Vladan Kuzmanović ◆ Ivan Ignjatović
Editors

CIVIL ENGINEERING 2021
ACHIEVEMENTS AND VISIONS

Proceedings of the International Conference celebrating 175th Anniversary of the Faculty of Civil Engineering, University of Belgrade

October 25–26, 2021, Belgrade, Serbia

Vladan Kuzmanović ◇ Ivan Ignjatović
Editors

Civil Engineering 2021
Achievements and Visions

Proceedings of the International Conference
celebrating 175th Anniversary of the

Faculty of Civil Engineering, University of Belgrade

October 25–26, 2021, Belgrade, Serbia

Publisher
UNIVERSITY OF BELGRADE – FACULTY OF CIVIL ENGINEERING

For the Publisher
Prof. dr Vladan Kuzmanović
Dean of the Faculty of Civil Engineering, University of Belgrade

Copublisher
ACADEMIC MIND, BELGRADE

For the Copublisher
Marko Vujadinović, Managing Director

Editors
Prof. dr Vladan Kuzmanović
Assoc. Prof. dr Ivan Ignjatović

Prepress and layout
Željko Hrček

Cover page design
Boris Popović

Printing
Planeta Print, Belgrade

Circulation
500 copies

ISBN 978-86-7466-895-5

PREFACE

Historia magistra vitae est. The 175-year history of the Faculty of Civil Engineering in Belgrade, a history made of research, teaching and building, shows us that knowing and respecting the achievements of our predecessors is an indispensable lesson in creating the future. That is why the scientific conference celebrating this great jubilee is called *Civil Engineering 2021 – Achievements and Visions*, and represents an ideal opportunity to look back at achievements in civil engineering, geodesy and geoinformatics at the beginning of the 21st century, but also to exchange visions of future development in these areas.

From the School of Engineering until today, the Faculty of Civil Engineering has changed its form, organization and curricula, but the essential concept of work has always been based on excellence. Excellence of a professor, researcher or designer. The concept of this conference is based on the excellence of twelve invited lecturers, from different research fields. Nine of them are from the institutions from abroad, while four of them are visiting professors of the Faculty of Civil Engineering, elected to these honorary titles in the jubilee year. All of them, in addition to excellent achievements in science and practice, have excellent cooperation with our faculty through joint projects or mentorships in doctoral studies. Three lecturers are active full professors of our faculty, so far distinguished by their scientific work, but also by creating a stimulating atmosphere for research, and forming a scientific youth with great potential. Short authors' biographies are presented before each work.

This collection contains 12 papers whose order follows the order of presentations at the conference. According to the topic and content, it consists of works in the fields of engineering mechanics and theory of structures (C. Butenweg, S. Živanović), materials and structures (S. Marinković, M. Veljković), geodesy and geoinformatics (B. Bajat, G. Huevelink), hydraulic and environmental engineering (D. Prodanović, S. Đorđević), geotechnics (L. Zdravković), roads, railways and airports (D. Lo Presti), construction management and technology (Ž. Popović) and fundamental sciences applied in civil engineering (E. Castro). Of course, three such broad scientific fields, such as civil engineering, geodesy and geoinformatics, consist of numerous specialized scientific fields and it is impossible to cover them all through only 12 papers. However, the topics of the conference – achievements and visions, gave the lecturers the task to go beyond the current scientific preoccupation and research focus. Looking beyond the current limits of their scientific fields, the authors were asked to prepare papers from an engineers' point of view that contribute to science without boundaries.

In addition to giving an overview of current knowledge, the conference papers should also serve as points of thought, as an inspiration for scientific discussions at the conference itself, but we also believe in the years after it. For researchers looking for sources of scientific ideas, and especially for those who are just beginning their scientific career - doctoral students, papers from this conference should spark creation, scientific breakthrough and personal contribution to the development of scientific thought in the thematic areas of the conference, and beyond.

As editors, we owe a great gratitude to the members of the Scientific and Organizing Committee who helped in the selection of lecturers, review of papers and successful organization of the conference. At the time of a pandemic, it was a serious task. We found a way to thank the previous generations of our professors, who also excelled in scientific research, through the formation of the Honorary Board. We had great institutional support from the Serbian Academy of Sciences and Arts, with which we share a common history through 20 members of the Academy – professors of the Faculty of Civil Engineering in Belgrade. Of course, many thanks to the Serbian Chamber of Engineers for sponsoring the conference and recognizing the significance and magnitude of this event.

Respecting the principles of the Open Science Platform, the proceedings will be freely distributed in both printed and electronic versions and made available to all, with the intention of reaching as wide an audience as possible.

The Editors

ПРЕДГОВОР

Historia magistra vitae est. Историја Грађевинског факултета у Београду дуга 175 година, историја саздана од истраживања, подучавања и грађења, учи нас да је познавање и поштовање достигнућа претходника неизоставна лекција у креирању будућности. Зато научна конференција поводом овог великог јубилеја носи назив *Грађевинарство 2021 – достигнућа и визије,* и представља идеалну прилику да се осврнемо на достигнућа у грађевинарству, геодезији и геоинформатици почетком 21. века, али и да разменимо визије будућег развоја ових области.

Од Инџинирске школе до данас, Грађевински факултет је мењао форму, организацију и наставне планове, али је суштински концепт рада увек био базиран на изврсности. Изврсности професора, истраживача или пројектанта. Концепт ове конференције базиран је на изврсности дванаест предавача по позиву, из различитих научних области. Од укупног броја, деветоро је са афилијацијом институција из иностранства, од тога четворо гостујућих професора Грађевинског факултета који су у та почасна звања изабрани управо у години јубилеја. Сви они, осим изврсних достигнућа у науци и струци, имају и одличну сарадњу са нашим факултетом кроз заједничке пројекте или менторства на докторским студијама. Троје предавача су активни редовни професори нашег факултета, који се истичу својим досадашњим научним радом, али и креирањем подстицајне атмосфере за бављењем истраживачким радом, те формирањем научног подмлатка са великим потенцијалом. Кратке биографије аутора приказане су испред радова.

Овај зборник садржи 12 радова чији поредак прати редослед излагања на самој конференцији. Према тематици и садржају, заступљени су радови из области техничке механике и теорије конструкција (C. Butenweg, С. Живановић), материјала и конструкција (С. Маринковић, М. Вељковић), геодезије и геоинформатике (Б. Бајат, G. Huevelink), хидротехнике и водноеколошког инжењерства (Д. Продановић, С. Ђорђевић), геотехнике (Л. Здравковић), путева, железница и аеродрома (D. Lo Presti), менаџмента и технологија грађења (Ж. Поповић) и базичних наука примењених у грађевинарству (E. Castro). Наравно, три тако широке научне области као што су грађевинарство, геодезија и геоинформатика, састоје се из бројних ужих научних области и није их све могуће сагледати кроз само 12 радова. Ипак, теме конференције – *достигнућа и визије*, дале су задатак и предавачима да изађу из оквира тренутне научне преокупираности и истраживачког фокуса и да, неспутани лабавим границама научних области, у припреми радова буду само инжењери у науци која не познаје границе.

Осим што дају преглед тренутних знања, радови са ове конференције треба да служе и као мисаоне тачке, као инспирација за научне дискусије на самој конференцији, али верујемо и у годинама након ње. За истраживаче који трагају за изворима научних идеја, а нарочито за оне који тек почињу своју научну каријеру – студенте докторских студија, радови са ове конференције треба да буду провокација из које ће проистећи креација, научни продор и лични допринос развоју научне мисли у тематским областима конференције, па и изван њих.

Као уредници, велику захвалност дугујемо члановима Научног и Организационог одбора који су помогли у одабиру предавача, рецензији радова и успешној организацији конференције. У доба пандемије то је био озбиљан подухват. Начин да захвалимо претходним генерацијама наших професора, који су се и сами истицали у научноистраживачком раду, нашли смо кроз формирање Почасног одбора. Велику институционалну подршку имали смо од Српске академије наука и уметности, са којом делимо заједничку историју кроз 20 чланова Академије – професора Грађевинског факултета у Београду. Наравно, велико хвала Инжењерској комори Србије за покровитељство над конференцијом и препознавање значаја и величине овог догађаја.

Поштујући принципе платформе за Отворену науку, зборник ће бити слободно дељен у штампаној и електронској верзији и доступан свима у намери да допре до што шире публике.

Уредници

Honorary Committee

Prof. Petar Anagnosti
Prof. Vojo Andjus
Prof. emeritus Dragan Budjevac
Prof. Branislav Ćorić
Prof. Marko Ivetić
Prof. Dušan Joksić
Prof. Vladimir Mićić
Prof. Mihailo Muravljov
Prof. Serafim Opricović
Prof. Mira Petronijević
Prof. Živojin Praščević

Scientific Committee:

Prof. Branislav Bajat
Prof. Dragan Blagojević
Assoc. Prof. Ljiljana Brajović
Assoc. Prof. Dejan Divac
Assoc. Prof. Ivan Ignjatović
Prof. Miloš Kovačević
Prof. Vladan Kuzmanović
Prof. Snežana Marinković
Prof. Zlatko Marković
Assoc. Prof. Goran Mladenović
Assoc. Prof. Marija Nefovska-Danilović
Prof. Jasna Plavšić
Prof. Zdenka Popović
Prof. Dušan Prodanović

Organizing Committee:

President

Prof. Vladan Kuzmanović
Dean of the Faculty of Civil Engineering

Members

Assoc. Prof. Jelena Dobrić
Assist. Prof. Aleksandar Đukić
Assist. Prof. Nenad Fric
Assist. Prof. Radovan Gospavić
Assoc. Prof. Ivan Ignjatović
Assist. Prof. Sanja Jocković
Assoc. Prof. Milan Kilibarda
Assist. Prof. Luka Lazarević
Assist. Prof. Miroslav Marjanović
Assist. Prof. Zorana Petojević
Assoc. Prof. Miloš Stanić

Organizer

University of Belgrade
Faculty of Civil Engineering

In cooperation with

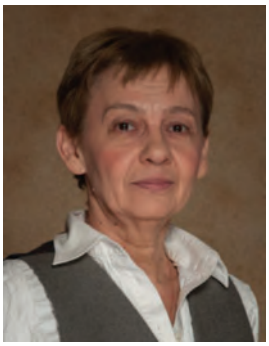
Serbian Academy of Sciences and Arts

Supported by

Serbian Chamber of Engineers

CONTENTS

Snežana Marinković



Snežana Marinković is a Full professor at the University of Belgrade, Faculty of Civil Engineering (FCEUB) since 2012, Head of the Chair for Materials and Structures (2015-2020), Vice-Dean for Science and Research at the FCEUB (2006–2012), member of the Council of the University of Belgrade (2015-2018) and member of the Council of Scientific Field of civil engineering - urbanism sciences of the University of Belgrade (since 2015).

Contact information: e-mail: sneska@imk.grf.bg.ac.rs

https://orcid.org/0000-0002-1808-392X

Prof. Marinković graduated from FCEUB in 1986, completed the magisterium at the FCEUB in 1991, and obtained a PhD at FCEUB in 2001 in the field of analysis and design of prestressed concrete structures. She has been working at FCEUB since 1987 as a teaching assistant (1987-2002), assistant professor (2002-2007), associate professor (2007-2012) and full professor since 2012, teaching a large number of subjects in the field of Concrete Structures. She has mentored four defended PhD theses and three on-going, and a large number of MSc and BSc theses. She was a member of several PhD defense committees at FCEUB and at the Faculty of Technical Sciences, University of Novi Sad. She co-authored four printed textbooks. She was the Chairwoman of the FCEUB's Teaching commission (2006–2012) and the Commission for PhD studies (2005–2012).

Prof. Marinković works in the area of analysis and design of concrete structures, especially in the area of application of green concretes and life cycle assessment. She is the founder and leader of the Concrete Structures Research Team (Chair of Materials and Structures, FCEUB). She has published 7 international book chapters, 19 articles in ISI-JCR-SCI journals, over 70 articles at international and national conferences, 19 papers in domestic journals, among others. Her works are cited 1193 times (h=14). She was a member of the Technical Committee and Head of the Serbian national group in the *fib,* and member of two RILEM Committees. She held several keynote lectures at international conferences, and was a member of the Scientific Committee of many international conferences. She serves as a reviewer for a number of ISI-JCR-SCI journals. She is a participant in several national projects, and was a Coordinator in one of them. From 2015, she is a member of the Home Scientific Committee for traffic, urbanism and construction of the Serbian Ministry of Education, Science and Technological Development. She is Editor-in-Chief of Building Materials and Structures Journal.

Prof. Marinković co-authored a number of conceptual and detailed designs, expertises, studies, and technical controls. A special mention should be given to the Detailed structural design of the roof and main columns of the Belgrade Arena in New Belgrade, for which, as one of the designers, she received the Award for greatest achievement in structural engineering in Yugoslavia in 1999, by the Yugoslavian Association of Structural Engineers (YASE). She achieved several short-term visits to technical universities in developed countries, especially through participation in COST actions. She was a coordinator from the Serbian side in projects funded by the Swiss National Science Foundation (SCOPES projects) and German Government (DFG projects), as well as within the project of Scientific and technological cooperation between Serbia and Portugal.

Prof. Marinković was the General Secretary of the Association of Structural Engineers of Serbia (2006–2010), and member of its Presidium since 2010. She received Award of the Belgrade Chamber of Commerce for the best graduate thesis in 1985/86 and the best PhD thesis in 2001; YASE Awards: for the greatest achievement in structural engineering in 1999, and the greatest scientific achievment in 2000/01.

Concrete structures in the 21st century and further

S. Marinković

University of Belgrade, Faculty of Civil Engineering, Chair of Materials and Structures

Summary

In less than 200 years reinforced concrete developed into cost-effective, versatile and durable construction material with a global production that makes it the most widely used one today. Modern times brought further development to cement composites with very high mechanical and durability performance, and improved cracking behavior and ductility, such as fiber and textile reinforced concrete, ultra-high performance concrete, and self compacting concrete. At the same time, various types of green concretes and technologies were introduced in attempt to improve the concrete sustainability aspect. Design of concrete structures grew from intuitive to performance-based design in 21st century. This work presents the development of design and construction of concrete structures through history, as well as current status and future perspectives.

Keywords: history of concrete structures, fibre and textile reinforced concrete, high and ultra-high performance concrete, 3D construction printing, performance-based design

1. INTRODUCTION

Concrete is the most widely used construction material today. The global production of concrete is extremely high – roughly 33 billion tons of concrete are produced globally each year, or over 4.7 tons per person per year ISO/TC 071 (2016), only second to water in mankind's consumption. Figure 1 shows that growth rate of cement production in last 70 years (approximately half of this amount is used in concrete) was much higher than population and steel production growth rate. In developed countries, structural concrete constitutes over 50% of all concrete, and developing countries will follow this trend (UNEP (United Nations Environment Programme), 2017). The demand for concrete is expected to increase, by 12-23% by 2050 compared to 2014, due to the rise of population and housing and infrastructure development needs, especially in Asia (IEA, WBCSD, 2018). All this started more than 2000 years ago in the Roman Empire.

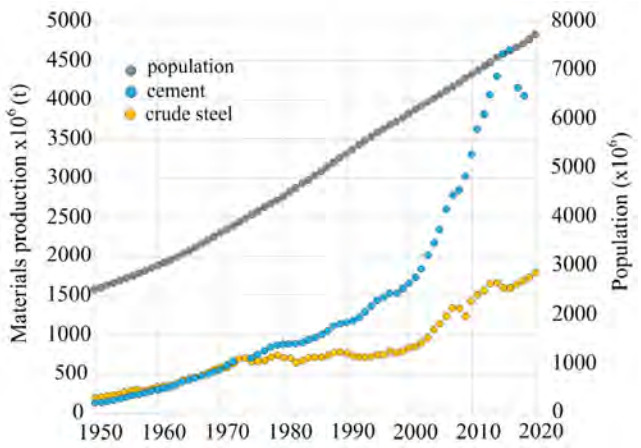


Figure 1. The growth of population, cement production and crude steel production over 1950-2020 period (data derived from (CEMBUREAU, 2017; 2018; 2019; 2020; UN Population Division, 2021; UNEP (United Nations Environment Programme), 2017; World Steel Association, 2021))

2. AT THE BEGINNING (EARLY CONCRETE STRUCTURES)

In ancient times various types of mortars were used as binders in stone and masonry constructions. From 3000 BC Egyptians used mud mixed with straw to bind dried bricks. They also used gypsum and lime mortars in the pyramids. Greeks, Cretans & Cypriots used lime mortars while Chinese used cementitious materials to hold bamboo together in their boats and in the Great Wall (700 to 200 BC). Babylonians & Assyrians used bitumen to bind stones and bricks (300 BC).

However, the first use of a truly cementitious binding agent (as opposed to the ordinary lime commonly used in ancient mortars) occurred in southern Italy in about the second century BC. A special type of volcanic sand called pozzolana, first found near Pozzuoli in the bay of Naples, was used extensively by the Romans in their cement. This volcanic sand reacts chemically with lime and water to solidify into a rocklike mass, even when fully submerged

(Shaeffer, 1992). The Romans used it for bridges, docks, storm drains, and aqueducts as well as for buildings.

Roman concrete largely differs from modern-day concrete because it was not produced with fluid consistency that could flow into a mold or a formwork. It was constructed in layers by packing mortar by hand in and around stones of various sizes. Since Romans often used rubble as aggregate, this first concrete might be more correctly termed cemented rubble. This assembly was usually faced with clay bricks on both sides, unless it was below grade. It was known that the bricks had little structural value so they were used to facilitate construction and as surface decoration. Having no reinforcement, Roman concrete could depend only on its tensile strength which was low.

The most famous structures from this period are Pantheon and Colosseum in Rome. These structures were intuitively "designed" and constructed. For instance, the Pantheon's dome is of spherical shape with clear span of 43.3 m. The spherical shell is a structural shape best fitted to a material with low tensile strength. Besides, the builders of the Pantheon used weight-reducing, waffle-slab-like voids in a dome as well as decreasing density concrete higher up in the walls and in the dome itself, in order to reduce the weight to be carried, Figure 2 (Radić *et al.*, 2008). The Pantheon's dome constructed in the second century AD is the largest unreinforced concrete dome ever built, and it is still standing.

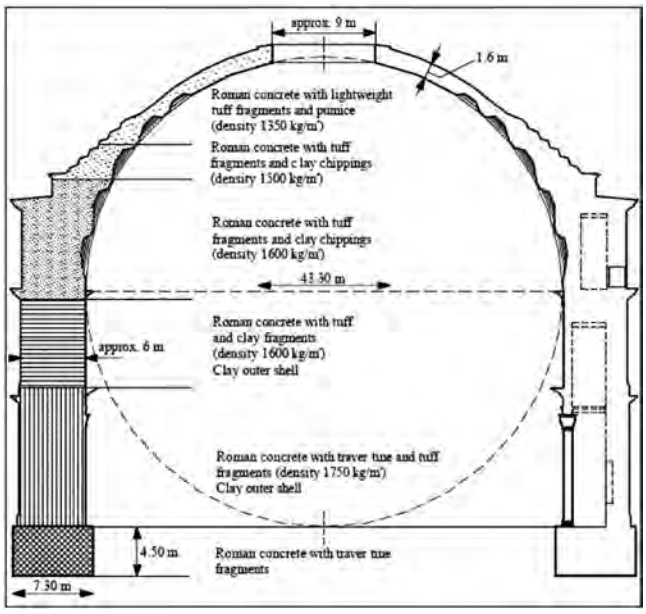


Figure 2. The temple Pantheon in Rome, construction lasted from 118 to 128 AD (Radić *et al.*, 2008)

Most certainly, the pozzolana cement made this type of construction possible. It was used throughout the Rome/Naples area but is not seen in northern Italy nor elsewhere in the Roman Empire. In the Middle Ages the "concrete making" skill was lost probably due to the lack of availability of similar pozzolans throughout the world. Stone and brick masonry continued to be the dominant construction materials for most of the world's significant buildings for many centuries.

The Industrial Revolution in Europe in the late 18th century brought structural steel as new construction material but also new developments and progress in cement and concrete were achieved. The precursor to modern-day cement was created in 1824 when Joseph Aspdin, an English mason, invented "improved cement" which he named Portland cement after the famously strong building stone quarried in Portland, UK. He was the first to use high temperatures to heat finely grinded limestone and clay until the mixture calcined, grinding it and then mixing it with water. Until 1860 the cement chemistry was completely understood and production technology improved - the era of Portland cements of modern composition began. This cement enabled concrete technology and properties as we know them today.

3. REINFORCED CONCRETE STRUCTURES – THE RISE OF EMPIRE

Plain concrete being an artificial rock has very low tensile strength. The use of steel reinforcing practically made this material capable of bearing both compressive and tensile stresses and enabled new structural forms such as cantilevers, monolithic frames, flat plates and slabs, and thin shells.

Approximately at the same time (1848 and 1849) first reinforced concrete was introduced by two Frenchmen, Jean-Louis Lambot and Joseph Monier, but neither of their inventions was actually a structure. J.L. Lambot was a farmer who constructed several small rowboats reinforced with iron bars and wire mesh (1848). J. Monier was on the other hand a gardener, who made a flowerpot for orange trees from reinforced concrete – concrete tubs with wire reinforcing (1849). He understood that this material had a great potential and generally deserves the credit for making the first practical use of reinforced concrete in structures in 1849 to 1867. He acquired first French patent in 1867 for iron reinforced concrete tubs, then followed by his pipes, tanks, flat plates, bridges, and stairways. He apparently had no quantitative knowledge regarding its behavior or any method of making design calculations - his patents relayed on intuition.

At the beginning, reinforced concrete was used mostly for industrial buildings, as this "new" material did not have the social acceptability of stone or brick. Buildings' structures were in the form of monolithic frames, mainly developed by Ernest L. Ransome in the United States, and by Francois Hennebique in France. These two were the most responsible for the successful promotion of reinforced concrete structures at the end of nineteenth century. Besides, Hennebique was probably the first to use stirrups

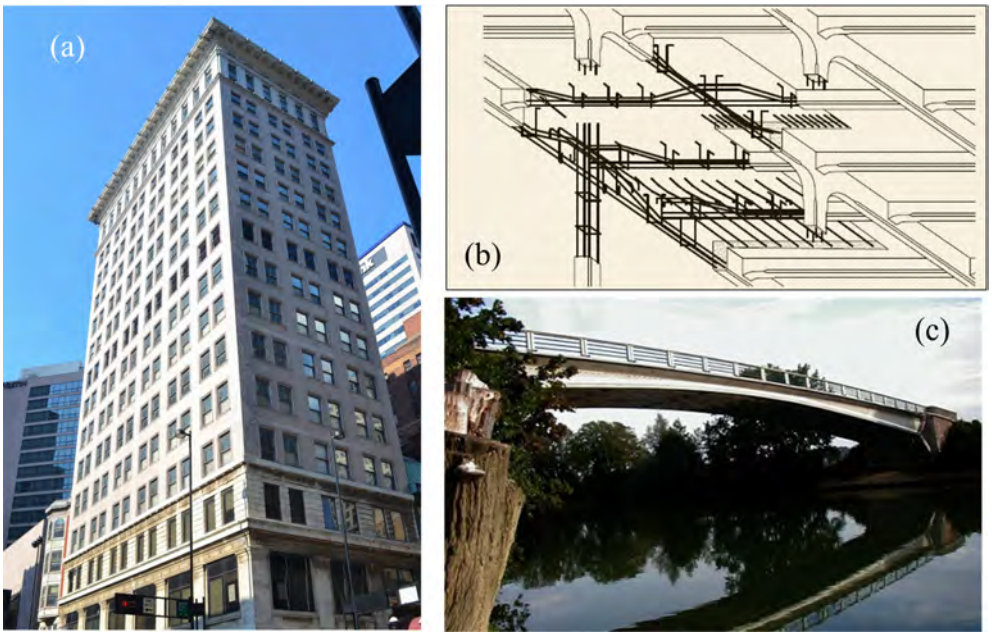


Figure 3. (a) The Ingalls Building, Cincinnati, USA (1904), (b) reinforcement layout from the Hennebique patent (1892) (Radić *et al.*, 2008), (c) The Luzancy Bridge, France (1946)

and bent-up bars to reinforce beams against shear, Figure 3 (Radić *et al.*, 2008). The Ingalls Building, a landmark structure in Cincinnati, USA, was built in 1904 using a variation of the Ransome system, Figure 3. It was the first concrete skyscraper, reaching 16 stories (63 m).

The first ideas on prestressing of reinforced concrete structures arose in 1930s. Eugène Freyssinet in France and Franz Dischinger in Germany developed this new technology mainly in bridge design and construction. It opened up new possibilities in spans, forms and aesthetics of reinforced concrete structures. The Luzancy Bridge over Marne River, designed by Freyssinet, was the first prestressed bridge to be built in France in 1946. This 55-metre span two-hinged rigid prestressed concrete frame is still in use, Figure 3.

From that time onward, reinforced concrete experienced an outstanding development as construction material, both in material and construction technology, as well as in the design theory. Today reinforced and prestressed concrete are successfully applied in all types of construction, building, industrial and infrastructure. Twice as much concrete is used in construction around the world than the total of all other construction materials. No alternative for concrete as a major global construction material currently exists that can be applied at sufficient scale.

There are several reasons for such popularity: easily available and low-cost constituent materials, good mechanical properties and fire resistance, any desired structural shape can be obtained. Concrete is durable with low maintenance cost – in some applications it is designed to last for hundred years. In a word, concrete provides the best ratio between the cost and quality compared to other construction materials. However, ordinary reinforced concrete is not suited for long-span structures for its relatively large ratio between the self weight and strength – this problem can be overcome by applying prestressed concrete or special structural systems such as two-chord girders for instance, Figure 4.

4. THE 21ST CENTURY

Modern age requires fast, quality and cost-effective construction as well as functional, robust, durable, and sustainable structures. In the 21st century care for the life on this planet and its survival imposed a sustainability imperative upon all human activities. Construction industry is no exception to this rule and traditional materials, technologies and practices have to be changed to meet those requirements. In the last 40–50 years concrete composition was significantly improved and new types of concrete were developed owing to invention of various chemical admixtures or application of mineral additions, polymers, and fibers. At the same time, mixing and casting methods as well as innovative construction technologies are introduced.

4.1. Materials – special concretes

The aim of the research on concrete composition is an improvement of its physical and mechanical properties, durability and, in the last few decades, reducing the impact on the environment. Invention of fibre reinforcement in 1970s, superplasticiziers in 1980s and introduction of silica fume as pozzolanic additive, enabled development of concrete with high compressive and tensile strength, improved cracking behavior and ductility, and with excellent

Figure 4. The roof structure of sports hall Belgrade Arena (1992) – two-chord externally prestressed reinforced concrete girders, spans 132.7 and 102.7 m, covering the area of 15000 square meters

resistance to various environmental actions causing deterioration of concrete during its service life.

4.1.1. Fibre reinforced concrete

The research on fibre reinforced concrete (FRC) started in 1960s in USA with investigation on tensile strength of concrete with uniformly distributed closely spaced short lengths of wire reinforcement (Di Prisco *et al.*, 2009). This research was the basis for a patent on steel fibre-reinforced concrete (SFRC) in 1969 and in 1970. Since that time, many new types of fibres were produced and a wide range of research on FRC properties was performed. RILEM Technical Committee TC162-TDF produced design guidelines for typical structural elements made of SFRC in 2002 and 2003 and afterwards recommendations were produced by other countries (Di Prisco *et al.*, 2009). While FRC was introduced in USA ACI 318 design codes in 2008, European codes still don't include this type of concrete.

FRC is a type of concrete in which fibres – mostly made of steel, glass or plastic (polymer) materials – are added in order to improve its cracking and fracture behavior. If fibres are short and added to the concrete during the mixing process they will be randomly distributed and embedded in the mortar, i.e. in the cement matrix, Figure 5.

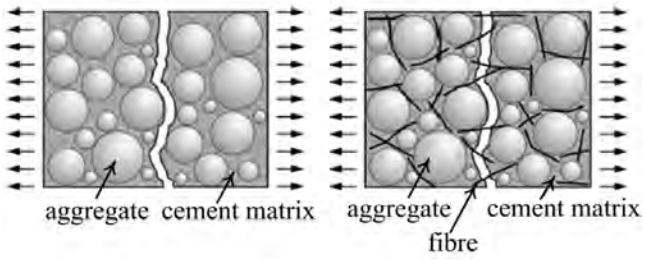


Figure 5. Crack in conventional and fibre renforced concrete (Fib, 2009)

By integrating ductile fibres with a high tensile strength in the matrix, concrete cracking can be reduced, i.e., a fine cracking pattern can be achieved, while cracking resistance is only slightly increased. The stress-strain behavior of FRC under uniaxial tension is shown in Figure 6. Brittle behavior of concrete is changed to strain softening or strain hardening behavior, depending on the fibre content, its geometry and the mechanical properties, as well as the bond between fibre and concrete matrix.

As can be seen the main advantage of FRC compared to non-reinforced concrete is its higher energy capacity, ductility and toughness – adding fibres generate a post-cracking residual tensile strength in combination with a large tensile strain. The higher the pull-out resistance of the fibres, and the longer it is retained as the strain increases, the slower is the decrease in the transferable tensile load and the more significant increase in concrete energy capacity.

The fibres used for FRC are mainly produced from steel, carbon, alkali-resistant glass, and polymer materials, Figure 7. Natural organic and mineral fibres, such as wood, cotton, bamboo, cellulose etc, were also investigated for internal curing (Dávila-Pompermayer *et al.*, 2020) and self-healing in concrete (Rauf *et al.*, 2020). However, natural fibers have an aging problem – they deteriorate in course of the service life (Li *et al.*, 2021), which has to be carefully taken into account in concrete structures.

Steel and carbon fibres have high tensile strength and elastic modulus, but the bond strength of carbon fibres is much lower. Glass fibres show a pronounced brittleness and lower modules of elasticity. Polymer or plastic fibres made out of polypropylene, polyvinyl or polyester are characterized by a sufficiently high strength, good bond behavior and very low modulus of elasticity. Besides, when the concrete is subjected to very high temperatures, the plastic fibres will melt and provide space for the expanding or evaporating water in the concrete. Large scale spalling can therefore be prevented.

The amount of fibres which can be incorporated in the mix depends on the composition and workability of the

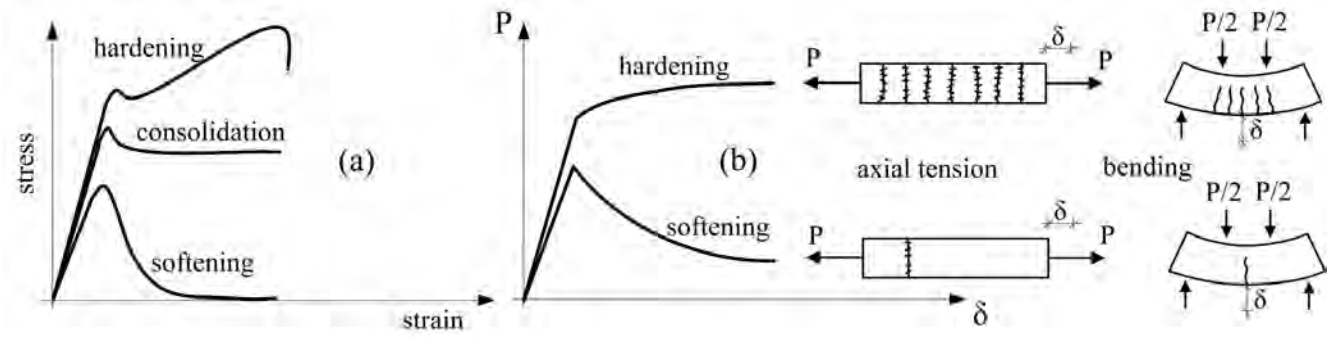


Figure 6. Strain softening and hardening (a) at section level and (b) softening and hardening behavior of structure in tension and bending (in the case of softening behavior the deformations localize in one crack while in the case of hardening behavior multiple cracking occurs before reaching the peak load value)

fresh concrete, the fibre properties and the mixing technology, and in practice ranges from 0.5–3.0% of concrete volume (Fib, 2009). To avoid the formation of fibre agglomerations it is recommendable to use less coarse aggregate with maximum particle size of 8 mm. Water/cement ratios between 0.4 and 0.5 turned out to be favorable for the production of FRC (Fib, 2009). Because of a low ratio of coarse aggregates and presence of fibres, the amount of required water is higher then in non-reinforced concrete. It is recommendable to use plasticiziers to avoid the high cement content.

In general, the compressive strength of FRC shows a minor increase with increase of the fibre content, but the stress-strain curve shows increased strain at failure and energy capacity because micro-cracking is restrained. The increase in the tensile strength is more pronounced. The strain at failure and fracture energy is significantly higher and may increase to several times the fracture energy of non-reinforced specimens. For this reason, higher resistance to dynamic loading and impact stress is generally also observed. This is, besides minimizing shrinkage cracking, also a reason why FRC is often applied for industrial floors or foundations. The shear strength of FRC can be enhanced due to "bridging effect" of fibres (at cracked state, the fibres establish a bridging of crack edges and inhibit the advanced opening of the crack), but to a limited extent. In this respect, the fibres act similarly to a shear reinforcement, but are less effective at an identical reinforcement ratio (Fib, 2009).

The use of fibres has a minor influence on the shrinkage and creep deformation of concrete. It does not prevent the formation of cracks that occur as a result of imposed and internal stresses. However, instead of few single cracks a large number of microcracks are produced and thus the crack widths can be reduced to a tolerable degree (Fib, 2009). Compared to non-reinforced concrete, SFRC has slightly higher resistance to elevated temperatures. This is achieved by the higher degree of cohesion created by the steel fibres. Polymer fibres (especially polypropylen fibres) are added in a targeted way to increase the fire rating of concrete, for the reasons explained above (Fib, 2009).

Durability related properties are similar to those of non-reinforced concrete since capillary and absolute porosity are not affected by fibres. In the carbonated edge zone of concrete elements, single fibres may corrode. However, this corrosion does not cause any significant damage, since spalling usually do not occur because the bursting pressure of the corrosion products created around the fibres is not sufficient (Fib, 2009). However, a part of the cross-section is lost due to corrosion as in ordinary reinforced concrete. The abrasion resistance depends on the fibers content. At a steel fibre ratio of 1.0 % of concrete volume, a significant increase in the hardness

Figure 7. Types of fibres in FRC: a) steel, b) carbon, c) glass and d) polypropylene fibers

of the concrete surface and wear resistance was observed (Fib, 2009).

FRC is usually applied in the construction of industrial floors, precast tunnel segments, pipelines, as shotcrete in repair works etc., Figure 8.

Figure 8. FRC precast tunnel lining segments (up) and FR shotcrete in underground structure (down)

Instead of random, three-dimensional short fibres distributions, it is possible to insert long fibres in the direction of the anticipated tensile stresses which is characteristic for **textile reinforced concrete (TRC)**. TRC is generally a concrete reinforced with non-metallic reinforcement, whether in form of bars, 2D mesh or 3D mesh, Figure 9. Textile reinforcement usually consist of tows (group of long continuous fibres bundled together in parallel) which are then further grouped to directionally oriented, 2D mesh or 3D mesh forms. Carbon, alkali resistant glass and basalt fibres have been most commonly used in TRC. Concrete must be produced without coarse aggregate to match the mesh openings, usually with addition of silica fume and fly ash.

This concrete can be used for strengthening of reinforced concrete structures (Scheerer *et al.*, 2019) while application in new structures is at the current state-of-the-research limited to innovative initiatives, mostly in Germany. While there is lot of work on behavior and application potential of TRC (Hegger and Voss, 2008; Scheerer *et al.*, 2017) for wider application much more research on the behavior of structural elements is needed (Bielak *et al.*, 2019; Portal, 2015).

4.1.2. High strength and ultra-high performance concrete

With invention of plasticiziers and superplasticiziers (chemical admixtures which enable reducing of the amount of water in the concrete mix, while maintaining the required workability) it was possible to produce concrete with compressive strength higher than 50–60 MPa, which was the limit for ordinary, normal strength concrete (NSC). Since 1980s new type of concrete was developed and named High Strength Concrete (HSC) with compressive strength up to 120 MPa, Figure 10.

Efforts were continuously focused on developing concretes with even higher compressive strength. In 1990s reactive powder concrete (RPC) was developed using components with increased fineness and reactivity and thermal curing to obtain very dense micro and macro structure. This composite is characterized by very high cement content, very low water/binder (w/b) ratio, use of silica fume (SF), fine quartz powder, quartz sand, superplasticizier and steel fibers (Azmee and Shafiq, 2018). In the late 1990s, the first ultra-high performance concrete (UHPC) developed through RPC technology was commercialized under the name Ductal®. In 1997, the world's first RPC structure as shown in Figure 11 was built for pedestrian bridge in Sherbrooke, Canada. Since 2000s much progress has been made on the development of UHPC. With further developments of the concrete technology, engineers realized that the advanced concrete, besides the high strength, should also have other excellent properties, which led to the term UHPC instead of UHSC.

UHPC is usually defined as a concrete having a characteristic compressive strength above 120 MPa. Water/binder ratio of such concretes is usually from 0.15 to 0.25 (Abbas *et al.*, 2016), and highly reactive additions such as silica fume (SF) have to be added to the mix of aggregate, cement, water, and admixtures. The effect of SF in concrete mix is dual, micro filler effect and effect of pozzolanic reaction. Having very fine particles, it fills the voids between the coarser cement particles enabling in that way the maximum possible packing density of all granular constituents. For this purpose, the amount of SF is required to range from 10 to 30 % of the cement mass (Abbas *et al.*, 2016; Fib, 2009). Secondly, addition of SF reduces the amount of calcium hydroxide $Ca(OH)_2$ present in the interfacial zone between aggregate and matrix. The almost pure SiO_2 in silica fume reacts with $Ca(OH)_2$, thus forming calcium silicate hydrates (CSH) of much higher strength while the porosity decreases in the bulk and in particular in the interfacial zone. All these effects result in a significant increase in compressive strength (Fib, 2009). Beside SF, quartz powder with particle sizes similar to cement is predominantly used in heat-treated UHPC. The quartz particles are inert at room

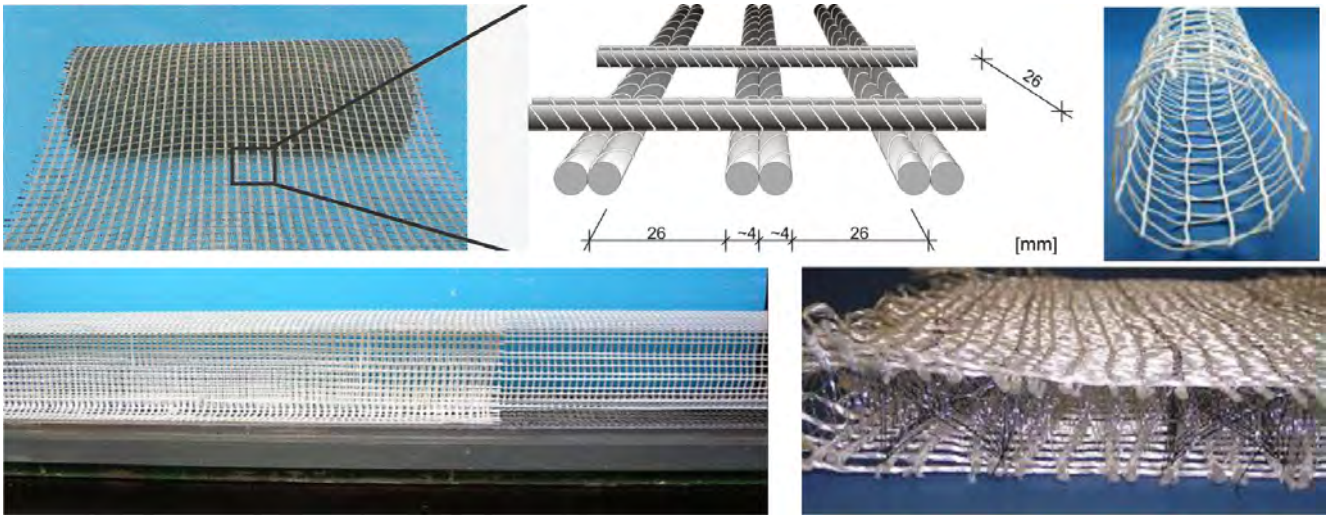


Figure 9. 2D and 3D mesh of textile reinforcement

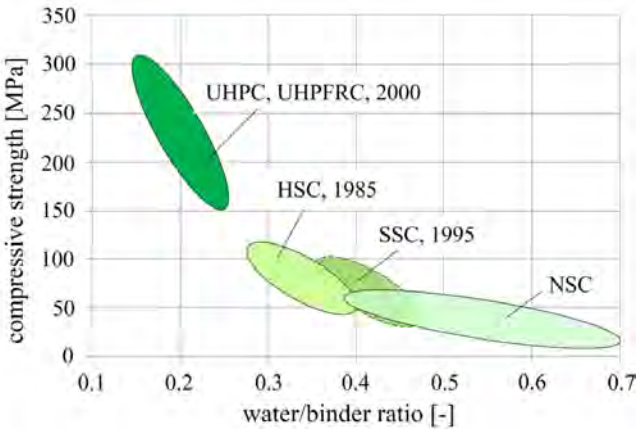


Figure 10. Development of the compressive strength of concrete as a function of the water/binder (w/b) ratio for different types of concrete (Müller, 2007)

Figure 11. Sherbrooke pedestrian bridge in Canada (1997) – span 60 m, precast prestressed RPC (200 MPa)

temperature but at high temperatures react with $Ca(OH)_2$ to form CSH phases (Fib, 2009).

Portland cement contents used in UHPC are normally very high and range from 600 to 1000 kg/m^3 (Fib, 2009). Aggregates with sufficiently high strength and adequate grain size distribution for high packing density should be applied. The workability of the UHPC can only be obtained by applying high amounts of high performance plasticiziers (up to 5% of the cement mass) (Fib, 2009). Typical composition of UHPC mix consists of cement (27–40% weight), SF (6–12% weight), quartz powder (7–14% weight), sand (35–45% weight), water (4–10% weight), and superplasticizier (0.5–3% weight) (Abbas *et al.*, 2016).

The compressive and tensile strength of UHPC depend strongly on the curing type. In ordinary curing conditions at temperature of 20°C, a maximum compressive strength of UHPC of approximately 200 MPa may be achieved. At curing temperature up to 250°C the strength will reach approximately 400 MPa. A compressive strength of 800 MPa may be achieved if additional mechanical pressure is applied (Fib, 2009). The uniaxial tensile strength and the flexural tensile strength of UHPC reach values ranging approximately from 10 to 60 MPa (Fib, 2009). On deformation level, UHPC shows an elastic and brittle behavior - the stress-strain relationship of UHPC is linear up to 80–90% of the ultimate load level (Abbas *et al.*, 2016). The modulus of elasticity of UHPC is significantly higher compared to NSC – it reaches approximately 50 to 60 GPa when the compressive strength is 200 MPa. The modulus of elasticity of RPC having a compressive strength up to 800 MPa may approach 75 GPa (Fib, 2009).

High volume content of hardened cement paste results in poor rheological properties. The autogenous shrinkage of UHPC is very pronounced and may reach values beyond 0.0012 (1.2‰) (Fib, 2009). For UHPC cured in ambient conditions and loaded at 28 days at stress levels 40-60% of compressive strength, estimated final creep coefficient varies between 0.8 and 1.5 (Fib, 2009; Flietstra, 2011; Russell and Graybeal, 2006). However, UHPC ex-

hibits much lower creep coefficient ranging from 0.3 to 0.7, if cured at high temperature (Russell and Graybeal, 2006).

All UHPC durability related properties are superior compared to NSC. Physical characteristics of concrete are mainly determined by its capillary porosity. This porosity is very low in UHPC, making absolute porosities so low that transport processes almost stop. As a consequence:

- water permeability is very low, approximately 10 and 60 times lesser than that of HSC and NSC, respectively (Abbas *et al.*, 2016; Mitrović *et al.*, 2021); for instance, the water permeability of a C190 UHPC was tested at three different temperatures and was found to be in a range between 4.0 to 5.0 $\times$ 10^{-15} m^2/s, which corresponds to the permeability of dense natural stone (Fib, 2009)
- carbonation depths measured on UHPC are almost negligible (Abbas *et al.*, 2016; Mitrović *et al.*, 2021)
- the chloride diffusion coefficient for UHPC ranges between 0.2 to 2.0 $\times$ 10^{-13} m^2/s and is lower for an order of magnitude than that of NSC (around 1.0 $\times$ 10^{-12} m^2/s) (Abbas *et al.*, 2016; Mitrović *et al.*, 2021)
- excellent behavior under freeze-thaw attack (Abbas *et al.*, 2016; Fib, 2009; Mitrović *et al.*, 2021)
- abrasion wear resistance is several times higher compared to NSC (Fib, 2009; Mitrović *et al.*, 2021).

As already mentioned, UHPC is a brittle material as natural stone, only with much higher tensile strength. To compensate for brittle behavior, fibres are added to concrete mixture resulting in a special type of concrete named **ultra-high performance fibre reinforced concrete (UHPFRC)**. It has been proven that the steel fibres amount of approximately 2.5 % of concrete volume leads to best results, both in view of fresh and hardened concrete properties (Fib, 2009). As in FRC, the fibers prevent macrocracking by holding the crack edges together already at the micro-cracking stage and thus hinder the cracking progress. As a result, a significant additional strain capacity is obtained while the increase in strength is minor. Similar to ordinary FRC, the fibre content and its properties determines whether the resulting behavior is of softening or of hardening type.

After two decades of development, several commercial UHPC formulations have been developed worldwide such as Ductal®, Dura®, BCV®, BSI®, Cemtec® (Azmee and Shafiq, 2018). UHPC has been used for both structural and non-structural precast components in many countries, Figures 12 and 13.

However, the UHPC outstanding physical and mechanical properties come with costly and sophisticated technology. The high material cost, together with possible limited available resources severely limits its commercial development and application in modern construction industry,

Figure 12. Shawnessy railway station, Canada (2004) – the thinnest concrete shell ever constructed (UHPFRC shell, 6 m × 5.15 m in layout, thickness 18 mm, 150/25 MPa UHPFRC compressive/tensile strength)

especially in developing countries. Lack of design codes also contributes to UHPC low acceptability. Apart from the two French national standards for UHPC published in 2016 (Azmee and Shafiq, 2018), no other standards have been adopted yet for this type of concrete. Besides, there is a question of its sustainability. UHPC allows a significant reduction of the cross-section of members while retaining a high load bearing capacity, and provides prolonged service life. Still, very high contents of cement and SF and high energy consumption result in high CO_2 emissions, and make the sustainability of UHPC questionable.

4.1.3. Self-compacting concrete

Self-compacting concrete (SCC) is a type of concrete that de-aerates and flows without the application of compaction energy, while staying homogenous during the whole placing process until hardening (Fib, 2009). This type of concrete was first invented in Russia and Japan in order to minimize negative effects of the placing process on the concrete properties, by omitting the external compaction process (Ozawa *et al.*, 2004).

In order to enable the concrete to de-aerate properly without any additional compaction work during the placing process, the dynamic viscosity of the mixture must be low enough to provide that entrapped air bubbles rise to the concrete surface. Today, two procedures are available to achieve this (Fib, 2009):

- *powder type self-compacting concretes* are characterized by an approximately 30% higher paste content compared to conventional vibrated concrete. This ensures a high fluidity of the mix. The paste itself consists of cement, large amounts of mineral additives such as fly ash or limestone powder, water and chemical admixtures (above all superplasticizers). In order to prevent segregation, the water/cement ratio and water/binder ratio are kept very low, leading to high strength values at the hardened state. Due to the high paste content, powder type SCC generally exhibits a reduced modulus

Figure 13. Jean Bouin Stadium, France (2013) – roof and façade self-supporting lattice Ductal® UHPC panels (roof panels 8/9 × 2/2.5 m, thickness 3.5/4.5 cm, façade panels 9 × 2.5 m, thickness 11cm)

of elasticity in comparison to conventional concretes of equal strength (Fib, 2009).

- *stabilizer type self-compacting concretes* have lower paste content (approximately 10% higher than conventional concrete) but significantly higher water/binder ratio. By adding superplasticizer the viscosity of the mixes, especially at low shear rates is strongly reduced hence allowing the entrapped air to leave the concrete. At rest, however, strong segregation would occur unless this is prevented by adding stabilizing admixtures, so-called viscosity modifying agents (VMA). At the hardened state stabilizer type SCC practically doesn't differ from conventional concrete of equal strength. The increased water/binder ratio allows for the production of low and normal strength concretes with self-compacting properties.

Different composition of powder type SCC in comparison to conventional vibrated concrete leads to different hardened state behavior. The increased paste content results in a reduced modulus of elasticity and higher shrinkage and creep deformation. Modulus of elasticity is approximately 10-20% below the prediction of Model Code 2010 (MC2010) (Fib, 2013) for a conventionally compacted concrete of equal compressive strength (Fib, 2009), while shrinkage and creep exhibits an increase of approximately 10-20% (Muller and Kvitsel, 2006). However, this is well within the scatter band of MC2010, which is defined to be ±30% for shrinkage and creep predictions.

The durability of SCC is characterized by an identical or improved behavior compared to conventional concrete of equal strength, regarding the carbonation and chloride diffusion behavior. Higher paste content in powder type SCC can lead to a poorer freeze-thaw behavior. According to investigations by Ludwig (2001) the freeze-thaw resistance of powder type SCC strongly depends on the type of filler used. For SCC containing fly ash an equivalent or even higher freeze-thaw resistance was detected. However, concretes containing limestone powder showed an increased weathering in the freeze-thaw test.

4.1.4. Green concrete

It is well recognized today that human activities cause major environmental impacts. Among others, the construction industry places a large burden on the environment: huge consumption of natural resources, especially non-renewable fossil fuels, large emissions of substances that pollute the air, soil, and water, as well as the generation of substantial amounts of inert and toxic waste. In particular, the concrete industry plays a leading role in this process: concrete production is not especially harmful per unit of concrete; however, the global production of concrete is very high leading to a high total impact. For these reasons it is of crucial importance to find way(s) of greening the concrete industry, i.e., to decrease its impact on the environment. These efforts can be divided into three major groups: (1) recycling of construction and demolition waste as a way to reduce the amount of waste and consumption of natural resources, (2) on the material level by introducing green concretes, and (3) on the structural level, by design for longer service life and reuse.

Green concretes are developed with the following aims:

- to reduce the natural aggregate's (NA) consumption and waste generation – for the production of these green concretes recycled concrete aggregates (RCA) are used instead of NA; they are called recycled aggregate concretes (RAC);

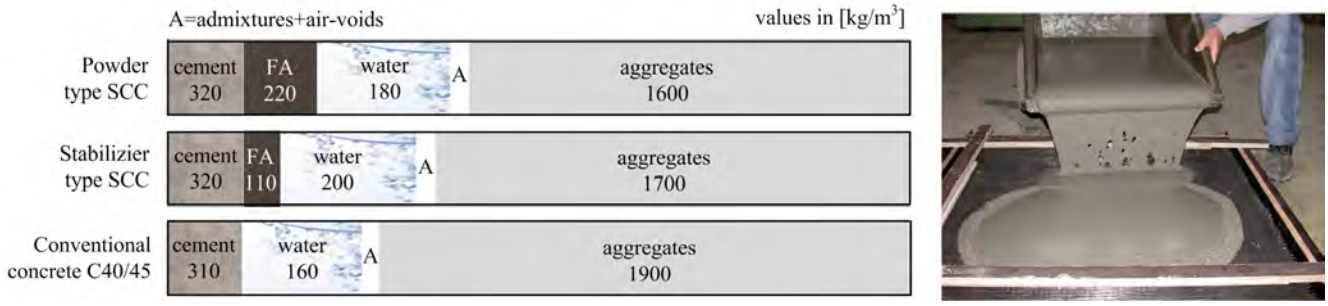


Figure 14. Mixture composition of a powder type SCC, a stabilizer type SCC and a conventional vibrated concrete C40/45 (Fib, 2009) (left) and self-compacting UHPC (right)

- to reduce CO_2 emissions which originate mostly from the cement production, namely by using a low content of cement.

Part of the cement is usually replaced by the so called supplementary cementitious materials (SCMs) with low embodied CO_2. The most common SCMs are reactive wastes from other industries; for instance, fly ash (FA), the by-product of the electricity production in coal-burning power plants, can be used for this purpose because it has pozzolanic properties. If concrete contains more than 35% of FA in the total cementitious materials mass, it is usually called high-volume fly ash concrete (HVFAC). Granulated blast furnace slag (GBFS), a by-product of pig-iron production in blast furnaces, is also commonly used. Fillers, inert or weakly reactive fine particulate materials, can partially replace cement or other reactive SCMs; the most commonly used is limestone filler. The cement content in concrete can also be reduced by appropriate mix design methods such as particle packing methods (Sunayana and Barai, 2017). These mix design methods enable better concrete properties with a smaller amount of cement. Recently, research has focused on the possibility of replacing large parts of cement with inert fillers and super plasticizers, in combination with particle packing methods (John *et al.*, 2018; Radović *et al.*, 2021). A concrete with a simultaneous replacement of cement and natural aggregates has the largest potential for a decrease of environmental impact, if engineering properties required for structural use can be obtained. For instance, concrete produced with RCA instead of NA and a partial replacement of cement with FA is named FARAC.

At the end of this 'cement replacement' line stands alkali activated concrete in which the cement binder is completely replaced by alkali activated materials rich in silicon and aluminum. Different natural and waste materials are activated with alkaline solutions, usually with a combination of sodium hydroxide and sodium silicate solutions. If low calcium FA is used, such concrete is named alkali activated fly ash concrete (AAFAC) which uses caustic sodium hydroxide and usually needs curing at elevated temperatures.

In Figure 15 results of LCA (Life Cycle Assessment (Marinković *et al.*, 2017)) performed for conventional (CC) and four different green concrete mixes are presented. Global Warming Potential (GWP) was calculated for 49 CC, 51 HVFAC, 54 RAC, and 30 FARAC mixes, collected from previous research. The analysis was performed for the production stage of the concrete life cycle (cradle-to-gate), while GWP was expressed as function of the concrete compressive strength (f_{cm}) to enable comparison, applying regression analyses for different concrete types. As it can be seen, CC mixtures have the highest values of this indicator, while GWP of HVFAC mixtures is almost 50% lower.

Whichever SCM, filler, aggregate type, or mix design method is used, the aim is to keep the structural performance of concrete at a desired level. Up to now extensive research was performed regarding the properties of HVFAC, RAC and FARAC and regarding the behavior of structural elements containing them (Marinković *et al.*, 2021). Although the production technologies are readily accessible and properties are well investigated, their application in structural concrete is not yet standardized (or it is allowed but with strong limitations canceling possible benefits). They are slowly entering the structural codes and, for instance, the new version of Eurocode 2 (CEN, 2004) planned for 2021 introduces RAC as a viable structural concrete.

4.2. Construction technologies

There is a saying in construction world: construction hasn't changed much since the times of pyramids – it is all about moving workers and materials. Indeed, the construction industry is probably the industry which experienced the smallest technology change in the last 50 years. The main reason is that the product (building, bridge, or highway) is too large to be produced and assembled in the factory. The highest level of industrialization and automation, until recent introduction of 3D printing technology, was precasting (prefabrication) of structural elements. Nevertheless, precasting together with prestressing technology enabled quality and fast construction of

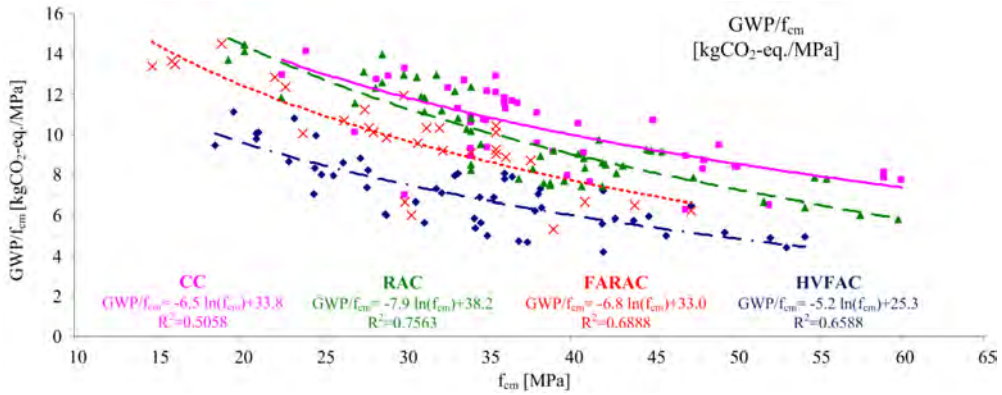


Figure 15. Normalized GWP of CC, HVFAC, RAC, and FARAC mixes per 1 m^3 of concrete

bridges and large building projects. Advancements in construction machines also contributed. For instance, for the construction of the Burj Khalifa skyscraper in Dubai, new super high-pressure trailer concrete pump was created that set the world record for vertical concrete pumping at 606 m. The Burj Khalifa has been the tallest building and man-made structure from 2009 until now, with total height of 829.8 m, Figure 16. Its primary structure is made of reinforced concrete. World's tallest bridge structure, Millau viaduct built in 2004 in France, with a height of structure equal to 343 m, and The Oceanographic building in Spain, are also shown in Figure 16.

With 3D printing technology introduced since 2000, radical changes in construction begin that parallel the changes automation and robotics brought to the industrial complex in the 70s. 3D printing is an additive technology, which means that a product is created from the ground up in a layer-by-layer manner. It was initially utilized for quickly and accurately creating prototype parts. Prior to the adoption of the building information modeling (BIM), 3D printing was already used by architectural companies to build scale models. It wasn't long before technology advanced to more ambitious construction purposes. For more than a decade, 3D printing has been used in several ambitious initiatives and projects in construction (Hossain *et al.*, 2020), Figure 17.

Due to already developed large printers, 3D printing can be used to create construction components or to print entire buildings and bridges. It may allow faster construction (3D printing has already shown that it can build a home in a matter of days) and building complex designs that are otherwise unattainable, or too expensive or labor-

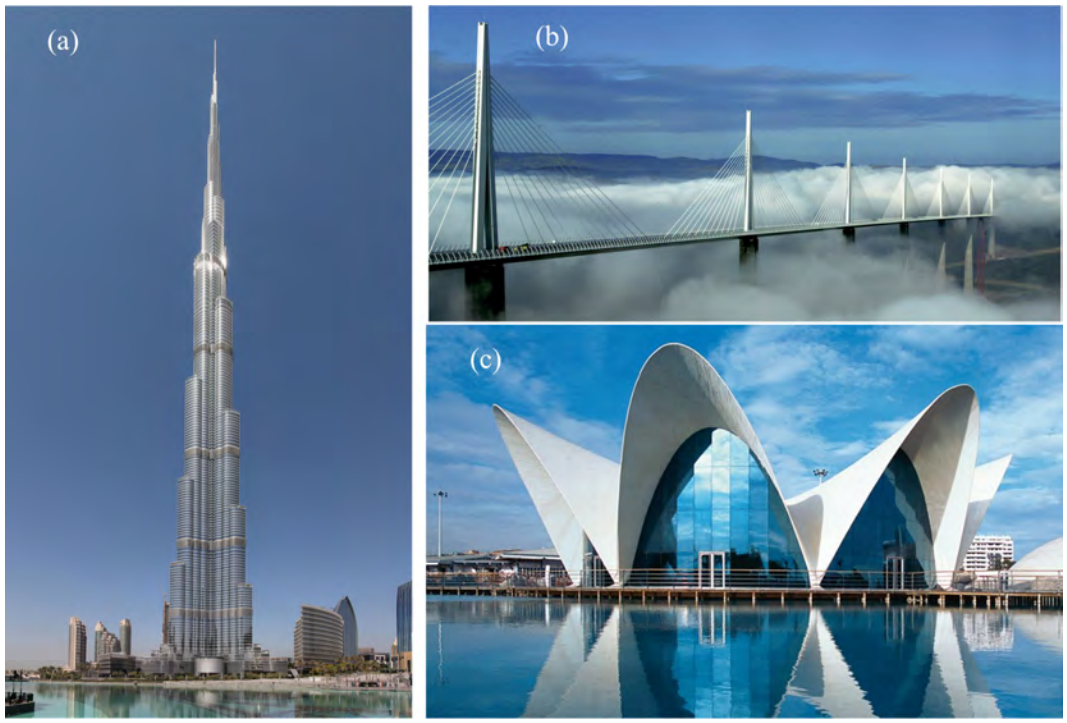


Figure 16. World's famous 21st century concrete structures (a) Burj Khalifa, Dubai, UAE (2009), total height 829.8 m, (b) Millau viaduct, France (2004), height 343 m, (c) The Oceanographic, Valencia, Spain (2003), 6.0–22.5 cm thick FRC shell

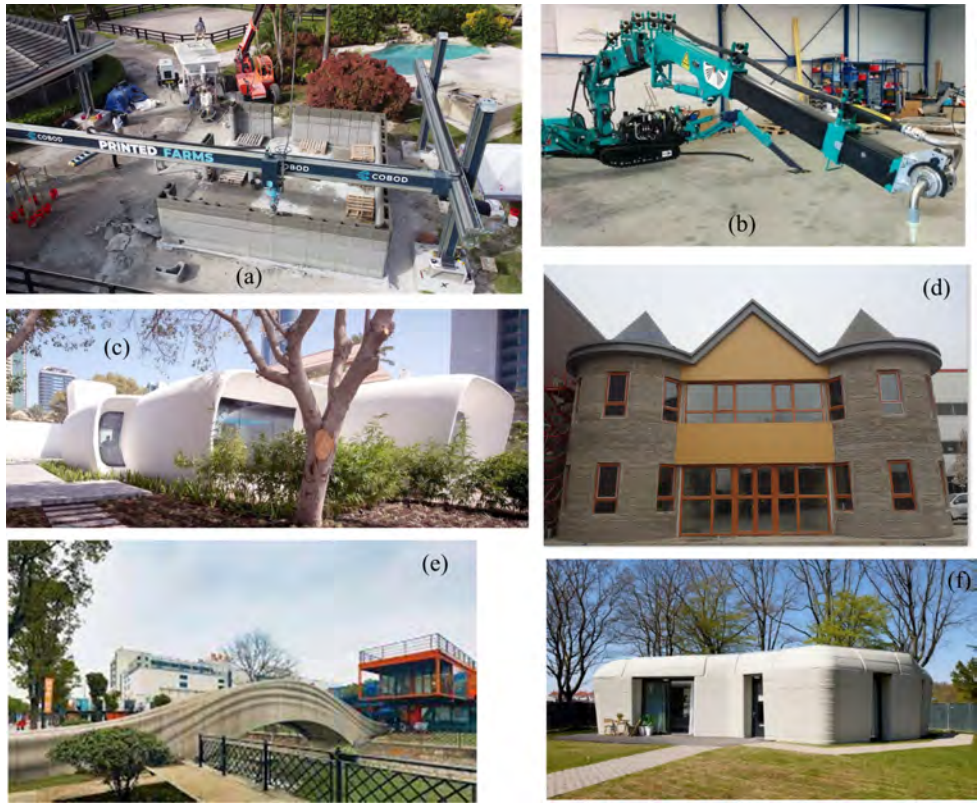


Figure 17. 3D large printers and construction projects in world (a) large 3D printer on site; (b) and at plant; (c) first 3D printed office building in Dubai, UAE, 2016 (6 m high, 36 m long and 12 m wide, constructed in 17 days); (d) two-story house, Shanghai, China, 2016 (entirely printed on site in 45 days with reinforcement and services embedded in walls); (e) longest 3D printed bridge, Shanghai, China, 2016 (span 26 m); (f) first 3D printed home in Europe within Project Milestone, Eindhoven, Holland, 2021 (where people actually live).

intensive to create by conventional construction means. It may enable lowering labor costs and producing less waste as well as construction to be undertaken in harsh and dangerous environments not suitable for a human workforce such as in space. Although this technology has enormous potential, there are many unresolved questions and it is questionable whether it will replace traditional concrete on a large industrial scale. It requires complex equipment (large 3D printers and the logistics) which cost of purchasing or renting is high, specific skilled labor, and adequate weather conditions on site. There is also the question of liability that may come with using robots rather than humans to perform certain construction tasks, how printed structures will incorporate services and reinforcement etc. Nevertheless, the potential of 3D printing is great. While the construction industry may never reach a point where it's used exclusively, it's only a matter of time that the technology will be improved and advance significantly.

4.3. Design

Design of concrete structures have progressed over last 170 years from intuitive, through working stress method to performance-based design which uses limit states concept and reliability.

Modern design codes (CEN, 2004; Fib, 2013) use a **performance-based approach** meaning that a structure or a structural component is designed to perform in a required manner during their entire life cycle. Performance is evaluated by verifying the behavior of a structure or a structural component against the specified performance requirements. The performance requirements are satisfied if all relevant performance criteria are met during the service life at the required reliability level. Further, performance criteria are quantitative limits defining the border between the desired and the adverse behavior, relevant for the specific aspect of performance. The performance requirements refer to the fulfillment of the essential demands of the stakeholders (founders, the owners, the residents, the users, the contractor, the design and construction team, the government, etc).

In MC 2010 (Fib, 2013) for instance, the concept of Limit State Design is applied to carry out performance-based design for serviceability and structural safety. Performance criteria for serviceability and structural safety are specified by serviceability limit states criteria, ultimate limit states criteria, and robustness criteria. The durability criteria are understood to be implicitly involved in the requirement that structures are designed for structural safety and serviceability for a predefined service life.

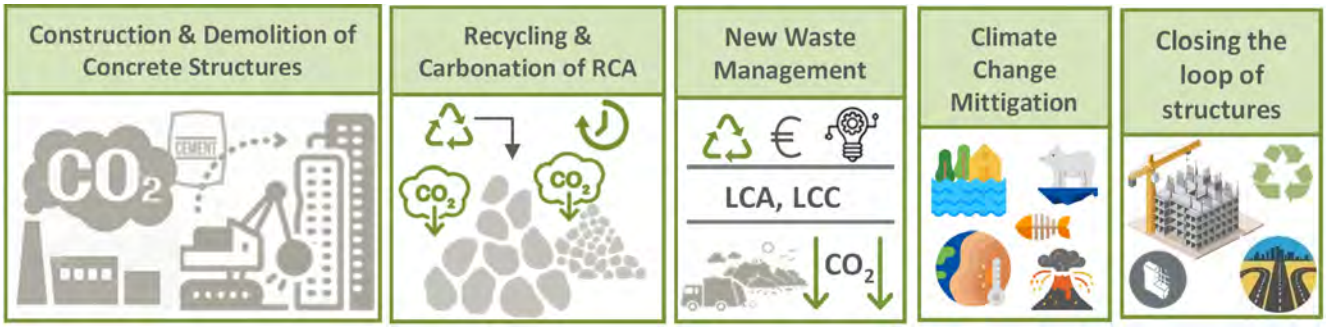


Figure 18. Closing the low-carbon concrete loop by recycling and fast carbonation of RCA

The serviceability limit states are related to functionality of the structure regarding the comfort of using the structure, and corresponding criteria are deformations, vibrations and local damage. The ultimate limit states address life safety, protection of the structure and environment, and protection of operations, and corresponding criteria are resistance of critical regions, fatigue, and stability. For the first time, MC2010 defines performance requirements for sustainability as related to impact on the environment, and impact on the society, which are defined as the influence of the activities, from the design to disposal, on the environment and on society, respectively. Performance requirements for environmental impacts include impact on human health, impact on social property, impact on biodiversity, and impact on primary productivity. Performance requirements for impacts on society are not defined, since at present, there are no general criteria for evaluating quantitatively the impact of a structure on the society. Therefore, MC 2010 recommends that depending on the purpose of the structures and the intents of designers, appropriate rating or target indexes should be set (with the weighing).

5. FURTHER

Concrete came a long way from Pantheon temple to Burj Chalifa in less than 200 years. With introduction of steel reinforcement, it developed into cost-effective, versatile and durable construction material with mechanical properties good enough for most of the structural purposes. In last 50 years it advanced even more to cement composites with very high mechanical and durability performance such as for instance UHPFRC. The intention to eliminate steel reinforcement as main cause of reinforced concrete deterioration led to introduction of short fibres or textile reinforcement, made of glass, carbon or polymer, i.e., FRC and TRC. The work on their further development continues and results will be seen in near future. For instance, adding nano-particles produced from silicon dioxide (SiO_2), aluminum oxide (Al_2O_3), iron oxide (Fe_2O_3) or titanium dioxide (TiO_2) to obtain higher particle packing density and accelerate cement hydration (Abbas *et al.*, 2016), or various hybrid solutions for enchasing FRC and UHPFRC behavior (Labib, 2020; Li *et al.*, 2019; Meng *et al.*, 2021; Raza *et al.*, 2020). etc.

All these modern composites must have hardening behavior (ductility), both in compression and tension, and adequate tensile strength to replace the role of steel reinforcement in concrete. However, their cost and impact on the environment is higher compared to conventional reinforced concrete. Research on reducing the cost and improving the sustainability of modern cement composites is needed to enable their future wider implementation. For instance, several studies were performed to modify the UHPC mixtures with the use of local raw materials and waste products in the effort of reducing the amount of Portland cement, steel fiber and SF (Aghdasi and Ostertag, 2020; Ferdosian and Camões, 2021; Li *et al.*, 2020; Yazıcı, 2007). With a lower cost and impact on the environment, UHPC and TRC will be much easier to be accepted by the construction market and arouse the interest of involved stakeholders.

A lot of research was focused on finding more sustainable solutions for conventional reinforced concrete and several types of green concretes were developed. Among them, the most promising are RAC and concretes with large part of cement replaced with various fillers such as limestone powder. Recycling represents a way to convert the concrete waste into a resource for new concrete and therefore is required for closing the concrete loop in the **Circular Economy** context. Beside the economic benefit, this concept contributes to environmental impacts decrease by minimizing resource use and waste generation. Moreover, concrete demolition waste crushed into RCA can be used as carbon sink - RCA re-absorbs CO_2 through carbonation process at much faster rate compared to concrete built in the structure. At the same time, carbonation leads to an improvement of RCA mechanical performance by reducing the porosity of residual cement paste fostering in that way broader application of RCA in new construction and closing the low-carbon concrete loop, Figure 18. The technologies for fast carbonation of RCA are expected to be developed on much larger industrial scale in near future.

The other prerequisite for wide application of new concretes and new types of reinforcement in structural engineering is that production quality control and structural design are regulated. So, on design level, modern composites with steel and non-metallic reinforcement, whether in form of short fibers, textile or conventional bars, as well as green solutions should enter the design codes in future. In majority of these new solutions, the state-of-the-research is developed enough to enable regulations and codes to be produced. Durability aspects will also be connected to limit state criteria, thus enabling the designer to conduct a service life design based on the load actions (e.g. carbonation or acid attack) and the resistance to this action (e.g. concrete cover, diffusion characteristics etc.). The sustainability performance requirements and limit states criteria should be incorporated in international standards in future. All these should enable the classification of concrete not only according to its compressive strength and workability class as it is now, but also according to required compacting way (whether the concrete needs compaction work or is self-compacting), architectural and aesthetic aspects, as well as the environmental impact.

3D printing overall technology is still relative in its infancy (Bukvić *et al.*, 2020) when talking about the concrete construction industry. Despite the current flaws, 3D printing is in position to be a viable solution that offers key benefits for our building's future and its fast development is expected. This robotic technology can enable us to build habitats on Mars or Luna. In fact, the mixture compositions for Luna and Mars concrete are already prepared. The world is not enough.

REFERENCES

Abbas, S., Nehdi, M. and Saleem, M. (2016). Ultra-high performance concrete: Mechanical performance, durability, sustainability and implementation challenges. *International Journal of Concrete Structures and Materials*, *10*(3), 271–295. doi: 10 .1007 / s40069 -016 -0157 -4

Aghdasi, P. and Ostertag, C. P. (2020). Tensile fracture characteristics of green ultra-high performance fiber-reinforced concrete (G-UHP-FRC) with longitudinal steel reinforcement. *Cement and Concrete Composites*, *114*, 103749. doi: 10 .1016 / j .cemconcomp .2020 .103749

Azmee, N. M. and Shafiq, N. (2018). Ultra-high performance concrete: From fundamental to applications. *Case Studies in Construction Materials*, *9*, e00197. doi: 10 .1016 / j .cscm .2018 .e00197

Bielak, J., Adam, V., Hegger, J. and Classen, M. (2019). Shear capacity of textile-reinforced concrete slabs without shear reinforcement. *Applied Sciences*, *9*(7), 1382. doi: 10 .3390 / app9071382

Bukvić, O., Radonjanin, V., Malešev, M. and Laban, M. (2020). Basic fresh-state properties of extrusion-based 3D printed concrete. *Građevinski materijali i konstrukcije*, *63*(4), 99–117. doi: 10 .5937 / GRMK2004099B

CEMBUREAU. (2017). *Activity report 2017, Cembureau – The European Cement Association.* http :// www .cembureau .eu / library / reports /. Brussels, Belgium.

CEMBUREAU. (2018). *Activity report 2018, Cembureau – The European Cement Association.* http :// www .cembureau .eu / library / reports /. Brussels, Belgium.

CEMBUREAU. (2019). *Activity report 2019, Cembureau – The European Cement Association.* http :// www .cembureau .eu / library / reports /. Brussels, Belgium.

CEMBUREAU. (2020). *Activity report 2020, Cembureau – The European Cement Association.* http :// www .cembureau .eu / library / reports /. Brussels, Belgium.

CEN. (2004). *En 1992-1-1 (eurocode 2): Design of concrete structures – part 1-1: General rules, rules for buildings, bridges and civil engineering structures.* CEN, Brussels.

Dávila-Pompermayer, R., Lopez-Yepez, L.G., Valdez-Tamez, P., Juárez, C.A. and Durán-Herrera, A. (2020). Lechugilla natural fiber as internal curing agent in self compacting concrete (SCC): Mechanical properties, shrinkage and durability. *Cement and Concrete Composites*, *112*, 103686. doi: 10 .1016 / j .cemconcomp .2020 .103686

Di Prisco, M., Plizzari, G. and Vandewalle, L. (2009). Fibre reinforced concrete: new design perspectives. *Materials and structures*, *42*(9), 1261–1281. doi: 10 .1617 / s11527 -009 -9529 -4

Ferdosian, I. and Camões, A. (2021). Mechanical performance and post-cracking behavior of self-compacting steel-fiber reinforced eco-efficient ultra-high performance concrete. *Cement and Concrete Composites*, *121*, 104050. doi: 10 .1016 / j .cemconcomp .2021 .104050

Fib. (2009). *Structural Concrete: Textbook on behaviour, design and performance* (Vol. 1). fib, Lausanne.

Fib. (2013). *fib Model Code for concrete structures 2010 (MC 2010)*. fib, Lausanne.

Flietstra, J. C. (2011). *Creep and shrinkage behavior of ultra high-performance concrete under compressive loading with varying curing regimes* (Master's thesis, Michigan Technological University). doi: 10 .37099 / mtu .dc .etds / 236

Hegger, J. and Voss, S. (2008). Investigations on the bearing behaviour and application potential of textile reinforced concrete. *Engineering structures*, *30*(7), 2050–2056. doi: 10 .1016 / j .engstruct .2008 .01 .006

Hossain, M., Zhumabekova, A., Paul, S. C. and Kim, J. R. (2020). A review of 3D printing in construction and its impact on the labor market. *Sustainability*, *12*(20), 8492. doi: 10 .3390 / su12208492

IEA, WBCSD. (2018). *Technology roadmap low-carbon transition in the cement industry.* France.

ISO/TC 071. (2016). *Strategic Business Plan ISO/TC 071 (ISO Technical Committee for concrete, reinforced concrete and prestressed concrete)*.

John, V. M., Damineli, B. L., Quattrone, M. and Pileggi, R. G. (2018). Fillers in cementitious materials—experience, recent advances and future potential. *Cement and Concrete Research*, *114*, 65–78. doi: 10 .1016 / j .cemconres .2017 .09 .013

Labib, W. A. (2020). Evaluation of hybrid fibre-reinforced concrete slabs in terms of punching shear. *Construction and Building Materials*, *260*, 119763. doi: 10 .1016 / j .conbuildmat .2020 .119763

Li, N., Zhan, H., Yu, X., Tang, W. and Xue, Q. (2021). Investigation of the aging behavior of cellulose fiber in reclaimed asphalt pavement. *Construction and Building Materials*, *271*, 121559. doi: 10 .1016 / j .conbuildmat .2020 .121559

Li, P.P., Brouwers, H.J.H., Chen, W. and Yu, Q. (2020). Optimization and characterization of high-volume limestone powder in sustainable ultra-high performance concrete. *Construction and Building Materials*, *242*, 118112. doi: 10 .1016 / j .conbuildmat .2020 .118112

Li, Y., Tan, K. H. and Yang, E.-H. (2019). Synergistic effects of hybrid polypropylene and steel fibers on explosive spalling prevention of ultra-high performance concrete at elevated temperature. *Cement and Concrete Composites*, *96*, 174–181. doi: 10 .1016 / j .cemconcomp .2018 .11 .009

Ludwig, H.M. (2001). Dauerhaftigkeit selbstverdichtender betone. *Selbstverdichtender Beton, König, Holschemacher, Dehn (Hrsg.), Bauwerk Verlag, Berlin*, 11–23.

Marinković, S., Carević, V. and Dragaš, J. (2021). The role of service life in Life Cycle Assessment of concrete structures. *Journal of Cleaner Production*, *290*, 125610. doi: 10 .1016 / j .jclepro .2020 .125610

Marinković, S., Dragaš, J., Ignjatović, I. and Tošić, N. (2017). Environmental assessment of green concretes for structural use. *Journal of Cleaner Production*, *154*, 633–649. doi: 10 .1016 / j .jclepro .2017 .04 .015

Meng, K., Xu, L. and Chi, Y. (2021). Experimental investigation on the mechanical behavior of hybrid steel-polypropylene fiber reinforced concrete under conventional triaxial cyclic compression. *Construction and Building Materials*, *291*, 123262. doi: 10 .1016 / j .conbuildmat .2021 .123262

Mitrović, S., Popović, D., Tepavčević, M. and Zakić, D. (2021). Physical-mechanical properties and durability of ultra-high performance concrete (UHPC). *Građevinski materijali i konstrukcije*, *64*(2), 109–117. doi: 10 .5937 / GRMK2102109M

Müller, H. S. (2007). Zum baustoff der zukunft. In T. Fletcher and A. Deletić (Eds.), (pp. 195–221). Beuth Verlag, Berlin, Germany.

Muller, H. S. and Kvitsel, V. (2006). Schwerpunkte-panorama-betontechnologie-kriechen und schwinden von hochleistungsbetonen. *Beton-Hamburg*, *56*(1-2), 36–42.

Ozawa, K., Maekawa, K. and Okamura, H. (2004). High performance concrete with high filling capacity. In E. Vasquez (Ed.), *Admixtures for concrete-improvement of properties: Proceedings of the international rilem symposium* (pp. 51–62). Chapman & Hall, London, UK.

Portal, N. W. (2015). *Usability of textile reinforced concrete: Structural performance, durability and sustainability* (Doctoral dissertation, Chalmers University of Technology, Gothenburg, Sweden). https :// www .researchgate .net / publication / 284038837 _Usability _of _Textile _Reinforced _Concrete _Structural _Performance _Durability _and _Sustainability.

Radić, J., Kindij, A. and Mandić, A. (2008). History of concrete application in development of concrete and hybrid arch bridges. In *Proc. of The Chinese-Croatian Joint Colloquium on Long Arch Bridges, Croatian University of Zagreb, University of FuZhou, Chorwacja* (pp. 9–118). Brijuni, Croatia.

Radović, A., Marinković, S. and Savić, A. (2021). Compressive strength of green concrete with low cement and high filler content. *Građevinski materijali i konstrukcije*, *64*(2), 93–108. doi: 10 .5937 / GRMK2102093R

Rauf, M., Khaliq, W., Khushnood, R. A. and Ahmed, I. (2020). Comparative performance of different bacteria immobilized in natural fibers for self-healing in concrete. *Construction and Building Materials*, *258*, 119578. doi: 10 .1016 / j .conbuildmat .2020 .119578

Raza, S. S., Qureshi, L. A., Ali, B., Raza, A., Khan, M. M. and Salahuddin, H. (2020). Mechanical properties of hybrid steel–glass fiber-reinforced reactive powder concrete after exposure to elevated temperatures. *Arabian Journal for Science and Engineering*, *45*(5), 4285–4300. doi: 10 .1007 / s13369 -020 -04435 -4

Russell, H. G. and Graybeal, B. A. (2006). *Material property characterization of ultra-high performance concrete* (Tech. Rep.). Publication no. FHWA-HRT-06-103, Federal Highway Administration (FHWA), U.S. Department of transportation. https :// www .fhwa .dot .gov. (Accessed 2 August 2021)

Scheerer, S., Chudoba, R., Garibaldi, M. P. and Curbach, M. (2017). Shells made of textile reinforced concrete–applications in Germany. *Journal of the international association for shell and spatial structures*, *58*(1), 79–93. doi: 10 .20898 / i .iass .2017 .191 .846

Scheerer, S., Zobel, R., Müller, E., Senckpiel-Peters, T., Schmidt, A. and Curbach, M. (2019). Flexural strengthening of RC structures with TRC—Experimental observations, design approach and application. *Applied Sciences*, *9*(7), 1322. doi: 10 .3390 / app9071322

Shaeffer, R. E. (1992). *Reinforced concrete: preliminary design for architects and builders*. McGraw-Hill College.

Sunayana, S. and Barai, S. V. (2017). Recycled aggregate concrete incorporating fly ash: Comparative study on particle packing and conventional method. *Construction and Building Materials*, *156*, 376–386. doi: 10 .1016 / j .conbuildmat .2017 .08 .132

UN Population Division. (2021). *World population prospects: The 2019 revision, united nations department of economic and social affairs.* https :// population .un .org / wpp / Download / Standard / Population /. (accessed 7 July 2021)

UNEP (United Nations Environment Programme). (2017). *Eco-efficient cements: Potential economically viable solutions for a low-CO2 cement based materials industry*. Paris.

World Steel Association. (2021). *Worldsteel in figures 2021*. https :// www .worldsteel .org / steel -by -topic / statistics / World -Steel -in -Figures .html. Brussels, 2015. (accessed 7 July 2021)

Yazıcı, H. (2007). The effect of curing conditions on compressive strength of ultra high strength concrete with high volume mineral admixtures. *Building and environment*, *42*(5), 2083–2089. doi: 10 .1016 / j .buildenv .2006 .03 .013

Gerard Heuvelink

Prof. Gerard Heuvelink is a senior researcher with ISRIC – World Soil Information and a special professor in Pedometrics and Digital Soil Mapping with the Soil Geography and Landscape group of Wageningen University. He is a former president of the Netherlands Soil Science Society (2004-2007) and Web of Science Highly Cited Researcher in 2019 and 2020.

Contact information: e-mail: gerard.heuvelink@isric.org

https://orcid.org/0000-0003-0959-9358

Prof. Heuvelink holds an MSc in Applied Mathematics (1987) and a PhD in Environmental Sciences (1993). His PhD research on error propagation in spatial modelling with GIS marks the start of his scientific career in spatial uncertainty analysis, with applications in the soil domain as a main focus. He has also contributed substantially to the development of pedometrics, which is the part of soil science that uses mathematical and statistical methods for the study of the distribution and genesis of soils. He chaired the Pedometrics Commission of the International Union of Soil Sciences from 2003 to 2006 and was president of the Netherlands Soil Science Society from 2004 to 2007. He received the Richard Webster medal for the best body of work that advanced pedometrics in 2014 and chaired the Organising Committee of the 25-th anniversary Pedometrics Conference in Wageningen in 2017.

Prof. Heuvelink is also the recipient of the Peter Burrough medal of the International Spatial Accuracy Research Organisation (2020). He has published over 150 articles on geostatistics, spatial uncertainty analysis and pedometrics in peer-reviewed international scientific journals and was recognized by the Web of Science Group as a Highly Cited Researcher in 2019 and 2020, with more than 8500 citations ($h = 40$).

He is a Deputy Editor of the European Journal of Soil Science and editorial board member of Geoderma, Spatial Statistics and three more ISI journals. Over the years, he supervised more than 35 MSc students and 20 PhD students. He teaches courses in statistics, geostatistics, pedometrics and spatial uncertainty analysis at Wageningen University and the Graduate School for Production Ecology & Resource Conservation. Prof. Heuvelink is also a visiting professor of the Institute of Geographical Sciences and Natural Resources Research, University of Chinese Academy of Sciences, China.

Spatial interpolation for mapping the environment

Gerard Heuvelink

Soil Geography and Landscape group, Wageningen University, Netherlands

Summary

A common problem in the Earth and environmental sciences concerns the spatial interpolation of variables from point measurements. Examples of such variables are precipitation, temperature, air quality, soil and water pollution, and aboveground biomass. The history of spatial interpolation goes back to the early days of geoinformatics, when basic algorithms such as nearest neighbour interpolation and inverse distance weighted interpolation were developed. Geostatistics took it a step further by developing kriging interpolation, which weighs the measurements in accordance with their degree of spatial correlation and has the additional advantage that it quantifies the interpolation error. Over the years many different kriging algorithms have been developed. Currently regression kriging is very popular. It typically produces more accurate maps than ordinary kriging because in addition to information derived from nearby point measurements, it also benefits from information contained in explanatory environmental variables, as for example derived from remote sensing. Extensions of regression kriging to space-time interpolation have also been made. Machine learning methods, such as random forest, gradient boosting and artificial neural networks, have recently become a serious competitor of kriging. These methods rely solely on the relation between the dependent and explanatory variables and are very flexible in their model structure. This can improve the map accuracy substantially, if at least sufficient training data are available. A very recent improvement of machine learning for spatial interpolation is given by a technique known as Random Forest Spatial Interpolation. This technique closes the spatial interpolation circle, by revisiting the original ideas and incorporating measurements from neighbouring locations in the set of explanatory variables of a random forest model. In this chapter I review the historic development of spatial interpolation techniques and highlight the contributions to this field by researchers of the Geodesy and Geoinformatics group of the Faculty of Civil Engineering, University of Belgrade.

1. INTRODUCTION

Many variables in the Earth and environmental sciences can only be measured at a finite number of point locations, while researchers and end users may need values of these variables for all locations in an area of interest. Examples of such variables are groundwater level, soil organic matter content and the grade of valuable minerals such as copper, nickel and gold in ore deposits. Other variables can more easily be measured exhaustively in space through remote sensing, but these indirect measurements are often only proxies of the true variables and ground-truth observations are still needed for mapping. Examples of such variables are precipitation, temperature, aboveground biomass and air quality. In all of these cases there is a need for spatial interpolation.

Spatial interpolation has a long history that goes back to the early days of geo-information science. Methods for spatial interpolation have evolved substantially during the past decades and have become much more refined and elaborate, among others through the use of remote sensing information as 'covariates' in addition to point observations. All developments aim to use the available information optimally to obtain a map of the variable of interest that is as close as possible to the truth. Thus, the quest for 'optimal spatial interpolation' has been a topic of research by geographers and environmental scientists for many decades and continues as new methods are developed and new data come available. In particular, research is also directed to quantification of the spatial interpolation error.

The purpose of this chapter is to give a brief historic account of spatial interpolation methods and provide an outlook of future developments. Deterministic interpolation methods are the oldest and presented first. Next, geostatistical interpolation is discussed and finally, machine learning for spatial interpolation is addressed. Contributions to these more recent developments by researchers of the Geodesy and Geoinformatics group of the Faculty of Civil Engineering, University of Belgrade will be emphasized.

This chapter does not aim to provide a comprehensive and detailed account of spatial interpolation methods. There are many text books that provide detailed explanations and illustrations of the various methods discussed, such as Li and Heap (2008) and Burrough *et al.* (2015).

2. DETERMINISTIC SPATIAL INTERPOLATION METHODS

It is useful to start with a mathematical formulation of the basic spatial interpolation problem. Let $z = \{z(x), x \in D\}$ be the variable of interest defined in a geographic domain D, where x refers to geographic location and is taken here as a two-dimensional coordinate. We will further assume that a suitable geographic projection is used, such that distances between points are properly defined and satisfy basic conditions of a normed, metric space.

Let z be measured at a finite number of locations x_i, $i = 1, \ldots, n$, and let the goal be to predict $z(x_0)$, where x_0 is an arbitrary location in D, from the observations $z(x_i)$:

$$\hat{z}(x_0) = f\Big(z(x_1), \ldots, z(x_n)\Big) \tag{1}$$

Different methods can now be used to come up with a function f.

2.1. Nearest neighbour spatial interpolation

Perhaps the simplest spatial interpolation method is to let the prediction at location x_0 be equal to the nearest observation. This may be written as a weighted linear interpolation model:

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i \cdot z(x_i) \tag{2}$$

with all weights λ_i set to zero except for the nearest observation, which has a weight equal to one. Figure 1 gives a graphical illustration for a simple case of $n = 8$ observations. The interpolation results in a map with so-called 'Thiessen polygons' or 'Voronoi diagrams'. The interpolated value is constant within a polygon and makes discrete jumps when crossing the border between two polygons (i.e., when the nearest observation changes). This discontinuity may not be very realistic for real-world phenomena but nearest neighbour maps do support a visual exploratory analysis of spatial data and can also be used for variables measured on a categorical scale, such as soil type or land use.

2.2. Inverse distance weighting

A slightly more advanced approach is to use Eq. (2) and let all weights have a positive value that is directly derived from the distance between the observation and prediction points:

$$\lambda_i = \frac{c}{d_i^p} \tag{3}$$

where $d_i = \|x_0 - x_i\|$ is the distance between x_0 and x_i, p is a power (often taken as 2, i.e. inverse *squared* distance weighting) and c a normalizing constant to ensure that the weights add up to one. Figure 1 shows an example. While being very simple and straightforward, inverse distance interpolation is still often used in geo-information science and can in fact have remarkable good performance in specific cases (Sekulić, Kilibarda, Heuvelink *et al.*, 2020).

2.3. Splines

Splines are piecewise polynomials whose parameters are chosen such that a smooth surface is obtained that passes through the measurements. In the case of cubic splines, third-order polynomials are used and their parameters are chosen such that the resulting surface is continuous and twice differentiable, guaranteeing smooth surfaces that have neither discontinuities nor abrupt changes in the surface slope. Splines are often used in one-dimensional applications but also work in two-dimensional space and are therefore routinely used in spatial interpolation. There are many different variants of splines, one of which is thin plate spline, which aims for a trade-off between the smoothness (integral of the squared second derivative over the geographic domain) and the goodness (closeness of the interpolated surface to the measurements). Figure 1 shows an example. Note that sacrificing goodness for smoothness implies that the method no longer yields an exact interpolator (i.e., interpolations at measurement locations may differ from the measurements themselves).

2.4. Other methods

Other commonly used deterministic spatial interpolation methods are radial basis functions, which have similarity to splines; triangular irregular networks, which divide the study area in triangles that are modelled as linear planes (e.g. Delaunay triangulation); and trend surfaces, which fit a single polynomial to the entire geographic domain.

Note that all deterministic interpolation methods are inspired by the first law of geography, which states that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). This explains why in all these methods neighbouring observations get larger weights than more distant observations. But otherwise the choice of weights is still fairly arbitrary, and not derived from theory or optimized in any way (except if one uses cross-validation to calibrate the interpolation model, see Section 5).

3. GEOSTATISTICAL INTERPOLATION METHODS

Geostatistics has its roots in mining (e.g. Journel and Huijbregts, 1978), but has become increasingly popular in other scientific domains since the 1980s and is nowadays

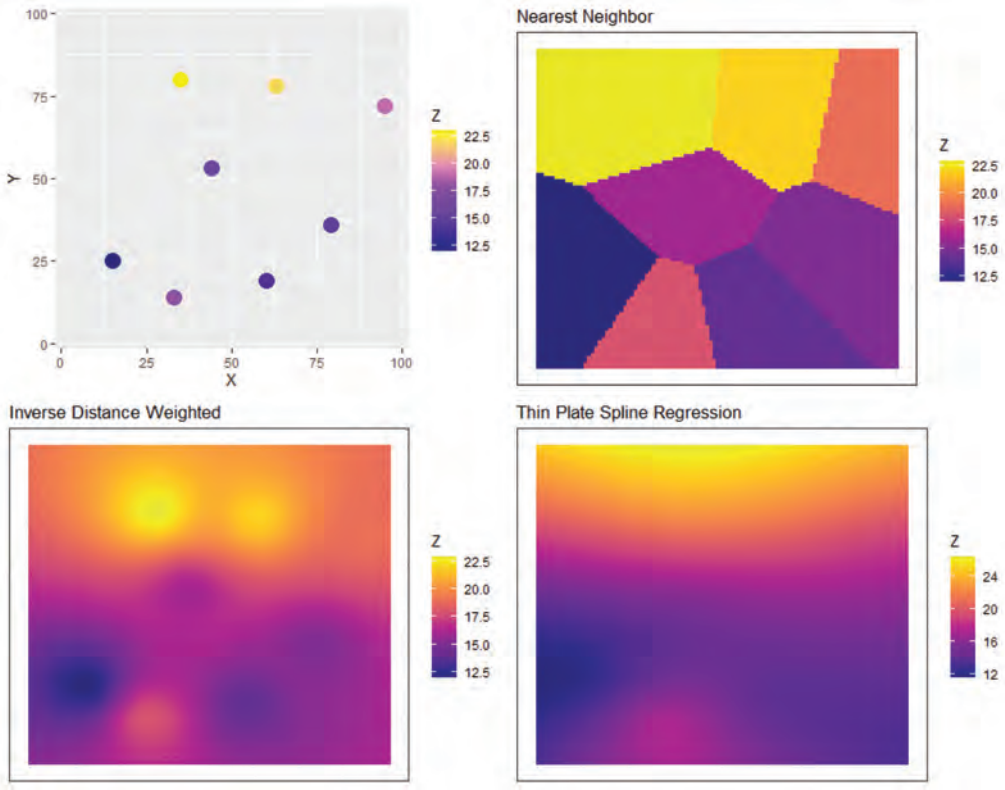


Figure 1. Simplified synthetic example of deterministic spatial interpolation. Top left panel shows the measurements, the other panels show interpolations with three different methods.

routinely used in various fields, such as soil science, climatology, hydrology and geography. Geostatistical interpolation is known as kriging, named after the South-African mining engineer Danie Krige (1919-2013). A fundamental difference between deterministic spatial interpolation methods and kriging is that the latter is model-based, meaning that the spatial interpolation is derived from a statistical model of spatial variation. Given this model, the spatial interpolation is no longer arbitrary but 'optimal', meaning that it has the smallest expected squared prediction error among all possible interpolations.

The basic geostatistical model assumes that the variable of interest z is a realisation of a random field Z, which typically satisfies:

$$Z(x) = m(x) + \varepsilon(x) \tag{4}$$

where m is a spatial 'trend' and ε a stochastic residual. Note the distinction between lower case $z(x)$ (a realisation, a deterministic number) and upper case $Z(x)$ (a random variable, characterized by a probability distribution). In the case of ordinary kriging, m is assumed to be constant. The stochastic residual is then taken as a random variable that has zero mean and constant variance, and whose spatial correlation is characterized by a semivariogram:

$$\gamma(h) = \frac{1}{2} E\Big[\big(Z(x) - Z(x+h)\big)^2\Big] \tag{5}$$

where h is the separation distance between two locations and E refers to mathematical expectation.

Note that Eq. (5) assumes that the semivariogram γ only depends on the distance between locations and not on the locations themselves. This so-called second-order stationarity assumption is invoked in order to be able to estimate the semivariogram from measurements. But even then semivariogram estimation requires at least 100 measurements (Webster and Oliver, 2007, Section 6.1.2), implying that geostatistical interpolation cannot be used in cases where only a few dozens of measurements are available. Note also that the separation distance h in Eq. (5) is a vector that has two (or three) geographical coordinates. Modelling is often simplified by assuming isotropy, in which only the Euclidean distance matters, i.e. by assuming $\gamma(h) = \gamma(\|h\|)$.

If sufficient measurements are available then the semivariogram can be estimated from them using the methods of moments or maximum likelihood estimation (Webster and Oliver, 2007). This is a key step in any geostatistical analysis, because the semivariogram is needed in kriging.

3.1. Ordinary kriging

Ordinary kriging predicts $Z(x_0)$ using Eq. (2), but unlike inverse distance interpolation the weights λ_i are not derived directly from the distances between observation and prediction points, but from their degree of spatial

correlation, as characterized by the semivariogram. This is well-explained in text books (e.g. Chiles and Delfiner, 2012; Webster and Oliver, 2007). Similar to inverse distance weighted interpolation, closer measurements typically have larger weights than distant measurements (because they have a higher correlation), but the spatial configuration of the measurements also matters. Clustered measurements tend to get lower weights than isolated measurements, which intuitively makes sense because clustered data have redundant information.

Ordinary kriging typically produces a smooth interpolated surface that reproduces the large scale pattern in the measurements. This is particularly the case when the semivariogram exhibits a substantial amount of short-distance spatial variation (i.e., a nugget variance). Thus, ordinary kriging usually does not show 'pockets' or 'islands' of high or low predictions around a measurement location as obtained in inverse distance interpolation. Figure 2 shows an example of ordinary kriging of the long-term annual precipitation in Turkey from measurements at 205 meteorological stations.

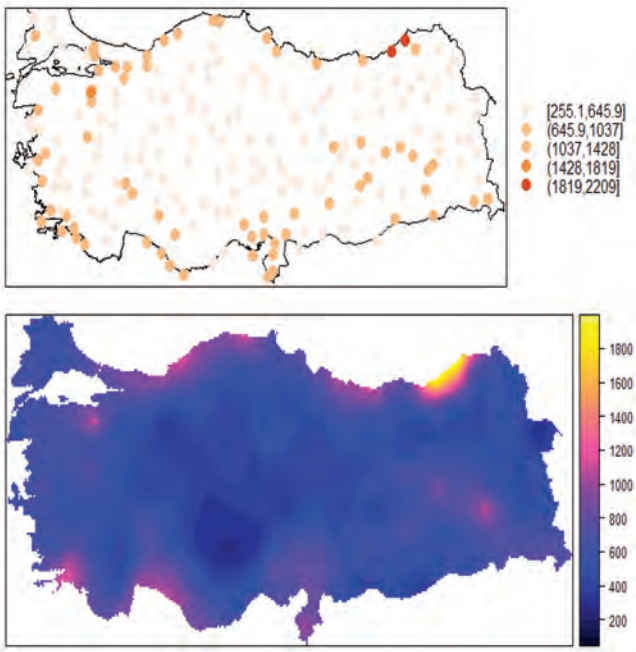


Figure 2. Ordinary kriging of long-term annual precipitation in Turkey. Precipitation measurements (mm) at 205 stations (top) and ordinary kriging map (bottom).

An important advantage of geostatistical interpolation over deterministic spatial interpolation is that it quantifies the interpolation error, by means of the kriging standard deviation. Assuming that the kriging error is normally distributed, this allows to compute the lower and upper limits of a prediction interval, such as shown for ordinary kriging in Figure 3. Note that the kriging standard deviation, also shown in Figure 3, tends to be relatively small close to sampling locations and large further away from them. Figure 3 also indicates that the ordinary kriging map of annual precipitation shown in Figure 2 is very uncertain, which is no surprise given that only 205 measurements were used to predict annual precipitation for an entire country.

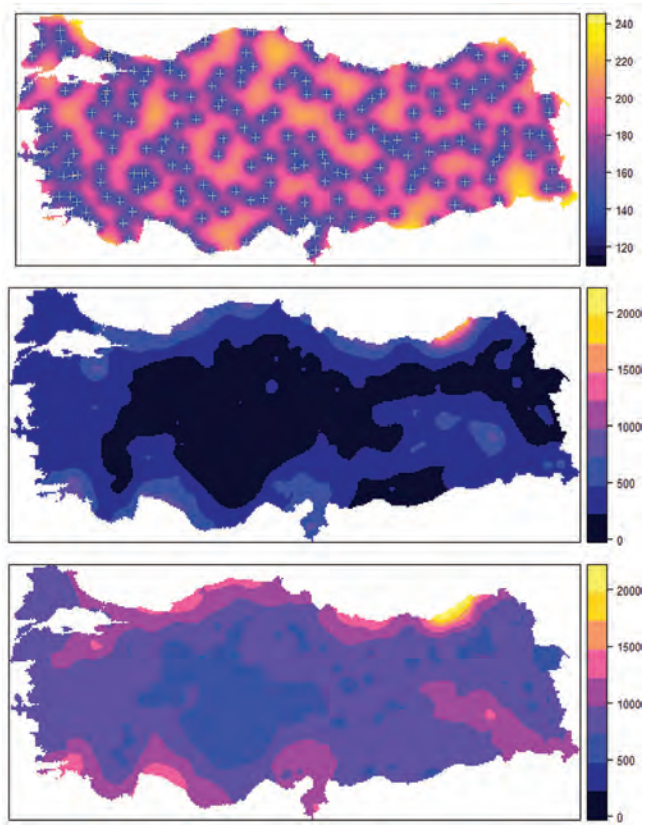


Figure 3. Interpolation uncertainty derived from ordinary kriging of long-term annual precipitation (mm) in Turkey. Ordinary kriging standard deviation (top), and lower (middle) and upper (bottom) limits of a 90% prediction interval. White crosses in standard deviation map are measurement locations.

Ordinary kriging was popular in the 1980s and 1990s, but nowadays its use has become almost obsolete. This is because it only relies on the measurements $z(x_i)$, as did the deterministic interpolation methods presented in Section 1. But with the advent of remote sensing there is often much more information available that can be used to improve spatial interpolation. For example, we know that soil pH depends on soil forming factors such as parent material, climate, terrain morphometry and land use (Jenny, 1994), and since information or 'proxies' about these factors are exhaustively available as maps, they should be used to improve the interpolation. This can be done in geostatistics using a technique known as regression kriging.

3.2. Regression kriging

Regression kriging is also known as kriging with external drift and universal kriging, with subtle differences between these methods that are not important here. The

crucial difference between regression kriging and ordinary kriging is that in regression kriging we no longer assume that the trend in Eq. (4) is constant. Instead we assume that it is a function of explanatory variables, also known as 'covariates'. Using the example of annual precipitation in Turkey, it makes sense to include elevation and distance to the coast as covariates. The top panel in Figure 2 shows that annual precipitation tends to be higher closer to the coast, and from climatology we also know that there is a correlation between annual precipitation and elevation. Thus, we might arrive at the following model:

$$Z(x) = \beta_0 + \beta_1 \cdot f_1(x) + \beta_2 \cdot f_2(x) + \varepsilon(x) \qquad (6)$$

where the β_i are regression coefficients and the f_i are covariates (i.e., elevation and distance to coast). The stochastic residual in Eq. (6) may be spatially correlated and is modelled by a semivariogram as in ordinary kriging, but it will typically have a smaller variance because part of the spatial variation in the dependent variable is explained by the covariates. Figure 4 shows the regression kriging prediction of the annual precipitation in Turkey. Note that the ordinary kriging and regression kriging maps are fairly similar because in this specific case the covariates do not explain much of the spatial variation in annual precipitation. However, Figure 4 clearly shows that the regression kriging map is less smooth and has fine-scale spatial detail. These are the result of the influence of the elevation map.

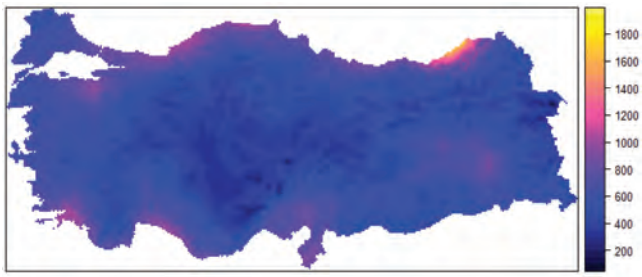


Figure 4. Regression kriging prediction of long-term annual precipitation (mm) in Turkey, using elevation and distance to coast as covariates.

Regression kriging has become the geostatistical interpolation 'workhorse' since the start of this century. More elaborate regression kriging models can make use of many covariates, such as outputs of remote sensing instruments and mechanistic models. In such case, careful attention should be paid to model selection, that is the procedures used to decide which covariates are included in the trend and which are not, and how.

The Geoinformatics group of the University of Belgrade has contributed much to the development and application of regression kriging. Among others, they addressed regression kriging in space-time and applied this to global estimation of daily temperatures (Kilibarda *et al.*, 2014). Figure 5 shows an example.

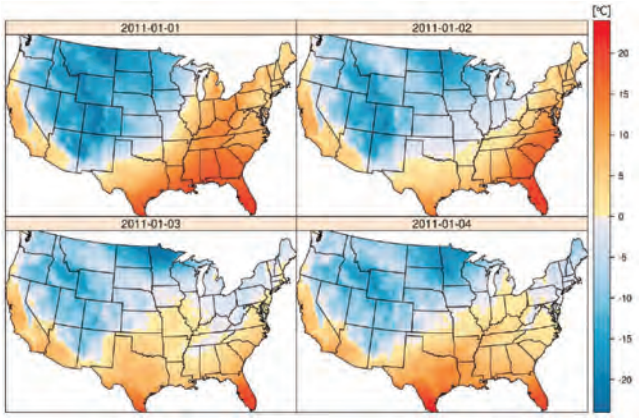


Figure 5. Regression kriging maps of mean daily temperature for the first four days in January 2011, covering coterminous USA in Albers equal-area projection (from Kilibarda *et al.*, 2014).

4. MACHINE LEARNING FOR SPATIAL INTERPOLATION

While geostatistical interpolation can incorporate covariates to help explain spatial variation, much emphasis is still on modelling the (residual) spatial variation through a semivariogram. However, with the rapid development of remote sensing and other instruments (such as UAVs), the number of covariates and their predictive power has grown that much over the past decades that focus has shifted from modelling the stochastic residual to improved modelling of the trend. This development received a boost with the maturing of machine learning models (Hastie *et al.*, 2009; James *et al.*, 2013), which proved to be very flexible in extracting relevant information from the covariates. It is beyond the scope of this chapter to provide a comprehensive review of machine learning algorithms for spatial interpolation. Instead we describe the main principles, name a few techniques that are often used, and illustrate their use with a few examples.

Machine learning, also termed statistical learning, refers to a set of tools for modelling relationships in large and complex data sets. It models the relation between a dependent and explanatory variables in a very flexible way, much beyond simple linear models such as the trend in Eq. (6), and uses the data to 'learn' the relationship. It often makes use of tree-based structures but not always, and can be used both for classification and regression. It does not explicitly take the spatial context into account, but merely predicts the dependent from the covariates, be it spatial variables such as temperature, soil organic matter and ozone concentration, or non-spatial variables such as stock market index, cancer type or fraud risk.

Common machine learning methods used for spatial interpolation are random forest, support vector machines, artificial neural networks and gradient boosting (James *et al.*, 2013). These methods have also been used for mapping by the Geoinformatics group of Belgrade Univer-

sity (e.g Bajat, Hengl *et al.*, 2011; Kovačević *et al.*, 2010; Marjanović *et al.*, 2011; Samardžić-Petrović *et al.*, 2016). Since machine learning methods are very flexible, there is a risk of overfitting and hence a data set is typically split in three: training, validation and test data. The training data are used to calibrate the model, the validation data are used to optimize meta-parameters and prevent overfitting, and the test data are used to evaluate the prediction performance (see also Section 5 below). The splitting of the data set into three subsets is often done repeatedly, in a nested k-fold cross-validation mode (e.g. Pejović *et al.*, 2018).

Figure 6 shows an example from Čeh et al. (2018) where the random forest algorithm was used to interpolate prices and sales ratios of apartments in the city of Ljubljana. Random forest and gradient boosting have also been used to map soil properties at continental and global scales (e.g. Hengl *et al.*, 2017; 2021; Poggio *et al.*, 2021). Figure 7 shows an example. The use of machine learning for environmental modelling and mapping is also a key research theme of the Geoinformatics group of Belgrade University, among others demonstrated by the GeoMLA conference held in Belgrade in 2016 (Tadić-Percec *et al.*, 2018).

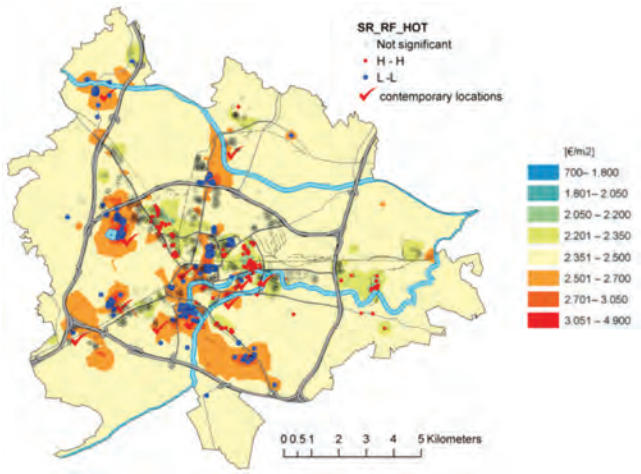


Figure 6. Map of sales ratio of apartments in the city of Ljubljana as obtained using random forest modelling and mapping (from Čeh et al., 2018).

It was mentioned before that an important advantage of geostatistical interpolation over deterministic interpolation is that it quantifies the interpolation error. Machine learning seems less-developed in this respect, although there are some machine learning algorithms that produce spatially explicit measures of the prediction accuracy. One such method is Quantile Regression Forests (Meinshausen, 2006), which derives the quantiles of the predictive probability distribution, so that prediction error standard deviations and prediction intervals can be computed. Figure 8 gives an example. Note that machine learning and kriging can also be combined by first using machine learning to model the trend, and next modelling the residual with ordinary or simple kriging. The kriging step then quantifies the interpolation error, although one should be aware that this approach ignores uncertainty about the trend.

In spite of the recent development, good performance and many successes obtained with machine learning for spatial mapping, in essence it is not a spatial model and does not take the geographic locations and spatial context into account when calibrating a model and making predictions. This suggests that there should be ways to further improve these methods. Several approaches have been proposed and tested, but arguably the most innovative idea to let machine learning better exploit spatial context and neighbourhood information was presented in Sekulić, Kilibarda, Heuvelink *et al.* (2020). The idea is simply to include the nearest observations and their distances from the prediction location as covariates of the machine learning model (Figure 9). It is interesting to observe that the nearest observations are often picked up by the machine learning algorithm as important covariates and thus improve map accuracy. In a way this brings us back to the deterministic interpolation methods discussed in Section 1, but then in a much more flexible way that can also benefit from external, environmental covariates.

A different approach to incorporate spatial context in machine learning mapping has recently emerged, known as deep learning or convolutional neural networks (Goodfellow *et al.*, 2016). The key element of this approach is that spatial predictions are not only derived from the values of the covariates at the prediction location, but also from covariate values in a local neighbourhood. The method is promising and multiple applications in the environmental sciences have been done (e.g. Ng *et al.*, 2019; Yu and Liu, 2021; **?**, [), but it is computationally intensive and requires a very large training set.

5. CHOOSING BETWEEN INTERPOLATION METHODS AND STATISTICAL VALIDATION OF INTERPOLATED MAPS

Many spatial interpolation algorithms have so far been reviewed in this chapter, which naturally leads to the question which algorithm should be used in any given situation. Clearly the choice between methods depends on many factors, some of which are quite pragmatic. For instance, users who are not familiar with geo-information science, geo-computation and software implementations of machine learning may be reluctant to go for complex machine learning algorithms. Similarly, geostatistical interpolation requires experience with geostatistical theory and methods. If such experience is lacking, it may be best to go for simpler interpolation methods, such as nearest neighbour or inverse distance weighted interpolation.

Another important consideration is data availability. Clearly one cannot use methods that heavily rely on covariate information if such information is not available or

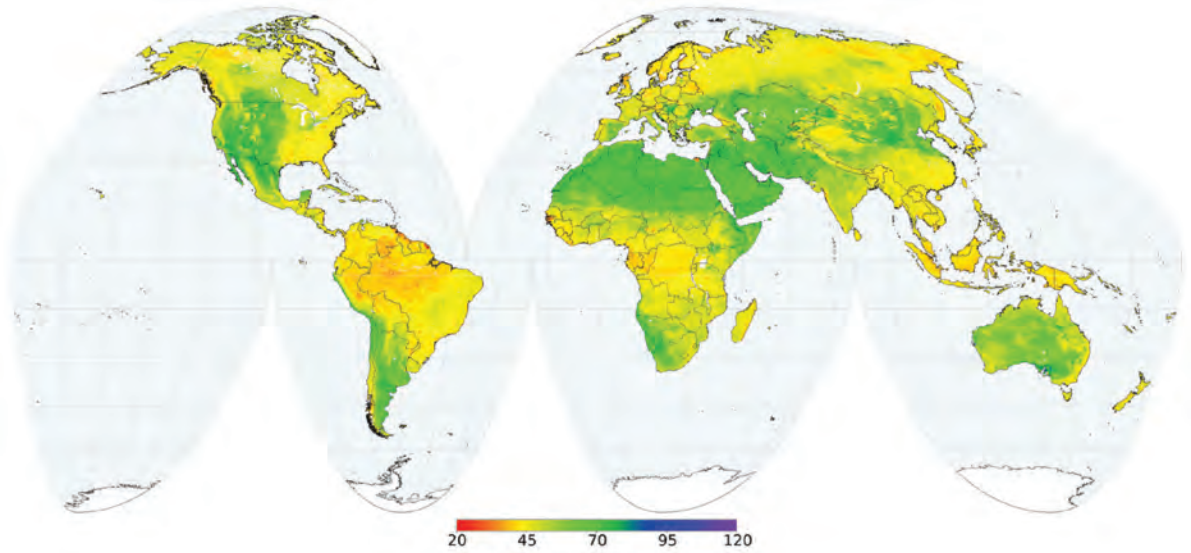


Figure 7. Random forest interpolated map of global soil pH (10 pH) for the 60 to 100 cm depth interval (from Poggio *et al.*, 2021).

useful. This may for instance happen if one maps soil pollution in an urban area (Qiao *et al.*, 2018), although for sufficiently large areas one may usually benefit from climate information, topography and distance to the pollution source (Pejović *et al.*, 2017). Kriging requires sufficient measurements of the dependent variable to be able to estimate the semivariogram reliably, while machine learning methods typically perform well only when a very large data set is available for model calibration. These methods are products of 'big data science', and may work well in data rich cases (for example, we can easily collect thousands of pictures of cats and dogs to let a machine learning model learn to distinguish between cats and dogs), but if every single observation is expensive (such as in ore exploration and groundwater monitoring) then machine learning may not be the best choice.

Computation time may also be a limiting factor. Geostatistical and machine learning methods can be very time consuming (e.g. Sekulić, Kilibarda, Heuvelink *et al.*, 2020) and require state-of-the art hardware and professional infrastructure. Much of this can nowadays be addressed by using high performance computing facilities from external sources such as Amazon and Google, but use of these facilities is costly and requires skilled personnel.

Perhaps the most important consideration when choosing between spatial interpolation methods is their prediction accuracy. As noted before some methods provide an internal measure of the interpolation error (i.e., the

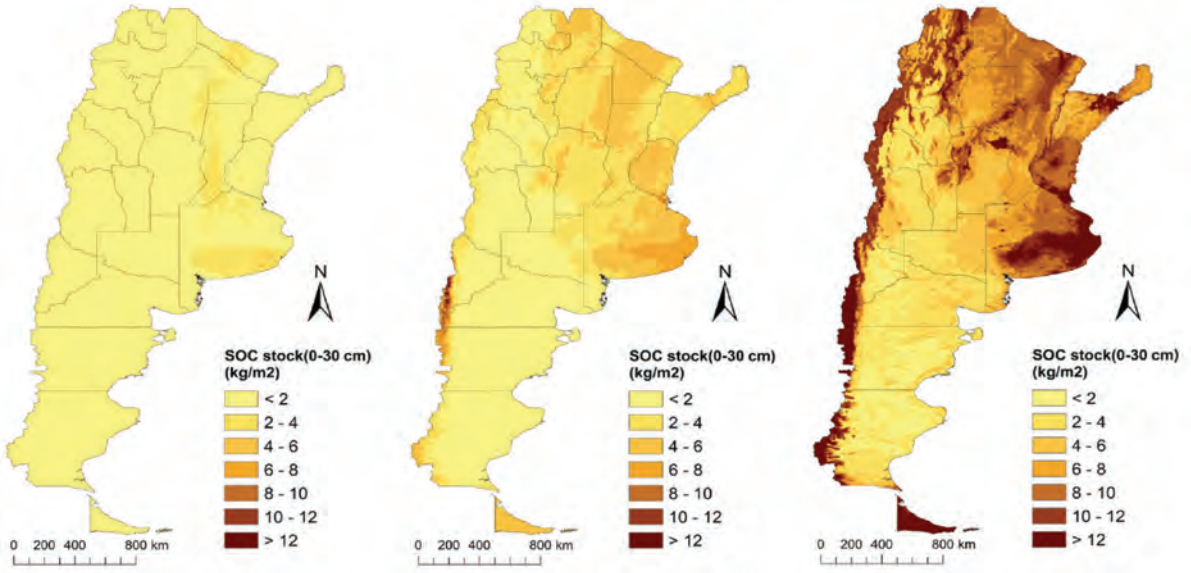


Figure 8. Maps of the 5 (left), 50 (centre) and 95 (right) percentiles of soil organic carbon stock for Argentina for the year 2017 as obtained with Quantile Regression Forest (from Heuvelink *et al.*, 2021).

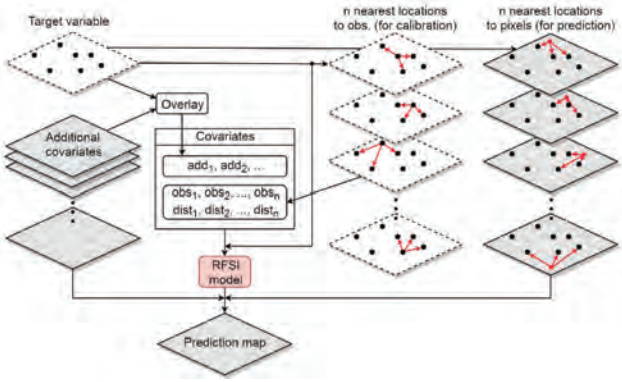


Figure 9. Schematic overview of the Random Forest Spatial Interpolation algorithm (from Sekulić, Kilibarda, Heuvelink *et al.*, 2020).

kriging standard deviation and quantiles as obtained with Quantile Regression Forests), but these are built on assumptions that make them less reliable. Thus, there is also a need for an external accuracy assessment using independent observations obtained with probability sampling.

5.1. Probability sampling and design-based statistical inference for map accuracy evaluation

Accuracy assessment of an interpolated map can best be done using an independent data set obtained with probability sampling (Stehman, 1999; Stehman and Foody, 2009; Wadoux *et al.*, 2021). Taking the simplest case of simple random sampling, unbiased estimates of common map accuracy metrics such as the Mean Error, Mean Absolute Error, Root Mean Squared Error, Map Purity, Model Efficiency Coefficient and the Prediction Interval Coverage Probability can easily be derived from the sample. If these metrics are computed for multiple spatial interpolation methods, then the one that has the best performance may be preferred. Moreover, if probability sampling is used then confidence intervals can also be obtained for all metrics, which allows to test whether the best performing interpolation method has a statistically significant better performance than other interpolation methods.

5.2. Map accuracy assessment with convenience sampling and cross-validation

While map accuracy assessment with independent data obtained with probability sampling is preferred, in many practical cases one does not have the luxury to collect validation data in this way. One is therefore forced to use the existing data for the accuracy assessment, either by data splitting or cross-validation. In data splitting the total data set is (randomly) divided into a subset used for model calibration and prediction and a subset for evaluation of the model performance. Cross-validation does essentially the same, but then repeatedly so that all observations are used both for model calibration and prediction as well as for evaluation of the map accuracy. For instance, in k-fold cross-validation the total data set is randomly divided in k equally sized subsets, after which each fold is put aside and used for accuracy evaluation one by one, each time using the other $k-1$ folds for model calibration and prediction. Figure 10 gives an example of accuracy assessment using leave-one-out cross-validation, where the number of folds k equals the number of observations.

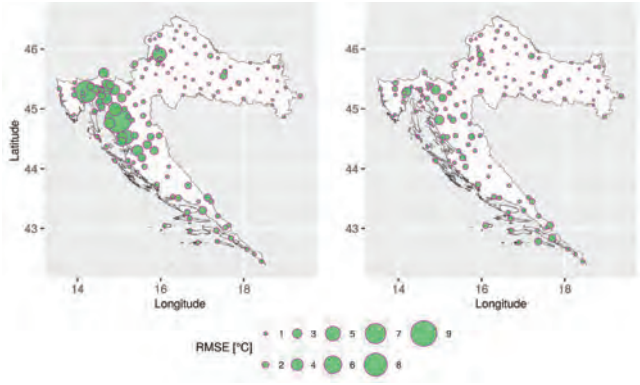


Figure 10. Annual average Root Mean Squared Error (RMSE) of mean daily temperature predictions for Croatia as obtained using spatio-temporal regression kriging and leave-one-out cross-validation (from Sekulić, Kilibarda, Protić *et al.*, 2020).

Cross-validation is often used in practice to evaluate the accuracy of interpolated maps, but one should be aware that there is no guarantee that the so-obtained estimates of the map accuracy metrics are unbiased. In particular, default cross-validation may be criticized if the data are highly spatially clustered, because in such case the estimates will be dominated by the performance in data-rich subareas, where the interpolation error is usually smaller. To avoid over-optimistic estimates of the accuracy metrics a method known as spatial cross-validation has been proposed (e.g Brenning, 2005; Roberts *et al.*, 2017), but these methods start from the wrong premise that spatial autocorrelation must be avoided and tend to produce over-pessimistic results (Wadoux *et al.*, 2021). The spatial mapping research community is still in need for cross-validation methods that provide realistic and reliable estimates of map accuracy metrics.

6. CONCLUSIONS AND OUTLOOK

Spatial interpolation for mapping the environment has come a long way, starting out with simple deterministic methods from the early years of geo-information science to the advanced geostatistical and machine learning methods of today. What can we conclude from these developments and what developments can we expect in the near future?

One important conclusion is that the past decades have shown that spatial interpolation can benefit greatly from

covariate information. We are much beyond the stage that measurements at sampling points are the only source of information. We now have clever algorithms that extract useful information from covariates and in this way explain a large portion of the spatial variation of the dependent variable. This development will likely continue in the future since covariate information will only grow and become more detailed, while data science will continue to refine and develop statistical learners that optimally extract information from covariates. However, it is important to recognize that these approaches can only deliver if there are sufficient observations of the dependent variable as required for model calibration. More research is needed into how many observations are needed to properly train a machine learning model and how this depends on the specific algorithm and its hyperparameters. There will always be cases in which the number of observations is too small to apply machine learning or geostatistical interpolation, so that deterministic interpolation methods are still useful.

Simple deterministic interpolation methods will also stay because advanced spatial interpolation methods require skilled and experienced staff that may not always be available. Even though spatial interpolation tools will become more user-friendly and may render a geostatistical or machine learning interpolated map with just one mouse-click, there is a serious risk of getting erroneous results if these methods are used by non-experts. Some common risks are overfitting and extrapolation. Many machine learning algorithms have in-built routines that protect against overfitting, but not all. For example, the number of effective parameters of a geographically-weighted regression model can be very large and hence it is important to ensure that this is in balance with the number of training data (Bajat, Krunić *et al.*, 2011). Extrapolation can occur both in geographic space and in feature space (Kilibarda *et al.*, 2015). Geographic extrapolation is easily detected but feature space extrapolation is not obvious and is a serious problem that has not yet received sufficient attention. One solution is to exclude subareas where extrapolation in feature space occurs by using the concept of Area of Applicability (Meyer and Pebesma, 2021). Even better would be to develop spatial sampling designs that avoid geographic and feature space extrapolation (e.g. Wadoux *et al.*, 2019).

Given the need for skills it is important that future generations of geo-scientists and geo-engineers are trained in advanced methods of spatial interpolation and the technical implementations. Fortunately many groups across the world have the infrastructure and know-how to teach courses on the subject. One excellent example of that is the Geodesy and Geoinformatics group of the Faculty of Civil Engineering of the University of Belgrade.

Many of the advanced spatial interpolation methods also require high-performance geo-computing facilities and the technical know-how that goes with that. Technical developments are rapid and it is only a matter of time until we can map the global environment at 10 or 1 m resolution. But we may ask ourselves whether our focus should be on increasing spatial resolution or on improving the map accuracy. We should realize that map accuracy may be more important than map resolution and think of ways to improve it, such as by improving spatial sampling designs and covariate preparation and selection.

The abundance of spatial interpolation methods, simple and advanced, also requires a framework to decide which method to use in which situation. Several criteria were presented in Section 5, the most important of which is the accuracy of the interpolated map. It is important to be able to determine the map accuracy from a sample of ground-truth data, and here further developments are needed, because we have seen that standard cross-validation methods may be deficient, particularly in case of clustered data. Not mentioned so far but important too is that the ground-truth data that are used to assess map accuracy may have measurement errors. These measurement errors should be accounted for when the map accuracy is estimated. For example, when the Mean Squared Error is computed by comparing interpolated values with ground-truth validation data, then the difference between them is not only caused by interpolation error, but also by measurement errors in the validation data. In fact, these measurement errors are also present in the calibration data, and ideally should be included in the model fitting (Delhomme, 1978; Van der Westhuizen *et al.*, 2021). For instance, measurements with large measurement error should carry less weight during calibration than measurements with small measurement errors, but exactly how much is not obvious in a machine learning context. This is another area where future research may lead to improvements in spatial interpolation methods.

Some spatial interpolation methods quantify the interpolation error. Examples of these mentioned in this chapter are kriging and Quantile Regression Forest. It would be useful to extend these 'internal' uncertainty quantifications to other machine learning algorithms. Interestingly, the spatial pattern of the uncertainty maps as obtained with Quantile Regression Forests may be quite different from that obtained using kriging, while validation with independent observations show that both approaches provide a realistic assessment of the map accuracy (e.g. Vaysse and Lagacherie, 2017). It is intriguing that both uncertainty maps are deemed valid while being very different. We should look into ways that better evaluate how well these maps that characterize the spatial interpolation error.

It would also be useful to evaluate how spatial interpolation can benefit from the output of mechanistic models. For instance, maps of crop yield, erosion and air quality can benefit from the output of dynamic, mechanistic mod-

els if such models are available. One obvious solution is to include these model outputs in the trend of the geostatistical model defined in Eq. (4) or use them as covariates in a regression kriging or machine learning model (e.g Sekulić, Kilibarda, Heuvelink *et al.*, 2020). Razavi (2021) provides an interesting discussion on bridging the gap between process-based models and deep learning models for mapping the environment. This seems particularly relevant to space-time interpolation because many machine learning algorithms may not satisfy basic physical laws of conservation of mass and energy. Incorporating process-based knowledge may thus not only lead to improved interpolations but also increase our knowledge about the physical, chemical, biological and social processes that are the cause of spatio-temporal variation.

Finally, it is interesting to note the role that the measurements have played in spatial interpolation over the years. Initially, they were used for direct interpolation and measurements nearby were more important than remote measurements. This is a characteristic of both deterministic and geostatistical interpolation methods. But in recent years with the advancement of machine learning the measurements were mainly used to fit a model that predicts the dependent variable from covariates, and spatial context and geographic distance between measurement and prediction locations were no longer explicitly used. We now see a revival of weighing observations according to geographical distance with the development of Random Forest Spatial Interpolation (Sekulić, Kilibarda, Heuvelink *et al.*, 2020; Sekulić *et al.*, 2021), where measurements are used both to train a machine learning model as well as provide local neighbourhood information. I hope and expect that this extension of machine learning and explicit recognition of the first law of geography will blossom in the years to come.

We have come a long way but this final section made clear that there are many open questions and research topics that need to be considered in future research. I am confident that researchers of the Geodesy and Geoinformatics group of the Faculty of Civil Engineering, University of Belgrade will continue to make a mark and make substantial contributions to this important research field.

REFERENCES

Bajat, B., Hengl, T., Kilibarda, M. and Krunić, N. (2011). Mapping population change index in southern Serbia (1961–2027) as a function of environmental factors. *Computers, Environment and Urban Systems*, *35*(1), 35–44.

Bajat, B., Krunić, N., Kilibarda, M. and Samardžić-Petrović, M. (2011). Spatial modelling of population concentration using geographically weighted regression method. *Journal of the Geographical Institute "Jovan Cvijic", SASA*, *61*(3), 151–167.

Bajat, B., Pejović, M., Luković, J., Manojlović, P., Ducić, V. and Mustafić, S. (2013). Mapping average annual precipitation in Serbia (1961–1990) by using regression kriging. *Theoretical and Applied Climatology*, *112*(1), 1–13.

Brenning, A. (2005). Spatial prediction models for landslide hazards: review, comparison and evaluation. *Natural Hazards and Earth System Sciences*, *5*(6), 853–862.

Burrough, P. A., McDonnell, R. A. and Lloyd, C. D. (2015). *Principles of Geographical Information Systems*. Oxford university press.

Chiles, J.-P. and Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons.

Delhomme, J. P. (1978). Kriging in the hydrosciences. *Advances in Water Resources*, *1*(5), 251–266.

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. MIT press.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S. and Kempen, B. (2017). Soilgrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, *12*(2), e0169748.

Hengl, T., Miller, M. A., Križan, J., Shepherd, K. D., Sila, A., Kilibarda, M., Antonijević, O., Glušica, L., Dobermann, A., Haefele, S. M., McGrath, S., G.E. Acquah, J. Collinson, L. Parente, M. Sheykhmousa, K. Saito, J.M. Johnson, J. Chamberlin, F.B.T. Silatsa, M. Yemefack, Wendt, J., R.A. MacMillan, I. Wheeler and J. Crouch. (2021). African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. *Scientific Reports*, *11*, 6130.

Heuvelink, Gerard BM., Angelini, M. E., Poggio, L., Bai, Z., Batjes, N. H., van den Bosch, R., Bossio, D., Estella, S., Lehmann, J., Olmedo, G. F. and Sanderman, J. (2021). Machine learning in space and time for modelling soil organic carbon change. *European Journal of Soil Science*, *72*(4), 1607–1623.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer.

Jenny, H. (1994). *Factors of Soil Formation: a System of Quantitative Pedology*. Dover Publications.

Journel, A. G. and Huijbregts, C. J. (1978). *Mining Geostatistics*. Academic Press.

Kilibarda, M., Hengl, T., Heuvelink, G.B.M., Gräler, B., Pebesma, E., Perčec Tadić, M. and Bajat, B. (2014). Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *Journal of Geophysical Research: Atmospheres*, *119*(5), 2294–2313.

Kilibarda, M., Tadić, M. P., Hengl, T., Luković, J. and Bajat, B. (2015). Global geographic and feature space coverage of temperature data in the context of spatio-temporal interpolation. *Spatial Statistics*, *14*, 22–38.

Kovačević, M., Bajat, B. and Gajić, B. (2010). Soil type classification and estimation of soil properties using support vector machines. *Geoderma*, *154*(3-4), 340–347.

Li, J. and Heap, A. D. (2008). A review of spatial interpolation methods for environmental scientists. *Geoscience Australia, Record 2008/23.*.

Marjanović, M., Kovačević, M., Bajat, B. and Voženílek, V. (2011). Landslide susceptibility assessment using svm machine learning algorithm. *Engineering Geology*, *123*(3), 225–234.

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, *7*, 983–999.

Meyer, H. and Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution*, *12*(9), 1620–1633.

Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S. and McBratney, A. B. (2019). Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma*, *352*, 251–267.

Pejović, M., Nikolić, M., Heuvelink, G.B.M., Hengl, T., Kilibarda, M. and Bajat, B. (2018). Sparse regression interaction models for spatial prediction of soil properties in 3D. *Computers & Geosciences*, *118*, 1-13.

Pejović, M., Bajat, B., Gospavić, Z., Saljnikov, E., Kilibarda, M. and Čakmak, D. (2017). Layer-specific spatial prediction of as concentration in copper smelter vicinity considering the terrain exposure. *Journal of Geochemical Exploration*, *179*, 25–35.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E. and Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL*, *7*(1), 217–240.

Qiao, P., Lei, M., Yang, S., Yang, J., Guo, G. and Zhou, X. (2018). Comparing ordinary kriging and inverse distance weighting for soil as pollution in Beijing. *Environmental Science and Pollution Research*, *25*(16), 15597–15608.

Razavi, S. (2021). Deep learning, explained: Fundamentals, explainability, and bridgeability to process-based modelling. *Environmental Modelling & Software*, *144*, 105159.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D., Wintle, B., Hartig, F. and Dorman, C. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929.

Samardžić-Petrović, M., Dragićević, S., Kovačević, M. and Bajat, B. (2016). Modeling urban land use changes using support vector machines. *Transactions in GIS*, *20*(5), 718–734.

Sekulić, A., Kilibarda, M., Heuvelink, G.B.M., Nikolić, M. and Bajat, B. (2020). Random Forest Spatial Interpolation. *Remote Sensing*, *12*(10), 1687.

Sekulić, A., Kilibarda, M., Protić, D. and Bajat, B. (2021). A high-resolution daily gridded meteorological dataset for Serbia made by Random Forest Spatial Interpolation. *Scientific Data*, *8*(1), 1–12.

Sekulić, A., Kilibarda, M., Protić, D., Tadić, M. P. and Bajat, B. (2020). Spatio-temporal regression kriging model of mean daily temperature for Croatia. *Theoretical and Applied Climatology*, *140*(1), 101–114.

Stehman, S. V. (1999). Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing*, *20*(12), 2423–2441.

Stehman, S. V. and Foody, G. M. (2009). Accuracy assessment. In *The SAGE Handbook of Remote Sensing* (pp. 297–309). SAGE London.

Tadić-Percec, M., Kilibarda, M., Kilibarda, M., Tadić-Percec, M., Hengl, T., Luković, J. and Bajat, B. (2018). Preface to special issue "GeoMLA Conference – Geostatistics and Machine Learning Applications in Climate and Environmental Sciences". *Geofizika*, *14*, 38–22.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*(sup1), 234–240.

Van der Westhuizen, S., Heuvelink, G.B.M. and Hofmeyr, D. (2021). Measurement error-filtered machine learning in digital soil mapping. In *EGU General Assembly Conference Abstracts* (pp. EGU21–9704).

Vaysse, K. and Lagacherie, P. (2017). Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma*, *291*, 55–64.

Wadoux, A. M.-C., Brus, D. J. and Heuvelink, G.B.M. (2019). Sampling design optimization for soil mapping with random forest. *Geoderma*, *355*, 113913.

Wadoux, A. M.-C., Heuvelink, G.B.M., De Bruin, S. and Brus, D. J. (2021). Spatial cross-validation is not the right way to evaluate map accuracy. *Ecological Modelling*, *457*, 109692.

Webster, R. and Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*. John Wiley & Sons.

Yu, M. and Liu, Q. (2021). Deep learning-based downscaling of tropospheric nitrogen dioxide using ground-level and satellite observations. *Science of The Total Environment*, *773*, 145145.

Lidija Zdravković

Lidija Zdravković is a Full Professor in the Department of Civil and Environmental Engineering at the Imperial College London, UK (since 2013), Head of the Geotechnics Division since 2014, a member of the British Geotechnical Association since 1997, a member of the Institution of Civil Engineers since 1997. She is a UK representative on Technical Committees TC 103 - Numerical Analysis, since 2010, and TC 221 - Tailings and Mine Wastes, since 2019, for the International Society for Soil Mechanics and Geotechnical Engineering.

Contact information: e-mail: l.zdravkovic@imperial.ac.uk

https://orcid.org/0000-0003-3092-0628

Prof. Zdravković graduated in 1988 in the Faculty of Civil Engineering, University of Belgrade (FCEUB), Department of Structural Engineering, as the best student and the winner of the *Prof. Ilija Stojadinović Award*. She defended her master's thesis in 1993 at the FCEUB and received her PhD in 1996 from Imperial College London, both in the field of geotechnics. After graduating, she worked in the FCEUB as a teaching assistant (1988-1992). She passed the professional exam and participated in several industry projects.

She continued her career in the Department of Civil and Environmental Engineering at Imperial College London as a post-doctoral student in 1996-1999. She was employed as an Assistant Professor (1999-2007), Associate Professor (2007-2013) and Full Professor since 2013. She gives lectures and tutorials to graduate and postgraduate students for several courses in the field of geotechnical engineering. She was a Departmental committee member for the organization of graduate studies 2004-2012 and for scientific research activities 2005-2016, and a Board member for the doctoral program "Sustainable Civil Engineering" 2013-2019. As Head of the Geotechnics Division, she has been a member of the Departmental Management Committee since 2014.

Her research activities are related to the numerical simulations of the behavior of geotechnical structures, and the characterization of thermo-hydro-mechanical behavior of soils through experimental research. She has contributed to the development of: thermo-hydro-mechanical formulations of the finite element governing equations for simulating the behavior of saturated and unsaturated soils; constitutive models according to the principles of critical state soil mechanics, kinematic hardening / softening and soil stiffness at small deformations; boundary conditions for sophisticated simulations of construction and operation of geotechnical facilities, as well as development of equipment, control and data acquisition systems for laboratory testing of soils.

Prof. Zdravković is a member of the BGA and ICE since 1997. She is a UK representative on ISSMGE TC 103 Numerical Analysis since 2010, and TC 221 Tailings and Mine Wastes since 2019. She was a member of the editorial boards of the international journals Geotechnique (2003-2006, 2014-2017) and Computers and Geotechnics (2010-2018) and serves as a reviewer for several international journals. She has mentored over 30 doctoral dissertations at Imperial College. She is the author and co-author of over 200 technical papers, as well as 2 monographs on the theory and application of the finite method in geotechnical engineering She has delivered over 30 keynote, plenary and guest lectures. Prof. Zdravković has received awards for high quality scientific papers, including the ICE Telford Gold Medal, 2002, and the BGA Medal in 2008, 2010, 2012 and 2020. As co-investigator on the PISA (Pile Soil Analysis) project, she won the BGA Fleming Award 2017, for innovation in the design of geotechnical structures. She was awarded the 2019 Medal from the President of Imperial College for outstanding contribution to student education, in the category of doctoral research mentors (President's Medal for Excellence in Education, category: outstanding Excellence in Research Supervision).

Geotechnical engineering: achievements and challenges over the past century

Lidija Zdravković

Department of Civil and Environmental Engineering, Imperial College London, UK

Summary

This paper presents a brief perspective on the development of the geotechnical engineering discipline since the early 20th century. The main challenges of working with soil as an engineering material are highlighted first, followed by author's outlook on the main transformative developments in this field over the past century that have enabled modern geotechnical design to grow, underpinned by the integrated application of advanced characterisation of the mechanical behaviour of soils and advanced computational analysis. Some of the main achievements are briefly discussed in this respect, as well as the current and future challenges.

1. BACKGROUND

Geotechnical engineering is an integral part of the civil engineering discipline, concerned with the engineering use of earth materials (soils and rocks) for the design, construction and maintenance of the natural and built environment. All civil engineering construction involves the soil in some way: structures rest on the soil via different foundation systems; they are constructed in the soil, such as tunnels and deep excavations; or they are constructed from the soil, in the form of earth dams and infrastructure embankments.

Soil itself, being a naturally produced rather than a man-made material, poses many challenges in both geotechnical engineering research and practice. The root causes are (i) the multi-phase nature of soil as a material and (ii) its highly-nonlinear response when subjected to thermo-hydro-mechanical (THM) perturbations. With respect to the former, although the soil is usually considered as a three-dimensional (3D) continuum in design, it is essentially a multi-phase material, containing an arrangement of solid particles as one phase and void (pore) space among the particles (Figure 1). The void space defines the state of the soil, as air (a second phase) in the voids classifies the soil as dry, water (a third phase) in the voids makes it fully saturated, while it is unsaturated if a mix of air and water is present in the voids. Phases interact with each other due to the non-rigid arrangement of the solid phase, which causes deformations in the ground and requires different mathematical formulations to describe the consequential relationships between the applied structural and environmental loads and the resulting volumetric and shear deformations in the ground. This relationship is highly nonlinear (Figure 2), as observed in many experimental studies of different soil types, leading to failure with increasing deformations when no further increase in load is possible. The multi-phase nature of soils also brings the challenge of stress being transferred into the ground in two parts, one through the solid phase (skeleton of solid particles) and the other through the fluid phase (air or water, or both) in the void space, which is not the case with other construction materials for which only a single stress tensor applies. Moreover, being a 3D continuum, the state of the ground is characterised with six components of stress (three direct and three shear stresses), six strain components (also three direct and three shear strains) and three displacement components (coinciding with coordinate directions). If hydraulic or thermal coupling is also considered, then the pore fluid pressure (air and water) and / or temperature in the ground also need to be defined.

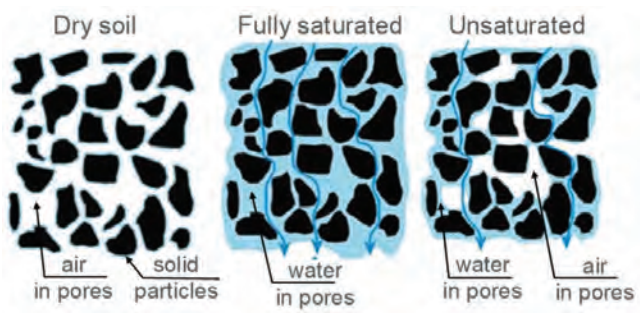


Figure 1. Multi-phase nature of soils and classification of their possible states

Although earthfill construction has been historically widely applied in society since early ages (e.g. embankments and dams for flood protection), the formalism in design and construction did not exists until the early 20th century, being principally based on empiricism and prior experience until then. It is widely accepted that Terzaghi's *effective stress* principle presents a corner stone and origin of modern soil mechanics and geotechnical engineering. Terzaghi (1966) defined that a *total stress* tensor, $\{\boldsymbol{\sigma}\}$, applied to an element of saturated soil, is transferred through the soil in two parts: one developed in the water in voids, termed a *pore water pressure*, u, as a scalar quantity; and the other developed in the solid particles, termed the *effective stress* tensor, $\{\boldsymbol{\sigma'}\}$, the relationship between them being $\{\boldsymbol{\sigma}\} = \{\boldsymbol{\sigma'}\} + u$. The effective stress principle demonstrated that any deformations in the soil and changes in its shearing resistance are caused exclusively by changes in the effective stress. Based on this Terzaghi further developed the theory of consolidation (one-dimensional, 1D), recognising that the flow of water in the voids, defined by Darcy (1856), was a time-related process which depended on the soil's permeability. This represents an additional challenge in geotechnical design, as both stresses and strains in the ground are transient (changing over time), following any construction process, therefore the design must assess both short- and long-term conditions in the ground. The consolidation theory was subsequently generalised for a three-dimensional (3D) domain by Biot (1941).

Following these developments from the 19th to early 20th century, the first mathematical framework adopted in the mechanics of soils (and most other materials) was that of linear elasticity (e.g. Timoshenko, 1934), where the ground was treated as a semi-infinite elastic medium. This framework requires only two parameters to describe a stress-strain relationship in a material, a Young's modulus, E, and a Poisson's ratio, ν, and it enabled early calculations of ground movements and, hence, assessment of serviceability states in the ground. A comprehensive summary of elastic solutions for application in soil and rock mechanics was given in Poulos and Davis (1974). Clearly, a framework like this is inadequate for describing failure in soils and is also contrary to the material nonlinearity shown in the Figure 2 example. However, the testing of soils was limited to consistency characterisation with Atterberg limits, mainly applicable to clays. It was Casagrande (1932) who standardised Atterberg tests for consistency-based soil classification (liquid and plastic limits, plasticity index), which have been used ever since. However, experimental procedures for mechanical characterisation of soils were limited.

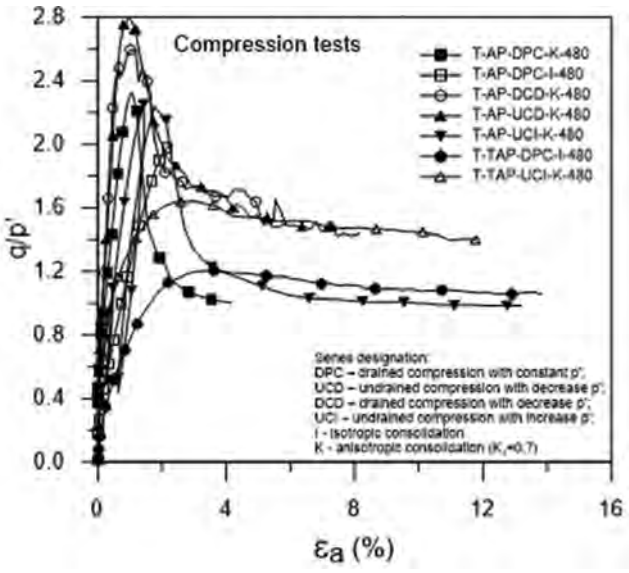


Figure 2. Nonlinear stress-strain behaviour (data from Pedro *et al.*, 2018)

Theoretical frameworks also started to develop at that time for the assessment, principally, of soil capacity / ultimate limit states in the design of geotechnical structures. As this was the time before computer power started to emerge, the calculation process was mainly undertaken by hand. The soil mechanical behaviour was simplified to being rigid (undeformable) prior to failure and failing along shear planes, mobilising one of two failure criteria, depending on the soil type: the Mohr-Coulomb failure criterion which uses effective stresses and the drained shear strength of the soil (effective cohesion, c', and angle of shearing resistance, ϕ') to determine the stress state at failure; and the Tresca failure criterion which uses total

Table 1. Classical methods of geotechnical analysis – theoretical requirements (from Potts and Zdravković, 1999)

Method of analysis		Theoretical requirements for accurate solution				
		Equilibrium	Compatibility	Constitutive behaviour	Boundary conditions	
					Force	Displacement
Limit equilibrium		S	NS	Rigid with a failure criterion	S	NS
Method of characteristics		S	NS	Rigid with a failure criterion	S	NS
Limit analysis	Lower bound	S	NS	Ideal plasticity	S	NS
	Upper bound	NS	S	Ideal plasticity	NS	S

S – satisfied; NS – not satisfied

Table 2. Classical methods of geotechnical analysis – design requirements (from Potts and Zdravković, 1999)

Method of analysis		Design requirements		
		Stability	Ground movements	Effect on adjacent structures
Limit equilibrium		S	NS	NS
Method of characteristics		S	NS	NS
Limit analysis	Lower bound	S	NS	NS
	Upper bound	NS	S – crude estimate	NS

S – satisfied; NS – not satisfied

stresses and the undrained shear strength of the soil, S_u, to determine stresses at failure. Three calculation methods, summarised in Table 1, were developed: limit equilibrium, stress fields (method of characteristics) and limit analysis (upper and lower bound). Only the upper bound limit analysis was able to provide a crude estimate of ground movements, with the remaining methods focusing on satisfying the requirement of equilibrium between external and internal forces in the ground. It is clear from Table 1 that these classical calculations do not satisfy all *theoretical requirements* (equilibrium, compatibility, constitutive behaviour and boundary conditions) to produce accurate solutions. Also, their ability to provide the main *design requirements*, as summarised in Table 2, is constrained to predominantly ensuring the stability of geotechnical structures. Assessment of ground movements and of the effects that new construction may have on existing structures and services is restricted.

Based on such state-of-the art until mid-20th century it was clear that soil mechanics and geotechnical engineering practice were very empirical, albeit founded on sound physical laws.

2. TRANSFORMATIVE DEVELOPMENTS FROM THE MID-20TH CENTURY

In the author's opinion, the main developments since the 1950s that have enhanced and modernised both research and practice in soil mechanics and geotechnical engineering were threefold: (i) new experimental apparatus and laboratory testing protocols for quantifying soils' hydro-mechanical behaviour and engineering design parameters; (ii) development of the *critical state* framework for the mathematical formulation of soil behaviour in terms of elasto-plasticity; and (iii) development of modern computers and associated numerical software.

2.1. Laboratory testing of soils

Rigorous geotechnical design requires quantification of soils' engineering properties, which are broadly classified as shear strength (defining soil capacity) and stiffness (defining soil deformability) under different loading paths. The most advanced derivation of such parameters is by testing specimens of soil in a well-instrumented testing apparatus which can measure applied stresses, changes in pore water pressure and resulting deformations of specimens. The Bishop and Wesley (1975) hydraulic triaxial apparatus (Figure 3) was one of the first to enable cylindrical soil specimens to be tested under a wide range of stress paths both in compression and extension, under drained or undrained conditions, and either at a controlled rate of loading or a controlled rate of strain. The objective was that such paths mimic some of the actual loading conditions in the field. To date the B&W cell has been the most widely used piece of testing equipment, with the obtained measurements interpreting the soils' drained shear strength, c' and ϕ'; undrained shear strength, S_u; Young's modulus, E; shear modulus, G; bulk modulus, K; and Poisson's ratio, ν.

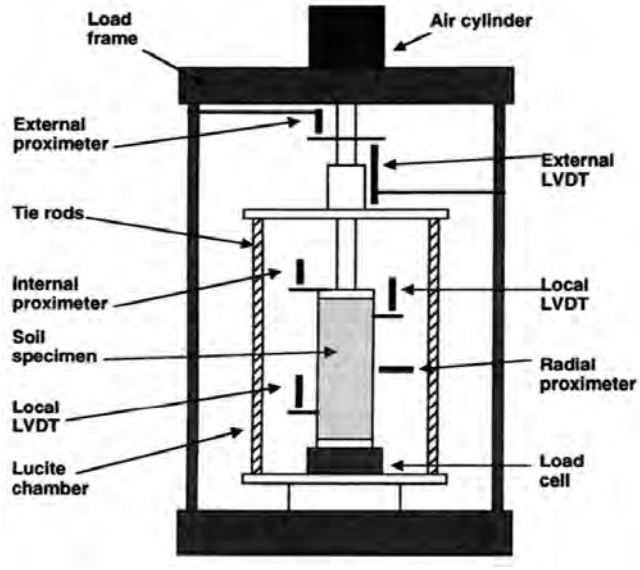


Figure 3. Schematic of a triaxial apparatus with local and external strain measurement

Triaxial compression and extension states represent specific axi-symmetric stress states in the ground, for which the major direct principal effective stress, σ'_1, is either vertical or horizontal, respectively, and the two horizontal stress are equal. However, the usual construction, excavation and loading processes in the ground induce additional shear stresses which contribute to the rotation of the major, σ'_1, and minor, σ'_3, direct principal stresses and result in the intermediate direct principal stress, σ'_2,

no longer being equal to one of the other two principal stresses. Both strength and stiffness of the soil under such paths are different and dependent on the orientation of the major principal stress, α, to the vertical, and on the magnitude of the σ'_2 stress (expressed via a parameter $b = (\sigma'_2 - \sigma'_3)/(\sigma'_1 - \sigma'_3)$). Quantification of the soil strength with respect to variations in and is only possible with laboratory testing using hollow cylinder devices, Figure 4, (e.g. Hight *et al.* 1983; Ishihara and Yasuda 1975). Both triaxial and hollow cylinder apparatus developments were then extended from monotonic to cyclic / dynamic loading, facilitating for the first time soil characterisation for earthquake engineering and soil dynamics design problems.

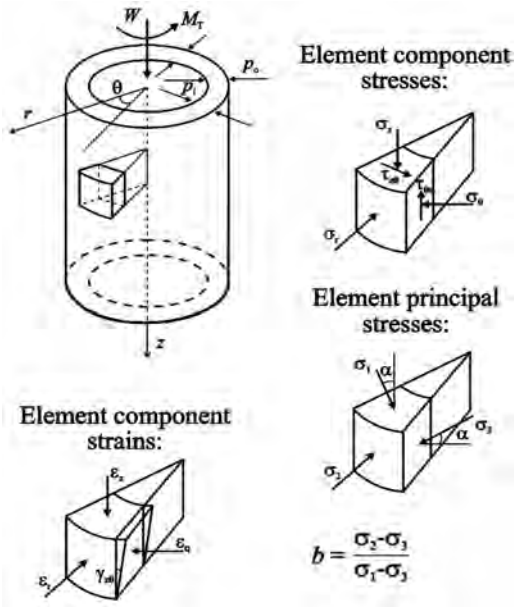


Figure 4. Schematic of a hollow cylinder specimen, with stress and strain components

While the measurement of soil shear strength was becoming more reliable, the measurement of a specimen's deformation in testing apparatus still relied on transducers positioned outside the cell, hence measuring not only the straining in the soil but also the compliance of the cell itself. The breakthrough came in the mid-1980s when measuring instruments were developed that could be mounted directly on the soil specimen (see Figure 3), hence measuring only the soil deformation (e.g. Goto *et al.* 1991; Jardine *et al.* 1984; Lo Presti *et al.* 1994). Such instruments improved the accuracy and resolution of strain measurements, being generally characterised as measuring the soil response at *small strains*. For the first time the true nature of soil stiffness was understood and quantified as nonlinear, being high at the start of the loading path, then degrading with increasing deformation in the soil (Figure 5, from Atkinson and Sallfors 1991). Numerous experimental studies have since demonstrated the small strain nature of stiffness (both shear and bulk components) in most soils, as well as its dependency on stress level and void ratio. This experimental development alone proved crucial for achieving accurate calculations of ground movements and the resulting confidence in the serviceability limit state design of geotechnical structures. Measurements of strain hysteresis during cyclic and dynamic loading would also become crucial in quantifying energy dissipation in the soil.

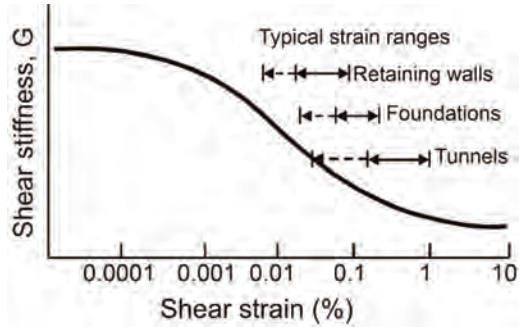


Figure 5. Small strain stiffness in soils

2.2. Critical state soil mechanics (CSSM) framework

The failure criteria of classical soil mechanics, Mohr-Coulomb and Tresca, define the failure envelopes / limit stress conditions in the soil. When interpreted in classical elasto-plasticity, both criteria assume associated conditions, in that the direction of plastic straining (the flow rule) is normal to the failure envelope. The Tresca criterion therefore represents the state of zero volumetric plastic strain, while the Mohr-Coulomb criterion represents constantly negative / dilative volumetric plastic strains. The latter implication does not apply to all soil conditions (e.g. some soils are contractive, with positive volumetric strains) and also implies an infinite mobilised shear stress in the soil under confined conditions. Both criteria are very limited in their representation of real soil behaviour.

The critical state framework was the first to link soils' strength and stiffness to their current volume. The framework (Schofield and Wroth, 1968) was derived from triaxial experiments on reconstituted clays (e.g. Henkel, 1960), subjected to sets of drained and undrained effective stress paths at different stress levels and overconsolidation. The applied constant volume and the changing volume paths demonstrated the uniqueness of the so-called *critical state line* (CSL), as a line of all failure states irrespective of the test path followed. The critical state therefore defines the failure state in a soil at which large shear distortions occur with no change in stress or in specific volume. The limiting surfaces of the possible stress space were identified as the Roscoe surface (a boundary to normally and lightly-overconsolidated soils wet of critical state) and the Hvorslev surface (a boundary to heavily overconsolidated soils dry of critical state).

The constitutive models first developed in this framework, the Cam clay (Schofield and Wroth, 1968) and the modified Cam clay (Roscoe and Burland, 1968), were able for the first time to simulate both contractive and dilative volumetric strains (ε_{vol}), depending on the soil state (wet or dry of critical, respectively), as well as the true ultimate strength (i.e. the critical state strength), Figure 6. The CSSM framework has endured the test of time, with most of the advanced constitutive models for saturated and unsaturated soils in use today having the critical state formulation and being extended to account for advanced facets of soil behaviour, such as nonlinearity, anisotropy, strain rate, creep, state and structure / fabric.

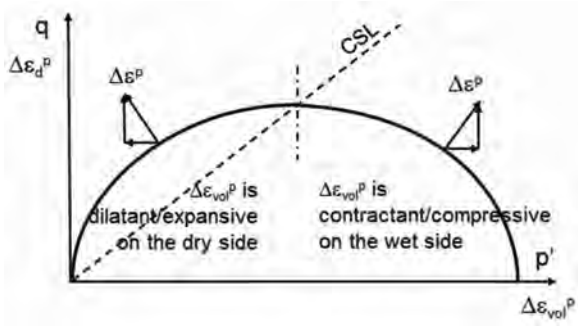


Figure 6. Volumetric behaviour of Cam clay models

2.3. Development of modern computers

The improved understanding of real soil behaviour, that came with advanced laboratory testing, and the first critical state constitutive models mentioned above, were impossible to apply in geotechnical design using classical calculation methods summarised in Table 1. The development of computer power from the mid-20th century enabled development of numerical methods of analysis, the most prominent in all engineering disciplines being *the finite element method* (FEM) which solves partial differential equations in a two- or three-dimensional continuum. Originally developed for linear problems it was first applied in mechanical, aero and structural engineering since the 1960s. The inclusion of material nonlinearity, plasticity and time in the FEM formulation that was appropriate for geotechnical engineering appeared in a textbook of Zienkiewicz (1971), who is considered as one of the pioneers of FEM. The geotechnical FEM software started to develop from the 1980s and examples of subsequent textbooks focussing specifically on FEM in geotechnical engineering include Potts and Zdravković (1999; 2001); Smith and Griffiths (1988); Zienkiewicz *et al.* (1999), among others.

A finite element (FE) numerical model considers a 2D or 3D truncated soil continuum, discretised with a mesh of smaller (finite) elements. As such, it requires appropriate *boundary conditions* (mechanical, hydraulic, dynamic) to be applied on the truncated boundaries, which has been one aspect of numerical developments in this field (see e.g. Potts and Zdravković, 1999; 2001). Another essential area of development has been the constitutive modelling of soils, which has seen progressively more complex mathematical forms to represent real soil behaviour observed with the advent of advance laboratory testing of soils. These have ranged from simple elasto-plastic Mohr-Coulomb type models to critical state kinematic surface or bounding surface plasticity models, with isotropic and / or kinematic hardening and softening (see e.g. Potts and Zdravković, 1999; 2001). Finally, the overall solution process of the FE numerical model involves solving a large system of equations to obtain displacements in the ground in the first instance, and from those to calculate strains and stresses. The solution process is challenging for nonlinear problems, such as those in geotechnical engineering, leading to a number of algorithms being proposed in the literature, which have variable robustness and accuracy.

The current availability of geotechnical FE software and growing complexities of geotechnical problems have been increasingly promoting the use of numerical methods, FEM in particular, for design calculations. The classical analysis methods are also still in use for preliminary calculations and are today formulated in various computer-based software.

3. MAIN ACHIEVEMENTS IN THE FIELD

The combination of the above developments has enabled full boundary value problems to be analysed as part of a geotechnical design process, with improved accuracy and efficiency of the solution, using appropriate boundary conditions and constitutive models for soil behaviour in geotechnical FE software. Examples that follow are based on the author's experience and other research conducted at Imperial College. In particular, the FE analyses presented here have been performed with the Imperial College Finite Element Program (ICFEP), developed since the 1980s through research and practical applications, underpinning design solutions for a wide range of geotechnical problems (Potts and Zdravković, 1999; 2001).

3.1. Soil-structure interaction

The experimental evidence of soils' nonlinear small strain stiffness (as depicted in Figure 5) enabled the development of early relatively simple nonlinear elastic models to reproduce such stiffness degradation (Simpson et al., 1979; Jardine *et al.* 1986), which could be implemented in FE software. This has revolutionised the assessment of soil-structure interaction (SSI) at conditions of working load (i.e. serviceability limit states). In particular, accurate quantification of ground movements induced by construction processes became possible, as well as the assessment of the effects that a new construction may impose on

existing structures and services, contributing to more efficient and economical design solutions. Such FE analyses became especially important for developments in urban environments where construction tolerances are tight.

London, as capital of the UK, is a prime example of an urban environment where the underground space in the 1980s was already congested with a 100-year old network of metro tunnels, escalators and station shafts. The existing tunnels, up to 20 m deep, used to be constructed under the streets to avoid inducing ground movements, due to tunnel excavation, under existing structures. The first new tunnels constructed in the 1990s, which extended the existing Jubilee Line, were not designed to align with roads on the ground surface, as this would not have been economical. They were instead passing underneath built areas of central London, therefore introducing design concern for building damage due to tunnelling-induced ground movements. Additionally, the new line introduced the then deepest station box located at Westminster, which, at around 40 m depth, surpassed existing deep excavations in London which were around 20 m deep on average. Clearly, the serviceability limit state was the primary concern for design and the sensitivity of existing structures and services could not be assessed by classical calculations, emphasising the relevance and importance of using numerical tools, such as FEM.

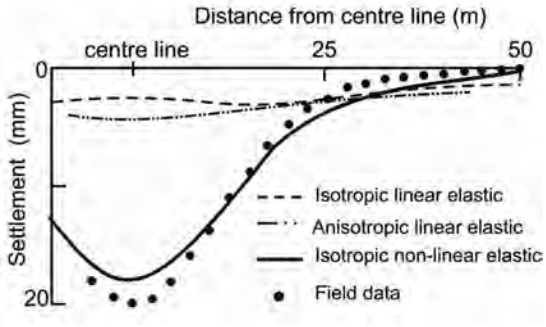


Figure 7. Predicted and measured settlement above the WB Jubilee Line tunnel at St James's park

Figure 7 is one example that clearly demonstrates the significance of reproducing the nonlinear stiffness of the soil in design calculations. The calculation process was the finite element analysis of the excavation of the Jubilee Line westbound (WB) tunnel in London clay, at the location of the greenfield site at St James's park (Addenbrooke *et al.*, 1997). The figure shows that the prediction of the ground surface settlement trough immediately after excavation was inadequate when the soil shear stiffness was simulated as linear elastic (be it isotropic or anisotropic). On the contrary, when a nonlinear stiffness degradation model (Jardine *et al.*, 1986) was introduced, calibrated on laboratory experimental data of stiffness degradation in London clay, the predicted settlement trough agreed very well with measurements. In a similar manner, the FE analysis using the same constitutive modelling of the soil, accurately predicted the horizontal wall movements at various stages of the Westminster station excavation when compared to field measurements (Figure 8). In effect, the FE analysis of the originally envisaged top-down construction of the station, performed before the construction started, showed that such sequence would have induced excessive ground movements which would have affected the surrounding buildings, in particular the sensitive masonry structures of the Big Ben clock tower and the Palace of Westminster (the seat of the UK Parliament). The actual construction sequence executed on site was derived from FE analyses of other possible solutions which resulted in acceptable magnitudes of ground movements.

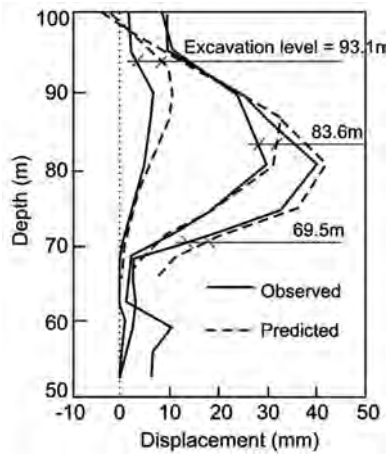


Figure 8. Predicted and measured wall deflections during construction of Westminster station

The modelling of nonlinear small strain stiffness has had very wide design implications for SSI evaluations, not only in underground construction. The North Sea oil and gas exploration in the 1980s started with gravity-based structures in shallower waters (e.g. Gullfaks C; Hight *et al.* 1988). The move to deeper waters was not possible with the same type of platforms and new solutions were explored, one being tension leg platforms (TLPs) which relied on foundations (piles or suction caissons) loaded in tension. The critical design aspect was that of possible magnitudes of the vertical pull-out movements (out of the seabed), implying the serviceability, rather than ultimate, limit state design. FE analyses with ICFEP were crucial in proving the design concept of first TLP platforms in the North Sea (Hutton TLP; Jardine and Potts 1988), where the seabed soil was a glacial clay and the foundation system comprised four groups of piles in tension, connected to the hull via steel tendons. The FE analysis, performed before the TLP installation and adopting the nonlinear stiffness model, predicted accurately the upward vertical pile movements with increasing template load (Figure 9), which were monitored at one of the legs (pile groups). The conventional analysis, assuming representative linear

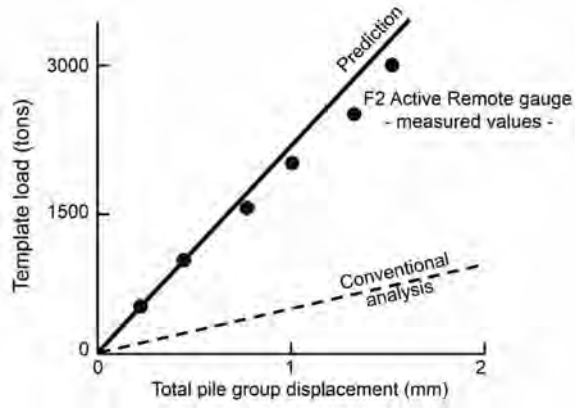


Figure 9. Predicted and measured pile group displacement at Hutton TLP

elastic soil stiffness, predicted unconservative pile movements.

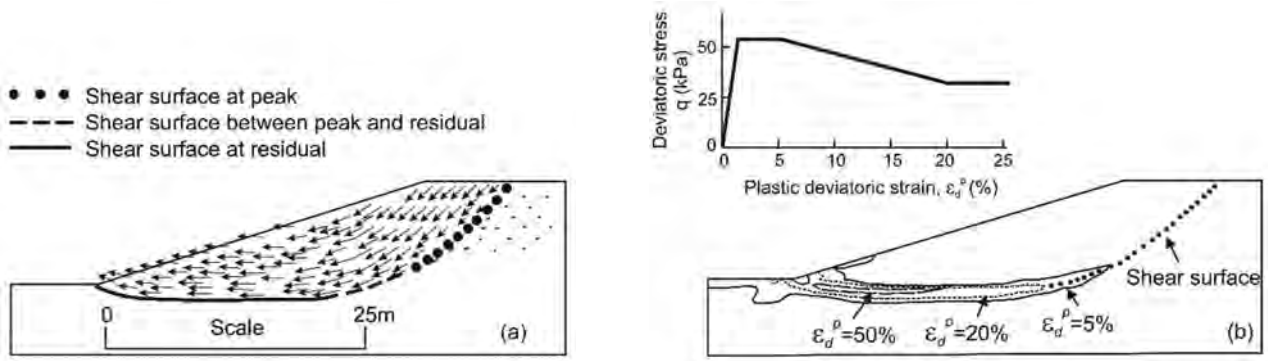
Examples cited above demonstrate that from the mid-20th century onwards the geotechnical engineering discipline started to transition from empiricism to rigorous numerical modelling, closely integrated with experimental characterisation of real soil behaviour and with field monitoring which provided essential data for the verification of numerical predictions.

3.2. Transient problems

Examples discussed in the previous section involved only mechanical FE modelling of the boundary value problems. Extension of FE software to account for hydro-mechanical coupling, solving both the equation of equilibrium and equation of continuity, has enabled exploration of transient states in geotechnical design, associated with a multiphase nature of soils.

Infrastructure slopes are one example where the lifecycle assessment of stability is crucial. This emerged as an issue in the 1970s and 80s in the UK when some existing cut slopes along motorways started to fail tens of years after their excavation. *Progressive failure* (i.e. progressive evolution of the failure surface in the slope and hence a progressive mobilisation of the soil's shear strength), which is a transient process, was suspected as a cause of instabilities, but was not possible to quantify with classical calculations that assume a constant value of mobilised soil strength. Namely, an excavation of a slope in a stiff plastic clay may be initially undrained, causing a depression of the initial phreatic surface (i.e. initial ground water table, GWT) and creating negative pore water pressures above it. Such a slope may appear stable initially, but the numerical studies conducted for the first time in the 1990s to analyse cut slopes in London clay (Kovacevic, 1994; Potts *et al.*, 1990; 1997), revealed the mechanism of a shear surface starting to develop in the slope immediately after excavation and continuing to progress up the slope for some time post-excavation. This was shown to be the result of pore water pressure equilibration with time, in conjunction with the brittle nature of stiff plastic clays. Depending on soil properties and slope geometry, the propagation of the shear zone may stabilise and never lead to failure, or it may lead to full failure at some time post-excavation. Figure 10, after Potts *et al.* (1997), shows (a) vectors of ground movements and (b) contours of plastic shear strains, ε_d^p, at failure in a London clay slope, predicted to happen 14.5 years after excavation. The London clay was modelled with a non-linear strain-softening Mohr-Coulomb model and the inset in Figure 10 shows the experimentally measured reduction of the mobilised shear stress, q, from its peak value at a deviatoric plastic strain, $\varepsilon_d^p = 5\%$, to its residual value at $\varepsilon_d^p = 20\%$.

The failure mechanism in Figure 10 clearly demonstrates the non-uniform mobilisation of the soil's strength along the shear surface, which has reached a residual value along some distance from the toe, and is below peak near the crest of the slope, in the last stable increment of the analysis that is represented in the figure. These early studies explained and quantified for the first time the importance of the brittle behaviour of stiff clays in the tran-



(a) **(b)**

Figure 10. (a) Vectors of ground movement and (b) Contours of plastic shear strains at failure, mobilised in the slope 14.5 years after its excavation

sient development of progressive failure in cut slopes. The stability was further shown to depend also on the slope's inclination and depth, initial stresses in the ground and hydraulic boundary conditions applied along the excavated surface of the slope. These findings assisted the subsequent development of new infrastructure slopes in London clay, as well as the redesign of existing slopes in conjunction with the widening of motorways. Similar recent FE studies by Zdravković, Potts and Taborda (2021), investigating additionally the effects of the small strain shear stiffness of London clay (small strain interpretations SS1 to SS4 in Figure 11) and different gradients and depths of excavated slopes, in conjunction with the high-speed rail line development in the UK, confirmed a clear design boundary of slope geometries that ensure design-life stability (more than 120 years) of cut slopes in London clay (Figure 11).

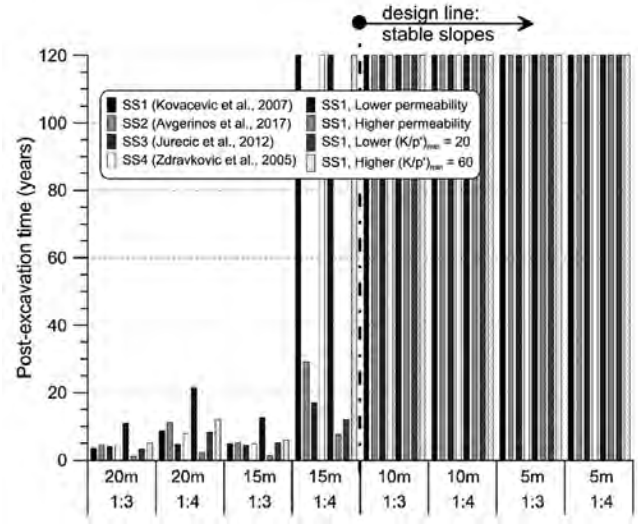


Figure 11. Summary of the predicted long-term stability of cut slopes in London clay

Recent developments of infrastructure slopes for railways have additionally been concerned with the magnitudes of the long-term base heave (hence serviceability limit state), due to the intended types of track to be placed, that cannot sustain large movements. This was a less critical design constraint in the past. As discussed above, the slope excavation is likely to be undrained (with respect to the clay's permeability and rate of excavation), and the magnitudes of ground movements in this phase are likely to depend on the mobilised shear stiffness in the soil. On the other hand, the long-term post-excavation movements, involving volumetric swelling, will mostly depend on the soil's bulk stiffness and permeability. Some design approaches suggest that if the placement of the track structure is delayed, usually by a year, then most of the base heave would have taken place, leaving negligible remaining operational heave (10 to 15 mm considered acceptable) over the design life of a slope (120 years). The

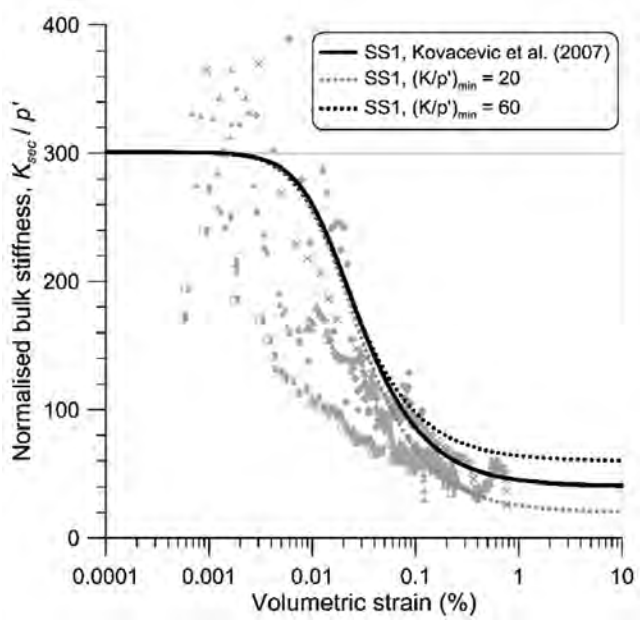


Figure 12. Experimental scatter and adopted modelling of London clay bulk stiffness

study of Zdravković, Potts and Taborda (2021), which varied the bulk stiffness and the permeability of London clay within the scatter of experimental data (Figure 12 and Figure 13, respectively) and analysed all the stable slope geometries identified in Figure 11, demonstrated that such design assumption was not appropriate for infrastructure slopes in stiff plastic clays. The magnitudes of operational heave were shown to be substantial and highly dependent on bulk modulus and permeability, and that small operational heave would manifest only in slopes of small depths (Figure 14). This indicated that the economy of the design could be substantially improved with a more accurate characterisation of these two soil properties as part of ground investigation, that may predict smaller heave. Otherwise, costly mitigating measures may need to be put in place to satisfy serviceability design requirements.

The capability to assess and quantify the transient nature of soil behaviour has further advanced the ability of research and practice to assess the whole life cycle of geotechnical structures, both short- and long-term, and contribute to safer and more efficient design.

4. CURRENT AND FUTURE CHALLENGES

Current and future challenges in soil mechanics and geotechnical engineering are being imposed by global aspirations to transition towards a sustainable and resilient society, which assumes exploitation of renewable energy sources, a zero carbon economy and infrastructure design that involves not only new build but, more importantly, life extension of existing infrastructure to combat the effects of climate change. All these variables and their interaction are highly complex, which even more necessitates

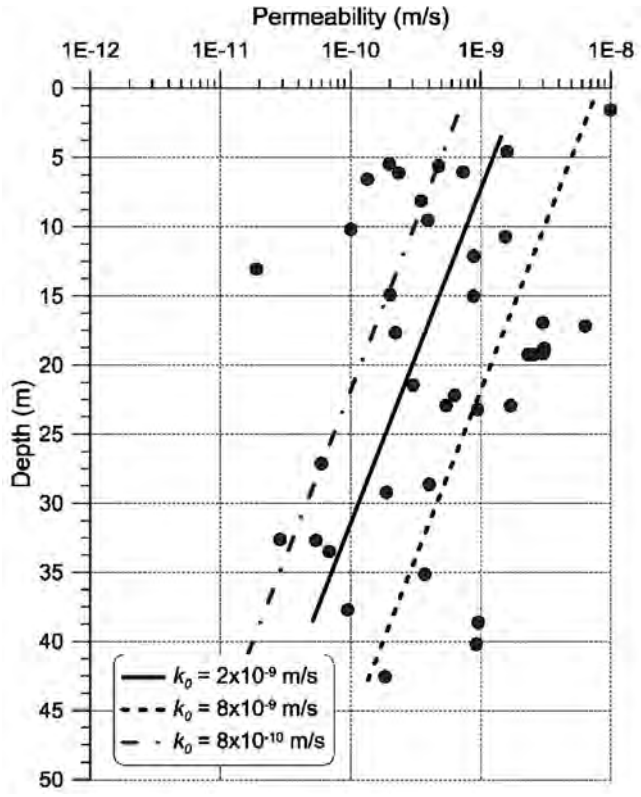


Figure 13. Experimental scatter and adopted modelling of London clay permeability profile

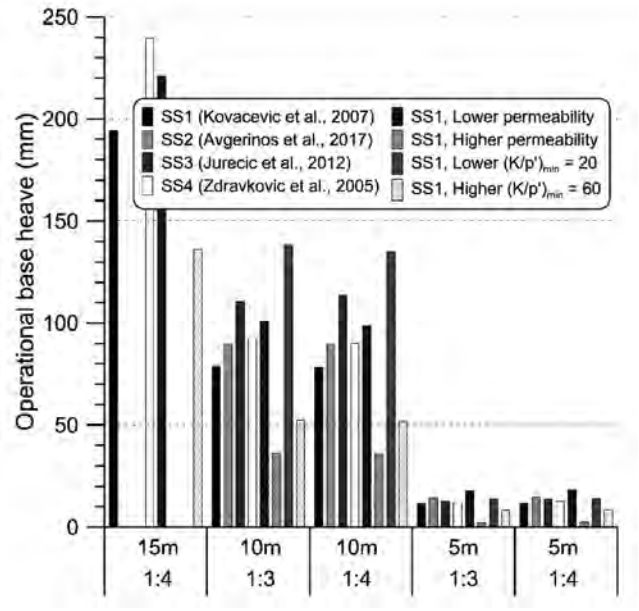


Figure 14. Summary of the predicted long-term operational base heave of cut slopes in London clay

the use of advanced numerical tools in the design process and at the same time promotes experimental campaigns that go beyond just mechanical loading of soil specimens, but also involve thermal and / or chemical perturbations. This section showcases some examples from the author's recent and current research which demonstrate the value of advanced calculations in producing safe and economical geotechnical design for the future.

4.1. Offshore wind as renewable energy

The development of offshore wind farms internationally, and in the North Sea sector particularly, has taken an exponential rise since 2015, following concerted collaborative research efforts between academia and offshore developers. In effect, the offshore wind industry was the driving force for the development of new design methods for monopile foundation systems to support wind turbine generators (WTG), having realised that the existing codified $p-y$ methodology (API, 2010; DNVGL, 2016) was not working well for large diameter wind turbine monopiles with low length to diameter ratios ($L/D < 5$). The $p-y$ methodology, used widely for oil and gas offshore developments in the 1980s and 90s, was based on empirical correlations derived from field tests on very slender laterally loaded piles ($L/D > 20$) in clays and sands (Reese *et al.*, 1974; 1975). Representing a pile as a beam supported by independent lateral nonlinear springs (i.e. a Winkler model) assumed that a soil reaction to such loading was produced only by mobilised lateral stresses in the ground, with representing a reaction in each spring and being the lateral displacement of each spring.

A new design methodology was developed as part of the PISA (**PI**le-**S**oil **A**nalysis) Joint Industry Project (JIP), led by the Carbon Trust and offshore wind developer Ørsted, with the consortium involving an academic working group and nine other offshore developers. Central to the new development was academic research, comprising Imperial College London, Oxford University and University College Dublin, with principal project outcomes featuring in the papers of Burd, Abadie *et al.* (2020); Burd, Beuckelaers *et al.* (2020); Byrne, Houlsby *et al.* (2020); Byrne, McAdam *et al.* (2020); McAdam *et al.* (2020); Taborda *et al.* (2020); Zdravković, Jardine *et al.* (2020); Zdravković, Taborda *et al.* (2020) and Burd, Taborda *et al.* (2020). Numerical analyses, ground investigation and in-situ pile testing were integral parts of the PISA project.

Considering the size of wind turbine monopiles, with current diameters of 8 to 10 m, full scale field testing from which to derive a design method, as in the original $p-y$ methodology, was clearly impractical. Even reduced scale field testing, as performed at the selected project sites of Cowden (Byrne, McAdam *et al.*, 2020) and Dunkirk (McAdam *et al.*, 2020), was expensive, raising also concerns of scale effects. The new design method was therefore envisaged to be developed on the basis of 3D FE analyses of monopiles. The key to achieving this successfully and to gaining the confidence of the industry partners and verification bodies on the feasibility of the new design method, was to demonstrate the ability of the FE modelling to deliver accurate predictions of monopile

response in given ground conditions. For this purpose site-specific FE analyses of medium-scale field tests at Cowden (Zdravković, Taborda *et al.*, 2020) and Dunkirk (Taborda *et al.*, 2020) were performed before the field testing took place, and *Class A* predictions (before the event) of pile behaviour were compared with subsequent field measurements. Cowden, as a site of low plasticity glacial clay till, and Dunkirk, as a dense marine sand site, were onshore sites chosen for field testing, as representative of sea-bed soils in some sectors of the North Sea.

Given that the foundation design had to represent accurately the monopile response at both working loads (serviceability limit states) and at failure due to extreme storms (ultimate limit states), the constitutive modelling of the two soils was required to reproduce accurately both the soils' small strain stiffness and ultimate shear strength. An extended critical state based modified Cam clay (MCC) model was selected for the modelling of Cowden till, while a state parameter-based critical state bounding surface plasticity model was adopted for the Dunkirk sand, both available in the software ICFEP that was employed in all FE analyses. The key to the successful development of the robust and accurate numerical model was (i) careful characterisation of ground conditions at both sites, integrating historic experimental evidence with additional testing performed for each soil during the PISA project, as documented in Zdravković, Jardine *et al.* (2020); and (ii) appropriate calibration of the two constitutive models against the experimental data, as detailed in Zdravković, Taborda *et al.* (2020) and Taborda *et al.* (2020). As an example, Figure 15 shows the calibrated small strain stiffness, as (a) maximum shear modulus, G_0, and (b) non-linear degradation, and (c) an undrained shear strength profile in triaxial compression, $S_{u,TXC}$ for Cowden till.

Piles tested at both sites were 0.762 m and 2.0 m in diameter, with varying between 3 and 10, and the horizontal load applied at the top of an extension piece 10.0 m above the ground level. Figure 16 shows examples of the applied horizontal load versus ground level displacement predicted by 3D FE analyses and measured at each of the sites, as well as those calculated by the $p-y$ method. The very good agreement between 'blind' *Class A* 3D FE predictions and measurements at both working and ultimate loads (the latter taken as the load mobilised when horizontal ground level displacement reaches 10% of pile diameter) demonstrated the proof of concept, in that the advanced site-specific FE analysis of laterally-loaded monopiles was capable of accurately predicting the monopile response. The same figure also demonstrates the inadequacy of the $p-y$ design method. The mechanism of ground movement around the monopile in Figure 17 revealed that, apart from the soil reaction in terms of the mobilised lateral stress in the ground, additional soil reactions are contributed by the mobilised vertical shear stresses at the pile-soil interface, the base shear force and the base bending moment.

The success of this step unlocked the final development of a new design method, which involved further 3D FE analyses with ICFEP of full scale monopiles ($D = 5$ to 10 m and $L/D = 2$ to 6) in similar ground conditions, but adjusted for an offshore environment. The FE results were used as direct input into the formulation and parametrisation of all four soil reactions, as part of the new simplified Winkler-type PISA design methods for low-plasticity clay tills (Byrne, Houlsby *et al.*, 2020) and dense sands (Burd, Taborda *et al.*, 2020). To generalise the PISA methodology for more common layered seabed profiles, additional 3D FE analyses were performed with ground conditions comprising homogeneous stiff plastic clay, soft clay and sands of different densities, the results of which were extracted and parametrised for definition of soil reactions for these materials. The final development of the PISA design method for layered soils is detailed in Burd, Abadie *et al.* (2020). Its application in offshore industry was shown to reduce the amount of steel by around 30% in a typical monopile, thus reducing substantially the cost of foundations. As a consequence, the increased rate of offshore wind farm developments in the UK has resulted in wind-generated energy accounting for 24% of total electricity generation in 2020, also substantially reducing the cost of wind energy.

4.2. Geothermal as renewable energy

Another form of renewable energy that is increasingly being researched and explored in practice is the thermal energy beneath the ground surface. The ground temperature is largely constant throughout the year below a depth of about 10–15 m and is approximately equal to the mean annual air temperature. As such, in the summer it is lower than the temperature at the ground surface and in the winter it is higher. Ground source energy systems (GSES) utilise this temperature difference, enabling storage of excess heat in the ground during the summer and its extraction for heating in the winter. GSESs are classified as open- or closed-loop systems. The former exchange heat with the ground by directly pumping water into or out of the ground, while the latter rely on a fluid circulating through pipes buried in the ground. An advantage of closed-loop systems is that the heat-exchanger pipes can be installed in structural elements buried in the ground, such as building foundations, retaining walls or tunnel linings, providing not only stability but also renewable heating and cooling of space in buildings.

The most extensively studied geothermal structures to date are thermo-active piles, for which full-scale field experiments reported by Bourne-Webb *et al.* (2009); Brandl (2006) and Laloui *et al.* (2006), have led to the development of a simplified qualitative framework for thermo-

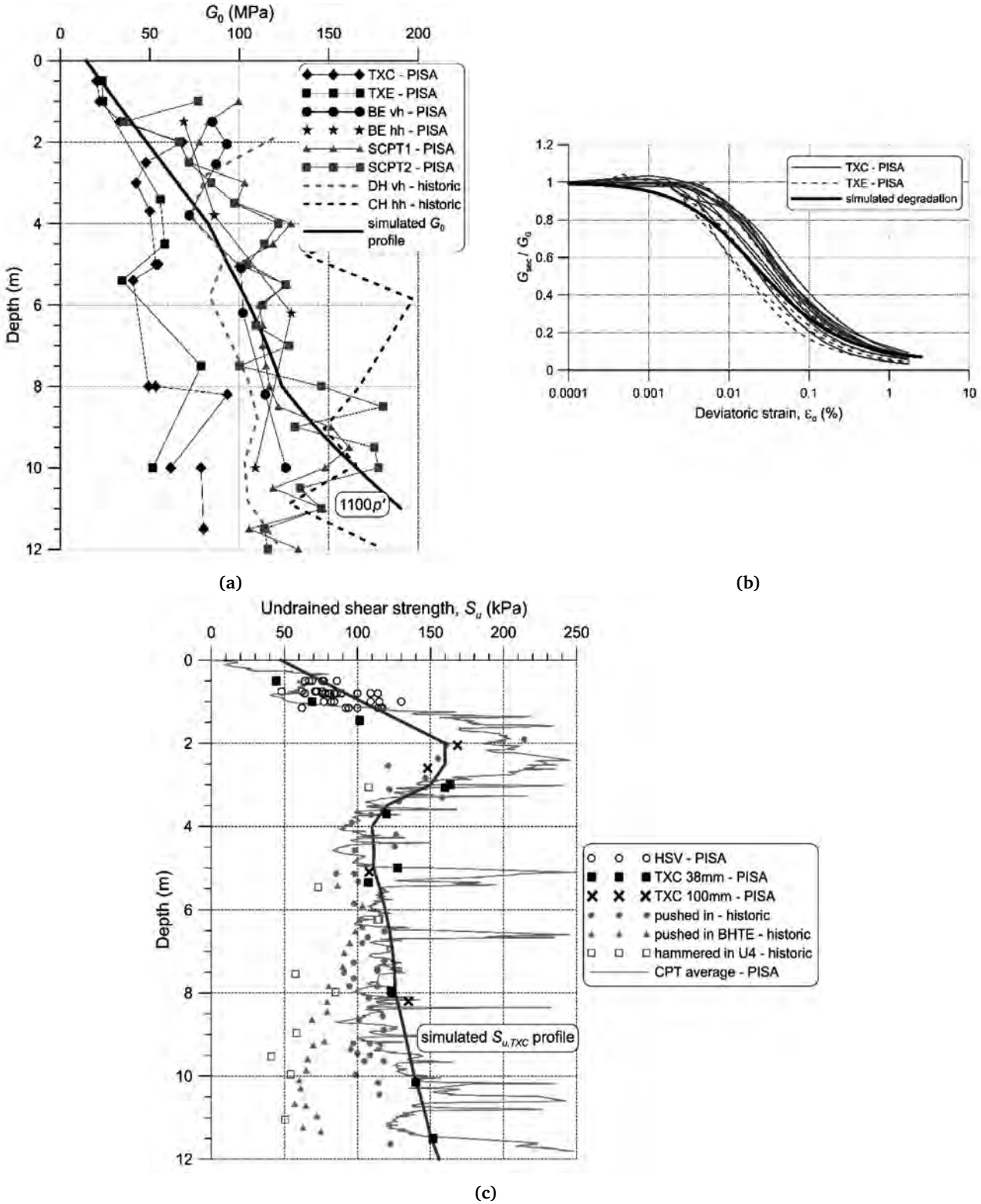


Figure 15. (a) Measured and interpreted profile for Cowden till;(b) Normalised shear stiffness degradation for Cowden till; (c) Measured and interpreted undrained shear strength profile for Cowden till

mechanical pile response (e.g. Amatya *et al.* 2012; Bourne-Webb *et al.* 2013). Physical, mainly centrifuge, experiments have provided further insights into the behaviour of thermo-active piles (e.g. Ng *et al.* 2015; Stewart and McCartney 2014; Yavari *et al.* 2014). Despite this body of research, the GSES technology currently presents design challenges for geotechnical engineers, particularly in predicting long-term temperature effects on both the structure and the surrounding soil, and usually requires advanced numerical analysis, as well as better understanding of soil behaviour under seasonal temperature changes (heating / cooling) induced by GSESs. The additional soil

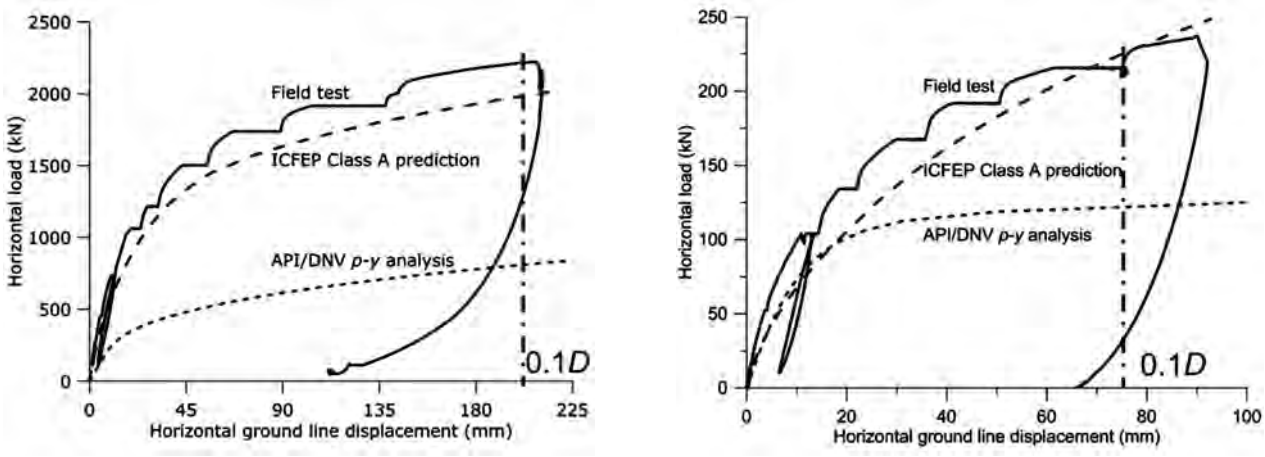


Figure 16. (a) Predicted and measured for Cowden test pile $D = 2.0$ m, $L/D = 5.25$; (b) Predicted and measured for Dunkirk test pile $D = 0.762$ m, $L/D = 5.25$

properties that become relevant for advanced calculations are usually thermal conductivity, volumetric heat capacity and coefficient of thermal expansion. Although very few issues have been reported regarding the serviceability of thermo-active piles (Loveridge and Powrie, 2013), this may be due to conservative design procedures rather than the lack of potentially adverse thermal effects. The magnitudes of thermally-induced axial stresses in the pile are of particular design concern, especially during the cooling phase (heat extraction) which causes shrinkage of the pile and hence tensile thermally induced stresses that can reduce into tension the total axial stress in the pile, leading to potential cracking of the pile.

The FE formulation for analysing this type of problems must be extended to include thermal coupling in addition to hydro-mechanical coupling in the governing FE equations, as well as appropriate thermal boundary conditions. The thermo-hydro-mechanical (THM) formulation and boundary conditions available in ICFEP were presented in Cui (2015); Cui *et al.* (2016) and Cui, Potts *et al.* (2018). Zdravković, Cui *et al.* (2021) summarised the whole process of assessing the modelling complexities of thermo-active piles and their accuracy. The accuracy of the THM finite element (FE) formulation was demonstrated by the back-analysis of the Lambeth College pile test (Bourne-Webb *et al.*, 2009), exploring at the same time the effects of the manner in which thermal loads were applied in the analysis (temperature or thermal flux over the pile body), as well as of the soil thermal properties (Gawecka, 2017; Gawecka *et al.*, 2017). The in-

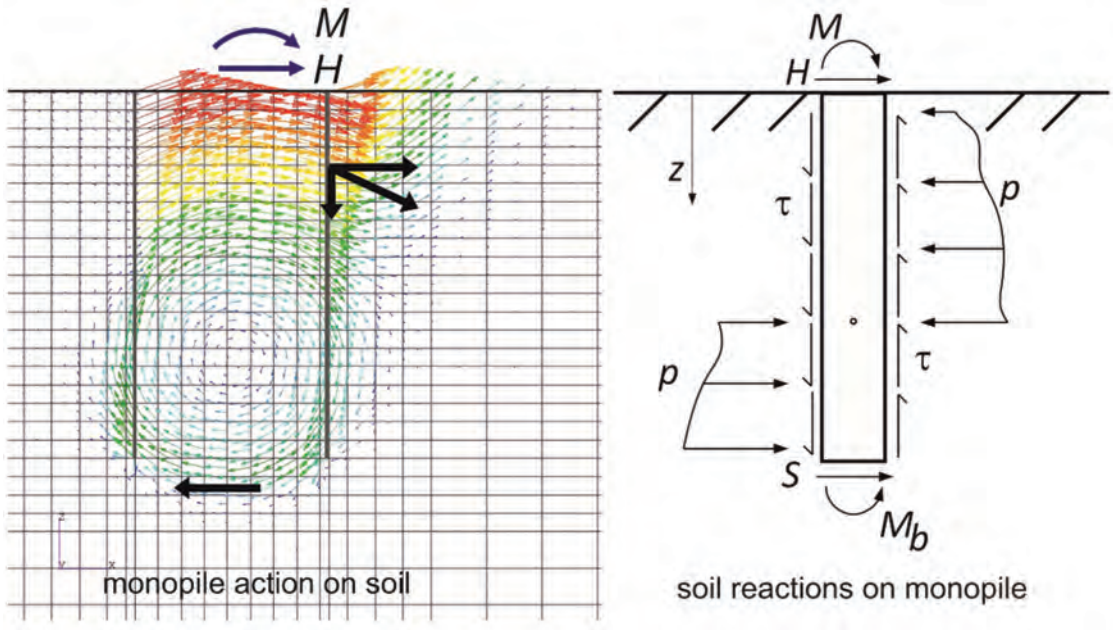
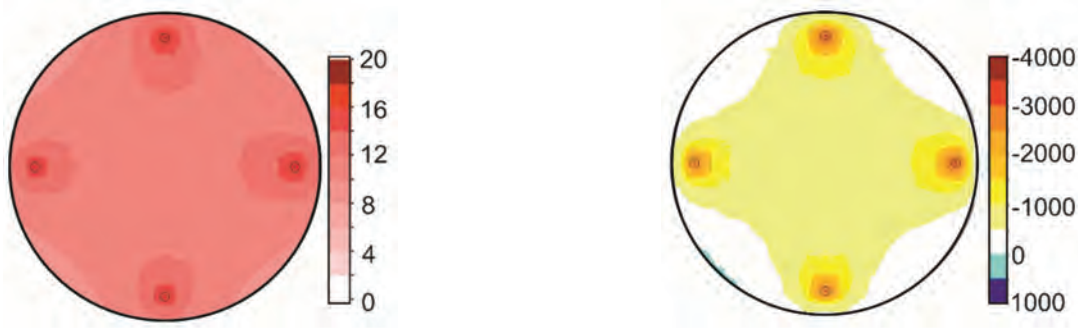


Figure 17. Mechanism of ground movements around a monopile and respective soil reactions



(a) **(b)**

Figure 18. Contours of (a) temperature change, ΔT (in °C; and (b) compressive thermally-induced axial stress, $\overline{\sigma}_{T,\max}$ (in kPa)

creasing complexity of the modelling was then discussed when introducing the heat-exchanger pipes directly in the analysis (Cui, Gawecka *et al.*, 2018; Gawecka *et al.*, 2018; 2020; Sailer, 2020; Sailer *et al.*, 2019) and capabilities of the newly proposed modelling approach were validated on two case studies of thermal response tests (Beier *et al.* 2011 and Loveridge *et al.* 2014).

Having demonstrated the accuracy and robustness of the modelling approach for thermo-active piles, a detailed three-dimensional (3D) THM modelling of a single thermo-active pile was performed (Liu *et al.*, 2020) assuming one, two or three U-loops that represent possible arrangements of heat-exchanger pipes installed in a bored pile, which is usually 600 mm or 900mm in diameter for thermo-active pile foundations in London clay. In the analyses the pile was initially loaded to working load (1800 kN and 2900 kN, respectively for a 23 m long pile 600 or 900 mm in diameter) in a London clay ground profile that was prescribed an initial temperature $T = 19.5$°C, as measured at Lambeth College test site (Bourne-Webb *et al.*, 2009). This was followed by activating the flow of hot water in the pipes (heating cycle, or heat storage), simulated by injecting hot water at a temperature of 39.5°C (i.e. $\Delta T = 20$°C) for a duration of 5 months and applying appropriate boundary conditions to maintain circulation, during which the thermally-induced axial stresses in the pile were monitored. As an example, Figure 18 shows very clearly the non-uniform variation of the (a) temperature change and of (b) thermally-induced axial stress (when it reached a maximum in the heating cycle) in the cross-sections of a 900 mm diameter thermo-active pile, with a two U-loop arrangement of heat exchanger pipes. The small circles in each figure define positions of the pipes (inlets and outlets), while the cross-sections is taken at 17 m depth (which is the depth of a maximum thermally induced axial stress in the 23 m long pile).

Figure 19 plots the predicted evolution of the maximum thermally induced compressive axial stress, $\overline{\sigma}_{T,\max}$, for each U-loop arrangement in two pile geometries, during the 5 months of a heating cycle. The results reveal the highly transient nature of this evolution, as well as substantial peak values of $\overline{\sigma}_{T,\max}$. The importance of transient behaviour is further illustrated in Figure 20, where the calculated values of $\overline{\sigma}_{T,\max}$ were compared to those indicated in the current design charts proposed by GSHPA (2012) for the same pile diameters, loading conditions and temperature changes. It is evident that the peak values of $\overline{\sigma}_{T\,\max}$ predicted in the 3D THM coupled analyses, which would be used for pile design, are significantly larger than those suggested by the design charts and that this difference reduces during the transient period. A possible reason for such large differences (up to 280%) is in the different modelling approaches, as the design charts were developed using the load-transfer ($t - z$) method, which is unable to model the non-isothermal response of the soil.

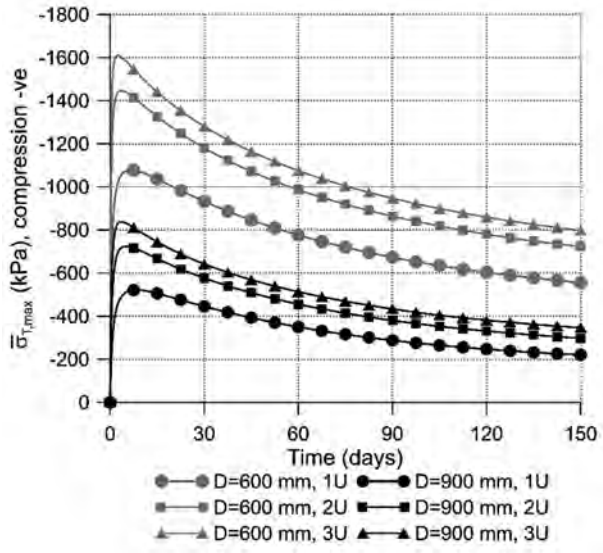


Figure 19. Predicted evolution of thermally-induced axial stress, $\overline{\sigma}_{T,\max}$, with time

The 3D THM analyses of Liu *et al.* (2020) were essential to understanding the complexities of the heat exchanger pipes / pile interaction and pile / soil interaction. They are, however, computationally demanding for regular application in design. A further study by Liu *et al.* (2020) proposed a simplified 2-step approach involving 2D FE analyses, which were demonstrated to predict the

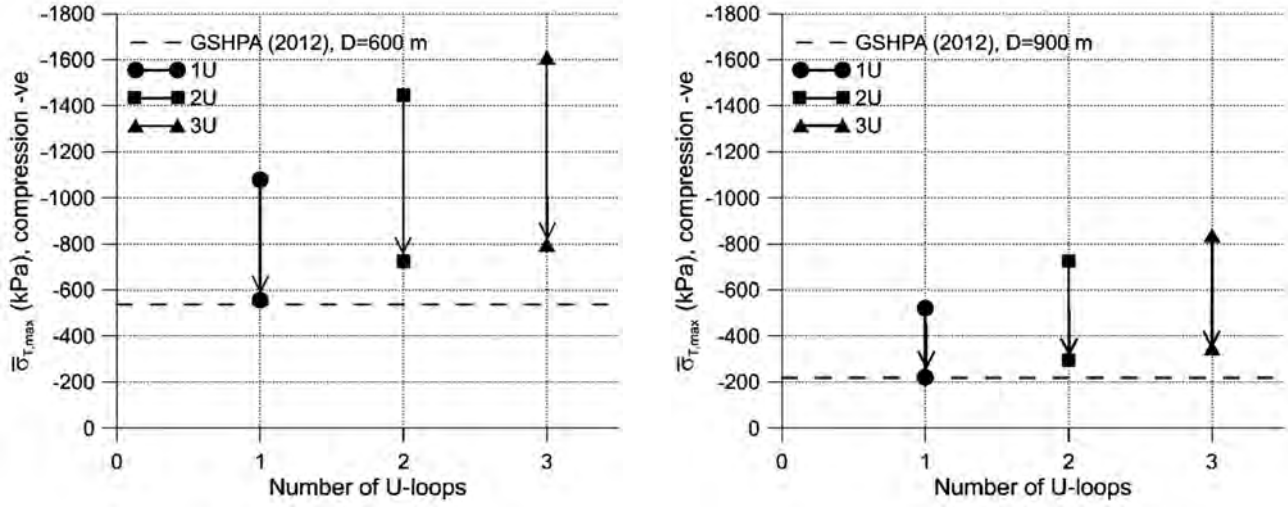


Figure 20. (a) Comparison of predicted $\overline{\sigma}_{T,\max}$ with GSHPA (2012) design charts for 600 mm diameter pile; (b) Comparison of predicted $\overline{\sigma}_{T,\max}$ with GSHPA (2012) design charts for 900 mm diameter pile

thermally-induced axial stresses in the pile with a similar accuracy to 3D analysis, but less computationally demanding.

4.3. Life extension of existing infrastructure

Ageing infrastructure is a common geotechnical issue, comprising systems that have already been in use for several decades, but have to stay in place for a foreseeable future. For example, foundation systems that have supported an urbanised site which is undergoing redevelopment and existing foundations must be re-used as they cannot be removed. Road and rail infrastructure and flood defences are other examples, the latter structures having to be raised to sustain the projected sea level rises as a consequence of climate change. Common to these examples is that their re-use / life extension usually involves higher loads compared to those originally imposed on the ground and the main design question is how much higher loads the soil can sustain after being already loaded for several decades.

The example discussed here considers earthfill embankments as common geotechnical structures for supporting networks of road and rail infrastructure and serving as flood defences along river banks and coastlines. They are often erected on soft clay deposits whose undrained strength limits the construction height of a single-stage embankment to 3 to 4 m. Bjerrum (1973) postulated that the shear strength anisotropy of soft clay deposits may be the key parameter for the short-term embankment design, as its structure imposes large rotations of principal stresses in the foundation soil, promoting the mobilisation of variable soil shear strength along a potential failure surface. In advanced FE modelling, the effect of shear strength anisotropy can be reasonably quantified through the application of appropriate soil constitutive models (e.g. Karstunen *et al.* 2005; Kavvadas and Amorosi 2000; Whittle and Kavvadas 1994; Zdravković *et al.* 2002).

However, during the transient post-construction period the foundation soil is subjected to the time-related process of consolidation, as discussed earlier in this manuscript, and additionally to creep, the former dissipating the construction-generated pore water pressures in the ground and the latter accounting for changes in the soil fabric. These processes can induce large settlements, thus requiring the raising of the embankment height so that it can continue to fulfil its purpose. The reduction of void space in the foundation soil leads to an increase of its undrained shear strength. Again, advanced numerical analysis and appropriate soil constitutive models are needed to quantify the likely gain in undrained shear strength and the likely height to which an embankment can be raised without failing (e.g. Bodas Freitas *et al.* 2012; 2011; Karstunen *et al.* 2005; Yin and Graham 1999).

Studying the trial road embankments constructed on soft Champlain deposits in east Canada and using the anisotropic critical state-based MIT-E3 constitutive model (Whittle and Kavvadas, 1994), implemented in ICFEP to represent the behaviour of this soft clay, (Zdravković *et al.*, 2002) were able to quantify the effect of undrained shear strength anisotropy in predicting accurately the embankment height at failure under fast construction. The work demonstrated the importance of taking account of this soil characteristic for the short-term design (i.e. before transient effects become dominant). Figure 21, which plots normalised horizontal ground movement at the embankment toe, shows an accurate prediction of the failure height (at 3.9 m, as observed in the field test) of the studied Sain-Alban trial embankment when employ-

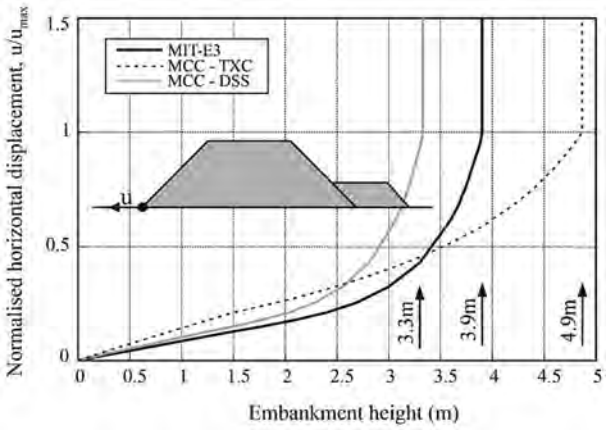


Figure 21. Predicted failure of Saint Alban embankment with different modelling approaches

ing the anisotropic MIT-E3 constitutive model in the analysis, and significant over-prediction (at 4.9 m) and under-prediction (at 3.3 m) when employing an isotropic modified Cam clay (MCC) model which can only represent an isotropic undrained shear strength, taken here as being that in triaxial compression (TXC) or in direct simple shear (DSS), according to Bjerrum (1973), respectively.

The same MIT-E3 model, however, when applied in a subsequent study of flood embankments in the Thames estuary (Zdravković *et al.*, 2019), where ground movements were monitored for a period of time post-construction, was not able to predict accurate settlements under the monitored embankment BANK 2. As shown in Figure 22a, the predicted settlement troughs at two time instances post-construction, were significantly lower than measured. This was attributed to the MIT-E3 model being able to simulate only consolidation post-construction, making it evident that, for the transient/long term modelling of the embankment, additional facets of time-dependent soil behaviour, such as creep or viscosity, needed to be considered. The fact that MIT-E3 takes account of soil anisotropy did not seem so important for the transient process.

A number of constitutive models have been proposed in the literature to simulate the time-dependent behaviour of soils (mainly clays) that follow isotach viscosity (Liingaard *et al.*, 2004). Most of these are elastic-viscoplastic models, based on Perzyna (1963) overstress theory. The model selected for the current study adopts the same concept, in combination with the Yin and Graham (1999) equivalent time (ET) framework and can therefore account for both consolidation and creep settlements. Details of the model, termed IC-ET, its generalisation in stress space and its implementation in ICFEP are given in Bodas Freitas *et al.* (2012; 2011). The model is isotropic and was therefore calibrated to reproduce the undrained shear strength in triaxial compression, similar to that reproduced by the MIT-E3 model. The main difference is that in the time-dependent framework the undrained shear strength in the analysis has to be consistent with the strain rate at which the shearing in the experiment was conducted. The finite element prediction of settlements under BANK 2, using the IC-ET model, was significantly more accurate when compared to field measurements in Figure 22b.

This has a direct impact on predicting the life extension of existing embankments in terms of raising their height, as shown schematically in Figure 23 with two possible scenarios. In both cases, the analyses of BANK 2 were extended for a period of 30 years, with the soil modelled either with the MIT-E3 or IC-ET models respectively. The initial height of this embankment was 2.8 m. Modelling the soil with MIT-E3 predicted that, after 30 years of consolidation only, the embankment could be raised to 3.9 m and 4.7 m under schemes A and B, respectively. On the other hand, allowing for 30 years of both consolidation and creep, he IC-ET soil model predicted that the embankment could be raised to a height of 5.3 m and 5.6 m for schemes A and B, respectively, being nearly a 100% increase of the original embankment height.

The above two case studies demonstrate the importance of accounting for soil shear strength anisotropy when designing the construction phase of earth embankments on soft soils. For long-term deformations, however, capturing creep deformations, in addition to consolidation seems to be significantly more important for the robust redesign of ageing infrastructure.

Examples in this section have highlighted the many challenges facing both geotechnical research and practice now and for the foreseeable future. Most of these have to do with long term assessments of stability and resilience of geotechnical structures. The likely questions that will require further research have to do with the cyclic loading resilience of offshore infrastructure; soil-atmosphere interaction in conjunction with climate change and its effect in particular on exposed infrastructure, such as slopes and embankments; safe long-term disposal of high-level nuclear waste in the natural environment; the need to expand geotechnical developments to more challenging ground conditions that have not been explored before; to mention but a few. These topics are already receiving attention, but convincing design recommendations still require further work.

5. FINAL REMARKS

Using a number of practical examples, this paper provides an overview of the development of the geotechnical engineering discipline and challenges that it has faced to date and those that still remain. It is evident that design solutions will need to continue to be addressed by robust and sophisticated numerical modelling, which will need to continue to be supported by equally advanced soil characterisation through laboratory and field testing. The field

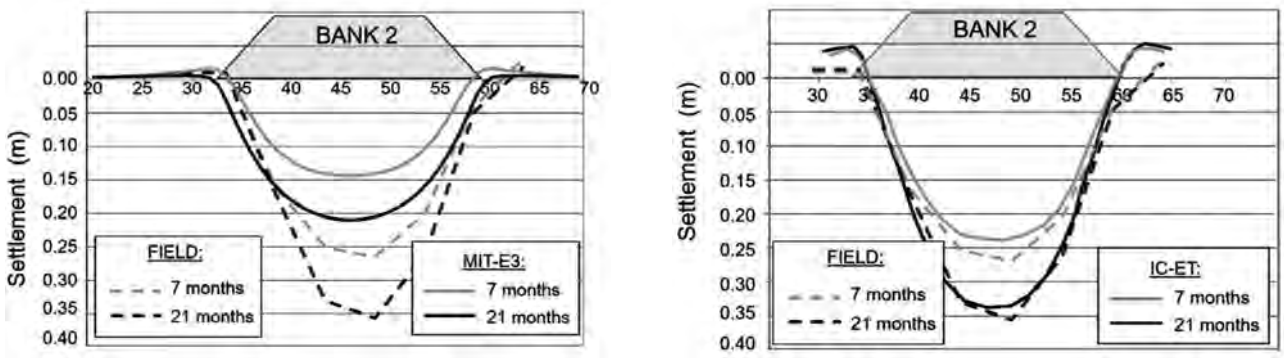


Figure 22. (a) Post-construction settlements of BANK 2; predictions accounting for consolidation; (b) Post-construction settlements of BANK 2; predictions accounting for consolidation and creep

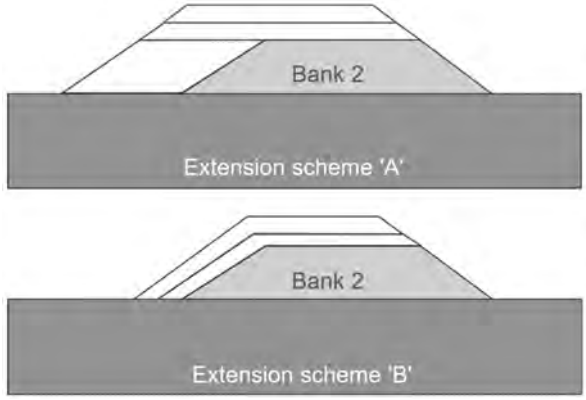


Figure 23. Simulated raising schemes applied to BANK 2

monitoring should become more frequent in the future, to enable further validation of numerical procedures.

It is emphasised that establishing a realistic ground model is challenging and requires careful integration of both field and laboratory data obtained from experimental campaigns. It is further emphasised that the selection of a constitutive model with which to simulate the soil is equally challenging and a function of understanding the nature of the problem to be analysed as well as the available experimental evidence.

All these processes require significant engineering judgement when deriving the numerical input, which in turn relies on an equal understanding of the numerical tools and of the real soil behaviour.

REFERENCES

Addenbrooke, T., Potts, D. and Puzrin, A. (1997). The influence of pre-failure soil stiffness on the numerical analysis of tunnel construction. *Géotechnique*, *47*(3), 693–712.

Amatya, B., Soga, K., Bourne-Webb, P., Amis, T. and Laloui, L. (2012). Thermo-mechanical behaviour of energy piles. *Géotechnique*, *62*(6), 503–519. doi: 10 .1680 / geot .10 .P .116

API. (2010). *Rp 2a-wsd–recommended practice for planning, designing and constructing fixed offshore platforms.* American Petroleum Institute Washington, DC, USA.

Atkinson, J. and Sallfors, G. (1991). Experimental determination of soil properties. *General Report to Session 1, Proceedings of the 10th European Conference on Soil Mechanics and Foundation Engineering*, *3*, 915–956.

Avgerinos, V., Potts, D. and Standing, J. (2017). Numerical investigation of the effects of tunnelling on existing tunnels. *Géotechnique*, *67*(9), 808–822.

Beier, R. A., Smith, M. D. and Spitler, J. D. (2011). Reference data sets for vertical borehole ground heat exchanger models and thermal response test analysis. *Geothermics*, *40*(1), 79–85.

Biot, M. A. (1941). General theory of three-dimensional consolidation. *Journal of applied physics*, *12*(2), 155–164.

Bishop, A. and Wesley, L. (1975). A hydraulic triaxial apparatus for controlled stress path testing. *Geotechnique*, *25*(4), 657–670.

Bjerrum, L. (1973). Problems of soil mechanics and construction on soft clays. *State-of-the-Art Report, Session 4, Proc. 8th ICSMFE, Moscow*, *3*, 109–159.

Bodas Freitas, T., Potts, D. and Zdravkovic, L. (2012). Implications of the definition of the ϕ function in elastic-viscoplastic models. *Géotechnique*, *62*(7), 643–648.

Bodas Freitas, T., Potts, D. M. and Zdravkovic, L. (2011). A time dependent constitutive model for soils with isotach viscosity. *Computers and Geotechnics*, *38*(6), 809–820.

Bourne-Webb, P. J., Amatya, B. and Soga, K. (2013). A framework for understanding energy pile behaviour. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, *166*(2), 170–177. doi: 10 .1680 / geng .10 .00098

Bourne-Webb, P. J., Amatya, B., Soga, K., Amis, T., Davidson, C. and Payne, P. (2009). Energy pile test at lambeth college, london: geotechnical and thermodynamic aspects of pile response to heat cycles. *Géotechnique*, *59*(3), 237–248. doi: 10 .1680 / geot .2009 .59 .3 .237

Brandl, H. (2006). Energy foundations and other thermo-active ground structures. *Géotechnique*, *56*(2), 81–122. doi: 10 .1680 / geot .2006 .56 .2 .81

Burd, H. J., Abadie, C. N., Byrne, B. W., Houlsby, G. T., Martin, C. M., McAdam, R. A., Jardine, R. J., Pedro, A. M., Potts, D. M., Taborda, D. M. et al. (2020). Application of the pisa design model to monopiles embedded in layered soils. *Géotechnique*, *70*(11), 1067–1082.

Burd, H. J., Beuckelaers, W. J., Byrne, B. W., Gavin, K. G., Houlsby, G. T., Igoe, D. J., Jardine, R. J., Martin, C. M., McAdam, R. A., Muir Wood, A. et al. (2020). New data analysis methods for instrumented medium-scale monopile field tests. *Géotechnique*, *70*(11), 961–969.

Burd, H. J., Taborda, D. M., Zdravković, L., Abadie, C. N., Byrne, B. W., Houlsby, G. T., Gavin, K. G., Igoe, D. J., Jardine, R. J., Martin, C. M. et al. (2020). Pisa design model for monopiles for offshore wind turbines: application to a marine sand. *Géotechnique*, *70*(11), 1048–1066.

Byrne, B. W., Houlsby, G. T., Burd, H. J., Gavin, K. G., Igoe, D. J., Jardine, R. J., Martin, C. M., McAdam, R. A., Potts, D. M., Taborda, D. M. et al. (2020). Pisa design model for monopiles for offshore wind turbines: application to a stiff glacial clay till. *Géotechnique*, *70*(11), 1030–1047.

Byrne, B. W., McAdam, R. A., Burd, H. J., Beuckelaers, W. J., Gavin, K. G., Houlsby, G. T., Igoe, D. J., Jardine, R. J., Martin, C. M., Muir Wood, A. et al. (2020). Monotonic laterally loaded pile testing in a stiff glacial clay till at cowden. *Géotechnique*, *70*(11), 970–985.

Casagrande, A. (1932). Research on the atterberg limits of soils. *Public roads*, *13*(8), 121–136.

Cui, W. (2015). *Development, implementation and application of thermo-hydro-mechanical coupling for soils in finite element analysis.* (PhD thesis, Imperial College London)

Cui, W., Gawecka, K., Potts, D., Taborda, D. and Zdravković, L. (2016). Numerical analysis of coupled thermo-hydraulic problems in geotechnical engineering. *Geomechanics for Energy and the Environment*, *6*, 22–34. doi: 10 .1016 / j .gete .2016 .03 .002

Cui, W., Gawecka, K., Potts, D., Taborda, D. and Zdravković, L. (2018). A petrov-galerkin finite element method for 2d transient and steady state highly advective flows in porous media. *Computers and Geotechnics*, *100*, 158–173. doi: 10 .1016 / j .compgeo .2018 .04 .013

Cui, W., Potts, D. M., Zdravković, L., Gawecka, K. A. and Taborda, D. M. (2018). An alternative coupled thermo-hydro-mechanical finite element formulation for fully saturated soils. *Computers and Geotechnics*, *94*, 22–30. doi: 10 .1016 / j .compgeo .2017 .08 .011

Darcy, H. (1856). *Les fontaines publiques de la ville de dijon: exposition et application...* Victor Dalmont.

DNVGL. (2016). *Dnvgl-st-0126: Support structures for wind turbines.*

Gawecka, K. A. (2017). *Numerical analysis of geothermal piles* (PhD Thesis). Imperial College London.

Gawecka, K. A., Potts, D. M., Cui, W., Taborda, D. M. and Zdravković, L. (2018). A coupled thermo-hydro-mechanical finite element formulation of one-dimensional beam elements for three-dimensional analysis. *Computers and Geotechnics*, *104*, 29–41. doi: 10 .1016 / j .compgeo .2018 .08 .005

Gawecka, K. A., Taborda, D. M., Potts, D. M., Cui, W., Zdravković, L. and Haji Kasri, M. S. (2017). Numerical modelling of thermo-active piles in london clay. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, *170*(3), 201–219. doi: 10 .1680 / jgeen .16 .00096

Gawecka, K. A., Taborda, D. M., Potts, D. M., Sailer, E., Cui, W. and Zdravković, L. (2020). Finite-element modeling of heat transfer in ground source energy systems with heat exchanger pipes. *International Journal of Geomechanics*, *20*(5), 04020041. doi: 10 .1061 / (ASCE) GM .1943 -5622 .0001658

Goto, S., Tatsuoka, F., Shibuya, S., Kim, Y. and Sato, T. (1991). A simple gauge for local small strain measurements in the laboratory. *Soils and foundations*, *31*(1), 169–180.

Henkel, D. (1960). The relationships between the effective stresses and water content in saturated clays. *Geotechnique*, *10*(2), 41–54.

Hight, D., Gens, A. and Symes, M. (1983). The development of a new hollow cylinder apparatus for investigating the effects of principal stress rotation in soils. *Geotechnique*, *33*(4), 355–383.

Hight, D., Potts, D. and Tjelta, T. (1988). Finite element analyses of a deep skirted gravity base. In *5th int. conf. on behaviour of offshore structures* (pp. 209–226). Trondheim, (Tapir).

Ishihara, K. and Yasuda, S. (1975). Sand liquefaction in hollow cylinder torsion under irregular excitation. *Soils and Foundations*, *15*(1), 45–59.

Jardine, R. and Potts, D. (1988). Hutton tension leg platform foundations: prediction of driven pile behaviour. *Geotechnique*, *38*(2), 231–252.

Jardine, R., Potts, D., Fourie, A. and Burland, J. (1986). Studies of the influence of non-linear stress–strain characteristics in soil–structure interaction. *Geotechnique*, *36*(3), 377–396.

Jardine, R., Symes, M. and Burland, J. (1984). The measurement of soil stiffness in the triaxial apparatus. *Geotechnique*, *34*(3), 323–340.

Jurečič, N., Zdravković, L. and Jovičić, V. (2013). Predicting ground movements in london clay. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, *166*(5), 466–482.

Karstunen, M., Krenn, H., Wheeler, S. J., Koskinen, M. and Zentar, R. (2005). Effect of anisotropy and destructuration on the behavior of murro test embankment. *International Journal of Geomechanics*, *5*(2), 87–97.

Kavvadas, M. and Amorosi, A. (2000). A constitutive model for structured soils. *Géotechnique*, *50*(3), 263–273.

Kovacevic, N. (1994). *Numerical analyses of rockfill dams, cut slopes and road embankments* (PhD Thesis). Imperial College London (University of London).

Kovacevic, N., Hight, D. and Potts, D. (2007). Predicting the stand-up time of temporary london clay slopes at terminal 5, heathrow airport. *Géotechnique*, *47*(5), 953–982.

Laloui, L., Nuth, M. and Vulliet, L. (2006). Experimental and numerical investigations of the behaviour of a heat exchanger pile. *International journal for numerical and analytical methods in geomechanics*, *30*(8), 763–781.

Lees, A. (2016). Contents and preliminary pages. In *Geotechnical finite element analysis: A practical guide* (pp. i–ix). ICE Publishing.

Liingaard, M., Augustesen, A. and Lade, P. V. (2004). Characterization of models for time-dependent behavior of soils. *International Journal of Geomechanics*, *4*(3), 157–177.

Liu, R. Y. W., Sailer, E., Taborda, D. M., Potts, D. M. and Zdravković, L. (2020). A practical method for calculating thermally-induced stresses in pile foundations used as heat exchangers. *Computers and Geotechnics*, *126*, 103743. doi: 10 .1016 / j .compgeo .2020 .103743

Lo Presti, D., Pallara, O., Costanzo, D. and Impavido, M. (1994). Small strain measurements during triaxial tests: Many problems, some solutions. In *First int. conf. on pre-failure deformation characteristics of geomaterials* (Vol. 1, pp. 11–16). Sapporo-Japan.

Loveridge, F. and Powrie, W. (2013). Pile heat exchangers: thermal behaviour and interactions. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering*, *166*(2), 178–196. doi: 10 .1680 / geng .11 .00042

Loveridge, F., Powrie, W. and Nicholson, D. (2014). Comparison of two different models for pile thermal response test interpretation. *Acta Geotechnica*, *9*(3), 367–384.

McAdam, R. A., Byrne, B. W., Houlsby, G. T., Beuckelaers, W. J., Burd, H. J., Gavin, K. G., Igoe, D. J., Jardine, R. J., Martin, C. M., Muir Wood, A. et al. (2020). Monotonic laterally loaded pile testing in a dense marine sand at dunkirk. *Géotechnique*, *70*(11), 986–998.

Ng, C. W. W., Shi, C., Gunawan, A., Laloui, L. and Liu, H. (2015). Centrifuge modelling of heating effects on energy pile performance in saturated sand. *Canadian Geotechnical Journal*, *52*(8), 1045–1057.

Pedro, A., Zdravković, L., Potts, D. and e Sousa, J. A. (2018). Geotechnical characterization of the miocene formations at the location of ivens shaft, lisbon. *Quarterly Journal of Engineering Geology and Hydrogeology*, *51*(1), 96–107.

Perzyna, P. (1963). The constitutive equation for work-hardening and rate sensitive plastic materials. In *Proc. vibrational problems* (Vol. 4, pp. 281–290).

Potts, D.M., Dounias, G. and Vaughan, P. (1990). Finite element analysis of progressive failure of carsington embankment. *Geotechnique*, *40*(1), 79–101.

Potts, D.M., Kovacevic, N. and Vaughan, P. (1997). Delayed collapse of cut slopes in stiff clay. *Géotechnique*, *47*(5), 953–982.

Potts, D. M. and Zdravković, L. (1999). *Finite element analysis in geotechnical engineering: Theory*. Thomas Telford.

Potts, D. M. and Zdravković, L. (2001). *Finite element analysis in geotechnical engineering: application*. Thomas Telford London.

Poulos, H. G. and Davis, E. H. (1974). *Elastic solutions for soil and rock mechanics*. Wiley New York.

Reese, L., Cox, W. and Koop, F. (1974). *Analysis of laterally loaded piles in sand, offshore technical conference, dallas, texas, paper no.* Offshore Technology Conference, Dallas, Texas.

Reese, L. C., Cox, W. R. and Koop, F. D. (1975). Field testing and analysis of laterally loaded piles om stiff clay. In *Offshore technology conference.*

Roscoe, K. and Burland, J. (1968). *On the generalized stress-strain behaviour of wet clay*. Cambridge University Press, New York.

Sailer, E. (2020). *Numerical modelling of thermo-active retaining walls* (PhD Thesis). PhD thesis. Imperial College London.

Sailer, E., Taborda, D. M., Zdravković, L. and Potts, D. M. (2019). Fundamentals of the coupled thermo-hydro-mechanical behaviour of thermo-active retaining walls. *Computers and Geotechnics*, *109*, 189-203. doi: 10 .1016 / j .compgeo .2019 .01 .017

Schofield, A. N. and Wroth, P. (1968). *Critical state soil mechanics* (Vol. 310). McGraw-hill London.

Smith, I. and Griffiths, D. (1988). *Programming the finite element method*. John Wiley & Sons, New York.

Stewart, M. A. and McCartney, J. S. (2014). Centrifuge modeling of soil-structure interaction in energy foundations. *Journal of Geotechnical and Geoenvironmental Engineering*, *140*(4), 04013044.

Taborda, D. M., Zdravković, L., Potts, D. M., Burd, H. J., Byrne, B. W., Gavin, K. G., Houlsby, G. T., Jardine, R. J., Liu, T., Martin, C. M. et al. (2020). Finite-element modelling of laterally loaded piles in a dense marine sand at dunkirk. *Géotechnique*, *70*(11), 1014–1029.

Terzaghi, K. (1966). *Soil mechanics in engineering practice*. John Wiley & Sons.

Timoshenko, S. P. (1934). *Theory of elasticity.* McGraw-Hill, New York.

Whittle, A. J. and Kavvadas, M. J. (1994). Formulation of mit-e3 constitutive model for overconsolidated clays. *Journal of Geotechnical engineering*, *120*(1), 173–198.

Yavari, N., Tang, A. M., Pereira, J.-M. and Hassen, G. (2014). Experimental study on the mechanical behaviour of a heat exchanger pile using physical modelling. *Acta Geotechnica*, *9*(3), 385–398.

Yin, J.-H. and Graham, J. (1999). Elastic viscoplastic modelling of the time-dependent stress-strain behaviour of soils. *Canadian Geotechnical Journal*, *36*(4), 736–745.

Zdravković, L., Cui, W., Gawecka, K., Liu, R., Potts, D., Sailer, E. and Taborda, D. (2021). Numerical modelling of thermo-active piles. In *Piling 2020: Proceedings of the piling 2020 conference* (pp. 35–59).

Zdravković, L., Jardine, R. J., Taborda, D. M., Abadias, D., Burd, H. J., Byrne, B. W., Gavin, K. G., Houlsby, G. T., Igoe, D. J., Liu, T. et al. (2020). Ground characterisation for pisa pile testing and analysis. *Géotechnique*, *70*(11), 945–960.

Zdravković, L., Potts, D. and Freitas, T. B. (2019). Extending the life of existing infrastructure. In *Vii ecsmge-2019 – geotechnical engineering foundation of the future.* Reykjavik, Iceland. doi: 10 .32075 / 17ECSMGE -2019 -1104

Zdravković, L., Potts, D. and Hight, D. (2002). The effect of strength anisotropy on the behaviour of embankments on soft ground. *Géotechnique*, *52*(6), 447–457.

Zdravkovic, L., Potts, D. and St John, H. (2005). Modelling of a 3d excavation in finite element analysis. *Geotechnique*, *55*(7), 497–513.

Zdravković, L., Potts, D. M. and Taborda, D. M. (2021). Integrating laboratory and field testing into advanced geotechnical design. *Geomechanics for Energy and the Environment*, *27*, 100216.

Zdravković, L., Taborda, D. M., Potts, D. M., Abadias, D., Burd, H. J., Byrne, B. W., Gavin, K. G., Houlsby, G. T., Jardine, R. J., Martin, C. M. et al. (2020). Finite-element modelling of laterally loaded piles in a stiff glacial clay till at cowden. *Géotechnique*, *70*(11), 999–1013.

Zienkiewicz, O. C. (1971). The finite element method. *London: McGraw-Hill*.

Zienkiewicz, O. C., Chan, A., Pastor, M., Schrefler, B. and Shiomi, T. (1999). *Computational geomechanics* (Vol. 613). John Wiley & Sons, Chichester.

Enrique Castro Rodríguez

Enrique Castro Rodríguez is an Associate Professor of at the Department of Applied Physics and Naval Technology of the Polytechnic University of Cartagena (Spain) and Coordinator of International Master Programs of the Technical University of Cartagena.

Contact information: e-mail: enrique.castro@upct.es

Prof. Castro received the BSc degree in physics in 2000 and the PhD in 2005 from the University of Granada. He carried out his PhD with a FPU grant at the Department of Physics of the University of Jaen, where he was working during an academic course as assistant lecturer. Afterwards, he obtained a "Juan de la Cierva" postdoctoral position at the Department of Applied Physics of the University of Granada. From 2009 he is employed at the Technical University of Cartagena. He has been the Vice-Dean of International Relationships of the Civil and Mining Engineering Faculty, Technical University of Cartagena.

Prof. Castro teaches thermal engineering and acoustics courses at Bachelor and Master levels and supervises MSc students in these areas. He conducts research on Heat Transmission, Structural Vibrations and Acoustic Emission, and collaborates intensively with industry in these fields. This collaboration has led to 5 research projects funded by private companies with Prof. Castro as a principal investigator. He has also participated in 16 public and private research projects and 5 educational projects for innovative teaching, being coordinator of two of them. His current projects include Detection and prediction of corrosion in the reinforcement of reinforced concrete structures (ECO), Development and simulation of predictive thermal models and thermal properties characterization of asphalt-HEATLAND and FLUITERM - New method for the determination of thermal conductivity in fluid, viscous or granular materials.

He is the coauthor of 21 international publications and co-inventor of one patent.

Applied Physics in Civil Engineering: Acoustic Emission and Cool Pavements

Enrique Castro Rodriguez

Department of applied physics and naval technology, universidad politécnica de cartagena, Spain

Summary

Applied Physics is a science field that study the application of physics to practical and engineering problems. Specific problems of an engineering field are usually solved more efficiently by engineers from that field, so applies physics scholars must focus on interdisciplinary problems or in applications where a high knowledge and comprehension of the physics of the systems is required.

To illustrate the contributions of Applied Physics to Civil Engineering, two very different applications are shown. The first one is Acoustic Emission technique, a non-destructive evaluation method which is based on elastic wave motion. It is a very useful evaluation method, but its results are difficult to interpretate and require specialised knowledge. A review of Acoustic Emission applications in Civil Engineering is presented.

In the second part, a specific application of Applied Physics in Civil Engineering is shown. It is an example of how Applied Physics can help other field when it requires physics basics that are beyond its usual needs. It is the realization of thermal analysis of the use of cool pavements in city roads to reduce the urban heat island effect.

1. INTRODUCTION

This paper has the objective to show some of the applications of Applied Physics in Civil Engineering. In order to carry out this objective, the first problem that we face is that it is difficult to define what Applied Physics is. Although it is widely used, Applied Physics is a term that can be considered ambiguous. At first, it can be seen as the part of the Physics that is more focused on engineering and practical applications, in opposition to Fundamental Physics that is dedicated to study the basics of reality and the fundamental principles of Physics. The problem is that Physics, as a field of study, is divided between many disciplines which are connected with a specific kind of systems or phenomena. In these disciplines, usually there are parts of them more connected to applications and others to the basics of the field. For example, Electromagnetism has countless applications, but Quantum Electrodynamics, which is the basics of Particle Physics and is fundamental to explain the behaviour at subatomic scales, has not yet, up to my knowledge, applications in any field of Engineering.

Although it is difficult to define Applied Physics as a science field, it is common accepted that it studies the application of Physics to practical problems, specially from other non-physics disciplines. Scholars from Applied Physics conduct research on Material Science, Biophysics, Acoustics, Instrumentation, etc. One thing that characterised Applied Physics is that it is an interdisciplinary science. It requires knowledge from other disciplines to be able to solve their practical problems and overcome their difficulties.

2. APPLIED PHYSICS IN CIVIL ENGINEERING

Civil Engineering, as any other engineering, makes extensive use of Physics. It is the basics of almost all of its disciplines, and the competences, knowledges and abilities that are learn by studying Physics can be applied to study other Civil Engineering courses. As examples of Physics basics in Civil Engineering disciplines, we have that Structural Calculation is based on Newton's Laws, Seismic Engineering requires knowledge from vibrations, the basics of Hydraulics is Fluid Mechanics and Road Design make uses of Dynamics. So, Physics is widely spread in Civil Engineering and civil engineers use it and have a proper knowledge of it, at least concerning their specialization.

In this context, what is the role of Applied Physics in Civil Engineering? All the Civil Engineering disciplines are well established and developed and researchers from Applied Physics can not to try to solve problems of Civil Engineering that can be more efficiently solved by a civil engineer. Instead, they must focus on physical problems and

applications that, due to their novelty, are not integrated into the Engineering disciplines. This kind of problems are, by definition, interdisciplinary and can be addressed by researchers from Applied Physics and Civil Engineering. As scholars from both fields usually know better one part of the problem or application, better results can be found when they work together.

Due to the wide range of Physics applications, an exhaustive list of them will require a length that is not available for this article. So, we are going to focus on two applications that are less known but enough interesting to learn about them.

First, Acoustic Emission applications will be shown. This is a materials and structures evaluation method that can monitor the internal processes that are taking inside them. Unlike other acoustic and vibrational non-destructive techniques, like Ultrasonic Inspection or Modal Analysis, it is a passive method that does not need a direct excitation of the material and can monitors processes in real time. It is at the same time is biggest drawback and advantage, because it cannot be repeated when the process has finished, but it allows to extract information about its dynamic characteristics. A proper understanding and utilization of Acoustic Emission requires good basics on mechanical wave motion theory, which is not very know between civil engineers apart from experts in Non-Destructive Evaluation, Seismic Engineering or Maritime Engineering.

Another application is how to use Heat Transmission basics in roads to reduce the urban heat island effect. This effect consists in higher temperatures in urban regions that in rural surroundings. This is produced by the lack of trees and vegetation, which absorbs and dissipate heat, by urban materials, which absorbs more solar radiation, by the urban geometry, which alters wind flow and create urban canyon, and by the presence of anthropogenic heat sources (Environmental Protection Agencys Office of Atmospheric Programs, n.d.). From Civil Engineering point of view, one of the measures that can be taken to mitigate this effect is the use of less absorbent materials in roads, the so-called cool pavements (Akbari *et al.*, 2015) and the design of urban roads and streets that reduces the urban canyon effect. The study of Heat Transmission is Civil Engineering degrees is not homogeneous. It is taught in countries where civil engineers can design the thermal isolation of buildings. But in others, like Spain, thermal isolation is responsibility of other engineers or architects.

3. ACOUSTIC EMISSION

Acoustic Emission (AE) refers to the generation of transient elastic waves due to the release of strain energy inside a material Ono (2014). This release uses to be very fast, so the elastic waves are on the ultrasonic range, between 30 kHz and 2 MHz. The processes that lead to the generation of elastic waves are very varied. They can be crack initiation and growth, plastic deformation, leak on pressurized fluids, corrosion, etc. The waves are produced when the process is happening and their emission is directly related with its dynamics, so the measurement of AE can be used to monitor and study the underlying process. By this way, AE technique is a passive non-destructive evaluation method that is used to monitor what is happening inside a solid, but, as opposite to other non-destructive techniques, it is difficult to repeat an experiment with the same sample because the generating processes are irreversible.

Unless the internal process that produces AE is spontaneous, it is necessary to apply an external stimulus over the test material to create it. It can be an external load, a rise on the internal pressure of a vessel, the flow of a fluid through a leak, a thermal excitation, etc. If possible, the external stimulus is measured together with the produced AE to relate both of them. The analysis of AE gives us information of the effect of the external stimulus on the induced internal process.

As a passive non-destructive evaluation technique, AE can monitor continuously a structure during long periods of time. It means that it can be used as Structural Health Monitoring method in critical structures. In this case the external stimulus uses to be the ambient conditions that can alter and degrade the structure.

An example AE waveform can be seen in Fig. 1. It corresponds to a single wave that arrives to the sensor from the source. This kind of signals are called burst-type and they are the most common type of signals in AE. They are also called hits, especially after they arrives to the sensor and are recorded by the AE equipment. They are characterised by a sudden rise followed by an exponential decay. They are produced by separate events like fractures or crack growth. The other kind of signals are continuous-type signals, which are longer and their beggining and end are not clear (Gallego and Martínez, 2016). This kind of signals are created by events which overlap between them or by a long event, like the leakage in a pressurized vessel.

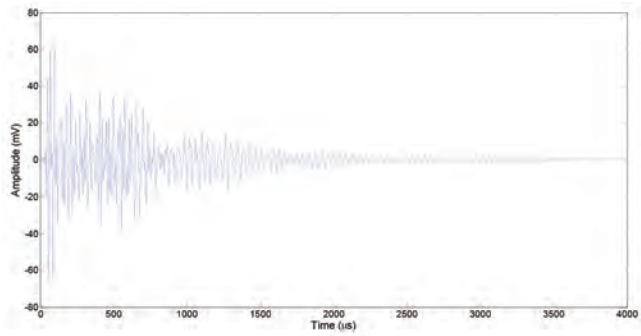


Figure 1. AE waveform from a Hsu-Nielsen source in a reinforce concrete specimen. It has been recorded with a VS30-SIC-46dB sensor with a resonant frequency of 30 kHz

The AE instrumentation consist of one or several AE sensors, one preamplifier per sensor, and one measurement systems with one channel per sensor. The amplitude of AE waves is very low, so the sensors need to be very sensitive and the electrical signal must be preamplifier early. It makes that the peamplifier must be connected near the sensor, and some sensors have attached preamplifier.

There are two types of AE sensors: broad band and resonant sensors. Broad band sensors are able to measure in a wider range of frequencies but their sensitive to amplitude is lower, while resonant sensors are more sensitive over a narrow range of frequencies around the resonant frequency of the sensor. The selection of the frequency range of the sensors depends on the application. For concrete and rock applications it is necessary to use low frequency sensors (30–100 kHz) due to the attenuation of high frequencies, while in low attenuating materials, like metallic ones, it is possible to use sensors in the frequency range of 100–400 kHz.

One of the principal advantages of AE technique is the possibility of detection the localization of AE sources. By using several sensors, knowing the wave velocity and assuming that is constant, the origin of an AE burst-type signal can be determined. This allow to know where is happening the internal process (for example a fracture) in real time and follow its evolution. This information is very valuable and can help to evaluate how it affects to the integrity of a structure.

The low amplitude of AE waves is its main drawback as non-destructive evaluation method. AE sensors must be very sensitive to measure them, they can typically detect pressures of around 1 μbar and displacements of 1 fm over 1 μs Ono (2014) (the atomic radius of hydrogen is 3.1×10^3 fm). It produces that many signals are recorded apart the interesting ones from AE sources. A researcher or operator must deal with this noise signals in order to keep only the valuable ones.

4. ACOUSTIC EMISSION IN CIVIL ENGINEERING

The utility of Acoustic Emission in Civil Engineering is the same than in other fields: To monitor internal processes in structures and materials. The principal motivation is to make a non-destructive evaluation of a structure, which is the same that in other engineering specialities. Other objectives can be to localise a flaw or damage, like in a railway, evaluate corrosion, monitor soil processes, etc.

The most distinguishing factor is the materials and structures under study. Concrete is of major importance as more critical structures are made of it, so fracture creation, crack propagation, strain energy and level of damage in concrete are in general the most studied issues with AE by civil engineers. The same can be done with other materials which are used in civil structures, like metallic or composite materials, although in this case the interest is shared with mechanical, aeronautical and naval engineers. Rocks and soils are also materials of almost exclusive interest of civil and mining engineers.

In the following sections, applications of AE to structural and material evaluation are summarised. This is not an exhaustive relation of all possible AE research lines and applications in Civil Engineering, because it is a broad field with many different aspects in which AE can be applied. But they are representative and varied examples of AE applications that can be of interest for a civil engineer.

4.1. Plastic Deformation

Elastic deformation does not produce AE, but as plastic deformation causes internal changes on the material, it produces AE. Most of the plastic materials follow the Kaiser effect, which is related with plastic deformation. The Kaiser effect can be described as the lack of AE when a load is applied until its previous stress levels are exceeded (Dobroň *et al.*, 2007). So, the first time that an increasing load is applied to a plastic material it produces AE. When the load is kept constant or it decreases, AE is residual. If a new load is applied, AE will not appear until the new load exceeds the previous one. The first application of the load produces internal changes that remain after the load is removed. Another application of the same load does not produce any internal change in the material until a higher pressure is reached. The Kaiser effect can be uses in several ways. First, if a structure is inspected periodically by applying a load on it, if no damage exists it will not produce AE because the pressure levels will not be higher than the previous one. But if damage occurs between proof tests, in the next inspection the stress will rise to higher levels than before in discontinuities due to damage and AE appears. This is known as Dunegan Corollary Drummond *et al.* (2007), that states that AE during testing reveals damage during the preceding operational period. The second application of Kaiser effect is when it is violated, so when AE is detected for applied loads lower than the maximum previous one. It indicates that the internal damage is very high, and the structure can fail. It is known as Felicity Effect, and its quantification by the Felicity Ratio, the ratio of the load of the first AE appearance in the new loading cycle to the maximum load of previous cycles, is used as damage parameter (He *et al.*, 2021).

Most of construction materials follows these effects, so they can be evaluated with it. They are used in damage diagnosis and monitoring in metal (Drummond *et al.*, 2007) and concrete structures (Li *et al.*, 2021), composite reinforcements (Ono, 2014), masonry structures and asphalt pavements (Behnia *et al.*, 2018), rocks (Zhang *et al.*, 2021) and soil (Mao *et al.*, 2019).

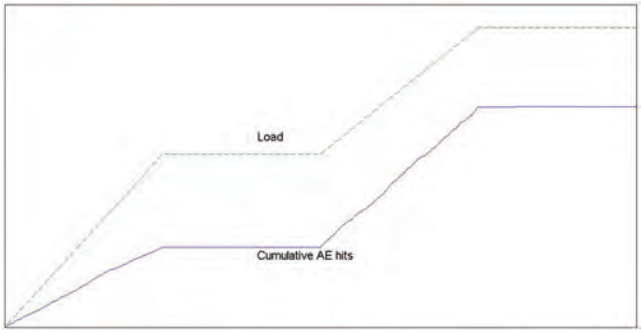


Figure 2. Representation of the Kaiser effect. Applied load and Cumulative AE hits versus time. Cumulative AE hits rise only when previous levels of applied load are exceeded

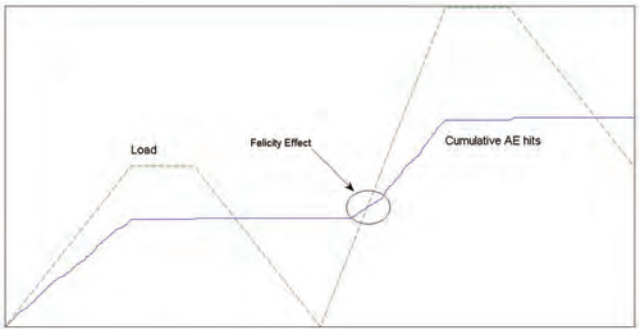


Figure 3. Representation of the Felicity effect. Applied load and Cumulative AE hits versus time. In this case there is an increase of Cumulative AE hits before the previous levels of applied load are reached. The Felicity effect is a violation of Kaiser effect and it can indicated the presence of damage.

4.2. Cracking

Detection and evaluation of crack formation and grow is one of the main advantages of AE. With sensors located at different points and the prior knowledge of ultrasound velocity in the material, it is possible to locate the AE sources and check if a crack is expanding. Linear localization requires two sensors, or more if the material is very attenuating. Planar localization can be done with three sensors and 3-D localization requires at least four sensors. Localization is done usually with triangulation algorithms based on arrival times difference between different sensors. Other approach is to use zonal localization, which is useful in very attenuation or heterogeneous mediums. The sensors are distributed between the structure, and it is determined which is the nearest zone to each one. When a sensor receives the first AE hit (or the one with the most amplitude), the source location is assigned to this sensor's zone. If other sensors are able to measure the same AE event, then the localization can be refined by determining the part of the zone which is nearest to the sensor that receives the second hit. With the complete hit sequence, it is possible to delimit with some accuracy the source location.

In concrete it is possible to determinate the type of AE source by means of moment tensor analysis (also known as SiGMA analysis) (Ohno and Ohtsu, 2010). It is based on arrival time and first motion amplitude of each AE hit and can determinate if it has been produced by a tensile, shear or mixed-mode crack. If SiGMA analysis cannot be carried out due to insufficient number of sensors, then some information can be deduced from parameter analysis. Some of typical AE signal parameters are amplitude, duration, rise time (time from the arrival to the maximum amplitude peak), ring-down counts (number of times that the signal exceeds the threshold) and signal strength. With some of these parameters, the following values are defined:

$$\text{RA value} = \frac{\text{Rise time}}{\text{Peak amplitude}}$$

$$\text{Average Frequency} = \frac{\text{Ring-down counts}}{\text{Duration}}$$

Using these parameters, it is possible to classify the cracks into tensile cracks and other-type cracks (Ohtsu, 2018). The classification is made looking at the combination of both parameters. If Average Frequency is predominant, then the crack is tensile type, while if RA value is the predominant one, it is other-type crack. The physical meaning of this classification comes from the fact that tensile cracks cause opposite movement of the crack sides, which results in AE waveforms with short rise time and high frequency (Behnia *et al.*, 2014). In shear type cracks, the movement of one crack side over the other produces friction between the sides after the initial crack. It produces a longer release of energy that it reflected in waveform of longer duration which results in longer rise time and lower frequency.

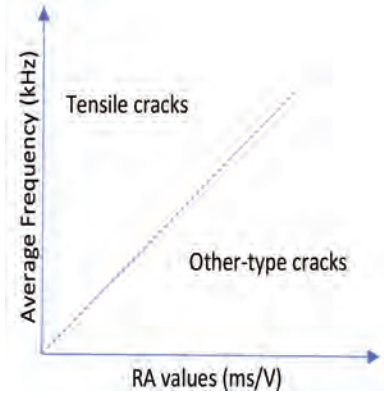


Figure 4. Typical crack type classification based on parameter analysis

4.3. Methodologies for evaluation of damage state

A lot of research has been conducted in order to evaluate the damage state of a structure with AE, which has led to different evaluation methodologies and damage parameters based of AE features. Cumulative AE hits plots are one of the primary ways to study damage evolution.

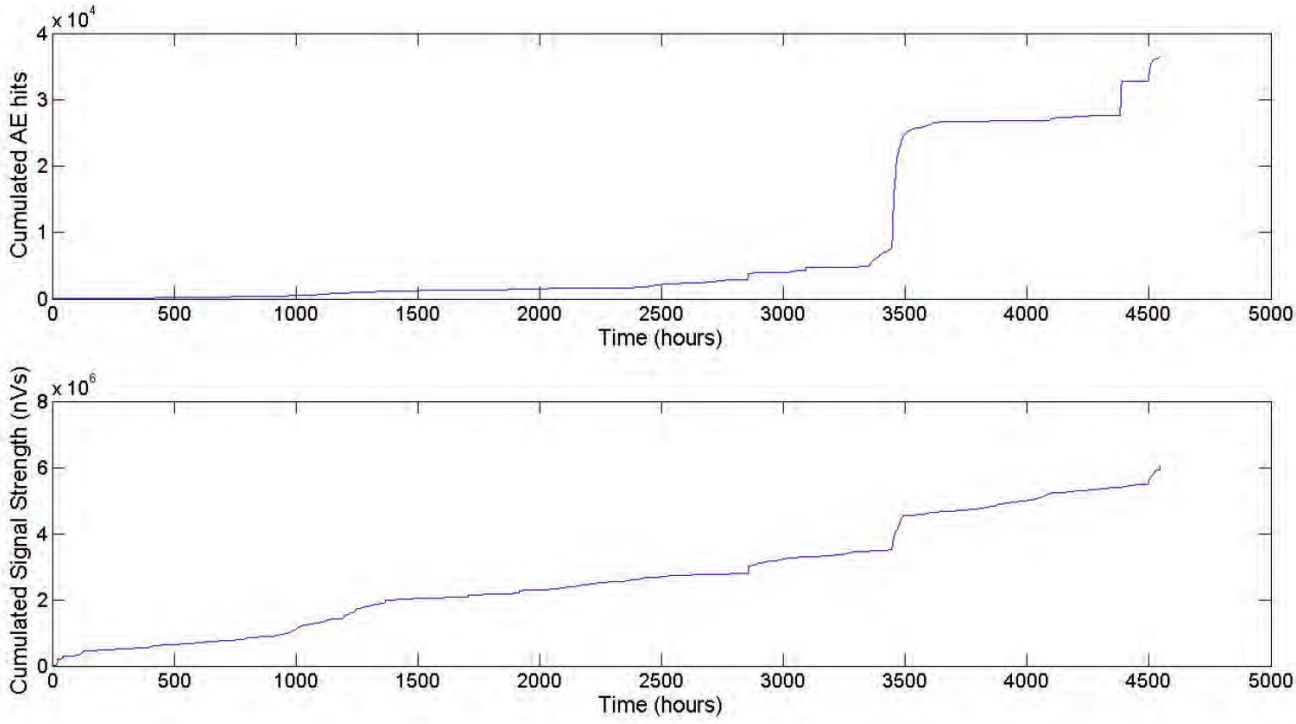


Figure 5. Comparation of cumulated AE hits (up) and cumulated Signal Strength (down) from the same sensor in a corrosion test in a reinforcement concrete sample

Changes in slope reflect different emission rates, which are connected with cracking processes and damage state, because zones with defects can produce AE in response to an increase in applied stress. Similar analysis can be done with cumulative AE energy plots and cumulative signal strength (CSS). AE energy, which can be identified with the part of elastic energy released during an event that is measured by the sensor, is regarded as the area under an AE signal. It is defined as:

$$E = \int_{t_1}^{t_2} V^2 \, dt$$

Where V is the voltage measured between the initial (t_1) and end (t_2) time of the hit. Signal strength is the area under the rectified AE signal (di Benedetti *et al.*, 2013). Cumulative AE energy and CSS plots are usually more interesting that cumulative AE hits plots because low amplitude and short hits, which usually are noise or reflections, affects very little on them.

As an example of evaluation by cumulative AE energy plots, it has been found that cumulated AE energy is proportional to hysteretic strain energy in reinforce concrete specimens under cyclic loading (Benavent-Climent *et al.*, 2009), so it could be used to monitor the damage state of a reinforce concrete structure under dynamical loads. Moreover, parameters based on cumulative AE energy have been proposed for predict the level of damage (Benavent *et al.*, 2010; Benavent-Climent *et al.*, 2009).

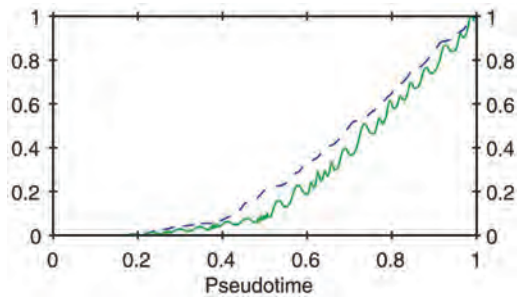


Figure 6. Normalised cumulated AE energy (dashed line) and normalised hysteretic strain energy (dotted line) from a RC beam-column connection

4.3.1. Evaluation methodologies based on Felicity effect

Evaluation methodologies that make use of Felicity effect are based on the use of damage parameters that quantify this effect. These methodologies require the application of load cycles in order that Felicity effect can emerge.

Felicity ratio (FR), that has been mentioned before, is defined as the ratio between the load that produces significant AE activity in the current load cycle and the maximum load in the previous loading history. It is also known as load ratio (Ohtsu *et al.*, 2010):

$$\mathrm{FR} = \frac{\text{Load of the first AE appearance in the current loading cycle}}{\text{Maximum load in loading history}}$$

In order to take into account only significant AE, a criterium has been stablished to know exactly which is the load of the first AE appearance. This criterium specified that the first significant AE appearance in the current cycle takes place when the historic index during the current loading cycle exceeds 1.4 (Behnia *et al.*, 2014). Historic index H(t) is defined by the following equation:

$$H(t) = \frac{N\sum_{i=k+1}^{N} S_i}{(N-k)\sum_{i=1}^{N} S_i} \tag{1}$$

Where N is the number of hits up to time t, S_i is the signal strength of hit i, k is an empirical factor who is a function of the number of hits and the type of material. For RC structures, it is given by the following function (ASTM, E, 2006):

$$k = \begin{array}{cc} \text{Not applicable} & N < 50 \\ N-30 & 51 < N < 200 \\ 0.85N & 201 < N < 500 \\ N-75 & 501 < N \end{array}$$

Other functions for k with a higher number of intervals for k values have been used by other authors (Behnia *et al.*, 2014).

A value less than 1 for the Felicity ratio is always an indication of degradation of the structure, which increase as the difference of FR to 1.

Calm ratio (CR) is the ratio of cumulative AE activity (typically AE hits or signal strength) during the unloading process to that of whole load cycle (Ohtsu *et al.*, 2010).

$$\text{CR} = \frac{\text{Cumulative AE activity during unloading}}{\text{Total AE activity during whole cycle}}$$

Experimental values of CR for different damaged concrete beams varies from 0.05 to 0.6 (Ohtsu *et al.*, 2010).

The combination of these two ratios can be used to classify damage in reinforced concrete beams in a qualitative way. Low FR values and high CR values indicates heavy damage, while low CR values and high FR values means minor damage. This classification is schematically shown in fig. 7 (Ohtsu *et al.*, 2010).

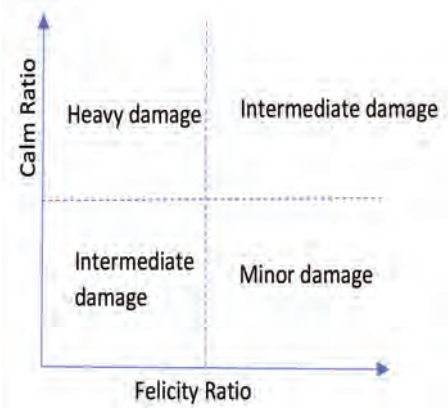


Figure 7. Qualitative Damage criteria by the combined values of Felicity and Calm ratios

4.3.2. Intensity analysis

Intensity analysis is a methodology to evaluate the structural severity of a defect (Fowler *et al.*, 1989). It compares the changes in the signal strength of AE during the test with the AE events that have more signal strength. The changes in signal strength are determined by the historic index $H(t)$ (1), which is an analytical method for determining the changes of slope of CSS. $H(t)$ compares the average signal strength of the last events to the average signal strength of the whole sequence of events. $H(t)$ increases sharply at the knee of the curve, and after the knee it declines until onset of failure, where it reaches a maximum. The higher signal strength values are taken into account to calculate the second parameter used in intensity analysis, which is known as Severity (S). It is the average signal strength of the J events having the largest value of signal strength. Its definition is shown in the following equation:

$$S = \frac{1}{J}\sum_{i=1}^{J} S_i \tag{2}$$

Where J is a constant that is derived empirically and that depends of the type of material. Its value is 50 for concrete (Nair and Cai, 2010) and 10 for metal (Fowler *et al.*, 1989). An increase in Severity means the onset of structural damage, being this increase more significant with the intensity of damage.

Intensity analysis is carried out by calculating the Historic Index and the Severity. The maximum value of Historic Index and the Severity are represented on a chart divided into six zones, as can be seen in fig. 8. It will give us information of damage intensity and if it is necessary to take immediate actions over the structure. The description of each zone is shown in table 1.

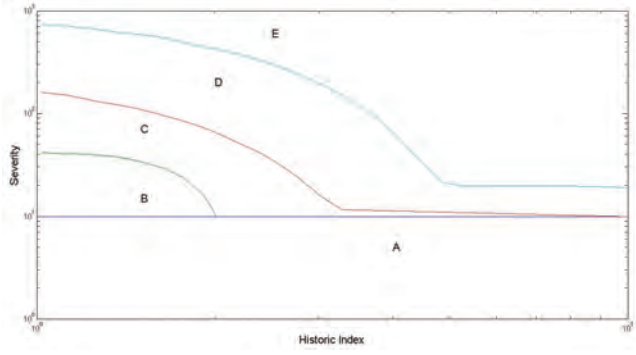


Figure 8. Intensity zones in concrete (Golaski *et al.*, 2002)

Table 1. Description of each intensity zone (Fowler *et al.*, 1989)

Intensity Zone	Description
A	Minor defect, not relevant
B	Small defect, noted for future evaluations
C	Defect, further evaluation required
D	Significant defect, further evaluation required
E	Major defect. The structure is in danger, it is necessary its closure and evaluate it with more non-destructive methods.

4.3.3. Seismic based methodologies

Acoustic emission and earthquakes are similar physical processes, because in both cases elastic waves are produced after a fast release of strain energy. Obviously, there is a big difference in scale because the sources of seismic waves are geological processes that produces waves of very low frequency, large wavelength and high amplitude, while AE sources are microscopic or small and produce high frequency waves with low amplitude. But the physical similitudes are enough to allow the adoption of evaluation methodologies that were first used in Seismology.

***b*-value.** The b-value is a crack identification method which is based on the statistical distribution of hit amplitudes. Usually, the number N of AE events with peak amplitude above a given level A shows an exponential decay with the increase of this level. So, a plot of the logarithm of N versus the logarithm of A shows a linear relationship between both of them (Manthei and Eisenblätter, 2008):

$$\log N = a - b \log A \tag{3}$$

This is the Gutenberg-Richter relation (Gutenberg and Richter, 1954) for the frequency of earthquakes, with magnitude of earthquakes instead of $\log A$. The slope b is the fundamental parameter of this analysis. It characterizes the frequency of low amplitude events in comparison to high amplitude ones. If b is small, there is a large portion of high amplitude events, while if it is large, the portion of high amplitude events is small.

It can be used to identify the size of cracks, because microcracks will release low energy hits, leading to high values of b, while macroscopic cracks produce higher energy hits, that cause lower values of b.

Improved *b*-value. Improved b-value (Ib-value) is an adaptation of b-value to special conditions of AE. In AE the number of events changes with time, and it rises when approaching final failure. If different numbers of AE events are use in b-value calculation, there will be inconsistencies between the different values. To avoid it, in Ib-value calculation the same number of events is considered. A range of 50 to 100 events has been considered an appropriate number (Shiotani, 2008).

Peak amplitudes vary from each monitoring due to distance of source to sensor, type of sensor, attenuation of the medium, so another modification is to stablish a method to choose the amplitude range for which the slope is calculated. If μ is the mean of the amplitude distribution and α its standard deviation, then the amplitude range for which the slope is calculated is between $\omega_1 = \mu - \alpha_1\sigma$ and $\omega_2 = \mu - \alpha_2\sigma$. α_1 and α_2 are two user defined constants. The value of these constants must be chosen so that the provided range $[\omega_1, \omega_2]$ presents a good linearly in the logarithm of the frequency-amplitude distribution during the test.

Once the amplitude range is determined, Ib-value is calculated by (Benavent *et al.*, 2010):

$$Ib = \frac{\log N(\omega_1) - \log N(\omega_2)}{(\alpha_1 + \alpha_2)} \tag{4}$$

With these modifications, Ib-value is not affected by monitoring conditions and it uniquely depends on the damage level.

It has been established that Ib-values lower than 0.05 implies the onset of macroscopic cracks, while when the values are higher than 0.05, the cracks are still microscopic. Usually, amplitudes are measured in decibels, so Ib-value must be multiplied by 20 to compare it with b-value. By this normalization, the boundary between microscopic and macroscopic cracks becomes 1.

It has been seen that an increase in the number of critical events (events with $Ib < 0.05$) that produces macroscopic fractures can be used as an indicator of the onset of yielding of steel bars in reinforced concrete structures under load cycles (Benavent *et al.*, 2010), which means that the size of cracking processes increases just before the yielding of steel reinforcing bars.

Minimum *b*-value. Minimum b-value analysis (Schumacher *et al.*, 2011) is a method to use b-value methodology to the inspection and evaluation of bridges. It combines the crack evaluation by b-value analysis with the information of passing vehicles, which are the applied load to the bridge. The pass of vehicles can be due to normal traffic conditions (ambient conditions) or chosen by researchers to have more controlled conditions (heavy trucks).

To determine the minimum b-value, each time a vehicle pass through the bridge, which can be considered a load cycle, the b-value is calculated and its minimum value is found for each sensor. Then, the mean and standard deviation of the minimum b-values of all sensors and load cycles of the same magnitude is obtained. This gives an averaged representation of the ongoing degradation of

in-service bridges, especially in a section of interest due to prior damage.

Energy *b*-value. Another modification of b-value is the energy b-value, which is more precise in the evaluation of damage evolution in concrete. It is based on the fact that acoustic emission from cracking can be mixed with other energy emissions, resulting in a noisy signal which peak amplitude does not correspond to the one emitted from cracking. Sagasta *et al.* (2018) proposed as solution the filtering and extraction of the elastic energy from cracks using signal processing techniques (wavelet transform), which they called true energy.

Instead of using the amplitude distribution in the calculation of b, it uses the true energy distribution. So, the Gutenberg-Richter relation becomes:

$$\log N(AEE) = a - b_e \log AEE \tag{5}$$

Where AEE is the true energy, $N(AEE)$ is total number of reconstructed signals with true energy higher than AEE and b_e is the energy b-value.

4.3.4. Fractal analysis of AE to evaluate energy dissipation

Cracking phenomena present multiscale aspects, which makes that energy dissipation during microcracking propagation occurs in a fractal domain (Carpinteri and Pugno, 2005). This domain has been found to be comprised between a surface and the specimen volume V. So, the total dissipated energy E_{ax} during cracking scales as:

$$E_{\max} = \Gamma V^{D/3} \tag{6}$$

Where D is the fractal exponent, whose value is between 2 and 3, and Γ is the fractal energy density, which can be considered a size-independent parameter.

This dissipated energy is released as elastic waves. So, *AE* energy measured during microcrack propagation is proportional to the dissipated energy. Following equation (6), the volume dependence of the total *AE* energy $E^{AE}_{\max}$ is (Benavent-Climent *et al.*, 2009):

$$E^{AE}_{\max} = \Gamma_{AE} V^{D/3} \tag{7}$$

Where Γ_{AE} is the fractal density of *AE*.

Its analysis can be used to obtain the damage level of a structure from AE data of a reference specimen (Carpinteri and Pugno, 2005). If the reference specimen is extracted from the structure, we have that

$$\Gamma_{AE} = \frac{E^{AE}_{\max}}{V^{D/3}} = \frac{E^{'AE}_{\max}}{V' D/3} \tag{8}$$

Where $E^{'AE}_{\max}$ and V' are the total *AE* energy and the volume of the reference specimen. So, we can substitute the fractal density in equation (6):

$$E^{AE}_{\max} = E^{'AE}_{\max} \left(\frac{V'}{V'} \right)^{D/3} \tag{9}$$

And we obtain the total *AE* energy that can is emitted before the collapse of the structure. It can be used to define a non-dimensional parameter η to describe the damage level of the structure (Carpinteri and Pugno, 2005):

$$\eta = \frac{E^{AE}(t)}{E^{AE}_{max}}$$

Where $E^{AE}(t)$ is the cumulated *AE* energy at time t.

The fractal exponent depends on the damage distribution. If the damage is distributed over in the whole specimen volume, we have that $D = 3$, while if it organized on a preferential fracture surface then D is near 2. It has been found that the fractal exponent is related with the b-value (Verstrynge *et al.*, 2021), with $D = 2b$ in most cases.

4.3.5. Multiparametric analysis

One of objectives of *AE* researchers it to find *AE* signatures that allows to distinguish between different damage *AE* sources, so damage identification will be carried clearly. There are many *AE* features that can be analysed and potentially used to look for the *AE* signature, so many times this search of an AE signature is carried out in an automatic way. Multiparametric analysis using different methodologies and approaches are usually done to study the huge amount of data from *AE* and correlate them to damage mechanism (Behnia *et al.*, 2014).

Pattern recognition is one of the first and principal approaches to multivariate analysis in *AE* evaluation. Cluster analysis of different *AE* parameters, for example using K-means algorithm, is used to group *AE* hits that have similar characteristics and check if each group is related with a different *AE* source.

Principal Component Analysis is a signal processing technique to reduce the dimensionality of data to allow the pattern recognition analysis and the interpretation of the data. It is used in combination with other techniques because it reduces the computational effort and to simplify the application of the other signal processing techniques.

Machine learning algorithms has been also extensively used in *AE* analysis, especially neural networks. They simplify the human effort in correlating *AE* features and damage sources, but, as other automatic techniques, requires a precisive labelling of *AE* events and the knowledge of the existing sources in order to be really effective and avoid errors.

4.4. Civil Engineering Applications

4.4.1. Concrete Structures

AE is used as non-destructive evaluation method in infrastructures in two different ways. First, critical struc-

tures are monitored in real time, acting *AE* as a structural health monitoring method. Second, periodical inspections are carried out to evaluate the damage state. The first one gives much information and if the structure is instrumented during its building and it is measuring from the beginning of its service, the health state will be known in every moment. The second one is less expensive and more appropriate for old infrastructures where it is not possible to implement a continuous monitoring.

Global health evaluation is carried out by analysing Kaiser and Felicity effects. This analysis requires the application of an external load, which is usually the most difficult issue of the experimental part. In the case of bridge inspection, this condition is met by passing heavy trucks over the bridge (Nair and Cai, 2010). For dams (Zhao *et al.*, 2021), changes in water level can be exploited as load cycles to study both effects. Intensity analysis is also very used for structural evaluation of infrastructures.

Crack monitoring is other of the main applications of *AE* in concrete structures. Its study with *AE* is very interesting due to the possibility of localization of new fractures and growing ones via difference of arrival times. Also, the determination of the crack type using moment tensor analysis or parameter analysis can provide valuable information.

Another specific application, related with crack monitoring, is assessment of corrosion in reinforce concrete structures. The increase of size of reinforcing steel due to corrosion produces cracking processes around it that can be detected by *AE*. A phenomenological model of corrosion loss in steel has been proposed (Sharma *et al.*, 2018). It has four phases of different rates of corrosion loss that depends on the time from the beginning of the corrosion process (fig 9). It has been found that *AE* cumulative parameters in reinforce specimens under a corrosion process follow a similar trend to this model, especially cumulative signal strength (fig 10). It means that AE can be effective in monitoring corrosion loss in reinforce concrete structures and cand detect it in early stages.

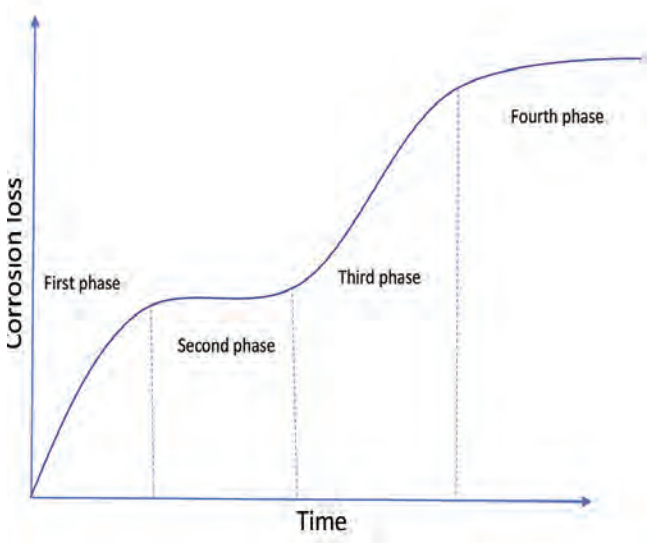


Figure 9. Phenomenological model of corrosion loss in steel

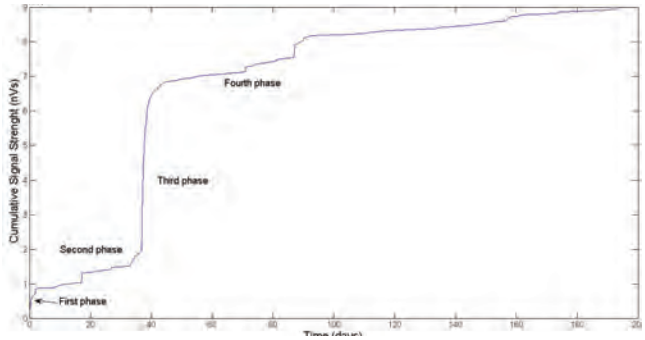


Figure 10. Cumulative Signal Strength of a reinforce concrete specimen under a accelerated corrosion

4.4.2. Cracks detection in rail tracks.

AE is possible the only non-destructive method that can be used to monitor online fracture creation and crack growth of in-service rail tracks. Fatigue cracks can be detected by the increase of *AE* hits and count rate under traffic loading. Also, it has been reported that variations in other *AE* parameters, like energy, rise time and average frequency are also indicators of the presence of cracks (Kostryzhev *et al.*, 2013).

The principal problem of the application of *AE* in rail tracks is the high level of noise. The movement of trains over the rails produces continuously acoustic emission that masks acoustic emission from cracks. So, many efforts have to be made to eliminate this noise and be able to analyse *AE* from cracks. Denoising of *AE* implies the use of advanced signal processing techniques, as Discrete Wavelet Transform (Hao *et al.*, 2018; Zhang *et al.*, 2015) or autoregressive models (Zhang *et al.*, 2018), to obtain filtered signals. Then, these signals can be studied with parametric *AE* analysis. Other approach is to filter complete *AE* hits by the relationship between their amplitude, frequency range and duration, distinguishing between low amplitude, frequency and duration hits, which are due to noise, from different combinations of these parameters that correspond to several *AE* sources: deformation, ductile crack grow and brittle crack grow (Kostryzhev *et al.*, 2013).

4.4.3. Asphalt pavements

The initiation and accumulation of microcracks, which can eventually lead to the formation of macrocracks, produces a gradually deterioration of asphalt pavements. These microcracks are caused principally by fatigue damage due to traffic loads and from non-uniform thermal stress due to fast changes of ambient temperatures.

Fatigue damage of asphalt mixtures can be described by Kaiser and Felicity effect. While this is difficult to use directly in in-service pavements because an external load must be applied, it has been used to characterised *AE* from fatigue damage and uses other parameters as damage indicators. With this approach, it has been found (Qiu *et al.*, 2020) that sources localization can map the cracks distribution and growing, and b-value analysis can give us information of crack size in every moment. Waveform analysis based on time-frequency distributions (normally using the wavelet transform) is applied to look for features that can differentiate between *AE* sources. Larger concentrations of energy in high frequency bands are related with cracking and crack propagation, while concentration in low frequency bands can be caused by the closing of cracks and friction between the faces of the crack (Qiu *et al.*, 2021).

Thermal cracks in asphalt pavements can be monitored using the concept of Embrittlement Temperature (T_{EMB}), which is defined as the temperature corresponding to the high-energy *AE* event above an energy threshold (Behnia *et al.*, 2018). With this parameter it is possible to know at which temperature cracking processes start in asphalt and the onset of thermal damage, allowing the estimation of low-temperature performance grade of virgin, short-term and long-term asphalt binders.

Visible thermal cracks are transverse to pavement, but there are spiral cracks inside the asphalt that can be only seen after the removal of top material. These cracks are the result of the stress field that appears with the constraints imposed by the granite block (Behnia *et al.*, 2017). Spiral cracks have a spiral thickness parameter, which controls how tightly and in which direction spiral is wrapped. It has been observed that spiral the thickness parameter is related to T_{EMB}.

The Kaiser effect has been found in thermal cycling in concrete asphalt, probably due to the periodical stress that appear with cycles of temperatures which increases when the differences of temperatures are higher. Felicity effect is also found here. But instead of indicates a severe damage in asphalt, it has been related with some kind of microcrack self-healing due to the adhesive nature of asphalt materials (Behnia *et al.*, 2018). The Felicity Ration (FR) is defined here as the ratio between the temperature when Felicity effects appears and the minimum temperature in the previous cooling cycle. It is used to quantify the self-healing behaviour of asphalt materials by the Healing Index:

$$\text{Healing Index } (\%) = 100(1 - FR)$$

which ranges from 0% (no healing) to 100% (fully healing).

5. HEAT TRANSMISSION APPLICATION IN CIVIL ENGINEERING: COOL PAVEMENT

5.1. Introduction

The development of urban areas, where buildings, roads and other infrastructures replace land and vegetation, leads to the formation of urban heat islands (UHI), a phenomenon whereby urban regions experience warmer temperatures than their rural surroundings.

UHI has a severe impact on the quality of life and the economy of urban citizens. It contributes to general discomfort, respiratory difficulties, heat cramps, exhaustion and heat-related mortality. It also increases water and energy consumption.

Many factors contribute to UHI formation, but one of the most important are the properties of urban materials, in particular solar reflectance, thermal emissivity and heat capacity. Urban surfaces of asphalt and other materials reflect less solar radiation and absorb more of the sun energy than those in rural settings. This higher absorption of heat increases surfaces temperature, which heats surrounding urban air, creating a warm bubble over the city.

So, one of the solutions to mitigate the UHI effect is to use building materials with higher solar reflectance and lower thermal emittance, especially in roof and pavements. They are the surface elements more exposed to solar radiation and constitutes a high percentage of urban area (the combined fraction of roof and pavement area of a dense city may approach to 90% Akbari *et al.* 2015), so the use of less radiance material on these elements can have an important impact in the mitigation of UHI phenomenon. This rise to the concept of cool pavement, which absorbs less solar heat because it is built with highly reflective materials.

The economic costs of replacing existing pavements in a city demands an evaluation of the effect on street temperatures of the use of cool pavement. Depending on street configuration and geometry, the change to a cool pavement can significantly decrease the ambient temperature or only a little. It must be evaluated previously to save unnecessary costs.

One example of implementations of cool pavements is the project “Life Heatland” (LIFE HEATLAND, 2021), cofunded by EU under LIFE programme and whose partners are Business Association of Investigation Technological Center of the Construction of the Region of Murcia (CTCON), CHM (a spanish construction company), The Regional Federation of Construction Entrepreneurs of Murcia (FRECOM), Slovenian Construction Industry Cluster (SGGCCS), and Murcia City Council. Its objective is the demonstration in a medium scale of the effectiveness of new cool pavement to mitigate the UHI effect. Six streets of the city of Murcia have been paved with the new reflective pavement and a another one with a normal one to

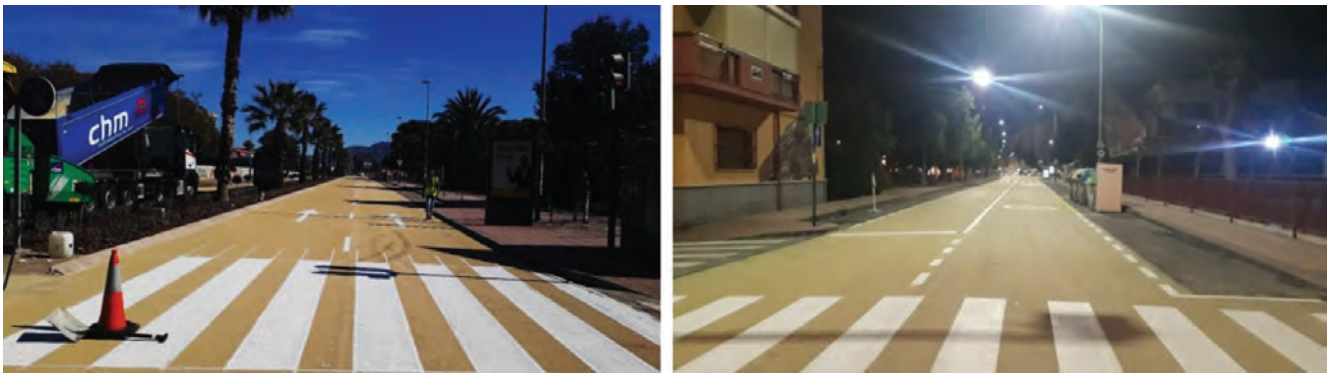


Figure 11. Cool pavement just after its installation in two different streets and in different light conditions: day (left) and night (right)

be used as a reference, being a total of 24000 m^2 of new pavement. Four metering towers have been installed in these streets to measure the air and pavement temperatures, humidity and concentration of different gases.

Experimental measures of temperatures can be used to check cool pavement performance in similar streets. But in order to evaluate the impact of paving new streets with different configurations or climatic conditions, a mathematical model of the combined heat processes must be elaborated and implemented in a simulation software.

5.2. Cool Pavement

The reflective properties of cool pavements are usually obtained by adding metallic oxide pigments. Another approach, combined with the use of reflective materials, is the implementation of porous pavements that make them permeable. This allows stormwater to soak into the pavement and soil, and the following evaporation of water cools them. It also has a non-expected consequence that improves the comfort and quality of life in cities, that it is the reduction of traffic noise due to the more porous surface (Environmental Protection Agencys Office of Atmospheric Programs, n.d.). Also, the higher reflectance increases urban visibility at night, which increases citizens security and can be used to reduce street lightning.

Solar reflectance of the cool pavement studied under LIFE HEATLAND project is 30% (Newsletter, 2020), which is almost four times that conventional pavements. It leads to a reduction of mean temperature of pavement between 7°C and 11°C (fig. 12) and a decrease of ambient temperature of around 2°C. It also has reduced traffic noise in 3 dB(A) and has increased the luminance of streets at night a 150% that with conventional asphalts.

Concerning the cost of the studied cool pavement, the material cost is 250% higher than conventional one, basically because the cost of the new additives. On other hand, it needs less thickness, so the cost of material transport and installation is estimated to be around 8% inferior than conventional asphalt. Also, the new pavement has 5 years more of lifetime than conventional ones.

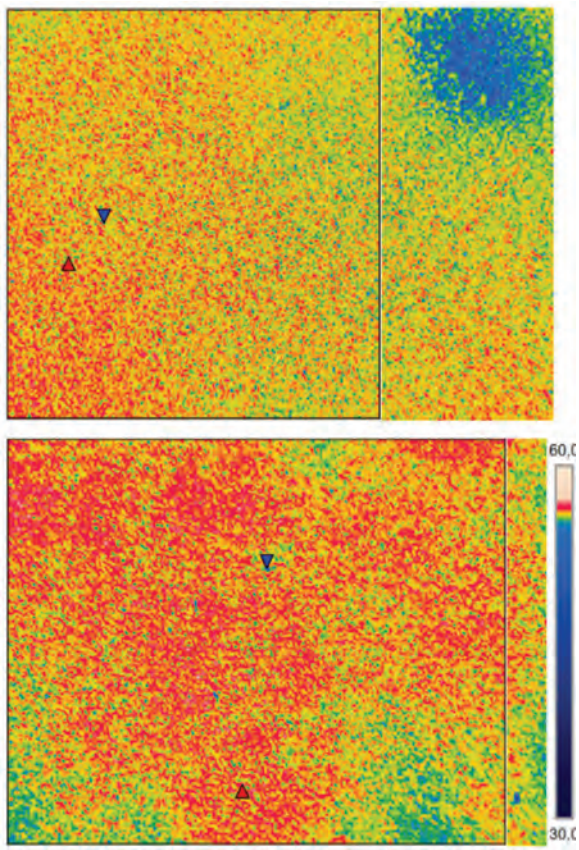


Figure 12. Thermography of cool pavement (left) and conventional one (right) under a solar radiation of 800 W/m^2 and a air temperature of 32°C. Their mean temperatures are 49.5°C for the cool pavement and 56.3°C for the conventional one.

5.3. Simulation software of thermal influence of pavement

To study the thermal influence of reflective pavement is streets located in cities with other climatic conditions, a simulation software has been developed. This software implements the coupled differential equations that describe the different processes involved in temperature distribution in the pavement and air.

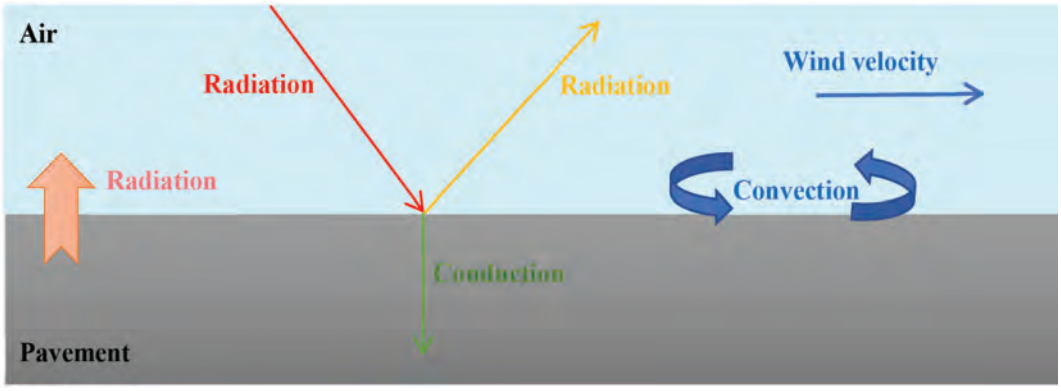


Figure 13. Combined heat transmission processes in air-pavement system

There are several combined heat processes that affect pavement temperature. Incident radiation from sun, reflected radiation from buildings, radiation from pavement, heat conduction in pavement, natural and forced convection between pavement and air. All these processes transfer energy to or from the pavement, so the variation of the internal energy of the pavement can be expressed as:

$$\Delta U = \rho S \Delta x c_p \Delta T = \sum_{i=1}^{N} Q_i \tag{10}$$

where ρ is the density of the pavement, S is its area, Δx is the pavement thickness, c_p the specific heat capacity, ΔT the variation of temperature and $Q_i$the heat from process i. This equation is derived and divided by S to obtain the temporal variation of pavement temperature:

$$\rho \Delta x c_p \frac{dT}{dt} = \sum_{i=1}^{N} \frac{Q_i}{dt} \tag{11}$$

5.3.1. Incoming solar radiation

There are three radiation processes: direct radiation from sun, reflected radiation from nearby buildings to pavement and reflected radiation from the pavement. So, the net radiation over the pavement is:

$$\text{Rad} = (1-\delta) \cdot (\text{Radsolar+Radfled}) \tag{12}$$

where δ is the reflected radiation rate of the pavement, Radsolar is the solar radiation over the pavement and Radfled is the reflected radiation from building or other sources. So, the total heat that is absorbed by radiation by the pavement is:

$$\frac{dQ}{dt} = \text{Rad} \cdot S \tag{13}$$

where Q is absorbed heat and S is the area of the pavement.

5.3.2. Outgoing radiation

The radiated heat from pavement is given by the Stefan-Boltzmann law. Taking into account the radiated heat that arrives from the ambient to the pavement, the net radiated heat of the pavement is:

$$\frac{dQ}{dt} = S\varepsilon\sigma(T^4 - T_r^4) \tag{14}$$

where σ is the Stefan-Boltzmann constant, ε is the asphalt emissivity, T the pavement temperature and T_r the reference temperature of ambient (del Cerro Velázquez *et al.*, 2008).

5.3.3. Convection

The last heat transmission process is convection between pavement and air, which is given by:

$$\frac{dQ}{dt} = Sh(T - T_a) \tag{15}$$

where h is the convection coefficient between asphalt and air, whose value depends on air conditions (humidity, velocity, temperature...).

5.3.4. Balance of energy in pavement-air boundary

All the processes shown before are related to energy transmission between the pavement and the environment. All of them are combined into eq. (11) to give:

$$\frac{dT}{dt} = \frac{\text{Rad}}{\rho \Delta x c_p} - \frac{\varepsilon\sigma}{\rho \Delta x c_p}(T^4 - T_r^4) - \frac{h}{\rho \Delta x c_p}(T - T_a) \tag{16}$$

5.3.5. Heat conduction in pavement

The energy transport through the pavement is controlled by the conduction heat equation:

$$\frac{dT}{dt} = \frac{k}{\rho c_p} \nabla^2 T \tag{17}$$

where k is the asphalt's conductivity.

5.3.6. Thermal gradient in air

Temperature gradient in air is used to calculate the temperature distribution. The determination of this gradient is a complex problem with many physical processes connected between them: wind velocity, which changes with altitude, the temperature, which changes the air density, atmospheric pressure, humidity, atmospheric stability and solar radiation. The equations that can be used to determine the thermal gradient in air can be found in (Businger *et al.*, 1971).

5.3.7. Implementation and results

To solve numerically the system of coupled differential equations the Network Simulation Method has been used (Sánchez-Pérez *et al.*, 2018; 2021), which is very suitable to solve coupled non-linear problem. It is based on the designing of an electric network which is mathematically equivalent to the physical problem. Then, the open software NgSpice (NgSpice Software, 2020) is used to obtain the numerical solution for each time.

A specific software for an easy determination of the temperature distribution has been programmed. An example of input parameter screen can be seen in fig. 14.

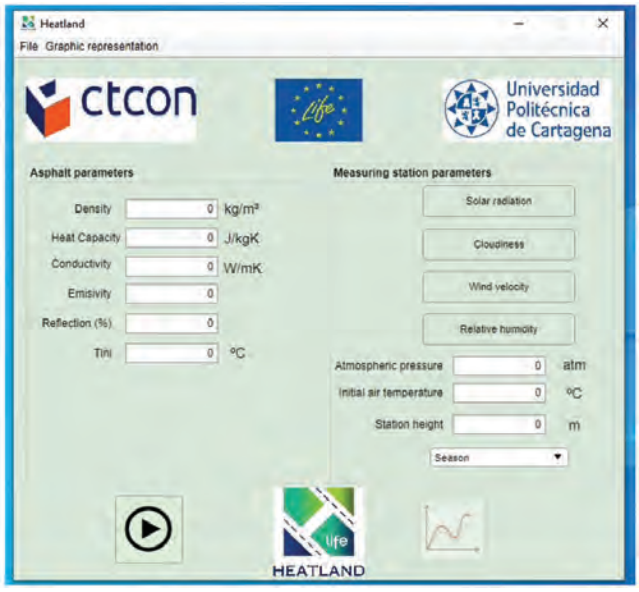


Figure 14. Input physical parameters screen

The results from the software have been validated with experimental data. In fig. 15 it can be seen the comparison between software results and experimental data for the asphalt and air temperature with time in two different winter days. The simulation results and experimental data have similar distribution, with a difference up to 1°C most of the time, and they are identical in many cases. The main discrepancies can be attributed to the occurrence of events that affect temperature measures but are not taking into account by the physical model. For instance, the shadow of an individual cloud or the presence of a heat source near the measurement point.

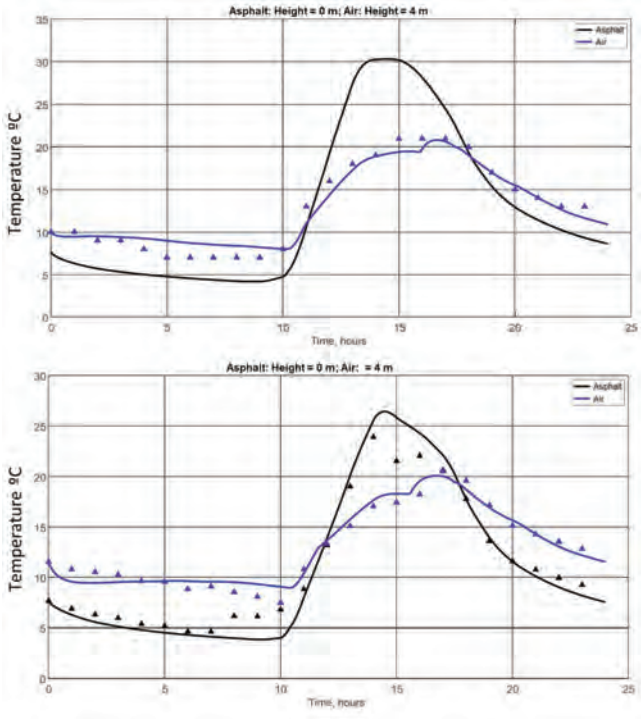


Figure 15. Comparison between simulation and experimental results in two different winter days. Solid line: Simulation results. Triangles: experimental data

6. CONCLUSIONS

Two different applications of Applied Physics have been presented. The Acoustic Emission phenomenon has been presented and a brief review of its application in Civil Engineering have been shown. It is a passive non-destructive evaluation method that provides information in real time of dynamic internal processes. It can be used to structural health monitoring of critical structures, to evaluate damage state by periodic inspections and to characterise materials.

The installation of cool pavement in Murcia urban roads has been shown. It has been demonstrated by an experimental thermal analysis that in is an effective measure to mitigate the urban heat island effect, although at the moment it has a higher cost than normal pavements. A software has been developed to analyse the effect of the installation of a cool pavement. It can calculate the decrease of temperature before the installation, so local governments can evaluate if it is worth the installation of a new cool pavement.

Acknowledgement

The author would like to thank Juan Francisco Sánchez Pérez and Francisco Miguel Moral Moreno for their assis-

tance in redacting the conclusions from LIFE HEATLAND project.

Part of this work has been supported by the LIFE Climate Change Adaptation programme 2014-2020 LIFE16 CCA/ES/000077).

REFERENCES

Akbari, H., Cartalis, C., Kolokotsa, D., Muscio, A., Pisello, A. L., Rossi, F., Santamouris, M., Synnef, A., WONG, N. H. and Zinzi, M. (2015). Local climate change and urban heat island mitigation techniques - the state of the art. *Journal of Civil Engineering and Management*, *22*(1), 1–16. doi: 10 .3846 / 13923730 .2015 .1111934

ASTM, E. (2006). *Standard terminology for nondestructive examinations", the american society for testing and materials.* ASTM International.

Behnia, A., Chai, H. K. and Shiotani, T. (2014). Advanced structural health monitoring of concrete structures with the aid of acoustic emission. *Construction and Building Materials*, *65*, 282–302.

Behnia, B., Buttlar, W. and Reis, H. (2018). Evaluation of low-temperature cracking performance of asphalt pavements using acoustic emission: A review. *Applied Sciences*, *8*(2), 306.

Behnia, B., Buttlar, W. G. and Reis, H. (2017). Spiral cracking pattern in asphalt materials. *Materials & Design*, *116*, 609–615. doi: 10 .1016 / j .matdes .2016 .11 .077

Benavent, A., Castro, E. and Gallego, A. (2010). Evaluation of low-cycle fatigue damage in rc exterior beam-column subassemblages by acoustic emission. *Construction and Building Materials*, *24*(10), 1830–1842. doi: 10 .1016 / j .conbuildmat .2010 .04 .021

Benavent-Climent, A., Castro, E. and Gallego, A. (2009). Ae monitoring for damage assessment of rc exterior beam-column subassemblages subjected to cyclic loading. *Structural Health Monitoring*, *8*(2), 175–189. doi: 10 .1177 / 1475921709102143

Businger, J. A., Wyngaard, J. C., Izumi, Y. and Bradley, E. F. (1971). Flux-profile relationships in the atmospheric surface layer. *Journal of the Atmospheric Sciences*, *28*(2), 181–189. doi: 10 .1175 / 1520 -0469(1971) 028{\textless}0181 : FPRITA{\textgreater}2 .0 .CO ; 2

Carpinteri, A. and Pugno, N. (2005). Fractal fragmentation theory for size effects of quasi-brittle materials in compression. *Magazine of Concrete Research*, *57*(6), 309–313. doi: 10 .1680 / macr .2005 .57 .6 .309

Del Carpinteri, A., Lacidogna, G. and Pugno, N. (2007). Structural damage diagnosis and life-time assessment by acoustic emission monitoring. *Engineering Fracture Mechanics*, *74*(1-2), 273–289. doi: 10 .1016 / j .engfracmech .2006 .01 .036

del Cerro Velázquez, F., Gómez-Lopera, S. A. and Alhama, F. (2008). A powerful and versatile educational software to simulate transient heat transfer processes in simple fins. *Computer Applications in Engineering Education*, *16*(1), 72–82. doi: 10 .1002 / cae .20159

di Benedetti, M., Loreto, G., Matta, F. and Nanni, A. (2013). Acoustic emission monitoring of reinforced concrete under accelerated corrosion. *Journal of Materials in Civil Engineering*, *25*(8), 1022–1029. doi: 10 .1061 / (ASCE) MT .1943 -5533 .0000647

Dobroň, P., Bohlen, J., Chmelík, F., Lukáč, P., Letzig, D. and Kainer, K. U. (2007). Acoustic emission during stress relaxation of pure magnesium and az magnesium alloys. *Materials Science and Engineering: A*, *462*(1-2), 307–310. doi: 10 .1016 / j .msea .2005 .12 .111

Drummond, G., Watson, J. F. and Acarnley, P. P. (2007). Acoustic emission from wire ropes during proof load and fatigue testing. *NDT & E International*, *40*(1), 94–101. doi: 10 .1016 / j .ndteint .2006 .07 .005

Environmental Protection Agencys Office of Atmospheric Programs. (n.d.). *U. reducing urban heat islands: Compendium of strategies.* Cool Pavements.

Fowler, T. J., Blessing, J. A., Conlisk, P. J. and Swanson, T. L. (1989). The monpac system. *Journal of acoustic emission*, *8*(3), 1–8.

Gallego, A., Molina and Martínez, E., González. (2016). *Emisión acústica. niveles i y ii (aend)*. Madrid: FC Editorial.

Golaski, L., Gebski, P. and Ono, K. (2002). Diagnostics of reinforced concrete bridges by acoustic emission. *Journal of acoustic emission*, *20*(2002), 83–89.

Gutenberg, B. and Richter, C. (1954). *Seismicity of the earth*. Princton Univ. Press.

Hao, Q., Zhang, X., Wang, K., Shen, Y. and Wang, Y. (2018). A signal-adapted wavelet design method for acoustic emission signals of rail cracks. *Applied Acoustics*, *139*, 251–258. doi: 10 .1016 / j .apacoust .2018 .04 .038

He, Y., Li, M., Meng, Z., Chen, S., Huang, S., Hu, Y. and Zou, X. (2021). An overview of acoustic emission inspection and monitoring technology in the key components of renewable energy systems. *Mechanical Systems and Signal Processing*, *148*, 107146.

Kostryzhev, A., Davis, C. and Roberts, C. (2013). Detection of crack growth in rail steel using acoustic emission. *Ironmaking & Steelmaking*, *40*(2), 98–102.

Li, S., Chen, X. and Zhang, J. (2021). Acoustic emission characteristics in deterioration behavior of dam concrete under post-peak cyclic test. *Construction and Building Materials*, *292*, 123324. doi: 10 .1016 / j .conbuildmat .2021 .123324

LIFE HEATLAND. (2021). *Project available online: https : // heatlandlife .eu /*. (accessed on 11 August 2021)

Manthei, G. and Eisenblätter, J. (2008). Acoustic emission in study of rock stability. In G. Christian and M. Ohtsu (Eds.), *Acoustic emission testing: Basics for research - applications in civil engineering* (pp. 239–310). Springer Berlin Heidelberg: Berlin, Heidelberg.

Mao, W., Yang, Y. and Lin, W. (2019). An acoustic emission characterization of the failure process of shallow foundation resting on sandy soils. *Ultrasonics*, *93*, 107–111. doi: 10 .1016 / j .ultras .2018 .11 .007

Nair, A. and Cai, C. S. (2010). Acoustic emission monitoring of bridges: Review and case studies. *Engineering Structures*, *32*(6), 1704–1714. doi: 10 .1016 / j .engstruct .2010 .02 .020

Newsletter. (2020). *Life heatland.* (April-September 2020)

NgSpice Software. (2020). http://ngspice.sourceforge.net/index.html.

Ohno, K. and Ohtsu, M. (2010). Crack classification in concrete based on acoustic emission. *Construction and Building Materials*, *24*(12), 2339–2346.

Ohtsu, M. (2018). Prospective applications of ae measurements to infra-dock of concrete structures. *Construction and Building Materials*, *158*, 1134–1142.

Ohtsu, M., Shiotani, T., Shigeishi, M., Kamada, T., Yuyama, S., Watanabe, T., Suzuki, T., van Mier, J., Vogel, T., Grosse, C. and et al. (2010). Recommendation of rilem tc 212-acd: acoustic emission and related nde techniques for crack detection and damage evaluation in concrete: Test method for damage qualification of reinforced concrete beams by acoustic emission. *Materials and Structures/Materiaux et Constructions*, *43*(9), 1177–1181. doi: 10.1617/s11527-010-9639-z

Ono, K. (2014). Acoustic emission. In T. Rossing (Ed.), *Springer handbook of acoustics* (pp. 1209–1229). Springer.

Qiu, X., Xu, J., Xu, W., Xiao, S., Wang, F. and Yuan, J. (2020). Characterization of fatigue damage mechanism of asphalt mixtures with acoustic emission. *Construction and Building Materials*, *240*, 117961. doi: 10.1016/j.conbuildmat.2019.117961

Qiu, X., Xu, J., Xu, W., Yang, Q., Wang, F. and Yuan, J. (2021). Diagnosis of damage evolution process for asphalt mixtures using pattern recognition with acoustic emission signals. *Construction and Building Materials*, *280*, 122536. doi: 10.1016/j.conbuildmat.2021.122536

Sagasta, F., Zitto, M. E., Piotrkowski, R., Benavent-Climent, A., Suarez, E. and Gallego, A. (2018). Acoustic emission energy b -value for local damage evaluation in reinforced concrete structures subjected to seismic loadings. *Mechanical Systems and Signal Processing*, *102*, 262–277. doi: 10.1016/j.ymssp.2017.09.022

Sánchez-Pérez, J. F., Marín, F., Morales, J. L., Cánovas, M. and Alhama, F. (2018). Modeling and simulation of different and representative engineering problems using network simulation method. *PLOS ONE*, *13*(3), e0193828. doi: 10.1371/journal.pone.0193828

Sánchez-Pérez, J. F., Mascaraque-Ramírez, C., Moreno Nicolás, J. A., Castro, E. and Cánovas, M. (2021). Study of the application of pcm to thermal insulation of uuv hulls using network simulation method. *Alexandria Engineering Journal*, *60*(5), 4627–4637. doi: 10.1016/j.aej.2021.03.058

Schumacher, T., Higgins, C. C. and Lovejoy, S. C. (2011). Estimating operating load conditions on reinforced concrete highway bridges with b-value analysis from acoustic emission monitoring. *Structural Health Monitoring*, *10*(1), 17–32. doi: 10.1177/1475921710365424

Sharma, A., Sharma, S., Sharma, S. and Mukherjee, A. (2018). Investigation of deterioration in corroding reinforced concrete beams using active and passive techniques. *Construction and Building Materials*, *161*, 555–569. doi: 10.1016/j.conbuildmat.2017.11.165

Shiotani, T. (2008). Parameter analysis. In G. Christian and M. Ohtsu (Eds.), *Acoustic emission testing: Basics for research - applications in civil engineering* (pp. 41–51). Springer.

U. Reducing Urban Heat Islands: Compendium of Strategies - Urban Heat Island Basics. (n.d.). *Specifications crio-9045*.

Verstrynge, E., Lacidogna, G., Accornero, F. and Tomor, A. (2021). A review on acoustic emission monitoring for damage detection in masonry structures. *Construction and Building Materials*, *268*, 121089.

Zhang, X., Feng, N., Wang, Y. and Shen, Y. (2015). Acoustic emission detection of rail defect based on wavelet transform and shannon entropy. *Journal of Sound and Vibration*, *339*, 419–432. doi: 10.1016/j.jsv.2014.11.021

Zhang, X., Hao, Q., Wang, K., Wang, Y., Shen, Y. and Hu, H. (2018). An investigation on acoustic emission detection of rail crack in actual application by chaos theory with improved feature detection method. *Journal of Sound and Vibration*, *436*, 165–182. doi: 10.1016/j.jsv.2018.09.014

Zhang, Y., Okere, C. J. and Su, G. (2021). Effect of loading rates on accurate in-situ stress determination in different lithologies via kaiser effect. *Arabian Journal of Geosciences*, *14*. doi: 10.1007/s12517-021-07674-3

Zhao, Z., Chen, B., Wu, Z. and Zhang, S. (2021). Multi-sensing investigation of crack problems for concrete dams based on detection and monitoring data: A case study. *Measurement*, *175*, 109137. doi: 10.1016/j.measurement.2021.109137

Christoph Butenweg

Christoph Butenweg is a Full Professor for Technical Mechanics and Structural Engineering at FH Aachen University of Applied Sciences (since 2016), General Manager of the SDA-engineering GmbH (since 2006), Herzogenrath and Member of the Executive Board of the Center for Wind and Earthquake Engineering, RWTH Aachen University (since 2016).

Contact information: e-mail: christoph.butenweg@fh-aachen.de

Prof. Butenweg graduated at the Ruhr-University Bochum in 1994 and received his PhD degree in 1999 at the University of Essen. He worked from 1999 to 2001 as a Structural Engineer in the engineering company „Karvanek-Thierauf“ in Essen. Between 2001 and 2016 he was employed as Chief Engineer at the Institute of Structural Statics and Dynamics, RWTH-Aachen. In 2006 he founded the engineering company SDA-engineering GmbH together with Dr.-Ing. Philippe Renault after many years of collaboration at the Chair of Structural Statics and Dynamics of RWTH Aachen University. In 2016 he was elected as Full Professor. Since 2021, Prof. Butenweg is elected as a visiting professor at the Faculty of Civil Engineering, University of Belgrade.

Prof. Butenweg authored and co-authored more than 250 national and international publications and four textbooks on Structural Statics and Dynamics. He is a former President of the German Society of Earthquake Engineering and Structural Dynamics (DGEB) (2011–2015), Member of the Executive Committee of the European Association of Earthquake Engineering (EAEE), Member of the Research Advisory Board of German Committee for Masonry Construction e. V. (DAfM), Member and qualified Structural Engineer of the Chamber of Engineers Bau NRW, Member of the Working Group Masonry WG1 of TC250/SC8 and Chairman of the international conference SEDIF (Seismic design of industrial facilities).

In addition, Prof. Butenweg is Leader of the Project Team 5 for the evolution of Eurocode 8-4/6, Chairman of the DIN Committee “NA 005-06-37 AA” and Member of the code committee construction engineering NA 005-51-06 AA, Technical committee earthquake, special problems (Preparation of the earthquake codes DIN 4149 and Eurocode 8 with National Annex for Germany).

Prof. Butenweg is a reviewer for the following journals: Bulletin of Earthquake Engineering, Advances in Civil Engineering, Engineering Mechanics, Bautechnik and Bauingenieur.

His teaching activities include the following undergraduate and graduate courses: Linear and nonlinear finite element methods, Structural Dynamics, Bionics, Earthquake Engineering, Technical Mechanics and Structural Engineering in general. His research expertise is related to complex problems in structural dynamics and earthquake engineering and computational structural mechanics.

Integrated approach for monitoring and management of buildings with digital building models and modern sensor technologies

Christoph Butenweg

FH Aachen – University of Applied Science and
Center for Wind and Earthquake Engineering – RWTH Aachen University, Germany

Summary

Nowadays modern high-performance buildings and facilities are equipped with monitoring systems and sensors to control building characteristics like energy consumption, temperature pattern and structural safety. The visualization and interpretation of sensor data is typically based on simple spreadsheets and non-standardized user-oriented solutions, which makes it difficult for building owners, facility managers and decision-makers to evaluate and understand the data. The solution of this problem in the future are integrated BIM-Sensor approaches which allow the generation of BIM models incorporating all relevant information of monitoring systems. These approaches support both the dynamic visualization of key structural performance parameters, the effective long-term management of sensor data based on BIM and provide a user-friendly interface to communicate with various stakeholders. A major benefit for the end user is the use of the BIM software architecture, which is the future standard anyway. In the following, the application of the integrated BIM-Sensor approach is illustrated for a typical industrial facility as a part of an early warning and rapid response system for earthquake events currently developed in the research project "ROBUST" with financial support by the German Federal Ministry for Economic Affairs and Energy (BMWI).

Keywords: integrated approach, monitoring systems, structural heath monitoring, integrated BIM-Sensor approach, warning and rapid response system

1. INTRODUCTION

Germany's seismicity in general is characterized as low or moderate. However, there are seismic events of moderate to strong ground motions with damage potential that arouse attention (Tailfingen 1978 $M_w = 5.7$, Roermond 1992 $M_w = 5.9$), especially as the exposed regions are densely populated, highly industrialized and have a very dense network of critical infrastructures. These infrastructures were mainly built without considering any seismic codes and provisions. As critical infrastructures according to German Federal Office for Civil Protection and Disaster Assistance are organizations and facilities with high significance for the society, their failure or impairment would result in lasting difficulties in supply chain, significant threats for the public security or other dramatic consequences (Bundesamt für Bevölkerungsschutz und Katastrophenhilfe, 2015). Although buildings are considered to be the most vulnerable component of the urban built environment to seismic hazard, the resilience of complex systems like transport and utility networks, gas and electric network systems, health care system, power plants and industrial facilities is even more critical, as they have an extreme socio-economic impact and can affect the regional prosperity (Pitilakis, 2014). Although the seismic code provisions in Germany are constantly improving (DIN EN 1998-1/NA, 2001) and the seismicity of the Lower Rhine Embayment was investigated in detail the recent years (Pilz *et al.*, 2020), an important part of the critical infrastructure was built with lower or none seismic code provisions. In combination with the high density of population and the existence of many industrial facilities, the seismic risk in this region increases significantly, despite the rather low possibility of occurrence of earthquake events with $M_w > 6.5$. According to Pilz *et al.* (2020), in this region there are 51 electricity supply substations, 108 hospitals, about 1500 industrial facilities, 4 industrial chemical parks and 8 road and railway bridges over the Rhine (Figure 1). Interactions between complex systems may result in rapid damage propagation. Therefore, there is an increased necessity for interconnected monitoring systems, that can detect damage and provide immediately this information to crisis management centres. Appropri-

ate countermeasures to prevent or mitigate damage can be triggered effectively only if sufficient information is available. Knowledge of the structural health status of critical nodes of the transport network, such as bridges, is beneficial for the definition the optimal routes for rescue teams and ambulances. Chemical plants are an important component of the infrastructure in the Lower Rhine region. The problem after the occurrence of an earthquake in these chemical plants with hazardous processes is the risk of accidental scenarios such as fire, explosion or dispersion of toxic substances due to leakage of content. Immediate countermeasures must be triggered to limit the consequences of such events in terms of fatalities, environment pollution and repair cost (Paolacci *et al.*, 2012).

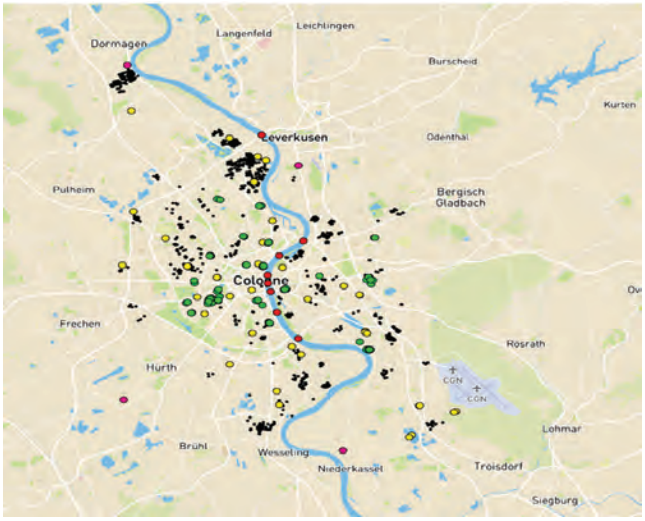


Figure 1. Distribution of critical facilities in Lowe Rhine Region: Electricity substations (yellow circles), hospitals (green circles), road & railway bridges over Rhine (red circles), industrial facilities (black circles), chemical industry parks (purple circles) (Pilz *et al.*, 2020)

It can be summarized, that there is a need to quickly assess the possible damage and safe operation of critical infrastructures after seismic events and to pass on the information to operators and authorities. This situation prompted the German Federal Ministry for Economic Affairs and Energy (BMWI) to start the development of a novel early warning and rapid response system for earthquake events in the Lower Rhine region within the project "ROBUST".

Fundamental element of the novel system is the automated interaction of smart sensors or sensor systems, which are capable not only to record motions or strains, but also to process the recordings decentral and to forward the results of the assessment. In the past decades, the breakthrough in software and hardware development of sensors and techniques for data processing enabled the extended and low-cost use of a remarkable number of seismic and monitoring sensors (Rainieri *et al.*, 2008). Reliable earthquake early warning systems were developed all over the world and especially in seismic prone countries as Japan, USA and Mexico (Allen *et al.*, 2009). At the same time, the important research areas of engineering diagnostics and post-earthquake damage detection gained more attention with promising results. Acknowledging the developments in both directions, the examined system tries to incorporate and extend their advantages and provide a useful tool for the protection of critical infrastructures (Wu and Beck, 2012). The system to be developed is going to be applied to a production unit of a chemical plant in the Lower Rhine Region in the western part of Germany. The innovative system consists of four basic components: (i) the seismic sensor network, (ii) the local monitoring systems within the facility, (iii) the communication infrastructure and (iv) the integration of the sensor data in BIM models of the monitored facility.

Regarding the first component, namely the seismic network for the detection of ground motions, the existing sparse regional network of the Geological Service in Krefeld is being extended with new "intelligent" seismic sensors. To this end an algorithm has been developed, aiming at the optimization of the location of the new sensors such, that while accounting for the geophysical conditions, reliable information on earthquakes relevant for the infrastructures can be obtained and communicated to maximize the lead time while, at the same time, to minimize the risk of false alarms. The novel sensors, developed by the German Research Center for Geosciences (GFZ), are capable to evaluate the ground motion and the spectral intensities in real time and send out messages via the communication network to the operators and authorities (e.g. civil protection) before the arrival of the damaging seismic shear waves.

The information on a detected earthquake triggers also the local monitoring systems to increase the sampling rate for the installed sensors, to generate detailed information for the evaluation of the structural response and condition of the technical installations. The exchange of information during the seismic event takes place via a communication network that consists of several decentralized nodes. Each seismic sensor and monitored infrastructure corresponds to a node within the communication infrastructure, responsible to send and receive the appropriate information. Simultaneously, the same communication infrastructure, under development by Fraunhofer Institute for Open Communication (FOKUS), will provide damage and operating conditions to the concerned parties (authorities, operators of critical infrastructures, civil protection). This allows the recipients to trigger automated or semi-automated safety measures such as emergency shutdowns, to activate fire alarm systems or to prevent the access to facilities with critical damage through a red traffic light. The kind and the intensity of the measures depend on the expected damage from the detected ground motion and can be updated relying on the most current measured

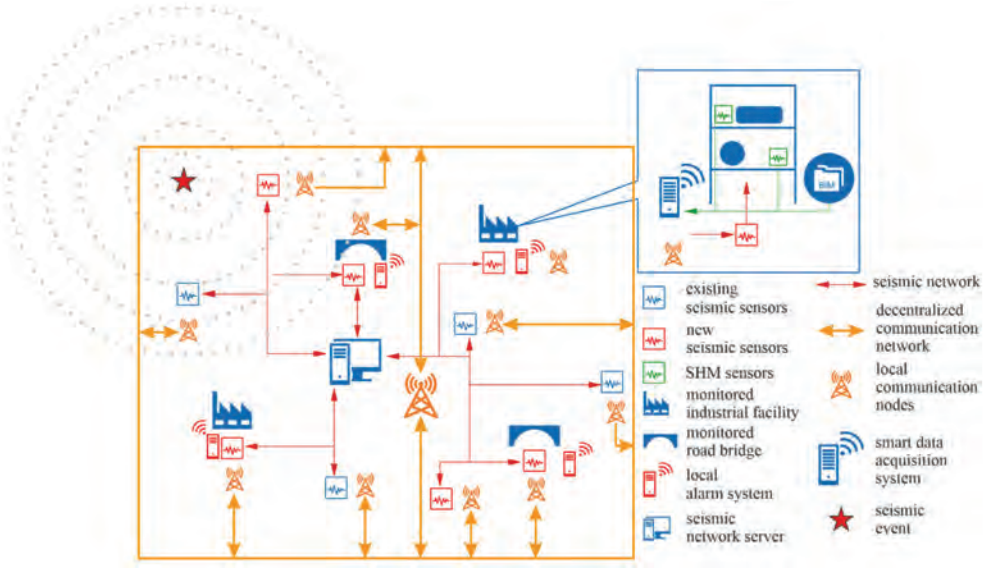


Figure 2. Earthquake early warning and rapid response system with local SHM integrated in BIM (Balaskas *et al.*, 2021)

data. After the assessment of the seismic performance of the monitored structures, structural health reports can be provided to the parties involved.

The local monitoring systems consist of a number of sensors (accelerometers and velocity sensors, displacement transducers, strain gauges, temperature sensors) and data acquisition devices. The sensors are placed in appropriate positions of the facility to record effectively the response of particular damage indicators. The prediction of damage is based on specific damage indicators of the structural system and technical installations. In nearly real time, the measurements from the monitoring sensors are evaluated either decentral and locally by intelligent data acquisition systems or remotely by linked computer networks. The results of the damage identification are integrated in two steps in the communication platform. Initially, right after the earthquake, only basic information is forwarded (e.g. information that a part of a facility has been collapsed). A few minutes after the event, more detailed information based on a further assessment of the sensor recordings are communicated, including a graphical representation of the damage.

One fundamental objective and innovative part of the project ROBUST is the spatial representation of the damage distribution of the monitored industrial facility right after an earthquake which becomes possible through the integration of the local structural health monitoring (SHM) systems in the BIM model of the facility. This requires the registration of the sensors by type and location in the BIM model. For this purpose, it is intended to use the DESITE BIM (DESITE BIM, n.d) software package, which provides the opportunity to link sensor measurement data with BIM models. As BIM models with integrated sensor information are platform independent, they can be easily used on any user compatible end device like tablets or smartphones. This facilitates the usage of the models and sensor data results through BIM viewers on tablets or smartphones by different end users (e.g. facility operators, civil protection).

Figure 2 shows the overall approach with existing and new seismic network nodes, local SHM systems integrated in BIM, local communication nodes and applications of the end users as local alarm systems and smart data communication systems.

2. EARTHQUAKE EARLY WARNING AND REACTION SYSTEM

2.1. General Information

The Lower Rhine Embayment in the west part of Germany is one of the most important areas of earthquake recurrence north of the Alps, facing a moderate level of seismic hazard in the European context but a significant level of risk due to a high population density and a large number of critical industrial infrastructures. As the seismic faults are directly crossing the study area, the lead time, i.e. the time between the onset of an earthquake and the arrival of the destructive seismic waves is in the order of a few seconds at most for events in the Lower Rhine Embayment. In this context, the project aims at designing a user-oriented hybrid earthquake early warning and rapid response system where regional seismic monitoring is combined with smart, on-site sensors, resulting in the implementation of decentralized warning procedures.

2.2. Seismic scenarios for the area of interest

The first step is the optimal arrangement of the regional seismic network arrangement based on the seismicity in the regarded area. The applied optimization approach

is based on representative samples of relevant seismic recordings. The design of an earthquake early warning network requires the consideration of all potential and known earthquake sources in the regarded area. Due to the sparsity of large earthquake recordings in the region, stochastic simulations of scenario earthquakes were performed using the finite-source ground motion simulation code EXSIM (Boore, 2009)] and the known seismicity in the area (Grünthal *et al.*, 2018). Overall, more than 700 realistic scenario earthquakes are considered (Figure 3). The proposed approach can be used for the generation of a large set of spectra for hypothetical sources located according to a seismic catalogue.

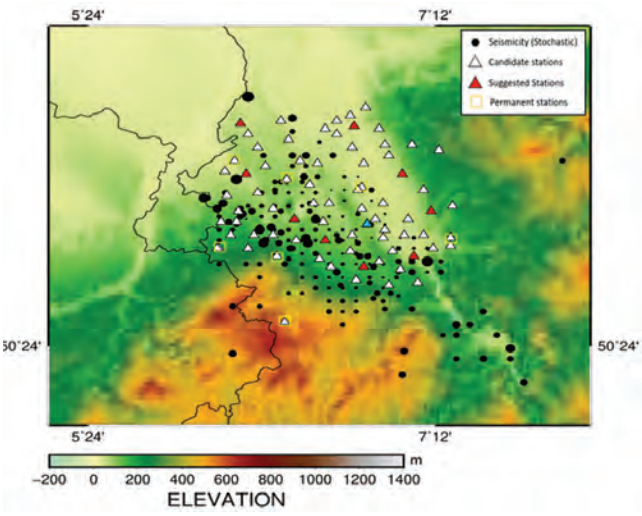


Figure 3. Potential seismic network in the Lower Rhine Embayment, Germany with existing permanent stations (yellow squares), candidate stations (white triangles) and suggested installations sites (red triangles). The simulated scenario earthquakes are shown as black dots with diameter proportional to earthquake magnitude (Balaskas *et al.*, 2021).

2.3. Optimal densification for real time assessment and early warning

The locations of the seismic stations were chosen with respect to the distribution of critical infrastructures in the Lower Rhine Embayment. For each location, stochastic seismograms were computed for all scenario earthquakes resulting in a database consisting of more than 30.000 traces representing ground acceleration. An important aspect of the earthquake early system performance is its ability to give not only timely, but also reliable warnings with respect to the expected shaking at the target site. The optimal layout of the network is the one that can predict the correct level of expected ground motion at the target site at a sufficiently long lead time for the highest number of scenario earthquakes. To quantify the network quality, an objective function according to Oth *et al.* (2010) is used. For finding the optimal station locations, a micro genetic algorithm was applied. Such algorithms are a specific subset of genetic algorithms with special search algorithm based on evolutionary principles to find optimal models with respect to a given objective function. By minimizing the objective function, the best network distribution of existing and additional seismic stations can be derived (Figure 3). As can be seen by the red triangles, by adding nine additional stations to the existing network, such optimal design will allow strong ground-motions with respect to the target site to be detected most quickly and most reliably.

2.4. Immediate reaction system

An earthquake early warning system can reduce the seismic risk on critical facilities more effectively combined with a reaction system. The term reaction system covers all automated measures that can be taken immediately after the detection of the first earthquake waves. Structural damage is not the only threat, especially for critical facilities. Accidents or side effects due to a seismic event such as fire, explosions or dispersion of toxic substances can have even more harmful consequences than the structural damage itself (Salzano *et al.*, 2009). Possible countermeasures to limit damage and prevent domino effects that can be performed in range of seconds are emergency shutdowns of processes, the closure of pipe systems, the activation of backup measures and the emergency stop of critical technical installations. In the context of the ROBUST project, the warning for an upcoming seismic event and the damage prediction is forwarded to the facility operators, who are responsible for triggering appropriate and plant specific measures.

2.5. Communication network

Essential part of the developed system is the decentralized communication network, which enables the interaction between seismic sensors and local monitoring sensor systems and forwards warning messages and structural health status reports user oriented to subscribed recipients. Typical users are facility managers, local authorities, civil protection or rescue teams. Each user has access to different level of information depending on his involvement. From the communication infrastructure point of view, all the participants are divided into two categories: (i) sensors and (ii) actors. Sensors “produce” information and actors “consume” information. The nodes are spatially distributed and one node corresponds to each seismic sensor. At the monitored facility there is one node responsible for the seismic sensor and the controller of the local monitoring system of the facility. Predefined abonnements specify the flow of information between the nodes.

At the idle period, the platform is responsible to control the connection of the devices, their health status and the plausibility of the sensor measurements. The responsible administrators of each subsystem are informed through automated messages about the health status of the whole sensor infrastructure. When a ground motion detected by the seismic network exceeds the acceleration threshold of a specific facility, the communication node of the corresponding seismic sensor forwards the information to the subscribed recipients. The communication network delivers the information of an upcoming event via messages and an indicative lamp system. Synchronously, the communication node of the monitored facilities triggers the measuring system. Right after the earthquake, the health status of the facility based on the assessment of the recording of the monitoring system is forwarded via the communication infrastructure to the interested parties. The first report includes a first assessment of the performance of the structure (partial or total collapse of the facility or parts, increased risk for fire or explosions, process interruptions). After a few minutes, a second more precise health status report is delivered. It contains information about the detection of damage and their spatial distribution. The second health status report is accompanied with information about the operability of the facility and provides a three-dimensional representation of the structural damage and its distribution. Furthermore, the health status of critical technical installations will be distributed.

3. SELECTION OF THE FACILITY AND DAMAGE INDICATORS

An industrial facility in the Lower Rhine Region was selected as a case study for the installation of the integrated SHM – BIM approach and their coupling with the seismic network by means of the communication network. The facility, a typical chemical production unit, cannot presented here, since the operator only provided the access to the high-risk facility with observance of the confidentially agreement.

Therefore, the basic approach is explained using a typical chemical production unit that was tested on a shaking table at EU-Centre in Pavia within the framework of the project SPIF project "Seismic Performance of Multi-Component Systems in Special Risk Industrial Facilities" (Butenweg *et al.*, 2020). This project was carried out within the framework of SERA (Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe) financially supported by the European Union. Figure 4 shows the test structure consisting of a primary steel structure supporting several process units. The main structure was conceived as a three-storey steel frame with flexible diaphragms made of crossbeams hinged to the frame beams. Figure 4 shows some views of the test structure on the shaking table. The structure is 3.7 m × 3.7 m in plan with a storey height of 3.1 m, which leads to a total height of 9.3 m. The structure was clamped to a reinforced concrete base plate. This latter has dimensions of 4.8 m × 4.8 m and a thickness of 0.4 m and was constructed with a concrete class C30/37. The experimental data of the test structure and the corresponding BIM model are also used in the ongoing ROBUST project as a test framework for development and implementation.

Figure 4. Test structure of the SPIF-project (Butenweg *et al.*, 2020)

For the selected facility, probabilistic earthquake risk analyses will be carried in advance, to define the consequences of different seismic hazard levels. The risk assessment considers the consequences of damage to the load bearing system and to the installed technical components as well. The vulnerability is described by vulnerability curves, which will be defined based on relevant literature (INDUSE-2-SAFETY, 2019). At first, the most critical parts due to potential seismic damage of the primary structural system are identified. Thereafter, the most important and sensitive components in the facility are identified and prioritized with respect to their impact on the plant and surrounding area and their importance for the production process. Typical components are storage tanks, vertical and horizontal vessels, processing equipment, pumps usually connected to each other by piping systems. Due to the complex interactions of single installations a multiplicity of accident chains can be generated (Nasserasadi and Ghafory-Ashtiany, 2008). In the next step the failure modes and the position of the expected damage are investigated in detail. After the basic identification of the facility, suggestions for further damage prevention or damage limitation will be provided and discussed with the facility operator (Paolacci *et al.*, 2012). Emergency actions mitigating the structural damage or domino effects resulting from structural damage will be investigated. For this reason, appropriate damage indicators and assessment methodologies, providing useful information about the global performance of the structure or local failures of critical component will be implemented. Possible damage indicators are typically peak floor accelerations (PFA), interstorey drift ratios (IDR), residual IDR and strains in critical connections (e.g. elbow or tee connections of pipelines) (Varelis *et al.*, 2012). The number

and types of the selected sensors (accelerators and velocity sensors, displacement transducers, strain gauges etc.) will be chosen with respect to the damage indicators specifically for each facility.

Immediate damage detection, especially for critical facilities, is essential to minimize their operation interruption due to shut down time. Traditionally, engineering inspections and nonlinear analyses are necessary to evaluate the capacity of the structures after seismic events. This may last weeks or months, accompanied with high economic costs due to shut downs. Modern structural health systems can provide fast and nearly automated damage detection based on changes of the dynamic behaviour of the monitored structures (Hwang and Lignos, 2018). There is a variety of methods both in time and frequency domain such as autoregressive methods, eigensystem realization algorithms, stochastic subspace identification and frequency domain decomposition methods, which can be used to assess the seismic performance by applying different indicators (Brincker and Ventura, 2015). Moreover, the arrangement of a geotechnical seismic sensor close to the monitored facility, allows the application of either output only or input/output methods (Avitabile, 2018). For different structural types, different types of damage indicators may be effective (Hwang and Lignos, 2018). The functionality of the measurement chain, selected damage indicators and assessment methods will be verified in benchmark scaled structures and substructures in laboratory in advance. After an earthquake, the fundamental information that should be extracted from the monitoring system concern the operational ability of the corresponding infrastructure depending on the distribution of damage. The smart data acquisition devices, which are capable of processing the data in real time and forward the crucial information to the communication network (National Instruments, 2021) plays an important role in the integration of local monitoring system to the overall early warning and reaction system.

4. INTEGRATION OF SHM IN BIM

The integration of local SHM systems in BIM is one the main objective of the research project ROBUST. BIM is an intelligent process, which promotes the interdisciplinary cooperation of different architects, engineers, constructors and facility operators from the planning phase, through construction to the facility operation (Delgado *et al.*, 2016; Sacks *et al.*, 2018). The integration of local SHM systems in BIM offers geospatial content to the measured and processed data and provides a better control and management of the facility (Rüppel, 2019; Valinejadshoubi *et al.*, 2017).

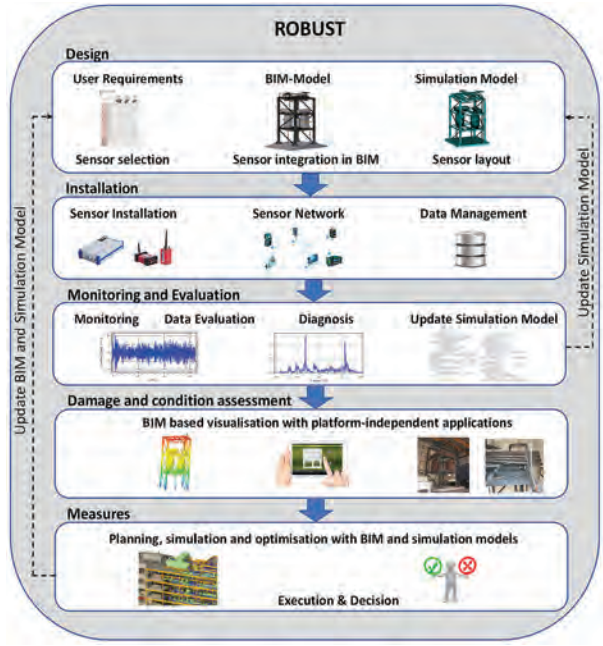


Figure 5. Integration of structural health monitoring in BIM

The integrated SHM-BIM approach is subdivided in the five consecutive working steps ***design, installation, monitoring and evaluation, damage and condition assessment*** and ***measures*** as illustrated in Figure 5. The overall approach starts with a detailed design phase in which the sensor integration in BIM and the optimization of the sensor layout with BIM-compatible simulation models take place. Intelligent sensors with their own processors will also be used for decentralized data evaluation with low data transfer. After physical installation of the sensor network in combination with efficient data management, con-

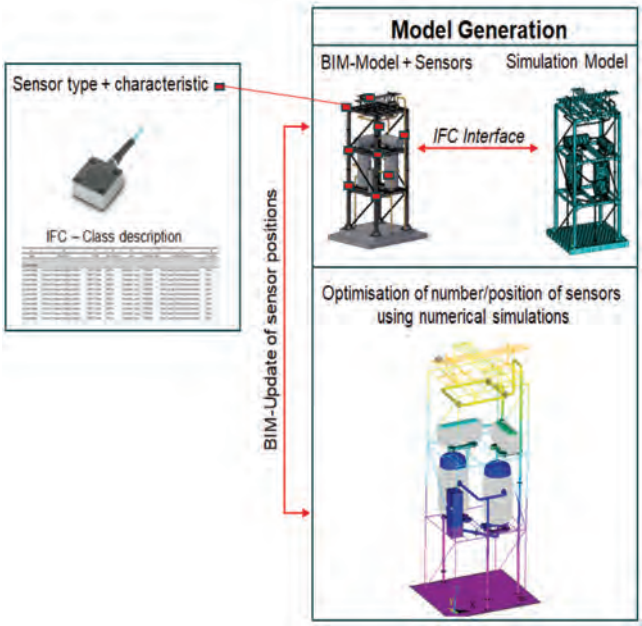


Figure 6. Optimal integration of sensors in BIM supported by simulation models

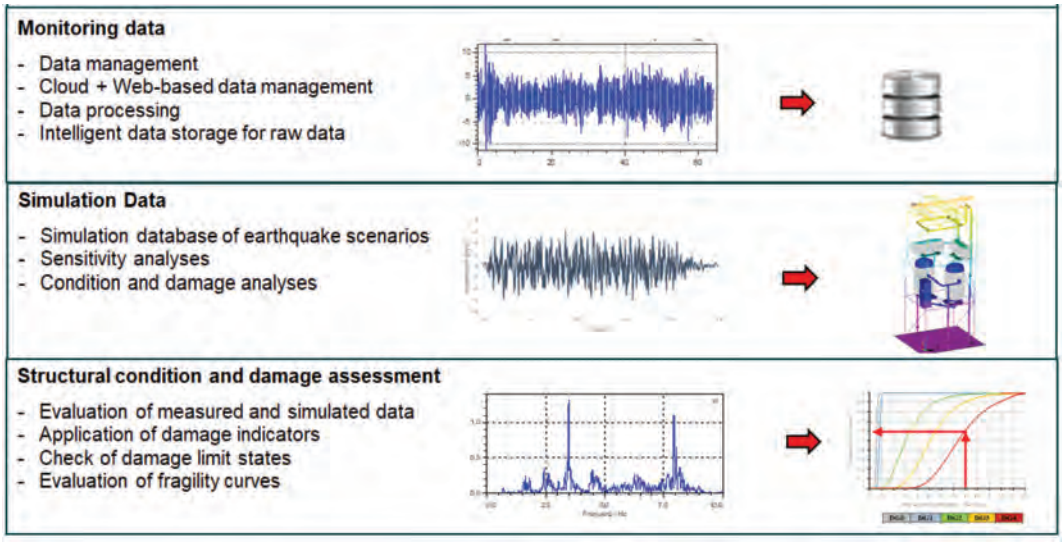


Figure 7. Monitoring, simulation and condition and damage assessment based on data evaluation

tinuous evaluation and processing of the real-time data for a BIM-based visualization of the measurement data is executed with the start of the condition monitoring. The condition assessment based on the measurement data forms the basis for planning of appropriate measures that can be improved and optimised with simulation models. The new approach can be applied to any kind of facility and different types of sensor systems. Based on the connection and continuous updating of the BIM and simulation models, the approach provides a reliable decision-making basis for operation, maintenance and repair. Due to the continuous changes in production units, the automated updating of BIM and simulation models is key aspect of the new approach.

The first working step "***design***" includes the development of a generalised solution approach for the visual and IT integration of sensors in BIM based on a specification for data communication between sensors and BIM. The visual integration of sensors in BIM is achieved by mapping the sensors as components in the BIM model with their specific positions. The IT integration of sensors in BIM will be realized by using IFC (Industrial foundation classes), which is the international standard for semantic building models (ISO 16739-1, 2018). It is planned to integrate the information of the sensors as part of the general building information. This requires an extension of the IFC description, as the current modelling options and extendable constructs are not sufficient (Chen *et al.*, 2014; Delgado *et al.*, 2016; Rüppel, 2019). The aim is a user-friendly integration of sensors, whereby intelligent sensors with their own processors and decentralised data evaluation are also envisaged. In addition, the type of the sensor network with interactions and communication protocols are defined and described. The informative value of a building monitoring highly depends on the sensor layout, in which the selection, positioning, specific functionality and the measurement data to be recorded from the sensors are determined. Furthermore, the time and duration of the monitoring must be taken into account. Thus, monitoring can take place over the entire life cycle or only within the scope of defined execution phases. Therefore, it is the aim to support the design process of the sensor layout through the coupling with BIM. For this purpose, corresponding simulation models are created on the basis of the digital building data, which are used for the system identification and the optimisation of the sensor layout. This approach is fully in line with the BIM philosophy to realize an optimal and cost-effective design through a more time intensive and improved planning. Figure 6 illustrates the visual and IT integration of sensors in BIM.

The next working step is the "***installation***" of the sensor system that includes the development of the data management and communication scheme of the sensor network. Software modules will be developed for the continuous data acquisition and transfer of sensor data to BIM and for data reduction and compression for long-term storage on central data servers. Data management also includes the optimization of data traffic through decentralized evaluations of measurement data on intelligent sensors. Furthermore, a concept for an adaptive adjustment of the sampling rates to the usage and environmental conditions is being developed. The data management also includes the efficient storage of time-dependent state variables (e.g. temperature, frequencies, ...) as a result of data processing.

Within the working step "***monitoring and evaluation***" the physically installed monitoring system is coupled with the BIM based facility model and efficient algorithms for the data evaluation of real-time measurement data (e.g. acceleration time histories), that are developed and implemented. The time dependent measurement data are used for frequency analysis and determination of maxi-

mum values at the measuring points of the structural system and technical installations. The results of the data analyses serve as a basis for condition assessment using appropriate damage indicators. In addition, the measurement data are used for system identification to update the simulation model. Herein, it is advantageous that the influence of system changes can be specifically investigated with the generated measurement data.

Figure 7 shows the monitoring and evaluation with data acquisition and processing, the simulation models for sensitivity analysis and simulation of earthquake databases and the structural condition and damage assessment.

In the working step "***damage and condition assessment***", the time-varying state variables are visualised spatially and building-related with the BIM model in a way that is easy for the user to understand and interpret the date. This is possible because the distribution of the sensors in the building is known through the integration of the sensors in BIM. The definition of the state variables and indicators to be visualised will be implemented in such a way that they can be adapted to the problem and in a user-oriented manner. Furthermore, tools are created for the end user that allows the generation of different views on the results. A challenge in this working step is the data storage and the representation of nearly real-time data. Figure 8 shows the user-oriented data visualization based on cloud or web-based databases.

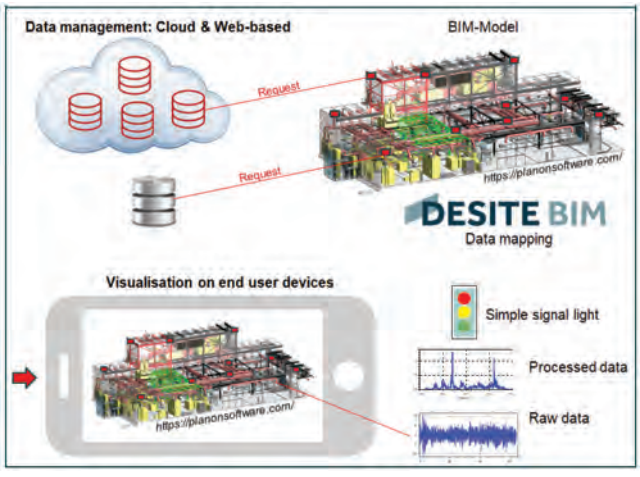


Figure 8. User-oriented data evaluation based on cloud or web-based databases

The last working step includes the optional implementation of measures as a result of the condition assessment and visualization of the results. The measures range from ad-hoc measures through building automation to changes in use and structural changes. Since the effects of the measures are complex, the simulation models are used to optimize them in terms of their effectiveness. Overall, this optimization step is intended to objectify the measures. With the approach pursued, the resulting measures are based on an integral evaluation that could not be achieved by defining measures based on local effects.

5. CONCLUSION

The contribution presents the application of an integrated BIM-SHM approach for a typical industrial facility as part of an early warning and rapid response system for earthquake events currently developed in the research project "ROBUST" with financial support by the German Federal Ministry for Economic Affairs and Energy (BMWI). The concept of a user-oriented earthquake early warning and rapid response system is based on interconnected decentralized intelligent sensors and local sensor systems integrated in BIM. The investigated system is prototypically applied in Lower Rhine Region, which is one of the most important seismic areas in Germany. The essential relay between the seismic network (responsible for the early warning), the local monitoring systems (responsible for the evaluation of the seismic performance of critical infrastructure) and the interested parties, namely the facility operators and the local authorities, is the decentralized communication network, which enables the transfer of warning, health status and critical information. The integration of SHM system in BIM models is a step forward in the management of industrial facilities. The possibility to visualize detected damage through BIM viewers on tablets or smartphones by different end users will support the understanding and the decision-making processes for facility managers and will allow to reduce production interruptions and as a consequence the economic losses. Furthermore, it is a useful tool to support civil protection measures and rescue teams.

ACKNOWLEDGEMENT

The author is grateful for the financial support by BMWI, which funded the ROBUST research project in the thematic program "Geo-Research for Sustainability (GEO:N)" in the framework program "Research for Sustainable Development (FONA 3) under the call "Natural Risks – Early Detection of Earthquakes and their Consequences". Furthermore, the author would like to thank Georgios Balaskas, Benno Hoffmeister from the Institute of Steel Construction, RWTH Aachen University, Marco Pilz from the Department of Seismic Hazard and Risk Dynamics, GFZ German Research Centre for Geosciences and Anna Bauer from Wölfel Engineering & Co. GmbH for their important contributions to this publication.

REFERENCES

Allen, R. M., Gasparini, P., Kamigaichi, O. and Bose, M. (2009). The status of earthquake early warning around the world: An introductory overview. *Seismological Research Letters*, *80*(5), 682–693. doi: 10 .1785 / gssrl .80 .5 .682

Avitabile, P. (2018). *Modal testing: a practitioner's guide*. John Wiley & Sons. Available: http://onlinelibrary.wiley.com/book/10.1002/9781119222989.

Balaskas, G., Hoffmeister, B., Butenweg, C., Pilz, M. and Bauer, A. (2021). Earthquake early warning and response system based on intelligent seismic and monitoring sensors embedded in a communication platform and coupled with bim models. In M. Papadrakakis and M. Fragiadakis (Eds.), *8th eccomas thematic conference on computational methods in structural dynamics and earthquake engineering* (pp. 987–998). Streamed from Athens, Greece, 27–30 June.

Boore, D. M. (2009). Comparing stochastic point-source and finite-source ground-motion simulations: Smsim and exsim. *Bulletin of the Seismological Society of America*, *99*(6), 3202–3216. doi: 10.1785/0120090056

Brincker, R. and Ventura, C. (2015). *Introduction to operational modal analysis*. John Wiley & Sons. Available: http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1017454.

Bundesamt für Bevölkerungsschutz und Katastrophenhilfe. (2015). *Bundesamt für bevölkerungsschutz und katastrophenhilfe*. https://www.bbk.bund.de/DE/AufgabenundAusstattung/KritischeInfrastrukturen/kritischeinfrastrukturen_node.html. Bonn.

Butenweg, C., Bursi, O., Paolacci, F., Marinković, M., Lanese, I., Nardin, C., Quinci, G., Butenweg, C., Marinković, M., Pavese, A. et al. (2020). Seismic performance of an industrial multi-storey frame structure with process equipment subjected to shake table testing. *Engineering Structures*, *243*. doi: 10.1016/j.engstruct.2021.112681

Chen, J., Bulbul, T., Taylor, J. E. and Olgun, C. G. (2014). *A case study of embedding real-time infrastructure sensor data to bim*. American Society of Civil Engineers (ASCE).

Delgado, D., M, J., Butler, L. J., Gibbons, N., Brilakis, I., Elshafie, M. Z. and Middleton, C. (2016). Management of structural monitoring data of bridges using bim. In *Proceedings of the institution of civil engineers-bridge engineering* (Vol. 170, pp. 204–218).

DESITE BIM. (n.d). https://thinkproject.com/de/produkte/desite-bim.

DIN EN 1998-1/NA. (2001). DIN EN 1998-1/NA, National Annex – Nationally determined parameters – Eurocode 8: design of structures for earthquake resistance – part 1: general rules, seismic actions and rules for buildings. *DIN Deutsches Institut für Normung e.V. Berlin*.

Grünthal, G., Stromeyer, D., Bosse, C., Cotton, F. and Bindi, D. (2018). The probabilistic seismic hazard assessment of germany—version 2016, considering the range of epistemic uncertainties and aleatory variability. *Bulletin of Earthquake Engineering*, *16*(10), 4339–4395. doi: 10.1007/s10518-018-0315-y

Hwang, S.-H. and Lignos, D. G. (2018). Assessment of structural damage detection methods for steel structures using full-scale experimental data and nonlinear analysis. *Bulletin of Earthquake Engineering*, *16*(7), 2971–2999. doi: 10.1007/s10518-017-0288-2

INDUSE-2-SAFETY. (2019). *Component fragility evaluation, seismic safety assessment and design of petrochemical plants under design-basis accident conditions (induse-2-safety): Final report*. Luxembourg: Publications Office of the European Union.

ISO 16739-1. (2018). *Industry foundation classes (ifc) for data sharing in the construction and facility management industries — part 1: Data schema*.

Nasserasadi, K. and Ghafory-Ashtiany, M. (2008). *Major contributing factors to seismic risk of industrial facilities*.

National Instruments. (2021). *Specifications crio-9045*. https://www.ni.com/pdf/manuals/376783f_02.pdf.

Oth, A., Böse, M., Wenzel, F., Köhler, N. and Erdik, M. (2010). Evaluation and optimization of seismic networks and algorithms for earthquake early warning–the case of istanbul (turkey). *Journal of Geophysical Research: Solid Earth*, *115*(B10), 321. doi: 10.1029/2010JB007447

Paolacci, F., Giannini, R. and De Angelis, M. (2012). Analysis of the seismic risk of major-hazard industrial plants and applicability of innovative seismic protection systems. *Petrochemicals*.

Pilz, M., Nievas, C., Prehn, K., Razafindrakoto, H., Schorlemmer, D., Weatherill, G. and Cotton, F. (2020). *Seismic risk analysis in germany: an example from the lower rhine embayment. final report*. GFZ German Research Centre for Geosciences.

Pitilakis, K. (2014). *Syner-g: typology definition and fragility functions for physical elements at seismic risk: Buildings, lifelines, transportation networks and critical facilities*. Dordrecht: Springer.

Rainieri, C., Fabbrocino, G. and Cosenza, E. (2008). Structural health monitoring systems as a tool for seismic protection. In *The 14th world conference on earthquake engineering* (pp. 12–17).

Rüppel, K., Uwe Smarsly. (2019). *Bim und sensorik an den beispielen brandschutz und bauwerksmonitoring, buch: Building information modelling*. Springer-Verlag.

Sacks, R., Eastman, C., Lee, G. and Teicholz, P. (2018). *Bim handbook: A guide to building information modeling for owners, designers, engineers, contractors, and facility managers*. John Wiley & Sons. Available: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119287568.

Salzano, E., Agreda, A. G., Di Carluccio, A. and Fabbrocino, G. (2009). Risk assessment and early warning systems for industrial facilities in seismic zones. *Reliability Engineering & System Safety*, *94*(10), 1577–1584. doi: 10.1016/j.ress.2009.02.023

Valinejadshoubi, M., Bagchi, A. and Moselhi, O. (2017). Managing structural health monitoring data using building information modeling. In *Smar 2017 - fourth conference on smart monitoring, assessment and rehabilitation of civil structures*.

Varelis, G. E., Karamanos, S. A. and Gresnigt, A. (2012). Steel elbow response under strong cyclic loading. In *The twenty-second international offshore and polar engineering conference* (pp. 487–496).

Wu, S. and Beck, J. L. (2012). Synergistic combination of systems for structural health monitoring and earthquake early warning for structural health prognosis and diagnosis. In *Health monitoring of structural and biological systems 2012* (Vol. 8348, p. 83481Z).

Slobodan Djordjević

Slobodan Djordjević is Professor of Hydraulic Engineering at the University of Exeter, College of Engineering, Mathematics and Physical Sciences (since 2012), Co-director of the Centre for Water Systems and Programme Director for MSc in Water Engineering. He has held visiting professor positions at The University of West Indies, Technical University of Munich and University of Belgrade.

Contact information: e-mail: s.djordjevic@exeter.ac.uk

https://orcid.org/000-0003-1682-1383

Prof. Djordjević graduated in 1984 from the Faculty of Civil Engineering, University of Belgrade (FCEUB), received his master's degree in 1988 and his doctorate in 2001. He received the Award of the "Jaroslav Cerni" Fund. He worked at the FCEUB as an Assistant (1991-2001) and Assistant Professor (2001-2002). He held tutorials in Fluid Mechanics, Hydraulics, Measurements in Hydraulic Engineering and Water Supply and Sewerage. As an Assistant Professor, he taught Hydraulics 2 and Ecological Engineering. He was the secretary of the Council of the Department of Hydraulic Engineering. In his graduate and postgraduate studies, he worked under the guidance of Prof. Miodrag Radojković, learning from him the principles and methods of numerical modeling in hydraulics and consistency in research. Prior to appointment as Assistant, he was on an internship at the Technical University of Delft and on professional training at the Florida State University, where he worked on modelling of transport processes in free-surface flows. During his employment at the FCEUB, beside teaching activities, Prof. Djordjevic mostly worked on the development and application of software for flood modelling and analysis and design of sewer systems, and he also received his doctorate in that field. He participated in the preparation and organization of five international scientific conferences held in Portorož and Dubrovnik. He worked on several drainage studies in former Yugoslavia.

Prof. Djordjević joined the University of Exeter in 2002, where he served in various leadership roles, most recently as the inaugural Associate Dean (International & Development) in the College. Research interest of Prof. Djordjević are in the development and application of advanced methodologies and tools for water management, simulation of drainage networks, pipe blockages and sewer ventilation, flood modelling and assessment of diverse impacts of flooding, scour around bridge piers, water quality modelling, tidal energy extraction, water-food-energy nexus, cascading effects between water, waste, energy, transportation and other infrastructure systems, nature-based solutions and drought risk management.

He is the co-author of 10 chapters in monographs and 80 publications in journals, and his works in the field of flood impact modeling are among the most cited works in that field. He supervised 20 doctorates. He co-chaired the international conferences ICFR2013 and CCWI2019. He serves as Editor-in-Chief of Water Supply Journal.

He maintains cooperation with colleagues from Serbia through joint publications, projects RAMB, RECONECT and others.

Impacts of flooding on road and rail transport systems

Slobodan Djordjević

College of Engineering, Mathematics and Physical Sciences, University of Exeter, United Kingdom

Summary

Most of the casualties during extreme weather events are caused by people trying to drive through flood water. In addition, other flood impacts on road and rail transport systems can be very significant. Therefore, to tackle this problem and to be better prepared for impacts and be in a stronger position to manage emergencies, there is a strong need to combine and integrate various technologies. This paper describes recent progress in merging flood modelling, traffic simulation and data processing techniques, aiming at improving the strategies and measures to deal with impacts of flooding on transport systems. Novel methodology and various applications are described, along with some directions for future research.

Keywords: flooding, impact assessment, road and rail transport, traffic simulation

1. INTRODUCTION

Flooding can cause casualties, displacement of people and other huge impacts. Recent events around the world and the projections of climate change and of urban growth indicate that the frequency and intensity of floods will increase in many regions, especially in cities. Flood risk management strategies are principally based on the assessment of diverse actual and potential impacts of flooding. The effectiveness of measures aimed to increase resilience to flooding can be quantified by comparing their cost with the reduction in various types of flood damage achieved by those measures. At first instance, typically this is done for the present situation and by using any available information on flood damage recorded during historical events. Ideally, this should be followed by an analysis of a range of scenarios covering future economic development, changes in rainfall patterns and design storms due to global warming, population growth and movement, and changes in land-use in rural areas and in urban fabric in cities. There are numerous methods for modelling those scenarios, different ways how such an analysis can be done and how the recommendations can be formulated but, importantly, the reliability of all those depends on the accuracy of flood impact assessment.

Flood impacts are typically classified using two criteria (Hammond *et al.*, 2015). The first criterion distinguishes between *tangible and intangible* impacts. Tangible damages are those that can be quantified in monetary terms, examples being the damage to property (Chen *et al.*, 2016) or the loss of profits if a business is disrupted (Chen *et al.*, 2013). Intangible impacts are those that cannot be readily quantified in monetary terms, such as the loss of life, the negative impact on the human health and mental well-being, and impacts on the environment. "Readily" in this statement indicates that in principle some intangible impacts can be monetised, for example using the monetary value of DALYs (Disability Adjusted Life Years), or the cost of dealing with flood-driven pollution. However, there is a degree of controversy and ethical issues may be raised if environmental impacts or a person's life or health are valued differently in different countries, which such monetisation inevitably involves.

The second common distinction is between *direct and indirect* damage (Hammond *et al.*, 2015). A direct damage is a loss caused by the immediate physical contact of flood water with humans, property and the environment, whereas indirect damages are induced by the direct impacts and may occur beyond the immediate spatial and temporal limits of the flood event. This classification can include further distinctions e.g. into primary, secondary and tertiary impacts and classes of costs. Regardless of the classification adopted, a comprehensive assessment of flood impacts requires multidisciplinary approaches that should involve economists, medics, social scientists, urban planners and, of course, engineers.

This keynote lecture is focussed on *impacts of flooding on road and rail transport systems*. Understanding of this category of flood impacts requires knowledge in several areas of Civil Engineering. For example, structural engineering and geotechnical engineering deal with evaluating local scour depth at bridge piers and other impacts of debris accumulation on bridges during flooding (Ebrahimi *et al.*,

2020; 2018). Roads and railway engineering is needed for the design and maintenance of transport systems with regard to drainage and in the context of traffic simulation during flooding (Evans *et al.*, 2020). Hydraulic engineering is required for flood modelling (Leandro *et al.*, 2011; Martins *et al.*, 2017).

Unlike some other areas of flood risk management, impacts of flooding on road and rail transport systems have not been studied in much detail in the past. There is only a relatively limited literature on this topic, despite the unfortunate fact the proportion of fatalities in vehicles is very high – especially in cities where most of the casualties are caused by people trying to drive through flood water. In addition, flood impacts on transport systems can also have significant cascading effects on economy, other services (e.g. water distribution, solid waste collection, energy supply and communication systems) and the environment (e.g. increased air pollution and oil spills from overturned of flooded vehicles).

There is a number of research questions of relevance in this context that will be addressed in this paper. What happens to traffic when drivers need to avoid the flooded streets? How to simulate the interactions between two highly dynamic phenomena – flooding and traffic? How to assess impacts? How mitigation and adaptation measures could alleviate direct impacts and cascading effects? How to communicate early warnings effectively? How to optimally make and communicate decisions during emergency in real time? How does functioning of drainage and transportation also interact with other critical infrastructure systems? What specific problems are involved in managing flood risk in underground railway tunnels and stations?

The methodologies and results presented in this keynote lecture are drawn from the outcomes of EU-funded research consortia that the Exeter Centre for Water Systems staff have been involved with over the past ten years – CORFU (Djordjević *et al.*, 2011; Hammond *et al.*, 2018) PEARL (Pyatkova, Chen, Djordjević *et al.*, 2019; Vojinovic *et al.*, 2014) and RESCCUE (Evans *et al.*, 2020; Velasco *et al.*, 2020). Our ongoing work and plans for future will also be indicated.

2. METHODOLOGY

The methodology for assessing impacts of flooding on road transport described here is completely novel. It has been developed gradually by enhancing and merging diverse tools for flood modelling, traffic simulation and data processing. So far it has been tested on seven case studies (Barcelona, Beijing, Bristol, Exeter, Marbella, Napier and St Maarten) focussing on different aspects of this problem and adjusting to data availability in study areas, in some cases covering the entire city and through traffic and in other cases looking only at a bounded part of the city. The evolution and some details of this methodology are described in this section.

2.1. Integration of flood modelling and traffic simulation

The key idea in integrating flooding and traffic simulation is that the former sets the conditions for the latter, by distinguishing between three cases for each road section illustrated in Figure 1 (Pyatkova, Chen, Djordjević *et al.*, 2019): (i) where flood depths are high (in excess of 30 cm), the road is closed and there can be no traffic, or (ii) at relatively small flood depths between 10 cm and 30 cm the road remains open but actual driving speeds are reduced proportionally to water depths, or (iii) in roads not affected at all by flooding the traffic is normal as in dry weather.

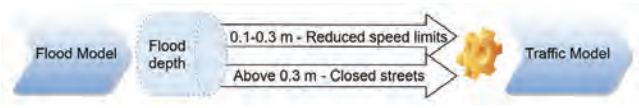


Figure 1. The general outline of the flood and traffic integration tool

For such set conditions, the road network is updated to exclude flooded roads and reduce vehicle speed in some roads. When these conditions are set using a single flood map, for example maximum flood depths during an event, we are talking about *static integration* of flooding and traffic simulation. An example of that is shown in Figure 2 with a part of simulated road network in the city of Marbella, Spain (Pyatkova, Chen, Butler *et al.*, 2019). The upper image (a) is a snapshot of traffic with all roads open and a smooth flow of vehicles, and the lower image (b) shows a static flood map overlain on a street map and a snapshot of traffic simulation indicating congestion around flooded areas. In these images, green blobs are buses and red and yellow blobs are private cars and taxis, respectively.

If flood event is described as a series of flood maps with varying water depths as the event unfolds, then some roads may get opened and closed (or have different degrees of speed reduction) during the event, which is termed *dynamic integration*. This is obviously a much more realistic approach, however also computationally more demanding, which will be discussed bit later.

2.2. Micro simulation of traffic and creation of individual journeys

The tool used in this research is the SUMO (Simulation of Urban Mobility) model developed by the German Aerospace Center, initially released in 2001 and originally dedicated for air traffic control. This open source, portable

Figure 2. Free flowing traffic and the congested traffic after closure of flooded roads

package has meanwhile found applications in all sorts of traffic in large networks and there is an established community of users and developers. In contrast to *mezzo simulation* of traffic that simulates traffic flows, what SUMO does is *micro simulation* of traffic, which tracks individual vehicle journeys. (An even more sophisticated approach is *nano simulation*, which is beyond the scope of the analysis presented here.)

In order to be able to simulate movement of individual vehicles, their *journeys* need to be created first for a given road network. This can be done with different degrees of complexity, which to a large extent depend on the availability and quality of traffic counting data and possibilities for field data based model calibration. In any case, even with the most advanced traffic monitoring system, creation of a large number of individual journeys (possibly millions a day in a big city) can never be as realistic as flood simulation using a high resolution fully coupled 1D/2D flood model driven by spatially varying rainfall from weather radar. Instead, creation of journeys should be seen as creation of *plausible scenarios* that could be meaningfully used to analyse and quantify potential impacts.

One simple approach to generating individual vehicle journeys is to identify residential parts of the city and the zones where people work (offices, industrial facilities, hospitals, educational settings) and then employ a "gravity-type" model to generate scenarios where people drive from residential areas to work during morning hours and in the opposite direction in late afternoon. Total numbers of such generated journeys could be controlled by observed data on actual traffic during the day. Within this approach, bus routes can be explicitly defined based on their real schedules.

A much more advanced way of creating individual journeys is to create the so called *Random Route Catalogue* with a rather large number of journeys and then *sample* individual journeys from that catalogue such that overall statistically they best match the actual traffic data (Evans *et al.*, 2020). This approach is schematically presented in Figure 3, which shows different types of data provided to the originally written Python script that merges different types of data and generates the Catalogue, from which then another Pyhthon scirpt samples the routes and feeds them to SUMO for traffic simulation. It should be noted that a journey is defined by start and finish points and the path between them that requires shortest time to get from start point to finish point using the roads currently open for traffic depending on flooding situation.

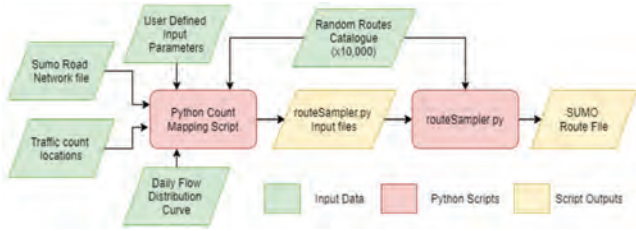


Figure 3. Flow chart describing generation of individual vehicle journeys by sampling from Random Route Catalogue

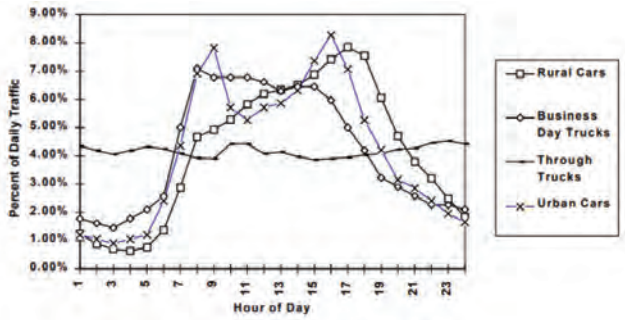
A typical result of generation of journeys is given in Figure 4, where percentages of four categories of vehicles (hourly, over 24 hours) are presented. Also in this image a small extract from traffic data is shown in a table with street name, information about location and lanes and (in the right most column) the daily count of vehicles passing through that location.

When a road on a journey changes its status, all affected paths between vehicle's current position and the finish point need to be recalculated, and that is the most challenging element of this procedure because it can get very computationally demanding. Another major problem is a gridlock scenario during periods of time when extensive and prolonged flooding completely blocks a part of the network and vehicles have no exit from it.

2.3. Input data for traffic simulation during flooding

The next important step in this methodology is setting up of input data for traffic simulation for flooding scenario. The flow chart in Figure 5 shows how the conversion process is carried out for flooding impact assessment, where the green boxes on the left half of this flow chart depict the normal "Dry Weather" data conversion process and the red boxes on the right side show the extra steps that the Phyton conversion script carries out for the "Flooded Scenario" traffic model data (Evans *et al.*, 2020).

The first step in Figure 5 is the conversion of OpenStreetMap data to the SUMO network format. At initial stages of this research, that step was made using the



	A	B	C	D	E	F	G	H	I	J	K
1	Count	Road	Loca	Northing	Easting	Direct	ADT	% Heavy Vel	La	Count St	Count D
2	1629	AHURIRI BYPASS SH50	220	5622709.78	1935924.38	Both lane	4921	25	Lates Count		23/11/2019
3	1857	ALEXANDER AVENUE	200	5618687.57	1934826.76	Both lane	1270	4	Lates Count		21/06/2018
4	1859	ALEXANDER AVENUE	480	5618565.83	1934576.98	Both lane	967	4	Lates Count		21/06/2018
5	1946	ALLEN BERRY AVENUE	260			Both lane	893	2	Lates Count		26/07/2019
6	1255	ALPERS TERRACE	397	5620738.63	1934931.75	Both lane	2242	6	Lates Count		04/09/2019
7	629	ALPERS TERRACE	293	5620678.46	1935011.26	Both lane	2541		Lates Count		15/09/2010

Figure 4. Example of journeys created by sampling from Random Route Catalogue and extract from street data with daily count of vehicles passing through different locations used in creating the Random Route Catalogue

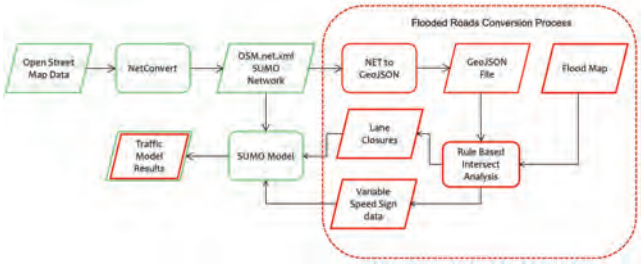


Figure 5. Conversion process for traffic simulation in flooding conditions

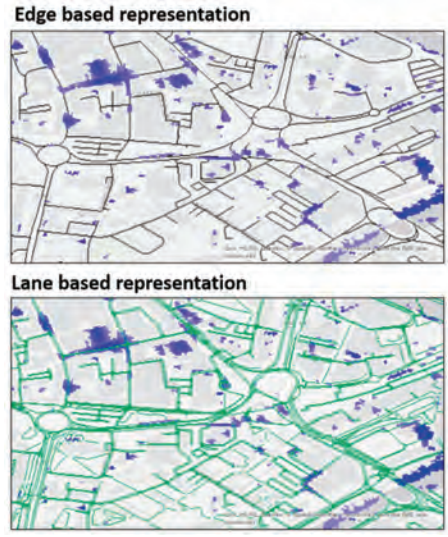


Figure 6. Edge-based and Lane Based representations of road networks

"Edge Based" conversion, which treats a street network as a system of links and nodes ("edges"), essentially similar to a typical representation of a pipe network, assuming bi-directional traffic flow, but not fully implementing the information about the number of lanes. Further improvement of that approach was introduced by Evans *et al.* (2020) by developing the "Lane Based" conversion, which employs fully the lane information, thus enabling modelling of partially flooded streets, for example when some lanes experience a flood depth different from flood depths in some other lanes on the same road. In the example shown in Figure 6, the Edge Based street data conversion resulted in 5,408 links with 10,816 Derived Lanes (twice as many, due to possibility of two-way traffic), whereas the Lane Based conversion resulted in 51,489 links and as many lanes, in other words one order of magnitude much more details. Importantly, this improvement also keeps the flow directions in each lane the same as described in the OpenStreetMap data.

3. RESULTS

In the first instance, comparison between traffic simulation during dry weather and during flooding scenarios enables quantification of local effects and global impacts across the study area. That can include a number of results directly from SUMO simulation, as well as further results derived by merging with additional types of information, e.g. types of cars, air quality conditions in the area etc. Some of these results are shown in this section.

3.1. Differences between traffic in flooding and no-flooding scenarios

By counting the total number of vehicles in traffic at any point in time during SUMO simulations, one can calculate how many more vehicles are taking longer journeys during flooding, possibly also distinguishing between different types of vehicles, e.g. buses, private cars (rural and urban), taxis, local ("business day") tracks, through trucks (passing through the city) etc. Figure 7 shows one such comparison for a flood event caused by intense rainfall between 7:00 and 7:30 that resulted in closure of some roads

starting at 7:00. The ten solid lines in this diagram show numbers of vehicles in traffic for ten different scenarios of randomised journeys during flooding; the black dashed line shows the average number of vehicles in traffic during dry weather. The increase in number of cars in traffic is up to 100% compared to no flooding scenario, and it varies during the event. It can be noted that – before 7:00 and after 10:45 – apart from "Flood Scenario 3" (green line) that took bit longer to clear the flooding-caused delays, all other flood scenarios are fairly similar to the dry weather scenario. It is also interesting to note that "Flood Scenario 2" (orange line) does not recover as quickly as other scenarios during the strongest impact of flooding, but once flooding is over it recovers (i.e. gets close to dry weather traffic) faster than four other scenarios.

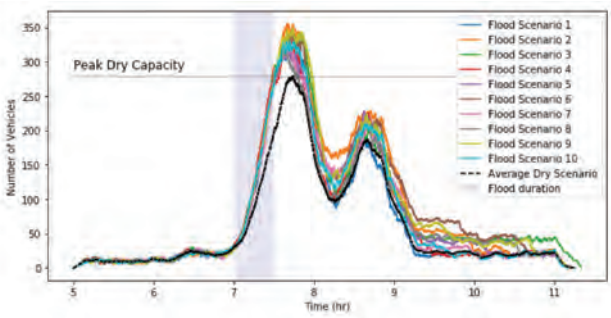


Figure 7. Number of cars in traffic for different flooding or no flooding scenarios

This example highlights the differences in randomised traffic flows. An analogy with modelling in hydraulic engineering could be that randomised inputs to traffic simulation here correspond to sampling of computational parameters (e.g. roughness coefficients) in computational hydraulics applications. Both approaches allow for assessment of uncertainties in input parameters.

3.2. Increase in fuel consumption, CO_2 emissions and air pollution

Post processing of traffic simulation results can include the assessment of increase in fuel consumption and consequent CO_2 emissions and air pollution. For such an assessment, information is needed (or assumptions have to be made) about type and age of vehicles in traffic and their average emissions of CO_2 or PMx. Figure 8 shows one such analysis, which indicates percentage increase relative to dry weather traffic with regard to the number of re-routed vehicles, time delays, increase in distance driven, fuel consumption, and CO_2 emission, all on an hourly basis and for the whole event. Clearly, this kind of analysis is more relevant for cities like Belgrade or Beijing that anyway have problems with traffic jams and air pollution.

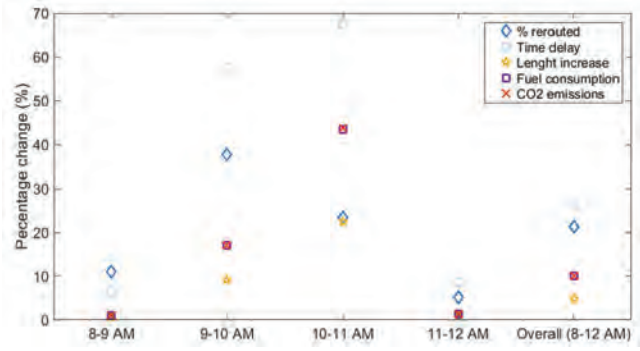


Figure 8. Increase in fuel consumption and CO2 emissions

3.3. Evacuation following a tsunami warning

Coastal cities at risk of tsunami usually have warning systems and evacuation plans. In order to analyse if evacuation is more efficient if it is done on foot or by car, the methodology described in this paper was used in a city in New Zealand for SUMO simulation of two scenarios: (i) movement of pedestrians representing evacuation on foot, and (ii) car traffic congested by a sudden increase in journeys after tsunami warning is issued. In both cases the assumption was that people are aware of the two safe zones that could not be reached by the tsunami wave and that those areas would be the finish points of their journeys, whether they are walking or driving. The results (not shown here) indicate that in this particular case evacuation on foot is lot more efficient and therefore could save lives. This examples illustrate the value of the ability to simulate both pedestrian and vehicle movement in emergency situations, which can ensure best possible preparedness for natural disasters.

4. OTHER IMPACTS OF FLOODING ON TRANSPORT SYSTEMS

4.1. Management of solid waste containers at risk of displacement by flooding

A very detailed study of potential impact of flooding on solid waste containers was conducted in Barcelona (Martínez Gomariz *et al.*, 2020). It started from the database of all solid waste containers in the city, classified into five categories of material collected (glass, paper and cardboard, packaging, organic waste and general waste) and with regard to their size (1.8–3.0 cubic metres). Flood simulations were ran for a range of return periods, which provided flow depths and flow velocities across the city. Those variables are then used for detailed analysis of hydrodynamic forces on each container, taking into account the density of material and varying the amount of stuff in the containers (empty, half-full, full). Based on the criteria for sliding and toppling, the containers at highest risk of being moved during flooding have been identified, which

enabled the city managers to optimally target any investment in fixing the containers. This analysis also enables the assessment of the cascading effect that waste containers floating in flood waters may present as an additional road blockage impacting on traffic in the city during flooding.

4.2. Flooding of underground train stations and tunnels

Recent extreme weather events caused flooding of underground train stations in a number of cities, including New York City (Hersher, 2021). The most dramatic one was flooding of underground trains in Zhengzhou in July 2021, which killed a number of people and left many stranded in underground trains for hours with water up to their necks. These tragic events raised questions about suitability of plans to build new underground train lines in some areas prone to flooding, and led to an increased interest in better mitigation and adaptation to climate change in cities that experience such problems. Comprehensive reviews of flood risk are being updated and a range of structural and non-structural measures are under consideration, all aiming to enable city managers to make informed decisions regarding future investments in both new and existing systems. Those reviews need to identify *most critical causes* of flooding, which can be very different from one place to another. For example, in London where water distribution pipes in general are relatively old, pipe bursts are found to be the main potential cause of flooding of underground stations. Obviously, tunnels and stations that are low laying and close to a river or sea coast are at higher risk of leakage from water bodies and may require regular pumping regardless of any surface flooding.

A number of very interesting lines of research is under way or planned for future, both aiming to provide better preparedness for extreme events and in terms of better handling of emergency situations in real time (Forero-Ortiz *et al.*, 2020). One of those is the development of tools for simulation of *ventilation* of underground train systems. This is a very complex flow phenomenon, because air is as a compressible fluid, and boundary conditions for such a simulation (at all entrances and vent openings at the surface) are numerous and their characteristics are not easily described in numerical models. Raising of entrances above possible flood levels reduces ease of movement for pedestrians coming in and out of stations, which is more of a problem in old city centres with constrained space and less so in new areas.

Ventilation ducts are often a conduit for flood waters to enter underground space, therefore those need to be positioned above possible flood levels. Of course, ventilation is necessary for the comfort of passengers, but it also needs to be flood safe. Typically, ventilation systems are better in recently built underground train networks that are adequately designed for large number of users (e.g. in Seoul), than in old systems that have parts originally designed for much smaller traffic flows (e.g. in London). Development of numerical models for simulation of air movement (transport of air across the network of tunnels) and ventilation (bi-directional exchange of air between the underground system and the atmosphere) is an exciting and open research avenue in Fluid Mechanics.

Aside from its relevance for management of underground rail transport systems discussed here, a corresponding problem in need of ventilation/air flow model is management of hydrogen sulphide and other toxic gasses in sewer networks. Sewer ventilation can lead to odour nuisance, while the lack of it exacerbates corrosion of pipes (Adkins *et al.*, 2018; Djordjević, 2021).

Another interesting area of research is designing optimal and clear *signage* for emergency situations in underground train stations, aiming to enable precise communication and reduce likelihood of panic, which is linked to behavioural psychology.

5. CONCLUDING REMARKS

Flood impacts on transportation systems can be significant. This paper presented a novel methodology for dynamic integration of various types of simulation models – flood modelling, micro simulation of traffic, and data acquisition and processing techniques – that offers complex and comprehensive analysis of various scenarios. This in turn can lead to better adaptation to climate change caused sea level rise and increased frequency and intensity of heavy rainfall events and subsequent flooding. The ultimate goal is to save lives and minimise negative impacts on the environment, economy and other infrastructure systems.

Scientific progress in this domain has been substantial, but there is a lot of space for further research in enhancing described methodologies and combining them with achievements in other disciplines. To name a few: stability criteria for flooded vehicles (Martínez-Gomariz *et al.*, 2018) could be used in addition to the 30 cm depth threshold for closure of roads, traffic light system could be optimised for flooding conditions, and warnings and information provided to drivers in real time during flood events can be improved.

Interesting analogies between various problems in assessment of flood impacts on transport systems and corresponding issues in computational hydraulics applications in simulation of pipe networks were drawn, illustrating the need for exchange of ideas and methods between different research topics.

Acknowledgments

The methodologies and results presented in this paper are outcomes of research funded by European Commis-

sion through Framework Programme 7 projects CORFU (Grant Agreement No 244047) and PEARL (Grant Agreement No 603663) and Horizon 2020 project RESCCUE (Grant Agreement No 70017). I am deeply indebted to my wonderful colleagues with whom I had the pleasure of working on these and other projects, in particular Dr Katya Pyatkova, Dr Barry Evans and Professor Albert Chen from University of Exeter, Dr Eduardo Martínez-Gomariz from Technical University of Catalonia and Professor Zoran Vojinović from IHE-Delft. They gave invaluable input to the results presented here and are key contributors to this paper.

REFERENCES

Adkins, S., Djordjević, S. and Savić, D. A. (2018). Wastewater system ventilation–a friend or adversary? In *International conference on urban drainage modelling* (pp. 712–716).

Chen, A. S., Hammond, M. J., Djordjević, S., Butler, D., Khan, D. M. and Veerbeek, W. (2016). From hazard to impact: Flood damage assessment tools for mega cities. *Natural Hazards*, *82*(2), 857–890.

Chen, Y., Fingleton, B., Pryce, G., Chen, A. S. and Djordjević, S. (2013). Implications of rising flood-risk for employment location: a gmm spatial model with agglomeration and endogenous house price effects. *Journal of Property Research*, *30*(4), 298–323.

Djordjević, S. (2021). Modelling of phenomena related to air flow in sewer systems, keynote lecture. In *19th congress of serbian association for hydraulic research*. Belgrade.

Djordjević, S., Butler, D., Gourbesville, P., Mark, O. and Pasche, E. (2011). New policies to deal with climate change and other drivers impacting on resilience to flooding in urban areas: the corfu approach. *Environmental Science & Policy*, *14*(7), 864–873.

Ebrahimi, M., Djordjević, S., Panici, D., Tabor, G. and Kripakaran, P. (2020). A method for evaluating local scour depth at bridge piers due to debris accumulation. In *Proceedings of the institution of civil engineers-bridge engineering* (Vol. 173, pp. 86–99).

Ebrahimi, M., Kripakaran, P., Prodanović, D. M., Kahraman, R., Riella, M., Tabor, G., Arthur, S. and Djordjević, S. (2018). Experimental study on scour at a sharp-nose bridge pier with debris blockage. *Journal of Hydraulic Engineering*, *144*(12), 04018071.

Evans, B., Chen, A. S., Djordjević, S., Webber, J., Gómez, A. G. and Stevens, J. (2020). Investigating the effects of pluvial flooding and climate change on traffic flows in barcelona and bristol. *Sustainability*, *12*(6), 2330.

Forero-Ortiz, E., Martínez-Gomariz, E., Cañas Porcuna, M., Locatelli, L. and Russo, B. (2020). Flood risk assessment in an underground railway system under the impact of climate change—a case study of the barcelona metro. *Sustainability*, *12*(13), 5291.

Hammond, M., Chen, A. S., Batica, J., Butler, D., Djordjević, S., Gourbesville, P., Manojlović, N., Mark, O. and Veerbeek, W. (2018). A new flood risk assessment framework for evaluating the effectiveness of policies to improve urban flood resilience. *Urban Water Journal*, *15*(5), 427–436.

Hammond, M. J., Chen, A. S., Djordjević, S., Butler, D. and Mark, O. (2015). Urban flood impact assessment: A state-of-the-art review. *Urban Water Journal*, *12*(1), 14–29.

Hersher, R. (2021). *Nyc's subway flooding isn't a fluke. it's the reality for cities in a warming world.* https://www.npr.org. U.S. National Public Radio.

Hilly, G., Vojinovic, Z., Weesakul, S., Sanchez, A., Hoang, D. N., Djordjevic, S., Chen, A. S. and Evans, B. (2018). Methodological framework for analysing cascading effects from flood events: The case of sukhumvit area, bangkok, thailand. *Water*, *10*(1), 1–26.

Khan, D., Veerbeek, W., Chen, A., Hammond, M., Islam, F., Pervin, I., Djordjević, S., Butler, D. et al. (2018). Back to the future: assessing the damage of 2004 dhaka flood in the 2050 urban environment. *Journal of Flood Risk Management*, *11*(s1).

Leandro, J., Djordjević, S., Chen, A., Savić, D. and Stanić, M. (2011). Calibration of a 1d/1d urban flood model using 1d/2d model results in the absence of field data. *Water Science and Technology*, *64*(5), 1016–1024.

Martínez-Gomariz, E., Gómez, M., Russo, B. and Djordjević, S. (2018). Stability criteria for flooded vehicles: A state-of-the-art review. *Journal of Flood Risk Management*, *11*, S817–S826.

Martínez Gomariz, E., Russo, B., Gómez Valentín, M. and Plumed Gómez, Á. (2020). An approach to the modelling of stability of waste containers during urban flooding. *Journal of flood risk management*, *13*(S1), 1–18.

Martins, R., Leandro, J., Chen, A. S. and Djordjević, S. (2017). A comparison of three dual drainage models: shallow water vs local inertial vs diffusive wave. *Journal of Hydroinformatics*, *19*(3), 331–348.

Pyatkova, K., Chen, A. S., Butler, D., Vojinović, Z. and Djordjević, S. (2019). Assessing the knock-on effects of flooding on road transportation. *Journal of environmental management*, *244*, 48–60.

Pyatkova, K., Chen, A. S., Djordjević, S., Butler, D., Vojinović, Z., Abebe, Y. A. and Hammond, M. (2019). Flood impacts on road transportation using microscopic traffic modelling techniques. In *Simulating urban traffic scenarios* (pp. 115–126). Springer.

Russo, B., Sunyer, D., Velasco, M. and Djordjević, S. (2015). Analysis of extreme flooding events through a calibrated 1d/2d coupled model: the case of barcelona (spain). *Journal of Hydroinformatics*, *17*(3), 473–491.

Stevens, J., Henderson, R., Webber, J., Evans, B., Chen, A., Djordjević, S., Sanchez-Munoz, D. and Dominguez-Garcia, J. (2020). Interlinking bristol based models to build resilience to climate change. *Sustainability*, *12*(8), 3233.

Velasco, M., Russo, B., Monjo, R., Paradinas, C., Djordjević, S., Evans, B., Martínez-Gomariz, E., Guerrero-Hidalga, M., Cardoso, M. A., Brito, R. S. et al. (2020). Increased urban resilience to climate change—key outputs from the resccue project. *Sustainability*, *12*(23), 9881.

Vojinovic, Z., Sanchez, a., Alves, A., Abebe, Y., Medina, N., Mark, O., Domingo, N. S., Djordjevic, S., Chen, A., Pyatkova, K. et al. (2014). Adapting to climate change through holistic risk assessment – the pearl approach. In *11th international conference on hydroinformatics hic 2014*. New York City.

Wang, Y., Chen, A. S., Fu, G., Djordjević, S., Zhang, C. and Savić, D. A. (2018). An integrated framework for high-resolution urban flood modelling considering multiple information sources and urban features. *Environmental modelling & software*, *107*, 85–95.

Branislav Bajat

Branislav Bajat is a Full Professor at the Faculty of Civil Engineering, University of Belgrade (FCEUB) since 2016, President of the Council of the FCEUB, member of the Professional board of natural and mathematical sciences of the University of Belgrade, and member of the Scientific board for energy, mining and energy efficiency of the Ministry of Education, Science and Technological Development of the Republic of Serbia.

Contact information: e-mail: bajat@grf.bg.ac.rs

https://orcid.org/0000-0002-4274-2534

Prof. Bajat graduated at the Faculty of Civil Engineering, University of Belgrade at the Department of Geodesy in 1989, where he received magister's degree in 1996 and PhD in 2004. He continued his professional education during his research studies in 2002 and postdoctoral studies in 2004 at the Technische Universität Hamburg (TUHH). He started his professional career as the Associate Engineer in 1989 at the Institute of Geodesy of the Faculty of Civil Engineering, where he was elected as a Trainee Assistant in 1990, Teaching Assistant (1997), Assistant Professor (2005) and Associate Professor for the Engineering Geodesy and Modeling and Management in Geodesy (2011). He was elected as Full Professor in 2016. In the period 2012-2015 he was the director of the Institute of Geodesy and Geoinformatics of the Faculty of Civil Engineering, and the Head of the Department of Geodesy and Geoinformatics during 2015-2018. Since 2020, he has been the President of the Council of the Faculty of Civil Engineering.

Besides the Faculty of Civil Engineering, he is also engaged as a Professor at the Faculty of Geography, University of Belgrade. He has also lectured at the Department of Geoinformatics, Palacký University in the Czech Republic, and at the Department of Applied Geoinformatics and Cartography, Charles University in Prague as a Visiting Professor. He actively participated in the development of doctoral studies in the study program Geodesy and Geoinformatics, where he was a mentor on four defended doctorates.

Prof. Bajat is a member of the editorial boards of the journal Geodetska Služba, Građevinski kalendar and the international journals SPATIUM and Geodetski vestnik, as well as a member of the Publishing Council of the Institute of Architecture and Urbanism of Serbia. He is a reviewer in more than 30 international scientific journals.

His main scientific research areas are the application of new spatial and spatio-temporal statistical methods in geosciences and environmental sciences. His special field of interest is working on projects related to the application of geoinformation technologies in various engineering fields. So far, he has published more than 90 papers in international and national scientific journals, monographs and conferences.

From topographic surveying to geomatics – geospatial technology trends

Branislav Bajat

University of Belgrade, Faculty of Civil Engineering, Department of Geodesy and Geoinformatics

Summary

In recent years the development of information technology, associated with the introduction of satellite navigation and remote sensing techniques, has resulted in replacing analogue photogrammetric plotters, classical cartography, optical theodolites and levelling instruments, tachymetry, plane tables, steel chains and cloth tapes, and other "old fashioned" equipment with new methods and technologies in the capture, storage and management of spatial data. As a consequence, the geodetic profession has had that opportunity to expand its competence and services. This paper presents the latest achievements in geodesy and geoinformatics with regard to technological, methodological and application issues at the global level, as well as examples of accomplishments in this field related to the Serbian case.

1. INTRODUCTION

Since the beginning of the 21st century, significant technological developments have taken place in the world, especially in spatial data acquisition techniques (primarily satellite Earth observation technologies), which provide us with large and complex high spatial and temporal resolution data sets, in which traditional data processing applications are not applicable.

Geospatial data is the main driver of today's information society and it is an integral part of everyday life for modern people. It represents a key element in the following processes: decision making and resource management, data exchange, communication, etc. Sustainable development cannot be ensured without the use of this data. In this sense, geospatial data is used in an increasing number of scientific disciplines that aim to aid the understanding of spatial phenomena, but also in economics to improve the production process and/or to support decision making. The question is no longer what the potential of geospatial data is, but what the best way is to consolidate, analyze and distribute this data to end users in the form of new knowledge.

1.1. Geodesy, surveying engineering, geomatics and geoinformatics, are they the same professions?

Geodesy, like many scientific fields, is technology driven science. However, at the same time, it has developed as an engineering discipline because of its practical applications. The last decades of the 20th century brought great changes in the field of technologies that deal with the acquisition of spatial data. The introduction of Global Navigation Satellite Systems (GNSS), advances in the development of space technologies for Earth Observation (EO), image processing, the Internet of things (IoT), new mobile devices, etc. left a big mark in the disciplines that are considered as classical geodetic and surveying specialties, like cartography, photogrammetry, state surveying, cadaster, engineering, physical geodesy, etc. The focus in the development of these disciplines has shifted from data measuring and processing methods to new technologies, data distribution methods and visualization.

This inevitably reflected on standard terminology in this area. Geodesy in Serbian literature (inherited from German) is divided into "higher geodesy" (höhere Geodäsie), which is concerned with measuring Earth on the global scale, and "surveying" which is concerned with measuring specific parts or regions of Earth.

In the last decade of the 20th century, the terms geodesy and surveying were increasingly replaced by the term geomatics, which is a modern discipline that integrates the acquisition, modeling, analysis, and management of georeferenced data (Basic *et al.*, 2003). According to ISO standard 19122 "geomatics is a discipline concerned with the collection, distribution, storage, analysis, processing, and presentation of geographic data or geographic information".

Location remains the main factor in geomatics, which is used to join a wide range of data for spatial analysis and visualization. Geomatic engineers apply geometric

principles to spatial information, managing different spatial data infrastructures, at the local, regional, or global level. The extensive accessibility and the use of modern technologies, such as global navigation satellite systems (GNSS), remote sensing, and geographical information systems (GIS), increase the precision and efficiency of the profession previously recognized as geodetic engineering.

Geomatics is an umbrella term, under which traditional surveying and more modern technology-induced sub-disciplinary groups like remote sensing, GIS, global navigation satellite systems fall. Geomatics is also noticed as a more appropriate term for describing modern technology-based surveying.

Unlike the term geomatics, which in non-European countries (America, Canada, Australia and others) has completely replaced the term geodetical engineering, the term geoinformatics appeared in Europe at the beginning of the 21st century. Even though these two terms are often mixed up, there are considerable differences between them.

According to the Encyclopedia of Information Science and Technology, the term geoinformatics refers to academic disciplines or jobs that deal with geographic data, whose end result is an understanding and interpretation of human interaction with the environment. Linguistically, the term geoinformatics is derived from two words; "geos", which refers to the term "Earth" and "informatics" relating to a wider range of activities (e.g., informatics, software engineering and computer graphics). Although there is no official definition as a comprehensive explanation of this term, geoinformatics can be considered as "a discipline whose focus is the theory of modeling storage, management and processing of geospatial data, as well as the development of geoinformation systems and necessary information and communication technologies" (Hake *et al.*, 2002), or "science that develops and uses the infrastructure of information sciences in solving tasks in the field of geosciences and related engineering disciplines". Following from the above-mentioned definitions, geoinformatics can be defined as a relatively broad term which includes various technologies, approaches and processes, whose ultimate goal is the interpretation of geoinformation related to the spatial environment, in order to make comprehensive decisions. Geoinformation represents various types of data, obtained from GIS, remote sensing and socio-economic studies, from which maps or other forms of reports can be generated that enable better interpretation, management and decision-making on human activities related to the environment. Special roles of geoinformatics are the analysis and synthesis of geospatial data with the use of GIS technology, which facilitate the high-quality cognition and visualization of geospatial data using modern technologies like 3D surface models, scene animation, and virtual reality (Šíma, 2007). All these roles covered by geoinformatics require additional knowledge that experts have to possess, especially in the field of informatics, as well as the areas of modeling and data analysis (deep learning, data driven modeling, etc.). Areas in which the fields of geoinformatics and geomatics overlap and are distinct from each other with regard to knowledge and skills in informatics are shown in Figure 1.

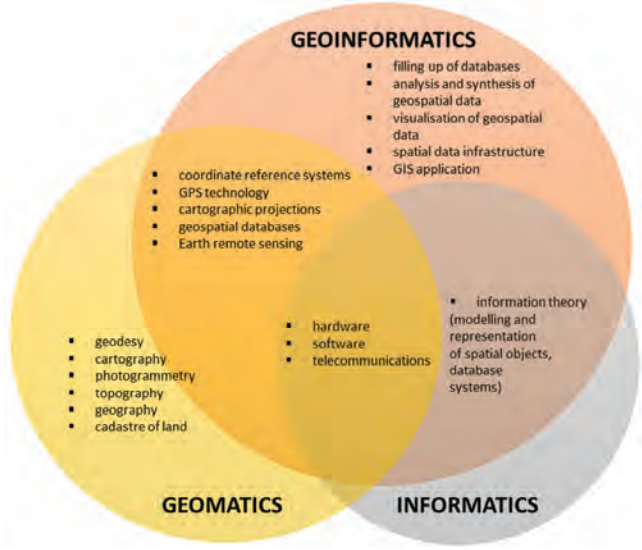


Figure 1. Relation between geomatics and Geoinformatics (adopted from Šíma, 2007)

2. GLOBAL NAVIGATION SATELLITE SYSTEMS (GNSS)

One of the greatest technological achievements from the end of the twentieth century is the development of global positioning systems (GPS). Although, like most innovations from the 20th century it was primarily developed for military purposes, the possibility of GPS use in the civilian sector, i.e. its commercial use, was quickly noticed. Nowadays, GPS receivers are an integral part of various instruments that basically contain sensors for other purposes, enabling recording of their geolocation. In geodesy and land surveying they are an indispensable tool for precise 3D positioning.

The term Global navigation satellite system (GNSS) refers to various satellite constellations that provide positioning, navigation, and timing services on a global or regional level.

The *Navigation Satellite Timing and Ranging System* (NAVSTAR) system (commonly known as GPS) was created by combining the TIMATION and 621B programs, motivated by the needs of the US military for precise positioning, navigation and time transfer. The key moment was the decision of the Department of Defense to direct financial resources to the development of a new satellite navigation system instead of modernizing the TRANSIT system. The basic architecture was approved in 1973, and in 1978 the first GPS satellites were launched. The system

consists of 24 active satellites and it was declared fully operational in 1995.

The *Globalnaya Navigatsionnaya Sputnikovaya Sistema* (GLONASS) was developed by the Soviet Union as an experimental military communications system during the 1970s. When the Cold War ended, the Soviet Union recognized that GLONASS had commercial applications, through the system's ability to transmit weather broadcasts, communications, navigation and reconnaissance data. The first GLONASS satellite was launched in 1982 and the system was declared fully operational in 1993. After a period in which GLONASS's performance declined, Russia committed to bringing the system up to the required minimum of 18 active satellites. Currently, like NAVSTAR, GLONASS constellation consists of 24 satellites.

The European Satellite Navigation System *Galileo* started its initial services on December 15, 2016. Galileo is based on a constellation of 30 satellites and two ground control centers (Fucino in Italy and Oberpfaffenhofen in Germany), providing information concerning the positioning of users in many sectors such as transport, social services, the justice system and customs services, public works, search and rescue systems, and leisure.

BeiDou, or BDS, was originally established as a regional GNSS called Compass owned and operated by the People's Republic of China. China is currently expanding the system to provide global coverage, with 35 satellites by 2020.

In addition to systems that are operational at the global level, it is also worth mentioning regional navigation systems, with limited areal coverage, such as: a) Indian Regional Navigation Satellite System (IRNSS)/Navigation Indian Constellation (NavIC), a regional navigation system owned and operated by India. IRNSS is an autonomous system designed to cover the Indian region and 1500 km around the Indian mainland. The system consists of 7 satellites and was declared operational in 2018. In 2016, India renamed IRNSS as the Navigation Indian Constellation; and b) Quasi-Zenith Satellite System (QZSS) maintained by Japan. QZSS accompanies GPS to improve coverage in East Asia and Oceania. It consists of 4 satellites and expansion to 7 satellites is planned for its autonomous capability by 2023.

GNSS technology is also embedded in other devices which are intended for the collection of geospatial data, like LiDAR and digital photogrammetry.

3. EARTH OBSERVATION

The era of satellite observation of the Earth began on October 4, 1957, when the Soviet Union launched the first artificial satellite, Sputnik-1. Since then, the number of satellites in orbit has increased exponentially, so that today there are more than 2,000 active satellites that observe the Earth (Ferreira *et al.*, 2020).

The term Earth observation technology includes different remote sensing techniques and methods related to gathering data about planet Earth's physical, chemical and biological systems, usually involving satellites carrying sensor devices. Nowadays, EO serves for monitoring the status of, and changes in, the natural and man-made environment. EO technologies provide reliable and repeat-coverage datasets, which in association with other conventional methods provide a unique resource of information concerning the planet. EO technologies have increased accuracy to sub centimeter levels.

Individual surfaces reflect and absorb radiation differently in different parts of the electromagnetic (EM) spectrum, which is the basis for their mutual differentiation (e.g., determination of land cover) (Artiola *et al.*, 2004; Ma *et al.*, 2017). Earth observation satellites are equipped with various active and passive sensors. Active sensors emit radiation towards the Earth's surface and then register the return radiation that has been reflected or refracted. In contrast, passive sensors do not emit their own radiation, but register the radiation that the Earth's surface emits (its own energy) which reflects from the Sun (Andries *et al.*, 2019).

Passive optical systems register radiation within the visible (400 nm – 700 nm), near infrared (700 nm – 1000 nm), shortwave infrared (1000 nm – 2500 nm), and longwave infrared (8000 nm – 14000 nm) parts of the EM spectrum. Some satellite missions/sensors, in addition to the optical/infrared part of the EM spectrum, also register radiation in thermal-infrared (TIR) channels. The TIR spectral range implies a wavelength range between 3000 nm and 14000 nm, but sensors due to atmospheric radiation penetration most often register radiation in the ranges 3000 nm – 5000 nm and 8000 nm – 14000 nm (Jensen, 2006). Table 1 shows the most important passive optical satellite missions/sensors. The main selection criteria are the current and future availability of data, but also the importance and presence of missions/sensors in practice.

Active radar satellite systems, known as Synthetic Aperture Radar (SAR), emit radiation to the Earth's surface in the microwave part of the EM spectrum and then measure the backscatter radiation. These systems can collect data regardless of weather conditions or time of day (Liang and Wang, 2019). A number of microwave bands of the EM spectrum are in use, while their central wavelengths also vary from sensor to sensor. In general, SAR systems are designed to emit/register EM radiation in an X-channel ($\sim$ 3 cm), a C-channel ($\sim$ 6 cm), an L-channel ($\sim$ 25 cm) or a P-channel ($\sim$ 70 cm) (Jensen, 2006; Liang and Wang, 2019). SAR systems, like optical systems, have been present in orbits for several decades, with different spatial, spectral and temporal resolutions. SAR systems emit polarized and coherent EM radiation and then register the amplitude and phase of the returned radiation. Table 2 shows a list of the most important SAR systems.

Table 1. Overview of passive optical satellite missions/sensors with the most important characteristics

Satellitemission/sensor	No. of channels	SR [m]	PR [day]	Cover [km^2]	Open access	Availability
AVHRR *	4–6	1100	0.5	2399	Yes	1979-
VIIRS *	22	375–750	1	3060	Yes	2011–
MODIS *	36	250–1000	0.25-0.50	2330	Yes	2000-
SENTINEL-3 *	30	300-1000	1–2	1440	Yes	2016-
FIREBIRD *	5	42.4-356	3-4	178–211	No	2012-
LANDSAT *	4–11	15–100	16	185	Yes	1972–
ASTER *	14	15–90	16	60	Yes	2000-
SENTINEL-2	13	10–60	5	290	Yes	2015-
SPOT	4–5	2.5–20	26	120	No	1986–
RAPIDEYE	5	6.5	1-5.5	77	No	2008-2020
ZIYUAN III	4	2.1-5.8	5	50	No	2012–
PLANETSCOPE	4	3.7	1	24	No	2016–
IKONOS	5	1–4	1.5-3	11.3	No	1999-2015
QUICKBIRD	5	0.61–2.24	2.7	16.5	No	2001–
WORLDVIEW	4-17	0.31–2.40	1–4	17.6	No	2007–
SKYSAT	5	0.9–2.0	0.15	8	No	2013–
PLEIADES	5	0.5–2	4-5	20	No	2011–
GEOEYE	5	0.46-1.84	2.1-8.3	15.2	No	2008-

(SR – spatial resolution, TR – time resolution, * – possess thermal channels)

Table 2. Overview of active radar satellite missions/sensors with the most important characteristics

Satellite mission/sensor	Channel	Polarization	SR [m]	TR [day]	Cover [km^2]	Open access	Availability
ERS-1/2	C	VV	20	35	10–100	Yes	1991–2011
JERS	L	HH	20	44	175	Yes	1992–1998
Radarsat	C	HH/HV/VV **	1–100	5–10	18–500	No	1995–
TerraSAR-X i TanDEM-X	X	HH/HV/VV *	1–20	11	5–100	No	2007–
Cosmo-Skymed	X	HH/HV/VV *	1–100	0.5	10–200	No	2007–
SAOCOM	L	HH/HV/VV **	7–100	10–100	30–350	No	2015–
Sentinel-1	C	HH/HV/VV *	5–100	12	80–400	Yes	2015–
ALOS-2	L	HH/HV/VV **	3–100	14	25–350	No	2015–
ICEYE	X	VV	1-3	18–22	5–30	No	2018–
PAZ	X	HH/HV/VV **	0.25–40	11	4–200	No	2018–
Biomass	P	not yet announced					2022–

SR – spatial resolution, TR – time resolution, * – collect radiation in one or two polarizations, ** – collect radiation in all or any polarization (polarimetric)

3.1. Modern space-based geodetic observations and studies

Space-based geodetic observations can be categorized into four basic techniques: positioning, altimetry, interferometric synthetic aperture radar (InSAR), and gravity studies.

Building on this idea, scientists have developed advanced positioning techniques through Global Navigation Satellite Systems (GNSS). GNSS encompasses various satellite navigation systems, such as the United States' NAVSTAR (commonly known as GPS), Russia's GLONASS, and Europe's Galileo. Although these satellite systems were designed mainly for navigation, they were found to be very useful for precise positioning, with accuracy levels of less than a centimeter. GNSS also provides very high temporal resolution measurements (second by second, or even faster), yielding key observations of time-dependent processes in the lithosphere, atmosphere, and ionosphere.

In contrast to measuring three-dimensional (3-D) changes by means of positioning techniques, altimetry comprises only changes in surface elevation. Space-based radar and laser altimetry techniques accurately measure a satellite's height above the Earth's surface, which is then converted to the height of the surface above a reference ellipsoid.

Altimetry measurements are conducted by releasing pulses toward the Earth's surface every several milliseconds, resulting in circular ground measurements (footprints) along the satellite track. Because of the large footprint (diameter > 75 meters), altimetry measurements are suitable for measuring flat surfaces. Radar altimetry was actually designed for measuring the height of ocean water surfaces but was also found to be useful in measuring

changes in ice cap elevation and water levels in rivers and lakes.

A powerful method for detecting surface change is the InSAR technique. This method compares pixel-by-pixel SAR phase observations of the same area acquired from roughly the same location in space to produce digital elevation models (DEMs) or surface displacement maps with high spatial resolution (typically from 5 to 100 meters). Such maps, termed interferograms, are obtained from repeat orbit observations and can reach centimeter-level precision. Changes in the Earth's surface at the local level (areas less than 100 km^2) are best determined by a combination of high temporal resolution (GPS) and high spatial resolution of InSAR observations.

InSAR is very effective in detecting localized subsidence and uplift of land surfaces in response to natural or anthropogenic causes. Some of the more successful applications are monitoring: subsidence due to compaction of sediments; aquifer-system response to groundwater pumping and artificial recharge; extraction of fluids in oil fields; and excavation in mines and tunnels. InSAR and GPS are also powerful tools for studying surficial processes such as landslides and soil moisture content.

Accurate determination of the Earth's shape by measuring the geoid is another important contribution using precise tracking of altimetry and gravity mission satellites. Altimetry measurements over oceans, combined with gravity models, can yield detailed information on the bathymetry of the ocean floor. For example, they can help detect the location of unknown seamounts, which show up as localized gravity anomalies reflecting lateral mass differences between seamounts and ocean water. Sea level rise and glacial melting are of immediate interest to communities around the world, particularly in the context of global climate change. At the same time, geodetic monitoring of water level changes in wetlands, rivers, and lakes is very important for hydrological models that can serve as decision support tools for water resource managers.

In addition, satellite orbits are very sensitive to lateral variations in the Earth's gravity field. Precise measurements of satellite orbits by ranging (distance) and other technologies yield accurate determination of the geoid shape and its variations over time. The geoid is defined as the equipotential surface of Earth's gravity field that best fits the global mean sea level. It describes the mass distribution within the Earth, from which one can infer Earth's dynamic structure. The new generation of gravity satellites is also very sensitive to short-term changes in the geoid, reflecting near-surface mass redistribution such as ice melts or large-scale seasonal changes in water budgets. Geoid measurements are also very important for the adjustment of GPS-determined heights with a standard mean sea level datum.

Before the beginning of the space geodesy era, sea level changes were observed based on tide gauge records, acquired only along coastlines. Currently, based on the global coverage of altimetry satellites, sea level changes are observable over entire ocean basins. In addition, the determination of relative movements of land from the sea can be resolved by monitoring the subsidence or uplift of gauge stations by GPS measurements.

Many continental-scale atmospheric and ionospheric phenomena can also be measured through space geodesy. In addition to precise positioning, GPS can monitor changes in ionospheric total electron content (TEC) and the amount of water potentially available in the atmosphere for precipitation, because signals transmitted from the GNSS satellites are sensitive to the water content in the atmosphere and to TEC in the ionosphere. Cognizance of these variables is very useful in numerical weather prediction systems, increasing their ability to forecast heavy rain and hurricane intensity, as well as to forecast adverse space weather (Wdowinski and Eriksson, 2009).

SAR data, analyzed using Interferometric SAR (InSAR) techniques, can be used to model millimeter-to-centimeter scale deformation of the Earth's surface over regions tens to hundreds of kilometers across. These displacement fields are becoming essential guides for studies of tectonics, earthquake focal mechanisms, volcano behavior, hydrology, and public safety related to Earth hazards.

Since space geodetic measurements are present in many different fields, geodesy today connects various scientists for interdisciplinary research that helps mitigate the influence of the forces of nature on our growing population, as well as the effect of the population on Earth's surface.

During the 50-year history of the space geodetic era, the innovative development of space technologies has yielded numerous geodetic methods. Space geodesy provides observations at various spatial and temporal levels with a corresponding variety of applications at the local, continental, and global level. Acquiring continuous data in various temporal resolutions (from seconds to a timespan of several days) for any location, space geodesy techniques provide scientists with data sets of repeated observations that run over decades.

3.2. EO product CORINE Land Cover (CLC) for Serbia

Corine Land Cover (CLC) provides consistent information on land cover and land cover changes across Europe. This inventory was initiated in 1985 (reference year 1990), and it established a time series of land cover information with updates in 2000, 2006, 2012, and 2018. Serbia joined the mapping programme in 2005. CLC products are based on the photointerpretation of satellite images by national teams of participating countries – the EEA member and cooperating countries – following a standard methodology

and nomenclature with the following base parameters: 44 classes in the hierarchical three level Corine nomenclature; minimum mapping unit (MMU) for status layers is 25 hectares; minimum width of linear elements is 100 meters; minimum mapping unit (MMU) for Land Cover Changes (LCC) for the change layers is 5 hectares. For the CLC1990 and CLC2000 campaigns, Landsat imagery was used for photo interpretation. The following campaigns, CLC2006 and CLC2012, were based on IRS P6 LISS, Spot 4/5 and RapidEye data, while the latest update, CLC2018, leveraged the availability of Sentinel-2 data (Figure 2). The resulting national land cover inventories were further integrated into a seamless land cover map of Europe. Land cover and land use (LCLU) information is important not only for land change research, but also more broadly for monitoring environmental change, policy support, the creation of environmental indicators and reporting. CLC datasets provide important datasets supporting the implementation of key priority areas of Environment Action Programs of the European Union as protecting ecosystems, halting the loss of biological diversity, tracking the impacts of climate change, assessing developments in agriculture and implementing EU directives.

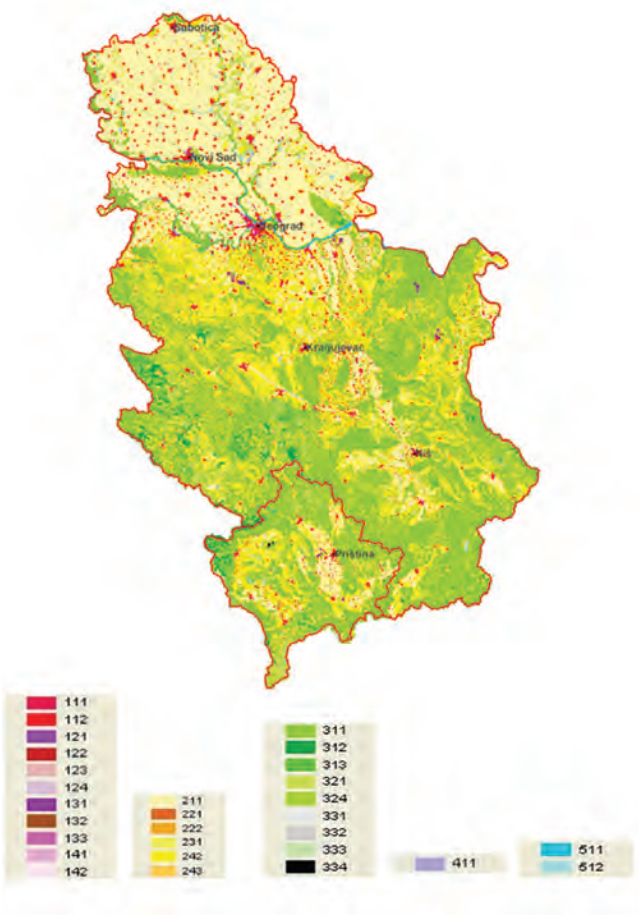


Figure 2. CORINE Land Cover for Serbia with Level 3 classes

4. LIDAR AND PHOTOGRAMMETRY

Since the beginning of the 21st century, the Light Detection and Ranging (LiDAR) technique, as an active remote sensing technique, has risen up as an efficient and reliable tool for 3D point cloud collection. LiDAR is a remote sensing technique that measures distance by illuminating and measuring reflected laser pulses from the target using optical sensors. Nowadays, LiDAR is one of the most effective and reliable means of spatial data collection. The beginning of this technique goes back to the 1980s, when NASA developed laser instruments which were the predecessors to modern LiDAR systems (Hoge *et al.*, 1980).

LiDAR includes a set of 3D points, called point clouds, which sample the surface of natural or artificial objects. A point cloud is basically a large collection of points that are placed on a three-dimensional coordinate system. Due to different platforms where LiDAR sensors are mounted, this spatial data acquisition technique is divided into the subdivisions: Airborne Laser Scanning (ALS), Terrestrial Laser Scanning (TLS) and Mobile Laser Scanning (MLS). They share many features, especially those resulting from laser-ranging technology, however they differ in terms of data capture mode, typical project size, scanning mechanism and obtainable accuracy and resolution.

ALS technology typically combines LiDAR sensors and an aircraft platform to rapidly acquire highly accurate surface information over a large area. Initially LiDAR sensors were mounted on light jet airplanes and helicopters. Great developments in automation, navigation, robotics and miniaturization have enabled the mass production of comparably low-priced unmanned aerial vehicle (UAV) drones that are capable of carrying low weight sensors in a range of several hundred grams (Figure 3a).

This has greatly reduced the cost of flying and thus highlighted ALS as a cost-effective, fast, efficient and inexpensive topographic survey technique for generating dense and accurate elevation data across landscapes and shallow water areas. A topographic survey includes a Digital surface model (DSM), Digital terrain model (DTM), Canopy height model (CHM) and intensity images. Depending on the application level, these data can be divided into: infrastructure and environmental surveys (land survey, urban planning, open-pit mining, transport, pipelines, power lines); forestry and agriculture (potential irrigation area, water embankments, forestry inventory); and natural hazards (flood risk mapping, deformation monitoring, gravity risks, glaciology hazards) (Li *et al.*, 2020).

Terrestrial laser scanners are ground-based laser scanning instruments which enable the rapid collection of three-dimensional (3D) point clouds. TLSs use ground-based remote sensing systems, which are single-handed instruments mounted on static tripods (Figure 3b), or the detached part of geodetic total station. These systems scan in all directions, including upwards. Once scans of a single

(a) (b) (c)

Figure 3. a) UAV ALS equipment (© *Bwag/Commons*) b) single-handed TLS instrument c) 3D mobile laser scanning truck

Figure 4. Point clouds of a tunnel entrance and railway line (railway Beograd-Bar)

area are complete, the tripod is moved to another location to scan from another angle or capture data from a new area.

TLS is used in engineering geodesy, particularly in dams, monitoring tunnels (Figure 4), and for surveying areas susceptible to land erosion or landslides. The ability of terrestrial laser scanners to rapidly capture the complex surfaces of historical buildings and sculptures makes it a preferred technology in cultural heritage and archaeological studies (Figure 5). Terrestrial laser scanners are also used for gathering data from engineering structures, building interiors, and areas of dense vegetation.

MLS is a land surveying method that uses laser systems mounted on moving vehicles (Figure 3c). MLS provides stop-and-go and on-the-fly modes of the observation, and they are most commonly used in corridor mapping of road and railway environments. Point cloud files greatly speed the design process by providing real-world context in which you can re-create the referenced objects or insert additional models.

4.1. Recent Digital photogrammetry trends

Photogrammetry is a mature, remote sensing technique which dates from the end of the 19th century. From that time until today, it has passed through 4 phases of development: Plane Table Photogrammetry (triggered by the invention of photography) from about 1850 to 1900; Analogue Photogrammetry (invention of precision optics and mechanics and airplanes) from about 1900 to 1960; Analytical Photogrammetry (invention of computers) from about 1960 to the present and Digital Photogrammetry (microelectronics, semiconductors from the late 1980s to the present (Gruen, 2021).

The introduction of digital photogrammetry resulted in a substantial shift in data acquisition and data handling, especially in the automatization of data processing procedures, which were until then manually performed and very laborious. Nowadays, orientation/geo-referencing can be done partially automatically, while digital surface modeling and ortho-image generation are fully automated, and object extraction and modeling is possible in a semi-automated mode.

Photogrammetric sensors mounted on UAV or geodetic terrestrial total stations make the costs of this equipment and data acquisition and processing significantly cheaper, especially in terrestrial applications, but also in UAV-based aerial projects. UAV digital photogrammetry is mostly applied in small-scale thematic mapping, while classical photogrammetry is applied in parallel with this for large scale mapping.

Even though orthophoto images are widely used products in photogrammetry, it was always used for 3D measurements and 3D modeling. As with the LiDAR technology, photogrammetry provides 3D point clouds. The difference that distinguishes photogrammetry from Lidar is that it provides clouds that have an RGB value for each point, resulting in colorized point clouds (Figure 6).

Figure 5. 3D model of Studenica monastery produced from point clouds

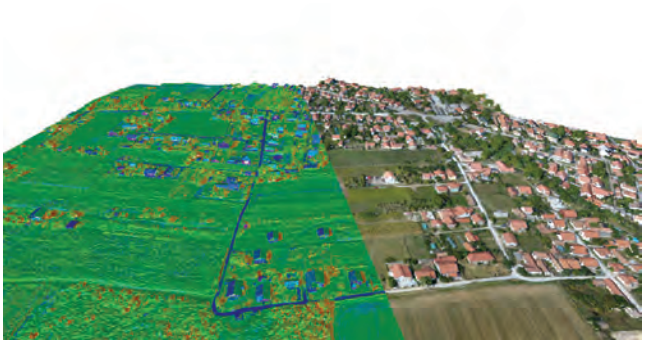


Figure 6. 3D point cloud (left) and 3D view of orthophoto image (right) of Požarevac town obtained by UAV photogrammetry

Today digital photogrammetry covers a very wide range of different applications: 3D cloud topographic mapping, precise agriculture, 3D/4D city modeling, BIM, and others. Specifically interesting are applications in close range domains of photogrammetry (terrestrial and UAV systems) like Cultural Heritage mapping, structural health monitoring, industrial quality control, human body measurement, etc.

5. STRUCTURAL HEALTH MONITORING

During recent decades, the need for quality inspection, structural assessment, and monitoring in the construction industry has significantly increased. Engaging traditional methods based on physical and classical geodetic methods for periodic inspection and condition assessment is quite tedious, costly, time-consuming and potentially risky. Besides, any carelessness in construction maintenance or postponed actions may result in large future expenses, especially in the case of structures with special purposes (Rashidi *et al.*, 2020).

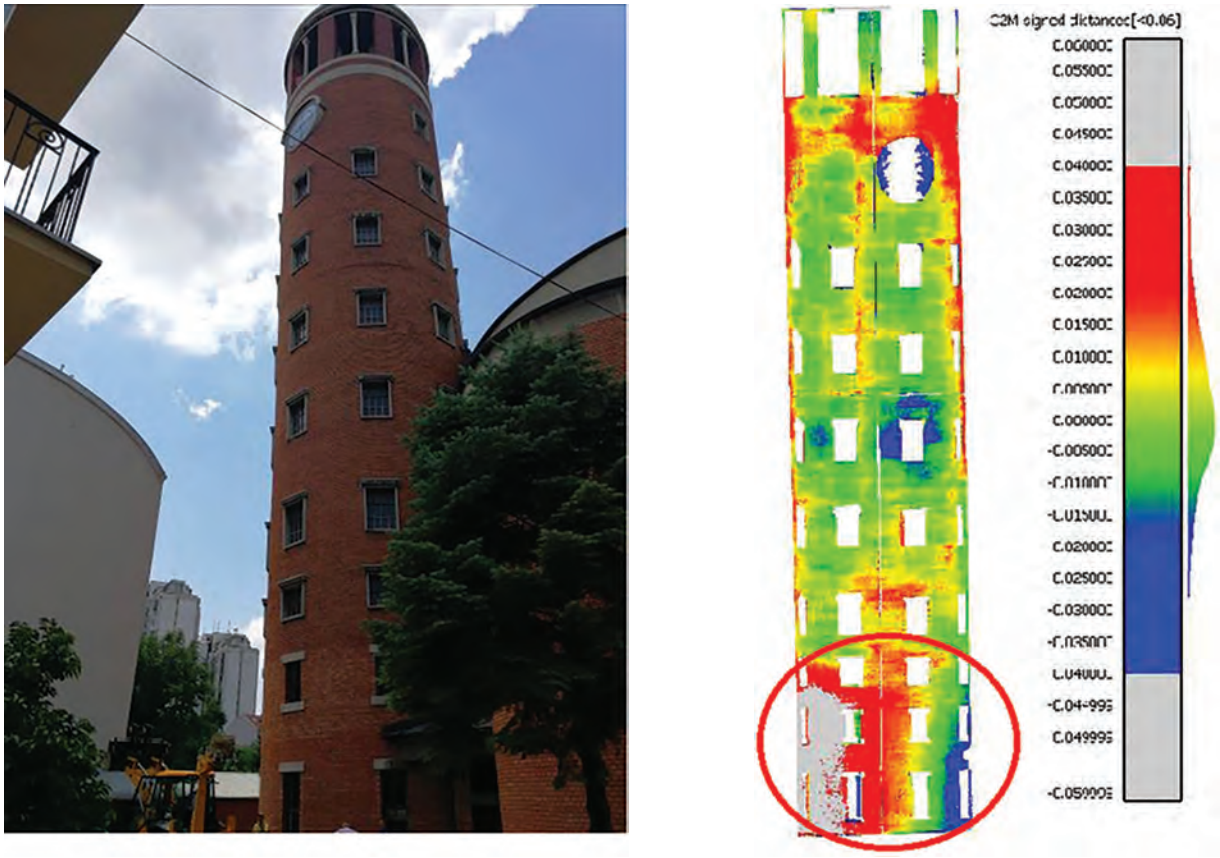
Geodetic monitoring of an object and its surroundings means presenting the structure by discrete points in such a way that they characterize the object, and the movements of the points represent the movements and distortions of the object. Therefore, the geometry of the object is modeled. Conventionally, modeling the deformation process means to observe the characteristic points at certain time intervals in order to monitor properly the temporal course of the object's movements (Beshr *et al.*, 2015).

Structural Health and Response Monitoring is an advanced method of monitoring structural status and performance without otherwise affecting the structure itself. Structural Health and Response Monitoring uses several types of sensors embedded in, or attached to, a structure to detect any exceedance of the allowed performance criteria, as well as to identify and verify structural behavior. The emerging use of Structural Health and Response Monitoring, particularly in the last decades, is a result of the increasing need for monitoring advanced designs and materials, as well as the improved management of existing structures.

Terrestrial laser canners (TLS) technology allows non-contact records of the behavior of an object being monitored (Figure 7). The scan rate of TLS is up to 1 million points per second, which significantly reduces the measuring time and increases the quantity of information obtained about the object. The accuracy of determining the 3D coordinates of single measured points by currently commercially available laser scanners is a few millimeters. The precision can be increased using suitable data processing. Based on an approximation of the specific parts of the monitored object by fitting geometric primitives to point clouds, using regression algorithms, the data processing can be significantly automated (Erdélyi *et al.*, 2018).

Satellite radar interferometry is a remote sensing technique able to measure displacements of the Earth's surface, tectonic deformations, volcanic deformations (Figure 8), ground subsidence or lift from various causes (most often gas extraction or groundwater activity), landslides and glacial motion, and to monitor the stability of infrastructure and buildings (Lesko *et al.*, 2018).

Airborne Lidar and photogrammetry are both viable methods for capturing point clouds for the 3D modeling of man-made hard structures. Although both methods produce point clouds, the manner of capturing data differs in many ways, resulting in point clouds with differing characteristics.



(a) **(b)**

Figure 7. a) The tower of St. Anton Padovansky Church in Belgrade b) Deviations of the tower with respect to an ideal cylinder, determined on the basis of TLS monitoring

6. HIGHER AND SPACE GEODESY

A comprehensive definition of higher geodesy includes the measurement of variations in Earth's shape, gravity field, and rotation, in studies of study dynamic processes in the realms of geophysics, hydrology, the cryosphere, atmospheric science, oceanography, and climate change (Freymueller *et al.*, 2021). The special branch of higher geodesy is physical geodesy, which is more concerned with gravitational field and determining the physical shape of the Earth.

In the six decades of the space era following 1957, when the first the artificial satellite Sputnik was launched, space geodetic technologies developed rapidly. There are two notable trends in the evolution of geodesy and its contributions to Earth sciences. First, scientific attention has progressed from stationary processes to time-varying and transient processes. Second, formerly distinct sub-fields within geodesy and geomatics are more converging in the sense of solving the same tasks, such as in the use of combined deformation measurements from Global Navigation Satellite System (GNSS), Interferometric Synthetic Aperture Radar (InSAR), and gravity changes to study mass transport. The accuracy enhancements and accessibility of space geodetic observations have led to an explosion in the range of Earth science disciplines that benefit from information coming from geodesy.

In the twenty-first century, all higher and physical geodetic studies will be predominantly based on geodetic measurements from space. Space geodesy uses a set of techniques established on precise distance, or phase measurements transmitted or reflected from extraterrestrial objects, such as quasars, the Moon, or artificial satellites. The distance or phase measurements between Earth's surface and objects in space are very precise and available for almost any location on the Earth's surface. This means that positioning, surface elevation and gravity field, and their changes over time can be determined precisely with global Earth coverage.

The first space geodetic measurements were based on existing astronomical equipment such as the radio telescope, and on the analysis of early satellite orbits. These observations yielded the first space-based observation of Earth's gravity field and tectonic plate movements. Early space geodetic measurements, beginning in the 1980s, had accuracy levels of between 5 and 10 centimeters. These measurements were conducted across the entire globe and yielded the first direct observations of tectonic plate motion. Soon after, satellites intended for geodetic measurements were developed. At the same time, data

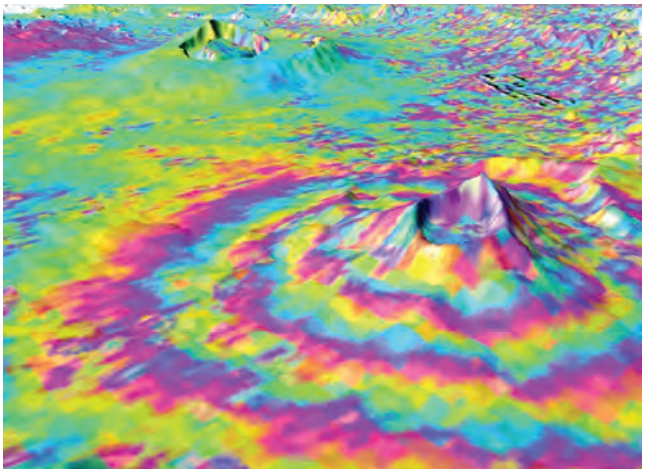


Figure 8. Envisat Advanced Synthetic Aperture Radar interferogram over the Kenyan section of the Great Rift Valley shows small surface displacements that are not visible to the naked eye of the Longonot Volcano (front right) (ESA, CC BY-SA 3.0 IGO)

obtainable from other satellite missions, such as GPS or SAR satellites, enable very precise positioning and detection of earthquake and magma-induced crustal deformation, subsidence, glacial movements, and wetland surface water level changes.

6.1. Higher and physical geodesy in Serbia

6.1.1. Establishment of new geodetic reference frames

Geodetic reference frames are fundamental for all positioning applications on the Earth, together with precise orbit determination for a wide range of satellite missions. The first significant efforts in this direction in Serbia were made in the 1990s, which coincided with the reorganization of the geodetic service and the formation of the Republic Geodetic Authority (RGA), which initiated basic geodetic works on the territory of Serbia. Members of the Institute of Geodesy and Geoinformatics from the Faculty of Civil Engineering in Belgrade have taken active roles in most of these project designs.

The realization of reference geodetic networks in Serbia started with the initiative of the International Association of Geodesy (IAG), and within a project by the European Reference Frame with the GPS measuring campaign “Balkan 98”. As a result of this campaign, the reference networks of Serbia and Montenegro (Former FRY), Bosnia and Herzegovina and Albania, are included in the European Terrestrial Reference Framework (ETRF) with 29 points. The reference network for Serbia and Montenegro consisted of eight points; six points are on the territory of the Republic of Serbia and two points are on the territory of Montenegro. The reference network project for the territory of Serbia and Montenegro was made by the Institute of Geodesy at the Faculty of Civil Engineering in Belgrade. The project was carried out by the Republic Geodetic Authority of Serbia and the Real Estate Directorate of Montenegro with logistical support from the Yugoslav Army.

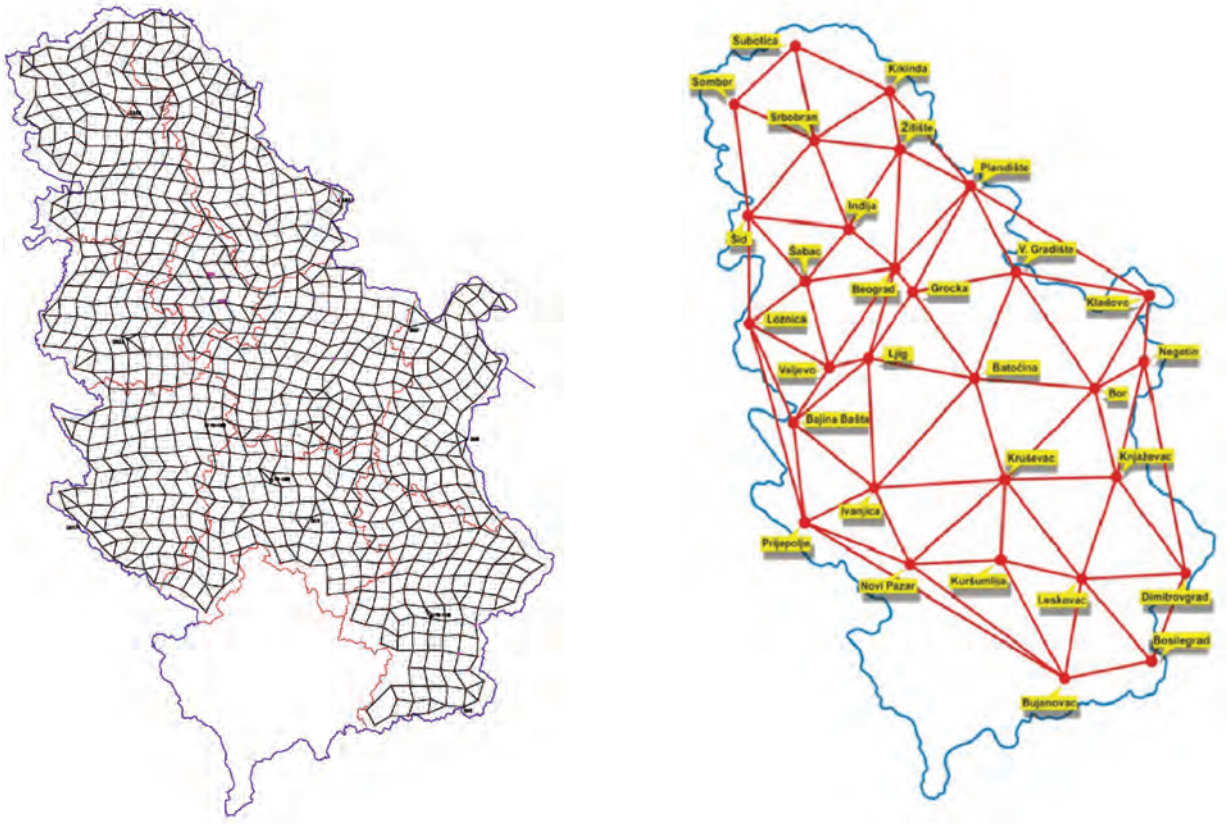
The coordinates of the points were determined in the reference system ITRF 96, epoch 1998.7, by standard deviation in position ±2 mm and 6.5 mm in height. In order to monitor the temporal evolution of the spatial coordinate system of the basic reference network (UREF Serbia), in July 2004 the EUREF 2004 measurement campaign was carried out in the ITRF96 measurement system for the 1998.7 epoch and an agreement was reached with an accuracy of 0.007 m for 10 points in the AGROS permanent station network.

6.1.2. Creation of a passive reference network for the territory of Serbia (SREF)

The project design was realized in the period from 1996 to 1998. A total of 838 points were stabilized and determined at an approximate distance between points of 10 km (Figure 9a). GPS measurements were performed by the method of relative static positioning. These measurements measured 1662 GPS vectors. The accuracy achieved for determining the coordinates was 0.60 cm in position and 1.01 cm in height. A complete passive network consisting of passive geodetic points at a mutual distance of 10 km was completed in the year 2003, in which the average accuracy of the coordinates of the points of the network is centimeter level. The implementation of the SREF network made it possible to establish a reference network that is in line with European standards and in the European reference system. The main purpose of the network was to support the completion of the land survey and real estate cadaster on the territory of Serbia. The passive SREF network was remeasured in 2010 with the same quality, which opened the opportunity for some preliminary studies on horizontal Earth crust deformations on the territory of Serbia.

6.1.3. Active Geodetic Reference Basis of Serbia (AGROS)

Establishment of the Active Geodetic Reference Basis of Serbia (AGROS), a permanent service of precise satellite positioning for the territory of Serbia and a network of permanent GPS stations began in December 2001 (Figure 9b). In February 2002, cooperation was established with the European Academy of the Urban Environment (EAUE), with the formation of a network of 16 permanent GPS stations in Central and Eastern Europe called EUPOS (European Positioning Determination System). It was then decided that, by harmonizing technical standards, the existing AGROS project would become a subproject of the EUPOS system. The project design was completed in 2003,



(a) **(b)**

Figure 9. a) Passive reference network for the territory of Serbia (SREF); b) Active Geodetic Reference Basis of Serbia (AGROS)

and AGROS was fully implemented in 2006 (in December 2005 it started being used commercially). AGROS consists of 31 permanent GNSS stations and it is based on Virtual Reference station (VRS) and Area Correction Parameter (ACP) concepts. Four stations of the AGROS network are included in the European Network of GNSS Stations – EPN (EUREFF Permanent GNSS Network. The main purpose of AGROS is to speed up the realization of topographic and land surveys for the real estate cadaster and engineering geodesy. It provides accuracy of 2-3 cm for Real Time Kinematic positioning.

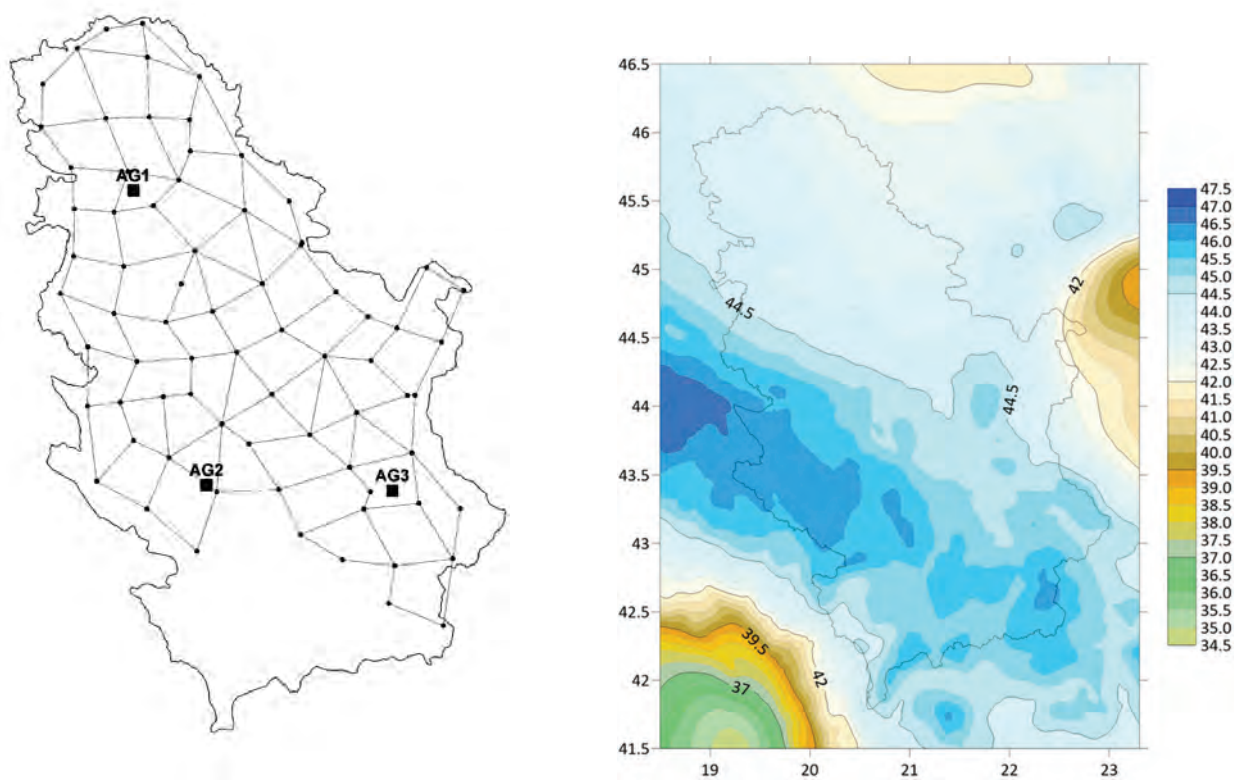
6.1.4. Network of the absolute acceleration of the Earth's gravity and New Basic Gravimetric Network (NBGN) of the Republic of Serbia

The project was carried out at the beginning of 2008. Field work was done within the "Capacity Building for Serbia: Real Property Registration and Cadaster Project". The main purpose of this network was to make a reference gravimetric framework for creating the basic gravimetric network of the Republic of Serbia.

Based on a project designed in 2005, the Basic Gravimetric Network was completed in the year 2007. During the period October 10-26, 2007, a Swedesurvey group (Lantmäteriet) performed a campaign to determine the absolute measurements at three points in Serbia using the Micro-g LaCoste FG5 absolute gravimeter (Odalović *et al.*, 2012). The NBGN was planned and implemented, with 78 points at an average distance of approximately 50 km, as a system of 60 closed polygons with 137 connections (Figure 10a). The main purpose of the network was to establish a basis for the future leveling network of Serbia. Realization of this network began in 2016, with planned completion in 2022. The *NBGN* will also be used for geophysical exploration on the territory of Serbia.

6.1.5. New Serbian state reference system ETRS89

During 2010, the Republic Geodetic Authority implemented the GPS measurement campaign EUREF Serbia 2010 as part of EUREF, with the aim of introducing a new state ETRS89 reference system in the Republic of Serbia. For the implementation of the new system, the conventional reference framework of the ETRF2000 reference system ETRS89 was adopted, which is in line with recommendations by the EUREF Technical Working Group (ETWG). Spatial coordinates for 12 points in the national reference system ETRS89 were determined which defined the coordinate system, and they were also a basis for determining the coordinates of other points of the SREF network in ETRS89 and ETRS2000. The campaign included measurements at 20 EPN stations, 48 stations of national permanent networks (30 in Serbia, 18 in Macedonia, Bul-



(a) **(b)**

Figure 10. a) New Basic Gravimetric Network of Serbia; b) Geoid of the Republic of Serbia

garia and Hungary) and 19 points of existing networks (12 in Serbia and 7 in Macedonia) during 2010. For the datum definition, only points that are in the system of the European network of permanent stations are useful, namely points that are declared as class points A, which means guaranteed accuracy of 1cm in the ETRS89 system and station velocities of less than 1 mm per year epoch. The EUREF campaign in Serbia in 2010 determined coordinates with an accuracy of about 2 mm for all three coordinate axes.

6.1.6. Geoid of the Republic of Serbia

The application of AGROS made it possible to determine the preliminary geoid for the territory of the Republic of Serbia. From 2006 to 2012, ellipsoidal heights of the second level of high accuracy (NVT2), as well as orthometric and normal heights of SREF points, were determined, which made it possible to determine discrete undulations at over 1100 points that are correctly distributed for the territory of Serbia.

At the same time, using gravimetric survey data (about 80,000 points) from the EGM96T global geopotential model (adapted to the acceleration anomalies for the territory of the Republic of Serbia) and digital terrain model with a resolution of one arc in both directions, the geoid for the Republic of Serbia was determined in 2009 with an accuracy of 15 cm.

In the period from 2009 to 2011, gravimetric data from the countries of the former Yugoslavia were taken over, so that in 2011 a recalculation of the geoid area was performed using data related to neighboring countries and the new EGM2008 model (Figure 10b).

6.1.7. Reference leveling network of Serbia

The ongoing project to create a Reference leveling network for Serbia started in September 2017. All of the field measurements are almost finished and the network is expected to become usable in 2023. The reference leveling network consists of a system of closed polygons, evenly distributed throughout the territory of the Republic of Serbia. The average perimeter of each polygon is about 115 km, with an average distance of 5 km between points. Such an arrangement of leveling lines and points is a good basis for use in engineering geodesy and land surveying, as well as a basis for its later densification, if necessary. The final structure of the reference leveling network should consist of 35 fundamental benchmarks of the Second order high accuracy leveling network (NVT2), 162 points of SREF, 1767 newly established points, 364 lines and 138 polygons (Figure 11). Because the new leveling lines follow the NVT2 lines as much as possible, it will be a good opportunity for preliminary studies on recent vertical crustal movements on the territory of Serbia.

Figure 11. Reference leveling network of Serbia

7. DIGITAL AND WEB CARTOGRAPHY

With the rapid development of ground observation and web technology, cartography is undergoing a more obvious development. Computer techniques have made a great contribution to cartography. Since the 1990s, map making has been transformed from manual to digital computer operation. Integrated digital cartography has become the main approach of map making and publication. In particular, the integration of geographic information systems, geo-databases, and remote sensing have greatly improved the efficiencies, contents, varieties and currency of digital maps.

The difference between map users and map designers is disappearing, and the methods used to explore and evaluate maps are also changing (Griffin *et al.*, 2017). The growth of Internet GIS technology and the World Wide Web have made it possible to access voluminous GIS data and maps in near real-time whenever they are available (Tsou, 2009). Massive volumes of geospatial data are obtainable and collected online, and are used in web applications and maps for different purposes. In the last decades of the 20th century, the main novel issue in cartography was the appearance of digital cartography – the transition between analog maps printed on paper and digital maps deliverable in new digital formats. Thanks to the progress of the Internet, the beginning of the 21st century brought new challenges, transforming cartography into a web-based tool for representing and assessing geographic information in interactive and mobile devices (Coetzee *et al.*, 2020).

In the multitude of definitions that describe web mapping, this term can most simply be described as the process of designing, implementing, generating and delivering maps on the World Wide Web (Neumann, 2008). From this definition it is clear that web mapping uses the web - multimedia platform, to deliver maps using the benefits of this technology, not neglecting the need to develop studies and theories based on digital cartography. The progress of web mapping started right at the beginning of the web, mainly in the form of published online static maps in the 1990s. In this relatively short time span of thirty years, web mapping has already passed through nine distinct eras (Veenendaal *et al.*, 2017). Each of these web mapping eras is based on one or more important technological developments that define that era (Figure 12).

The static web mapping era started in the early 1990s and it was based on basic HTTP and HTML technologies realized through HTML clickable images, and hyperlinks. The greatest achievement of this era was the possibility of sharing maps with multiple users beyond paper-based maps. Dynamic maps created using Dynamic HTML (DHTML), Common Gateway Interface (CGI), Java applets and servlets, plugins and ActiveX technologies gave users the opportunity to select layers and submit a request to view the resulting map according to the user's preferences and selections. The services web mapping era represents an upgrade in which Java Applets, CORBA, Microsoft.NET technologies, etc. were used to create replaceable components with well-defined interfaces to facilitate a flexible and interoperable environment (Veenendaal *et al.*, 2017).

The development of Asynchronous JavaScript and XML (AJAX) technology in association with image tiling technology represented the beginning of the Interactive Web Mapping Era. User interaction with the map was enhanced by allowing online maps to be delivered to a user in a continuous and fast way when the user was simultaneously interacting with the map interface. The start of Web 2.0 (in 2005) represented the beginning of the Collaborative Web Mapping Era. Users were enabled to make an active contribution to web mapping production. As a result, OpenStreetMap – a mapping platform for volunteered data capture – was founded in which users were free to share, create and adapt the database in open access format. For the first time, new terms such as Crowdsourcing and Volunteered Geographic Information (VGI) were introduced. As a consequence, users have increasingly become creators and not just consumers of web maps.

2005 also saw the beginning of the Digital Globes Era (also named virtual globe or virtual earth). The appearance of Google Earth, BING maps, Cesium, NASAWorldWind and others exposed maps and digital earth imagery

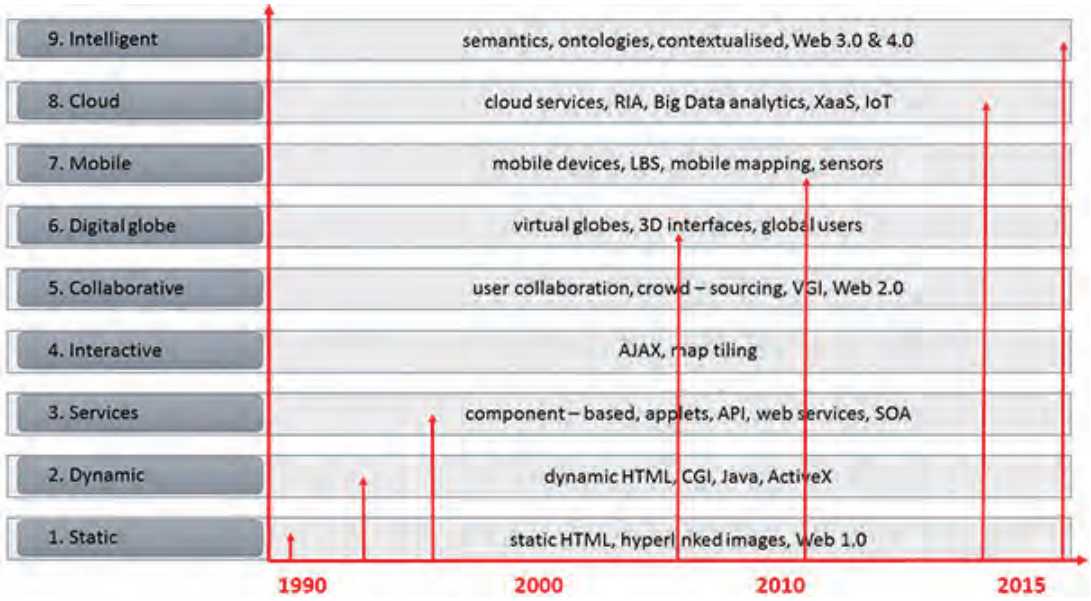


Figure 12. Outline of web mapping eras (Adopted from: Veenendaal *et al.*, 2017)

in real-time to the whole world's population through desktop or mobile devices, opening up the geospatial world of mapping worldwide via the Internet. The development of various free software packages like MapServer, GeoServer, etc., as well as Open Geospatial Consortium (OGC) standards, facilitated the expansion of web cartographic applications (Agrawal and Gupta, 2020).

The increasing use of smartphones and similar handheld devices facilitated access to and interaction with online maps regardless of location and time. Thanks to location-based services (LBS), the use of cell phone tower triangulation techniques and GPS-enabled smartphones, the outdoor, and even indoor, location of users can be identified. The Mobile and Location-Based Web Mapping Era provided new facilities in emergency services, marketing, navigation, traffic monitoring, weather monitoring, health, risk assessment, sports, tracking, etc. Linking all available sensors with mobile devices directly or indirectly connected to the Internet (Internet of things (IoT)) provided the generation with an enormous amount of data, which necessitated new methods for storage, software, services and infrastructure of data. The response to these challenges was the Cloud Web Mapping Era that supported the expansion of digital earth technologies, as well as complex geospatial applications.

Web services like Amazon's cloud storage, Microsoft Azure cloud ESRI ArcGIS Online, GeoNode, GIS Cloud, and CartoDB Cloud, and open-source platforms like Hadoop, Hive, and MongoDB, also made available and distributed parallel computing environments for performing geospatial big data analytics.

The appearance of the Web 3.0 generation of Internet technology, also known as the semantic web, which relies heavily on the use of machine learning and artificial intelligence (AI), brings us to the ongoing Intelligent Web Mapping Era. This era includes the representation of semantic information, automated and semi-automated building of ontologies, composition of semantic web services, natural language processing and data discovery (Zhang *et al.*, 2015).

Web mapping is still developing, and there is still some room for additional development in its usability, service benefits, quality and innovation challenges.

8. MODERN REAL ESTATE CADASTER

Real estate and utility cadastral registration and surveying, commonly known as the cadaster, is the most recognizable classical geodetic discipline. A cadaster considers the records of real estate, land and properties data maintained in graphical and textual formats. Although cadastral registration and surveying can be considered a technical profession, throughout history every country in the world has had some uniqueness in its own cadastral development, very much related to the social circumstances and historical heritage of that country.

However, globalization, social changes and technological progress in the 20th century spurred the work and research on with regard to developing modern global cadastral systems, which would take the place of traditional systems. In July 1998, the Commission for Cadaster and Land Management of the International Federation of Surveyors (FIG) announced the document "Cadaster 2014", which projected the trends and developed the vision of what cadastral systems might look like in the 21st century (Kaufmann and Steudler, 1998).

The Real Estate Cadaster presented in the Cadaster 2014 study includes tools for the adequate and simple

registration of real estate and rights to them. One of the main conclusions of this study is that the implementation of such a modern cadaster requires the support of modern information technologies (Kaufmann, 2002). This approach resulted in cadastral survey information today often being a base element in Geographic Information Systems (GIS) or Land Information Systems (LIS) used to assess and manage land and built infrastructure.

In a modern environment, it makes sense to develop a comprehensive cadaster on Web technology that will provide the stakeholders with access to different information systems by governing directories, controlling access rights, integrating data from different sources, and managing the rules to be applied (Kaufmann, 2017).

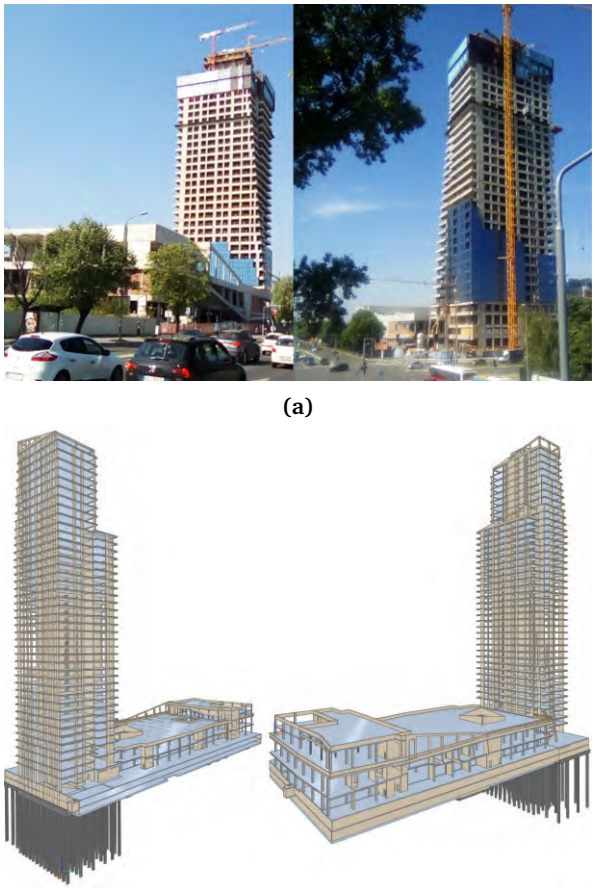
The growth of densely populated urban areas has caused traditional cadastral registration systems to face many difficulties in representing complex and multilevel property situations on 2D maps. These challenges, combined with the rapid development of 3D technologies, have forced the research and progress of 3D cadaster systems.

Although Cadaster 2014 does not explicitly mention the 3D cadaster, it recognizes the need to modernize traditional cadastral systems and transform them into more modern systems that will rely on information technologies. Currently, the real estate cadaster is based on 2D cadastral maps and procedures that do not support the unequivocal registration and visualization of complex 3D property situations and complex objects, especially in urban areas.

During the past two decades, much attention has been paid to the 3D cadaster issue, raising worldwide awareness of the need for 3D cadastral systems. The 3D registration of real properties, unequivocal registration of complex 3D property situations, and their visualization are necessary in a 3D cadaster to increase transparency in land administration and property registration (Višnjevac *et al.*, 2019).

Given that the 3D Real Estate Cadaster will be one of the main parts of the national geospatial data infrastructure (NDSI), it is necessary to take into account the INSPIRE Directive when developing and implementing the 3D Real Estate Cadaster. The 3D cadaster should represent 3D geometries of cadastral objects and handle their spatial relationships in a 3D space.

In order to ensure a smooth transition between the 2D and 3D cadaster, it is necessary to provide appropriate standards. At the moment there is no specific standard committed to 3D cadasters. However, an important contribution in the land and urban administration is the Land Administration Domain Model (LADM) of ISO 19152 (Van Oosterom *et al.*, 2013). The LADM standard provides a formal conceptual model for recording and managing land administration data. This standard is not dedicated to 3D cadasters, but it can be adopted to model 3D property and the semantic information associated with rights, restrictions, and responsibilities regarding the property, providing in such way an extendable basis for the development and improvement of efficient and effective land administration systems.



(a)

(b)

Figure 13. a) Photo of the construction phase of the "West 65" project (Belgrade) b) The BIM model of the construction project (Adopted from: Petronijević *et al.*, 2021)

An important research trend is coupling Building Information Modeling (BIM) technology with the cadaster. BIM contains accurate geometric data and building information (Figure 13), which can support the management of land and property information in high-rise buildings (Petronijević *et al.*, 2021). The linkage between BIM and the 3D cadaster is a new research area that has become a focus of various investigations.

9. STRATEGY AND OBLIGATIONS OF SERBIA IN THE FIELD OF GEOSPATIAL DATA INFRASTRUCTURE

The importance of geospatial data and geoinformation in modern society is best illustrated by the adoption of the INSPIRE (INfrastructure for SPatial Information in Europe) directive, which was adopted by the Council of Europe and

the European Parliament in March 2007, in order to develop, support and harmonize activities related to environmental management policy. One of the main tasks set for all EU countries was the establishment of National Geospatial Data Infrastructures – NSDIs, which are integrated geospatial data systems. These infrastructures should enable users to comprehensively identify and access spatial information obtained from a variety of sources, from the local, through the national to the global level.

In the process of accession to the European Union within the framework of Chapter 27, which refers to the environment, the Republic of Serbia has committed itself to implementing the INSPIRE directive, for which the first steps in legal terms were taken in the 2009 amendments to the Law on State Survey and Cadaster, in which the Republic Geodetic Authority (RGA) was given jurisdiction over the formation and management of the NSDI. The next step in the implementation of the INSPIRE directive was adoption of the Law on National Geospatial Data Infrastructure, which entered into force in April 2018.

By recognizing the RGA as a key holder of work in the construction and maintenance of the national geospatial data infrastructure, the potential of the geodetic profession itself as a leader in this type of work was recognized. In the part of the strategy that refers to measures and activities for increasing the quality of services in the field of geospatial data and registration of real estate rights in the official state records, RGA marked the development and adjustment of the personnel structure in the geoinformation sector as one of the strategic goals.

10. GEODESY AND BIG DATA CHALLENGE

With increasing data volumes and availability of open access to all data sources, geodesy/geoinformatics has become a data-rich field. Data are continuously collected from thousands and thousands of GNSS stations around the world, and the number of installed instruments continues to grow. Multiple InSAR satellites already repeatedly record the footage of most of Earth's land surface in high resolution timescales. Even if some parts of the world are still poorly sampled, very soon in the future we will have data everywhere, at any time (Freymueller *et al.*, 2021).

The volume of data obtained from geodetic observation techniques has increased dramatically. In the same way, numerical environmental models and other geophysical models important for geodesy come with continually growing space/time resolutions. The data obtained from a multitude of sources requires advanced approaches to efficiently handle and couple the vast amount of available geodetic data. Definitely, new space and earth observation techniques are tackling challenges related to extensive data collection ("big data"). Strategies and methodologies from the fields of data science and machine learning have shown great potential not only in this context, but also when applied to more limited data sets, in order to solve complex non-linear problems in geodesy.

11. FUTURE PROSPECTIVE IN GEODESY AND GEOINFORMATICS

The application of Earth Observation (EO) is rapidly changing as a result of exponential advances in sensor and digital technologies. Recent decades have witnessed extraordinary developments in ICT, including the Internet, cloud computing and storage, which have all led to radically new ways of collecting, distributing and analyzing data about the Earth.

The airborne, satellite and mobile platforms are adding huge volumes of geodetic data at the regional and global level, multiplying the data challenges related to data analysis and new numerical modeling and interpretation of big data. In particular, they are facing simultaneous challenges and opportunities related to expansive data collection ("big data"). Also, requirements for new environmental numerical models and other geophysical models important for geodesy come with ever growing resolutions and dimensions, as well as the combination and extraction of information from multiple inhomogeneous data sets.

This requires new experts, whose knowledge will cover a wide range of disciplines, starting with modern techniques of spatial data acquisition, through the knowledge of information technology, software, and techniques (machine learning, enhanced learning, etc.) for processing and modeling complex data sets. They will be practically experts in the analysis and processing of data (Data Science) in an area that is considered a key profession of the 21st century, with special emphasis on the creation and management of geospatial data. At this moment, the interest in geospatial data in the world is best shown by the statistics released by Google, which refer to publicly available databases (there are about 25 million of them at the moment), revealing that databases related to geospatial data are in first place.

As a response to these challenges, the Department of Geodesy and Geoinformatics of the Faculty of Civil Engineering at the University of Belgrade is the first higher educational institution in Serbia to have taken on the responsibility, from this academic year, of accrediting geodesy and geoinformatics as two parallel separate curriculums that will educate professionals to meet the demands of these fields in the near future.

REFERENCES

Agrawal, S. and Gupta, R. D. (2020). Development of soa-based webgis framework for education sector. *Arabian Journal of Geosciences*, *13*(13), 1–20.

Andries, A., Morse, S., Murphy, R., Lynch, J., Woolliams, E. and Fonweban, J. (2019). Translation of earth observation data into sustainable development indicators: An analytical framework. *Sustainable Development*, *27*(3), 366–376. doi: 10 .1002 / sd .1908

Artiola, J. F., Brusseau, M. L. and Pepper, I. L. (2004). *Environmental monitoring and characterization*. Elsevier Academic Press, Amsterdam; Boston.

Basic, T., Medak, D. and Pribicevic, B. (2003). Quo vadis, geomatica? *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences.*, *34*(6/W11), 26–29.

Beshr, A. A. E.-W. et al. (2015). Structural deformation monitoring and analysis of highway bridge using accurate geodetic techniques. *Engineering*, *7*(08), 488–498. doi: 10 .4236 / eng .2015 .78045

Coetzee, S., Ivánová, I., Mitasova, H. and Brovelli, M. A. (2020). Open geospatial software and data: A review of the current state and a perspective into the future. *ISPRS International Journal of Geo-Information*, *9*(2), 90.

Commission. (2021). *7. available online.* http :// www .faridesm .ir / cad2014 _eng .pdf. (accessed on 20 July 2021)

Erdélyi, J., Kopacik, A. and Kyrinovic, P. (2018). Multi-sensor monitoring of suspended steel bridge structure. In M. et al. (Ed.), *Advances and trends in geodesy, cartography and geoinformatics* (pp. 27–33). Taylor & Francis Group, London.

Ferreira, B., Iten, M. and Silva, R. G. (2020). Monitoring sustainable development by means of earth observation data and machine learning: a review. *Environmental Sciences Europe*, *32*(1), 1–17. doi: 10 .1186 / s12302 -020 -00397 -4

Freymueller, J. T., Bendick, R., Borsa, A., Newman, A., Brooks, B., Fu, Y., Kinsman, N., Larson, K., Plag, H.-P. and van Dam, T. (2021). *Measuring the restless earth- grand challenges in geodesy.* https :// www .unavco .org / community / publications _and _reports / community -vision / geodesy _science _plan / measuring -the -restless -earth .pdf. (accessed on 2 August 2021)

Griffin, A. L., Robinson, A. C. and Roth, R. E. (2017). Envisioning the future of cartographic research. *International Journal of Cartography*, *3*(sup1), 1–8.

Gruen, A. (2021). Everything moves: The rapid changes in photogrammetry and remote sensing. *Geo-spatial Information Science*, *24*(1), 33–49.

Hake, G., Grünreich, D. and Meng, L. (2002). *Kartographie: Visualisierung raum-zeitlicher informationen (8. ausgabe)*. De-Gruyter, Berlin/New York.

Hoge, F., Swift, R. N. and Frederick, E. B. (1980). Water depth measurement using an airborne pulsed neon laser system. *Applied Optics*, *19*(6), 871–883.

Jensen, J. R. (2006). Remote sensing of the environment an earth resource perspective, 2nd edition. In Pearson (Ed.), .

Kaufmann, J. (2002). Cadastre 2014: A vision for a future cadastral system. In *In proceedings of 1st congress on cadastre in the european union.* Granada.

Kaufmann, J. (2017). From a traditional to a comprehensive cadastre. In *Cadastre: Geo-information innovations in land administration* (pp. 13–22). Springer.

Kaufmann, J. and Steudler, D. (1998). A vision for a future cadastral system. *Working group 1 of FIG Commission*, *7*.

Lesko, M., Papco, J., Bakon, M., Czikhardt, R., Plakinger, M. and Ondrejka, M. (2018). The use of satellite radar interferometry to monitor landslides in slovakia. In M. et al. (Ed.), *Advances and trends in geodesy, cartography and geoinformatics ii.* Taylor & Francis Group, London.

Li, X., Liu, C., Wang, Z., Xie, X., Li, D. and Xu, L. (2020). Airborne lidar: state-of-the-art of system design, technology and application. *Measurement Science and Technology*, *32*(3), 032002.

Liang, S. and Wang, J. (2019). *Advanced remote sensing: terrestrial information extraction and applications*. Academic Press.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P. and Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 277–293. doi: 10 .1016 / j .isprsjprs .2017 .06 .001

Neumann, A. (2008). Web mapping and web cartography. In *Encyclopedia of gis* (pp. 1261–1269). Springer.

Odalović, O., Starcević, M., Grekulović, S., Burazer, M. and Aleksić, I. (2012). The establishment of a new gravity reference frame for serbia. *Survey Review*, *44*(327), 272–281.

Petronijević, M., Višnjevac, N., Praščević, N. and Bajat, B. (2021). The extension of ifc for supporting 3d cadastre ladm geometry. *ISPRS International Journal of Geo-Information*, *10*(5), 297.

Rashidi, M., Mohammadi, M., Sadeghlou Kivi, S., Abdolvand, M. M., Truong-Hong, L. and Samali, B. (2020). A decade of modern bridge monitoring using terrestrial laser scanning: Review and future directions. *Remote Sensing*, *12*(22), 3796. doi: 10 .3390 / rs12223796

Šíma, J. (2007). Geomatics and geoinformatics in modern information society–projection of new trends into their curricula at the university of west bohemia in pilsen. In *Scientia est potentia–knowledge is power, fig commission 2 symposium, prague (czech republic)* (pp. 7–9).

Tsou, M. (2009). *Recent developments in internet gis.* https :// www .geospatialworld .net / article / recent -developments -in -internet -gis /. (accessed on 10 August 2021)

Van Oosterom, P., Lemmen, C. and Uitermark, H. (2013). Iso 19152: 2012, land administration domain model published by iso. *FIG Working Week 2013 in Nigeria–Environment for Sustainability, Abuja, Nigeria, 6–10 May 2013*.

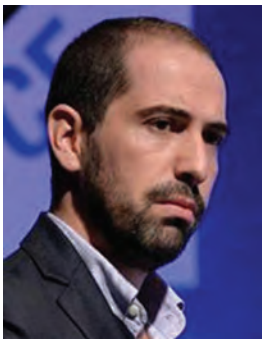
Veenendaal, B., Brovelli, M. A. and Li, S. (2017). Review of web mapping: Eras, trends and directions. *ISPRS International Journal of Geo-Information*, *6*(10), 317.

Višnjevac, N., Mihajlović, R., Šoškić, M., Cvijetinović, Ž. and Bajat, B. (2019). Prototype of the 3d cadastral system based on a nosql database and a javascript visualization application. *ISPRS International Journal of Geo-Information*, *8*(5), 227.

Wdowinski, S. and Eriksson, S. (2009). Geodesy in the 21st century. *Eos, Transactions American Geophysical Union*, *90*(18), 153–155.

Zhang, C., Zhao, T. and Li, W. (2015). *Geospatial semantic web*. Springer International Publishing: Cham, Switzerland.

Davide Lo Presti



Davide Lo Presti is an Assistant Professor and holder of the "Rita Levi Montalcini fellowship" at the University of Palermo, Italy and a Visiting Academic at the University of Nottingham, UK. In the last 10 years, Dr. Lo Presti initiated and coordinated several research projects/programmes looking at smart solutions for sustainable transport infrastructures, with a total budget of more than 12M EUR provided mainly by the European Commission (FP7 and H2020) and the Conference of European Directorate of Roads (CEDR).

Contact information: e-mail: davide.lopresti@unipa.it/davide.lopresti@nottingham.ac.uk

https://orcid.org/0000-0002-5125-8074

Prof. Lo Presti graduated from the University of Palermo, Faculty of Engineering in 2005, and obtained his master's thesis in 2007. He defended his PhD thesis in collaboration with the University of Nottingham in 2011. He held several research positions at the University of Nottingham, Nottingham Transportation Engineering Centre (Research Associate, Senior Research Fellow, Principal Research Fellow), during 2011-2019. He worked as a Visiting Researcher at TU Delft, University of Washington, University of Sao Paolo and University of California, Davis.

His expertise is Sustainable Engineering, with a focus on transport infrastructure's technology development and asset management. His research is both applied and fundamental and involves investigation of effective techniques for recycling waste/secondary materials into civil engineering applications such as roads, airports and railways. Also, with the goal of investigating a sustainable asset management for preserving existing infrastructures as well as envisioning the future ones, Prof. Lo Presti is specialising in smart technologies, life-cycle management and multi-criteria decision making.

Within the above fields, since 2010, Prof. Lo Presti has authored more than 80 publications mainly in international scientific journals. He has been involved in training and technology transfer to small companies as well as big enterprises and association of industries. His works are more than 1800 times cited ($h = 21$). He serves as a Deputy Panel Chair of "Proceedings of the Institution of Civil Engineers - Transport" Journal and a member of the Editorial Board of: the Journal of the American Asphalt Pavement Technologist (AAPT), Advances in Civil Engineering and Cogent Engineering Journal. So far, he supervised 7 post-doctoral researchers, more than 15 PhD theses and a number of MSc theses. He was an invited speaker, plenary speaker or Chair at more than 40 international conferences, and (co)organized more than 15 international conferences. Prof. Lo Presti serves as a reviewer in more than 15 international journals.

Dr. Lo Presti has extensive experience in coordinating international and multi-disciplinary projects. From 2011, he secured more than 12 million € in research grants, currently coordinating 3 research projects/programs for more than 5M€. He is Task Group Leader within three Technical Committees (RAP, WMR and PAR) within the Asphalt materials cluster of the International Community RILEM.

Towards smart transport infrastructure

Davide Lo Presti

University of Palermo (IT)
University of Nottingham (UK)

Summary

The world's transport network has developed over thousands of years, emerging from the need of allowing more comfortable trips to roman soldiers to the modern smooth roads enabling modern vehicles to travel at high speed and to allow heavy airplanes to take off and land safely. However, in the last two decades the world is changing very fast in terms of population growth, mobility and business trades creating greater traffic volumes and demand for minimal disruption to users, but also challenges such as climate change and more extreme weather events. It is having in mind these key challenges, that the SMARTI European Training Network carried out research, training and dissemination to contribute towards the achievement of the paradigm shift to *SMARTI: Sustainable Multi-Functional Automated and Resilient Transport Infrastructures* (http://smartietn.eu). This paper represents a summary of the four years effort produced by the consortium and provides three main findings, the principles to follow towards achieving this change, the details of the SMARTI and finally a series of prototypes developed within individual research projects advancing the research and development of specific technology concepts, decision making framework and/or asset management methodologies for roadways, railways and airports. As a result, this work aims at stimulating a cultural change amongst academia and stakeholders of the transport infrastructure industry, responding to skills shortage and addressing the need to form engineers versed with a multi-disciplinary approach, able to work within a collaborative international environment and, above all, knowledgable of the main challenges of our times.

Keywords: Smart Infrastructure, transport infrastructure, Sustainability, Multi-functionality, Automation, Resilience, SMARTI

1. INTRODUCTION: WHAT IS A SMART TRANSPORT INFRASTRUCTURE?

In order to meet the both European strategic energy and climate targets for *Smart, sustainable and inclusive growth, Smart, green and integrated transport* and ensure the transition to a low-carbon economy by 2050, Europe needs to invest in research and training to modernise infrastructures and management practices and avoid spending billions of Euros on maintaining the existing European surface transport infrastructure networks. In fact, some infrastructure has experienced a fast rate of deterioration as a result of poor design/construction or more demanding loads than anticipated, and many transport infrastructures are getting close to the ends of their design lives. It is for this reason that *European policies forecast a future transport system that is resilient, resource-efficient, climate- and environmentally-friendly, safe and seamless for the benefit of all citizens, the economy and society*. This is seen to happen with "smart" and integrated solutions in transport infrastructure design, maintenance and operations that are estimated to contribute to a 25% reduction in fuel and energy consumption for today's vehicles[1]. However policies do not indicate clear solutions for this problem and, although it seems that 'Smart' appears to be a panacea that will improve society in a multiplicity of ways, there is no common definition in its usage and there is no policy that specifies it. There is a universal thread to the way in which 'smart' is used, so **what is a 'Smart Infrastructure'?**

In October 2011, The Royal Academy of Engineering in UK held a roundtable meeting where professional from politics, engineering, communications, higher education, energy, transport, business and innovation, with the aim to trying answer this question[2].

This group provided a first definition stating that:

> *"Smart infrastructure' responds intelligently to changes in its environment, including user demands and other infrastructure, to achieve an improved performance."*

[1] FEHRL, National Road Research Centres in Partnership (2012), 10 reasons why transport infrastructure research is needed in Horizon 2020, FEHRL news 2-August-2012. http://www.fehrl.org/index.php?m=33&a=content&id=704

[2] The Royal Academy of Engineering "Smart infrastructure: the future" – January 2012

Figure 1. Smart Infrastructure principles according to The Royal Academy of Engineering in UK

According to this definition an infrastructure should act, based on information captured from sensors. Data collection is at the core of the smartening of the infrastructure and on this basis the system would then allow appropriate analysis of the collected data. Furthermore, the system uses a feedback loop of data, which provides opportunities to increase its performance. The resulting evidence will then lead to more informed decision-making. Results can then be communicated to humans and/or automatically translated into actions from the system.

Within this meeting applications of smartness that have benefited different industries were identified and potential barriers to a smarter future were also examined. As a conclusion, the group of stakeholders stated that: "*The infrastructure found within the UK is becoming smarter and this presents many economic and societal benefits, however there are barriers preventing further integration and further smartening of systems. To ensure that this smartening and integration continues smoothly, organisations need growing their skills and developing better communication with each other, with users and with government [. . .].*

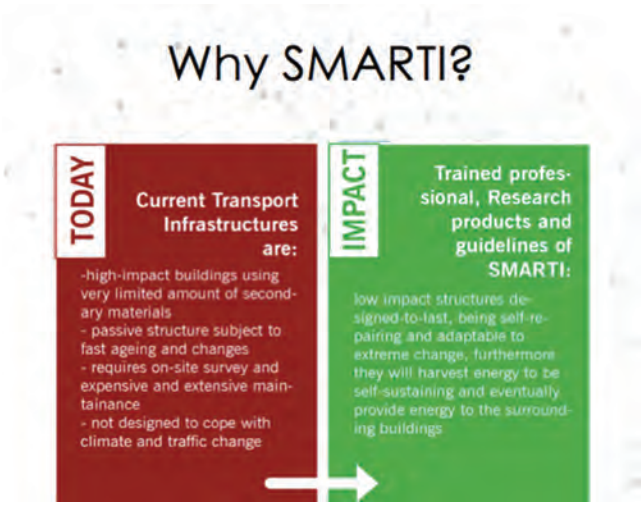


Figure 2. WHY smart infrastructures are needed

This first experience stimulated a multi-sectorial roundtable around Europe and having close discussions with partners and key stakeholders it was decided to extend the wave of change from the UK to the whole continent for empowering Europe to overcome these barriers and deliver long-term benefits with a wider vision. In fact, in the last two decades the world is changing very fast in terms of population growth, mobility and business trades creating greater traffic volumes and demand for minimal disruption to users, but also challenges such as climate change and more extreme weather events. It's from this urgent need that within our consortium, it was deemed necessary to go beyond the definition of the Royal Academy that considered a smart infrastructure as an "intelligent" system that can help human improving their efficiency in its management, and move towards:

> *A SMART Transport Infrastructure is a system designed, constructed and managed by adopting smart (elegant) solutions allowing the infrastructure to cope/adapt to the main challenges of our times*

It is then of paramount importance that transport infrastructure must adopt smart solutions that trained professionals should find by first of all learn about these key challenges (i.e. sustainability). On this basis, engineers should then embrace innovations to conceive transport infrastructure that adapt and last.

1.1. SMARTI European training Network

The above-mentioned vision of Smart Transport Infrastructure adopts a holistic approach to achieve 'revolution, not evolution in transport infrastructure engineering and playing with the word it was defined as *SMARTI: **S**ustainable **M**ulti-functional **A**utomated **R**esilient **T**ransport **I**nfrastructure*. It is with this ambitious goal and in close consultation with key stakeholders, that was concluded that in order to achieve long-term benefits it is necessary to shape a large multi-sectorial and collaborative training-through-research-initiative aiming at ***creating the professionals that can exploit it***. This new generation of professionals will represent the people that will actually design, construct, operate and maintain the transport network of the future. Industry and the private sector, were well aware of the problems and convinced of the real opportunity offered by a joint academy-industry training-through-research programme aimed at conceiving and promoting the paradigm shift to the *SMARTI*. This inspired the shaping of *SMARTI ETN: **S**ustainable **M**ulti-functional **A**utomated **R**esilient **T**ransport **I**nfrastructures European training network* a large collaborative training-through-research-initiative where experts of the four main pillars **SMAR)** including outstanding researchers and innovators, developers of high-tech sensors, remote communications, advanced monitoring equipment, software developers and world-wide expert trainers, could work alongside and transfer knowledge to the stakeholders of the transport infrastructure **(TI)** sector.

Figure 3. SMARTI ETN programme

The overarching goal of SMARTI ETN was then providing high quality interdisciplinary, inter-sectorial and international training to young researchers, thereby contributing to the development of a new generation of multidisciplinary engineers, fully able to work within the field of smart transport infrastructures. In order to develop a network of such engineers, SMARTI ETN provided: (a) in-depth training by research at both academic and industrial host sites through joint academy-industry recruitment and secondments to other research centres and stakeholders; (b) sound multidisciplinary and inter-sectorial scientific training and understanding of environments at the stakeholders via courses and secondments; (c) a programme of complementary skills training and network wide training events to develop competencies and career options. The project also provided outstanding opportunities for Transfer of Knowledge (ToK) between the stakeholders including SMEs, large organisations and academia.

SMARTI ETN also aimed at the following sub-objectives:

- *creating a platform* where training in smartening of transport infrastructure will be finally structured and provided at graduate level. It will not be anymore sporadic and scattered amongst various disciplines throughout Europe.
- providing outstanding opportunities for *Transfer of Knowledge*, through the fellows and between the transport infrastructure stakeholders and innovators acquainted with smartening of systems
- generate a summary of the "*SMARTI findings for early exploitation*" ***in Europe*** and ***in developing countries*** summarizing scientific results and best practices for ready use in EU and to be disseminated world-wide.

SMARTI ETN core was formed by eight institutions allocated in six countries: United Kingdom, Ireland, Denmark, France, Spain and Italy. Among the eight institutions, three universities, one public research centre and four private companies are involved. Each institution hosts one or several fellows providing all the necessary support for the development of their project. (http://smartietn.eu). The programme ended in August 2021 and as a main result it allowed developing and refining the following outcomes that are further detailed in this paper:

- "SMARTI Principles" containing the founding principles to acquire the necessary skills and support the further development and implementation of SMARTI in Europe and beyond.
- "SMARTI Vision", illustrating the details of the features that transport infrastructure should acquire to smarten their characteristics.
- "SMARTI Prototypes", a series of technology concepts, asset management methodologies and decision-making frameworks developed along the lines of the SMARTI vision.

2. SMARTI PRINCIPLES

SMARTI programme created a multidisciplinary European Training Network where developers of high-tech sensors, remote communications, advanced monitoring equipment, software developers and world-wide expert trainers, could work alongside and transfer knowledge to the stakeholders of the transport infrastructure sector. This inter-sectorial and multi-disciplinary group went also through a much-needed training-through-research programme that enabled the transfer of knowledge between the scientists as well as the 15 early-stage researchers. This process allowed to develop the SMARTI prototypes and contributed in shaping a list of definitions and principles that are here reported to support further development in the field:

1. ***Smart solutions for transport Infrastructure can be conceived by acquiring multi-disciplinary competences through international and collaborative research***
   As highlighted earlier, the challenges faced by transport infrastructure managers are mounting rapidly as transport demand increases and funding becomes limited, while the political push for safety and sustainability intensifies. These challenges call for integrative strategies that will require innovation in the many different functions in the transport sector. Implementation of innovation is fuelled by insights obtained acquired with a multi-disciplinary approach and collaborative research. Such insights can come by solutions originally aimed at different applications, hence research results are often "on the shelf" for quite some time before the situation arises in which they are eventually mobilised and adapted for use (i.e. solar panels for energy harvesting). Also, this approach allows acquiring new techniques to carry out novel decision making with already existing data/results/measurements of the asset (i.e. machine learning for data analysis).

As a result, transport infrastructure managers must have a research and development sector that requires researchers with a multi-disciplinary expertise to be constantly attentive to opportunities for innovation where already existing knowledge can be used.

2. ***These SMART (elegant) solutions should be engineered at different levels and applied with a life-cycle thinking approach.***
   The transport infrastructure engineer ecosystem must change. In order to conceive smart solutions, multi-disciplinary research is needed at all levels ranging from materials to infrastructure components, to then going to even bigger scale from network level to cities/province/regions. Furthermore, researchers should promote innovations at any stage of the life-cycle of the infrastructure hence, from the design and planning, including material/product supply, to construction and use/management of the transport infrastructure asset.

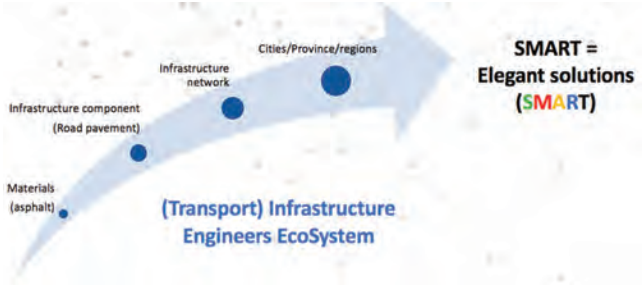


Figure 4. Engineers' ecosystem to smarten transport infrastructure

3. ***Development and Implementation of smart solutions can be facilitated by following the technology readiness levels (TRLs)***
   Collaborative Research with long-term and short-term goal is needed for a strategic development of the solutions and to facilitate the migrations from other sectors as well as experience benefits and limitations at the appropriate scale.
   In the 70s, NASA developed the Technology readiness level (TRL) scale as a type of measurement system used to assess the maturity level of a particular technology. TRL scale uses the parameter that evaluates the maturity of a technology according to a series of indicators that go from 1 (the basic principles are documented) to 9 (the technology is released, and industrial production is started".
   TRL is useful to indicate the research project result's maturity level and define what steps should be taken in order to bring the research result to the market. Hence, Assessing the TRL of each new technology and/or research project results will help stakeholders understanding the necessary step needed to implement it also in the transport infrastructure field.

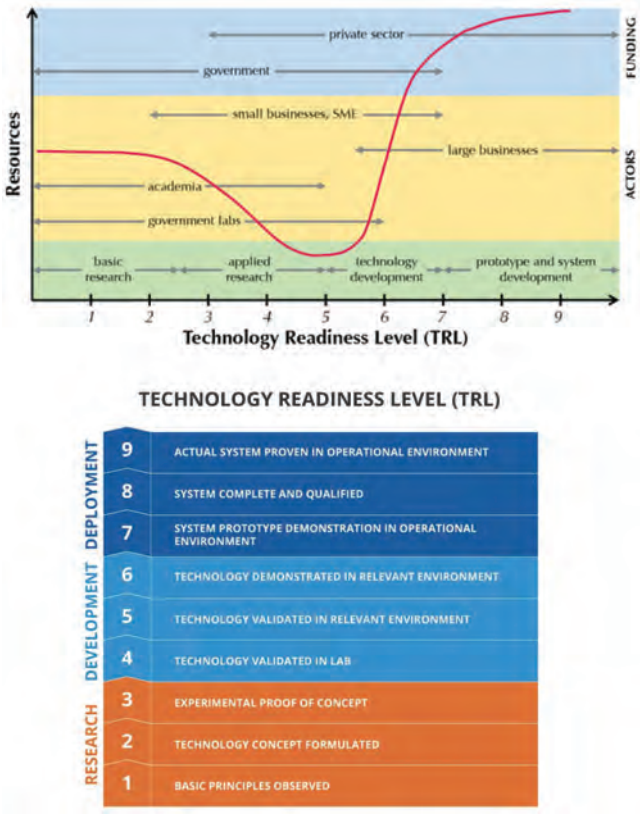


Figure 5. Technology readiness level scale

In particular, the following steps should be considered to "smarten" the infrastructure by adapting existing technologies from other fields, or developing new ones. Here it's a step-by-step process to be adopted for research and development towards smart transport infrastructure with an example for "automated" transport infrastructure

- State of the art of current best practice in construction and technology (i.e state of the art of infrastructure monitoring)
- Transfer of technologies and methods from other sectors (i.e. structural health monitoring of building and/or other structures)
- New research on the adoption of early stage and emerging techniques and products (i.e. sensors for responsive infrastructure)
- Demonstration of all the technologies and systems developed. (i.e. accelerated testing and/or live lab)

3. SMARTI VISION

The SMARTI vision depicts the most probable world in which we might travel in 30 years; it is based on the experience of the partners and on extensive analysis of existing visions and very recent research programs that

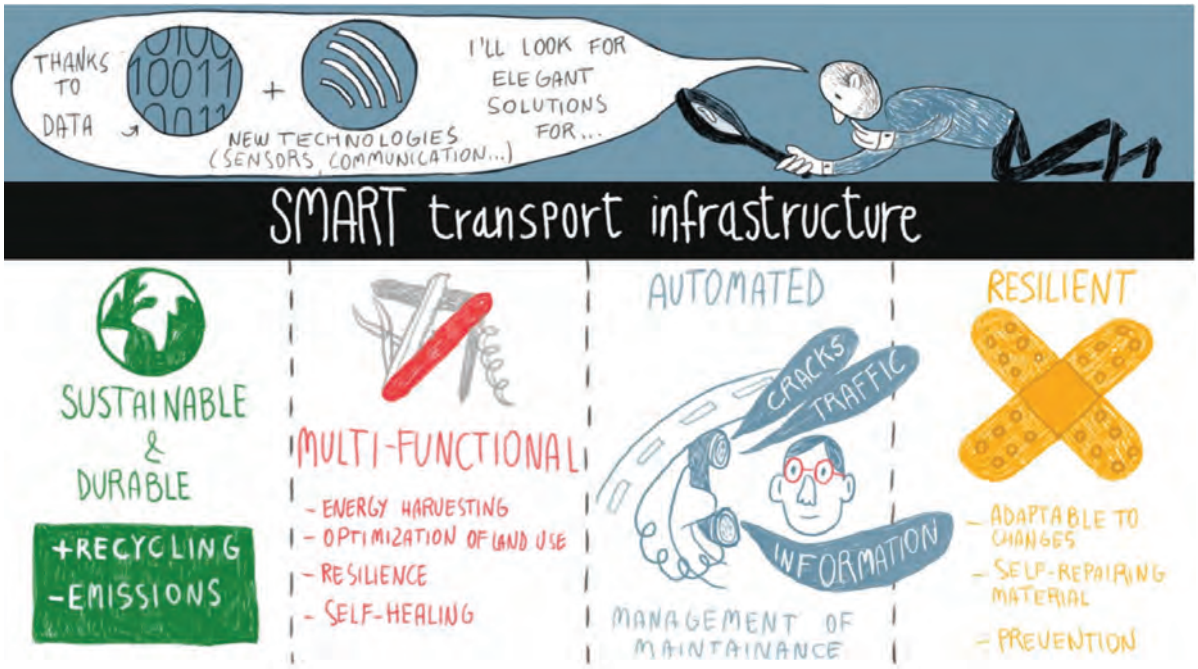


Figure 6. SMARTI vision

still need to exploit the results (INTRO[3], SEERP V and FOREVEROPEN[3]) or that are about to start INFRAVATION[4]. This vision is shaped around **four main pillars** representing at the same time the fundamental challenges for designers and features that transport infrastructures (TI) should acquire towards their smartening across Europe in new or existing roads, railways and airports, whether at local or at network level:

***S**ustainable:*	design to last, maximise recycling, minimize the impact
***M**ulti-functional*:	conceived not for transport purposes only and towards optimization of land use
***A**utomated:*	equipped with sensors to allow proactive communication towards a more intuitive use and a simplified management
***R**esilient:*	conceived to self-repair and be adaptable to changes due to natural and anthropogenic hazards

Sustainable transport infrastructure *are designed-to-last while maximizing the re-use and recycle of secondary materials and minimumising the impact on the planet*. It is not accidental that the first word of our vision for Smart infrastructure is the S of Sustainability. There is no doubt that providing solution for more sustainable built environment is the main challenge of our times, so big that it can be considered as the main container of all the other aspects.

In our definition of more sustainable transport infrastructure, two are the main goals that would need to be achieved: ***Design to last via maximising the re-use/recycle of secondary material*** in transport infrastructure is now starting to happen extensively and in several cases research efforts aimed at maximising the re-use of Reclaimed asphalt (RA) and recycled ballast (RB)[5][6]. Recycled tyre rubber (RTR) is another alternative material that is very resilient and can be reutilised in many transport-engineering applications such as rubberised asphalt, as anti-vibration mats under the ballast, as a component of rubber/plastic composite sleepers, as sleeper's rubber pads and also in the form of rubberised asphalt to create a sub-ballast. Other secondary materials such as steel slug, asphalt roofing shingles, waste ashes, etc., have already been studied and used in practical applications with environmental benefits, but again limitations and issues have also been highlighted. With regards to ***Minimising the impact on the planet,*** efforts are being made for reducing the energy consumed in the production of materials. For instance, low-temperature asphalt represents substantial energy savings and an associated mitigation of emissions, although more research is needed to draw definitive conclusions regarding their performances when coupled with high-amount of secondary materials. The Sus-

[3] FP7 – INTRO – co-operative systems for intelligent roads – http://intro.fehrl.org/

[4] INFRAVATION - ERA-NET Plus for road infrastructure innovation – INFRAVATION: An Infrastructure Innovation Programme – www.infravation.net

[5] CEDR – EranetRoad call 2012 – AllBack2Pave – http://allback2pave.fehrl.org

[6] FP7 PEOPLE Marie Curie ITN – SUP&R ITN: Sustainaibale Pavements & Railways – www.superitn.eu

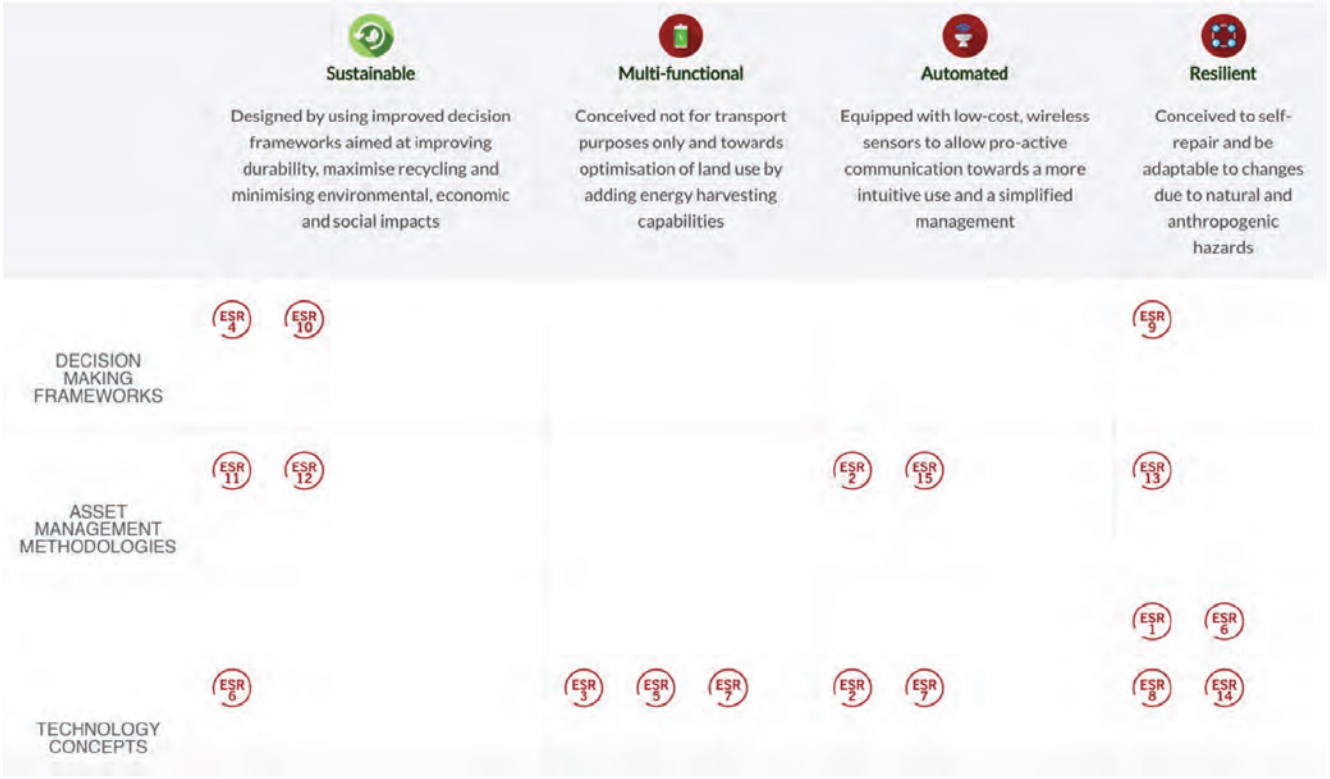


Figure 7. SMARTI Prototypes developed within the ETN

tainable transport infrastructure should also be conceived to reduce their impact in the environment and society by being economically-viable. In this sense, techniques such as Life-cycle Assessment (LCA), Life-Cycle Costing (LCC), Sustainability rating Systems (SRS) and the circular economy concept aim at the quantification of these impacts, however their use within the transport infrastructure field is still at the early stage.

Multi-Functional transport infrastructure *will be built having in mind the optimization of land use. Transport infrastructure occupancy of the planet earth surface is extraordinary-wide. This could be justified if these are conceived not-for transport use only but to serve population with other functionalities.* An immediate example is having infrastructure that can capture energy to power lighting, signage or even the electric vehicles themselves. This is a quite recent research area that focuses on avoiding energy losses within infrastructures by transforming it into other form of usable energy. Sources of energy at road and runway pavements and railway surfaces can primarily be identified as **solar radiation and geothermal energy** harvestable through embedded water pipes or circulating air. Furthermore, **mechanical energy** can be harvested from the stresses caused by traffic through piezoelectric sensors. These technologies have similar potential in airports, where runway could harvest energy for de-icing purposes, or in railway trackbeds where energy could be harvested, especially in remote areas for nearby lightening and signalling.

Automated transport infrastructure *are featured with the capability of monitoring its own condition, simplify its use and management.* With this system roads can be frequently monitored to detect changes in structural integrity that may not only foreshadow a future crack/distress manifestation but also allow for more accurate scheduling of preservation actions. There are basically two types of sensory and communications technology: *embedded sensors and moving vehicles/crowd monitoring*. **Embedded sensors** can be placed in specific places where is crucial having primary data and that can be interpolated to gain information related to other section of the infrastructure. Embedded systems must be sparse, low cost and can be either detachable/resuable and/or sacrified within the infrastructure. This information is then collected from asset managers via wireless readers that are either manually operated or mounted on a moving vehicle moving up to a speed. The managers can then analyse these data and interpret results to then produce their recommendations and maintenance strategies. Infrastructure monitoring with **moving vehicles** such as dedicated vehicles and/or crowd monitoring with in-service vehicles, or private user vehicles, should be implemented to allow a redundant sensing system from one side and transport of information on the other side. Especially with crowd monitoring, having an information that from the infrastructure reaches the vehicle and then reaches other vehicles (cooperative mobility) will enable the deployment of advanced (e.g. dynamic) guidance and management systems, tailored to respond

to in situ requirements, in effect improving the use of the infrastructure as well as the reliability and efficiency of the network management. With such a redundant sensing system, not only it is possible to improve the real-time management in transportation towards remote asset management, possibly also with the help of 5G communication, but it is also possible to improve design practices and introduce models for predictive maintenance based on real-time data. This type of systems can be developed for roads as well as for airports and railway applications.

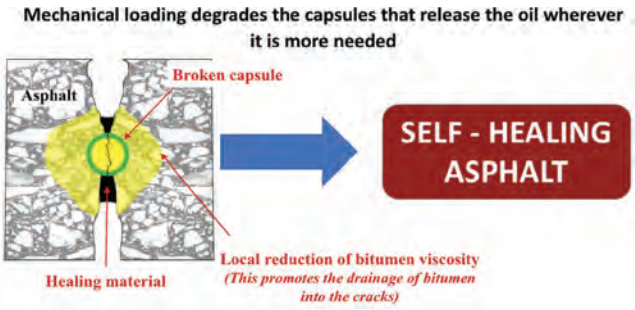


Figure 8. Self-healing asphalt mixtures with Encapsulated rejuvenators (ESR1)

Resilient transport Infrastructure *will be conceived to maintain services in face of natural and anthropogenic hazards and with the ability of fast adapting to change and eventually self-repairing*. Extreme natural or man-made events (e.g., floods, earthquakes, terrorist attacks) can have an enormous impact on the ability of a network to fulfil its function, reducing the connectivity and increasing travel time. The extension of transport networks in Europe makes it impossible to upgrade all assets (pavements, trackbeds, bridges, tunnels, etc.) **to sustain increased risks posed by climate changes, increased traffic growth or increased risk of terrorist attacks**. In order to make the best decisions on investments, it is paramount to define a consistent measure of the network resilience to these events, even considering that they can only be in very general manner. In this sense, development of decision-making frameworks, accurate traffic change prediction and technologies allowing fast adaption of infrastructures to change, are needed to protect our built environment and guarantee transport of goods and people.

4. SMARTI PROTOTYPES

To achieve the implementation of the SMARTI vision, multi-disciplinary research in the adoption/development of the smart solutions for smartening the transport infrastructure is needed at all levels ranging from materials to infrastructure components, to then going to even bigger scale from network level to cities/province/regions. Furthermore, researchers should promote innovations at any stage of the life-cycle of the infrastructure hence, from the design and planning, including material/product supply, to construction and use/management of the transport infrastructure asset. Along these lines and within some of these areas, the 15 early-stage researchers of the SMARTI ETN have developed smart solutions here identifies as "Prototypes". Each prototype has been conceived in form of a 1) technology, 2) asset management methodology or 3) novel decision making framework and the development has been carried targeting different pillars as in the following picture.

Within SMARTI ETN, 15 inter-cconected projects have been developed and in this paper the details of some individual and/or cluster project are reported to provide an idea of the research and development undertaken to smarten transport infrastructure. Hence, what follows is an overview of prototypes grouped in three levels: Materials, Infrastructure components, infrastructure network and or cities, carefelly chosen to have examples from roadways, railways, airports and also ports

4.1. MATERIALS: Self-healing asphalt mixtures with encapsulated rejuvenators (technology concept)

Description of the innovation: Encapsulated bitumen rejuvenators are capsules containing a bitumen rejuvenator able to soften bitumen and heal the cracks when they appear in the asphalt because of the fatigue effect of traffic. The capsules are dependent on aggregate gradation, therefore, there will be an optimised capsule design for every asphalt mix. The tests required to find the optimised conditions were also defined in this project. Figure 1 shows a 3D capsule image captured by Computed Scanning Technology.

Targeted infrastructure component: Asphalt roads.

Targeted SMARTI Pillar: Resilience.

Final TRL: 4

Future vision for implementation of the innovation in Europe: it is expected that in five years this technology will obtain the first results about its performance in an operational environment like road sections. Depending on these results obtained from large scale, this technology could be applied for specific types of roads or specifics parts of a road where the effect of the encapsulated bitumen rejuvenators is more needed. To apply this method in Europe, three steps are required:

1. System prototype demonstration in operational environment.
2. Costs of raw materials required in capsules production are expensive. However, the materials used in labs usually are more expensive than the materials used in the industry because of the high level of purity of them. Accordingly, it is necessary to reduce the costs of the raw materials and analyse the economic feasibility of this technology.

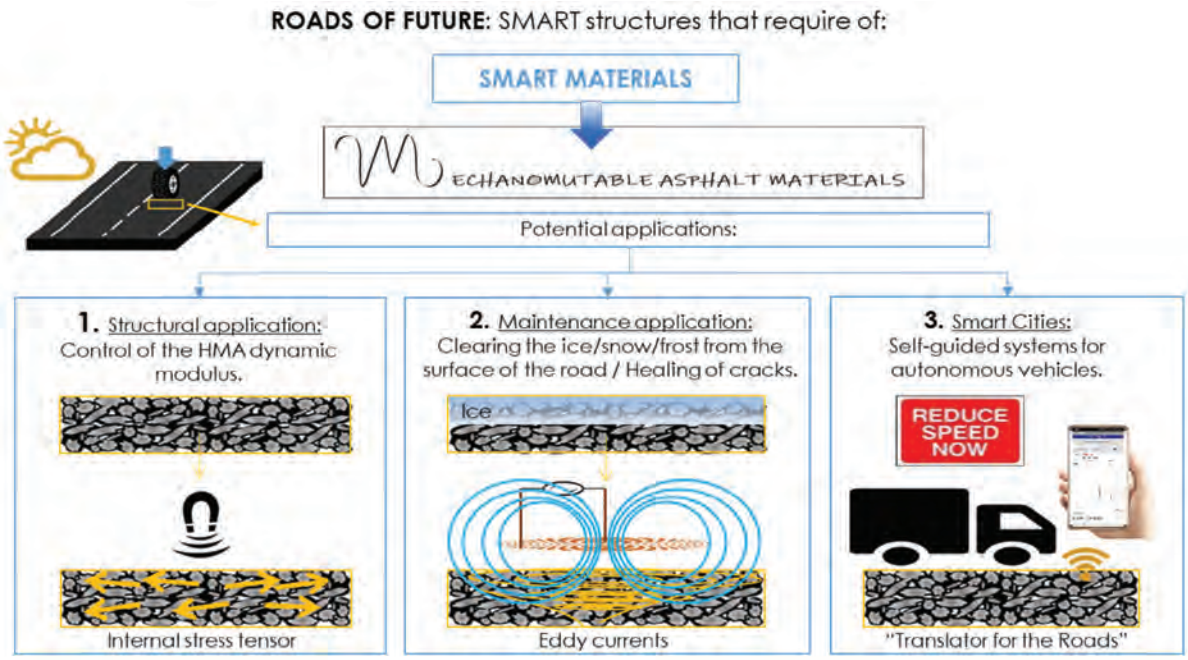


Figure 9. Mechanomutable Asphalt Materials (MAMs) for smart road pavement (ESR14)

3. Conduct Life Cycle Assessment (LCA) to evaluate the environmental impacts resulting from implementing this technology.

4.2. MATERIALS: "Mechanomutable" asphalt materials (technology concept)

Description of the innovation: This project was addressed to develop the Mechanomutable Asphalt Materials (MAMs). The MAMs are new and smart asphalt materials for flexible pavements. Therefore, MAMs can be part of roads, airports and ports. Specifically, during SMARTI project, MAMs had been studied to be used as alternatives to clear the ice/snow from the surface of the road, to enhance the healing of asphalt materials, to improve the performance of asphalt materials during service and to assign a role to the roads inside the industry of the autonomous vehicles. Therefore, after SMARTI, MAMs could be mainly thought to be used in cities and highways.

MAMs are composed of a bituminous matrix with magnetically responsive materials activated by the effect of external magnetic fields. Understanding that asphalt materials modified with magnetically responsive materials have a capacity to emit magnetic field signals, the term MAMs was extended to encoded asphalt materials. In this sense, various dosages and types of magnetically responsive materials were used to produce different magnetic field signals easily to be detected by magnetic field sensors, with the aim to evaluate the possibility to give a role to the road in the guidance of the autonomous vehicles. Through the encoded roads could be possible to increase the reliability of implementation of the autonomous vehicles, increasing the road safety, reducing the traffic congestions, offering systems which can continue working despite the climate conditions and which can include elderly and disabled people in the use of the transportation infrastructure. Considering also that, time-varying magnetic fields can produce temperature increases in the asphalt materials modified with electrically responsive materials via the Joule effect, the term MAMs was again extended to electroconductive asphalt materials with two capacities: (1) enhance the healing capacities of the asphalt materials that composes them and (2) increase the superficial temperature of the flexible pavements in order to clear the ice/snow from the surface of the road.

There are three main potential applications of MAMs in the construction of smart pavements: (1) improving the performance of the materials of the pavements, (2) providing the encoded roads required in the implementation of the self-guided vehicles and traffic management systems and (3) helping in the maintenance and rehabilitation of the roads to ensure the road safety and service of the pavements. Each one of the mentioned applications were studied at laboratory scale in the 14th SMARTI ETN project.

Future vision for implementation of the innovation in Europe: MAMs is still a technology under development, therefore, at this moment is still a project under research. In this sense, by the moment is focused to research centres, but it is expected that soon, and after more research, it could be also addressed to be used by public agencies and industry. For a future implementation of MAMs in Europe, it could be required:

1. Complementary laboratory tests at mixture scale. The results obtained until now should be deeply studied.

2. Real scale tests experiments, in order to validate and complete the guidelines required for a proper use in real projects.

From the ESR's point of view, in 2025 the research efforts related with this project could have been reached enough advance to define a proper guideline for the use of MAMs in each one of the studied applications. Maybe some real scale projects could be under execution or almost finished. In 2040, maybe the use of the technology will have been included in the standards and protocols of daily used in the construction and maintenance of transportation infrastructure.

4.3. INFRASTRUCTURE COMPONENT: Railway Component Monitoring Model (decision-making framework)

Description of the innovation: In this project, a predictive tool based on the Multi-Criteria Decision-Making approach has been developed; the tool is formatted in Excel software and it is also programmed in a Python code. This project is a collaborative research involving AECOM and the University of Nottingham. Since the first month, the earthwork field has been narrowed down considering for the analysis just embankments. Instability of these assets is notoriously more complicated to be detected before failure than for cuttings. Then, the project is attempting to reinforce the asset management failure control, linking the causes of embankment instability (geotechnical parameters highlighted through the literature review) to visible/detectable signs of instability (track geometry parameters).

A deep literature review study allowed to divide the parameters responsible of embankment instability for three main macro categories: Geotechnical issues, external factors and soil mechanics properties. All the parameters play a role in the embankment instability; nevertheless, the chosen ones need to address the same criteria:

- Scientific evidence: information gathered has to be proved by scientific researches published in recent decades. The search for relevant publications has to be performed using Web-based scientific search engines, databases and electronic libraries.
- Availability: data collected for the study needs to be openly available in a public repository that issues datasets, and/or generated at a central, large-scale facility, available upon request. If data is subjected to third party restrictions, the availability needs to be previously agreed. Data storage and how data can be reached must be well defined at an early stage.
- Data Coverage: data collected and involved in the study must be available for all the sites evaluated. If one of the datasets is unavailable for one asset, then the parameter has to be discarded from the analysis.
- Measurability: the way in which parameters are measured/evaluated must be clear, objective and unbiased. The clarity is especially needed to be addressed when qualitative parameters are involved.
- Updatability: data considered must be routinely collected and so updatable in order to repeat the analysis in the future.

This narrowed down the parameters for the backward analysis to four: type of soil, seasonal deformation, rainfall, vegetation. For the analysis, the parameter showing the best correlation with the track movement is the type of soil, while the rainfall doesn't show a strong effect on the position of the track. This is unexpected as the water content is considered one of the principal causes of instability.

The main steps for this method are:

- Construction of the hierarchy containing 'goals' (as in the AHP literature), the alternatives for reaching it, and the criteria for evaluating the alternatives.
- Establish priorities among the factors - pairwise comparisons of the factors (intensity of importance).
- Check the consistency of the judgments.
- Reach a solution.

At first the model was shaped in Python. It was decided though to build the model in Excel, so to make it user friendly and to enlarge the accessibility to any decision makers regardless competence and knowledge. The ongoing work (very last step) consists on the model validation, examining new assets. A period of secondment in SNCF (Paris) was supposed to happen from March 2020 till May 2020 with the purpose of collecting data coming from a different railway culture and so further test the model. Unluckily, due to the pandemic, this plan was cancelled, and the model is going to be tested with data coming from more embankment assets all around the UK.

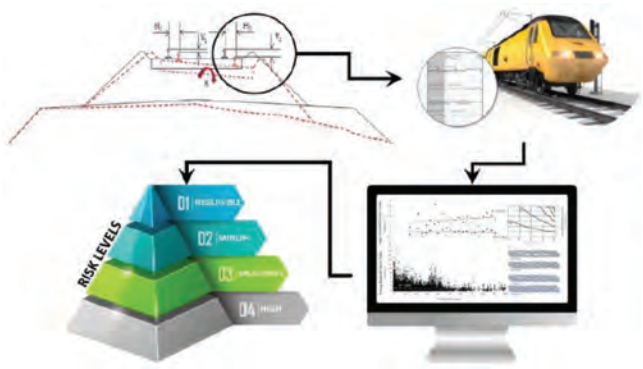


Figure 10. A sketch explaining the developed model (ESR4)

Targeted infrastructure component: railway embankments.

Targeted SMARTi Pillar: Sustainable.

Final TRL: 3

Future vision for implementation of the innovation in Europe: To implement the project, it would be interesting to see how the model behave when inputs are collected and analysed in a different way, following standards and procedures not common in the United Kingdom. Moreover, considering four parameters plus the track geometry is a good point to start from; it would be interesting to find a way to enlarge the range of indicators to be included in the decision-making process.

4.4. INFRASTRUCTURE COMPONENT: Remote structural health monitoring for railway trackbeds (technology concept)

Description of the innovation: In this project, a gauge to monitor deformations in railway tracks will be developed. The gauge will be a smart sensor (piezo electric) to collect vibrations from the rail track continuously to help monitor the deterioration of the track in real time. Collected data will be transferred by wireless means to a server where it can be analysed and used to predict failures before they occur on tracks. An algorithm is to be developed to be used for track maintenance.

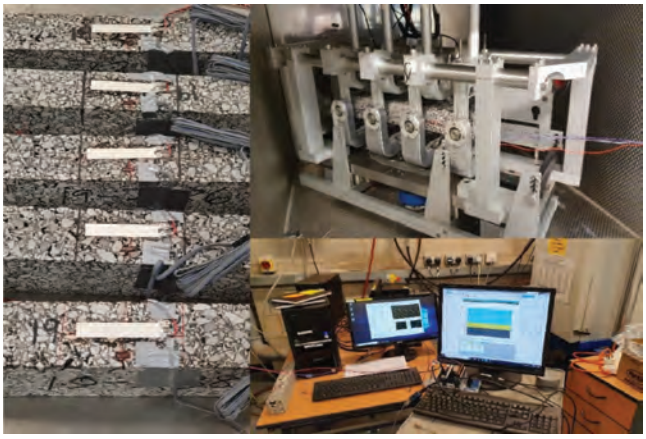


Figure 11. The developed gauge being tested and validated using 4PB fatigue test (ESR5)

Targeted infrastructure component: Railways

Targeted SMARTi Pillar: Multifunctional and Automated

Final TRL: 4 to 5

Future vision for implementation of the innovation in Europe: This technology needs to be tried first in laboratories using different designs and calibrated to see how it will perform in a real track. Data analysis, an optimal gauge should then be tried on rail lines where data could be collected remotely. However, since most lines are already built, the gauge must be inserted through an already built track, hence measures will be developed to be able to achieve that. This gauge should be used in newly constructed tracks and track renewals to monitor subgrade compressive strains. Hopefully this low-cost sensor would be an alternative or replacement for costly track stiffness measurement using Falling Weight Deflectometers, particularly where there are failures and where improvements are needed to provide optimum and cost-effective gauge to collect and transmit deterioration data. With this type of sensing system, a novel track monitoring model for predictive maintainance can be developed using the data collected.

4.5. INFRASTRUCTURE COMPONENT: Hybrid pavement for solar energy harvesting

Description of the innovation: Energy harvesting is the process by which energy is captured and stored exploiting an external source. In the field of road engineering, the main exploitable energy source is the solar radiation. The road energy harvesting systems are able to convert the sunlight into electricity thanks to the solar cells placed under a semi-transparent layer (photovoltaic roads), or they can convert the solar energy in thermal heat by means of solar thermal systems. The thermal gradient of the pavement can be exploited by thermoelectric generators (TEGs), by heat pipes or by heat-transfer fluids (i.e. water) pumped into a medium (asphalt solar collectors, porous layer or air conduits). This prototype is given by the union between a photovoltaic road and a porous medium. The "hybrid road" is a pavement able to generate electricity and, at the same time, harvest energy by means of water pumped in the porous layer.

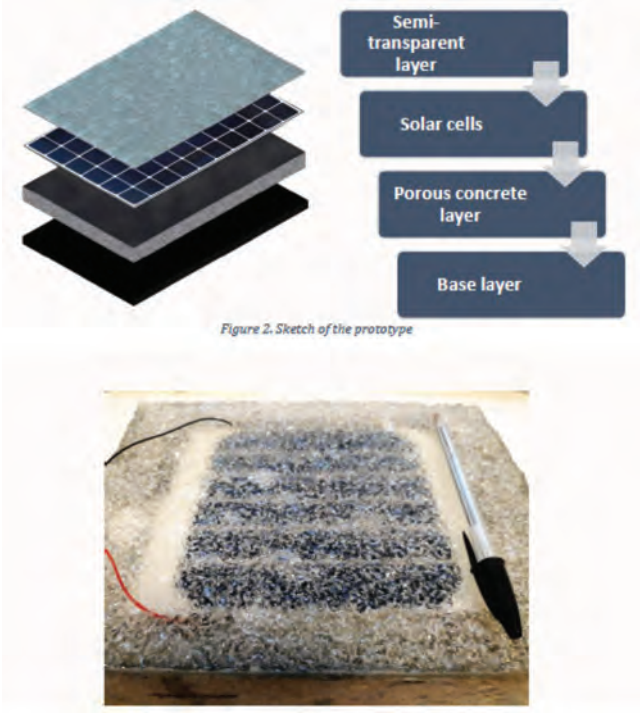


Figure 12. The developed prototype with a detail of the transparent layer (ESR3)

Targeted infrastructure component: Roads, Airport pavements

Targeted SMARTi Pillar: Multifunctional.

Final TRL: 3

Future vision for implementation of the innovation in Europe: The first step for the implementation of the "hybrid road RA2ROAD" is the identification of certain areas, well exposed to the solar radiation, where the use of existing surfaces for energy harvesting could be advantageous. The second step is the construction of a full-scale prototype, its validation in an operational environment and an accurate cost-benefit analysis. In Europe, some pilot applications could be bike paths or sidewalks, which don't require high mechanical performances. Going further, the system could be produced in prefabricated slabs, which could be easily replaced in case of breakage

4.6. INFRASTRUCTURE COMPONENT: Automated road pavements (technology concepts and asset manageemnt)

Description of the innovation: Pavement instrumentation topic is the most wide research topic in the field of pavement health monitoring. Hence, within SMARTI ETN three projects (ESR2, ESR7 and ESR15) worked very closely analyzing different solutions towards an automated monitoring of road pavements. In particular:

- ESR2 worked on nano-scale strain gauges
- ESR7 worked on piezo-floating gate sensors
- ESR15 worked with strain gauges and geophones

The picture below illustrates the vision of automated road pavement defined in SMARTI. This paragraph though will present only the details of the ESR15 project.

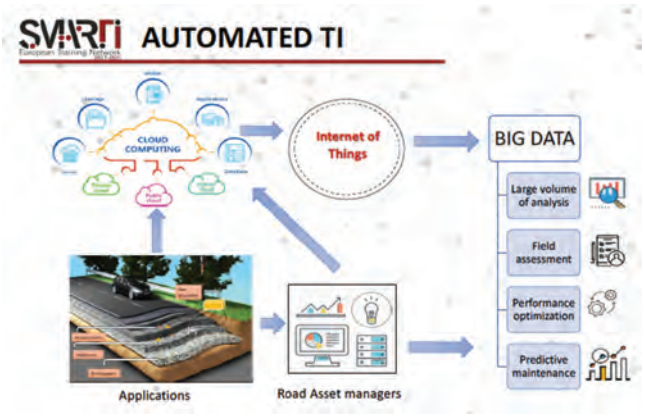


Figure 13. SMARTI Automated road pavement vision (ESR2,7,15)

ESR15 project focused on different procedures and technologies to analyze and calculate the deformations of the pavement structure. There are certain non-destructive tools to measure the response of the pavement continuously, which is achieved by using relatively small embeddable sensor technology. This would allow the measurements to be recorded without the disruption of the traffic. Horizontal and vertical strain sensors and temperature probes are the most commonly used devices to instrument in the pavement and determine the pavement conditions. Some applications have also used piezoelectric sensors to monitor the fatigue damage of pavements and calculate stresses and strains in pavement layers. Furthermore, new types of sensors like geophones and accelerometers are designed and deployed to measure the inertial vibrations of the pavement and used in research studies to compute the deflection basin under the passage of wheel loading. An array of geophones have also been used in laboratories to measure the inertial vibrations in the pavement under a known vehicle loading. A mobile load simulator and back calculation tool are used to calculate the responses and measure the pavement deteriorations. Accelerometers and deflectometer are also used for measuring the pavement deflections.

In this project, a nondestructive analysis of the roads is done with the sensors to continuously monitor the pavement. An efficient method is developed to process and improve data retrieved from the sensors. This is an effective method for the calculations of pavement properties with the sensors.

The continuous non-destructive testing of the pavement is an efficient method for analysis of the roads. This is achieved successfully with sensors that are deployable even after the construction of the pavement. The solution consists of sensors embedded in the pavement, measuring the overall deflection of the pavement under real traffic and a wireless data acquisition system transmitting the data to the road manager. In this study, two types of geophones and accelerometers are used to measure the vertical displacement during the load passage.

The first objective of this study is to evaluate the possibility of using the geophones and accelerometers. Their respective measurements are converted into vertical displacements to find the deflection basin induced in the pavement during the passage of vehicle. For this purpose, effective signal processing tools are developed to improve the signal of the sensor and quantify the deflections that are caused in the pavement. Another aspect is interpreting the data, which is achieved by some back-calculation tools.

The sensors that are studied in this work are two types of geophones and accelerometers that measure the vertical displacement under the wheel loading. In order to test the technology first the sensors are selected and tested under three different test settings:

1. In the laboratory under a hydraulic vibrating table.
2. In the accelerated pavement test facility with different loads and speed (representing the actual load and speed of the vehicle).
3. On a real motorway road and the measurements made under real heavy traffic.

In the first calibration tests, simulation of the pavement deflection is replicated using a vibrating table and applying different amplitudes and velocities. The data received from the sensor is treated in order to improve the accuracy. During this step, a processing method is developed and applied to the data received from the sensors, mainly to convert the physical values of these sensors into the vertical displacements and enhancing the accuracy of the signal. The procedure consists of a passing the data signal through a filter and integrating and amplifying the signal and applying the Hilbert transform to get an accurate displacement signal. Two main aspects were observed that:

- The processing method was depended on the speed and load of the vehicles
- Each sensor had different amplification factor

An accelerated pavement testing facility is available in IFSTTAR, France where the sensors are instrumented 1cm from the top layer of the pavement. During the accelerated tests, these sensors are subjected to different loads and the speed of the vehicle and different wheel position with respect to position of the sensors. The processing method developed in the laboratory tests data is applied to the data recieved from the APT tests. The displacement data is also used to back-calculate the layer moduli of the pavement. Here, a scheme is developed for the reverse calculation using the data from geophones and accelerometers. Finally, the same procedure of signal processing is applied to the data received from the instrumented geophones at the Ax motorway in France. With the instrumentation of the sensors, we were able to achieve the following:

- Better results of the deflection values from the sensors
- Accurately identifying the trucks type
- Identifying the speed
- Also possibility of the estimating the load of the trucks
- Reverse calculation of the layer moduli

Targeted infrastructure component: Roads/Highways.

Targeted SMARTi pillar: Multifunctional: because the sensors used are able to identify more than one application that includes the speed, vertical deflection and the type of the load. In addition, the wireless technology is expandable and can include other type of sensors. Also, Automated: because the system of wireless monitoring directly sends the data to the data bases of the manager without physically being present at the site.

Final TRL: 6

Future vision for implementation of the innovation in Europe: The next development steps include mostly the enhancement and polishing of the existing technology of continuous monitoring. Some of the future steps are as follows:

- More experimental setup for the instrumentation of the road with the developed technology and monitoring the issues that comes with it.
- The inclusion of the different types of the sensors to the wireless data acquisition system and the enhancement of the processing techniques for the different sensors to be used with the wireless data acquisition system.
- Complete wireless transmission among the sensors and the data acquisition system and improving the effective data transfer.
- Improving the durability of the system and making it more resilient against the environment.
- Making the technology cost effective.

It is expected that by 2025, the implementation of instrumentation with various application and automation of the data processing on system, and by 2040, multifunctionality together with automation and continuous health monitoring.

4.7. INFRASTRUCTURE NETWORK: Optimising the dynamic structural evaluation of airport pavements (asset management)

Description of the innovation: For analyzing high frequency continuous measurements, a forward model needs to be fast to predict the deflection of pavements. Therefore, tools like FEM could not be used; an alternative formulation was searched in the literature. This formulation was however not implemented. Therefore, in this project this implementation was carried out and validated to test its result.

A hypothesis was setup. A RWD equipment that works for flexible pavement is heavy enough to generate significant response that will be picked up by the measuring sensors at hand. But rigid pavements are stiffer, therefore in this research more focus on modelling aspect was attended to. Ideal experiments with stationary sensors show the potential of the technology.

By developing the implementation of the model formulation via Matlab and Mathematica, an extremely complex and long solution was automatically generated and thus was ready to predict the deflection of jointed rigid pavement, after which the validation of implementation was carried out numerically via open source FEM software. An experimental validation was carried out independently where the measurement data and the structural parameters were provided by a Research institute in France known as STAC. The measurements came from an instrumented rigid pavement database. This validation was successful and thus the next step was to setup experiments to do measure the load transfer. This was carried out by preparing test setups by using geophones, line laser sensors customized for the project application. The implementation is useful for continuous evaluation technology to be used for rigid pavements, be it on airports or highways. Only jointed plain concrete can be evaluated by use of this model. This model can replace FEM solution when

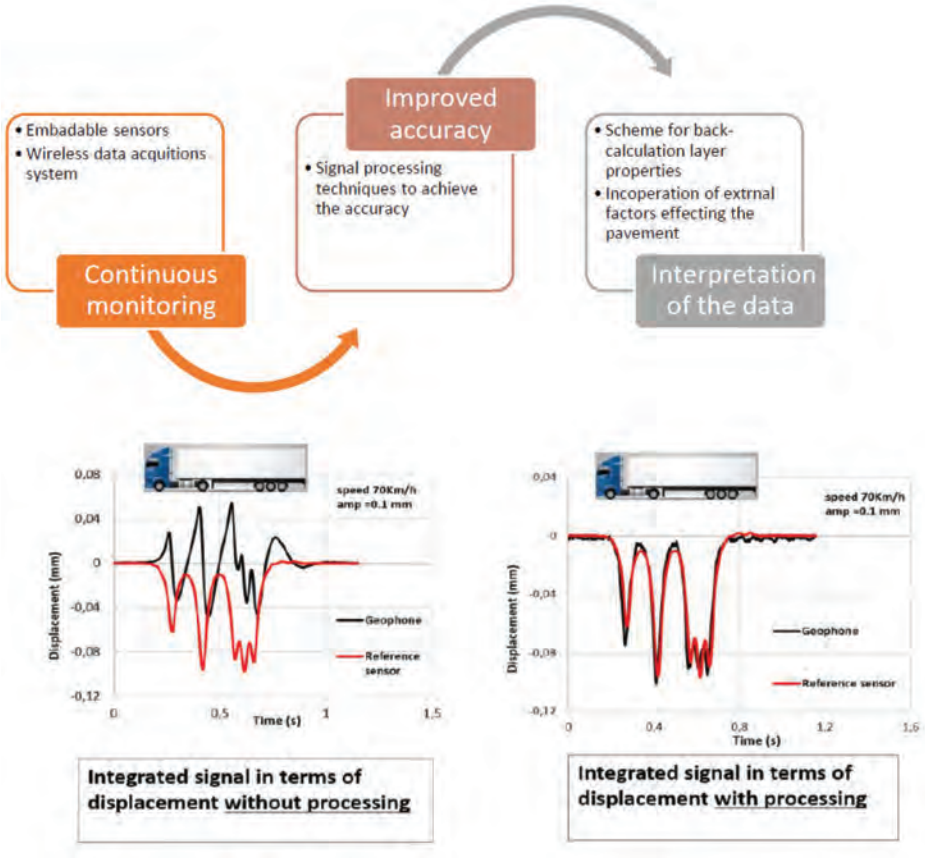


Figure 14. Improving accuracy of data mining thorugh geophones (ESR15)

computing resources are limited or an inverse analysis is needed.

Name of the innovation: Implementation of a 3D semi analytical forward model to assist in measurement of load transfer of jointed concrete pavement by continuous measurement technologies.

Targeted infrastructure component: Roads/Highways and airports where rigid pavement infrastructure exists.

Targeted SMARTi pillar: Automated.

Final TRL: 5

Future vision for implementation of the innovation in Europe: This project focused on a specific application of technology that is not popular or underutilized by industrial partners. Therefore, it may not be implemented in nowadays. In future, it is possible that better structural assessment of pavement will be automated. In this scenario, this project can help that automation for this particular application of technology for rigid pavements.

Generally, to implement the developed method in Europe, the following steps are required:

1. Further experimentation with Industrial Partners who are developing the RWD/ own an RWD device.
2. Modification in device/RWD technology based on previous step.
3. Test the improvement and apply the research in this project on a network level.

4.8. INFRASTRUCTURE NETWORK/CITIES: Resilient road networks (decision-making framework)

Description of the innovation: Ths project focuses on developing indicators & tools to predict, assess and reduce the impact of disruptive events on society Transport models, representing the complex real-world transport and land use systems, provide a powerful tool to predict transport systems performance under different conditions and assess the impact of transport infrastructure options. These transport models are thus essential to plan, design and manage transport networks and land use. However, transport planning, modelling and design standards currently give priority to transport efficiency (i.e. the efficient movement of vehicles through a transport network under normal conditions) rather than to resilience (i.e. the movement of vehicles under disrupted conditions). The rising interest on the concept of resilience and the research on road network resilience should lead to the incorporation of "resilience thinking" into transport planning, modelling and design standards.

As the main function of road networks is to provide mobility (i.e. connect trip origins and destinations in a timely manner) the resilience of road networks should characterise their ability to deliver this function under a multitude of potentially unpredictable conditions. Indeed, the

Figure 15. Optimising the dynamic structural evaluation of airport pavements (ESR11)

resilience concept accepts the possibility that a wide range of disruptive events may occur but are not necessarily predictable, and focus on anticipating and enhancing the performance of the disrupted system rather than preventing or mitigating the loss of assets.

However, assessments of the potential impact of disruptive events on road networks generally focus on a few scenarios illustrating the proposed methodology or the worst-case scenario using a game theory approach. Hence, the studies available in the literature do not uncover the potential impact of all possible disruption scenarios as suggested by the resilience paradigm. This is unfortunate as such findings could be used by practitioners and public authorities for understanding the potential impact of a wide range of scenarios. For example, considering that a wide range of events (including pavement maintenance, road accidents, and flooding) can unpredictably disrupt any parts of the network, the identification of multiple worst-case scenarios rather than the single worst-case scenario would be more useful in practice.

Similarly, most studies focus on a few case studies such that their findings may only be valid for the specific networks studied. Further investigations are thus required to evaluate the effectiveness of these approaches and the generality of their findings. The project was conducted to develop a resilience assessment framework to support decision-makers in predicting, assessing, and ultimately reducing the impact of disruptive events on road networks. To this end, the research carried out to develop this framework aimed at answering the following questions: What makes a road network resilient? How should we manage road networks for greater resilience?

To reach this aim, four research objectives were considered:

- develop more accurate and meaningful methods for quantifying road network resilience.
- test and validate these methods on multiple case studies including abstract and real road networks.
- use the case studies to explore the effects of network characteristics including size, topology, and demand distribution on network resilience.
- use the case studies to explore the effects of link-repair strategies on network resilience.

In contrast to the studies available in the literature, the research project considered the impact of multiple potential disruption scenarios and a variety of case studies to provide a clearer understanding of the generality of the results and conclusions.

Figure 16. Indicators & tools to increase road network resilience to disruptive events (ESR9)

Targeted infrastructure component: Road networks and cities.

Targeted SMARTi Pillar: Resilience.

Final TRL: 2

Future vision for implementation of the innovation in Europe: The proposed guidelines and modelling tools need to be adapted to a commercial setting and validated in a relevant environment. To this end, potential users (transport operators) should be approached to get relevant data and feedback in order to test and refine the guidelines and modelling tools. Moreover, the proposed guidelines and modelling tools should be further developed to reach TRL 9 and more importantly meet the needs of the potential users. Similar research and innovation projects should be carried out to reach this aim. This effort should lead to the incorporation of the "resilience thinking" approach into transport planning, modelling and design standards in both Europe and developing countries.

4.9. INFRASTRUCTURE NETWORK/CITIES: Sustainable Road Management through Low-cost techniques (asset managemnt)

Description of the innovation: This project is using innovative and Low-cost technologies/techniques for road pavement analysis. Structure-from-Motion (SfM) techniques are used for pavement distress identification and therapy exploiting image-based 3D modelling techniques for pattern recognition and subsequent analysis. Imagery has been collected using low-cost equipment such as cell phones and low-cost cameras and drones which help reduce costs associated with conditional surveys.

The project also utilizes deep learning models for detecting pavement distresses using refined labels based on the environmental context of the situation. In addition to this, the project has aimed to utilize limited historical databases by leveraging flexible data analytical methods such as gradient boosting algorithms and deep learning methods to predict future intervention occurrences and locations. Combined these are being used to optimize aspects of pavement management strategies for urban road networks.

There were a few technical issues encountered during the project. These included the following:

1. Data collection – waiting time to get data from the municipality and also lack of available data for some aspects of the maintenance pipelines, solutions had to be sought to use smaller data sets and different data sets.
2. Equipment limitations – waiting times for procuring equipment such as a drone to carry out surveys; solutions had to be sought using different cameras.
3. Survey limitations – quarantine period impacted the ability to carry out a lot of surveys in different environmental contexts; secondment period had to rely on data previously recorded.

Accordingly, the solutions explored during the project can be broken into three specific sections. The first of these is the development of models and workflows to understand the environmental/area context of the region under analysis to predict the likely locations of future maintenance interventions. To do this, varying algorithms were explored such as gradient boosting and tabular deep learning models. The second aspect was on creating a workflow to carry out a hotspot analysis on the specific roads in the network which are considered critical as identified from the first part. To do this, deep learning object detection models were considered using imagery from smartphones and webcams attached/equipped to vehicles traversing the road networks. Finally, the third part was on creating workflows and optimizing techniques to generate 3D models of the pavement distresses on the network and both specific points in the network using point and shoot cameras and over a longer section using drone imagery. Additionally, once the models were generated, segmentation techniques were explored to isolate the distresses, measure them and analyze them to feed the information to the asset management system and to make decisions on what type of therapy is required. Once again, considering, the three aspects of the project, the results obtained can also be considered using these three sections. Firstly, a prediction model was produced utilizing readily available historical data from the municipality and statistical offices in Palermo and Italy. The model was formed to identify the next series of streets in the municipality where interventions should take place and also the time interval. The model was also validated using historical data. Secondly, the project was able to produce detection models that could identify the location of specific distresses along a road network.

These models were environmentally specific to urban road networks in Sicily and secondary roads in France. Using these models surveys can be run using smartphones in vehicles to obtain a list of distresses occurring along a road or network and the specific location of the labelled distresses. This would give a practitioner an idea of the condition of the road and also the critical points along the road where attention is required. Finally, the project was able to produce workflows that can replicate 3D models of pavement distresses. The project has been able to authenticate the accuracy of the techniques for pavement distress identification and analysis by carrying out numerous surveys on distressed pavements using different phones, cameras and drone imagery and comparing results to laser technologies and high-quality camera imagery. Additionally, segmentation models were developed to isolate the distresses within the 3D model for analysis and understanding of the shape and structure of the distresses.

Targeted infrastructure component: Roads/Highways

Targeted SMARTi pillar: Automated

Final TRL: 5-6

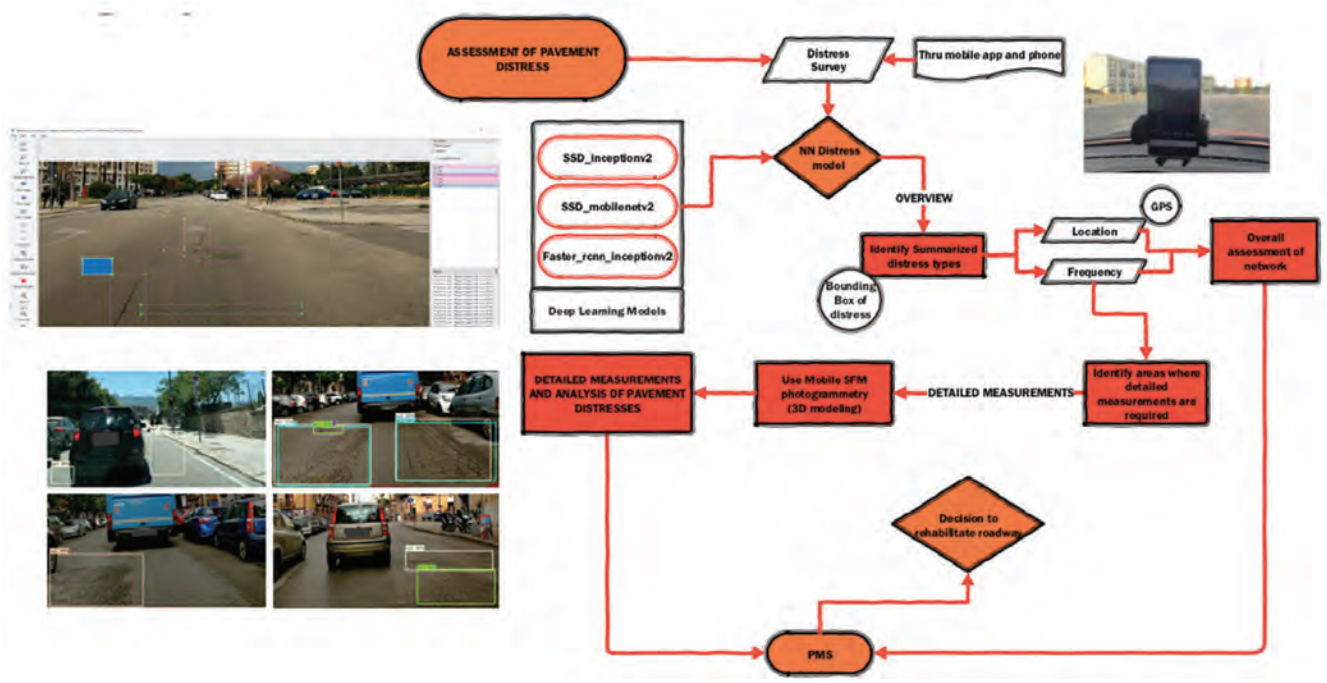


Figure 17. Framework and results for 2D imagery (ESR12)

Future vision for implementation of the innovation in Europe: Concerning the majority of the techniques and tools developed under the project, the tools are country and more specifically environmentally sensitive and specific and therefore the direct application of the tools or models in this phase should only be in the area or city of the case study. The workflows developed, however, could be utilized to generate new models and systems for other countries in Europe and this is important for future studies and other implementations. The context of the application is very important and thereby any future implementation needs to understand the 'business case' before any application is done. For future development even in the region of application (Sicily), there needs to be a larger engagement with the municipality and more surveys are needed before commercialization or implementation is done as there needs to be a high level of confidence in the developments.

For the future vision of the project, it is difficult to project what will happen in 5 years, much less 20. This is because the field of artificial intelligence and image modelling is constantly updated and advanced. In the last 2 years alone, the advances have been vast and it would be impossible to predict that far in the future as it is likely that new modelling techniques, equipment and tools might become available that could supersede what has been done to date. That being said, the workflows developed should still provide guides on some aspects of practical implementation and the advice of being environmentally context-specific should remain key regardless of whatever new technologies arise. With the workflows in place on how to manage the data and what questions are needed to be asked, those can further be adapted to provide newer guidelines with the more advanced tools that are expected to be available in the future. Additionally, should the engagement continue with the municipality beyond the project then it is possible and hopeful that models could be more robust and reliable and by 2025 a tool is implemented within their system.

5. CONCLUSIONS

Policy makers identify the smart infrastrctures as a must-to-happen, a sort of panacea of issues related to the built environment and our societies. However, whenever it is asked "what is a smart transport infrastructure", colleagues between academia and industry often provide different replies or different prospectives. This is an historical period characterized by fast and drastic climate changes and exponentially increased traffic. The existing infrastructures have not been designed to adapt to those, so they have experienced a fast rate of deterioration. As a result, extensive and expensive on-site surveys and maintenance interventions are needed to allow people and goods to travel. Europe needs investing in research of innovative solutions but even more on the formation of professionals able to envision the smartening of the infrastructures and exploit it through the modernisation of design, construction and management practices for the coming years. In response to this urgent need that a consortium of 28 partners, 18 of which from industry (9 SMEs) from 8 European countries, USA, Asia and Australia had the un-missable opportunity of a 48 months timeframe for learning-by-doing, discussing definitions, opportunities and barriers on the strategically important topic of

Smart Infrastrcture, while training 15 talented graduates versed towards the smartening of transport infrastructures in Europe and beyond.This was possible thanks to the great framework of the format offered by the H2020 Marie Curie European Training Networks (nowadays called Doctoral Networks) that allowed to form the SMARTI ETN a joint academy-industry training-through-research programme aimed at conceiving and promoting the paradigm shift to Sustainable Multi-functional Automated Resilient Transport Infrastructures.This platform allowed performing cutting-edge research and training in an international, multi-disciplinary and multi-sectorial environment, that led to define what is presented in this paper as:

- *"SMARTI Principles" towards cultural change and improved research and development*
  - Smart solutions for transport Infrastructure can be conceived by acquiring multi-disciplinary competences through international and collaborative research
  - SMART (elegant) solutions should be engineered at different levels and applied with a life-cycle thinking approach.
  - Development and Implementation of smart solutions can be facilitated by following the technology readiness levels (TRLs)
- *"SMARTI Vision" with definition and features of smart transport infrastructure*
  - *Sustainable TI:* design to last, maximise recycling, minimize the impact
  - *Multi-functional TI*: conceived not for transport purposes only and towards optimization of land use
  - *Automated TI:* equipped with sensors to allow proactive communication towards a more intuitive use and a simplified management
  - *Resilient TI:* conceived to self-repair and be adaptable to changes due to natural and anthropogenic hazards.
- *"SMARTI Prototypes" for the smartening of roadways, railwasy and airport*
  - Technology concepts
  - Asset management methodologies
  - Decision making frameworks

As a result, Europe can now count on a new generation of highly-skilled and appealing researchers, professionals able to envision the smartening of Transport Infrastructures and prepared for a career both in industry and academia. The joint supervision of scientists from academia and from industry has been crucial for the professional development of the fellows as well as for carry out the research. The training programme combined in thematic taught modules, consisting of core, specialist and transferable skills, and secondments in industry and/or academia, were fundamental to deliver all the necessary skills in a short space of time, providing a rounded education to the fellows. This type of education has shown to be of fundamental importance towards filling the gap of skills shortages in the field of smart transport infrastructure. In fact, the whole SMARTI ETN programme, as well as this paper, aim at stimulating a cultural change amongst academia and stakeholders of the transport infrastructure industry addressing the need to form engineers versed with a multi-disciplinary approach, able to work within a collaborative international environment and, above all, knowledgable of the main challenges of our times.

Acknowledgements

A special thanks goes to all the colleagues that have worked within the SMARTI ETN by producing scientific products as well as helping with the management and/or organization of the thematic training. Furthermore, the fellows must be acknowledged not only for the support in the writing of this paper but mainly for the hard work profused in 4 years.

The SMARTI ETN project has received funding from the European Union's Horizon 2020 Programme under the Marie Curie-Skłodowska actions for research, technological development and demonstration, under grant n.721493.

Milan Veljković

Milan Veljković is Chair Professor of Steel and Composite Structures at Delft University of Technology, Netherlands. He has coordinated more than 50 research projects and partnered in 11 international projects, with a total budget of more than 12M EUR. He is a member of 4 editorial boards of international journals.

Contact information: e-mail: m.veljkovic@tudelft.nl

https://orcid.org/0000-0003-4791-2341

Prof. Veljković started his academic career as an Assistant at the University of Sarajevo, Department of Civil Engineering (1985-1991). He worked as a Research Engineer at the Institute for Structures and Materials, University in Sarajevo (1987-1990). After that, he was a Research Engineer (1991-1993) at Luleå University of Technology, Sweden. At the same university he attended PhD studies during 1994-1996. After receiving a PhD from Luleå University of Technology, he held several academic positions: Lecturer (1997-2003), Senior Lecturer (2003-2007) and Chair Professor (2007-2015). In 2015 he was elected as a Chair Professor of Steel and Composite Structures at Delft University of Technology. He is a visiting professor at the University of Belgrade.

Beside teaching activities, he held following positions: Member of the board of Civil Engineering Department (2000-2002), Director of a research school (2003-2005), Director of Centre for Risk analysis and Risk Management (2005-2015) and Director of Program for Fire Protection Engineers (2005-2014) at Luleå University of Technology. He was a Member of Board of Department of Structural Engineering (2015-2017) at TU Delft, and since 2018 he is Member of Board of Department Engineering Structures.

The main focus of Prof. Veljković's research is related to: connections/joints in buildings, bridges and supporting structures for wind turbines, development of technical aspects for the implementation of the circularity in construction, prediction methods for behaviour of connections in bridges and offshore structures exposed to fatigue loading, material modelling, experimental and numerical analysis, 3D printing using WAAM (Wire and Arc Additive Manufacturing) for construction sector.

Prof. Veljković actively participates as a member of technical committees, including: Swedish Technical Committee for Standardisation, Dutch Technical Committee for Standardisation, European Standardization committee CEN250/SC3 Steel Structures, Standardization committee CEN250/SC4 Composite Structures. During 2015-2018, he was a Chairman of Project Team for the development of the 2nd generation of Eurocodes, Part EN1993-1-8.

He coordinated more than 50 research projects and partnered in 11 international projects, with a total budget of more than 12M EUR. He was a member of management committees of 4 COST Actions and Evaluator of research projects of Sweden, Czech Republic, Norway, Belgium, Slovenia and a couple of funding schemes of European Union.

Prof. Veljković published more than 300 articles in international journals and conference proceedings. His papers are more than 2200 times cited ($h = 24$). He is a member of 4 editorial boards of international journals. For his outstanding contributions, Prof. Veljković received Sigge Thernwalls Stora Byggpris in 2015 for his contribution for development of education and research at LTU, Sweden and TC12 Risk Analysis and Safety of Large Structures and components, 2021, Robert Moskovic Award, in the recognition of an outstanding contribution to Integrity of Steel Structures.

Emerging technologies for future steel structures

Milan Veljkovic

Delft University of Technology, Netherlands

Summary

Three areas and corresponding projects relevant for future of development of steel structures are addressed: reusable steel-concrete composite structures, innovative bolted connection in offshore supporting structure for wind turbines and material model of 3D printed plate made using Wire Arch Additive Manufacturing process. The final goal of all three projects is to improve sustainability by reducing CO_2 equivalent emission and reduce costs for the market implementation of innovations. For two firstly mention projects the goal is feasible while the third area is not sufficiently mature for the market implementation. None of these projects would be possible without advanced Finite Element Method and successful validation by experiments at different scales: material experiments, component and full-scale experiments. Emerging technologies, some of which are addressed in the paper, clearly illustrate increasing needs and opportunities for highly educated engineers.

Keywords: reusable composite floor system, wedge connection, 3D printed carbon steel, WAAM, GMAW

1. INTRODUCTION

The text about emerging technologies contains personal reflections based on the experience of creating recently started research projects at the research group of Steel and Composite Structures, TU Delft. A large majority of financially demanding research projects are awarded if societal relevance and/or industrial objectives are addressed. Financing from a funding organisation, either national or at the European level, are then possible. In some exceptional cases, a company is interested in contributing to such research projects. But in that case, the hourly rate of the staff is higher.

Three projects based on technological innovations supported by the experimental program and followed by the use of advanced prediction methods are provided below to illustrate an implementation of new technologies and justification of societal needs.

The most recently finished PhD project which have been financed by RFCS and RWO has focused on the reusable composite floor system. The main innovation of this project is the use of relatively large and heavy concrete decks and oversized holes for the connection of steel beam and concrete deck by a demountable connector. Buildings and construction sector have a big potential for improvement of sustainability because they "accounted for 36% of final energy use and 39% of energy and process-related carbon dioxide (CO_2) emissions recorded in 2018, 11% of which resulted from the manufacturing of building materials and products such as steel, cement and glass" (World Green Building Council, 2021).

An increase of CO_2 equivalent in the construction were identified from 2017 to 2018, driven by population expansions (World Green Building Council, 2021).

Innovations of connections in supporting structures for offshore wind turbines are driven by a new generation of wind turbines that have a capacity beyond 10 MW. Some of the existing technical solutions, such as L-flange connection, are technically not feasible for challenging diameters of monopile structures. Competitiveness of supporting structures is important to accomplish goals in the energy transition and climate objectives of the Netherlands and EU. One of such projects is related to use of a patented wedge connection which enables the assembling of monopiles (tubular tower) 8 m in diameter. The solution which is shown below is already used in a test tower for a wind turbine made by one world-leading turbine producer. The second offshore project provides an alternative to jacket structures, making them lighter by introducing the wrapped composite joint using FRP stripes. Such a solution utilises bonding as the main load transferring mechanism instead of welding of circular hollow sections. Therefore, an adverse effect of welding by creating notches and residual stress in joints, which strongly influence the relatively low fatigue detail category, is prevented. This project will not be addressed in the paper because of the limited space (and time) available, but a reference is made to a spin-off company in charge of implementing innovation (TREE COMPOSITES, 2021).

Above mentioned projects have accomplished a patent right owned by a TUD, by a company or by both parties.

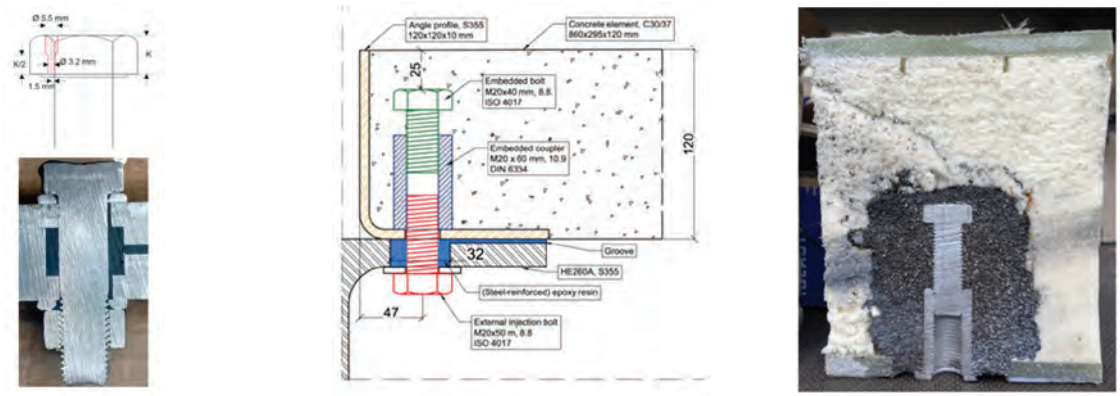
The use of WAAM (Wire Arch Additive Manufacturing) in construction sector is still in an early phase of development. An iconic bridge was successfully manufactured in 3MXD (MX3D, 2021) and assembled in Amsterdam in 2021, demonstrating the potential of 3D printing technology. Industrial application is constantly upscaling, for example, by manufacturing a large crane hook. The most recent crane hook has a loading capacity of 350t (Huisman, 2021). However, a search for a ground-breaking product that will improve the sustainability of construction sector is continuing worldwide, including a start-up project in our group. The investigation is supported by a leading steel producer through ArcelorMittal Global R&D Bars & Wires. For the time being, only material data are publicly available. A short overview of the material tests, basic material data for design and for advanced FEA modelling were derived on specimens fabricated by 3D robots using carbon wires in GMAW technology.

2. REUSABLE COMPOSITE STRUCTURES

Easy execution and deconstruction, the potential for reuse of components and structures, and the infinite number of recycling are the main characteristics of steel structures. However, structures made only of steel are not possible, except in industrial applications. In many other applications of buildings and bridges steel must be combined with concrete, timber or FRP, to establish a composite action and assure effective use of building materials. Welded headed studs are traditionally used in steel-concrete composite structures. Intensive research in the area of composite steel-concrete structures is spread over the world, for example, Australia (Lee and Bradford, 2013), USA (Kwon, 2008), UK (Lam *et al.*, 2013), Luxemburg (Kozma, 2020; Kozma *et al.*, 2019), Serbia (Pavlović, 2013). One of the main objectives in this research is focused on the replacement of welded studs by pretension bolts or by other demountable connectors. These efforts are seen afterwards as a good way to facilitate the transition of the construction sector to a circular business model primarily enabling easy deconstruction of composite structures. One of the research objectives in recently finished European projects REDUCE (Reuse and demountability using steel structures and the circular economy, RFCS Project No. 710040) and national project INTERMOD (Two-scale modelling of composite “steel-concrete- reinforced resin” interaction, NWO Project T16045). The main objective of these projects was to gain knowledge in using demountable shear connectors embedded in large and heavy prefabricated concrete decks. Decks are connected to steel beams using “hand tightened” external bolts, The holes in the beam flange were oversized. More than 5 times larger clearance between the bolt and the bolt hole was required compared to requirements in traditional steel structures. The clearance was filled by injected two-component resin. An additional study was made even using reinforced (by steel shots) injected resin.

Injected bolts are in used in the Netherlands for decades and especially applied in renovation of riveted or bolt structures and with a normal clearance hole, see Figure 1a. The basic rules for such application are provided in EN1993-1-8:2005 (CEN, 2005).

A new function was allocated to injected bolts in the design of the reusable steel-concrete composite structure and therefore, new knowledge was necessary. Combination of two building materials in a structural component leads to a new requirement to tolerances for the bolt clearance and to a new design of injected bolts. The main motivation for such design requirements was to allow for geometric imperfections of all structural members made of different building materials. Rapid execution of heavy floor decks and easy demounting were additional requirements that call for larger tolerances in reusable structures. The most suitable design was accomplished using a coupler with a bolt embedded in the concrete deck. The deck was connected to the beam by external bolts having a hole



(a) Traditional **(b)** Injected bolt in confined condition **(c)** Semi-confined condition

Figure 1. Injected bolts in steel-steel connection, steel-concrete and steel-FRP composite structures

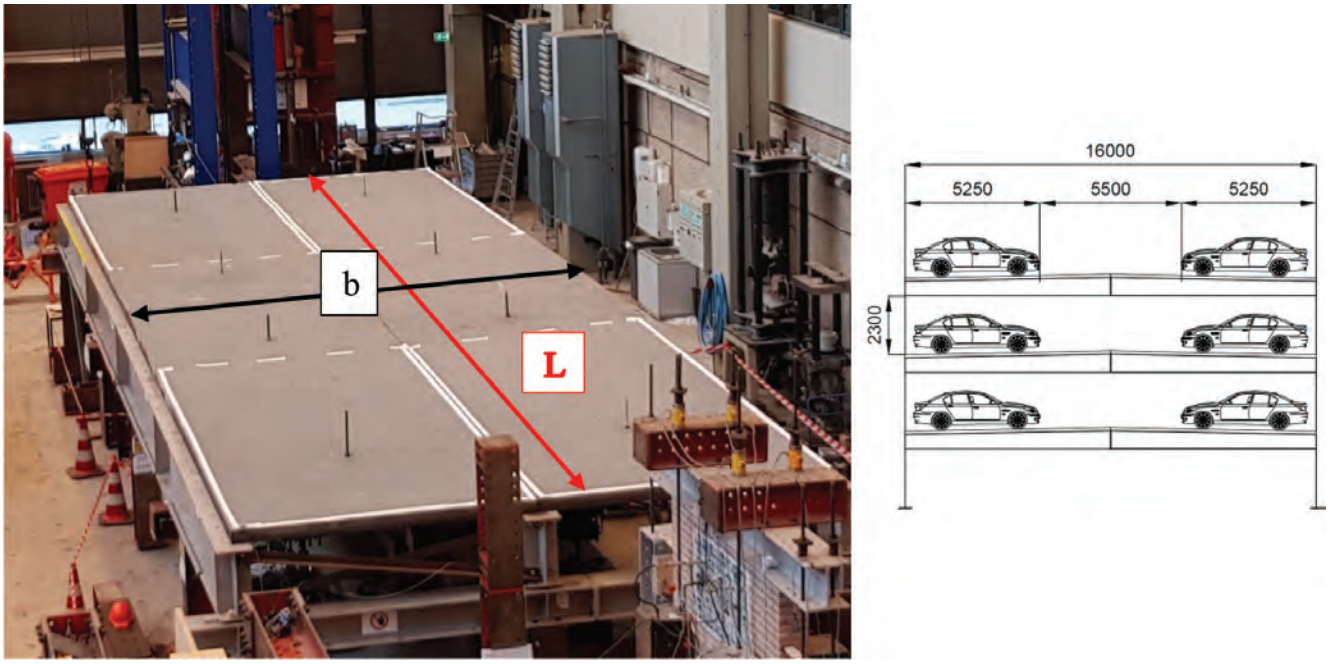


Figure 2. Specimens of composite floor system to check the feasibility of executing and reuse

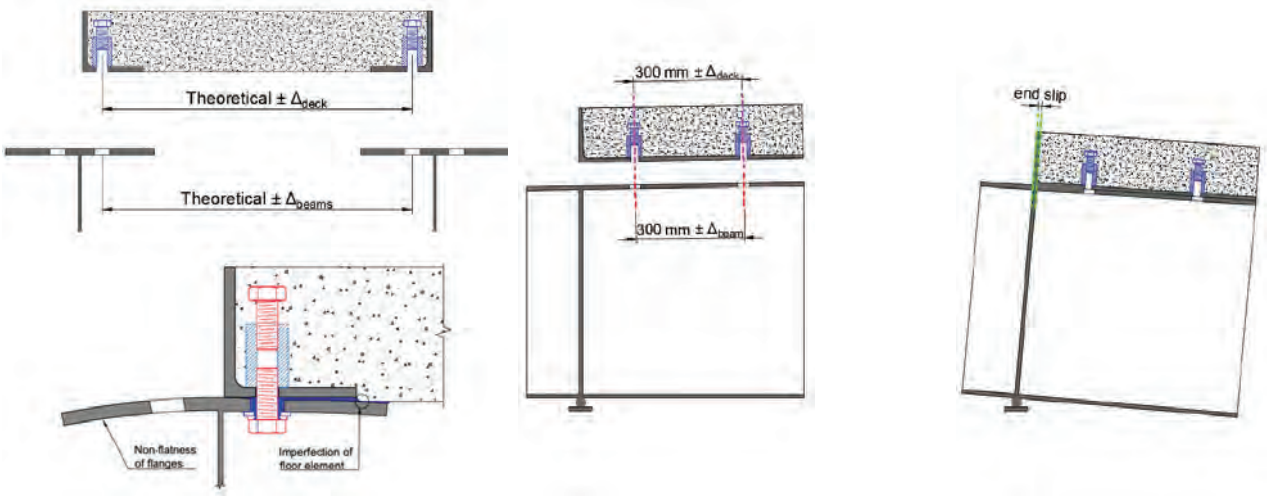
in the head for injection of oversized clearance. A case study car park building using 4 concrete decks of 6 tonne each was analysed in (almost) full scale in the laboratory, see Figure 2, $b = 5,2$ m and $L = 14,4$ m. The span of steel beam in the car park building is 16 m, but that was not possible to place in crowded Stevin2 lab. The same solution for the connector in reusable composite steel-concrete prefabricated deck floor system was proved to be sufficient even for composite slabs using the same test set-up shown in Figure 2. However, mechanical properties were thoroughly investigated only for the prefabricated solid deck, see the main findings below (Nijgh, 2021).

A resin injected bolt coupler connector was used in the full-scale experiments. It consists of three main components: (i) a bolt and coupler embedded in the concrete deck, (ii) an external injection bolt that has a hole in the head and (iii) injection material. Two components epoxy resin mix SW404 and HY2404 was injected through a hole in the bolt head to fill the complete bolt-hole clearance. The coupler was made of higher steel grade (10.9) compared to the external bolt to make it more resistant in the case of overloading. This is a precaution to ensures the reusability of the concrete decks and to minimise the creation of waste. The embedded bolt provides the pull-out resistance of the connector, whereas the coupler transfers the shear force from the deck to the external bolt connected to the steel beam via the injected material. The steel-concrete floor system provides full confinement to the injected material.

However, in the case of steel-FRP composite deck, the shear connector and injected material are not fully confined, see Figure 1c. The connection between the FRP deck and steel beam was a technical challenge before this connection was tested. A lack of generic and feasible connection technology between FRP decks to steel beams was a hinder to the use of such structural components. A clear demand for a light and durable deck was identified by RWS (Ministry of Infrastructure and Water Management) for use in movable bridges. Preliminary results (Csillag and Pavlović, 2021; Olivier *et al.*, 2021) provided indications that such a light-weight movable bridge might be technically possible.

The size of hole clearance is influenced by imperfections from the manufacturing process of steel beams and concrete deck, execution tolerances, deformations during the execution of the floor system, i.e. the interface slip, but also caused by the requirements on the speed of construction, see Figure 3. The maximum measured hole spacing deviation between the flange and couplers was 2 mm. The end slip of 1.45 mm after the fourth slab was installed was predicted by FEA (Gîrbacea, 2018). The hole clearance of 12 mm was estimated necessary for easy assembling and disassembling based on the lab measurements of specimens delivered for the feasibility study.

Numerous assembly attempts in the laboratory confirmed that an increase of construction tolerances greatly reduced the execution time. This will certainly hold at the construction site where the frame execution tolerances (European Committee for Standardization, 2011) should be added. Due to the oversized holes, the bolt shank was not in direct contact with the hole wall after the execution. Without additional measures, the composite action would be accomplished only when the uneven gaps are closed by interface slip. An obvious solution was to inject the bolt clearance by resin injected through the bolt head to achieve composite behaviour. The overhead injection



(a) cross-section tolerances **(b)** longitudinal tolerances **(c)** execution deformation

Figure 3. Imperfections and deformations during execution of the heavy floor system

was easily performed by unskilled workers (students) at heights up to 2.5 meters using a readily available manual caulking gun at a pace of 30 seconds per bolt see Figure 4c. The epoxy resin fully filled the hole clearance with minor air inclusions as shown in Figure 4d. To prevent the adhesion of the resin to the steel beams and bolts, a release agent was successfully applied Figure 4e.

Mechanical properties of the connector were established in push-out experiments. Details of this investigation are available in (Nijgh, 2021). Representative results of the load-slip curves are provided below to illustrate the big potential of the connector. Figure 5 shows the relation between the nominally applied force per connector P and the averaged connection slip s for the resin and steel-reinforced resin injected specimens, see Figure 6.

The most relevant properties are summarised below:

- The experimental results demonstrated that a similar resistance Pu was obtained if the injected material (resin or reinforced resin) because the confinement was so strong even for the clearance of 12 mm that the shear resistance of the injection bolt was governing.
- The secant stiffness at 0.4Pu of the steel-reinforced resin-injected specimens was on average 80% higher compared to that of the resin-injected specimens. Frictional and adhesive effects caused the scattering of nominally identical specimens, but experimental results qualitatively demonstrated the benefits of implementing injection bolts in demountable shear connectors. Finite element analysis more precise predicted the beneficial effects of (steel-reinforced) resin-injected connections showing the stiffness for the non-injected connection (k_{sc} = 30.5 kN/mm) compared to the secant stiffness of the resin-injected (k_{sc} = 82.1 kN/mm) and steel-reinforced resin-injected specimens (k_{sc} = 102.7 kN/mm (Nijgh, 2021).
- The experimental results of push-out tests showed two subgroups of specimens (Nijgh, 2021), (i) those that failed after moderate slip (5-8 mm) and those that failed after substantial slip (> 15 mm), irrespective of the injectant. A sensitivity FE study led to the hypothesis that poor concrete vibration led to a lower strength in the region around the shear connector and contributed to the deviation between the two groups (Nijgh, 2021).
- An experimental study of different types of connectors in push-out specimens consisting of steel beam and FRP deck is shown in Csillag and Pavlović (2021) and Olivier *et al.* (2021). The specimens resembled push-out tests used for steel-concrete composite structures. The fully-confined shear connectors demonstrated a ductile behaviour, while the semi-confined shear connectors did not accomplish the ductility requirement according to EN 1994-1-1 (Lee and Bradford, 2013) provisions. For

(a) hole clearance **(b)** drilled head **(c)** injection equipment **(d)** filled gap **(e)** cylinder

Figure 4. Installation process of injected bolts from the hole clearance to the disassembling

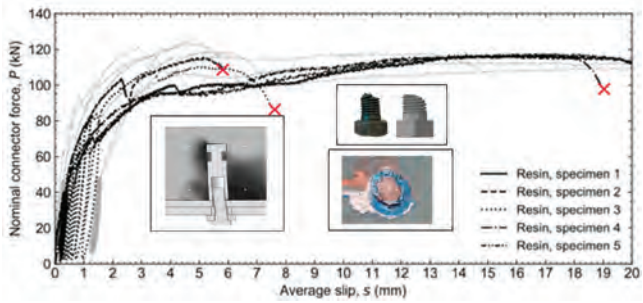


Figure 5. Load-slip results of pushout experiments using resin and the failure mode

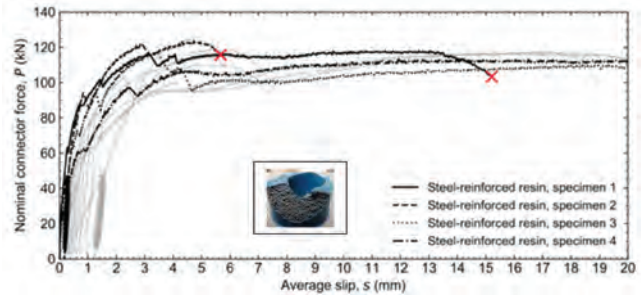


Figure 6. Load-slip results of pushout experiments using steel-reinforced

the specimens with FRP decks and "partially-preloaded" bolts, 25% lower secant stiffness at 0,7 Pu was experimentally observed compared to the specimens with concrete deck and non-preloaded bolts.

3. WEDGE CONNECTION, AN INNOVATION FOR OFFSHORE WIND TOWERS

Offshore supporting structures for wind turbines typically feature several types of connections: grouted connections, L-flange bolted connection, slip and double slip connections are most commonly used.

The grouted connection has a troubled past. The first grouted connections were based on the traditional straight grouted connection. The space in between tower segments was filled with grout (a cementitious material). Upon hardening of the material, this joined the two tower segments. In 2009 it came to light that due to deformations, the grout started to degrade, and very expensive repair works had to be carried out to prevent long term instability. This mainly happened with the early offshore wind farms with 2 and 3 MW turbines.

Following the earlier design flaws, the grouted connection was upgraded to a design whereby the grout is kept in place between two conical tubular segments, although other designs involving so-called shear keys are also used. This connection is now certified according to DNV-OS-J101, a design standard for foundation design.

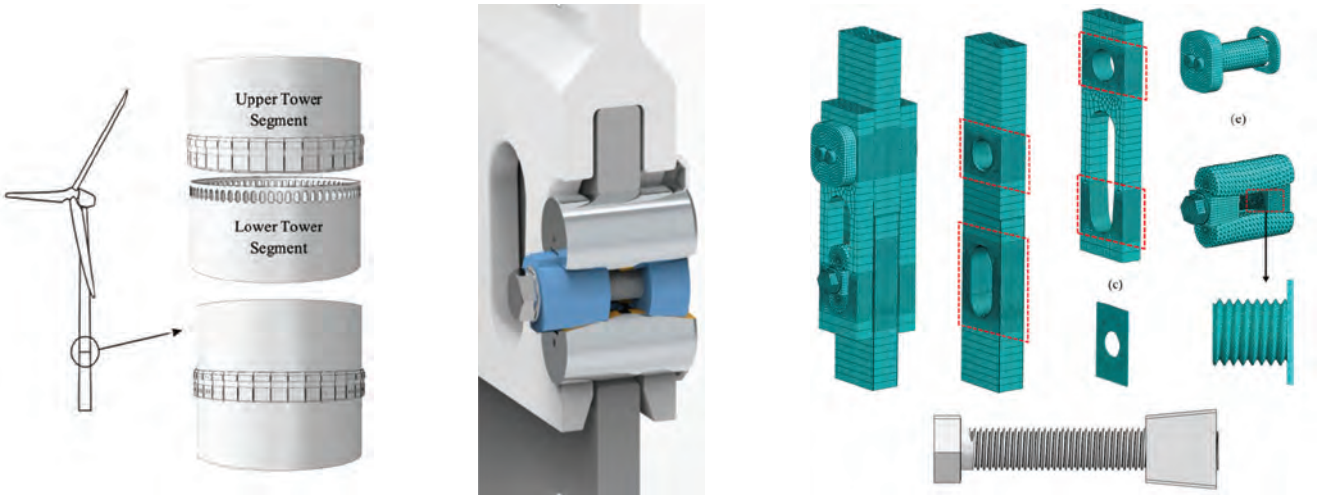
The grouted connection requires significant waiting time, up to 24 hours, for the grout to cure, and it reaches its full strength, typically after a month. The connection also requires sensitive and expensive grout seals (> 15 kEUR). Due to the fact that it takes time for the grout to reach its strength, it is relatively difficult to use it underwater (pile to sleeve connection) or in the splash zone (transition piece and monopile), since the waves exert loads on the pile immediately upon installation, thus requiring expensive fixation measures during curing. Due to the overlapping steel (typically 1.5 times the diameter in length), the price of the grouted connection goes up exponentially with increasing the diameter.

Bolts in the L-flange connection require a predefined level of pretension to be maintained during the lifetime of the connection. Bolts have various drawbacks, but most important is that the standard bolted L-flange connections have reached the limit of their capacity and cannot be used for the next generation WTG's. Bolted connections are also not well suited for floating installation which requires a fast and safe instant connection.

Other concepts, such as the (double) slip joint are difficult to produce and expensive due to a large amount of steel needed. In order to simplify the installation while increasing speed and safety, the industry was urgently looking for better solutions. Recent research of C1 Wedge Connection^{TM} (C1CONNECTIONS, 2021), jointly done by the company and the research group, is showing promising results primarily because the transfer of forces from an upper to a lower tower segment is centrical and well-controlled by the robust wedge principle. A trend of the connection development is depicted in Figure 7 where simplification of the connection is shown by integrating the pin connection with the upper segment of a tower. Such design allows easy installation, especially important for offshore structures where "every minute saved" in the execution process reduces the tower costs importantly.

The wedge connection has substantial advantages compare to L-ring flange connection, illustrated by following findings:

- The same ultimate resistance, about 810 kN is predicted by FEA performed on a segment specimen of the L flange connection and the wedge connection shown in Figure 8a and Figure 8b, respectively. The same "tower" thickness of 40 mm is used in both specimens, but the bolt size of M18 steel grade 8.8 is used to pretension the wedge connection instead of M42 grade 10.9, which is necessary in the L-flange connection. Two machined L-flanges of 125 mm thickness are required because the bolts are positioned eccentrically. Such thickness is governed by the design requirement of the pretensioned bolts. Very stiff flanges should be kept in contact because of structural and environmental conditions. Tight tolerances are imposed on the flatness of the flanges.
- A ductile failure of the lower tower segment is governing failure mode, and brittle failure of the bolt governs

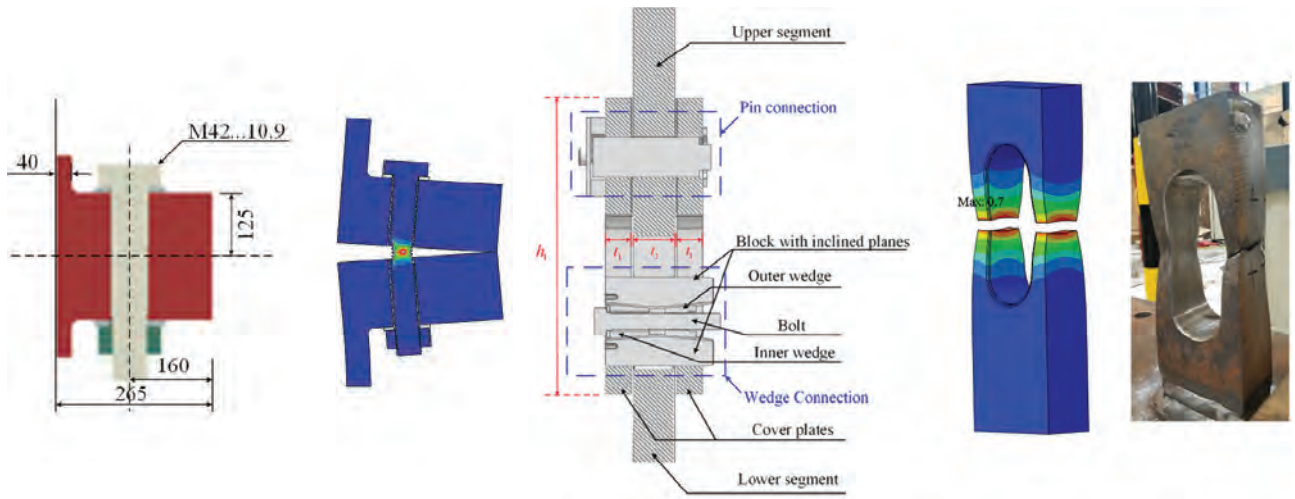


(a) First generation **(b)** Third generation prototype **(c)** FE model for a strength and feasibility analysis

Figure 7. Development of the wedge connection and FEA model used in the research

the resistance of the wedge and L-flange connection, respectively. These critical sections govern the failure mode at the fatigue limit state.

- Preliminary fatigue experiments of the first generation of the wedge connection have shown at least 5 detail categories higher fatigue resistance compared to the fatigue class of bolts. This is not a surprise having in mind the fatigue of the based material in an oval hole of the wedge connection. Machining of the hole influence the fatigue resistance of the connection. A stress range in the bolts of the L-flange segment is much smaller than stress ranged in the tower segment as long as the L-flanges are in contact. A gap opening of the flanges leads to an increase in the bolt force. Understanding the failure mechanisms indicates the importance of maintenance and checking if an untightening of the bolts is present.
- Another benefit that is clearly on the side of the wedge connection is sensitivity to un-tightening. The bolt of the wedge connection is tightened by rotating the nut for 34 radians, while the tightening of the L-flange connections is accomplished by rotating the nut by 2 radians. This means that small untightening of the bolts in L-flange connection leading to a much higher loss of the pretension and eventually to the gap opening if the monitoring program is not appropriate.
- The wedge connection technology is expected to be made market-ready in about 2 years when the results, including detailed investigations of fabrications tolerances, will be accomplished. The connection is already tested in a down scale and even used in a full-scale test project already.
- Costs of L-ring flange connection and the wedge connections are compared for the 7,5m diameter for MT-



(a) L-flange connection, FEA **(b)** Wedge connection, FEA and experiment

Figure 8. FEA and experiments of L-flange connection and the wedge connection

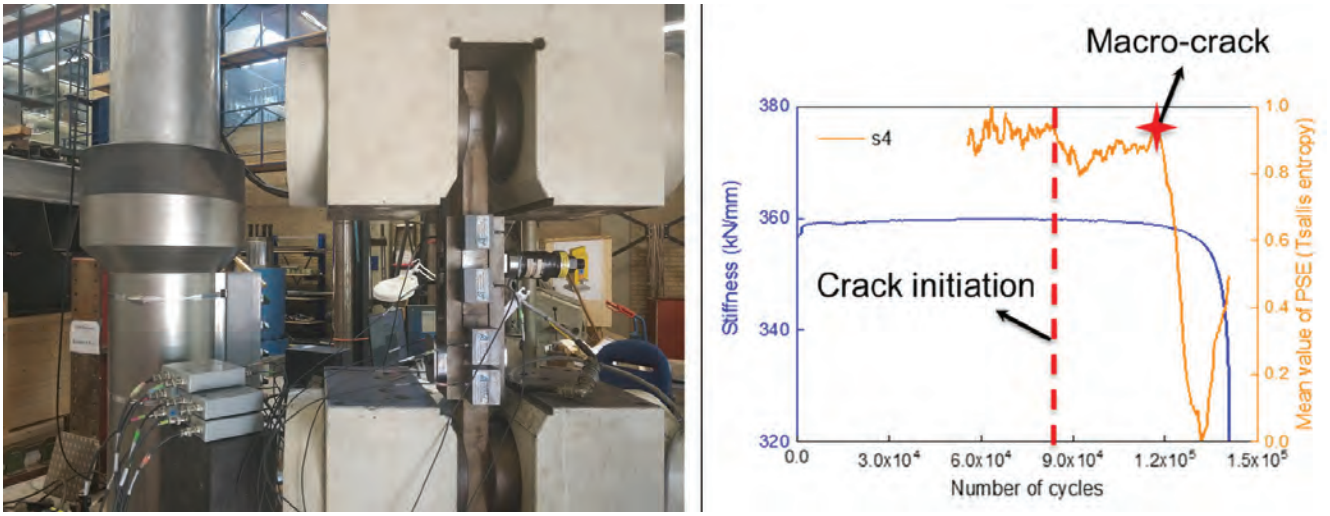


Figure 9. AE sensors attached to the specimen and indication of the crack initiation

TP connection suitable for the next generation of wind turbines. It is expected that at this given load level, the ring flange connection will not be technically feasible. An indication of the costs collected for the fabrication of the connection alone shows that the wedge connection cost is about one half of the L-ring flange connection cost. These figures are based on discussions with monopile suppliers, contractors, and utilities. The costs of fabrication of the connection alone are considered.

Development of the C1 Wedge Connection^{TM} is driven by offshore market opportunities.

The international market for offshore wind is expected to grow to 10 GW annually by 2026. A realistic assumption is that this growth is sustained until 2030.

Such a positive forecast for the market opportunity is opening possibilities for further research even in new areas with unpredicted success. One of such areas is the use of acoustic emission in the monitoring of the early-stage cracks. This research is completely new for our group, and therefore, cooperation with Aerospace Non-Destructive Testing Laboratory at TUD was established (Xin *et al.*, 2021).

FEA analysis was performed in advance to identify the critical area for AE monitoring. The conventional AE sensors R15α sensor and Express-8 MISTRAS system were used for data acquisition, see Figure 9. In addition to the traditional AE features, AE feature based on information entropy has been proposed so-called PSE (Tsallis entropy). PSE provides robust and reliable information for AE monitoring, which is independent of user-defined acquisition threshold and noise. The mean value of the PSE is selected as the damage indicator for the damage identification. The early damage and stiffness degradation of the wedge connection can be effectively detected using the appropriate signal processing method see diagram in Figure 9.

4. WIRE ARC ADDITIVE MANUFACTURING (WAAM)

The interest of our group in WAAM technology started when a potential to optimise the production of structural parts was demonstrated using the stainless-steel wires firstly and then the carbon steel wires. The technology has become mature, leading to shorter fabrication times and less expensive total costs due to lower raw material costs.

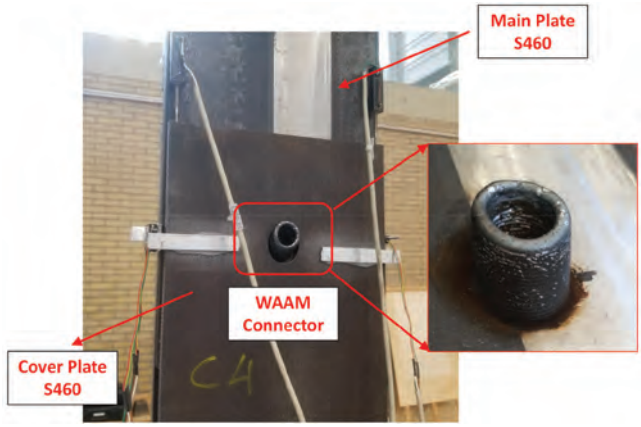


Figure 10. GMAW Pin Connection manufactured by hybrid manufacturing

Therefore, it is timely to conduct the material tests to evaluate the plastic flow and fracture of the Gas Metal Arch Welding (GMAW) steel plate. Such investigation would lead to the material characterisation necessary for a reliable prediction of structural behaviour and/or an efficient design of connections using GMAW fabricated plates. The coupon specimens were cut from the GMAW plates in different directions in relation to the printing orientation to investigate possible material anisotropy. The plates were printed at ArcelorMittal Global R&D Bars & Wires lab by a welding robot based on Wire Arc Additive Manufacturing (WAAM) procedure and by using the welding wire AWS 5.18 ER70S-6 Böhler K56.

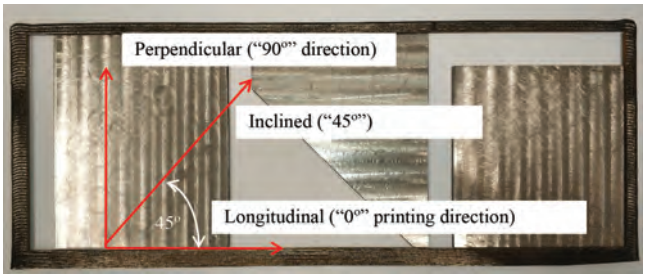


Figure 11. The milled plate used to cut specimens

Results of uniaxial coupon specimens, the stress-strain curve, were analysed in three stages: the elastic stage, the plastic stage and the coupled plastic-damage stage. The FE simulation is performed to calibrate the true stress and strain curves in different stages.

Together with ArcelorMittal Global R&D Bars & Wires, we are developing new knowledge for the optimal design of pin connections by hybrid manufacturing, see Figure 10. The main plate and cover plates were made of S460 steels, and the pin connectors were manufactured by GMAW.

Four plates, with a dimension of 400 mm × 150 mm × 3.72 mm, were fabricated using welding wire and getting approximately 1 mm thick layer. The chemical composition and mechanical properties of the steel wire are given in (Xin *et al.*, 2021). Obtained results clearly indicate that 3D printing technology was appropriate, showing material properties rather close to the rolled mild steel and minor anisotropy, $E = 200$ GPa, fracture strain above 25%, strength/yield stress beyond 1,3, for $f\,y = 420$ MPa. The roughness of the surface was just 0,2 mm for the thickness of 3,7 mm, obtained using the wire diameter of 3 mm. The printed plates were milled to 2.8 mm thickness, see Figure 6, to study the plastic behaviour of GMAW plates neglecting the surface roughness effect.

Material tests along three different orientations have been prepared on smooth dog-bone (SDB) specimens according to specification ASTM E8 (ASTM E., 2001) to save the material costs by using small coupon specimens. These tests were used to characterise the anisotropic in-plane plastic behaviour of WAAM steel plates at room temperature.

The deformation of specimens is recorded by two-dimensional Digital Image Correlation (DIC), while the force is recorded via a universal electronic material testing machine with a capacity of 100.0 kN. The gauge length used to calculate engineering strain is defined as 25.0 mm based on ASTM E8 (ASTM E., 2001). The tensile tests were conducted in a displacement-controlled mode with a speed of 0.5 mm/min.

Summary of material characteristics influenced by the printing direction is given below:

- The average elastic modulus along 45° and 90° orientation is 8% and 4%, respectively, which are less compared to the printing direction "0° direction", respectively.
- The average yield strength along 45° and 90° orientation is 6% and 7.0% larger than the 0° direction, respectively.
- The average ultimate strength along 45° and 90° orientation is 1% and 2% larger than the 0° direction, respectively.
- The average uniform elongation along 45° and 90° orientation is 8.0% and 12% larger than to the 0° direction, respectively.

The isotropic behaviour could be assumed for engineering applications using the parameters shown in Table 1. According to these values, the 3D printed material complies with all requirements imposed on structural steel (Tarus *et al.*, 2021):

Table 1. Material properties for design of 3D printed GMAW carbon steel

Milled:	As printed:
• $E = 198851$ MPa	• $E = 197330$ MPa
• σ_y: 430 MPa	• σ_y: 420 MPa
• σ_u: 570 MPa	• σ_u: 545 MPa
• σ_u/σ_y: $1.33 > 1.10$	• σ_u/σ_y: $1.30 > 1.10$
• $\varepsilon_u/\varepsilon_Y$ :: $38 > 15$	• $\varepsilon_u/\varepsilon_Y$: $27 > 15$
• $\varepsilon_u = 0.16 > 0.15$	• $\varepsilon_u = 0.16 > 0.15$
• $\nu = 0.27$	• $\nu = 0.25$

The failure mode of dog-bone specimens is shown in Figure 12. Obvious necking is observed before the final fracture appears for the specimens in all three different orientations, as it would be in the case of mild steel specimens. The obtained fracture plane, which is inclined (slant) in the thickness direction for all specimens, indicating that the shear failure is dominated in the thickness direction. Different from the flat fracture mode of the mild steel, the slant fracture is not only in the appearance of fracture pattern but also in the micro-void growth and coalescence. The shear localisation mode, namely large micro-void shape but relatively small volume changes, happened in the microscale (Weck and Wilkinson, 2008) along the thickness direction. The plastic flow in the localised shear bands lead to the ultimate fracture of GMAW plates is observed in all loading orientations.

5. PLASTIC FLOW OF GMAW PLATES

The uniaxial stress-strain relationship of GMAW specimens is modelled in three distinct regions: the elastic stage, the plastic stage, and the coupled plastic-damage stage. The coupled plastic-damage stage is further decomposed into the plastic-dominated zone and the damage-dominated zone. The elastic stage is controlled by the elastic strain and Young modulus. The plastic and coupled

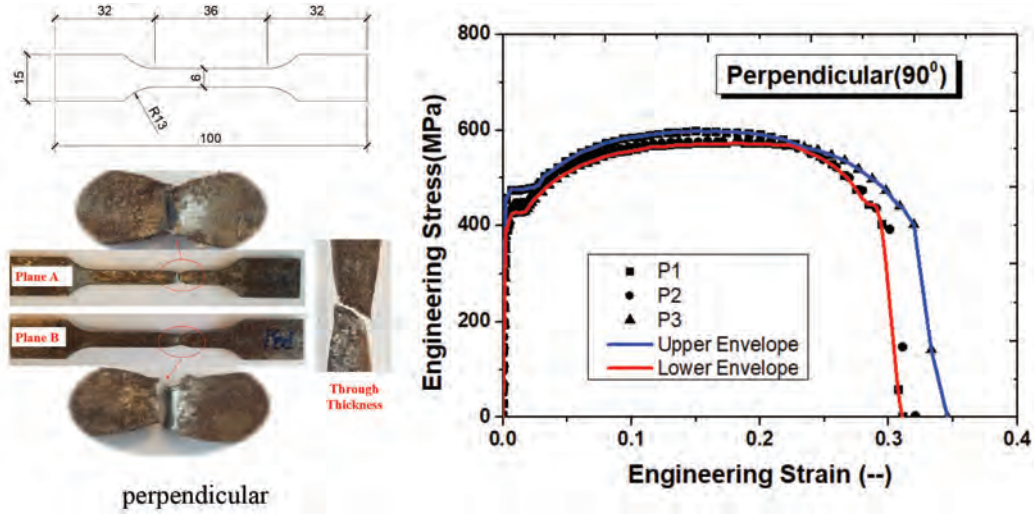


Figure 12. An example of the coupon test results

plastic-damage stages are presented in Figure 13. Necking and damage effects are not considered for true stress.

The relationship of the uniaxial true stress-strain is easily obtained in the plastic stage (Tarus *et al.*, 2021). The main challenge and focus point is in modelling of the relationship of uniaxial true stress-strain in the coupled plastic-damage stage. The calibration process is sequential, meaning that the effects of damage-dominated zone is considered separately from the calibration in plastic-dominated zone. Details are shown in (European Committee for Standardization, 2011).

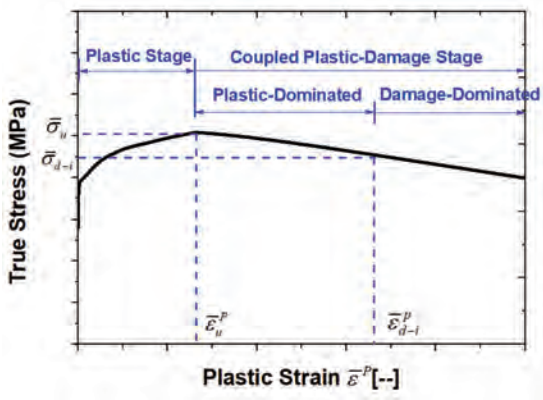


Figure 13. Plastic and coupled plastic-damage stages

The material is exposed to the plastic stage when the equivalent plastic stage is $0 \leq \bar{\varepsilon}^p \leq \bar{\varepsilon}^p_u$, where $\bar{\varepsilon}^p_u$ is the plastic strain when the true stress without considering necking and damage effects reaches the peak, see Figure 13. In the plastic stage, the uniaxial plastic strain and true stress are simply obtained through the engineering strain-engineering stress relationship, assuming conservation of the volume. The engineering strain ε_{en} and engineering stress σ_{en} relationship is simply converted to the true stress σ vs. true strain ε relationship using well known Eq. (1) and (2).

$$\varepsilon = \ln\left(1+\varepsilon_{en}\right) \tag{1}$$

$$\sigma = \sigma_{en}\left(1+\varepsilon_{en}\right) \tag{2}$$

The coupled plastic-damage stage reached when $\bar{\varepsilon}^p_u > \bar{\varepsilon}^p$. The point of maximum true stress is the onset of the necking. When $\bar{\varepsilon}^p_u < \bar{\varepsilon}^p \leq \bar{\varepsilon}^p_{d-i}$, the plasticity is dominated in the coupled plastic-damage stage. The weighted function, according to (Y., 2021) shown in Eq. (3), is used to predict the true stress after necking. The measured engineering stress-engineering strain relationship is considered as a target in calibrating the weight constant W, $0 \leq W \leq 1$, using the finite element simulation.

$$\bar{\sigma}^{\text{neck}} = \bar{\sigma}_u \left[W\left(1+\bar{\varepsilon}^p - \bar{\varepsilon}^p_u\right) + (1-W)\left(\frac{\left(\bar{\varepsilon}^p\right)^{\bar{\varepsilon}^p_u}}{\left(\bar{\varepsilon}^p_u\right)^{\bar{\varepsilon}^p_u}} \right) \right] \tag{3}$$

where W is a weight constant, $0 \leq W \leq 1$.

The damage is dominated in the coupled plastic-damage stage for strains $\bar{\varepsilon}^p > \bar{\varepsilon}^p_{d-i}$, see Figure 13. The true stress in the damage-dominated zone could be obtained by the Eq. (4) because the damage effects after necking are not considered in the definition of true stresses. The damage evolution law, expressed in Eq. (5), is assumed in determining the dependence of the damage scalar d on the equivalent plastic strain $\bar{\varepsilon}^p$. The experimentally established engineering stress-engineering strain relationship is considered as the objective of the calibration. The parameter B is varied in the finite element model until the calculated engineering curves agree well with the experimental results. The stop criterion of the parameter B calibration is when the difference between the numerial predicted and experimental engineering stress-strain relationship is within 5%. Noted that Eq. (4) and (5) could be repeated several times by adjust the $\bar{\varepsilon}^p_{d-i}$ until the predicted engineering stress-strain relationship agreed well

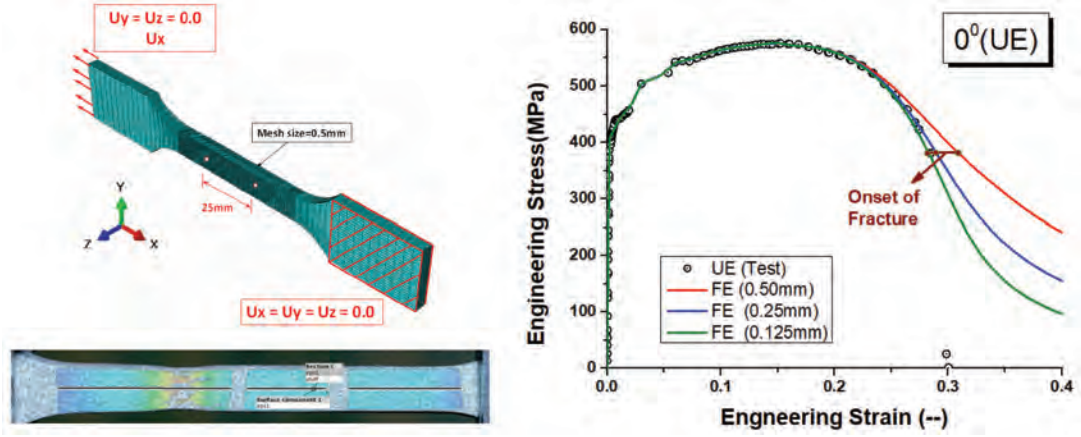


Figure 14. Finite element model and validation of analysis by various element size

with numerical results.

$$\overline{\sigma} = (1-d)\overline{\sigma}^{\text{neck}} \tag{4}$$

$$d = \begin{cases} 0 & \overline{\varepsilon}^{p} < \overline{\varepsilon}^{p}_{d-i} \\ 1-\exp\left[-B(\overline{\varepsilon}^{p} - \overline{\varepsilon}^{p}_{d-i})\right] & \overline{\varepsilon}^{p} \geq \overline{\varepsilon}^{p}_{d-i} \end{cases} \tag{5}$$

The finite element model and boundary condition used for parameters calibration of coupled plastic-damage stage are shown in Figure 14. One end of the coupon specimen model is fixed, and the other end is used to introduce the displacement. The specimen is modelled by solid element C3D8. The quasi-static simulation is implemented using ABAQUS/EXPLICIT solver with a total time 1 second for the mass scaling. Note, no physical meaning of the time step duration exists.

The time increment and the mesh size influence FE results. A larger variation of results is obtained in the plastic stage and coupled plastic-damage stage.

When the time increment is less than 1×10^{-5}, all the models lead to identical engineering stress and strain curves until the force reaches the peak value. The maximum force represents the onset of the necking. However, the predicted engineering stress of the model in the softening stage, the post necking behaviour, using the time increment of 1×10^{-5} second, is larger compared to the time increments of 1×10^{-6} second and 1×10^{-7} second. The differences in engineering stress and strain predictions using the time increment 1×10^{-6} second and 1×10^{-7} second are very small. Hence, the time increment 1×10^{-6} is proposed for the calibration based on 0.5 mm element size in the middle part of the model.

Details of the material model for advanced FEA are shown in Table 2. It is important to underline that these parameters are calibrated to experimental results of coupon specimens and for one triaxiality stress stage, using the certain mesh size, element type and time step. The results for the upper and lower boundary of experimental results in three directions are given.

Table 2. Calibrated parameters for coupled plastic-damage stage

Description		$\overline{\sigma}_u$	$\overline{\varepsilon}^{p}_{u}$	W	$\overline{\varepsilon}^{p}_{d-i}$	B
		(MPa)	(–)	(–)	(–)	(–)
0°	LE	667.27	0.1677	0.00	0.1920	0.30
	UE	670.90	0.1600	0.28	–	–
45°	LE	694.07	0.1981	0.00	0.2428	0.05
	LE	700.20	0.1981	0.35	–	–
90°	LE	695.26	0.19921	0.0	0.3892	0.20
	UE	708.92	0.1713	0.60	–	–

6. CONCLUDING REMARKS

Steel is used in various applications in civil engineering, from supporting structures to convert renewable energy to electricity to buildings and infrastructure, for its strength, stiffness, durability, and versatility. It is well known that it is endlessly recyclable and reusable. But it has challenges when exposed to corrosion, elevated temperatures, and fatigue. Hybrid structures in which a combination of two or more building materials are optimally used offer infinite opportunities for improved sustainability of the built environment. Improvements in computational methods for predicting the behavior of materials and structures together with the use of new methods based on AI, help engineers to create new applications and improve best practices solutions tremendously.

The use of steel per capita is seen as one of the key indicators of the wellbeing of people within a geographical region. However, steel making is linked to the emission of CO_2 equivalent gases, which negatively influence our natural environment. The increase in population is directly linked to an increase in steel production, independently if steel is produced either from raw material or from recycled steel.

The biggest paradigm change of our time, related to steel as a material, is in the production of “fossil-free” steel, which is actively exploring by leading steel-making

companies. That may bring new opportunities for the use of steel in construction sector and in several related industries, and increase needs and market opportunities for highly educated engineers.

REFERENCES

ASTM E. (2001). *Standard test methods for tension testing of metallic materials.* Annu B ASTM Stand ASTM.

C1CONNECTIONS. (2021). *Technology.* https://c1connections.com/technology/. (visited on 17.09.2021)

CEN. (2005). *En 1993-1-8: Eurocode 3: Design of steel structures - part 1-8: Design of joints.* Belgium.

Cheng, L., Xin, H., Groves, R. M. and Veljkovic, M. (2021). Acoustic emission source location using lamb wave propagation simulation and artificial neural network for i-shaped steel girder. *Construction and Building Materials*, *273*, 121706. Retrieved from https://www.sciencedirect.com/science/article/pii/S0950061820337107 doi: https://doi.org/10.1016/j.conbuildmat.2020.121706

Csillag, F. and Pavlović, M. (2021). Push-out behaviour of demountable injected vs. blind-bolted connectors in frp decks. *Composite Structures*, *270*, 114043.

European Committee for Standardization. (2011). *En1090-2: Execution of steel structures and aluminium structures - part 2: Technical requirements for steel structures.*

Gîrbacea, A. (2018). *Assessment of demountable steel-concrete composite flooring systems* (Unpublished master's thesis). Delft University of Technology.

Huisman. (2021). *Huisman scales up 3d printing of crane hooks.* https://www.huismanequipment.com/en/media_centre/press_releases/163-155_Huisman-scales-up-3D-printing-of-crane-hooks. (visited on 17.09.2021)

Kozma, A. (2020). *Demountable composite beams: Design procedure for non-conventional shear connections with multilinear load-slip behaviour* (PhD Thesis). University of Luxembourg.

Kozma, A., Odenbreit, C., Braun, M. V., Veljkovic, M. and Nijgh, M. (2019). Push-out tests on demountable shear connectors of steel-concrete composite structures. In *Structures* (Vol. 21, pp. 45–54).

Kwon, G. U. (2008). *Strengthening existing steel bridge girders by the use of post-installed shear connectors* (PhD Thesis). The University of Texas at Austin.

Lam, D., Dai, X. and Saveri, E. (2013). Behaviour of demountable shear connectors in steel-concrete composite beams. In *Composite construction in steel and concrete vii* (pp. 618–631).

Lee, M. S. and Bradford, M. A. (2013). Sustainable composite beam behaviour with deconstructable bolted shear connectors. In *Composite construction in steel and concrete vii* (pp. 445–455).

MX3D. (2021). https://mx3d.com/. (visited on 17.09.2021)

NEN. (2011). *Nen-en 1994-1-1: Eurocode 4: Design of composite steel and concrete structures - part 1-1: General rules and rules for buildings.*

Nijgh, M. (2021). *A multi-scale approach towards reusable steel-concrete composite floor systems* (PhD Thesis). Delft University of Technology.

Nijgh, M., Girbacea, I. and Veljkovic, M. (2019). Elastic behaviour of a reusable tapered composite beam. *Engineering Structures*, *183*, 366–374.

Olivier, G., Csillag, F., Tromp, E. and Pavlović, M. (2021). Conventional vs. reinforced resin injected connectors' behaviour in static, fatigue and creep experiments on slip-resistant steel-frp joints. *Engineering Structures*, *236*, 112089.

Pavlović, M. (2013). Resistace of bolted shear connectors in prefabricated steel–concrete composite decks.

Tarus, I., Xin, H., Veljkovic, M., Persem, N. and Lorich, L. (2021). Evaluation of material properties of 3d printed carbon steel for material modelling. *ce/papers*, *4*(2-4), 1650-1656. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/cepa.1469 doi: https://doi.org/10.1002/cepa.1469

TREE COMPOSITES. (2021). *Design and manufacturing of innovative composite solutions for offshore challenges.* https://www.treecomposites.com/. (visited on 17.09.2021)

Weck, A. and Wilkinson, D. (2008). Experimental investigation of void coalescence in metallic sheets containing laser drilled holes. *Acta Materialia*, *56*(8), 1774-1784. Retrieved from https://www.sciencedirect.com/science/article/pii/S1359645407008592 doi: https://doi.org/10.1016/j.actamat.2007.12.035

World Green Building Council. (2021). *Global status report 2017.* https://www.worldgbc.org/news-media/global-status-report-2017. (visited on 17.09.2021)

Xin, H., Tarus, I., Cheng, L., Veljkovic, M., Persem, N. and Lorich, L. (2021). Experiments and numerical simulation of wire and arc additive manufactured steel materials. *Structures*, *34*, 1393-1402. Retrieved from https://www.sciencedirect.com/science/article/pii/S2352012421007670 doi: https://doi.org/10.1016/j.istruc.2021.08.055

Y., L. (2021). Uniaxial true stress-strain after necking. *AMP Journal of technology*, *5*(1), 37–48.

Željko Popović

Zeljko Popović is a Senior Vice President in Mubadala Investment Company, Abu Dhabi, UAE, specialized in management of real estate and infrastructure projects, with particular interest in construction contracts, consultancy agreements and claims for extension of time and additional costs. He worked on complex projects worldwide, and he is a member of Charted Institute of Arbitrators (MCIArb), Royal Institution of Chartered Surveyors (MRICS), American Society of Civil Engineers (ASCE) and Project Management Professional (PMP).

Contact information: e-mail: zelkop011@gmail.com

Zeljko Popović is currently involved in management of real estate development portfolio, covering around 10 million square meters of development land consisting of 40 plots and islands, with 6 million square meters of building gross floor area, having total market value in excess of $1 billion.

Dr. Popović graduated from the University of Belgrade, Faculty of Civil Engineering (FCEUB) in 1988 at the Department of Structural Engineering. He received his MSc degree in 1992 from the FCEUB at the Department of Construction Management and Technology for research on elemental cost model for infrastructure projects, as well as PhD degree in 2002 with the thesis related to procurement and tender evaluation using database visualization techniques and fuzzy logic. Later, he proceeded to obtain further education in construction law and arbitration and obtained an LLM/MSc degree in 2015 from Robert Gordon University, School of Law in Aberdeen.

He started his career as an assistant at the FCEUB (1988-1993) and then worked with contracting companies as site engineer, construction manager and project manager on industrial and building projects (1993-2001). He joined the consultancy company Parsons Brinckerhoff and specialized in site supervision, project management and contract administration of high-value master developments under all types of FIDIC contracts (2001-2011), then became a senior manager in a property development and investment company being responsible for design, construction and commercial management of residential and hospitality assets (2011 to date). Projects worth mentioning include luxury hotels in Russia, power plant and 400kV power stations in Abu Dhabi, major infrastructure at Palm Jumeirah in Dubai, land reclamation and waterfront community in Abu Dhabi, luxury high-raise residential and hotel buildings of 40-50 floors at various locations in the UAE.

Dr. Popović authored several books (in Serbian): “Construction Project Management” (four editions as a co-author, with the latest one published in 2021) and “Construction Claims” (2009). His book “Claims in FIDIC Contracts under UAE Law” was published in English (2016). He also authored more than 50 papers in various professional magazines and seminar proceedings and developed SQL/VB software applications for contract administration and project document control.

Dr. Popović volunteers to mentor under and post graduate students of the Higher Colleges of Technology in Abu Dhabi, Department of Civil Engineering and participates as an external examiner and member of the Industry Advisory Committee for civil engineering at the Colleges.

Delay and disruption claims in FIDIC construction contracts – current practice and future trends

Željko Popović

Mubadala Investment Company, UAE

Summary

Construction claims are requests for compensation in additional time and/or money submitted in relation to the project. Contractor claims are particularly frequent in practice, especially those arising from delay and disruption events on a project. In this paper, we analyze root causes of contractor's time-related claims and discuss key trends and potential improvements, including quality of project programmes and reporting, recommended contract procedures based on FIDIC contract forms, optimal choice of techniques for analysis of extension of time and associated additional costs. Further analysis focuses on peculiarities and importance of local laws and regulations, which can critically influence the outcome of claim assessment, both the entitlement and quantum. Some practical advice and recommendations originating from the latest practice manuals and own experience are provided to assist practitioners engaged on construction projects.

1. INTRODUCTION

Construction claim can be defined as a request by either party to the contract for compensation, which can be in the form of the additional payment and/or an extension of time (EOT). In most contracts, cost and time parts of the claim are to be submitted separately.

A claim on a construction project can typically be made:

- in accordance with a particular contract clause giving the entitlement to additional money and/or time (e.g. unforeseen ground conditions, late provision of the site);
- because one party has failed to fulfil an obligation stated in the contract (e.g. delayed payment for works, failure to provide required technical inputs);
- based on applicable legislation, regardless of whether it is mentioned in the contract (e.g. structural liability and other statutory requirements).

Although claims can be submitted by both parties, contractor claims are more common. Particularly complex and notorious are contractor's delay and disruption claims resulting in simultaneous requests for EOT and additional payment due to loss of time and/or productivity. Therefore, it seems logical to discuss them from technical, commercial and legal perspectives.

2. RISKS AND ROOT CAUSES OF DELAYS IN CONSTRUCTION PROJECTS

The extensive CIOB Report (2009)[1] addressed the inefficiency of time management in the 21st Century, based on data collected from around 2,000 projects. CIOB criticized the improper use of network programming, lack of resource/cost allocation in construction programmes and poor contract administration. In 33% of building and 75% of civil engineering projects, contractors were responsible for delays and only 20% of them were familiar with a delay being declared even if the contract required it! Reasons for failing to notify delays included 'possible catch up' (41%), blame of others (10%), good relationships (37%); only 12% of reasons were contractual.

(EC Harris, 2013)[2] reported that top two causes of construction disputes were: improper contract administration and failure to make interim extension of time awards. Dispute resolution times seem to be the longest in the Middle East, 14.6 months in average, as compared to 12.8 months globally. The average time for resolution of disputes has gradually increased from 9.1 months in 2010 and 10.6 months in 2011. The average dispute value is around USD 30 m.

[1] Chartered Institute of Building, Managing the Risk of Delayed Completion in the 21st Century, Report, 2009

[2] EC Harris, Global Construction Disputes: A Longer Resolution, Report, 2013

The UAE Study (2010)[3] investigated 42 potential delay factors with 50 local companies. Client-related factors (change orders, slow decision making) and poor contract practice (planning, estimating and contract administration) contributed to delays. Another **UAE Survey (2012)**[4] found that 90% of 63 surveyed projects were delayed in average by 8.3% of the planned duration due to poor scope management and lack of proper planning/control.

HKA Crux Reports (2019–2021)[5] are particularly relevant as they come from the company that specializes in claim consultancy. HKA findings from 1,185 projects correlate with CIOB's and highlight that claims mostly arise from poor scope management, design delays and contractor's improper project management practices. Staff training and claim techniques are also mentioned.

CIOB recommended further standardization, training and accreditation measures, which are yet to be finalized. Other professional organizations, such as SCL[6], RIBA[7] and RICS,[8] have taken similar initiatives and published their own professional guides.

Author's own study (2016)[9] covered several building projects in the UAE and findings are summarised on Tab. 1. It was found that programmes do not capture all contract requirements and methods of work in spite of their complexity (10,000+ activities). Long ($> 28d$) and short ($< 7d$) activities are not optimised, while excessive total floats ($> (28d)$ invite for disputes about float ownership (no strategy for time contingencies). Project reporting lacks visualisation, while delay analysis methods mostly rely on the original (as-planned) programme, even if it is flawed. "Parallel' delays" of the employer and contractor are not investigated and accounted for.

Above studies reveal several **root causes of time-related problems in construction,** with the following three topping the list:

1. Poor quality of contractor's programme including maintenance of site records and progress reporting;
2. Inappropriate contract frameworks and procedures for time management and claims;
3. Sub-standard EOT claim analysis and presentation techniques.

Each of the above is analyzed highlighting practical issues and emerging trends.

3. QUALITY OF THE PROGRAMME AND PROGRESS REPORTING

3.1. Project programme

Programme approach should comply with "good practice" guides (SCL,[10] PMI[11]) and experience:

- Programme should be based on network scheduling methodology, the *Precedence* method (using activity nodes and logical relationships) and scheduling software (e.g. *Oracle Primavera)*.
- Programme must contain enough detail to communicate the intended methodology for *each trade* in *each area of the work*, supported with:
  - Narrative and sketches explaining the method of execution: logistics plans, method statements, cycle analysis, workflows, etc;
  - Captured contractual requirements: Provisional Sums, milestones, review/approval periods, etc;
  - Activity ID structure: (ProjectCode)-(Phase)-(Discipline)-(Floor/Location)-(ActivityNo);
  - Calendars with holidays and reduced working hours (summer, Ramadan);
  - Productivity rates used to calculate durations;
  - Manpower histograms;
  - S-curves (effort, costs, Earned Value);
  - Work Breakdown Structure, up to 6 levels,[12] such as: (1) Phase, (2) Building/Location, (3) Discipline/Trade, (4) Activity Groups, (5) Detailed Activities, (6) Micro Activities; level 5 is appropriate for most contracts;
  - Derived reports, e.g. Line of Balance, to explain movement of crews;
- Baseline Gantt Chart, showing: Activity ID/Description, Original Duration, Early/Late Start/Finish, Total Float, plus updates: Current Duration, Actual Start/Finish, %Complete. The longest (critical) path to be highlighted. Programme presentations may further benefit from 3D BIM (Building Information Modelling) methods/software.
- Quality checks: activities (no missing tasks), logic (meaningful), calendars (resources, activities), leads / lags (no unreasonable waiting times), constraints

[3] Omayma Motaleb and Mohammed Kishk, An investigation into causes and effects of construction delays in UAE, Egbu, C. (Ed) Proceedings 26th Annual ARCOM Conference, 6-8 September 2010, UK, pp. 1149-1157

[4] Arun Bajracharya and Mohammad Halloum, Cost and Time Overrun Revisited: A Study on the Infrastructure Construction Projects in Abu Dhabi – UAE, ICCIDC–III Conference, July 4-6, 2012, Bangkok, Thailand

[5] HKA, Claims and Dispute Causation – A Global Market Sector Analysis (Crux Report s, 2019–2021)

[6] Society of Construction Law (SCL), Delay and Disruption Protocol, Feb 2017, Second Edition

[7] Royal Institution of British Architects (RIBA), Good Practice Guide: Extensions of Time, RIBA Publishing, 2008

[8] Royal Institution of Chartered Surveyors (RICS), Extensions of Time, 1st Ed., RICS Professional Guidance, 2014

[9] Z Popovic, Claims for Extension of Time in FIDIC Construction Contracts, A Practical Approach under UAE Law, Higher Colleges of Technology, Abu Dhabi, UAE, 2016

[10] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Ed., Feb 2017

[11] Project Management Institute, Practice Standard for Scheduling, 2nd Edition, PMI, USA, 2011

[12] CIOB and Hill International, Masterclass – Manage, Analyse and Avoid Construction Delay & Disruption in a Changing Economic Climate, Masterclass led by Keith Pickavance, Abu Dhabi, UAE, 2-3 June 2009

Table 1. Programme quality study (UAE projects)

Item	Projects									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Project Details										
Project type	Residential	Offices	Residential	Residential	Residential	Residential	Residential	Residential	Hotel	Residential
Configuration and built-up area (m^2)	3B + G + 3P + HC + MEP + 38F (77,939 m^2)	6B + G + 3P + MEP + 34F (90,713.3 m^2)	3B + G + 25F (65,257 m^2)	1B + G + 1P + 10F(115,099m2)	3B + G + 11F (16,754 m^2)	3B + G + 9F, 3B + G + 7 (38,889 m^2)	3B + G + 39F (45,619 m^2)	2B + GF + 5P + 34F + R (37,433 m^2)	5B + 3B + 35F (70,200 m^2)	1B + GF + 28F + R (36,865m^2)
Contract duration	24 m	22 m	22 m	24 m	18 m	22 m	24 m	24 m	36 m	22 m
Conditions of contract	FIDIC 1987	FIDIC 1987	FIDIC 1999	FIDIC 1999	FIDIC 1999	FIDIC 1987	FIDIC 1987	FIDIC 1987	FIDIC 1999	FIDIC 1999
Baseline Programme										
Method statement	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WBS level (1-6)	L5	L5	L5	L5	L5	L4	L5	L5	L5	L5
No of activities	4,037	3,678	8,190	11,767	2,709	2,536	6,741	6,218	6,895	3,345
No of open ends	2	2	2	2	2	2	5	5	2	2
No/% of long activities ($> 28d$)	156 (3.8%)	133 (3.6%)	215 (2.6%)	2,037 (17.3%)	59 (2.1%)	137 (5.4%)	1,084 (16.1%)	1,036 (16.6%)	384 (5.5%)	261 (7.8%)
No/% of short activities ($< 7d$)	842 (20.8%)	724 (19.7%)	4,213 (51.4%)	2,998 (25.5%)	2,068 (76.3%)	564 (22.2%)	1,813 (26.9%)	1,493 (24%)	1,885 (27.3%)	804 (24%)
No/% of excessive total floats ($> 28d$)	1,533 (37.9%)	1,324 (36.0%)	2,836 (34.6%)	6,774 (57.5%)	1,012 (37.3%)	1,103 (43.5%)	4,726 (70%)	5,237 (84%)	4,263 (61.8%)	1,267 (37.9%)
No/ % of critical activities	120 (2.9%)	121 (3.3%)	416 (5.0%)	741 (6.3%)	334 (12.3%)	86 (3.4%)	185 (2.7%)	135 (2.2%)	475 (6.9%)	1,087 (32.5%)
Milestones provided	19	21	72	39	11	11	65	65	48	49
Resources, costs	Loaded	Loaded	Loaded	Loaded	Loaded	High level	Loaded	Loaded	Loaded	Loaded
Different calendars	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d	Yes, 5d-6d
Provisional Sum and time contingency	PS–Nil. Cont.=float.	PS–Nil. Cont.=float.	PS–Yes. Cont.=float.	PS–Yes. Cont.=float.	PS–Yes. Cont.=float.	PS–Yes. Cont.=float.	PS–Yes. Cont.=float.	PS–Yes. Cont.=float.	PS–Yes. Cont.=float.	PS–Yes. Cont.=float.
Number of baseline iterations; comments	3; Authority approvals, missing scope, finishes	4; Authority approvals, missing scope, finishes	6; Authority approvals, missing scope, finishes	4; Authority approvals, missing scope, finishes	5; Incorrect appr. sequence; key trades not properly linked	3; Finishes, MEP not as per industry standards	5; Trades not linked properly, conflicts	5; Trades not linked properly, conflicts	6; Approval sequencing, missing scope, finishes	5; Approval sequencing, missing scope, finishes

Table 2. Progress reporting and EOT claims study (UAE projects)

Item	Projects									
	1	2	3	4	5	6	7	8	9	10
Progress Reporting										
Actual dates presented,% complete shown	Yes; Physical %	Yes; Physical %	Yes; Physical %	Yes; Physical %	Yes; Physical %	Yes; Physical %	Yes; Physical %	Yes; Physical %	Yes; Physical %	Yes; Physical %
Logic adjustments	Yes	Yes	Yes	Yes	Yes	Nil	Yes	Yes	Yes	Yes
Float monitored	Weekly	Weekly	Weekly	Weekly	Weekly	Monthly	Weekly	Weekly	Weekly track	Weekly track
S-curve (cost, effort)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cost report, EV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Summary dashboard	By Engineer	By Engineer	By Engineer	By Engineer	By Engineer	By Engineer	By Engineer	Yes	By Engineer	By Engineer
Risk workshops (time)	Occasional	Occasional	Occasional	Occasional	Occasional	Weekly	Occasional	Occasional	Occasional	Occasional
Submissions tracked	Yes	Yes	Yes	Yes	Yes	Nil	Yes	Yes	Yes	Yes
Weekly reports	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
EOT Claims (Time)										
EOT claims submitted	7	7	3	1	12	3	9	9	2	1
EOT claims notified	31	31	3	3	53	3	89	81	2	1
Delay analysis method used to calculate EOT	Windows analysis	Windows analysis	Windows analysis	Simple time impact anal.	Impacted as planned	Impacted as planned	Impacted as planned	Impacted as planned	Simple time impact anal.	Simple time impact anal.
EOT claimed (days)	151d	212d	299d	85d	339d	264d	403d	399d	104d	118d
EOT approved (days)	151d	212d	215d	Reviewed	339d	264d	222d	217d	82d	91d
EOT entitlement – number of accepted items	3	3	1	1	7	Variations; authorities	3	3	1	Delay in site possession
EOT entitlement – number of rejected items	28; contr.des., delays not critical,records	28; contr.des., delays not critical,records	2; delay not critical, lack of substantiation	–	21; contr.des., delays not critical,records	2; delay events not critical, subst.	12; contr.des., delays not critical,records	12; contr.des., delays not critical,records	1; works in Contractor's original scope	Nil
Quality of submission (basis, cause, effect)	Good, as per ind. standard	Good, as per ind. standard	Reasonable	Reasonable	Up to ind. standards	Poor	Reasonable; non-stand.	Reasonable; non-stand.	Reasonable, standard	Reasonable, standard
EOT Claims (Cost)										
Costs claimed (AED)	–	–	17 m	Not yet	5.4 m	5 m	14 m	12 m	Not yet	Not yet
Costs approved (AED)	–	–	0	–	2.7 m	0	0; Concurrent	0; Concurrent	-	-
Entitlement (accepted)	-	–	Nil	–	Var., author.	Nil	Nil	Nil	–	–
Details, actual costs	–	–	Poor,no rec.	–	Eng. determ.	Poor, subst.	Poor	Poor	–	–

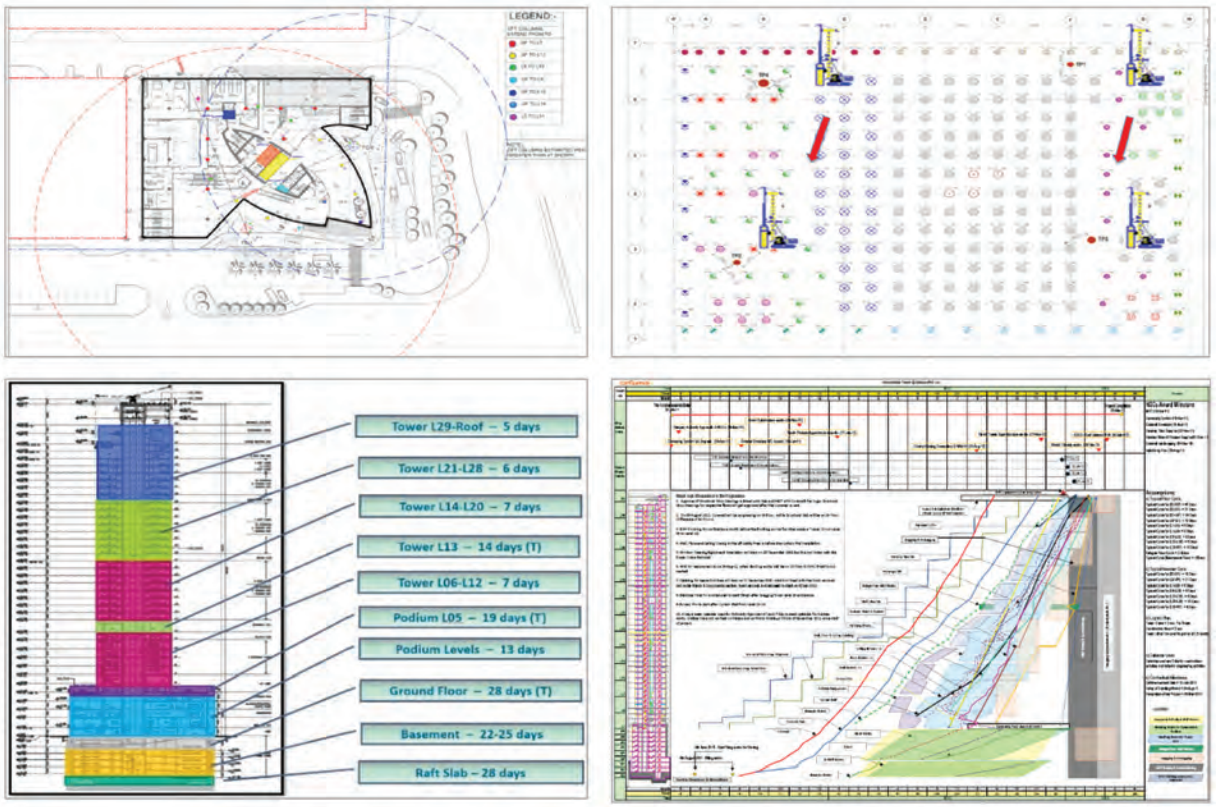


Figure 1. Standard scope and programme analysis techniques for a high rise project (44 floors, Abu Dhabi, 2017) – (1) Site layout and logistics (up left), (2) Optimization of crew movement for piling works (up right), (3) Study of cycles for concrete works (bottom left), (4) Line of Balance with crew movements through floors (bottom right)

(based on contract requirements), floats (no excessive floats, no negative floats), long durations ($> 28d$ to be justified), resource loading (resources, costs), near critical path (float ≈ 0).

- Contractor's time contingencies should be included as separate activities, as opposed to relying on excessive activity floats.

Examples of the above approach are shown on Fig. 1, for illustration.

Further visualisation may include linking of activities in the 2D programme with 3D model components by importing IFC files as illustrated on Fig. 2 (example of *Powerproject BIM* software). There are also integrated packages that allow for direct creation of three-dimensional programmes. Such visual enhancements are useful for programme presentations (e.g. creating fly past videos) and quality checks (e.g. simulation of parallel works during execution), but they do not negate the need for "good practice" illustrated on Fig. 1.

Current FIDIC Conditions of Contract (2017-2019) still do not specify guidelines for BIM-assisted project management, but offer *'Advisory Notes to Users of FIDIC Contracts Where the Project is to Include Building Information Modelling Systems'*, announcing more detailed guidelines in the future. The New Engineering Contract (NEC4 2017) forms, however, provide a clearer path for inclusion of BIM through *Secondary Option Clauses*, *Option X10: Information Modelling*, with guidance for *BIM Protocol* and *Employer's Information Requirements*.

Unfortunately, linking of the programme (contractor's responsibility) with the BIM model (designer's responsibility) on a project (paid by employer) may complicate things. There is a first-in-history case **Trant v Mott MacDonald (2017)**[13] where a designer (Mott) refused to give BIM model access to the employer (Trant) due to financial disagreements, which brought a GBP55m power-generation project in the Falkland Islands to a complete stop, and court resolved by granting access against some deposit payments in order to rescue the project.

Insurance facilities for BIM are also in early stages. Today's BIM practice relies mostly on BIM Level 2 model, which still maintains lines of responsibility and complies with standard Professional Indemnity (PI) insurance polices. (Note: BIM *Level 0* comprises shared 2D drawings, *Level 1* includes isolated 3D deliverables, *Level 2* involves 3D model collaboration using centralized data with potential cost/programming components, and *Level 3* involves holding all data with FM information on a fully integrated web-based system). *Digital Built Britain* (formerly *BIM Task Group*) issued *Publicly Available Specifi-*

[13] Trant v Mott MacDonald [2017] EWHC 2061 (TCC)

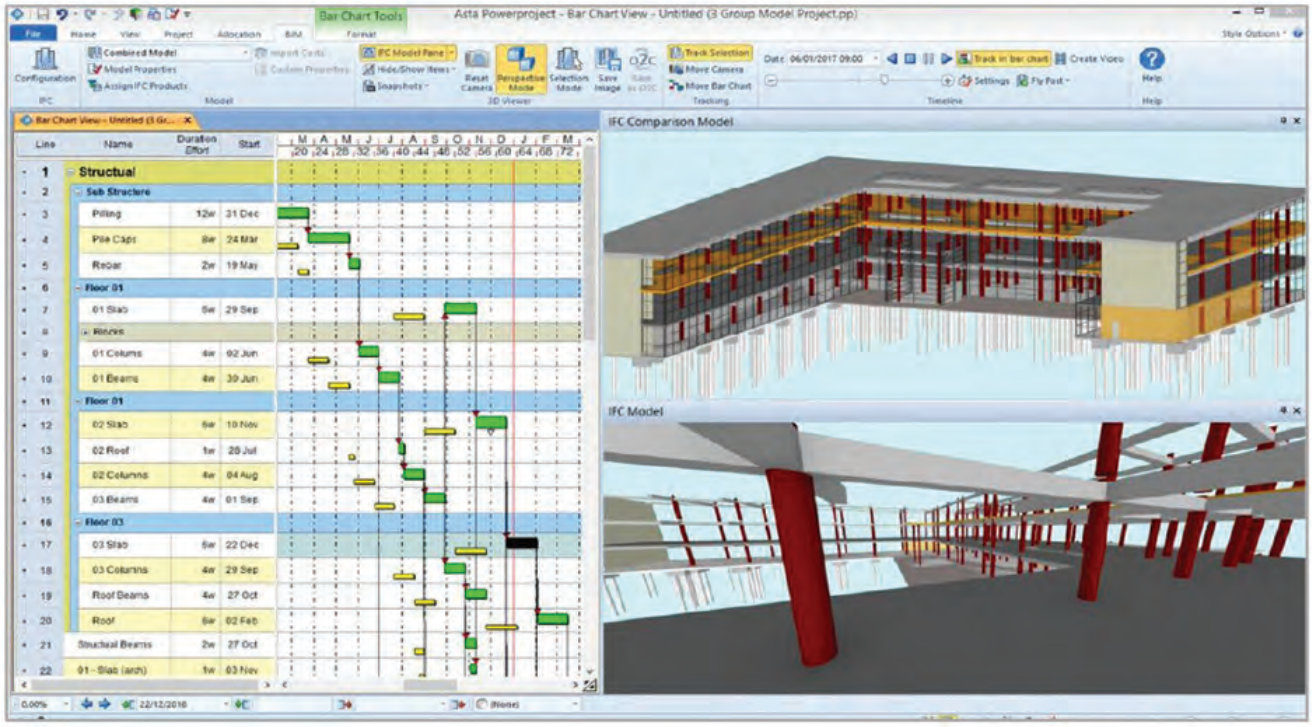


Figure 2. Programme integration with 3D BIM model (example - Powerproject BIM software)

cation (2013)[14] recommending that project designer remains responsible for BIM Level 2 model, however for Level 3 in the future, new PI Insurance strategies will be needed (e.g. single project insurance policy instead of traditional individual annual policies).

3.2. Progress reporting and contemporary records

The CIOB Report (2009)[15] revealed that 49% of contractors attempt to perform delay analysis on a Programme that is not updated! This is against the practice guides; according to the SCL Protocol (2017),[16] *'the programme should be updated to record actual progress and any EOT granted...then...used as a tool for determining EOTs'*. Practice guides offer comprehensive advice on how to maintain the programme so that it remains suitable for EOT calculations.

The critical path (the longest path that drives project completion) may change in the updated programme, and English courts have acknowledged the dynamic nature of the programme. In **Henry Boot Construction v Malmaison Hotel (1999)**,[17] assessment of delay was based on a revised programme where the work affected by the developer was not on the critical path.

In line with good practice, most conditions of contract, including FIDIC 2017, specify minimal industry standard for monthly report submissions. These normally cover progress charts/descriptions for each stage (design, construction, procurement, manufacture, delivery, erection, testing); photographs and other visuals; quality assurance and safety records; comparisons of actual and planned progress with corrective measures.

Standard conditions of contract usually do not prescribe reporting variables and formats as illustrated on Fig. 3 (from the same project planned on Fig. 1). Advanced reporting techniques (Earned Value, online dashboards) and visuals by site Web cameras and drone inspections (time-stamped photos/videos) are common practice. In average, a 30-minute drone flight could collect the same contemporary data that it may take a surveyor one full day,[18] which is 20 times more efficient strategy for gathering of contemporary records for future EOT claim analysis.

Web-based document management systems further facilitate claim research and increase availability of progress reports and other contemporary records, such as: correspondence, meeting minutes, inspection records, requests for information, instructions, drawings, material submissions, delivery documents, site safety records, and similar.

BIM-assisted progress reports can also be used to enhance clarity of the status of works and may help to iden-

[14] UK BIM Task Group, Publicly Available Specification–PAS, PAS 1192-2:2013, UK, 2013

[15] CIOB, Managing the Risk of Delayed Completion in the 21st Century, Research Report, 2009

[16] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Ed., 2017

[17] Henry Boot Construction (UK) Ltd v Malmaison Hotel (Manchester) Ltd [1999] 70 Con LR 32

[18] Driver Trett, Middle East & Africa Knowledge Report, 10 May 2021

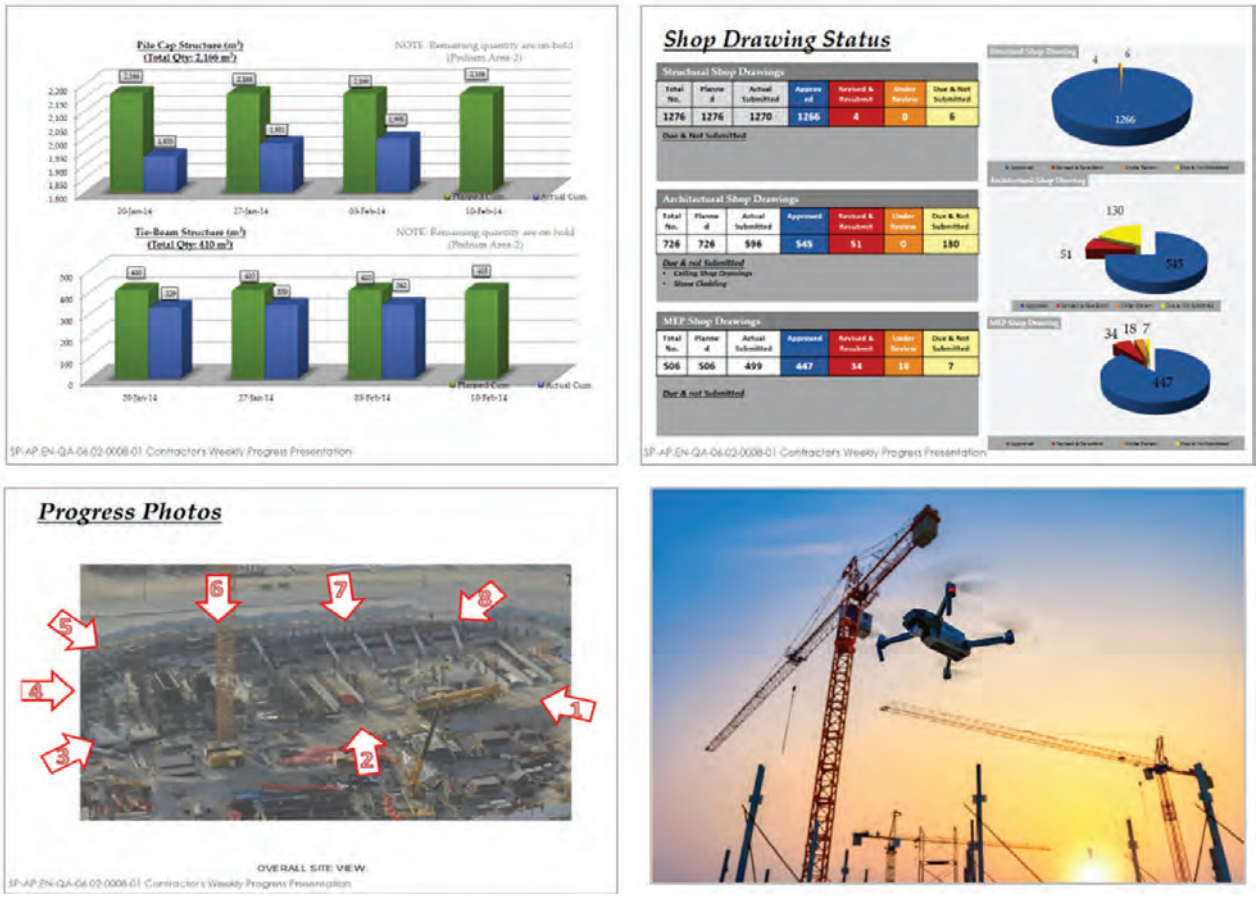


Figure 3. Sample progress reports for a high rise project with 44 floors in Abu Dhabi (2017) – (1) Planned v actual progress of shop-drawings and site works for each trade (upper row), (2) Visuals from Web cameras and drone inspections (bottom left); drone action (bottom right)

tify and visualize critical activities that need attention (examples on Fig. 4a-4b).

4. CONTRACT PROCEDURES FOR TIME MANAGEMENT AND CLAIMS IN FIDIC CONTRACTS

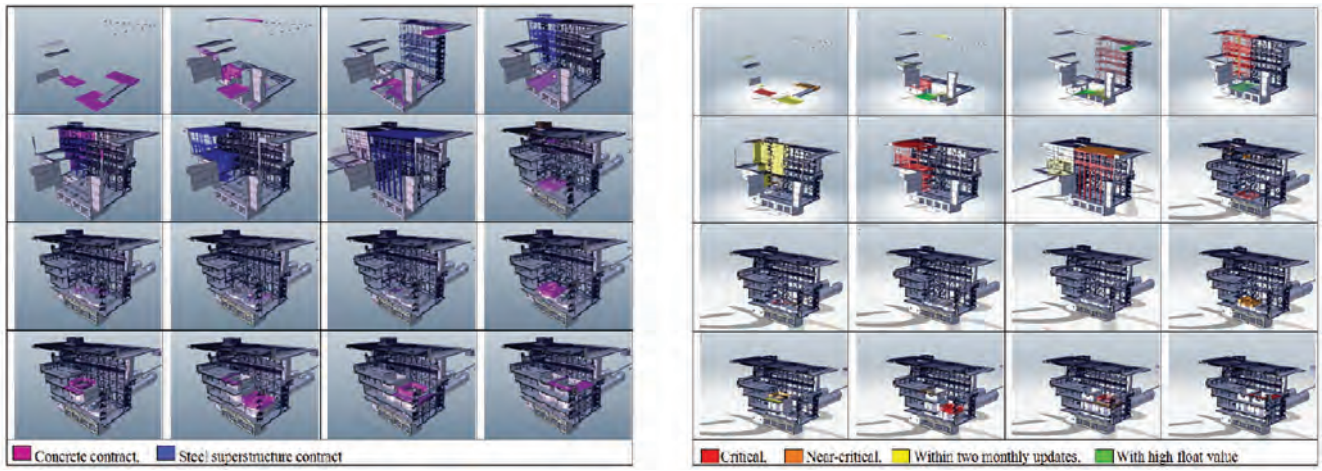
General Conditions of Contract published by FIDIC *(Fédération Internationale des Ingénieurs-Conseils)* are popular for international construction contracts and accepted by financial institutions worldwide. They are intended to be used in any jurisdiction and should contribute to claim prevention. The latest editions were published in Dec 2017 (Red, Yellow and Silver Books, classified by responsibility for design) and Feb 2019 (Emerald Book, underground works), the most popular being the Red Book for works designed by the Employer (referred to hereafter). All books are similar in regard to time and claim management provisions.

4.1. FIDIC golden principles

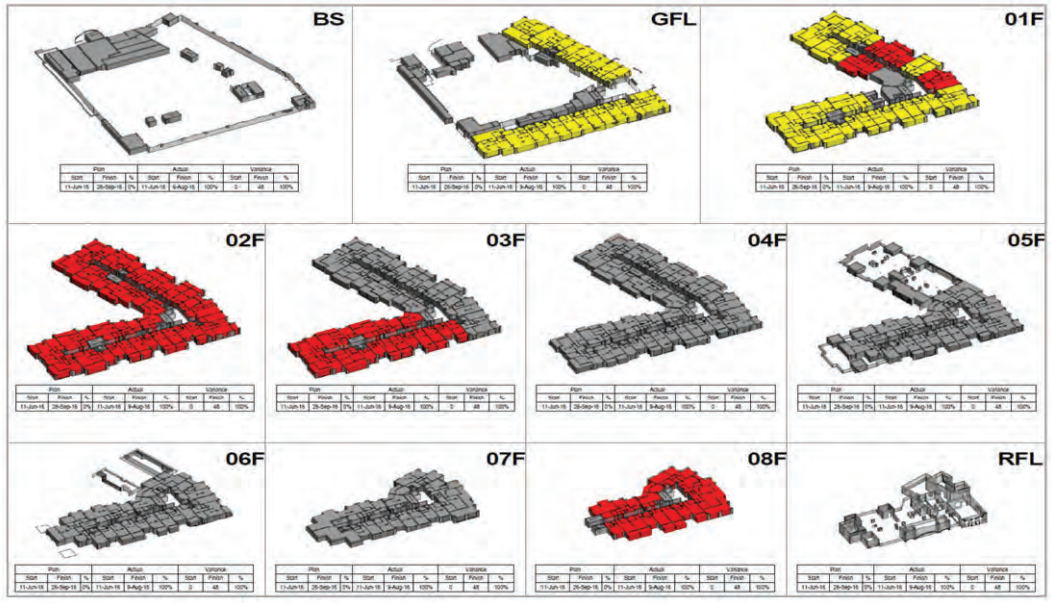
FIDIC contracts are based on fair and balanced risk/reward allocation between the Employer and the Contractor. An impartial Consultant (the Engineer) is managing the project and making decisions, including claim determinations. To emphasise this approach, FIDIC recently published five **FIDIC Golden Principles (2019)**[21]:

1. **The duties, rights, obligations, roles and responsibilities must be in line with the General Conditions, and appropriate to the requirements of the project**
   In regard to claims, the Engineer's third-party role should not be degraded by Particular Conditions. He should not be required to obtain Employer's approval before making fair determination of a Contractor's claim/EOT.
2. **The Particular Conditions must be drafted clearly and unambiguously**
   Tender negotiations should summarized in Particular Conditions, as opposed to combining various emails/meetings and attaching them separately, creating potential ambiguities in priority of documents.
3. **The Particular Conditions must not change the balance of risk/reward allocation from the General Conditions**

[20] Michel Guevremont, "4D Simulation for delay claims avoidance", Concordia University, Canada, March 2021.
[21] The FIDIC Golden Principles, First Edition 2019



(a) Example of 3D-visualization of critical activities: (1) works/trade criticality (zero total float) for concrete and steel works (left); (2) location criticality based on total float values (right)[20]



(b) Example of 3D-visualization of bathroom tiling (G+8 building in Abu Dhabi, 2017): areas not ready (grey colour), in progress (yellow), open for work but not started (red)

Figure 4

It is not acceptable to replace/modify General Conditions with intention to negate Contractor's right to claim, or impose high standards of proof (e.g. Contractor assuming all the risk for Unforeseeable physical conditions).

4. **All time periods specified for Contract Participants to perform their obligations must be of reasonable duration**
   Unreasonably short claim notices (say, 5 days instead of 28 days), or limiting the Contractor's right to suspend work (say, 3 months notice in lieu of standard 21 days).
5. **Unless there is a conflict with the governing law of the Contract, all formal disputes must be referred to a Dispute Avoidance/Adjudication Board (DAAB)**
   Unfortunately, it is common practice to delete all DAAB clauses, or restrict DAAB disputes.

FIDIC recommends avoiding substantial modifications of General Conditions, which would eventually no longer represent FIDIC balanced approach and may mislead the public.

4.2. FIDIC procedures for time management

FIDIC time management procedures are outlined in Clause 8 *(Commencement, Delays and Suspension)*.

Sub-Clauses 8.1 *(Commencement of Works)* and 8.2 *(Time for Completion)* define start and duration of works. The Commencement Date must be within 42d after the Contractor receives the Letter of Acceptance and the Contractor must commence *'with due expedition and without delay'*. Contractor who does not proceed diligently can face termination under Clause 15.2 *(Termination for Contractor's Default)*. The Employer is obliged to give right of

access to the Site within the specified time (2.1 *Right of Access to the Site)*.

It is important that Contractor must *'complete the whole of the Works, and each Section (if any), withing the Time for Completion for the purposes of taking over under Sub-Clause 10.1 (Taking Over the Works and Sections)'*, which includes all Tests on Completion (Clause 9). Inexperienced Contractors sometimes underestimate testing and commissioning of the Works requiring participation of local authorities and service providers. Time for completion of Sections of the Works, if any, must be specified in the Contract Data.

Sub-Clause 8.3 *(Programme)* requires the contractor to submit an initial programme within 28d after commencement. New to FIDIC, the Contractor must use the programming software specified by the Employer or, if not specified, software acceptable to the Engineer. This implies certain resources to be priced in the Tender. The Contractor must submit a revised programme, *'which accurately reflects the actual progress of the Works, whenever any programme ceases to reflect the actual progress'*. The Engineer reviews the initial and revised programmes and may issue a Notice to the Contractor within 21d (initial programme) and 14d (revised programme) if there are discrepancies with Contract Documents or actual progress, or if the programme is *'otherwise inconsistent with the Contractor's obligations'* (improper sequencing, scope omissions, mistaken durations, etc). In the absence of the Notice it is deemed that there is no objection by the Engineer and submitted programme becomes 'the Programme'.

Content of the Programme is specified in Sub-Clause 8.3 in more detail than in previous Red Book (11 items compared to 5), and includes:

- Key milestones and durations (Commencement Date, Time for Completion, Site possession, periods for submissions, sequence/timing of inspections, etc); FIDIC "day" is equal to "calendar day".
- Order in which the Contractor intends to carry out the Works, including Contractor's design activities, procurement, manufacture, delivery, construction, installation, and for any revised programme, the sequence and timing of required remedial work;
- All activities (to the specified level of detail), logically linked and showing the early/late start and finish dates and floats for each activity, and the critical path(s);
- For each activity, actual progress to date, any delay and its effects on other activities (but the Contractor cannot rely upon anything contained in the Programme or supporting report as a Notice given under the Contract);
- Supporting report with description of the stages and methods adopted for execution of the Works, estimated resources (staff, equipment), identification of change(s) to the previous programme and proposals to overcome the effects of any delays.

For the first time, FIDIC specifies that Contractor's programme *must be supported by a description of the methods and sequencing of the Works in line with intended execution strategy and construction techniques*. The Programme is clearly "owned" by the Contractor who is responsible for its management throughout the project. The Programme is a '"live" document adjusted from time to time and not part of the signed Contract (not a "Contract Document").

Sub-Clause 8.4 *(Advance Warning)* requires Parties to *'advise the other and the Engineer, and the Engineer shall advise the Parties, in advance of any known or probable future events'* which may adversely affect the work, productivity, Contract Price, or Time for Completion. Such claim prevention measures are obligatory, although without any time bar. Early warnings resemble other contract forms (NEC4). The Engineer can request a proposal to minimise adverse effects, but it may trigger a Variation (13.3.2 *Variation by Request for Proposal)*.

Sub-Clause 8.5 *(Extension of Time for Completion)* deals with the circumstances in which the Contractor will be entitled to an EOT and practical steps that the Engineer will be expected to take in determining the EOT. Those circumstances include:

a) Variations to the Works, subject to compliance with the procedures from Sub-Clause 13 *(Variations and Adjustments)*;

b) Causes of delay that are specifically mentioned in the Conditions of Contract, which from the Contractor's point of view and under the Red Book include those listed in Tab. 3. However, recovery of Contractor's additional costs and loss of profit is not guaranteed even if the related EOT is justified and approved.

c) Exceptionally adverse and unforeseeable climatic conditions, which is a new potential cause of Contractor's EOT claims as compared to previous editions, however such "exceptional" conditions must be proven to be *'not reasonably foreseeable by an experienced contractor by the date for submission of the tender'* as per the definition of "Unforeseeable" in Sub-Clause 1.1.87;

d) Unforeseeable shortages in the availability of personnel or goods (or materials supplied by the Employer, if any) caused by epidemic or governmental actions, which should pass the test of "unforeseeability" similar to the above climatic conditions;

e) Delays, impediments or preventions caused by the Employer (or his personnel and contractors), with a specific note on parallel or "concurrent" delays, where effects of delays caused by the Parties are overlapped. In such cases, *'the Contractor's entitlement to EOT shall be assessed in accordance with the rules and procedures stated in the Special Provisions (if not stated, as appropriate taking due regard of all relevant circumstances)'*. Experience suggests that there are many approaches

and options for resolving concurrent delays and these are discussed later.

Sub-Clause 8.6 *(Delays Caused by Authorities)* is, for some reason, separated and Contractor must prove that he complied with all local procedures and that such delay was "unforeseeable" if he wants to request an EOT due to local Authorities.

Sub-Clause 8.7 *(Rate of Progress*) allows the Engineer to notify if delays are not excusable (not listed in 8.5 and 8.6) and he may request for corrective measures by the Contractor.

Summary of standard FIDIC 2017 procedures for time management is shown on Tab. 4 below.

4.3. FIDIC procedures for claims management

Clause 20 *(Employer's and Contractor's Claims)* defines the concept of Claims and divides them into three categories: (1) Employer's Claims for additional payments from the Contractor (or reduction in the Contract Price), (2) Contractor's Claims for additional payment from the Employer and/or EOT, or (3) any other entitlement or relief (including in connection with any certificate, determination, instruction, Notice, opinion or valuation of the Engineer). For the first time, Employer's and Contractor's claims are combined within the same clause.

Sub-Clause 20.2.1 *(Notice of Claim)* emphasizes that failure to give Notice within 28d after the claiming Party became aware, or should have become aware, of the event would result in loss of entitlement. Un-notified claim faces Engineer's rejection, which is a strict provision. Its effect in practice remains to be seen, particularly in civil law systems where *"unfair"* business practices are not allowed.

Sub-Clause 20.2.2 *(Engineer's initial response)* allows the Engineer to notify within 14d about Claim rejection (if Notice is not submitted on time) but still retain some flexibility and later review the fully detailed Claim (if the claiming Party disagrees with rejection and provides documents that justify late submission of the Notice of Claim).

Sub-Clause 20.2.3 *(Contemporary records)* deals with the definition of records required to support the Claim *('generated at the same time, or immediately after, the event')* and Engineer's right to inspect and monitor those records, which does not imply his acceptance. The existence and quality of contemporary records is another big point of discussion in practice.

Sub-Clause 20.2.4 *(Fully detailed Claim)* specifies the contents of a detailed Claim submission: (a) description of the event, (b) statement of the contractual/legal basis of the Claim, (c) contemporary records, and (d) detailed supporting particulars and calculations of EOT/money. Time for submission of fully detailed Claim is up to 84d from the date of becoming aware of the event (may be extended by the Engineer). This is a strict time bar, similar to the Notice of Claim, which entitles the Engineer to reject the Claim within 14d, however he can still proceed and review the fully detailed Claim if the claiming Party justifies the late submission.

FIDIC introduces *"deemed"* reaction of the Engineer in cases where Notice of Claim or fully detailed Claim were submitted late but not rejected by the Engineer (it is deemed that submission was accepted). Such deemed reactions are a novelty in FIDIC forms in an attempt to ease administrative omissions.

Sub-Clause 20.2.5 *(Agreement or determination of the Claim)* deals with Engineer's duties to review submissions and seek compromise with the Parties, if possible, or determine in accordance with his powers (3.7 *Agreement or Determination)*.

Sub-Clause 20.2.6 *(Claims of continuing effect)* provides a relief for long claim situations, allowing the initial Claim submission to be "interim", with regular monthly updates and final Claim submitted within 28d after the end of long-term effects.

Above review demonstrates many administrative requirements and obstacles that Contractor must comply with in order to get his rights, which implies additional site resources. The Engineer's obligations are also more sophisticated and require broad technical, commercial and legal skills to cope with project issues. Overall, there is an impression of further "legalisation" of construction business, where administrative/procedural duties may critically influence the outcome of work. Contractors typically respond by increasing the volume of submissions, which may further overload the Engineer. Final impact of these provisions remains to be seen.

Detailed contract procedures for submission, review and resolution of EOT and financial Claims in FIDIC 2017 contracts is shown on Fig. 5.

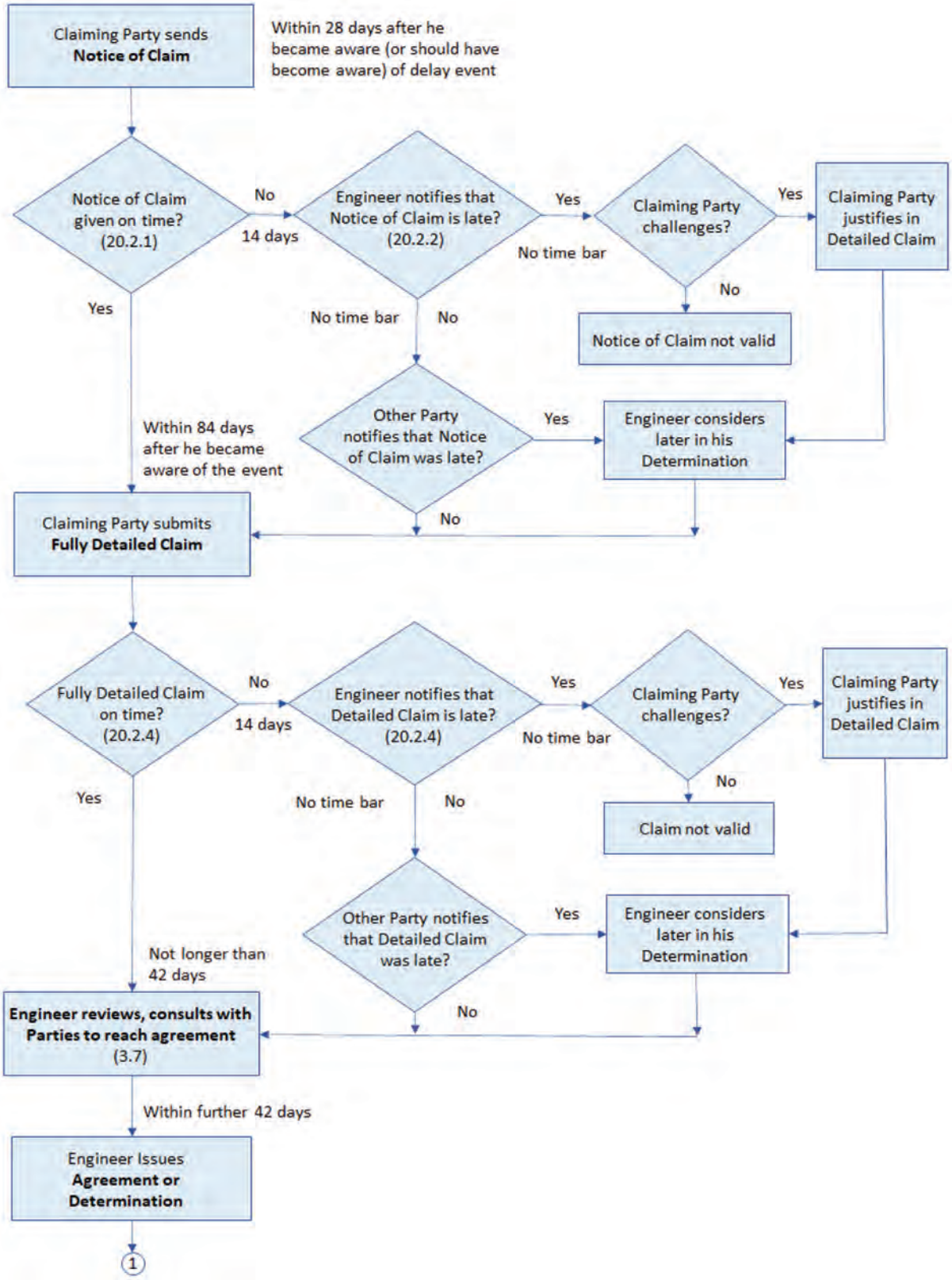
5. EOT CLAIM ANALYSIS TECHNIQUES FOR ADDITIONAL TIME AND COST

FIDIC 2017 Conditions of Contract, similar to many other contract forms, do not stipulate the preferred method for analysis of delays and calculation of EOT. Some recommendations, learned from experience and arbitration cases are summarized below.

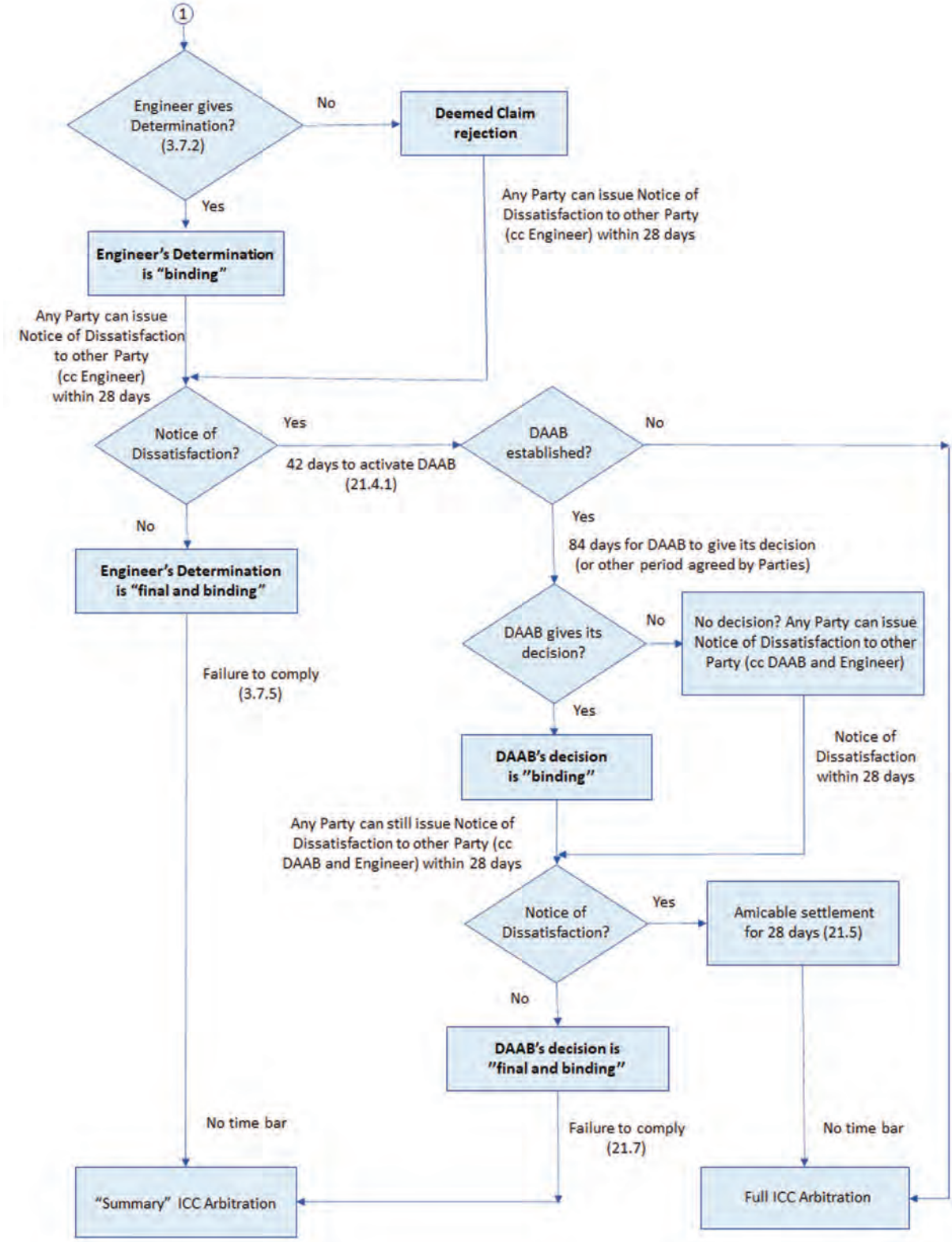
5.1. Quantification of additional time

Methods for delay analysis are reviewed and compared on Tab. 5.

Types and volumes of input data obviously vary between different methods, as shown on Tab. 6 (Contemporary Records, however, are always required).



(a) Claim submission, review and resolution procedures in FIDIC 2017 (part 1)



(b) Claim submission, review and resolution procedures in FIDIC 2017 (part 2)

Figure 5

Table 3. FIDIC 2017 Clauses that entitle the Contractor to raise claims for additional time and/or money

	Sub-Clause (Red Book)	Potential Event	Contractor's Entitlement		
			EOT	Cost	Profit
1.9	Delayed Drawings and Instructions	Delay in supplying the necessary drawing or instruction	✓	✓	✓
1.13	Compliance with Laws	Employer's delay in obtaining permits	✓	✓	✓
2.1	Right of Access to the Site	Late granting of access to construciton site	✓	✓	✓
4.6	Co-operation	Instruction for unforeseeable additional co-operation with others	✓	✓	✓
4.7	Setting Out	Error in setting-out data	✓	✓	✓
4.12	Unforeseeable Physical Conditions	Unforeseeable physical conditions at construction site	✓	✓	
4.15	Access Route	Changes to an access route	✓	✓	
4.23	Archaeological and Geological Findings	Fossils and other findings	✓	✓	
7.4	Testing by the Contractor	Delay in tests due to Employer	✓	✓	✓
7.6	Remedial Work	Remedial work prior to Taking-Over due to Employer	✓	✓	✓
8.5	Extension of Time for Completion	Variations, exceptional climatic conditions, shortages in personnel/goods, actions of Employer, his personnel or contractors	✓		
8.10	Consequences of Employer's Suspension	Suspension instructed by Employer (other than Contractor's default)	✓	✓	✓
10.2	Taking Over Parts	Employer taking over and/or using a Part	✓	✓	✓
10.3	Interference with Test on Completion	Disrupting interference with Tests on Completion	✓	✓	✓
11.7	Right of Access after Taking Over	Unreasonable delay by the Employer in permitting access		✓	✓
11.8	Contractor to Search	Search for defects that are not Contractor's fault		✓	✓
13.3	Variation Procedure	Contractor required to submit proposal for Variation which is not accepted		✓	
13.6	Adjustments for Changes in Laws	Changes in Laws after Base Date	✓	✓	
16.1	Suspension by Contractor	Suspension due to non-payment	✓	✓	✓
16.2	Termination by Contractor	Termination due to Employer	✓	✓	✓
16.3	Contractor's Obligations after Termination	Lost time and money during the notice period of 14d		✓	✓
16.4	Payment after Termination by Contractor	Losses and damages resulting from termination		✓	✓
17.2	Liability for Care of the Works	Rectifications due to 'forces of nature' or due to others	✓	✓	✓
18.4	Consequences of an Exceptional Event	Contractor's lost time and cost due to Exceptional Event	✓	✓	

Application of delay analysis methods is not always straightforward and results may vary, which may influence the outcome of EOT claim as demonstrated on a simple example below.

Example

The example project comprised construction of reinforced concrete bridge with two 18 m spans and access roads, under the FIDIC Red Book contract. The Time for Completion was 9 months. Contractor's method of work assumed sub-contracting the access road and prefabricating the main bridge girders off-site, and these activities had time reserves. Execution of works on site was on the critical path in the As-Planned (Baseline) Programme (Fig. 6).

Windows Analysis method was adopted for the initial review due to the nature of delay events and changes in the critical path. Four 3-month windows were defined for retrospective analysis.

Window 1 (0-3 months) included three delay events: (1) the Employer delayed access to site for 15d, which critically affected Contractor's mobilization; (2) the Contractor was 30d late in submitting sub-contractor prequalifi-

Table 4. Summary of time management procedures in FIDIC 2017

8.1	Commencement of Works	• Commencement Date within 42d after the Letter of Acceptance • The Engineer must give a Notice stating such date 14d in advance
8.2	Time for Completion	• The Contractor must complete the "whole of the Works" including details of Clauses 9 (Tests on Completion) and 10.1 (Taking Over the Works)
8.3	Programme	• The Contractor submits the initial Programme in 28d, revisions in 14d • Programme is a very detailed document (includes methods of work, sequencing, logical links, critical path, reserves, progress and delays) • The Engineer notifies about errors in the Programme (21d initial, 14d revisions) • The Contractor is the "owner" of the Programme, which is a "live" document
8.4	Advance Warning	• Parties should notify about potential adverse events (preventive action) • The Engineer can request for corrective measures (subject to a Variation)
8.5	Extension of Time for Completion	• Circumstances for potential Contractor's EOT claims: a) Variations to the Works (subject to compliance with Variation procedures) b) Causes of delay that are specifically mentioned in Contract Sub-Clauses c) Exceptionally adverse and unforeseeable climatic conditions (with proof) d) Unforeseeable shortages of personnel or goods (epidemic, government) e) Delays of the Employer's staff/contractors ("concurrent" delays possible!)
8.6	Delays by Authorities	• The Contractor must comply with local procedures and prove "unforeseeablity"
8.7	Rate of Progress	• The Engineer may notify and request for corrective measures (any other cause)

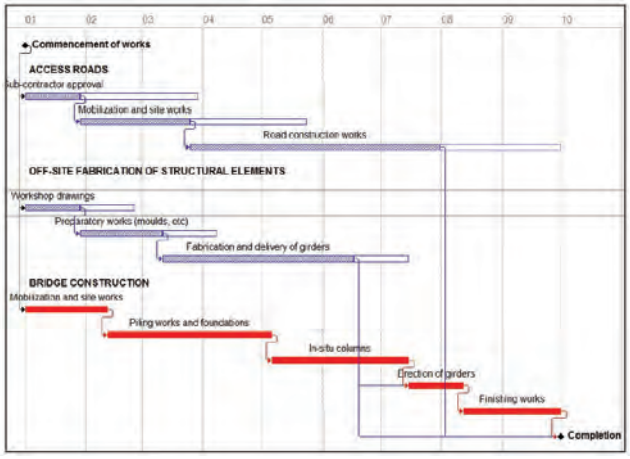


Figure 6. As-Planned Programme (Time for Completion is 9 months)

cation documents, but this activity was not on the critical path; and (3) after the completion of piling works, the Engineer instructed additional soil investigation to verify sub-soil condition, and approved an EOT of 15d.

Window 2 (3-6 months) comprised two delay events: (1) the Engineer instructed a change in access roads as directed by local authorities, resulting in additional works and non-critical delay (60d); (2) the Contractor himself delayed the prefabrication and delivery of bridge girders due to lack of materials (30d), but this was not on the contemporary critical path. The expected Completion Date did not change in Window 2.

Window 3 (6-9 months) witnessed no Employer's delay events, but progress reports recorded slow execution of road construction works due to lack of Contractor's resources (15d delay). This delay was on the contemporary (actual) critical path.

In Window 4 (9-12 months, As-Built Programme) there were no recorded delay events.

The analysis of Windows 1-4 is shown on Fig. 7 and summary of results is on Tab. 7.

Collapsed As-Built method was performed on the As-Built Programme (Window 4 above), which was used as a starting point to subtract the effects of all three Employer's delay events: (1)delay in providing access to site (15d); (2) additional soil investigation (15d); and (3) change in the route of permanent access road to the bridge (60d). The resulting Programme on Fig. 8 shows that project would have been completed 38d earlier but for these Employer's delays.

Impacted As-Planned method was performed on the initial As-Planned Programme, which was impacted by the same three Employer's delay events, and the cumulative effect of these delays on the Time for Completion appeared to be 15d, as per Fig. 9 below.

Tab. 8 tabulates different results produced by three retrospective delay analysis methods.

Differences in results suggest that any specific claim should be analyzed from different perspectives. None of the methods automatically favour the Employer or the Contractor. Window Analysis is able to assess the "net" effect of delay events using the "contemporary" (actual) critical path within each window, and should be used wherever possible, even though there is still no precise defi-

Table 5. Delay analysis methods

Method	How to Apply	Pros	Cons
As-Planned vs. As-Built	• Planned and As-Built Programmes are compared • Difference between completion dates of Planned and As-Built Programmes is a basis for EOT • Conclusions must be clearly supported (cause & effect analysis, narrative linking the delay events with EOT claim)	• Easy to understand • Best for simple programmes • Common sense approach is used in addition to software • Can be used when the As-Planned Programme and its updates are not perfect	• Requires experience and knowledge in order to establish and compare critical paths • Requires detailed review of project documents • Different experts may reach different conclusions, causing debates
Impacted As-Planned	• As-Planned Programme is analyzed • Delay events are inserted into the As-Planned Programme and critical path is analyzed to ascertain their impact • The resulting delay to completion date, if any, is a basis for EOT claim	• Simple to understand • Relatively fast to apply and review results • Suitable for smaller projects and where critical path remains unchanged (or little changed) over the duration of the project	• Does not consider changes in the network logic and contemporary critical path(s) • Not for projects with lots of variations and adjust-ments in methods of work (which is common) • Not suitable for analysis of concurrent delays
Time Impact Analysis	• Delay events are inserted as and when they occur into the As-Planned Programme • After each insertion, critical path is re-analyzed to assess impact of these events on the critical path • This method is best used during the course of the project to predict how delay events will affect the critical path and completion date	• Very useful to demonstrate effects of delay events in real time when they occur • Encourages a proactive approach to assessing EOT claims • Facilitates regular Programme updates and progress reporting	• As-Planned Programme should be realistic and fault-free • Requires detailed and accurate progress information • Can be challenging to pursue regularly on complex programmes, requires site resources
Collapsed As-Built	• As-Built Programme is first reconstructed and then analyzed • Delay events are isolated within the As-Built Programme and removed in the logical manner • Resulting earlier (theoretical) completion date is compared to actual; the difference represents a delay that is a basis for EOT • Delays are grouped based on responsibility and summarized	• The delays are factual (based on site documents) and actual (based on actual events and applied methods of work) • Analysis can be carried out even if the As-Planned Programme is not available	• As-built Programme must be reconstructed based on documents, which is a tedious and complicated work • Links in the As-Built Programme are introduced by the analyst, which may be subjective and cause of manipulation
Windows Analysis	• Focus is on a specified period or "Window" within the Programme (typically, one month) • Analysis is done on the updated Programme (data date is the beginning of the Window), starting from the As-Planned Programme • Within each Window, any of the delay analysis techniques can be applied to investigate events • Critical delays and responsibilities are tabulated for each Window to assess the EOT claim	• Due to regular updates at the beginning of each Window, analysis takes into account the actual or "contemporary" critical path(s), which reflects the actual methods of work • Aligned with standard monthly progress reporting procedures • Can be applied retrospectively	• Requires knowledge, experience and systematic approach • Document management system and quality of site records are critical (but it is also true for any other delay analysis method)

Table 6. Minimal inputs required for delay analysis

	As-Planned Programme	As-Built Programme	Progress Updates	Contemporary Records
As-Planned vs. As-Built	✓	✓		✓
Impacted As-Planned	✓			✓
Time Impact Analysis	✓	✓	✓	✓
Collapsed As-Built		✓		✓
Windows Analysis	✓		✓	✓

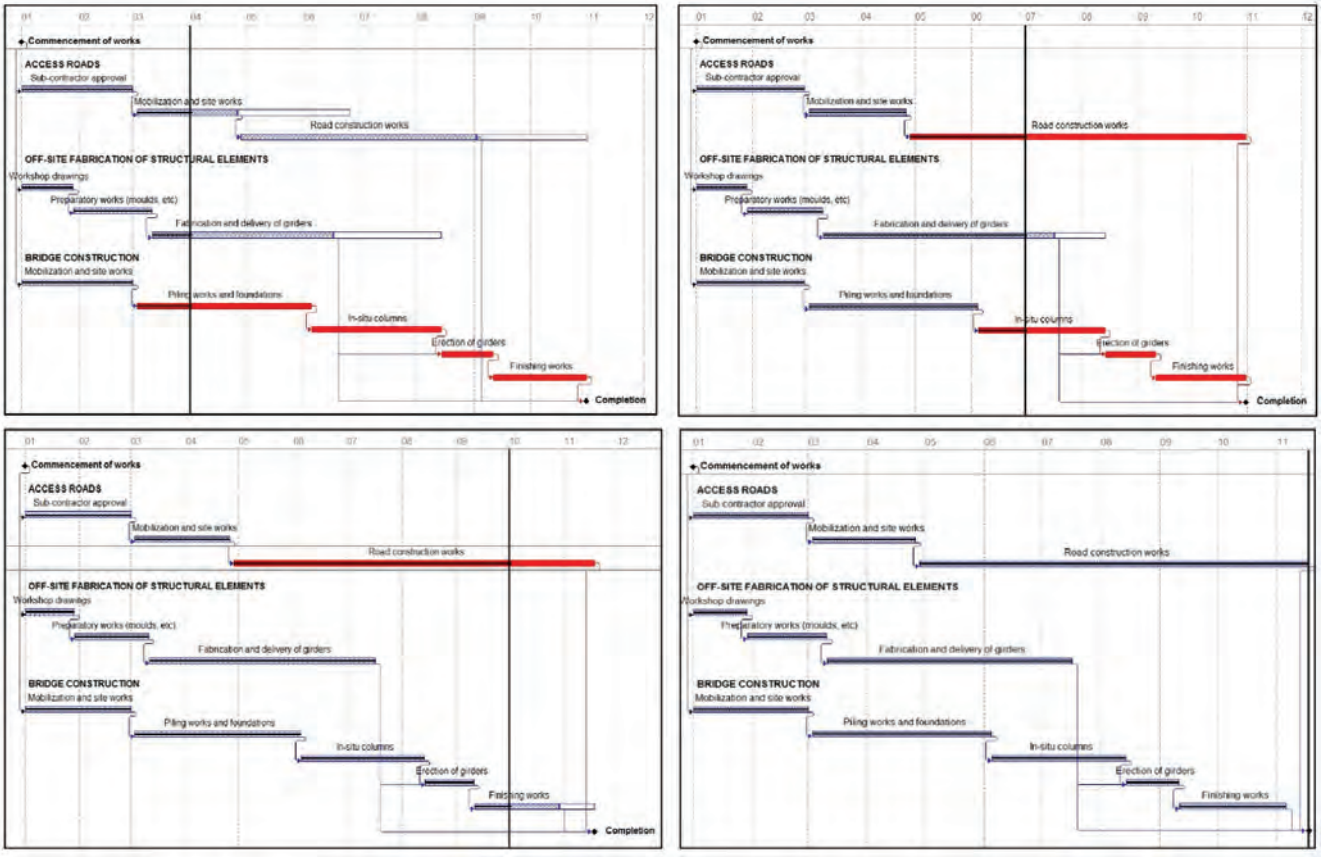


Figure 7. Windows Analysis: Window 1 (0-3 m, up left), Window 2 (3-6 m, up right), Window 3 (6-9 m, bottom left), Window 4 (9-12 m, the final As-Built Programme, bottom right)

nition amongst experts, see **Costain v Charles Haswell (2009)**.[22]

In FIDIC contracts, the windows analysis method can (and should) be applied progressively on a monthly basis using the updated programme. This preferred approach to delay analysis may elaborated in the Particular Conditions.

5.2. Quantification of additional cost

It is a common mistake to link time and financial compensation for delay. The SCL Protocol (2017)[23] states that *'entitlement to an EOT does not automatically lead to entitlement to compensation (and vice versa)'*. Additional monies may be recovered even without prolongation, and EOT may be granted without any additional money.

In FIDIC contracts, time and money claims should be notified and submitted separately, and Contractor's entitlement to Cost (defined in Sub-Clause 1.1.19 as *'all expenditure reasonably incurred including overheads and similar charges, but does not include profit'*) or Cost Plus Profit (defined in Sub-Clause 1.1.20 as *'Cost plus the applicable percentage for profit stated in the Contract Data or, if not stated, 5%'*) varies as per Fig. 1.

Delay costs are time-related costs associated with delays. There could also be additional **disruption costs** to cover the FIDIC-based **ICC Case 12654 (2005)**,[24] the tribunal explained that *'pure delay costs'* arise due to the pro-

[22] Costain Ltd v Charles Haswell & Partners Ltd [2009] EWHC 3140 (TCC)

[23] Society of Construction Law (SCL), Delay and Disruption Protocol, Feb 2017

[24] ICC Case No. 12654 (2005), published in the ICC Court of Arbitration Bulletin Vol 23/No 2-2012, page 38

Table 7. Windows Analysis Summary – Total project delay is 45d, out of which 30d are justified

Window	Delay Event	Event Duration	Respon-sibility	On Critical Path?	Project Delay	Justified EOT
No. 1 (0-3 m)	Late access to site	15d	Employer	Yes	15d	15d
	Late sub-contractor submission	30d	Contractor	No	–	–
	Additional soil investigation	15d	Employer	Yes	15d	15d
No. 2 (3-6 m)	Change in access road	60d	Employer	No	–	–
	Delay in prefabrication	30d	Contractor	No	–	–
No. 3 (6-9 m)	Lack of resources on site	15d	Contractor	Yes	15d	-
No. 4 (9-12 m) As-Built	No delay events	–	–	–	–	–
Total:					**45d**	**30d**

Table 8. Summary of delay analysis using three different retrospective methods

Method	Total Project Delay	Employer's Responsibility (EOT)	Contractor's Responsibility
Method 1 – Window Analysis	45d	30d	15d
Method 2 – Collapsed As-Built		38d	7d
Method 3 – Impacted As-Planned		30d	15d

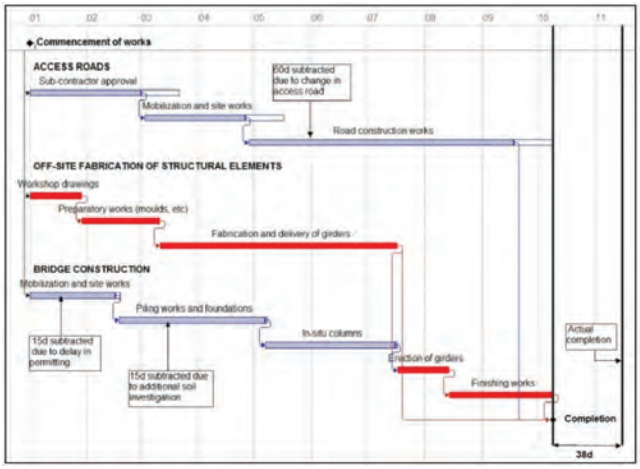


Figure 8. Collapsed As-Built - Completion would have been achieved 38d earlier without Employer's delays

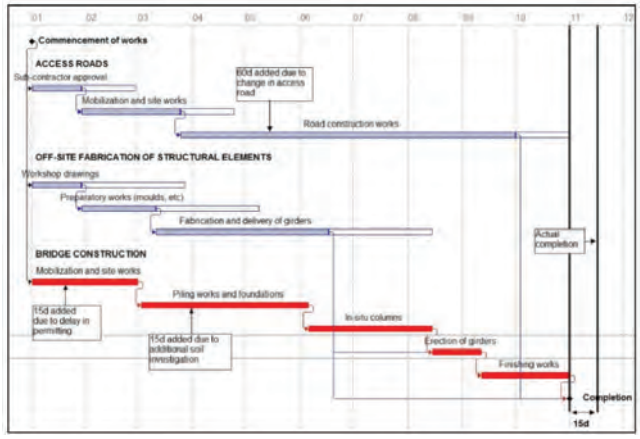


Figure 9. Impacted As-Planned – Employer's delays impacted the overall project completion date by 15d

longed use of resources, while *'disruption costs'* are linked to *'unproductive manner'* of work or *'out of sequence'* work.

Practice guides and own experience suggest several heads of claim for delay and disruption costs (sometimes also called loss and expense).

Site overheads (preliminaries)

Mobilization, running and demobilization costs may be identifiable within the 'Preliminaries' section of the Bills of Quantities (BOQ). Delay costs reflect the additional *running costs* during the EOT period (mobilization/demobilization costs are usually unaffected):

- site personnel (management, labour) and project-specific expenses (travel);
- site accommodation/offices, including furniture and IT equipment;
- site temporary facilities (stores, workshops, truck washing facilities);
- site safety and security facilities (fences, gates, safety barriers/ scaffolding, access equipment, cameras, etc);
- contractor's equipment (cranes, hoists, concrete plant, scaffolding, dedicated formwork);
- site temporary services (water, electricity, telecommunications, sewage, storm water drainage, cooling/heating), inclusive of maintenance/service charges;
- temporary protection of works and environmental measures (sampling, testing);
- site maintenance expenses (cleaning, removal of debris, government fees);

- extended insurances and bonds in accordance with the contract;
- other: extended storage, repeated testing and commissioning, extended warrantees.

Usually, actual costs incurred by the contractor need to be claimed in the absence of the prices from the BOQ, provided that resources were actually required on site (just presence of resources does not mean they are additional to the tendered ones!) and that contractor mitigated losses.[25] From the ***British Westinghouse v Underground Electric Railways* (1912)**,[26] the injured party must take all reasonable steps to mitigate losses, and cannot recover any loss which could have been easily avoided (mitigation will be further discussed).

Actual costs should rely on the contemporary records; unnecessary operating charges (repairs, parts, fuel) are not applicable, while rental costs should reflect the stand-by (idle) status. The costs should relate to the periods of delay and only to the activities that were actually delayed (if others progressed). The onus is on the contractor to prove that the costs claimed have been actually incurred and that every effort (mitigation) has been taken to minimise these costs.

Head office overheads

These are unrecovered costs of running the head office, such as: office rents, office running costs, salaries of office staff, legal/professional fees, administrative costs, depreciation.

Following **JF Finnegan v Sheffield City Council (1988),**[27] contractor's off-site overheads claim should show that *'the workforce, but for the delay, might have had the opportunity of being employed on another contract... during the overrun period.'* In **Peak Construction v McKinney Foundations (1970)**[28] the contractor was also asked to prove that there was other work/opportunity available which, but for the delay, the contractor would have secured.

The actual overhead costs must also be reasonable.[29] In **Tate & Lyle v Greater London Council (1981)**,[30] damages for head office expenses were justified, but could not be recovered because of the failure to keep proper records; the court wanted to know who at head office was involved and what was the record of their time spent on the overrunning project. **Babcock Energy v Lodge Sturtevant (1994)**[31] reinforced the contractor's entitlement to recover head office overheads based upon accurately recorded costs.

Head office overheads are sometimes calculated by **use of a formulas** (Hudson's, Emden's in UK, Eichleay in US), which have found limited judicial and arbitral approval. The use of formulas can lead to double-counting of overhead costs, especially where there are variations in works.

Loss of profit

In English law, economic loss must follow naturally from the breach or be within the reasonable contemplation of the parties at the time of contracting to be recoverable,[32] as per the second lag in **Hadley v Baxendale (1854)**;[33] loss of profit, being economic loss, can be treated as 'special damages' and included in the claim.

The contractor must show that he was prevented from earning profit elsewhere, as in **Peak Construction v McKinney Foundations (1970)**,[34] and demonstrate the actual loss of profit that he would have earned on other contracts had there been no delay. Rough estimates of profit margins are rarely sufficient. In **Wraight v PH&T (Holdings) (1968)**,[35] there was an employer's default and contractor was entitled to be reimbursed the amount of profit that he could earn if allowed to complete the works, less the amount saved because of the removal of his contractual obligation. In **Inserco v Honeywell Control Systems (1996)**,[36] the judge referred to Spon's industry norms and actual profit expectation in Inserco's business.

Actual loss of profit is difficult to prove and recovery may depend on the applicable law. In FIDIC forms, loss of profit is expressly excluded for some delay events (Fig. 1) but there is a default profit of 5% for recoverable cases as to prevent disputes.

Increased costs due to inflation

This type of claim must be a direct result of the prolongation; records must be kept to show the difference in labour/material prices between the actual and planned periods of work.

In FIDIC 2017, Sub-Clause 13.7 *(Adjustments for Changes in Cost)* allows for such price adjustments, provided that Schedule(s) of cost indexation are included in the Contract (otherwise, no adjustments for inflation).

[25] Ali Haidar and Peter Barnes, Delay and Disruption Claims in Construction (3rd ed, ICE Publishing, 2018)

[26] British Westinghouse Electrical Manufacturing Co Ltd v Underground Electric Railways [1912] AC 673

[27] J F Finnegan v Sheffield City Council [1988] 43 BLR 124

[28] Peak Construction (Liverpool) Ltd v McKinney Foundations Ltd [1970] 1 BLR 111

[29] Royal Institution of British Architects (RIBA), Good Practice Guide: Assessing Lass and Expense (2013), p 43

[30] Tate & Lyle Food and Distribution Ltd v Greater London Council [1981] 3 All ER 716

[31] Babcock Energy Ltd v Lodge Sturtevant Ltd (Formerly Peabody Sturtevant Ltd) [1994] 41 ConLR 45

[32] Ewan McKendrick, Contract Law – Text, Cases and Materials (5th edition, Oxford University Press, 2012)

[33] Hadley v Baxendale [1854] 156 ER 145

[34] Peak Construction (Liverpool) Ltd v McKinney Foundations Ltd [1970] 1 BLR 111

[35] Wraight Ltd v PH & T (Holdings) Ltd [1968] 8 BLR 22

[36] Inserco Ltd v Honeywell Control Systems Ltd [1996] 21 BLISS 3

This clause is frequently disallowed in practice, with some exceptions (e.g. copper rates on long-term power projects).

Increased costs for work under different climatic conditions

Loss/expense due to work in less favourable climatic conditions are recoverable in English law, as per **Bush v Whitehaven (1888)**.[37] The contractor needs to demonstrate that productivity was actually affected (disrupted), not just that the work was in less favourable climatic conditions.[38]

In FIDIC 2017, financial compensation for climatic Unforeseeable Physical Conditions is recoverable (Sub-Clause 4.12). It is necessary to prove that physical conditions were *'unforeseeable'* and provide supporting documents for actual costs, while the Engineer may examine other works for possible more favourable physical conditions and reduce costs accordingly to avoid taking advantage of varying site conditions.

Finance and interest charges

When a contractor incurs additional costs, this has to be financed either from his own capital or by borrowing. This is pure economic loss, similar to loss of profit, and may be claimed as 'special damages' under English law, subject to the remoteness test (damages directly linked to events).

The case of **FG Minter v Welsh Health Technical Services Organisation (1980)**[39] confirmed that interest on borrowed to cover contractor's loss/expense is recoverable under JCT contracts. In **Rees & Kirby Ltd v Swansea CC (1985)**,[40] the court assessed a similar claim on the basis of compound rather than simple interest. This was followed in Scotland in **Ogilvie Builders v Glasgow City DC (1994)**.[41]

In FIDIC 2017 contracts, finance and interest charges may be claimed under Sub-Clause 20.1(c) *'in connection with any certificate, determination, instruction, Notice, opinion or valuation'*, but applicable *Laws* may interfere, which is discussed later.

Disruption and forced acceleration

Disruption arises from the unproductive use of contractor's resources due to disturbance to the working methods, even without delay. Forced or 'constructive' acceleration (when contractor does not get the deserved EOT and is forced to accelerate) may invoke disruption as well.

There are several methods to quantify loss of productivity in a disruption claim. **The industry standard approach** (comparison of actual and standard/published productivity rates) and **the measured mile approach** (comparison of actual and non-disrupted productivity rates, but within the same project) are frequently used. In **Whittal Builders Company v Chester-Le-Street District Council (1985)**,[42] the court favoured the measured mile approach, which is also recommended by SCL Protocol. In **John Doyle Construction v Laing Management (Scotland) (2004),**[43] it was similarly acceptable to compare labour productivity actually achieved on site when work was largely free from disruption with disrupted productivity.

The principal problem is not so much in identifying the actual cost incurred, but in showing that, but for the disruptive event, the cost would have been less than actually turned out to be. This is often difficult to prove as contractor's own record of performance on a project is rarely perfect.

FIDIC 2017 contract forms do not elaborate too much on disruption. The Contractor should warn about *'events or circumstance which may adversely affect the work'* (8.4, Advance Warning), and may claim as *'entitlement or relief may be of any kind whatsoever'* (20.1 (c), Claims).

Sub-contractor EOT claims

Sub-contractor's EOT claims may be relevant, provided that they can be attributed to delays caused by the employer only. Passing sub-contractor claims directly to the employer could be very difficult, requiring analysis of sub-contract provisions, proof of payments to the sub-contractor and separate delay analysis.

Relevance of sub-contractor's claims in FIDIC 2017 Conditions of Contract critically depends on the applicable law, which is discussed later in more detail.

Claim preparation costs

Claim preparation is normally considered as part of project administration is generally not recoverable under English law, except where the issue proceeds to arbitration/litigation, see **James Longley v South West Thames (1984)**.[44] Compensation might be sought in arbitration where the request for further evidence involves unusually

[37] Bush v Whitehaven [1888] 52 JP 392

[38] Ali Haidar and Peter Barnes, Delay and Disruption Claims in Construction (3rd Edition, ICE Publishing, 2018)

[39] F G Minter Ltd v Welsh Health Technical Services Organisation [1980] 13 BLR 1, CA

[40] Rees & Kirby Ltd v Swansea City Council [1985] 30 BLR 1

[41] Ogilvie Builders Ltd v Glasgow City District Council [1994] 68 BLR 122

[42] Whittal Builders Company Ltd v Chester-Le-Street District Council [1985] 11 CLR 40

[43] John Doyle Construction Ltd v Laing Management (Scotland) Ltd [2004] Scots CS 141

[44] James Longley v South West Thames [1984] 25 BLR 56

heavy amount of managerial/specialist time,[45] see **Tate & Lyle v GLC (1981)**.[46]

FIDIC 2017 forms similarly convey that normal contract administration is part of Contractor's obligations and claim preparation costs should be excluded from submissions during the course of the project. For arbitration cases, recovery depends on actual circumstances.

5.3. Claim submissions

The **time-part of EOT claim** should contain these essential elements:

- Project and contract details;
- Delay event(s) and responsibility allocation;
- Contractual entitlement (contract clauses, law);
- Contractual procedures (warnings, notices, submissions);
- Cause and effect analysis (narrative of planned and actual events);
- Delay analysis (using appropriate method(s));
- Statement of claim (conclusion, days claimed);
- Substantiation (contemporary records).

The **cost-part of EOT claim** must show that cause of the delay is one that entitles the contractor, under the contract, to payment for the extra costs incurred. The factual evidence and the breakdown of costs must be clearly presented.

6. INTERPRETATION OF EOT PROVISIONS IN FIDIC CONTRACTS – PRACTICE AND TRENDS

The applicable laws may influence the way arbitrators interpret FIDIC EOT provisions. Laws are defined in Sub-Clause 1.1.49 to include all *'legislation, statutes, acts, decrees, rules, ordinances, orders, treaties, international law and other laws, and regulations and by-laws of any legally constituted public authority'*. Typical issues are discussed by comparing the English precedence law approach with the UAE civil law provisions. Supporting arbitration and court decisions and practical recommendations are highlighted.

Interpretation of contract terms

When the parties disagree about the particular clause, courts try to construe ('construct') the clause in order to give effect to it. English courts favour the "purposive" or "commercial" approach (contract wording is placed in context to be understood) as opposed to the 'literal' approach (subject only to the "contra proferentem" rule which interprets ambiguities against its creator). If the wording is ambiguous, the court may seek other evidence of intentions; in **Robertson v Jackson (1845),**[47] the phrase 'turn to deliver' the goods was constructed by referring to port procedures (custom of trade). In **Investors Compensation Scheme v West Bromwich Building Society (1998),**[48] the court looked for *'the meaning which a document...would convey to a reasonable man.'*

UAE construction contracts must comply the law where*'property is situated'* (UAE Civil Code, Article 19(2)). Furthermore, *'a mandatory provision (of law) shall take precedence over a contractual stipulation'* (Article 31), and *'ignorance of the law is no excuse'* (Article 29). If the contract is clear, it will be interpreted in accordance with its provisions (Article 259). If not clear, important becomes *'the common intention of the parties'* (Articles 257-266), a more subjective test than a "reasonable" man under English law.

The UAE law approach is qualified by Article 266(1): *'a doubt shall be interpreted in favour of the obligor'*, usually the contractor (especially if his bargaining power is inferior).[49] This is different than English "contra proferentem" rule, where the ambiguous clause is interpreted against the party who put forward the wording.

Good faith, ethical issues and role of the Engineer when assessing claims

In English contract law, there is no legal principle of good faith, for two reasons.[50] Firstly, parties are free to pursue their own goals in negotiating contracts. Secondly, there is concern that concept of good faith is too vague and uncertain. Nevertheless, in **Yam Seng v International Trade Corporation (2013)**,[51] the judge supported "fair dealing" and "reasonable standard of conduct" in a distributorship agreement regarding the provision of information. He stressed, however, that such approach depends on the context; in the construction context, he mentioned an example of consents/decisions that should not be withheld unreasonably.

In contrast, the UAE Civil Code implies Sharia principles of "good faith" and "fairness" (Article 246(1)): *'The contract must be performed in accordance with its contents, and in a manner consistent with the requirements of good faith'*. A failure to act in good faith is a breach of contract. Unlawful actions may include behaviours like

[45] John Murdoch, Will Hughes, Construction Contracts – Law and Management (4th ed, Taylor & Francis, 2008), p 234

[46] Tate & Lyle Food and Distribution Co Ltd v GLC [1981] 3 All ER 716

[47] Robertson v Jackson [1845] 2 CB 412

[48] Investors Compensation Scheme Ltd v West Bromwich Building Society [1998] 1 All ER 98

[49] International Comparative Legal Guides (ICGL), UAE Law, Adam Balchin and Euan Lloyd (Al Tamimi Co), 2014

[50] Fenwick Elliott LLP, Can you imply good faith into agreements made under English Law, Jeremy Glover, Mar 2013

[51] Yam Seng Pte Ltd (Singapore) v International Trade Corporation Ltd [2013] EWHC 111 (QB)

(Article 106): *'a) intentional infringement (of another's rights);... (b) contrary to Islamic Sharia, the law, public order, or morals;... (c) disproportionate harm that will be suffered by others;... (d) if it exceeds... custom'*. These provisions are sometimes relied upon by Contractors in construction claim disputes.

In FIDIC 2017, the Engineer *'shall consult with both Parties'* and *'make a fair determination'* (3.7), and could be liable for this. In **John Mowlem v Eagle Star Insurance (1992)**,[52] the architect was liable in tort under English law for wrongful interference with the contract. In Dubai's **Pacific Associates v Baxter (1988)**,[53] Engineer's unfair claim assessments under a FIDIC dredging contract were sanctioned.

Delay notices and warnings as condition precedent

Some contracts define claim notices as "condition precedent" and this is enforced in English law. The case of **Bremer Handelgesellschaft v Vanden Avenne Izegem (1978)**[54] defined that a notice provision is a condition precedent if: (1) it states the precise time for serving the notice, and (2) unless the notice is served within that time, the claimant loses his rights. In **Multiplex Construction v Honeywell Control Systems (2007)**,[55] the judge explained the purpose of delay notice: *'enables matters to be investigated while they are still current.'* and *'gives the employer the opportunity to withdraw instructions.'* This was confirmed in **Steria v Sigma Wireless Communications (2008)**,[56] with clarification that *'minutes of meeting prepared by third parties...did not constitute adequate notice'*.

FIDIC Sub-Clause 20.1 (Contractor's Claims) was clearly drafted as a condition precedent: *'If the Claiming Party fails to file Notice of Claim within... 28d, the claiming Party shall not be entitled'*. However, interpretations may vary under civil law provisions.

UAE Commercial Code[57] stipulates that commercial claims expire after 10 years (Article 95). This precedes any contractual time bars (Civil Code, Article 486(1)). Provisions of "good faith" and "unlawful exercise of a right" (Articles 246, 106) may also apply. Rejection of a valid claim due to a late notice or improper format (minutes of meeting) may be an act of bad faith, causing disproportionate harm to the Contractor. It may be an 'unjust enrichment' (Articles 318-319). **UAE Society of Engineers**[58] opined that claim notice as a condition precedent might not be enforceable under UAE law.

From the Employer's perspective, Contractor's failure to notify could equally be seen as an act of bad faith, depriving the Employer of any chance to apply corrective measures.

In FIDIC-based **ICC Case 15282 (2010)**,[59] the Contractor failed to comply with the notification time limit. Rather than dismissing the claim outright, the tribunal took a pragmatic approach and examined the Contractor's documentary evidence to determine whether this was sufficient, and rejected the claim on that basis and not on notice provisions.

It is therefore recommended that EOT claim should be reviewed based on its merits, rather than immediately rejected for non-compliance with the notice provisions alone. Insufficient notifications could be sanctioned as appropriate cost deductions in claim assessments.

Significance of contemporary records

Under English law, claim records were defined in the FIDIC-1987 case, **Attorney General for the Falklands Islands v Gordon Forbes Construction (Falklands) (2003)** as:[60] *'original or primary documents... produced or prepared at or about the time giving rise to a claim.'* It was not acceptable to simply produce witness statements after the event.

In the FIDIC-1987 **ICC Case 15282 (2010),**[61] the tribunal examined the claim and referred to the definition of 'contemporary records' in the *Falklands* case. The Claimant failed to present 'contemporary records' to support its claim and the tribunal rejected the claim.

FIDIC 2017 Sub-Clause 20.2.3 defines contemporary records similarly to the *Falklands* case *('prepared or generated at the same time, or immediately after, the event')*.

In the UAE, where "good faith" is enforceable, immediate claim rejections for procedural technicalities are not generally recommended. However, this does not negate the importance of maintaining contemporary records. In the **Dubai Court of Cassation Case (213/2008),**[62] the EOT was approved, but without additional costs due to the contractor's failure to demonstrate causation and evidence of actual idle resources and delay costs.

[52] John Mowlem & Co plc v Eagle Star Insurance Co Ltd [1992] 62 BLR 126

[53] Pacific Associates Inc and Another v Baxter and Other [1988] 44 BLR 33

[54] Bremer Handelgesellschaft mbH v Vanden Avenne Izegem nv [1978] 2 Lloyd's Rep. 113, per Lord Salmon

[55] Multiplex Construction v Honeywell Control Systems [2007] EWHC 447 (TCC)

[56] Steria Ltd v Sigma Wireless Communications Ltd [2008] 118 Con LR 177

[57] UAE Commercial Transactions Law No. 18 of 1993 ('Commercial Transactions Law', 'Commercial Code')

[58] UAE Society of Engineers, The legal effect of condition precedent described in Clause 20(1) of FIDIC 1999 in the context of England and UAE Courts, by Mahmoud Bader, 2014

[59] ICC Case No. 15282 (2010), published in the ICC Court of Arbitration Bulletin Vol 24/No 2 -2013, pp 53-54

[60] Attorney General for the Falklands Islands v Gordon Forbes Construction (Falklands) Limited [2003] 6 BLR 280

[61] ICC Case No. 15282 (2010), published in the ICC Court of Arbitration Bulletin Vol 24/No 2 -2013, pp 53-54

[62] Dubai Court of Cassation (213/2008) Commercial Appeal (19 January 2009)

The **UAE Evidence Act**[63], amended by **Law No. 36 (2006)**, allows electronic documents and signatures to serve as court evidence. **Abu Dhabi Court of Cassation (2010)**[64] held that electronic communication (emailed offer-acceptance) had the same evidentiary weight as physical document, highlighting the importance of electronic document management systems.

Approach to concurrent delays

The **SCL Protocol (2017)**[65] defines that *'true concurrent delay is the occurrence of two or more delay events at the same time, one an Employer Risk Event, the other a Contractor Risk Event, and the effects of which are felt at the same time' and 'both affect the critical path.'*

Court decisions under English law are inconsistent and there are three major approaches to granting an EOT.

The 'Malmaison approach' was originated in **Henry Boot Construction v Malmaison Hotel (Manchester) (1999)**[66]: *'If there are two concurrent causes of delay, one of which is a relevant event, and the other is not, then the contractor is entitled to an EOT...notwithstanding the concurrent effect of the other event'.* Under this approach, widely accepted in the industry, the contractor is given an EOT caused by the employer's event even if the contractor himself was responsible for another parallel delay event. The SCL Protocol similarly recommends that *'Contractor's concurrent delay should not reduce any EOT due.*[67]

The 'Dominant Cause approach" was introduced by Keating (1995)[68] in the absence of relevant law: *'If there are two causes, the plaintiff succeeds if he establishes that the cause for which the defendant is responsible is the effective, dominant cause.'* Which cause is dominant is decided by applying common sense standards. However, common sense was not sufficient in **H Fairweather v London Borough of Wandsworth (1987).**[69] The judge disagreed with the arbitrator's decision that the EOT should relate to the dominant cause and requested that each separate cause of delay is assessed individually.

The 'Apportionment approach' might apply in the absence of a dominant cause. In the Scottish case of **City Inn v Shepherd's Construction (2010),**[70] the judge decided '*to apportion the delay.'* The apportionment was also appropriate in **John Doyle v Laing Management (Scotland) (2004)**[71]. The apportionment approach has been criticized in England as being *'contrary to the principles within the SCL Protocol'* and in **Walter Lilly v MacKay (2012)**[72] the judge decided to follow the **Malmaison** case by granting an EOT for concurrent delays.

The cost of concurrent delay is clarified in SCL Protocol:[73] *'Contractor should only recover compensation if it is able to separate the additional costs caused by the Employer Delay from those caused by the Contractor Delay';* otherwise, *'the Contractor will not be entitled to recover those additional costs.'* Pickavance[74] and Knowles[75] similarly recommend that an EOT should be awarded **without monetary compensation**. FIDIC-based **ICC Case 12654 (2005)**[76] analyzed the Employer's delay in giving possession of the site and awarded the Contractor part of his claimed costs; the Contractor was able to demonstrate entitlement to some compensation.

The UAE Civil Code provisions of "good faith" (Article 246(1)) and "unlawful exercise of right" (Article 106) impose an obligation on the Engineer to take a fair account of Employer's concurrent delays. A judge/tribunal may consider responsibility of each party if *'the person suffering harm participated by his own act in aggravating the damage'* (Article 290), or may even 'apportion' liability *'if a number of persons are responsible for a harmful act'* (Article 291).

In the **Dubai Court of Cassation Case 266/2008**,[77] the employer claimed for delay and defects and the contractor made a counter claim for prolongation time/costs. The court expert found that a nominated sub-contractor, not under the main contractor's control, was a cause, and awarded an EOT. In another **Dubai Court of Cassation Case 1/2006**,[78] the contractor was not expeditious, but the employer delayed the commencement, instructed additional works and changed the use from residential to serviced apartments; the court granted an EOT for concurrent delay.

FIDIC 2017 introduces in Sub-Clause 8.5 *(EOT for Completion)* that concurrent delay is possible and should be addressed in Special Provisions, or otherwise assessed as per good practice. This guidance allows for amendment that contractor may receive an EOT for concurrent delay, but generally without prolongation costs; no party should benefit from its own mistakes.

[63] UAE Federal Law of Evidence No. 10 of 1992 ('UAE Evidence Act'), Law No. 36 of 2006 ('Electronic Transactions'

[64] Abu Dhabi Court of Cassation Case (2010), Decision on Using Electronic Communications as Evidence

[65] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Edition, Feb 2017, p 6

[66] Henry Boot Construction (UK) Ltd v Malmaison Hotel (Manchester) Ltd [1999] 70 Con LR 32

[67] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Edition, Feb 2017, p 6

[68] Donald Keating, Anthony May, Keating on Building Contracts (6th ed, Sweet & Maxwell, 1995), para 8.015-8.018

[69] H Fairweather and Co Ltd v London Borough of Wandsworth [1987] 39 BLR 106

[70] City Inn Ltd v Shepherd's Construction [2007] CSIH 68

[71] John Doyle Ltd v Laing Management (Scotland) Ltd [2004] BLR 295

[72] Walter Lilly and Company v MacKay [2012] BLR 503

[73] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Edition, Feb 2017, p 6

[74] Keith Pickavance, Delay and Disruption in Construction Contracts (3rd ed, LLP Publishing, 2005), p 352

[75] Roger Knowles, 200 Contractual Problems and their Solutions (3rd edition, Willey-Blackwell, UK, 2012), p 97

[76] ICC Case No. 12654 (2005), published in the ICC Court of Arbitration Bulletin Vol 23/No 2 -2012, p 38

[77] Dubai Court of Cassation Case 266/2008 (17 March 2009)

[78] Dubai Court of Cassation Case 1/2006 (16 April 2006)

Ownership of float in the programme

SCL Protocol (2017)[79] makes the float available to both the employer and the contractor (actually, to the project), and EOT should be granted on the 'first come, first served' basis: *'where there is remaining float in the programme...an EOT should only be granted to the extent that the Employer Delay is predicted to reduce to below zero the total float.'*

The SCL approach was followed in **Ascon Contracting v McAlpine Construction (1999)**.[80] McAlpine was the main contractor and Ascon was the structural works sub-contractor. The project was delayed by 9w, and McAlpine argued that 5w float was for their own delays. The judge rejected this and favoured the 'first come, first served' approach. In **Royal Brompton Hospital v Hammond (2002)**,[81] Judge Lloyd similarly held that the project owns the float, but recognised the potential unfairness to contractors when employer's delays occur first.

The 'first come, first served' approach is not universally accepted. **Thomas**[82] believes that *'any float in the contractor's programme is for the benefit of the contractor.* **'Pickavance**[83] discusses the American approach from **Natken v George Fuller (1972)**,[84] where neither total nor free floats should be used for employer's changes. **Knowles**[85] also suggests that float is for correcting contractor's own mistakes, although this is in contrast with SCL recommendations.

FIDIC does not expressly define the ownership of float. **Corbett and Richards**[86] suggest that wording of the FIDIC suite shifts the float ownership towards the Contractor, but this is not evidenced in arbitral awards.

In author's opinion, floats should be for the benefit of the contractor, but standard delay analysis methods usually consume them as delay events occur. To prevent this, the contractor should extract floats into "contractor's contingency/risk activities" and use them as needed.

Sub-contractor EOT claims

The sub-contractor should not claim for EOT directly against the employer as there is no contractual relationship between them. In FIDIC 2017, the Contractor *'shall be responsible for the acts or defaults of any Sub-Contractor'* (5.1, Subcontractors), including 'nominated sub-contractors'.

Nomination may disturb the line of responsibility. In English case of **Bickerton v North West Hospital Board (1970)**[87], the employer was liable to re-appoint a replacement nominated sub-contractor upon the termination of the original one. In contrast, in **Percy Bilton v Greater London Council (1982)**[88], prompt re-nomination by the employer was acceptable and the contractor was liable for delay of the new sub-contractor.

In the UAE civil law, similarly, the improper nomination might help the contractor. In the **Dubai Court of Cassation Case, 266/2008**,[89] the employer (not the main contractor, who had no control) was responsible for delays caused by a nominated sub-contractor. In the **Dubai Court of Cassation Case 213/2008**,[90] the contractor was not held liable for delays caused by the non-performing nominated sub-contractor, which he was forced to accept.

The UAE Civil Code allows a contractor to sub-let the works, unless prohibited by the contract or the nature of the works (Article 890) and a sub-contractor has no recourse against an employer (Article 891). However, nomination should not be imposed unreasonably.

Admissible delay costs and exclusion of consequential loss

Differences between 'direct' and 'consequential' losses are often discussed. English law follows that 'direct' or 'consequential' depends on whether a particular loss falls within the rules of natural flow from the breach and remoteness of damage. In **British Sugar v NEI Power Plant Projects (1997)**,[91] the Court of Appeal referred to **Hadley v Baxendale (1854)**[92] and interpreted that first leg (losses arising naturally) represented 'direct', while the second leg (special damages, which parties had in mind when entering into contract) represented 'consequential' losses.

In **Croudace Construction v Cawoods Concrete Products (1978)**,[93] a sub-contract to supply masonry blocks excluded liability for any 'consequential loss or damage'. The main contractor claimed for loss of productivity due to defective blocks and sub-contractor claims. The court held that, despite the clause, the main contractor was entitled to recover for all these items of loss.

[79] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Edition, Feb 2017, p 27

[80] Ascon Contracting Ltd v McAlpine Construction [1999] 43 BLISS 5

[81] Royal Brompton Hospital National Health Trust v Hammond etc [2002] BLR 255, [2002] All ER 801

[82] Reg Thomas, Construction Contract Claims (2nd Edition, Palgrave, UK, 2001), page 101

[83] Keith Pickavance, Delay and Disruption in Construction Contracts (3rd ed, LLP Publishing, 2005), p 335

[84] Natken & Co v George A Fuller & Co 347 F. Supp. 17 (W.D. Mo. 1972)

[85] Roger Knowles, 200 Contractual Problems and their Solutions (3rd edition, Willey-Blackwell, UK, 2012), p 88

[86] Cornerstone Seminars, Managing, Defending and Making Claims under FIDIC Contracts, Abu Dhabi, 2011

[87] Bickerton v North West Metropolitan Regional Hospital Board [1970] 1 WLR 607

[88] Percy Bilton v Greater London Council [1982] 1 WLR 794

[89] Dubai Court of Cassation Case 266/2008 (17 March 2009)

[90] Dubai Court of Cassation Case 213/2008 Commercial Appeal (19 January 2009)

[91] British Sugar plc v NEI Power Plant Projects Ltd [1997] 87 BLR 42

[92] Hadley v Baxendale [1854] 156 ER 145

[93] Croudace Construction Ltd v Cawoods Concrete Products Ltd [1978] 2 Lloyd's Rep 55

Similarly, in **McCain Foods v Ec-Tech (Europe) (2011)**,[94] 'consequential loss' was excluded from the liability in a contract for gas system installation, but employer's losses in production due to the failure of supplier were held not to be 'consequential'. Consequential losses in English law do not include those losses which are foreseeable at the time of signing the contract.

The UAE Civil Code defines loss or 'harm' (Article 283): *'(1) Harm may be direct or consequential; (2) If the harm is direct, it must unconditionally be made good, and if it is consequential there must be a wrongful/deliberate element'.* 'Consequential harm' may include loss of profit, but a 'deliberate element' needs to be demonstrated. **Dubai Court of Cassation** required a 'stronger' element than mere negligence.[95] Losses of profit have been awarded in UAE courts where it was foreseeable, similar to English courts. Moral damages for *'infringement of liberty, dignity, honour, reputation, social standing '* (Article 293(1)) may also be considered.

Express exclusions of 'consequential losses', which are common in contracts, may raise opposing opinions on what constitutes 'consequential losses', and should better be defined.

Pre-agreed delay damages

Under English law, the parties are free to pre-agree the liquidated damages (LDs), limiting the contractor's liability for delay damages and relieving the employer of proving the actual loss. English law establishes LDs as a genuine pre-estimate of damages, and not as a penalty, which is unenforceable. The penalty challenge was rejected in **Alfred McAlpine Capital Projects v Tilebox (2005)**[96] as the court found that $45,000.00/week was a reasonable pre-estimate close to the weekly rental value. In **Volkswagen Financial Services v Ramage (2007),**[97] however, the LDs clause was held to be a penalty and thus unenforceable

The UAE Civil Code provides legal ground for the parties to pre-agree the amount of LDs (Article 390): *'1) The parties may fix the amount of compensation in advance; 2) The court may... make the compensation equal to the loss and any agreement to the contrary shall be void.'* LDs have been previously reduced by UAE courts.[98] The **Abu Dhabi Case 25/24 (2004)**[99] clarified that LDs may be entirely set aside *'if the contractor succeeds in establishing the absence of loss.'*

In the **Dubai Court of Cassation Case 184/2008,**[100] the contract was for two buildings (AED4.9m, 12 months), but the work was initially delayed due to the redesign of foundations. The court approved an EOT, but also found that the contractor was responsible for further delays, so LDs were still applicable beyond the revised completion date.

FIDIC 2017 Sub-Clause 8.8 (Delay Damages) requires that Contract Data must state the daily and maximum damages (typically, 10% of the Contract Price) due from the Contractor if the Works are not completed in time, and this is usually viewed as an agreed pre-estimate of losses.

Acceleration of the works

SCL Protocol (2017)[101] defines that 'constructive acceleration' occurs when contractor is '*implementing acceleration measures to avoid the risk of liquidated damages as a result of not receiving an EOT' but warns that not all costs may be recoverable.*

Constructive acceleration was approved in **Motherwell Bridge Construction v Micafil Vakuumtechnik (2002),**[102] where Motherwell was a sub-contractor to Micafil under the FIDIC contract. Motherwell recovered premiums paid for labour overtime and disruption (10%), when they were pressurised by Micafil to keep work up to schedule without an EOT granted to them.

The **ICC Case 10847 (2003)**[103] resolved an EOT dispute under FIDIC-1987. The tribunal required from the Contractor *'to demonstrate that it had in fact accelerated because it had been denied and EOT and not merely that he had brought additional resources to the Site'.*

In the UAE, actions must be *'consistent with good faith'* and *'unlawful exercise of right'* is prohibited (Articles 246; 106). The Engineer has a duty to *'make a fair determination'* (FIDIC Sub-Clause 3.7). Any pressure on the Engineer not to award an EOT would be against the law: *'Any harm done to another shall render the perpetrator liable.'* (Article 282).

For 'constructive acceleration', there should ideally be an agreement or an instruction to accelerate. In FIDIC 2017, an agreement can follow the Contractor's proposal (13.2 Value Engineering), while re-sequencing can be instructed by the Engineer (13.1 Right to Vary). A contractor who is simply being harassed to accelerate, should try to get an instruction and/or negotiate an agreement, or at least notify the Engineer about the acceleration claim.

[94] McCain Foods GB Ltd v Ec-Tech (Europe) Ltd [2011] EWHC 66

[95] Julio Cesar Bueno (Editor), The project and construction review, Chapter 25 – UAE , Galadari, Dubai, p 292

[96] Alfred McAlpine Capital Projects v Tilebox Limited [2005] BLR 271

[97] Volkswagen Financial Services (UK) Ltd v George Ramage [2007] CTLC 119

[98] Al Tamimi & Co, UAE, Law Updates, The Middle East, construction and the law, by Adam Balchin, Feb 2013

[99] High Federal Court, Abu Dhabi, Case 25/24, 1 June 2004 (Civil)

[100] Dubai Court of Cassation Case 184/2008 (30 December 2008)

[101] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Edition, Feb 2017, p 41-42

[102] Motherwell Bridge Construction Ltd v Micafil Vakuumtechnik [2002] TCC 81 ConLR 44

[103] ICC Case No. 11847 (2003), published in the ICC Court of Arbitration Bulletin Vol 23/No 2 -2012, p 36

Mitigation of delays

Contractor's obligation to reduce the effect of delays may affect the EOT claim. In English law, **British Westinghouse v Underground Electric Railways (1912)**[104] set that injured party cannot recover any loss which could have been avoided by reasonable steps, which need not be costly as the other party is the wrongdoer; it is sufficient to issue notices and be 'reasonable'. In **DSND Subsea v Petroleum Geo-Services ASA (2000)**,[105] contractor's 'obligation to mitigate delays' did not include resequencing (commencing deep sea diving ahead of risers installation).

UAE civil law seems to follow similar principles through 'good faith' provisions. The Civil Code (Article 389) hints at mitigation when stating: *'If compensation is not fixed... the judge shall assess it in an amount equivalent to the damage in fact suffered at the time of occurrence'.*

FIDIC 2017 does not expressly mention mitigation, but Sub-Clause 7.1 specifies Manner of Execution *'in accordance with recognised good practice',* and 8.1 requires the Contractor to *'proceed with due expedition,'* suggesting optimal use of resources. Sub-Clause 8.4 (Advance Waring) reminds of cooperation. Mitigation may be clarified in the Particular Conditions.

Delays and suspensions due to non-payment

The employer has an obligation to pay for the works. In England, construction contracts under the **Construction Act 1996**[106] (s.104-106) must regulate the payments and contractor has the right to suspend work for non-payment (s.112(1)), entitling him to an EOT (s.112(4)). The **Local Democracy Act 2009**[107] provides for the recovery of costs due to suspension (s.112(3A)). Contractor's right to be paid can not be manipulated by under-certification. In **Henry Boot v Alstom Combined Cycles (2005)**,[108] the absence of a certificate didn't justify non-payment.

Under the UAE Civil Code, a contractor may rely on Article 247: *'each of the parties may refuse to perform his obligation if the other contracting party does not perform that which he is obliged to do.'* Article 414 similarly allows a *'person who is obliged to perform'* to *'refrain from so doing'* if the other person has not discharged his related obligations. However, suspension for non-payment is sensitive. **The Dubai Court of Cassation** held that 'good faith' is applicable to suspension; a party cannot suspend work if the other party has substantially discharged its obligations leaving only a minor part unperformed. Suspending for a minor breach is unlawful.[109]

FIDIC 2017 provides remedies where the Employer fails to make payment. The Contractor may recover financing charges (14.8 Delayed Payment), suspend/reduce the rate of work and claim for EOT/Cost/profit (16.1 Suspension by Contractor), or even terminate (16.2 Termination by Contractor). FIDIC allows for recovery of profit, which is more favourable than UAE law.

Global EOT claims

Global claims involve an overall evaluation of delay ('global delay') or loss/expense ('global cost') under a number of events, usually without cause-effect analysis. **SCL Protocol (2017)**[110] clarifies that global claim does not negate the need for records and analysis, and warns of risk of failing.

Under English law, global claims without a minimum of qualities are typically rejected. In **John Doyle Construction v Laing Management (Scotland) (2004),**[111] the judge considered global claims to be a risky business; any claim should establish: (1) the occurrence of an event, (2) that loss/expense was suffered, and (3) that loss/expense was caused by the event. In **London Underground Ltd v Citylink (2007),**[112] a global claim was rejected, but some sections for which causation and quantum could be established were allowed to be resubmitted.

Global claims are more often accepted in international cases (including FIDIC-based) where a reasonably fast and fair solution is needed.[113] The rules of common sense will normally be applied when dealing with such claims. Global approach might sometimes be justified for disruption claims, which are particularly notorious for cause-effect analysis.

Recently, in **Walter Lilly v MacKay (2012)**,[114] Judge Akenhead comprehensively reviewed the English law on global claims, and favoured the more pragmatic/commercial approach. He stated that there is nothing in principle "wrong" with global claims; however, there are added evidential difficulties to prove that the loss would not have been incurred in any other event (e.g. if the offer was so low that the loss would have occurred irrespective of the delay events).

[104] British Westinghouse Electrical Manufacturing Co Ltd v Underground Electric Railways [1912] AC 673

[105] DSND Subsea Ltd v Petroleum Geo-Services ASA and PGS Offshore Technology AS [2000] WL 1741490

[106] Housing Grants, Construction and Regeneration Act (UK)1996

[107] Local Democracy, Economic Development and Construction Act (UK) 2009

[108] Henry Boot Ltd v Alstom Combined Cycles Ltd [2005] EWCA Civ 814

[109] International Comparative Legal Guides (ICGL), United Arab Emirates – Construction & Engineering Law 2014, by Adam Balchin and Euan Lloyd, Al Tamimi & Company, 2014, para 3.10

[110] Society of Construction Law (SCL), Delay and Disruption Protocol, 2nd Edition, Feb 2017, p 27

[111] John Doyle Construction Ltd v Laing Management (Scotland) Ltd [2004] SLT 678

[112] London Underground Ltd v Citylink Telecommunications [2007] EWHC 1749 (TCC)

[113] Ali Haidar and Peter Barnes, Delay and Disruption Claims in Construction (3rd ed, ICE Publishing, 2018)

[114] Walter Lilly and Company v MacKay [2012] BLR 503

In UAE, in **Dubai Court of Cassation (213/2008),**[115] the project was delayed for 316d due to employer's actions. The court awarded an EOT, but refused to award expenses for idle resources without demonstrating the causation link between the fault and damages.

Although global claims are easier and less expensive to prepare, they impose a strong burden of proof on the contractor, especially under FIDIC conditions of contract. The global claim approach, however, is not forbidden and might be used in preliminary claim reviews.

7. CONCLUSION

Nature and root causes of time-related claims, especially those made by contractors, were analyzed and few key topics were discussed in detail: (1) quality of project programmes and reporting, (2) contract procedures for time management and claims recommended by FIDIC, (3) methods for analysis of extension of time (EOT) and associated costs, and, finally, (4) peculiarities in interpretation of EOT claim provisions in different jurisdictions. Illustrative examples and legal cases were used to demonstrate gaps in the current practice, establish trends and provide recommendations from the practitioner's point of view.

Overall, it is clear that proper management of EOT claims, more than ever before, requires a combination of technical, commercial and legal skills. Technical skills include scope management, project planning and progress reporting techniques (2D, BIM), management of site records. Commercial skills are necessary to analyze the claim scenario, select and apply methods for calculation of additional time/cost. Legal skills are increasingly required to comply with applicable laws, which FIDIC defines as a sum of legislation *and* by-laws of service providers, as these may affect the way FIDIC provisions are interpreted for claim purposes. Modern management of time-related claims can not succeed without combining all of the above skills.

In parallel with increased "legalisation" and "commercialisation" of construction projects, there is also a trend of promoting collaborative approach in construction contracts, as opposed to creating disputes. Such an approach emphasizes fairness and attracts investors, large contractors and banks towards using standard contract forms. Recently, in July 2021, FIDIC announced work on a new Collaborative Contract form for partnering procurement arrangements on international projects, showing the intention to further develop the trend. It is a fair approach and supporting best practices that motivated this article in order to remind of the ultimate project goal - to complete a project within the agreed duration and price.

[115] Dubai Court of Cassation (213/2008) Commercial Appeal (19 January 2009)

REFERENCES

Al Tamimi, E. (2003). *Practical guide to litigation and arbitration in the united arab emirates*. Kluwer Law International, UK.

Baldwin, A. and Bordoli, D. (2014). *Handbook for construction planning and scheduling*. John Wiley & Sons.

Barr, B. and Grutters, L. (2014). *Fidic users' guide*. ICE Publishing, UK. (originally by Brian Totterdill)

Chappell, D., Powell-Smith, V. and Sims, J. H. (2011). *Building contract claims*. John Wiley & Sons. (5th Edition)

Corbett, E. (n.d.). *Fidic 2017 – a practical legal guide*. Corbett & Co, UK.

Corbett-Jarvis, N. and Grigg, B. (2017). *Effective legal writing: a practical guide*. LexisNexis Butterworths.

Gibson, R. (2008). *Construction delays: Extensions of time and prolongation claims*. Taylor & Francis.

Haidar, A. and Barnes, P. (2011). *Delay and disruption claims in construction*. ICE Publishing, UK.

Ivković, B., Popović, Ž. and Stojadinović, Z. (2021). *Upravljanje projektima u gradjevinarstvu, (4th ed)*. Apostrof.

Jaeger, A.-V. and Hök, G.-S. (2010). *Fidic-a guide for practitioners*. Springer-Verlag, Germany.

Keith, P. (2005). *Delay and disruption in construction contracts (3rd ed)*. LLP Professional Publishing.

Kerzner, H. (2011). *Project management metrics, kpis, and dashboards: a guide to measuring and monitoring project performance*. John Wiley & Sons.

Knowles, J. R. (2012). *200 contractual problems and their solutions (3rd edition)*. Willey-Blackwell.

McKendrick, E. (2012). *Contract law: text, cases, and materials (5th edition)*. Oxford University Press (UK).

Mullen, J. and Davison, P. (2020). *Evaluating contract claims (3rd edition)*. Wiley Blackwell.

Murdoch, J. and Hughes, W. (2008). *Construction contracts – law and management (4th ed)*. Taylor & Francis, UK.

Popovic, Z. (2016). *Claims for extension of time in fidic construction contracts*. A Practical Approach under UAE Law, Higher Colleges of Technology, Abu Dhabi, UAE.

Robinson, M. D. (2011). *A contractor's guide to the fidic conditions of contract*. John Wiley & Sons.

Robinson, M. D. (2013). *An employer's and engineer's guide to the fidic conditions of contract*. John Wiley & Sons.

Solicitors, M. (2008). *Macroberts on scottish building contracts (2nd edition)*. Blackwell Publishing.

Stone, R. and Devenney, J. (2011). *The modern law of contract (9th edition)*. Routledge.

Uff, J. (2013). *Construction law (11th edition)*. Sweet & Maxwell, UK.

INDUSTRY STANDARDS, GUIDES AND REPORTS

AECOM. (2020). *Middle east construction handbook.*

Chartered Institute of Building (CIOB). (2009). *Managing the risk of delayed completion in the 21st century.*

Denton Wilde Sapte. (2006). *Construction law uae, first edition 2006.*

EC Harris. (2013). *Global construction disputes: A longer resolution.* (Report 2013)

HKA. (2021). *Claims and dispute causation - a global market sector analysis (crux report).*

Institution of Civil Engineers (ICE). (2011). *Ice manual of construction law (thomas telford, uk, 2011).*

International Comparative Legal Guides (ICGL). (2014). *Construction & engineering law 2014.* Al Tamimi & Co.

Law Business Research. (2011). *he project and construction review, chp 25 – uae.* Galadari & Associates, Dubai.

of Chartered Surveyors (RICS), R. I. (n.d.). *Damages for delay to completion (1st ed).*

Project Management Institute. (2011). *Practice standard for scheduling, 2nd edition.* PMI, USA.

Royal Institution of British Architects (RIBA). (2008). *Good practice guide: Extensions of time (riba, 2008).*

Royal Institution of British Architects (RIBA). (2013). *Good practice guide: Assessing loss and expense (riba, 2013).*

Royal Institution of Chartered Surveyors (RICS). (2014). *Extensions of time (1st ed, rics guidance note.*

Society of Construction Law (SCL). (2017). *Delay and disruption protocol, 2nd edition.*

World Steel Association. (2020). *Middle east construction handbook.* https :// www .worldsteel .org / steel -by -topic / statistics / World -Steel -in -Figures .html. Brussels, 2015. (accessed 7 July 2021)

Dušan Prodanović

Dušan Prodanović is a Full Professor at the Faculty of Civil Engineering, University of Belgrade, Head of the Hydraulic and Environmental Engineering Department since 2016, former Director of the Hydraulic and Environmental Engineering Institute, President of the Council of Scientific Fields of Civil Engineering 2018-2019, member of the Assembly of the Institute "Jaroslav Cerni" since 2017 and President of the Assembly since January 2021.

Contact information: e-mail: dprodanovic@grf.bg.ac.rs

https://orcid.org/0000-0003-0156-7271

Prof. Prodanović was born in 1960 in Crkvice (Zenica, BiH). He graduated from the Faculty of Civil Engineering, University of Belgrade (FCEUB) in 1985, at the Department of Hydraulic Engineering. He received his MSc degree in 1991 and his PhD in 1999 (both from FCEUB) in the field of hydroinformatics. He has been employed at FCEUB as an assistant (1986-2000), assistant professor (2000-2008), associate professor (2008-2013) and full professor since 2013.

He was advisor in 4 PhD and 1 MSc thesis and a number of graduate and master's theses. He also participated in the research and defense of 8 theses. He authored and co-authored 6 printed textbooks, was involved in the development of the laboratory for fluid volume measurement at FCEUB, as well as in the development of a number of teaching and laboratory installations, knowledge innovation courses, as well as in the establishment of a specialist course in 2017.

He works in the fields of: measuring hydrotechnical quantities in steady and unsteady flow conditions, development of sensor technology, development of measurement and control programs, metrology implementation, application of GIS in problems of drainage of urban and rural areas, development of new modules for connecting GIS packages with simulation programs, connecting databases with simulation models, problems of pre- and post-processing of measurements in various fields of hydraulic engineering and methods of assimilation of on-line measurements and simulation models.

He co-authored 277 papers at international and domestic conferences, 29 papers in peer-reviewed international journals, 55 papers in domestic journals, 9 chapters in international books, 9 national monographs and chapters. His works are cited 545 times ($h = 11$). He serves as an Editor of Water Supply Journal and Urban Water Journal and as a reviewer of 46 papers for 9 international journals. He participated in 7 international projects, developed lecture material at 2 universities and several international courses. He attended four universities, was a co-advisor on one doctoral dissertation (Imperial College, UK), and wrote chapters for four international monographs. He was organizer of the 9th International Urban Drainage Modeling Conference (Belgrade, 2012).

He was involved in 78 professional projects (studies, technical solutions and conceptual/main projects). He is a licensed engineer and a member of the Serbian Chamber of Engineers since 2003. He has participated in solving numerous problems in industry, and in the development of new structures. He serves as a Member of the Accreditation Body of Serbia, technical expert for flow (until 2020).

He was the director of the Institute of Hydrotechnics and Water Ecological Engineering for three terms, president of the Council of Scientific Areas of Construction and Urban Sciences (2018-2019), president of the Assembly of the Institute "Jaroslav Cerni" from January 2021 (member since 2017). He is a Deputy Chairman of the Main Scientific Committee for the regulation, protection and use of water, land and air. From 2002, he has managed three domestic scientific projects.

He received October Award of the City of Belgrade for graduate work (1985), October Award of the City of Belgrade for the best defended master's thesis (1991) and the first award of the Chamber of the City of Belgrade for the best dissertation (2001).

Flow measurements in non-standard conditions

Dušan Prodanović

University of Belgrade, Faculty of Civil Engineering, Chair of Hydraulic and Environmental Engineering

Summary

The flow rate of water, and the accuracy used to measure it, is important to know. Numerous international and national standards exist, defining the conditions for flow measurement. However, the needed "standardized flow conditions" are often impossible to satisfy in field and they can be used just as guidance. This paper presents four cases with such non-standard conditions, where the author had to perform the flow rate measurement. For each case the approach used is explained, the results obtained are given and the achieved accuracy is assessed. All measurements were performed using electromagnetic (EM) velocity measurement probes, specifically designed to suit local conditions. Parallel to direct velocity measurement, CFD (Computational Fluid Dynamics) was used to analyse the flow conditions and get better insight into the flow field, with an idea to extract the data from modelled flow field in order to perform the flow meter calibration. Throughout the presented cases, which cover last two decades, clear developments of both hardware and knowledge can be seen.

Keywords: flow measurement, non-standard conditions, Electromagnetic probes, CFD

1. INTRODUCTION

Knowledge about water flow rate, with acceptable accuracy, is fundamental for the whole hydro-ecological engineering as well as for other scientific fields. Different measuring techniques are developed to optimally suit the environmental conditions at the measurement site. Available measuring hardware varies in terms of accuracy, cost and robustness (Prodanović, 2007). In situations with the full-pipe (pressurized) flow, the flowing cross-sectional area is known in advance and usage of more accurate equipment is possible (with accuracy of even 0.5% or better, Baker 2002). For the free surface (open channel) flow, or a mixture of free surface and pressurized flow (mixed flow conditions), the flow cross-sectional area is variable and has to be measured together with velocity and with the minimum and maximum flow rate range of 1 : 1000 or even higher (Hager, 2010)). The expected accuracy of flow measurement in such conditions is much below 1%.

It is important to accurately measure flow rate for many reasons. Water (as well as other more precious fluids) has its price; distribution of water among users is not only an economical issue but more often a political one; water deficiency, or more often its abundance in floods, is a key development component for certain area; water quality is directly related to flow rate; technical parameters of the systems depend on flow and velocities, etc. Numerous international and national standards exist, defining the conditions for flow measurement and applicable measurement system. Each standard insists on providing "standardized flow conditions" that can easily be replicated and quantified. For the most measurement methods, the long straight upstream and downstream sections are needed, with stable and steady flow conditions throughout the whole reach, for example. Or, in most volumetric devices, the minimum working pressure has to be maintained to prevent cavitation.

In reality, the standards for certain type of flow measurement could be used just as a guidance, since the standardized flow conditions are hard to fulfil (Simonović, 1990). Such situations require adaptation for specific flow conditions (Bertrand-Krajewski *et al.*, 2021). For example, the international standard ISO 15769 (ISO, 2010) recommends specifications for Doppler ultrasonic (US) sensors when used for flow measurement in sewers. The straight approach section of constant flow conditions required by the standard is too long, and in most real situations the sensor is installed just upstream of a cascade in manhole. Another example are the measurements of the turbine hydraulic efficiency in hydropower plants (HPPs) are defined by guidelines (Performance Test Code, 2002) and standards (IEC 60041, 1999). Flow measurement has to be performed within the full pipe (tunnel) flow with a long straight reach, with predefined large number of current meters in cross section. However, in Kaplan or bulb turbines, with short intake structures, there is no possibility to fulfil the standard's recommendation.

To assess the flow measurement uncertainty in non-standard conditions, some redundant system is needed

(Chaundry, 2008). In most situations, the semi-integrative or velocity sampling temporary techniques could successfully perform this task. The most used ones are the ultrasonic (US) transit-time or Doppler profiling devices and electromagnetic (EM) flowmeter probes.

This paper presents several cases where the author was involved in the flow rate measurements in non-standard flow conditions for which there was no ISO or EN standard that can be directly applied. All measurements were conducted using EM type flowmeters, with different probe's design, fitted for specific conditions. Although the US devices are more common in recent years and the author is normally using them in certain more standard conditions, the EM probes were used because of their robustness, clear physically based principle of direct measurement of water's velocity, acceptable price and, maybe the most important reason, good cooperation with the company that produces high quality EM probes (Svet Instrumenata, 2021).

Throughout the presented cases, which cover about two decades of work, clear developments of both the hardware and knowledge of the experts from Hydraulic and Environmental Engineering department, Faculty of Civil Engineering at University of Belgrade is evident.

2. METHODS

2.1. Velocity measurement using EM

The operating principle of electromagnetic (EM) flow/velocity sensors is based on the Faraday's law of induction (Fig. 1, left side). The motion of the conductive fluid (Fig. 1, right side) through a transversal magnetic field generates a voltage (Shercliff, 1962). To allow for the stationary analysis of the electromagnetic induction phenomenon, some electric and magnetic properties of the environment are assumed (Michalski *et al.*, 2001). Originally, under these assumptions, Kolin (1936) has given the basic relationship for the EM theory (1):

$$\nabla^2 E = \mathrm{div}\left(\vec{V} \times \vec{B}\right) \tag{1}$$

where $\vec{V}$ is the streamwise velocity field, $\vec{B}$ is the magnetic induction and $\mathrm{div}(\vec{V} \times \vec{B})$ is treated as a charge distribution. The raw output signal is the voltage $E = E^1 - E^2$, induced between the electrodes of the EM sensor. The relations used in the electrical networks motivated an idea to describe how each part of the flow contribute to the output voltage E. Equation (1) can be written as integral within the control volume τ (2):

$$E = -\int_{\tau} \left(\vec{V} \times \vec{B}\right) \cdot \vec{j}\, d\tau \tag{2}$$

or with the weight function w (Shercliff, 1962) using the weight vector $\vec{W}$ (Bevir, 1970):

$$E = \int_{\tau} \vec{V} \cdot \left(\vec{B} \times \vec{j}\right) d\tau = \int_{\tau} \vec{V} \cdot \vec{W}\, d\tau \tag{3}$$

where the cross product $\vec{B} \times \vec{j}$ defines Bevir's weight vector $\vec{W}$, τ is the control (sampling, or integrating) volume of the EM sensor (Fig. 1, right side) and $\vec{j}$ is the virtual current vector (i.e., the current density set up in the liquid by driving an imaginary unit current between a pair of electrodes). Since the Faraday's law of induction is governed by the right-hand rule, the dominant contributor to the output E is the longitudinal component of the velocity vector, V_x, or streamwise component, which is needed for flow measurement.

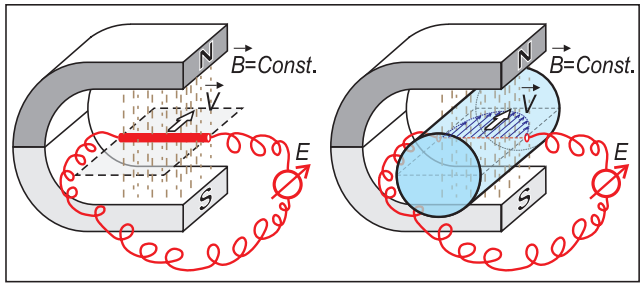


Figure 1. Basic Faraday's law of induction (left side) applied to (conducting) liquid flowing through the full pipe within the magnetic field (right side)

EM principle for flow measurement in full pipe allows accuracies of 1% to 0.5%, and with special attention, it can achieve the accuracy of 0.2%. However, the EM devices (or EM flow meters) are calibrated on flow rigs, mostly using volumetric systems. The actual velocity field and true magnetic field distribution within the profile (1)–(3)) are not important, since the calibration is carried out for integral component $E = f(Q)$.

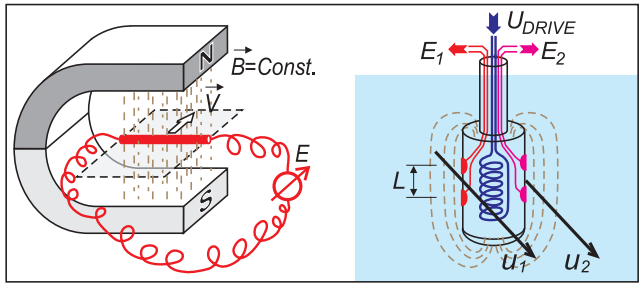


Figure 2. Same principle used for velocity measurement using inserted EM probe

The same Faraday's principle of induction can be applied from within the water, in order to measure the average velocities within the small sampling volume. Fig. 2, right side, presents the construction of such device (or EM probe): the EM coil used to generate the magnetic field is inserted into the water, and electrodes are relatively close, making the control volume τ (2) close to the inserted probe. There are numerous constructions of such device. Fig. 2, right side, presents the probe that measures the same velocity component in two adjacent control volumes, u_1 and u_2, producing electrical output signals E_1

and E_2. By averaging those two components, bigger sampling volume is used with more stable signal, or by differentiating two components, Reynold's shear stress can be calculated.

The EM probes are calibrated for velocity, using either standard towing tank where probe is moved with controlled speed through the still water (ISO, 2007) or in potential core of water jet. The basic accuracy of velocity measurement is 1-2%, while for better probes it can be up to 0.5%. The main benefit of EM probes is a wide measurement range of velocity, from several cm/s to 10 m/s. With longer averaging times (10–20 s) the EM probes can measure velocity down to 1 mm/s.

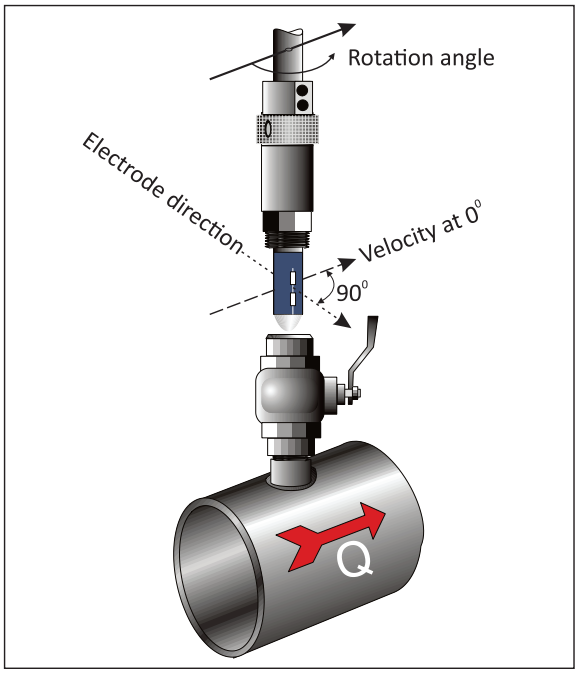


Figure 3. 1D EM probe (LOG-type) installed in pressurized pipe

The EM probe is not a contactless measurement device, i.e., it will disturb the true velocity field with its inserted body to a degree depending on the probe's shape. It is important to take this into account during both the calibration process and probe application. Different geometries of EM probes for velocity measurement, and different velocity components could be found on the market, and it is up to the user to select the optimal one:

- Cylindrical rod with 1D or 2D velocity measurement, that penetrates the fluid flow field (Figs 3–5). This type of probe is commonly named "LOG probe" since axisymmetric log-law velocity distribution is assumed and the flow is computed from measuring the velocity in one point.
- Specifically shaped LOG probes, for example one shown in Fig. reff26, or "MF pro – Water Flow Meter" made by OTT (2021).
- Flat "almost contactless" 1D, 2 × 1D or 2D probes designed to be fastened on pipe wall (Fig. 7), suitable for flow measurements in sewers.
- Full 3D spherical probes (Fig. 8) for measurement of all velocity components.

1D LOG-type EM probe is mostly used in water supply systems since it is cheaper and less accurate version of the EM full pipe flow meter. Probe can be inserted into the pipe under working pressure, without stopping the flow, and can be used to check actual velocity distribution by moving the probe along the pipe's diameter (so called "velocity profiling"). The probe's diameter is from 12 mm (for smaller pipes, to reduce the probe's blockage of the flow), mostly is 18 mm (with lengths in the range 300 mm – 1300 mm) and can be up to 50 mm for larger pipes with higher velocities, where oscillation of thinner probe is possible.

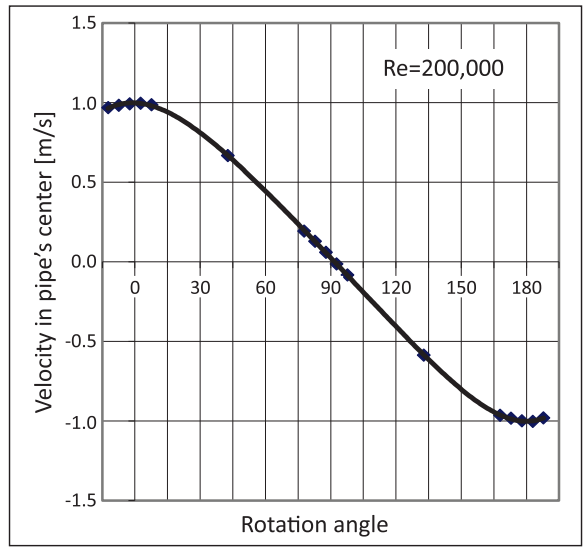


Figure 4. Angular 1D EM probe's sensitivity

Since the working principle of EM probes is based on vector multiplication of two fields (2), the cylindrical shape of probes allows the almost perfect cosines angular sensitivity (Fig. 4) covering direct and reverse flows. The true velocity direction can be achieved by rotating the 1D probe, searching for maximal velocity. If two more pairs of electrodes are added to the probe (Fig. 5), full 2D measurement is possible for velocity in the plane orthogonal to the probe's axis.

Main benefit of using EM LOG probes is that the user can either position the sampling volume in the area where mean velocity is expected, or move the probe along its axis and obtain the velocity distribution (velocity profiling). For permanent installation, special care has to be taken for possible debris and blockage of probe, so specific shapes are used in such situations (Fig. 6).

In sewer systems with highly polluted and dirty water, with lot of debris and junk, flat EM probes are developed that are installed on the inner side of pipe's wall (Fig. 7) or

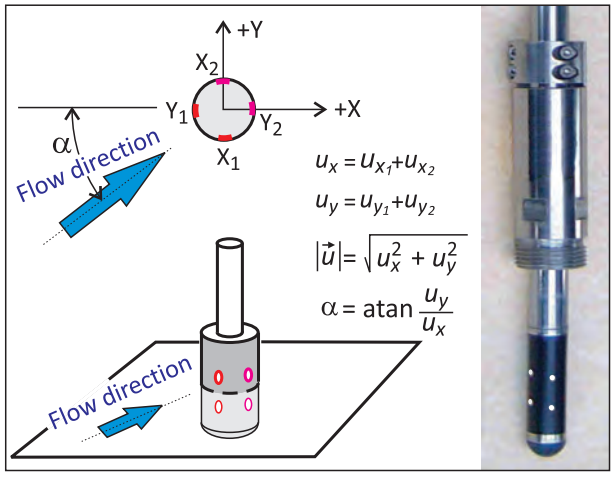


Figure 5. 2D EM LOG probe for velocity measurement (intensity and direction) in orthogonal plane

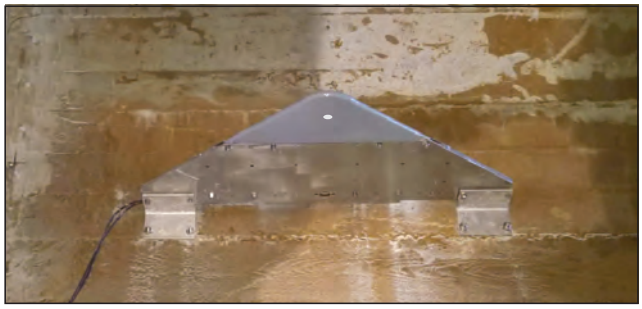


Figure 6. Different types of EM probes: 1D LOG-type for fixed installation in large tunnels

can be inserted from the outer side to be flush mounted, thus reducing the penetration. The length and width of the probes can be different, ranging from small ones (150–200 mm of length and 50 mm of width) for measurements in home sewer systems, medium size (Fig. 7) for larger sewers and sewer tunnels, and up to 700 mm (Fig. 18) for big tunnels.

Flat probes are mostly used for flow measurement in mixed conditions (pressurized and open flow) and they can be equipped with pressure sensor for water level measurement. The major drawback of flat EM flowmeter is limited depth of magnetic field and hence the near-by integration zone. For large pipes or tunnels, this can be a fraction of the "wet" cross section. Also, additional measures need to be taken in order to reach the desirable level of accuracy in mixed flow conditions (Ivetić, Prodanović and Stojadinović, 2018). One of the possibilities is to add more flat probes around the diameter, increasing the stability and reliability of flow measurement. Depending on the implementation, the price of such measuring site can be kept low if only one logger is used to drive all those probes (on the expense of losing the information on velocity distribution). To resolve the unknown velocity-pattern irregularity, the Site-Specific Calibration (SSC) is needed (Ivetić *et al.*, 2017b). It can be done using redundant

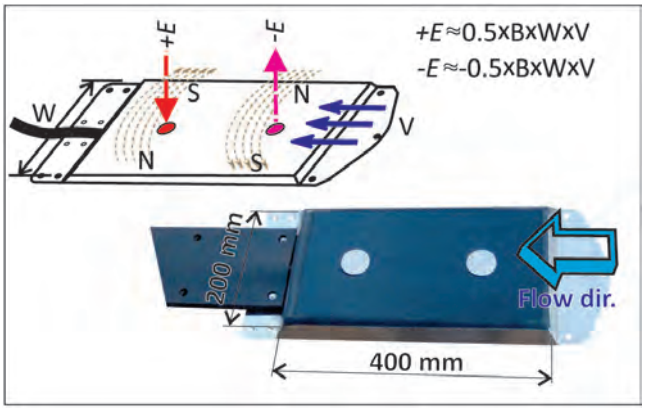


Figure 7. Modified version of EM probe suitable for flow measurement in sewer systems: Flat probe for wall mounting

temporal measurement system (Steinbock *et al.*, 2016) and combining with Computational Fluid Dynamics (CFD, Weissenbrunner *et al.*, 2016).

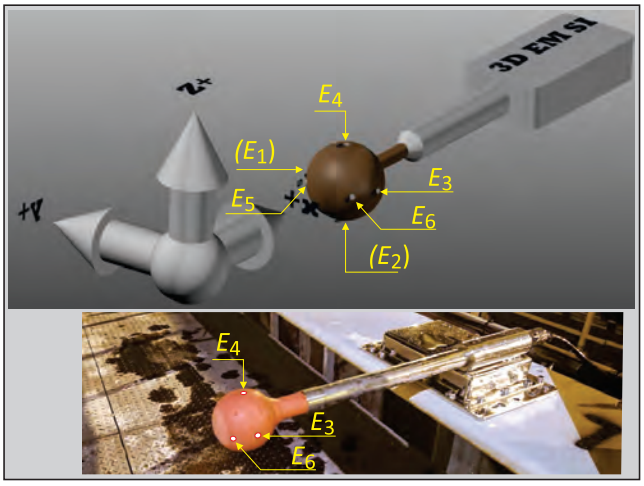


Figure 8. 3D EM probe with possibility to measure the streamwise (X) component with two separate pairs of electrodes.

The most advanced EM probe design has the possibility to measure (almost) 3D velocity field. With more than one magnet coil in the spherical head, and several electrodes, the probe can measure X, Y and Z component of the velocity (Fig. 8). Since probe's body will influence the velocity field when water is coming from behind the probe, the calibration chart is not ideal as for 2D probe (Fig. 4) and has to be established for each supporting frame used in certain application. The calibration of angular sensitivity is not as easy task as it might look, so the overall accuracy will depend on the effort involved.

To analyse the effect of flow separation on the spherical EM probe, the most important streamwise (X) component on the EM probe (Fig. 8) is measured using the electrodes E_1 and E_3 on 90° positions (electrodesE_2 and E_4 could be used also) and using electrodes E_5 and E_6 on 45° positions of the front side of the sphere (the probe is

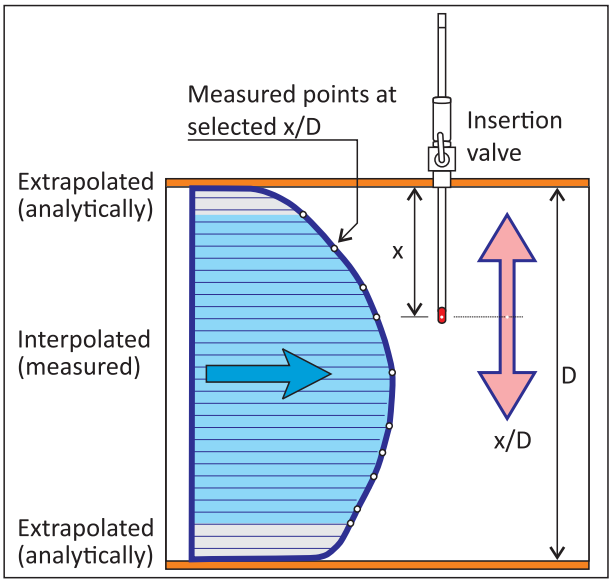


Figure 9. Basic setup for flow calculation, i.e., integration of measured velocities across the pipe's diameter.

called 3+D type). Comparing these two measurements, flow separation can be observed and taken into account.

2.2. Velocity profiling in unsteady flows

One of the most popular methods of flow measurement, with almost constant price/pipe diameter ratio, is point velocity measurement (or index velocity) and conversion of velocity into the flow rate. If EM probe is positioned at $0.121x/D$ from the pipe wall, assuming theoretically developed velocity profile, it will measure the true average velocity. In most cases, however, theoretical velocity profile is just a rough approximation of real conditions. Even in long straight pipes, the position of mean velocity is not constant, and it depends on the flow regime (Knudsen and Katz, 1958). The flow measurement error is difficult to estimate if the real velocity profile is unknown, and, according to the author's experience, the overall error is often an order of magnitude higher than expected (Prodanović, (Prodanović and Pavlović, 2003).

The velocity profiling (Fig. 9) is often used for accurate flow measurement in pipes where no flow meters are installed or for recalibration purposes of an existing flow meter that is too bulky to be calibrated on external calibration rig. To penetrate the pipe under the line pressure with the velocity sensor, an insertion hardware is needed (detail can be seen in Fig. 3). The probe is moved from one position to another, deeper into the flow field, recording the true velocities. As the first iteration, linear interpolation between the measurement points is suggested, and extrapolation using an analytical profile calculated for flow regime based on the measured velocities.

Although the process of velocity profiling should be straightforward, it is a rather time-consuming operation (in each measurement point the user has to wait for few minutes to perform good averaging), and the flow unsteadiness can introduce new errors. Fig. 10 presents a setup which can help in reducing the effect of unsteadiness. Since the mean flow is not constant during velocity profiling, correction to point velocities has to be applied before integration. An uncalibrated flow meter under the test can be used as an indicator of relative flow change.

In case where no fixed flow meter exists, second insertion-type meter can be used. The setup often used by the author is shown in Fig. 11 (Prodanović and Pavlović, 2003). It allows velocity profiling in two orthogonal directions and at the same time the correction for change in mean flow: during profiling along direction 1, probe 2 is fixed at one point and used for total flow monitoring,

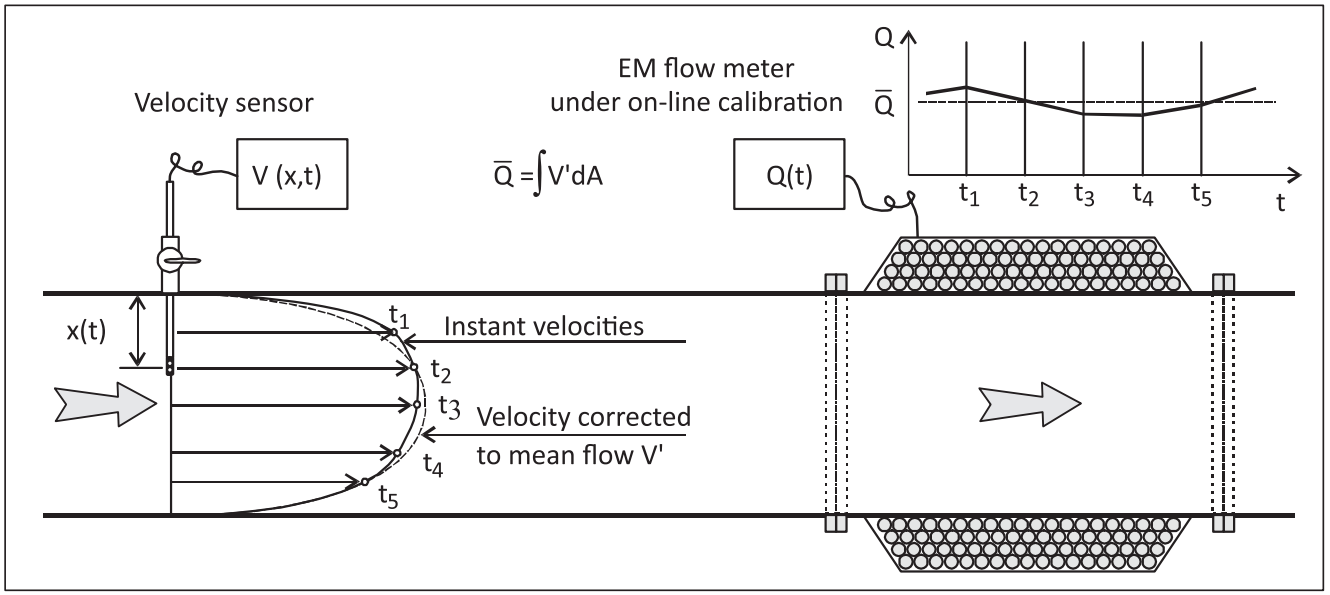


Figure 10. Recalibration of fixed flow meter by velocity profiling in unsteady conditions.

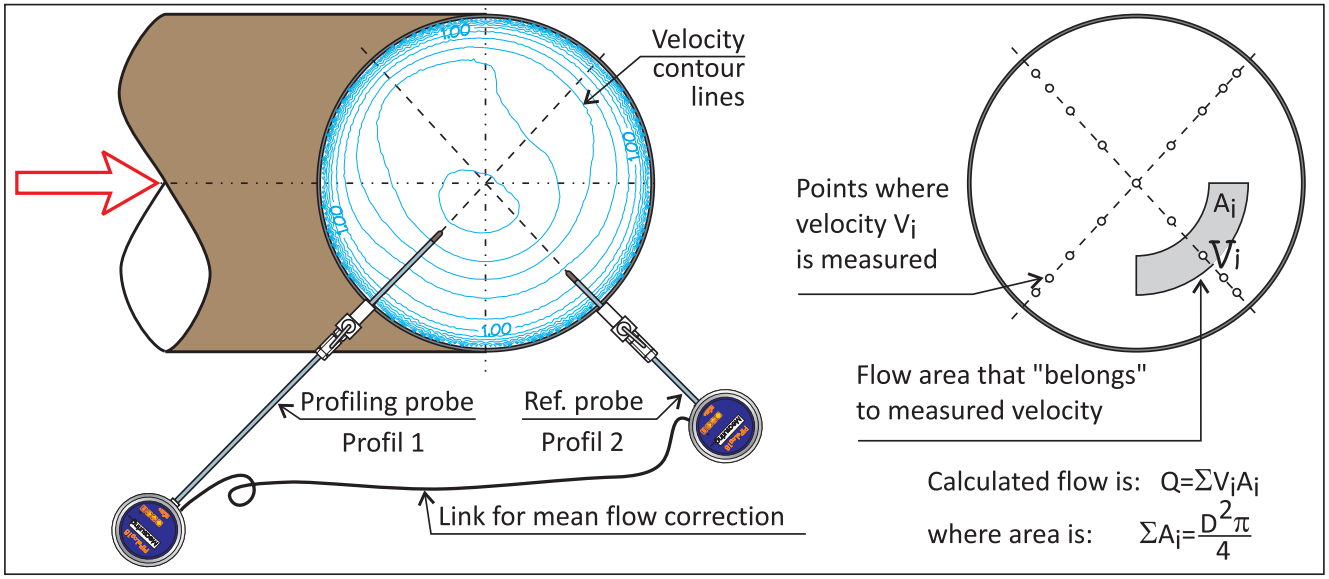


Figure 11. If there is no flow measurement, then second probe can be used as reference.

and vice versa for profiling in direction 2. Depending on the geometry used for velocity profiling, appropriate flow areas (Fig. 11, right side) have to be used for velocity integration.

3. CASE STUDIES

3.1. Case 1 – WTP Štrand, NS

The main source of drinking water in Novi Sad city (about 300.000 consumers) is central Water Treatment Plant (WTP) "Štrand" that treats groundwater from the Danube River banks. The capacity of WTP was 2×750 L/s in 2000 when the measurements were performed. Fig. 12 shows a schematic of WTP, which also includes the storage tanks of clean water under the filter stations that are not shown in the figure. Data for this case are taken from an earlier study (Prodanović and Ivetić, 2000), while today's situation is rather different.

At the inlets to the WTP, two old venturimeters were installed, which were out of function. Input to the new filter was measured using electromagnetic EM-900 flow meter, which was never calibrated, while input to the old filter station was not monitored. The bypass pipe leverages the operation of two filters, making the system highly dynamic.

Using the procedure described in section 2.2, with one EM probe of sufficient length for profiling and another shorter EM probe as the reference, the profiling was performed along two directions at EM-900 (position named $MP:N$ in Fig. 12). Signal from the reference EM was used to compensate for instabilities in flow since the signal from the existing EM-900 was hard to connect to the used logger due to grounding problems.

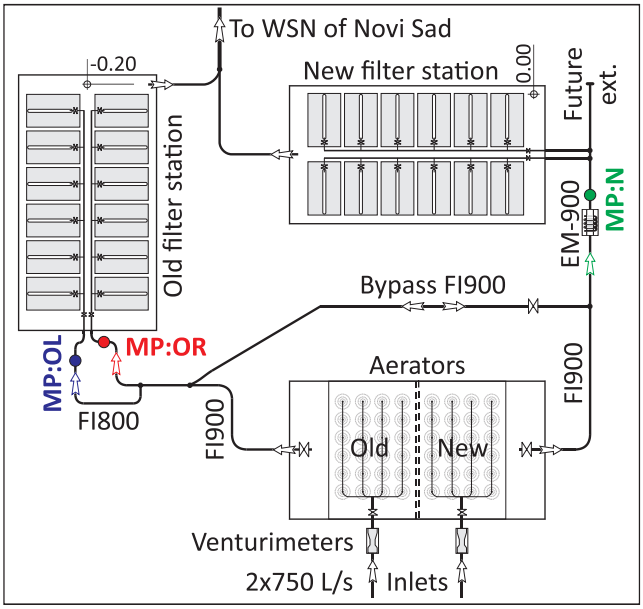


Figure 12. Schematic layout of the water treatment plant „Štrand" (as in 2000).)

Flow measurement on the old filter station was much harder to achieve. Figure 13 presents the axonometric presentation of the pipes: one common pipe with diameter FI900 (under the ground) is divided into the two FI800 pipes, which are raised above the ground and enter the building. There is no adequate place to measure the flow easily, neither in the common incoming pipe nor in two "smaller" pipes. Photo of the pipes entering the old filter station is given in Fig. 14.

Solution for the flow distribution measurement was to perform profiling along three directions (labelled 0°, 45° and 90° in Fig. 13), capturing the true velocity profile. Since there is no indicator of the overall flow rate,

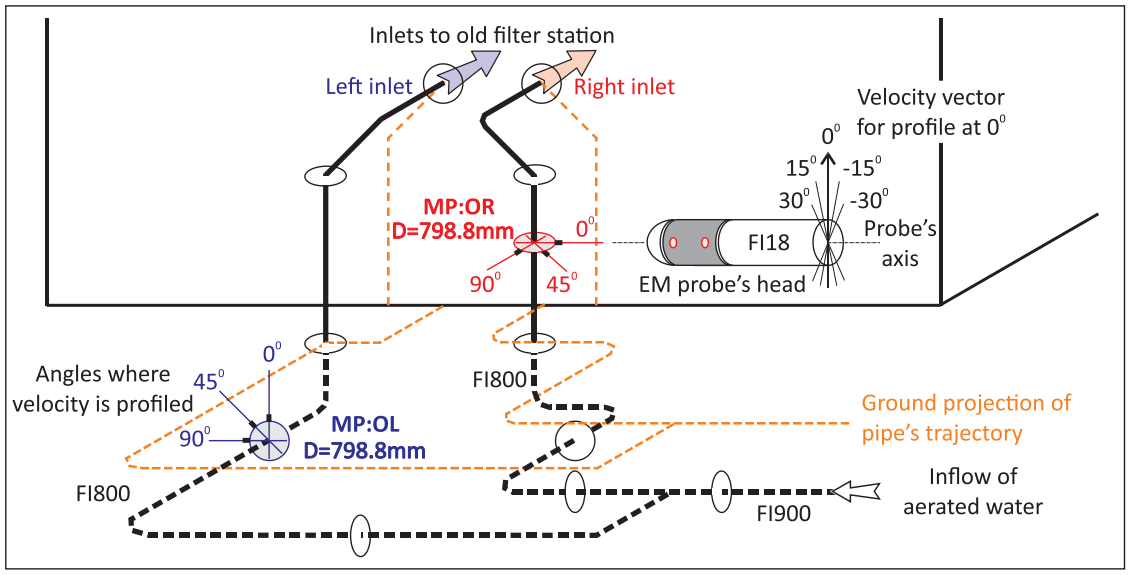


Figure 13. Two inlets to the old filter station (schematic drawing, pipe center lines shown with scaled cross sections)

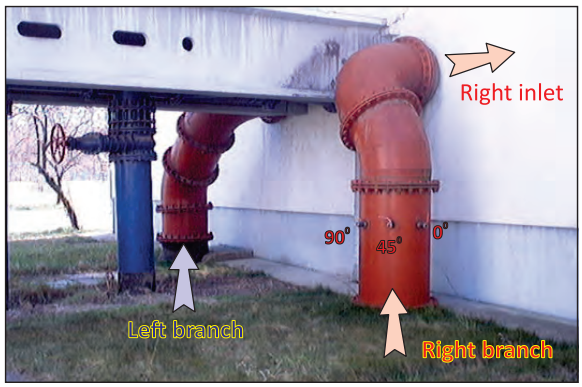


Figure 14. Photo of two pipes entering the old filter station

the methodology with two profiling probes was used (Prodanović and Ivetić, 2000). Measurements were performed at MP:OL position (connections were under the ground, some digging work was needed!) and MP:OR (connecting ball valves are visible in the Fig. 14).

3.2. Case 2 – Verification of built-in flow meters on large pipes in Topčiderska dolina

Flow measurement in large pipes (over 500–600 mm) is cheaper using Ultra Sound Transit Time (USTT) method then with bulky and heavy EM flow meters. The basic accuracy of USTT can be compared with EM flow meters (Danfoss, 2004) if flow conditions are satisfied. Misled by manufacturer's excellent specifications of the equipment, the common practice is to install the flow meter without any operational check of achieved accuracy. Main source of wrong expectation is that the USTT are named as "absolute sensors". However, they are absolute for average velocity measurement along one path between an electrode pair, but the connection between that measured velocity (or velocities for multiple path USTTs) and true flow rate depends on flow regime and local flow conditions.

The operation and accuracy of two newly installed USTTs flow meters in Belgrade's water supply system was checked during 2004 (Prodanović, 2004). Two large mains from the "Banovo Brdo" WTP, with FI1500 mm and average 2 m^3/s per pipe, are laid after a sharp turn through the "Topčider" valley (left part of Fig. 15), with a crossing over the "Topčiderka" river (Fig. 15, lower right corner). USTT's were installed at positions 11 and 12, relatively close to the sharp upstream elbow and downstream river crossing.

The only suitable place where flow rate could be checked was at positions 13 and 14, just downstream the river crossings, where a large manhole is located. Using the methodology of velocity profiling in unsteady conditions, where existing USTT flow meters were used as the reference, two types of current meters were used: the EM probes (Fig. 3) and SPECTRASCAN (Biwater Spectrascan, 1998) probes (micro turbine type). Due to physical limitations (small manhole), only one profiling line per pipe was used (Prodanović and Pavlović, 2004).

The first results indicated that the uncertainty of newly installed USTT flow meters is much higher than declared by the manufacturer. To confirm the findings, the CFD analysis of the flow conditions through the whole valley was performed and the results confirmed that the positions 11 and 12 are not optimal for USTT. Adequate profiles were available further downstream the valley (not shown of Fig. 15).

3.3. Case 3 – Flow measurement in large tunnels within the Trebišnjica system

The Trebišnjica River catchment in Eastern Herzegovina (Bosnia and Herzegovina) is one of the most complex

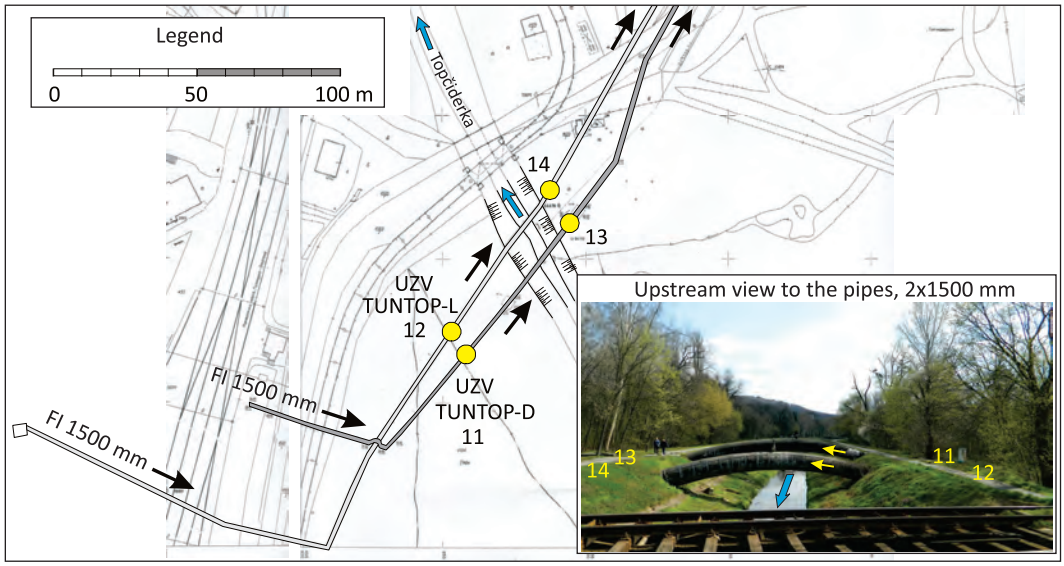


Figure 15. Two mains (FI1500 mm) along the "Topčider" valley with river crossing

karst areas in the Balkan region. It is characterized by complex karst landforms and drainage systems including karst poljes, ponors, springs, estavelles, developed underground connections, as well as underground bifurcation zones (Milanović, 2006). The area is a part of the Dinaric karst region, characterized by very irregular karstification caused by tectonics, compression, reverse faults and overthrusting structures (Jaćimović *et al.*, n.d.).

Figure 16. Catchment area of Dabarsko polje

In order to prevent flooding of karst poljes and utilize at the same time an exceptionally high hydropower potential, construction of a complex water resources system started in 1960s. The system has evolved in several stages, including construction of dams, tunnels and hydropower plants (HPPs), and is still evolving.

Important part of the system are two large tunnels that connect the Dabarsko polje and Fatničko polje (D-F tunnel), and Fatničko polje with Bileća reservoir (F-B tunnel). Two tunnels with their catchments (direct and indirect) are shown in Fig. 16.

To measure the flow rate in those tunnels is a non-standard task: the tunnels are large, flow is intermittent (during summer there is no flow, while during some periods tunnels are full of water but without flow rate) with high minimum-to-maximum ratio. The flow is in mixed conditions (mostly pressurized flow but can be also with free surface), and flow direction can be reversed. Additionally, water can carry large stones and gravel, together with other debris.

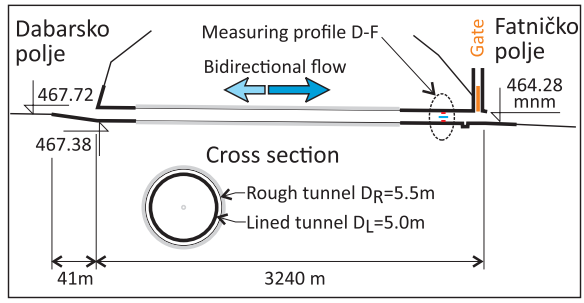


Figure 17. D-F tunnel with bidirectional flow

The flow through D-F tunnel mainly occurs in direction from Dabarsko to Fatničko polje with rates up to 60 m^3/s. The tunnel length is 3.24 km, diameter is 5 m in sections where it is lined and 5.5 m in other sections (Fig. 17). During some rainfall events, runoff response from the Fatničko polje catchment can be faster than that in Dabarsko polje and water can flow in opposite direction (with rates up to -40 m^3/s). The tunnel can be closed to prevent the reverse flow, which would be a loss for the hydropower system.

Measuring profile in the D-F tunnel is located 50 m before the exit to Fatničko polje. Details of the cross section

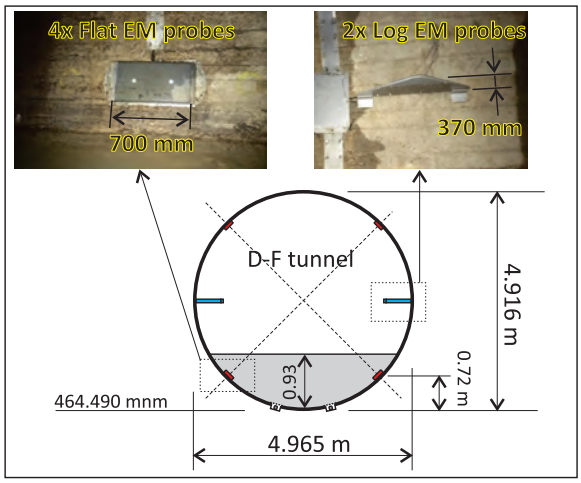


Figure 18. Measuring cross section in D-F tunnel

are presented in Fig. 18. Similar setup was used for other two measuring sites in the F-B tunnel.

Four large Flat EM probes were used per measuring cross section, to cover larger measuring volumes and improve the velocity integration if non-symmetrical velocity distribution occurs. Since flow in the tunnel can be with free surface (as shown on Fig. 18), two lower Flat EM probes were positioned below the expected minimum flow. Water level during free-surface flow is measured using two pressure sensors near the cross-section's bottom.

To assess the accuracy of "constructed" flow measurements, two approaches were tested:

- Usage of second, redundant type of measurement during certain period. The pair of LOG-EM probes is installed for one flow season (October-June) as a reference. These probes are less robust solution and are more susceptible to failure since they penetrate the flow profile, but they are closer to the mean velocity point for given flow conditions.
- CFD modelling of longer tunnel section, with detailed Flat EM representation. The model gives an insight into the expected velocity distribution for different flow conditions, helps in positioning of EM probes and estimating calibration parameters.

Second tunnel, from Fatničko polje to Bileća reservoir (Fig. 19) has 15.6 km length. The diameter is variable along the tunnel, depending on the slope and whether it is lined or not. At the inlet of the tunnel, there is a large dividing structure and inlet gate (Fig. 19, lower left part).

The measurement profile F-B_{IN} is positioned just upstream of the inlet gate and has 6.5 m in diameter. Expected flow rate is up to the 130–160 m^3/s (Water Institute „Jaroslav Černi", 2016).

Along the F-B tunnel, due to karst formations, water can infiltrate or exfiltrate, depending on the hydraulic conditions in the tunnel and surround terrain. Because of that, the outflow from the tunnel is not the same as the inflow and has different dynamics. The outflow measurement profile F-B_{OUT} is prepared upstream from the tunnel curve and contracting section (Fig. 19, lower right part, CFD model).

Fig. 19 shows one more flow measuring system just upstream of exit gate, labelled "Not operational flow measurement". In contracted section, the multi-trace US TT system was previously installed. Working conditions were such that velocities are too high with low pressures, so cavitation can easily develop. Mixture of air and water covered the US probes, prevented the transmission of ultrasonic pulses whenever flow rate exceeded 10-20 m^3/s. This location was therefore abandoned.

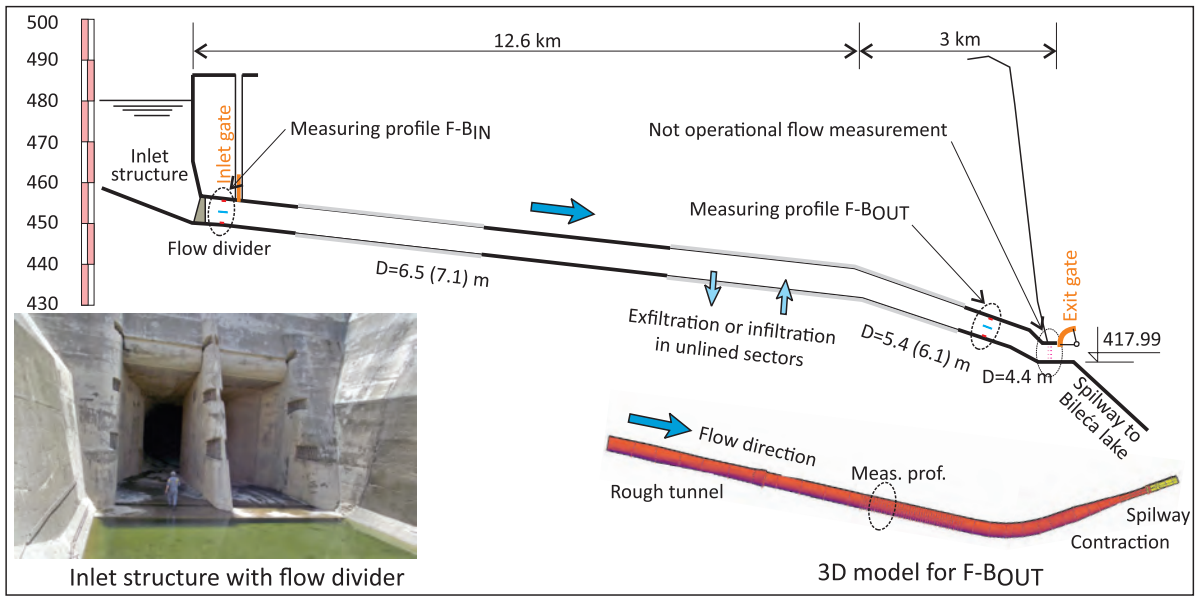


Figure 19. F-B tunnel with two flow measurement positions, at inlet and at outlet

Figure 20. F-B tunnel, inlet cross section F-B_{IN} before installation of equipment (left) and during the installation (right)

3.4. Case 4 – Flow measurement on intake of Djerdap 2 turbine

Hydraulic conditions around short structures in most cases prevents the existence of cross sections with parallel streamlines and fully developed turbulent flow profile. There is no "universal flow measuring device" that can be used in such conditions with adequate accuracy. This is the case with flow measurements at the intake of the aggregates of the Iron Gate 2 HPP (or HE "Djerdap 2", Fig. 21).

Figure 21. Iron Gate 2 (Djerdap 2) hydropower plant (taken from Google Earth, 2009; location: 44°18'24.61" N/22°33'53.54" E)

Both Serbian and Romanian HPPs at Iron Gate 2 are equipped with 8 turbines in the main plant and 2 turbines in the additional plant (total of 20 turbines). Total installed discharge is 8500 m^3/s (www .eps .rs). Figure 22 presents the longitudinal section through one bulb turbine (low head Kaplan type). The turbine has short intake part, close trash rack and fast service gate. The supporting wall below the turbine acts as a flow divider and straightener. The pressure taps are made in that wall, one at upstream part and two at left and right sides, creating the differential Winter-Kenedy (WK) flow measuring device. Standard procedure is to perform detailed measurements of relation between flow rate and measured differential pressure, for different working conditions, and determine the WK coefficients. Measurements could be done only on a scale model, in a controlled laboratory environment. The assumption is that the same WK coefficients will be applicable on the full-scale turbine.

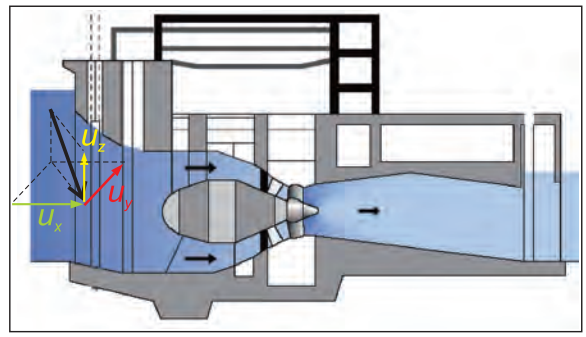


Figure 22. Bulb turbine with highly irregular conditions for proper flow measurement

Due to its position in the Danube (it can be seen from Fig. 21 that the Danube is not flowing orthogonal to Iron Gate 2 HPP) and leftovers of preconstruction works in riverbed, there is a significant incoming angle of water toward the turbines. The streamlines angle is larger at the Serbian side then at the Romanian side, and it depends on working conditions of spillways. Because of that, the incoming velocity vector is shown in Figure 22 with strong horizontal component parallel to the trash rack, U_Y (Prodanović *et al.*, 2011). The vertical component U_Z is typical for short turbine intakes.

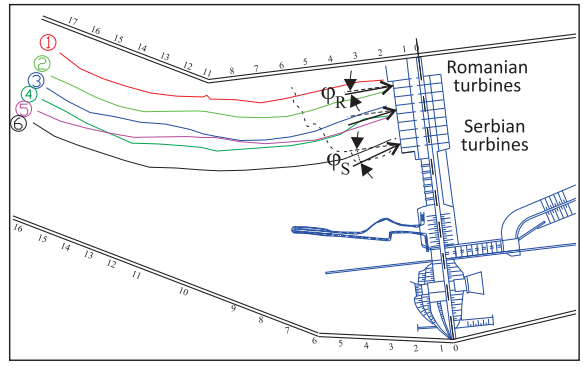


Figure 23. Scale model with streamlines (Water Institute „Jaroslav Černi", 2006)

To assess the influence of flow regime on operation of each turbine, the scale model was made (Fig. 23, Water Institute „Jaroslav Černi" 2006). Model clearly demonstrates that the Romanian turbines are better positioned than Serbian turbines, and that the curved in-

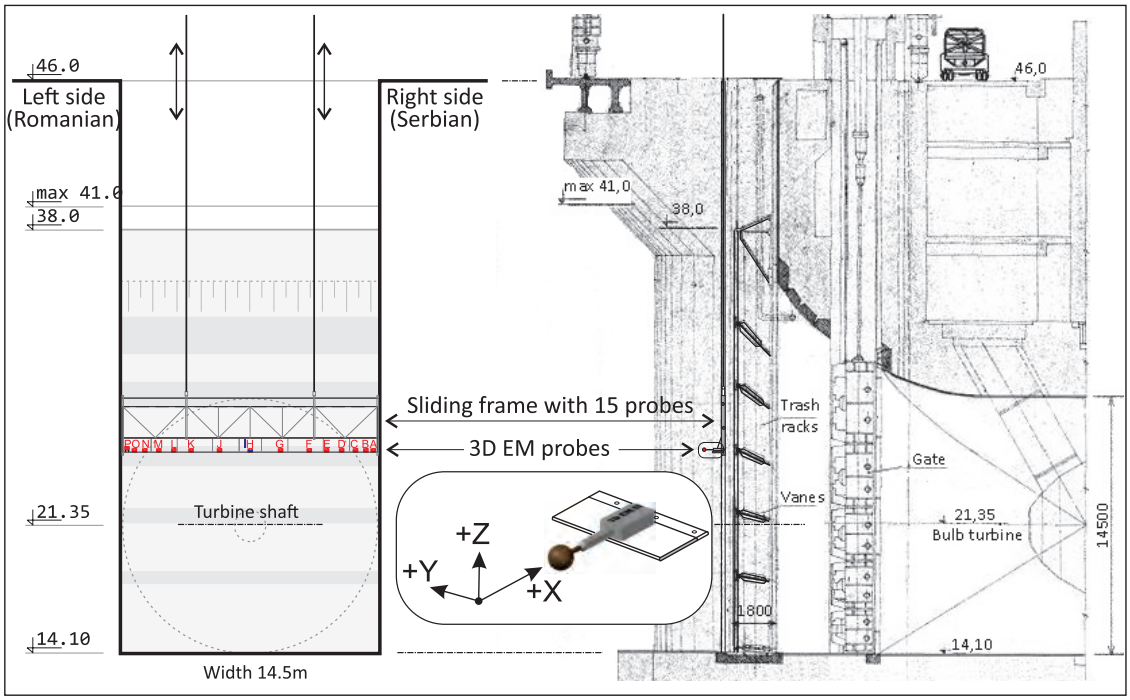


Figure 24. Suggested flow measurement method by profiling velocity field using one row of 3D EM probes == Streamwise view at the turbine's inlet (left side) and longitudinal section through the bulb turbine (right side)

coming streamlines are the reason for turbine's lower efficiency coefficients. However, the influence of curved streamlines on WK coefficients could not be assessed from the scale model. Flow rate measurements in situ and recalibration of WK coefficients was the only solution.

Since the HPP is equipped with Kaplan's short turbines and there is no regular cross section where flow can be measured using standard methods, available measuring methods for velocity field measurement were analysed by Prodanović *et al.* (2011). Figure 24. presents the suggested solution, which is the velocity field profiling using probes capable to measure all three velocity components at the same time. The equipment can be positioned in front of the trash racks. To reduce the blockage effect and to avoid increased head drop, a narrow sliding frame is suggested, with one row of probes installed.

3D velocity measurement was performed using newly developed EM probes (Fig. 8). Fifteen probes were installed on the frame, positioned to follow the expected velocity profile. The 3+D version of the EM probes were used, with which the streamwise velocity component is measured with increased accuracy. The frame is raised for approx. 27 m. The approach suggested in "Velocity profiling in unsteady flows" (section 2.2) was used, where the referent flow was taken from WK system and checked by turbine operational data (power, head loss, etc.). Since these complex measurements are conducted with numerous sensors working in real time, specialized software for data acquisition and for data analysis was developed (Water Institute (Water Institute „Jaroslav Černi", 2020).

4. RESULTS AND DISCUSSION

4.1. Flow measurement in complex unsteady situation (Case 1)

The recalibration of existing flow meter (Fig. 12, MP : N), equipped with a new electronic transmitter, was performed using two 1D EM probes (Fig. 25). Because of irregular position, two profiles (under angle of ±27° to vertical direction) were used for velocity distribution measurements. The genuine flow signal from the built-in flow meter was available only for manual reading since it was electrically complicated to connect to the used data loggers (grounding problems). Therefore, one longer EM probe was used for profiling, while the other EM probe ("reference") was used as flow indicator, measuring velocity at one point. This velocity is used in the normalization process. Output from the new transmitter was, as mentioned, manually recorded.

When switching from one profile (+27°) to another (−27°), both probes were fixed for certain period at the same position (same x/D), to create the correlation between the two points and use it in the velocity normalization process. Since EM probe gives instantaneous velocity at a point, which always has variations due to turbulence, it is a common practice to average the signal over certain period of time. The averaging time usually corresponds to the criteria that the variance of measured velocity is $\sigma_V \leq 1.0\%$. This criterion can be used, however, only for low turbulence intensity flow, where small vortices are predominant.

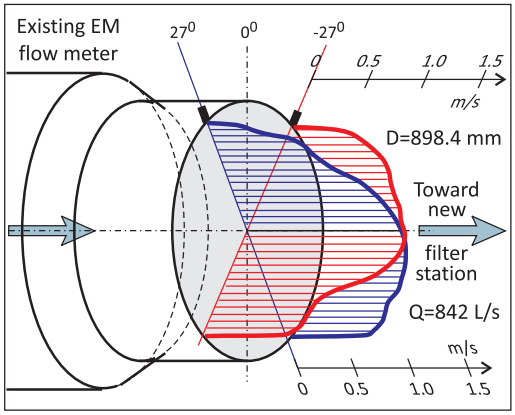


Figure 25. Recalibration of the built-in flow meter in dynamic conditions

Having large pipe's bend upstream of the EM under the recalibration (Fig. 12), slow helicoidal disturbance is recorded, with periods of about 2 minutes or longer. In Fig. 25, the upper diagram shows the 10-second averaged velocity in the reference position. Standard criteria of minimizing the standard deviation σ_V could not be used, so in order to provide the same conditions for all measuring points during profiling procedure, the reference signal was used to trigger the 1-minute averaging period on profiling probe (marked areas on lower diagram on Fig. 26). Obtained mean point velocities were then corrected for mean flow changes during profiling measurements (red and blue velocity profiles on Fig. 25), and then used for velocity integration.

Flow measurement at the inlets to the old filter station (Fig. 12) was even a bigger challenge. The flow conditions were far from standard, with several 90° bends in horizontal and vertical plane. At each pipe, three profiles for velocity distribution measurement per cross section were used (Fig. 27).

Daily variation of flow at two inlets is presented in upper part of Fig. 27. For each cross section ($MP : OL$ and $MP : OR$), profiling along three directions lasted for approx. 3 hours. Recorded velocities, normalized, are presented in Fig. 27 and show expected irregularity. This indicates that including one more profile (along 135° angle) would reduce uncertainty of the obtained results.

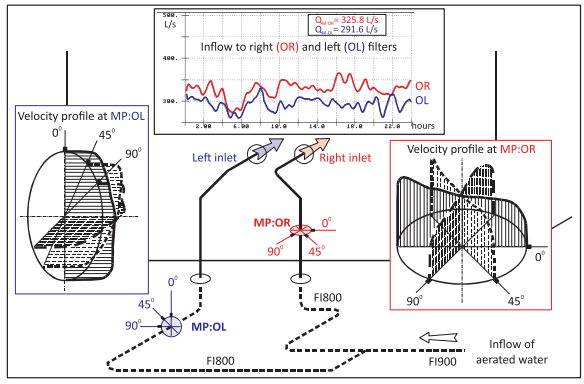


Figure 27. Velocity profiles at measuring points for two inlets to old filter station

Measurements at WTP "Štrand" were performed during the year 2000. At that time, the author did not have an operational 2D EM LOG probe, as shown in Fig. 5. Using the 1D EM LOG probe, oriented to measure orthogonal velocity component (at 0°, Fig. 28), the flow can be directly calculated. However, to "see" what the true velocity direction is, it is possible to rotate the probe around its axis. Fig. 28 presents the obtained result for one point in the 0° measured profile. The point was close to the wall and the EM probe was rotated in the range of angles of ±30°. The obtained maximum velocity was around the angle of +10°.

The problem associated to the approach of probe's rotation is the slow measurement process, which is undesirable due to global variations of inflow. If full 2D EM LOG probe had been available at that time, both components would have been recorded simultaneously, thus increasing the users' understanding of the full flow velocity pro-

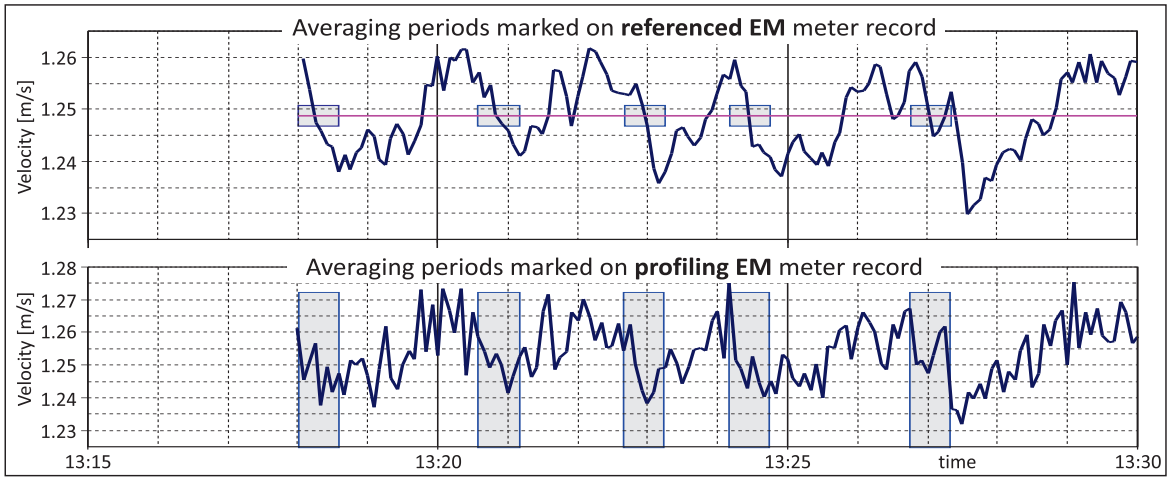


Figure 26. Velocities from reference probe and from profiling probe

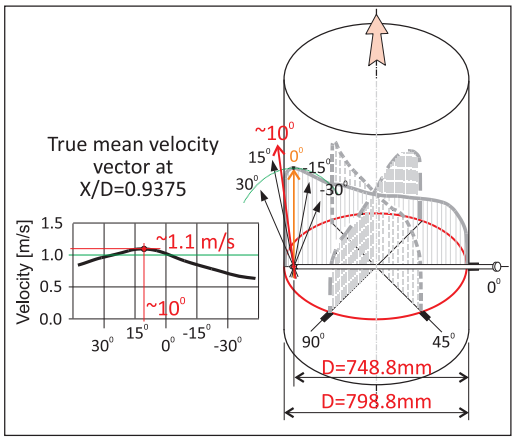


Figure 28. True angle of velocity vector near the pipe's wall

file. However, only the streamwise 1D component would then be used for flow computation.

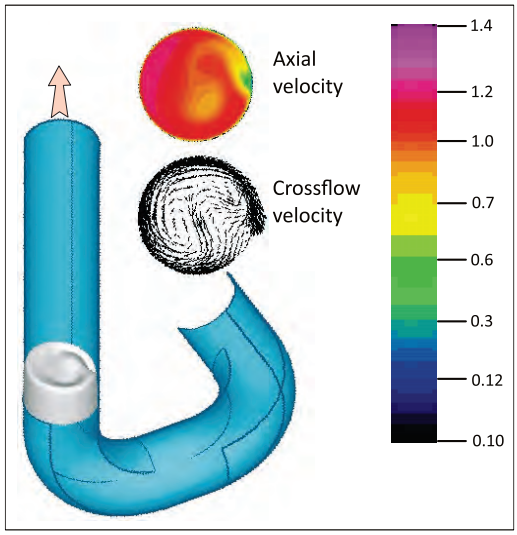


Figure 29. 3D flow simulation through doubled bend (Sontex Presentation CD, 2002)

To improve the understanding of expected velocity profile, the results of simulation of flow though the double bend done by Sontex Presentation CD (2002) is shown in Fig. 29. The bends are not in one plane; diameter is $D = 200$ mm and radius of curvature is $1.5D$, Reynolds number is $Re = 65,000$, turbulence model applied is $k\&e$ and the inlet profile (boundary condition) has ideal turbulent profile. The geometry is almost identical to situation at the WTP "Štrand" site, only with a different Reynolds number: in the presented simulation it was $Re = 65,000$ and on site it was $Re = 480,000$. The simulation result predicts helicoidal stream near the measurement profile, which was confirmed during the measurement.

4.2. Verification of large flow meters with the support of CFD (Case 2)

Two water mains, important for supplying the central part of the Belgrade city with drinking water, are laid cross the Topčider valley. Each pipe is 1500 mm of diameter, and they carry about half of total Belgrade's consumption. As described in section 3.2, the flow rate measurement in each pipe is done using USTT with two horizontal paths. Nominal accuracy of installed flowmeters is 0.5% (Danfoss, 2004). Flowmeters were installed upstream of pipe's crossing of the Topčiderka River. They were assembled in situ and only the geometry of US paths was checked. The overall flow measurement accuracy was not verified.

To verify the accuracy of USTT flowmeters, velocity profiling downstream the "Topčiderka" river was done. The existing manhole was used to drill a tap into the pipe and put the 1D EM LOG probe inside. Due to space limitations, only one direction for profiling was used. Also, since the pipe is large, it was impossible to move the EM probe to the end of the profile, so only 2/3's of the cross section was profiled.

To compensate the flow variations, the signal from existing USTT's logger was used. Measurement was repeated using the same tap connection, with the SPECTRASCAN (Biwater Spectrascan, 1998) probe, which uses the micro turbine for velocity measurement. Basic accuracy of the 1D EM LOG probe is 1% for velocity measurement, and accuracy of SPECTRASCAN is 2%.

Figure 30 presents the obtained measurements using two types of probes: green squares for EM probes and red circles for SPECTRASCAN corrected for mean flow variation (Prodanović and Pavlović, 2004). Since the measurements did not cover the whole profile, the three expected and possible profiles were extrapolated, and consequently, three possible flow rates were computed: Q_{13} in the range of 1.59–1.72 m^3/s and Q_{14} in the range 1.77–1.92 m^3/s.. At same time, the average $Q_{11}(= Q_{13}) = 1.59$ m^3/s and $Q_{12}(= Q_{14}) = 2.17$ m^3/s. Although the uncertainty of the computed control flow rates is high, it is obvious that USTT$_{11}$ is close to its nominal characteristics and that USTT$_{12}$ is overestimating the flowrate: even if the highest extrapolated value is used as the "accurate" flowrate, the error of USTT$_{12}$ is +10%.

Dashed line in Fig. 30 presents the theoretical velocity profile for flow rates as "seen" by USTT. The left diagram supports the conclusion that USTT$_{11}$ is "probably" within the nominal accuracy. However, the right diagram shows that USTT$_{12}$ is clearly out of the specified range. The word "probably" is used here because the overall recalibration method could be argued. Calibration was done with the instrument of lower accuracy (velocity measurement of 1–2%), and the integration methodology has even higher uncertainty (from diagrams it can be estimated to 3–4%) which gives combined uncertainty of 5%. Therefore, an instrument with a ten times lower accuracy was used to check the existing flow meter! However, having the difference between the obtained results even higher than the assessed 5% range, as in the case of USTT$_{12}$, leads to the

only conclusion that the tested flow meter is not within the specified characteristics.

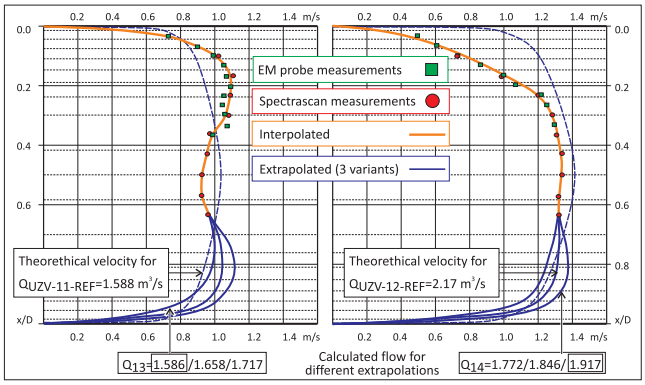


Figure 30. Measured velocity profiles downstream the river crossing, position 13 (left) and position 14 (right figure)

To have a third check of these conclusions with "not-so-perfect" insertion-type velocity meters, the CFD simulation of flow was performed. The idea was to inspect the assumed extrapolations in Fig. 30 and to look at the flow profile at the places where USTT were installed, trying to understand what the reason for the obtained results was.

The CFD model was created for the whole reach of the main where MM12 and MM14 are positioned (total length 450 m, pipe's diameter 1500 mm). Finite volume method was used, and flow area was divided in 610,836 volume elements, axisymmetric, with smaller volumes near the walls and around the bends (*Results of flow simulation in Topčider mains*, n.d.). Small detail of the model is given in Fig. 31, showing also computed streamwise velocities near the crossing over the Topčiderka River. Simulations were made by the FLUENT software (v. 6.1), with direct solution of turbulence based on the Reynolds-Stress-Method (RSM), applying stationary flow rate $Q = 1.917$ m^3/s. Several wall roughness values were tested ($\delta = 1.2$ mm, $\delta = 2.5$ mm and $\delta = 3.5$ mm), assuming uncompressible fluid and including the gravity force in the momentum conservation equation. The boundary condition at the downstream end was «OUTFLOW», where pressure and velocities are the same as previously calculated.

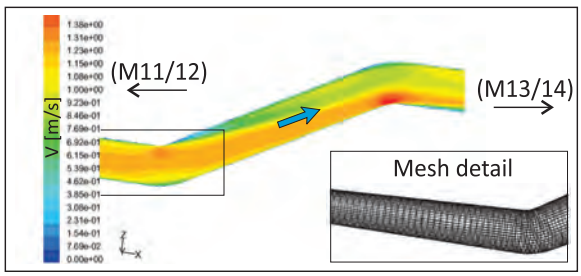


Figure 31. Detail of the CFD model, bend before the crossing over the Topčiderka River

Computed velocity profile at the position of USTT$_{12}$ (MM12) is presented at Fig. 32. The profile is not symmetrical nor fully developed. When the velocity profiles along the two US paths are linearly averaged (as USTT is doing), the computed mean velocity is greater than it should be and the installed flow meter shows greater flow rate. Since the whole reach where the USTTs are installed is under strong influence of the upstream bend, it was just a matter of luck which profile would provide positive/negative bias or accurate operation of installed meters.

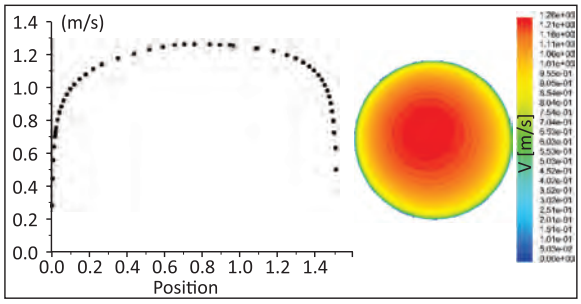


Figure 32. Velocity distribution at MM12 profile differs from an "ideal" symmetrical one

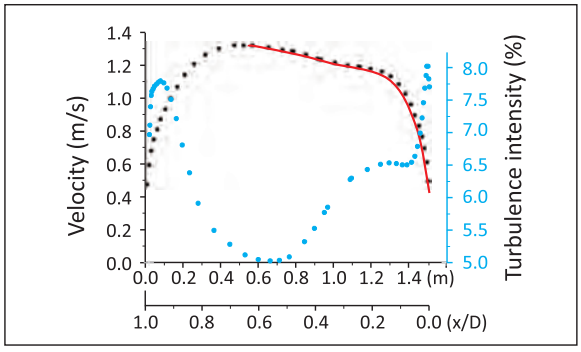


Figure 33. Results of CFD simulation for MM14 along the profiling direction: velocities (solid line represents measurements) and turbulence intensity

To check the assessed uncertainty of 3–4% for possible velocity extrapolation along the measured profile at MM14 (Fig. 30, right part, three possible solutions for extrapolation), the CFD result for the same cross section and the same direction of profiling is plotted in Fig. 33. Solid line shows velocities obtained by the measurement, and dotted line is the result of the CFD simulation. The shape of the line suggests that the applied extrapolation method with the highest flow rate is acceptable. The assessed uncertainty due to extrapolation is probably even lower then suggested!

Using the CFD simulation, it was possible to search for the best position where USTT should be installed. Close to the end of the Topčiderka River valley, the mains have a straight reach. The simulation shows that the velocity distribution is almost as theoretical one, with stable and symmetrical turbulence intensity. If USTTs were installed at that position, they would have (close to) the rated measurement error.

4.3. Site specific calibration in large tunnels (Case 3)

Calibration of a flow measuring device (permanently installed working meter) means that user can occasionally conduct parallel measurements by some higher accuracy equipment and use these measurements to calculate the calibration coefficients of the operating flow meter. In some situations, such a "site specific" calibration is hard to organize due to different reasons. e.g., due to large tunnel diameter and intermittent water flow as in the presented case of flow measurement in tunnels of the Trebišnjica system (section 3.3). In such cases, some other calibration methods are needed (Ivetić *et al.*, 2017b).

In case of the Trebišnjica tunnels, the permanent flow meter was constructed using flat EM probes attached directly to the walls (flush mounted, Fig. 18, left side). Such solution is robust and can resist the expected impact of debris and stones that can be dragged by the water. To improve the robustness of the solution, four probes were placed within one cross section. The main drawback of the solution with flush mounted EM probes is that velocity distribution near the wall is not equal to the mean velocity and depends on the flow conditions (Reynolds number).

According to the Pope (2008), velocity distribution in highly turbulent symmetrical pressurized flow can be calculated using exponential low:

$$V_x(z) = V\left(\frac{2z}{D}\right)^{1/n}\frac{(n+1)(2n+1)}{2n^2} \tag{4}$$

where $V_x(z)$ is streamwise velocity component, z is distance from the wall, D is pipe's diameter, and n is the exponent (coefficient) of power law. Pope (2008) suggested the following equation to calculate coefficient n as a function of Reynolds number:

$$n = 1 + \sqrt{\frac{Re}{50}} \tag{5}$$

The coefficient n can be calculated also using Nikuradze's equation (1932) with logarithmic relationship:

$$n = 0.5261(\log Re)^2 - 3.853\log Re + 13.1537 \tag{6}$$

For the place where flat EM probes are installed, the calibration coefficient is

$$C = \frac{V_{\text{mean}}}{V_{\text{meas}}} \tag{7}$$

where V_{meas} is the velocity measured by flat EM and V_{mean} is the mean velocity used to calculate the flow rate: $Q = V_{\text{mean}} \times A$ (A is the area of cross section). Looking to the red lines in Fig. 34, which represent the variation of the calibration coefficient with flow rate for flat EM probes, it is obvious that uncertainty is high, since two widely accepted

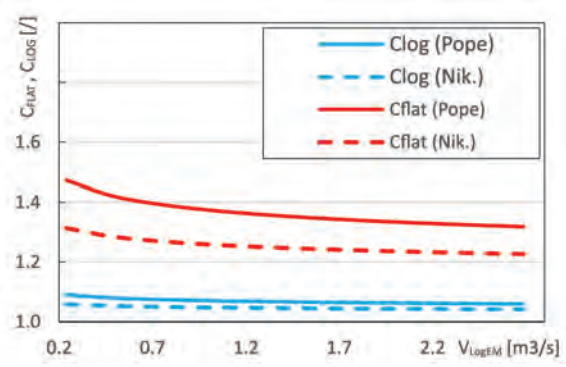


Figure 34. Comparison of Pope's and Nikuradze's equations for velocity distribution, for locations of flat EM and log EM probes

theoretical equations differ more than 10% and that dependence of coefficient with the flow rate is high.

The approach used in such situation was to temporarily install another set of EM probes having lower uncertainty (LOG based EM probes that penetrate into the flow field, Fig. 18, right side), but are less robust. The calibration coefficient for such probes is stable (Fig. 34, blue lines) regardless of the theoretical velocity distribution and flow regime. Also, the velocity field within the integration volume of EM probes (Fig. 35) is more homogenous for LOG type probes (as is in towing tanks used for probe's velocity calibration) than for flat probes at the tunnel's wall.

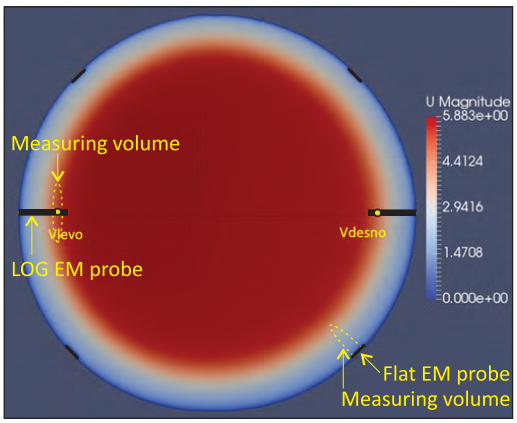


Figure 35. Comparison of measuring volumes for flat and LOG EM probes, for F-B tunnel

The setup with permanently installed four flat EM probes and temporarily added two LOG EM probes was used for all three measurement profiles: in D-F tunnel and two profiles in F-B tunnel. The calibration was performed using data collected for one season (December – Jun) for all flow regimes. After one year period, the temporarily installed LOG probes were removed from one profile and installed on another.

Some results for D-F tunnel during one season of calibration are presented in Fig. 36 (Ivetić, Prodanović, Cvitkovac *et al.*, 2018). The relative theoretical calibration coefficient for flat EM is plotted (corrected using LOG

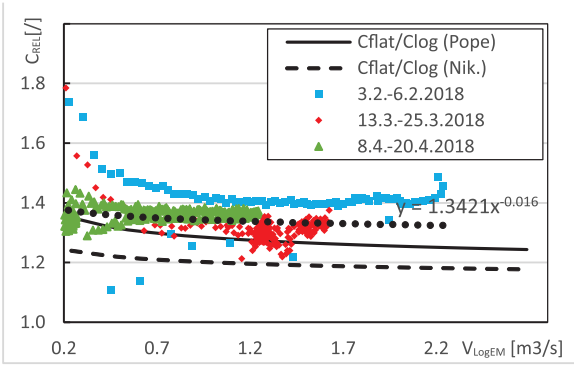


Figure 36. One season of flat EM calibration on DP-FP profile

EM probe data) for three events, including one with the reverse flow direction, from Fatničko polje to Dabarsko polje. Dashed and solid lines are for distributions given by equations (6) and (5) respectively. Final accepted velocity distribution based on recorded events is given by dotted line (Ivetić *et al.*, 2017a).

In situations where no secondary measuring systems could be installed (as LOG EM probes here), the authors tried to simulate possible output from the flat EM probe using equation (3). A detailed CFD model for all measurement profiles was created. Number of elements was between seven and eight million (10 times higher than in Case 2), since it was important to capture both long effects of flow through the tunnel (length about 100 m, 4 times shorter than in Case 2) and fine details around the EM probes (scale of few millimetres). The OpenFOAM finite volume model was used. Upstream boundary conditions were represented with given theoretical velocity distribution and the downstream conditions with constant pressure. Two turbulence models were tested, the $k-\varepsilon$ model and $-\omega$ (Shear Stress Transport model – SST). For the free-surface flow modelling, the Rigid-Lid approach was used (Ivetić, 2019).

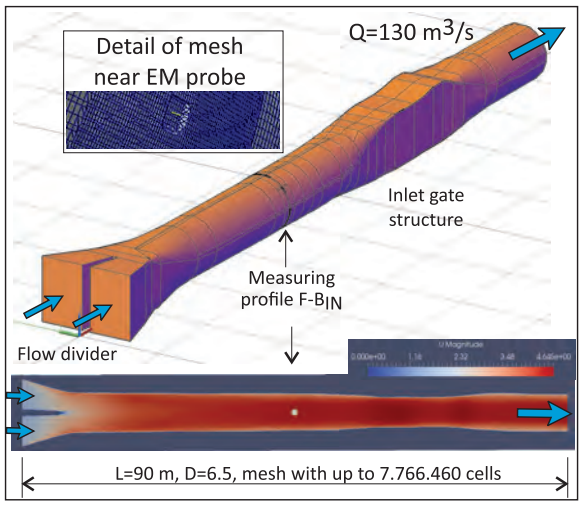


Figure 37. Detail of CFD model for F-B_{IN} measurement profile, between inlet flow divider and gate structure

Model of F-B tunnel inlet is presented in Fig. 37 (Water Institute „Jaroslav Černi", 2016). The influence of input flow divider on the velocity field can be seen. Slight flow instability was observed, which was averaged using four EM probes in cross section. Using results from the CFD model, the velocity distribution within the measuring probe's volume is plotted in Fig. 38, for lower EM1 and higher EM3 positioned in the left part of the cross section. Dashed line in the same figure shows the Bevir's weight vector (3), representing the strength of magnetic field versus distance from flat EM probe's surface (this curve can be obtained from EM probe manufacturer, or by direct measurement of magnetic field, Ivetić (2019); Stojadinović *et al.* (2018).

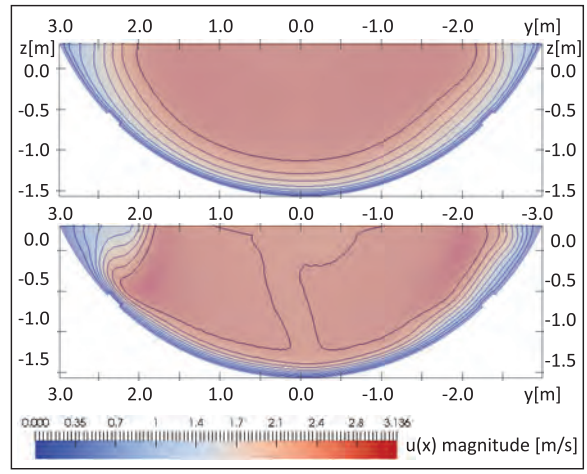


Figure 38. Relation between the strength of Bevir's weight vector (Eq. 2.3, dashed line) and velocity versus distance from flat EM1 and EM3 probes (solid lines)

To achieve the "site specific" calibration of flow meter, integration of two functions (dashed and solid line, Fig. 38) has to be performed. However, this solution assumes that velocity distribution computed using CFD has low uncertainty, with all local effects considered. To test this "hypothesis", numerous simulations were performed for all three measurement profiles in the Trebišnjica system. Dispersion of CFD results were very high because there were no referent measurements that could be used to "calibrate" the numerical solution. Therefore, the lack of higher quality data that can be used for calibration of flat EM probes (as mentioned at the beginning of this section), which we tried to bypass by employing CFD, is just translated to the problem of missing good data to calibrate the numerical model!

To illustrate high uncertainty of the uncalibrated CFD model, two simulations of free-surface flow regime are presented in Fig. 39. The two velocity distributions near the flat EM probes are computed for the same depth and flow rate but using different turbulence models. Which one is "better"? It is hard to say without detailed measurements. In similar situations, the scale model in labo-

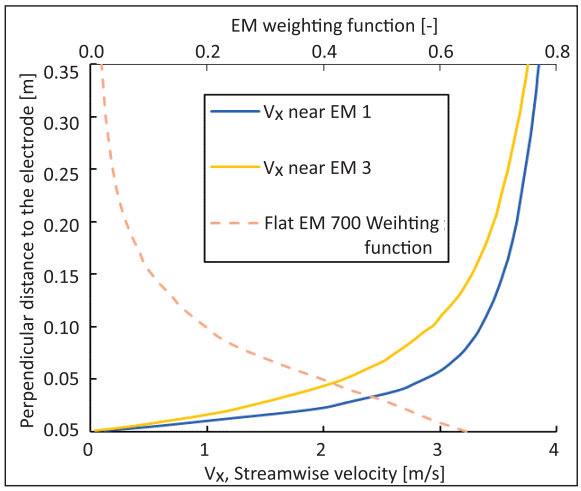


Figure 39. Example of two numerical solutions using different model parameters for the same flow conditions

ratory would be created, detailed measurements would be used to calibrate the CFD model that can then be used to predict the velocity distribution in the tunnel with much higher confidence.

Once when the flow measuring equipment in selected profiles were installed within the F-B tunnel (Fig. 19), it was possible to perform the water balance analysis. The tunnel is built in leaky karst formations, so it was expected that inflows and outflows would not be in balance, and that the difference would depend on water pressure within the tunnel.

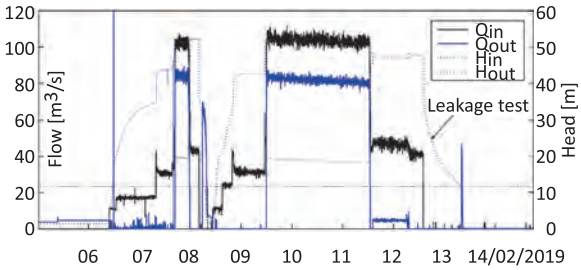


Figure 40. Water balance of F-B tunnel

During several days of February 2019, water balance test in F-B tunnel was performed. Fig. 40 shows inflow and outflow hydrographs using solid black and blue lines respectively, while pressure head is presented using dotted lines. Downstream pressure was controlled by exit gate. It can be shown that water loss ΔQ is exponentially related to the pressure head measured upstream of exit gate. During 12th of February, a test was performed with the downstream gate closed (small leakage under the exit gate can be seen). During that period, the inflow into the tunnel was about 40 m^3/s, representing the flow that fills karst caverns, thus representing the loss to the main flow.

4.4. Flow measurement in 3D unsteady velocity field of turbine intake (Case 4)

To perform the measurement of flow rate at the intake of turbines, where the conditions for standard techniques as described in IEC 60041 (1999) do not exist, the approach to map the whole velocity field is accepted. The turbine at "Djerdap 2" Hydro Power Plant (HPP) is a low head Kaplan turbine, which is sensitive to additional head losses that can occur if the whole cross section would be covered with current meters. Therefore, the horizontal profiling technique using a row of current meters with dynamic flow corrections is applied.

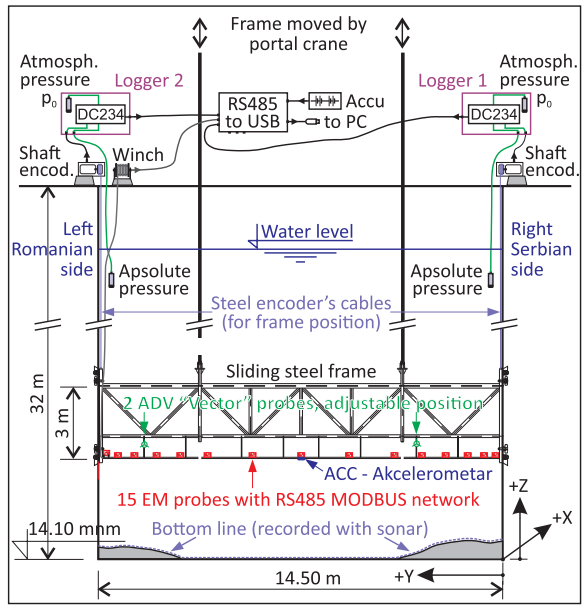


Figure 41. Measurement rig used for velocity field mapping

The method applied for velocity profiling is presented in Fig. 24. Details of the system are given in Fig. 41 (Ivetić *et al.*, 2021). The steel frame (14.5 × 3 m) is designed to minimize the flow disturbances and vibrations (accelerometer was added to central EM probe). The frame is hanged on portal crane (Fig. 42) and is positioned just upstream of the trash rack. The frame can be traversed vertically through the whole flow cross section, even during the highest flow rates.

Figure 42. Steel frame used during the measurement (in the most upper position)

Total of 15 3+D (streamwise velocity component is measured at two positions) EM velocity meters (Fig. 8, accuracy better than 1% for streamwise components in the range of 0–3 m/s) were mounted on the lower part of the steel frame in a horizontally symmetrical, but unevenly distanced pattern (Fig. 43, lower right part). Redundant and control velocity measurements were performed using two Nortek "Vector" (NORTEK, 2020) acoustic doppler velocity meters (ADV, accuracy 0.5% for all velocity components), mounted at height of 0.5 m above the EM meters (Fig. 43, upper left part). Continuous water level measurements were carried out via two fixed pressure transducers (accuracy better than 0.1%) installed at both sides of the intake. The position of the steel frame was precisely monitored with two UniMeasure "HX-EP" position transducers – shaft encoders (UniMeasure 2020, accuracy better than 0.025%) installed on the platforms (right and left) above the intake. Prior to the discharge measurements and the mounting of the equipment on the steel frame, four sonars were mounted on the positions of the EM meters (total of 15 positions can be used) for echo sounding the bottom-line profile.

The 3+D EM probes were calibrated in tawing tank prior to measurements, using the same geometry of supporting frame as one of the sliding steel frame. The influence of 3+D EM probes on ADV sensor is also checked in the towing tank.

All 3+D EM probes, two water level meters and two frame position transducers were connected in RS485 network with MODBUS protocol. All sensors have their internal loggers and accu-batteries for backup. The external large accu-battery was used to power the system, capable of two days mains free operation. The ADV meters were used as control redundant meters, and they were not connected to the RS485 network. They used the internal battery and data logger, with real-time clock to be synchronized to the rest of the system.

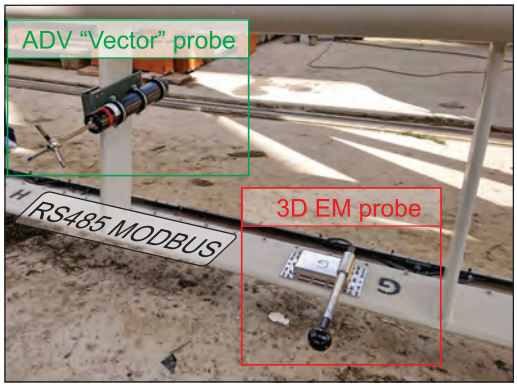


Figure 43. Steel frame with installed EM and ADV probes

Once powered on, the measurement system is continuously collecting data in real-time (except for the ADVs). Using specially developed Real-Time Monitoring software, data are processed and visualized. The sampling (for all the sensors) and visualization frequency is adjustable, with the highest frequency of 1 Hz.

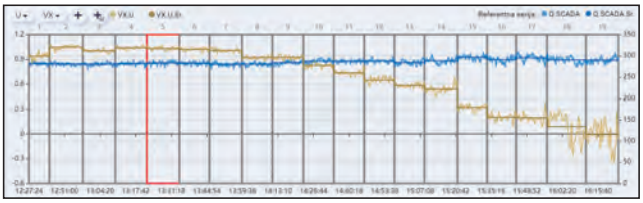


Figure 44. Turbine power during the velocity field traversing, used for unsteady turbine compensation (blue line) and streamwise component along the "U" vertical section (orange line) shown on Fig. 45

The measurement system can be operated in two modes, continuous and incremental. In the continuous mode, the steel frame is continuously traversed from bottom to the top with the lowest possible crane's speed of about 5 cm/s. This mode allows for the full velocity profiling to be taken in 6 to 7 minutes. In the incremental mode, the frame is traversed upwards between the semi-equidistant profiles (usually 17 to 20 profiles). In each of the profiles, the frame is kept still for about 10 minutes. In this manner, one discharge measurement can last for several hours. In both modes it is recommended to keep as stable as possible flow conditions at the examined turbine and two neighbouring ones.

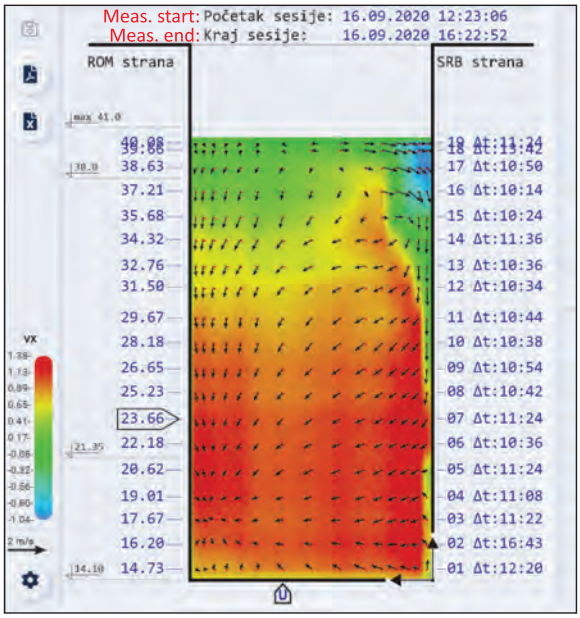


Figure 45. Velocity field, streamwise component colour coded, Y and Z components drawn by the vector (scales are in the lower left corner)

All measurements were synchronized with HPP's SCADA and ADVs using real time clock and merged off-line using Analysis software. This specifically tailored software performs data post-processing, correction for unsteady flow periods (Fig. 44), inspection of each probe's

operation, plot of velocity field (Fig. 45), comparison of EM and ADV results, plot of horizontal or vertical profiles (Fig. 46), selection the criteria for interpolation and extrapolation of velocities, measurement uncertainty assessment and discharge computation (Fig. 47).

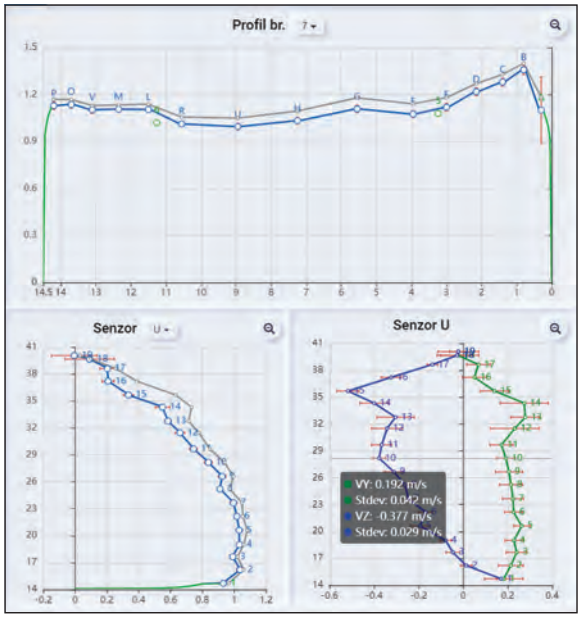


Figure 46. Velocity components with standard deviations: Streamwise velocity in horizontal cross section 7 (upper diagram) and selected components of velocity at vertical cross section U (lower diagrams)

Plot of velocity field (Fig. 45) gives the user a fast overview of the whole system operation. User can change the scales for each velocity component. Also, it is possible to plot velocity components and the standard deviation, where velocity instabilities could easily be observed. Selecting different horizontal profile (no. 7 at 23.66 m a.s.l., marked in Fig 45) and vertical section (U), the velocity profile can be visualised (Fig. 46), with standard deviation plotted as bars around each measured point.

Final step in data processing is the total flow calculation. The user can see the flow components for different areas of cross section (Fig. 47): the largest central part, which is interpolated between corrected EM measurements, and four extrapolated parts (user can select the type and parameters of the extrapolation functions). The bottom part will exclude the deposits at the bottom, if measured using the sonars.

Integral part of total flow calculation is the calculation of uncertainty budget. The uncertainty assessment procedure is developed according to GUM (Joint committee for Guides in Metrology, 2008) and all components are clearly presented to the user. As expected, the combined uncertainty is low (1.11%) for incremental type profiling and during measurement cycle when turbine operation is held constant (4 hours for the measurement shown in Fig. 45, starting at 12 : 23 : 06 and ending at 16 : 22 : 52). For continuous type of profiling, uncertainty of final calculated flow rate ranges from 2% to 5%.

5. CONCLUSIONS

The paper presented four cases of flow rate measurements, done by the author and his colleagues from the Hydraulic and Environmental Engineering department, Faculty of Civil Engineering at University of Belgrade. All selected cases include non-standard flow conditions where commonly used equipment and existing ISO or EN standard could not be used.

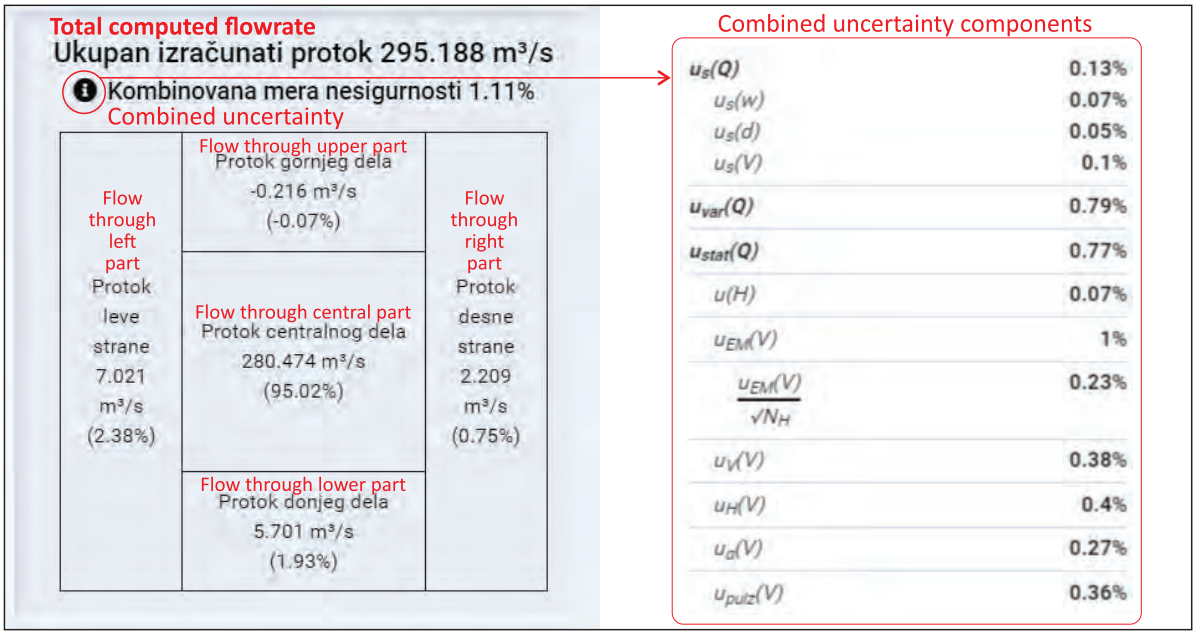


Figure 47. Example of calculated flow components with all components of uncertainty

The author has chosen to perform all presented measurements with EM type probes due to several reasons: EM measurement principle is physically based (2) and water velocity (not the velocity of particles within the water) is directly measured; all probes have full bidirectional measurement; intrinsic to EM principle of operation is the cosine rule (Fig. 4) which is essential for flow measurement in non-standard conditions where streamlines are not perpendicular to cross section; probe shape can be fitted for different purposes and flow conditions; it is possible to construct the 1D, 2D, 3D or 3+D (where streamwise component is measured at two positions, Fig. 8) probes; the EM probes are robust and can measure small velocities from few mm/s up to several m/s; and EM probes can work in harsh environments in sewers, even when covered with the sediments (due to space limitations, such a case is not shown here, but the details could be found in Ivetić *et al.* (2019); Prodanović *et al.* (2012) and Ivetić (2019).

Although the EM probes were used throughout all presented cases, one main drawback must be underlined: the EM probe has to be inserted into the flow field to conduct the measurement, thus it behaves as an obstacle. Measurement with inserted EM probes is not a contactless measurement (although Flat EM probes are close to the contactless type). Other techniques exist that allow true contactless velocity measurement, like PIV (Particle Image Velocimetry), laser Doppler, US Doppler or transit time. But all these contactless techniques do not measure true water velocity but the velocity of a certain tracer!

Finally, one interesting lesson has been learnt from the presented cases, related to the use of the CFD modelling in measurements. If sufficiently detailed modelling grid is prepared, CFD is expected to help the user to see true velocity field in advance, allowing the prediction (computation) of the calibration coefficient for the specific measurement position. However, it was shown that CFD can be used only to describe the flow field qualitatively, give an insight into the expected phenomena, help in selecting the best position for flow meter and give the proof for the results obtained using EM probes. If CFD model itself is not calibrated, it cannot substitute true calibration measurements.

Special thanks

All measurements presented in this paper are the result of team work of colleagues from Hydraulic and Environmental Engineering department, Faculty of Civil Engineering (FCE) at University of Belgrade, during last 30 years. The author wishes to give a special credit to Prof. Čedo Maksimović (Imperial College) and his visionary support for the field of measurements in hydraulic engineering, establishing the final-year course in mid 1980's. All developments in construction of EM probes and other electronic equipment were done by Mile Cvitkovac (Svet Instrumenata). Major field work in the first two cases, together with the help in laboratory tests and equipment calibration, was performed by Dragutin Pavlović (FCE). CFD model used in the second case and Flat EM probe were developed as a part of projects founded by the Serbian Ministry of science. Two last case studies were done under the projects conducted by Water Institute „Jaroslav Černi" (IJČ) for HET Trebišnjica. Damjan Ivetić (FCE) did a great job with his PhD work on EM behaviour when covered with sediments, design of flow measuring systems in cases 3 and 4, and in application of CFD. Many thanks to Predrag Vojt (IJČ) for his support in design and calibration of 3+D EM probes and ADVs for case 4, and for hard field work. Lastly, the author wishes to express gratitude to Branko Hrkić (FCE), retired craftsman and modeler whose support in laboratory and field work was invaluable, and his successor Ivor Koković, with a promising future.

REFERENCES

Baker, R. C. (2002). *An introductory guide to flow measurement* (Vol. 2). John Wiley & Sons.

Bertrand-Krajewski, J.-L., Clemens-Meyer, F. and Lepot, M. (2021). *Metrology in urban drainage and stormwater management: Plug and pray.* IWA Publishing. doi: https://doi.org/10.2166/9781789060119

Bevir, M. (1970). The theory of induced voltage electromagnetic flowmeters. *Journal of Fluid Mechanics, 43*(3), 577–590. doi: 10.1017/S0022112070002586

Biwater Spectrascan. (1998). *Spectralog data logger operation manual.* Hampshire, England.

Danfoss. (2004). *Sonokit product manual for two track mode.*

Hager, W. H. (2010). *Wastewater hydraulics: Theory and practice*. Springer Science & Business Media.

IEC 60041. (1999). *International standard: field acceptance tests to determine the hydraulic performance of hydraulic turbines, storage pumps and pump-turbines.* European Equivalent: EN 60041.

ISO. (2007). *ISO 3455:2007 Hydrometry – Calibration of current meters in straight open tanks.* International Organization for Standardization, Geneva (Switzerland).

ISO. (2010). *Iso 15769:2010 hydrometry – guidelines for the application of acoustic velocity meters using the doppler and echo correlation methods.* ISO – International Organization for Standardization, Geneva (Switzerland).

Ivetić, D. (2019). *Assessment of the liquid flow rate in complex flow conditions with flat electromagnetic sensors* (PhD Thesis). University of Belgrade, Faculty of Civil Engineering (in Serbian).

Ivetić, D. and Prodanović, D. (n.d.). Novel discharge measurement system at the turbine intakes of iron gate 2 hydropower plant: A system description.

Ivetić, D., Prodanović, D. and Cvitkovac, M. (2017a). Improved flow measurement using em flat probes in mixed flow conditions. In *Proceeding of 14th iwa/iahr international conference on urban drainage (icud2017)* (pp. 1734–1737).

Ivetić, D., Prodanović, D. and Cvitkovac, M. (2017b). Improvement of em flow meters accuracy through site-specific cfd calibration case study hps trebinje. In *9th eastern european young water professionals conference.*

Ivetić, D., Prodanović, D., Cvitkovac, M., Ivetić, D., Prodanović, D., Cvitkovac, M., Ivetić, D., Prodanović, D. and Cvitkovac, M. (2018). Water flow measurements in tunnels with combined flow conditions: case study tunnel dabarsko polje – fatničko polje. *Vodoprivreda*, *50*, 229–244 (in Serbian).

Ivetić, D., Prodanović, D. and Stojadinović, L. (2018). Bed-mounted electro magnetic meters: Implications for robust velocity measurement in urban drainage systems. *Journal of Hydrology*, *566*, 455–469. doi: 10 .1016 / j .jhydrol .2018 .08 .068

Ivetić, D., Prodanović, D., Stojadinović, L. and Pavlović, D. (2019). Bed-mounted electro magnetic meters: Assessment of the (missing) technical parameters. *Flow Measurement and Instrumentation*, *68*, 101588. doi: 10 .1016 / j .flowmeasinst .2019 .101588

Ivetić, D., Prodanović, D. and Vojt, P. (2021). Novel discharge measurement system at the turbine intakes of iron gate 2 hydropower plant: a system description. In *8th international conference: Contemporary achievements in civil engineering* (pp. 445–454).

Jaćimović, N., Dašić, T., Stanić, M., Milanović, P. and Djordjević, B. (n.d.). Distributed hydrological-hydraulic modeling of the karst polje water balance. In *Karst 2018: Expect the unexpected.*

Joint committee for Guides in Metrology. (2008). *Guide to the expression of uncertainty in measurement (GUM)*. International Organization for Standardization.

Knudsen, J. and Katz, D. (1958). Fluid dynamics and heat transfer.

Kolin, A. (1936). An electromagnetic flowmeter. principle of the method and its application to bloodflow measurements. *Proceedings of the Society for Experimental Biology and Medicine*, *35*(1), 53–56. doi: 10 .3181 / 00379727 -35 -8854P

Michalski, A., Starzynski, J. and Wincenciak, S. (2001). Electromagnetic flowmeters for open channels-two-dimensional approach to design procedures. *IEEE sensors journal*, *1*(1), 52–61. doi: 10 .1109 / JSEN .2001 .923587

Milanović, P. (2006). Karst of eastern herzegovina and dubrovnik littoral. *Monograph, Association of speleological organizations of Serbia, Belgrade*, *362*.

Nikuradse, J. (1932). *Vdi forschungsheft.*

NORTEK. (2020). *Vector – 300 m.* https : // www .nortekgroup .com / products / vector -300 -m.

OTT. (2021). *Mf pro - water flow meter.* https : // www .ott .com / products / water -flow -3 / ott -mf -pro -water -flow -meter -968 /.

Performance Test Code. (2002). Code 18: Hydraulic turbines and pump-turbines. *American Society of Mechanical Engineers, New York*.

Pope, S. (2008). *Turbulent flows, fifth printing.* Cambridge University Press.

Prodanović, D. (2004). *Water banace on raw and clean water system at banovo brdo wtp.* Study performed for Belgrade Water Works (in Serbian).

Prodanović, D. (2007). Selecting monitoring equipment. In T. Fletcher and A. Deletić (Eds.), (pp. 91–102).

Prodanović, D., Djačić, A., Branisavljević, N. and Rukavina, J. (2012). Laboratory tests of ultrasound and electromagnetic devices for flow measurements in sewer systems. *Aktualna problematika u vodoopskrbi i odvodnji, Bol, otok Brač*, 461–471 (in Croatian).

Prodanović, D. and Ivetić, D. (2000). Flow distribution measurement and performance analysis of water treatment plant. *Štrand, Novi Sad (in Serbian)*.

Prodanović, D. and Pavlović, D. (2003). Smart electromagnetic current flow meter. In *Xxx iahr congress.*

Prodanović, D. and Pavlović, D. (2004). Flow measurement on large pipes. In *25. conference on water and sewer systems '04.*

Prodanović, D., Pavlović, D. and Branisavljević, N. (2011). Flow measurement at the short structures in hydraulic complex conditions: He'djerdap 2'case study. *Vodoprivreda*, *43*(4-6), 103–115 (in Serbian).

Results of flow simulation in topčider mains. (n.d.).

Shercliff, J. A. (1962). *The theory of electromagnetic flow-measurement*. Cambridge University Press, Cambridge (UK).

Simonović, S. P. (1990). An expert system for the selection of a suitable method for flow measurement in open channels. *Journal of Hydrology*, *112*(3-4), 237–256.

Sontex Presentation CD. (2002). *Ruhrgas simulation.*

Steinbock, J., Weissenbrunner, A., Juling, M., Lederer, T. and Thamsen, P. U. (2016). Uncertainty evaluation for velocity–area methods. *Flow Measurement and Instrumentation*, *48*, 51–56.

Stojadinović, L., Ivetić, D. and Prodanović, D. (2018). Laboratory assessment of the flat electro magnetic meter's magnetic field. In *18th sdhi and sdh.*

Svet Instrumenata. (2021). http : // www .si .co .rs /.

UniMeasure. (2020). https : // unimeasure .com / wp -content / uploads / 2019 / 12 / HX -EP -SERIES -CATALOG -PAGES -1 .pdf.

Water Institute „Jaroslav Černi". (2006). *Hydraulic model of iron gate 2 hpp – report on the hydraulic model testing of the incident water flow conditions in the vicinity of the turbines* (Tech. Rep.). Belgrade (in Serbian).

Water Institute „Jaroslav Černi". (2016). *Flow measurement system on fatničko polje – reservoir bileća* (Tech. Rep.). Belgrade (in Serbian): Faculty of Civil Engineering – University of Belgrade, „Svet Instrumenata".

Water Institute „Jaroslav Černi". (2020). *Report on discharge measurement system at the intakes of the iron gate 2 hpp turbines* (Tech. Rep.). Belgrade (in Serbian): Faculty of Civil Engineering – University of Belgrade, „Svet Instrumenata".

Weissenbrunner, A., Fiebach, A., Schmelter, S., Bär, M., Thamsen, P. U. and Lederer, T. (2016). Simulation-based determination of systematic errors of flow meters due to uncertain inflow conditions. *Flow Measurement and Instrumentation*, *52*, 25–39.

Stana Živanović

Stana Živanović is an Associate Professor at the College of Engineering, Mathematics and Physical Sciences at the University of Exeter, UK, since September 2018. Before joining Exeter, she spent nine years at the University of Warwick (2009-2018), seven years (including PhD studies) at the University of Sheffield (2002-2009) and three years at the University of Belgrade (1999-2002).

Contact information: e-mail: s.zivanovic@exeter.ac.uk

https://orcid.org/0000-0001-9888-3806

Prof. Živanović graduated from the Faculty of Civil Engineering, University of Belgrade (FCEUB) in 1999, at the Department of Structural Engineering. She received her PhD degree from the Department of Civil & Structural Engineering, University of Sheffield, UK in 2006. Her academic journey started as an Assistant at the FCEUB, and led to the current position at Exeter after honing her research expertise in Sheffield and Warwick.

Her teaching activities span 20 years across four universities (Belgrade, Sheffield, Warwick and Exeter) in two educations systems (Serbian and the UK's) and at undergraduate and postgraduate levels. She has supervised numerous MSc and MEng theses and supervised/co-supervised six PhD students and five postdoctoral researchers. She was an external or internal examiner for a number of PhD theses (Universities of Liverpool, City London, Bristol, Warwick, Exeter, Oviedo, DTU Dublin) and contributed to two textbooks. She secured undergraduate research grants from the UK Institution of Structural Engineers to support her students final year projects on four occasions and received a few teaching awards by students' popular vote.

Prof. Živanović's scientific research expertise and interests are in modelling human-induced dynamic loading within civil infrastructure, human-structure interactions, human response to vibration and other stimuli and vibration performance of lightweight structures (such as fibre-reinforced polymer (FRP) composites, aluminium, and timber). Reconciling, often contradictory, demands for both vibration serviceable and sustainable structures is another area of growing interest. In collaboration with biomechanics and human motion scientists, she is investigating kinematics, kinetics and control of walking and other forms of human locomotion in the presence of vibration perturbations. She authored/coauthored more than 80 journal and conference papers. Her contributions are cited in Scopus 1560 times (h = 15). Prof. Živanović is an Associate Editor for *Structures and Buildings* journal, a member of Editorial Panel for *Vibration* journal, and served as an Associate Editor of *Shock and Vibration* journal. She is a co-organiser of a Vibration session at the IMAC conference organized by the USA Society for Experimental Mechanics. She has active collaborations with a number of prestigious research institutions, such as Monash University (Australia), Politecnico di Milano (Italy), Tongji University (China), TU Delft (Netherlands), University of Oviedo and Universidad Politécnica de Madrid (Spain). She has peer-reviewed both journal manuscripts for numerous journals and grant applications for national and international research councils and has delivered research talks in Europe, Australia and the USA.

In engineering and professional work, Prof. Živanović co-authored five applied research reports addressing the vibration issues of bridges and building floors for industry clients and contributed to one patent (a removable shear connector). She has collaborated with industrial partners, such as Arup, TSP Projects, Pipex px, Network Rail, Jacobs, and FiberCore Europe.

Human side of vibration serviceability equation: dynamic actions, reactions, and interactions

Stana Živanović

College of Engineering, Mathematics and Physical Sciences, University of Exeter, United Kingdom

Summary

This paper discusses vibration serviceability assessment of structures that evolved from a secondary design check in the last century to a governing design criterion at the present time. The focus is on human-induced dynamic actions, human reactions to vibrations and human-structure interactions. Research into understanding and modelling these aspects, supported by state-of-the art experimental facilities at the University of Exeter, is presented, and challenges for future research are formulated. Further understanding of how active humans respond to and interact with vibrations of the structures they occupy is identified as a major requirement for design of modern, slender, and serviceable structures in the future.

1. INTRODUCTION

Although the phenomenon of vibration intrigued scientists, including Galileo and Newton, for a long time, it was only with arrival of industrial revolution that complex relationship between vibration and humans started to gain more prominence. With mechanical vibrations becoming ever present in human lives, systematic research into impact of these vibrations on human health, comfort and wellbeing started to appear at the end of the 19th century, and advanced rapidly during the 20th century (Schiavi *et al.*, 2016).

The issue of vibration soon started to influence design of civil engineering structures. While satisfying both ultimate limit state (ULS) and serviceability limit state (SLS) criteria has traditionally been embedded in the design, it was the ULS which tended to govern for a long time. With development of high-strength materials, the design of civil engineering structures, especially bridges and grandstands, started to change towards more daring structural forms that not only serve the original purpose but also represent city landmarks and tourist attractions. Increasingly, these structures, typically operating at stress levels well below the safe limit, started to suffer from excessive vibrations due to dynamic actions. This is due to the increase in material strength resulting in structures that are lighter, less stiff, and less damped than their predecessors. In addition, the serious concern about climate change means that the structural engineering has an important role to play in reducing, and ultimately negating, the carbon emissions. First and obvious step in this process is to minimise the amount of construction materials used in any particular project, which leads to even lighter and less stiff structures, with even more pronounced vibration serviceability (VS) issues. It is then not a surprise that the vibration serviceability limit state (VSLS) typically govern the design of structures occupied by humans.

Excessive vibrations were reported on numerous structures, especially grandstands, footbridges, staircases, and long-span building floors, dynamically excited by structure users over the last 20 years. This development introduced multiple new challenges to structural engineers, most of which are associated with human's impact on the VSLS and none of which are studied in much detail in traditional civil structural engineering degree courses. Three challenges will be discussed in this paper: (1) how to accurately model dynamic actions by humans (such as walking, jumping, bouncing, and running), (2) what vibration response criteria should be used to avoid comfort/health/wellbeing issues for structure users and (3) whether and how structure users interact with oscillating floors and bridge decks? Note that bouncing is defined here as human's up and down body movement whilst both feet are in the contact with the floor, while jumping involves alternating contact and non-contact (flying) phases.

This paper aims to present research achievements in the three aspects, with particular emphasis on recent and ongoing work at the University of Exeter. The paper does not cover all details of research into VSLS as this would require more space; instead those in which the author has been most involved, past and present, are given the preference. Although every human activity generates a three-

dimensional force, only the largest – vertical – component will be analysed. In addition, most attention will be given to walking which is the most frequent human activity. Occasionally jumping, which normally generates the largest dynamic force among possible human activities, will be commented. Note that most people make between 1.4 and 2.5 steps every second, while jumping typically occurs in the 1.5–3.5 Hz frequency range. The approaches in researching activities reported in this paper provide a strong conceptual starting point for studying all the other activities.

After this introductory section, Section 2 describes developments in modelling human-induced dynamic loading. Next, human vibration comfort is addressed in Section 3, while Section 4 covers research into human-structure dynamic interaction. Section 5 formulates relevant research questions while Section 6 presents selected research in the author's research group. Section 7 summarises main conclusions.

2. HUMAN ACTIONS

Motion of the human body results in a dynamic force imparted on the supporting structure. Among the human actions, those that are repetitive, such as walking, running, jumping or bouncing, are of special interest as they have capability to either excite resonance response in a single structural vibration mode, or induce broad-band impact on the structure that results in a series of impulse responses composed of contributions of multiple vibration modes, or both. The structures that are prone to these effects are often landmark public structures such as footbridges, walkways, and corridors between buildings, at airports and shopping malls, long-span building floors, staircases, and cantilever grandstands.

Measuring the dynamic loading by humans is the first step towards successful modelling. Traditional focus is on recording the time-domain waveform of the dynamic force, as well as understanding its frequency content. Ideally, kinematics of the human body motion, that is the underlying cause of the generated force, should also be recorded.

Force plates and instrumented treadmills have been typically used for observing human kinetics, i.e. ground reaction force (GRF) imparted into the supporting structure while walking (Racic *et al.*, 2009; Živanović *et al.*, 2005). The typical shapes of GRFs for walking, jumping, running, bouncing and similar everyday activities are well known. An example of a GRF generated by a test subject walking on a split-belt treadmill is shown in Figure 1. Each foot generates an M-shaped force. These individual forces as well as their sum are presented in the figure.

Quasi-periodic nature of the total GRF can be observed in Figure 1. Removing the static component (human body

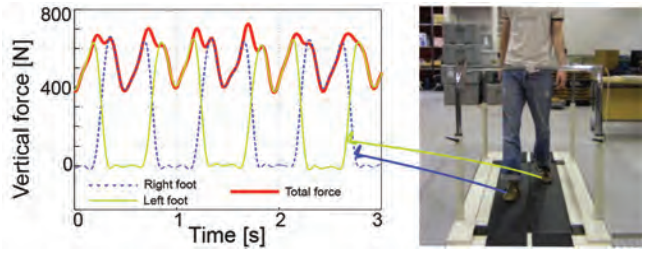


Figure 1. Vertical ground reaction force generated by walking on an instrumented treadmill

weight) from the signal results in a dynamic component whose time- and narrow-band frequency-domain signatures are shown in Figure 2. Note that the first harmonic (grey line in the figure) provides dominant contribution to the force signal. Its frequency corresponds to the frequency of walking. The second harmonic at twice this frequency is the second largest. It is these two harmonics that are most often considered capable of exciting resonance in footbridges and therefore creating vibration serviceability problems. This was the reason for their inclusion in the design guidelines, in a simplified form, as early as in 1978 (BSI, 1978).

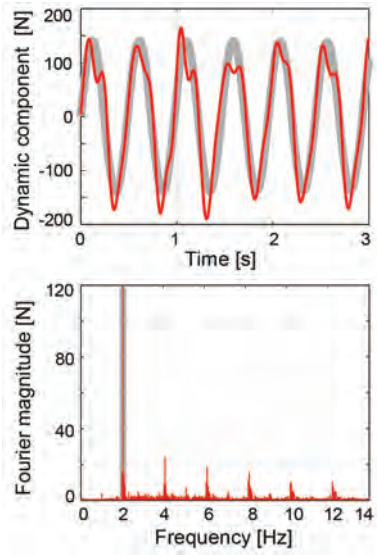


Figure 2. Time- and frequency-domain representation of the dynamic component of the vertical force

In essence, the walking (and all other activities) can be represented as a deterministic force, $F(t)$, written as a sum of any number of harmonics of interest, in this case N:

$$F(t) = W\left(1 + \sum_{i=1}^{N}\left(DLF_i \sin(2\pi i f_p t + \varphi_i)\right)\right) \quad (1)$$

where t is an independent time variable, DLF_i is a dynamic load factor (i.e. force amplitude divided by body weight W) for i-th harmonic, f_p is a pacing (activity) frequency, φ_i is the phase shift of the i-th harmonic, i is the

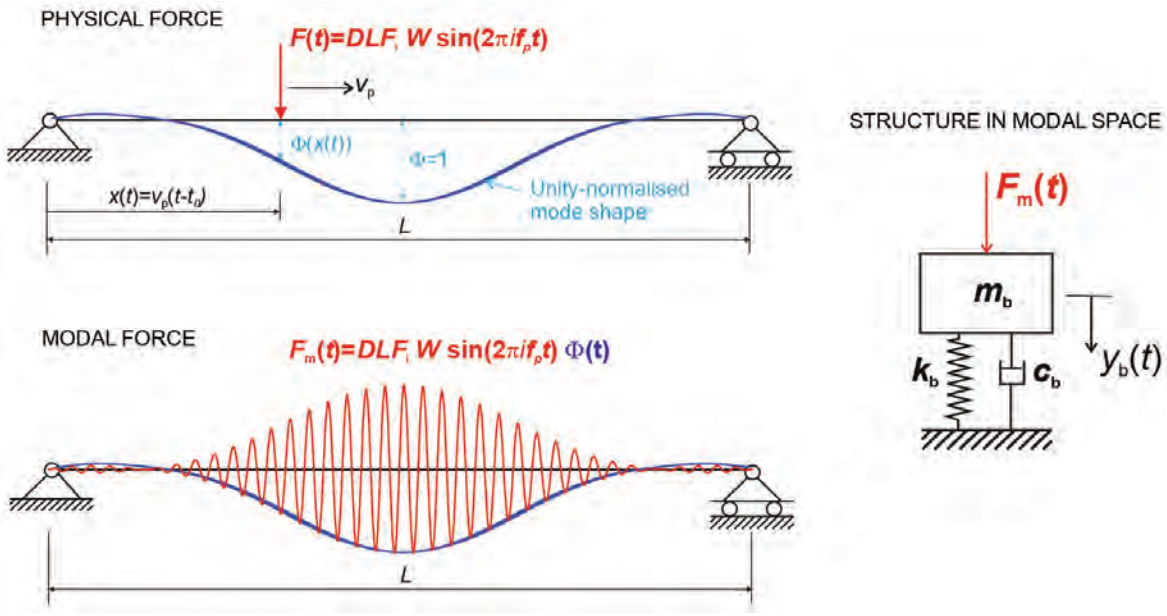


Figure 3. From physical model of the pedestrian's force $F(t)$, via modal force $F_m(t)$ to setting the modal model of the structure to calculate the modal response y_b. v_p is pedestrian's speed, Φ is the mode shape, t_o is the time instant when pedestrian enters the bridge and m_b, c_b and k_b are modal mass, damping and stiffness of the structure, respectively

order number of the harmonic, and N the total number of contributing harmonics. Each activity has a specific range of values for dynamic loading factors and activity frequency.

The model has been, and still is, the basis for the VS assessment due to most human actions, whereby specifics of each action are accounted for by appropriate choice of parameter values. For example, dynamic load factors for the first harmonic due to a person jumping takes value from as low as 0.2 for slow, low intensity, jumping at 1 Hz to 1.6 for fast, high-intensity, jumping at 3 Hz (McDonald and Živanović, 2017), while for a person walking this parameter can be as low as 0.1 at 1.4 Hz and as large as 0.7 at 2.2 Hz (Kerr, 1998).

The force generated by walkers changes in time and space. In case of footbridges, only the harmonic that can excite resonance in a particular mode of vibration is usually modelled. In modal space, this means that the physical force has to be weighted by the mode shape to calculate the modal force. This force, in turn, is applied to a single-degree-of-freedom (SDOF) model of the relevant vibration mode to calculate the vibration response at the point on the bridge that represents the antinode of the mode. The process is sketched in Figure 3 and it can easily be adjusted to account for multiple harmonics, multiple vibration modes and any time domain representation of the dynamic force. Note that despite the people generating the force at the discrete locations on the structure at which the feet are in contact with the ground, the modelling applies the forces in a simpler, spatially continuous manner, as the differences in the response prediction are negligible (Ahmadi *et al.*, 2019; Caprani and Ahmadi, 2016).

Reliable prediction of the vibration response of the structure is an ultimate aim of the VS assessment. While the process described in Figure 3 is essential starting point for researchers and engineers interested in the VSLS, it should be noted that the process of force modelling evolved significantly from the initial representation in Equation (1) over the last 20 years. There now exist sophisticated models that account for variations in pedestrian parameters within human population (in both single pedestrian and/or crowd load scenarios) that aim to predict probability of occurrence of a specific vibration response (Brownjohn *et al.*, 2004; Chen *et al.*, 2014; Tubino *et al.*, 2020; Van Nimmen *et al.*, 2020; Živanović *et al.*, 2007). This important move from deterministic to probabilistic estimation of the structure vibration has been a major conceptual shift in the VSLS. It provides the structure owners with means of considering the target structure users and their specific expectations of the structure behaviour. In this way, footbridges serving sensitive users (such as near hospitals) or providing the only pedestrian route in the area could be designed to stricter vibration criteria so that most users do not experience critical vibrations as opposed to a rarely used bridges in remote locations that might be allowed to vibrate more given infrequency of their use. Another important development in force modelling is that there are models, such as the one developed by Racic and Brownjohn (2011), which are capable of reproducing both time- and narrow-band frequency-domain representation of the force signal, removing the need for simplification of the original waveform. This particular model successfully borrowed strategies from numerical generation of electrocardiogram and

speech recognition to faithfully reproduce the original signals, demonstrating the benefits of applying the scientific knowledge developed in the research areas outside the civil engineering.

In summary, there exists a large body of research, only a small subset of which is cited in this paper, that provides sufficient information for successful modelling of human-generated dynamic load as it occurs on *rigid support surfaces*. Namely, majority of measurements that were used to develop the mathematical models were acquired on rigid, i.e. not-oscillating, supporting surfaces. As the structure occupied by a human is forced to oscillate due to the very dynamic force induced by this and other humans, then structural vibrations might, in turn, force people to adapt to them and therefore alter their typical body kinematics. As a result, the force generated by the pedestrian can, in some cases, differ substantially from the one that would occur had perceptible vibrations not happened, and therefore lead to predicting onerous vibration response. This phenomenon of human interaction with vibrating structures is currently a major research challenge and it will be subject of Section 4. Beforehand, human reaction to vibration has to be introduced.

Note that only the kinetics of the human actions, traditionally analysed in structural engineering applications, was considered in this section. However, to further understanding of human actions on flexible surfaces, underlying kinematics of body parts will need to be included. Figure 4 provides an illustration of vertical displacements of various body parts during walking; the trajectory of the body centre of mass (BCoM) being of most interest in the modelling.

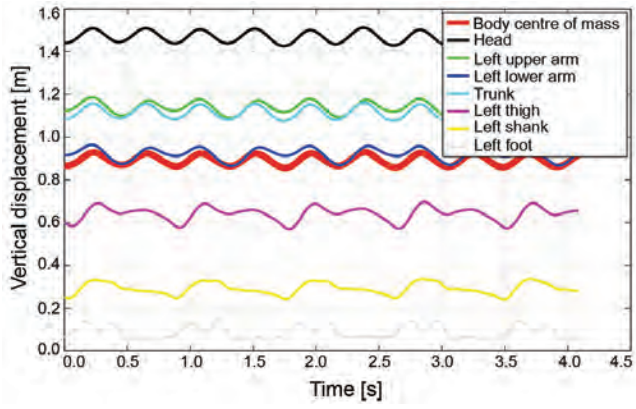


Figure 4. Displacement of human body parts while walking

3. HUMAN REACTION TO VIBRATION

Pedestrians crossing perceptibly vibrating structures are exposed to whole body vibration (WBV) that is entering the body via feet-deck interface. In this paper human reaction to vibration is defined as a subjective, psychological response to WBV. This response is crucial for deciding a comfort rating of the structure and it is an important measure of usability of the structure. The intention of this section is to introduce walking posture related comfort issues.

When it comes to design of slender pedestrian structures, such as footbridges, walkways, and even long-span floors with predominantly active workforce, the comfort of human users of these structures is central to successful design. It is well known that human response to WBV is a function of vibration magnitude, vibration frequency, and exposure time and that it differs for different postures and vibration direction (Griffin, 2012). Despite this, little research exists on *pedestrian*'s response to vibration, as opposed to standing/sitting/lying person's response that is frequently researched for the needs of transportation industry as well as in residential settings. As a result, the approach to defining vibrating limits for pedestrians on, say, footbridge structures, differs between different guidelines, and often does not consider all factors of interest. Ideally, for each vibration frequency, a vibration acceleration limit would be defined for a pedestrian as a function of the exposure time. Unfortunately, lack of experimental facilities for studying pedestrian response means that the current national and international design guidelines are underdeveloped and inconsistent. They often also lack rigour and, unsurprisingly, experimental verification.

As early as 1966, Leonard argued that designing footbridges for standing vibration receivers is not justifiable from economic point of view. Despite this, the International Standardization Organisation's guidelines for buildings and walkways defines the vibration limit for pedestrians by directly multiplying a frequency-dependent vibration perception threshold defined for standing people by a factor of 60 (ISO, 2007). In this way, the standard rightly acknowledges that pedestrians are less sensitive to vibration than the stationary (i.e. non-moving) people but it also incorrectly assumes that the walkers are most sensitive to the same frequency range as standing people, i.e. 4-8 Hz (Figure 5). However, recent evidence points out that the fundamental natural frequency (in the vertical direction) of a person walking is in the range 1.8-3.0 Hz (Ramírez *et al.*, 2013; Shahabpoor *et al.*, 2016; Wang *et al.*, 2017; Zhang *et al.*, 2016). This finding implies that the pedestrian is most sensitive to vibration in this range, rather than 4-8 Hz. At the other end of the spectrum, widely used guidance for design of footbridges, sets frequency-independent vibration limits (BSI, 2008; Sétra, 2006). Sétra sets three vibration limits, depending on the comfort requirements for a particular structure: 0.5 m/s^2 for maximum comfort, 1.0 m/s^2 for mean comfort and 2.5 m/s^2 for minimum comfort. BSI (2008) has the same lower limit, but it requires vibration to not exceed 2.0 m/s^2. Differently from Sétra's guidance, these limits are defined as function of environment in which the bridge resides (e.g. availability of alternative routes, fre-

quency of bridge use, and perceived stability/safety of the structure).

In addition to inconsistencies in accounting (or not) for frequency-dependency, none of the current design rules considers influence of exposure time on the pedestrian's rating of the structure's serviceability. For example, a typical low-mass composite footbridge that spans about 15 m is expected to satisfy the same vibration criterion as, say, a 60 m-span bridge despite its users' vibration exposure time being about four times shorter. This leads to unwanted conservativeness of the current design guidance when it comes to evaluating VSLS of short-span bridges and emphasises the need for further research and update of the current design rules (Uyttersprot and De Corte, 2021).

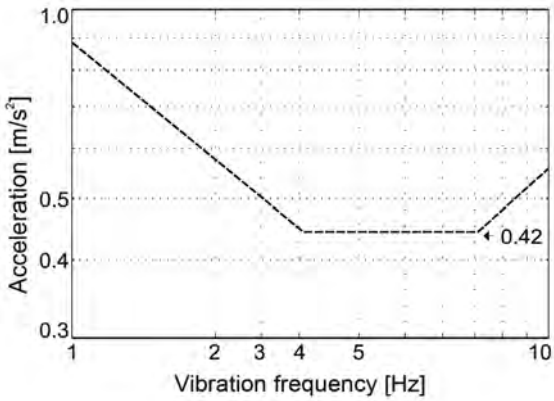


Figure 5. Vibration limit for walkers after ISO (2007)

4. HUMAN-STRUCTURE DYNAMIC INTERACTION

Understanding vibration behaviour of contemporary pedestrian structures is limited in cases when pedestrians are being influenced by structural vibrations. The current design guidelines that rely on "non-interactive models" are increasingly unreliable and outdated. For example, Figure 6 shows a predicted response (grey shade) of a 63 *m*-span footbridge to a pedestrian walking at the resonance frequency to be about 3.6 times larger than the maximum response measured (black shade). The pedestrian was likely influenced by vibrations rising above 0.6 m/s^2 resulting in significant discrepancy between the measured and the predicted responses. This has a serious consequence of misclassifying the bridge as being on the brink of "unacceptable structure" (predicted peak response of 2.5 m/s^2) instead of being classified in a more favourable category of "providing mean comfort" (measured response of 0.7 m/s^2), if solely the peak response (as defined by Sétra 2006) is used as a means of determining structure comfort. It is likely that the reduction in the actual response came as a result of pedestrian's either conscious or unconscious reaction/adaptation to vibrations. A sophisticated vibration prediction model would need to consider pedestrian's subjective experience of vibration and conditions that led to altering the usual locomotion.

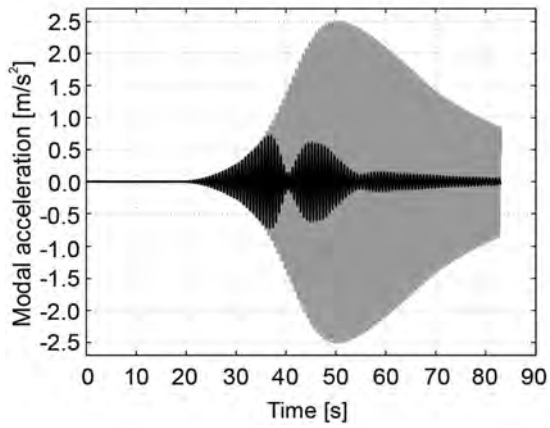


Figure 6. Calculated (grey) and measured (black) response to a pedestrian

Human body is a "structure" that has its own mass, damping and stiffness parameters. This dynamic system, when in passive (e.g. standing or sitting) posture, is known to interact with dynamic properties of the supporting structure, creating a new human-structure dynamic system (Sachse *et al.*, 2004). Depending on the human's location on the structure, exact posture, and human-to-structure mass ratio, the human could substantially increase the system damping and either increase or decrease (usually slightly) natural frequencies of the system. The cases of interaction in active postures, e.g. when running, walking (Macdonald, 2009; Van Nimmen *et al.*, 2014), bouncing (IStructE, 2008), and jumping (Gaspar *et al.*, 2020) have been documented only recently, and, for the vibration in the vertical direction, they often result in lower-than-initially-expected response.

Among all human actions, walking is among most complex to model – not only due to complexity of the underlying locomotion but also due to change in spatial position and therefore continuous change in vibration perceived by a walker. When close to location of supports of a bridge, a pedestrian is unlikely to be aware of any vibrations, whilst when traversing a location close to an antinode of a vibration mode, the vibration might become not only noticeable but also uncomfortable and even intolerable.

In addition to traditional interest of structural engineers in the human-generated GRF, it is expected that successful modelling of a pedestrian-structure interaction (PSI) also addresses the underlying kinematics of the person's motion. Two types of pedestrian models are dominating the research scene at this moment in time: (1) spring-mass-damping models – most often in form of a single-degree-of-freedom (SDOF) model, and (2) various variants of bipedal inverted pendulum models. In both cases, body kinematics is represented by lumping the body

mass into a single point that is meant to represent the BCoM.

The development of the first type of model has been inspired by a relatively mature use of the SDOF modelling in some structural engineering applications, such as modelling of a passive (standing or sitting) person (Ellis *et al.*, 1997; Sachse *et al.*, 2004; Zheng and Brownjohn, 2001) or an active person performing bouncing on grandstands (IStructE, 2008). When the SDOF system is "attached" to another SDOF system representing the structural vibration mode with known modal properties, a new 2DOF human-structure system is created (Figure 7). To determine the dynamic properties of the human DOF, a frequency response function (FRF) of a human-structure system is usually measured. An analytical FRF model that best fits the experimental data is then used to identify inherent mass, damping and stiffness of the human. An extensive literature search (Jones *et al.*, 2011) reported that the identified damping ratio for a standing human ranges between 33% and 69%, the natural frequency between 3.3 Hz and 10.4 Hz while the mass is close to the physical mass of the person. Identifying body dynamics parameters of an active person is more challenging due to continuous change in body posture and resulting alternations in tensing and relaxing various muscle groups. A reasonable compromise would be to identify an average damping, natural frequency, and stiffness of the equivalent SDOF model. The IStructE (2008) guidance for grandstands suggests a damping ratio of 25% and natural frequency of 2.3 Hz as dynamic properties of a bouncing human, with the mass in the model being taken as the actual mass of the person. These properties were derived by minimising the discrepancy between the force generated by the bouncer model and that measured on the rigid surface (Dougill and Parkhouse, 2006; IStructE, 2008).

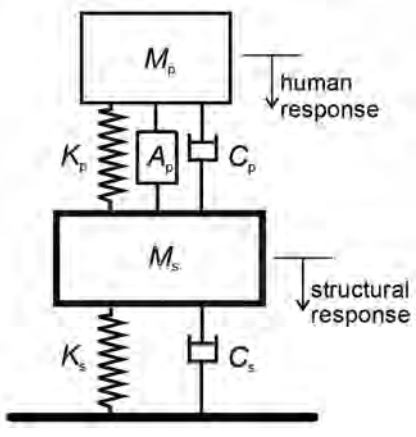


Figure 7. 2DOF modelof pedestrian (p) – structure (s) system

This approach has been extended to pedestrians with the aim to identify pedestrian-specific natural frequency and damping ratio. In addition, some authors added an internal actuator (A_p) to the classical SDOF properties to model a pedestrian (Figure 7). The actuator represents the muscle-generated force which drives the pedestrian and is also transferred to the bridge deck, similar to modelling bouncing people in the IStructE (2008) guidelines. Conceptually, this force, modified by the damping and stiffness of the legs, represents the force that pedestrian imparts on the vibrating structure. Naturally, the force is a function of the vibration amplitude and both stiffness and damping of the legs. Identification of parameters of this model (either with the internal actuator or with the actuator replaced with a force typical of walking on rigid ground) is usually performed by minimising the difference between the simulated and measured force at the pedestrian-structure interface. The identified pedestrian properties often differ depending on specifics of the optimisation procedure and vibration parameters used in the experimental part. Nevertheless, the indications are that the natural frequency (of the first vertical vibration mode) of the human body is between 1.8 and 3.0 Hz, as stated in the previous section. This is accompanied by a damping ratio in the 27-55% range. The critical evaluation of these models on as-built structures is not yet available, and they are still not ready for inclusion in VSLS guidelines. Note that, by its very nature, this model primarily aims to simulate pedestrian's contribution to the damping and natural frequency of the joint human-structure system, while it is less successful in reproducing the actual GRF and kinematics of the BCoM. Majority of research into this model does not analyse the last two parameters.

The inverted pendulum-based type of the pedestrian models, originally developed by medical researchers (Saunders *et al.*, 1953), have found their way into the civil engineering applications relatively recently. They range from simplest bipedal inverted pendulum model with rigid, massless legs (BIP), to relatively complex two- and even three-dimensional models whereby both stiffness and damping of the pedestrian's legs are included into the analysis. Usually, the two legs are assumed to have identical properties, which is a reasonable starting point for non-pathological gaits.

Several bipedal models of increasing complexity are shown in Figure 8. More detailed explanation of each is provided by Dang (2014). Their main feature is that the body mass is lumped into a single point representing the BCoM. As a result, the human kinematics is represented by the movement of this single point, and it is influenced by the geometry, damping and elasticity (when present) of the legs. The upper body is considered to be a passenger unit, whose contribution to the pedestrian kinematics is neglected (Perry and Burnfield, 2010).

The BIP model of the human includes point, no-slip, feet. This means that the geometry of each foot is modelled as a single point which, during a single step, stays at the same location on the supporting surface. The movement of the model can be described by a second order

Figure 8. Bipedal walking models: inverted pendulum, spring mass and spring mass with damper, respectively. Solid lines represent each model in its basic, point-foot, mode, while the dashed line represent addition of rocker-foot

differential equation, with initial conditions usually expressed via an attack angle (initial angle between the leg and the reference) and angular speed at the moment the foot touches the ground. At the end of each step, an upward impulse is applied to the BCoM to prevent it from falling and to help its transition to the next step, imitating the effect of a push (at toe-off) by a trailing leg. The rigid leg and no-slip nature of the model result in a vaulting motion of the BCoM, i.e. a series of arcs, with an abrupt change in the direction of the BCoM at each transition from one leg to another. This differs from a smooth BCoM trajectory seen in Figure 4, but the BIP model does approximate the single stance phase of walking locomotion and the first harmonic of GRF relatively well. However, the point foot and the rigid leg assumptions result in inability to represent forward progression of the centre of pressure in the foot and double support phase, respectively. It also overestimates the vertical excursion of the BCoM.

As foot and ankle are important in controlling the vertical excursion, Hansen *et al.* (2000) replaced the point foot by a rocker foot. This enabled the simulation of the forward progression of the centre of pressure from "heel" to "toe" reducing the total excursion of the BCoM (Lee and Farley, 1998). The model, however, still has issues of not modelling the double support phase. This drawback is resolved by introducing the elasticity of the legs in the spring-mass model (SMM) by replacing rigid legs in the original BIP model with elastic springs of stiffness k (Geyer, 2005). Differently from the previous two models, the SMM does not need external power – instead its kinetic energy of motion and potential energy of BCoM elevation and spring deformation are continuously exchanged, driving the model once set in motion. Conceptually, this model is, kinetically and kinematically, closer to human walking than the rigid-leg version. Addition of a rocker foot to the SMM with the point foot, as shown in the second picture in Figure 8, adds further advantages, equivalent to those already explained. A disadvantage is that computational effort might become prohibitively expensive.

Lack of the energy supply in the SMM is computationally convenient, yet it contradicts the research by Cavagna *et al.* (1976) who found that human body in walking dissipates energy. This is motivation behind introducing damper elements in the legs. However, in this case some external work has to be done to compensate for the energy loss. A control forces of various levels of sophistication have been proposed; this paper will not go into the details beyond stating that some suggestions for external work modelling can be found in work by Qin *et al.* (2013) and Lin *et al.* (2021). The spring constant for the leg is in the region 22-34 kN/m, and the damping coefficient in 50–230 Ns/m (Dang, 2014; Kim and Park, 2011). While the model with spring-damper legs (with or without rocker feet) improves qualitative agreement with the actual human locomotion, it is also more computationally demanding, and more difficult to evaluate and verify, due to larger amount of parameters involved which are not necessarily mutually independent.

Most of the models described in this section (Section 4) have underwent only a limited calibration against experimental data. More extensive evaluation of performance of these models on actual structures and their critical comparison are required if any of the models are to be adopted by VS design guidelines. For now, there is no consensus which model is most appropriate, from both conceptual and practical point of view, for use in civil engineering. It is of particular importance to identify the key mechanisms that are used for human's adaptation to vibrating surface and evaluate models' ability to represent these. This is a very much work in progress. Some candidate factors that are considered potentially important for inclusion in the pendulum-based models are foot clearance, and joint (knee and/or ankle) control. Crucially, making this final step in modelling humans on vibrating structures requires access to advanced experimental facilities, similar to those introduced in Section 6. Before introducing the facilities, a series of research questions that are expected to be investigated in near future are summarised in the next section.

5. RESEARCH QUESTIONS

The overview of the main components of the VS research presented so far has an aim of setting up the scene for entering the research area by new researchers and helping to understand the research challenges that await our research community short-to-medium term. Each of the three areas covered are at the different stages of development and have potential to be advanced in the years to come. Here I formulate questions in the areas of human response to vibration, human-structure interaction and uncertainty propagation in VS assessment that are essential for making required progress in the design against dynamic actions by humans.

5.1. Pedestrian's response to WBV

The questions within this challenge are:

- Which vibrations of the structure are perceptible by pedestrians?
- Which vibrations are acceptable, and which are unacceptable/intolerable by pedestrians?
- What vibration limits the structure should satisfy if to bit fit for the intended purpose?
- Apart from the vibration-related factors, should audio and/or visual stimuli also be included in setting up the VS criteria?

The first two questions are fundamental for providing scientific understanding of pedestrian as a vibration receiver. The third topic addresses a practical need for providing structural engineers with tools for assessing the acceptability of the structure, and it is likely to incorporate some information about the structure itself, including its significance and the conditions in which it resides. The last point has been rarely investigated, but would provide beneficial refinement of the vibration criteria, given that naturally noisy conditions (e.g. footbridges in or close to the airports) or more or less visually pleasant environments could alter human sensitivity to vibration. More generally, it is multi-modal sensitivity to structure and its environment which should be considered in a holistic way.

Answering these questions would provide more reliable prediction of structure's comfort and reduce economic harm caused by the instances of vibration rectification due to unexpected, excessive vibration once the structure is built.

It should be noted that this challenge is related not only to pedestrians, but also other active (moving) humans (such as those jumping or bouncing) as potential vandals on slender structures or simply as participants in the music or sport events on grandstands and in similar settings. More energetic the human is, less sensitive to vibration the human is expected to be. For example, people jumping are less sensitive to vibrations than people walking. Understanding their sensitivity to vibration (or lack of it) could help to identify potential self-imposed vibration limit. For example, it is likely that there is a vibration level that, once felt by an active person, would cause the concern about personal safety and lead to stopping the activity, in the same way as excessive sway of the London Millennium Footbridge on its opening day in 2000 had a large number of pedestrians reaching for the handrails and stopping walking (Dallard *et al.*, 2001).

5.2. Pedestrian's interaction with the structure

The main questions in the PSI area are:

- Under which conditions a pedestrian interacts with the structure and how is this reflected in human behaviour?
- What are the main effects of PSI on the pedestrian-structure system?
- What modelling approach is best suited to model PSI and provide reliable prediction of the vibration response of the structure?

The first question is again a fundamental question that requires discovering the drivers behind pedestrian's reaction and adaptability to structural motion. This is likely to be manifested in kinematics of human locomotion, whereby a pedestrian introduces adjustments to their walking in response to vibration perturbation. Addressing this question requires multi-disciplinary approach that includes utilising knowledge in psychology and biomechanics.

The second question is motivated by the ongoing discussion about the main effects of PSI on the system. Is the main effect in the influence on the system's damping and natural frequency (which is targeted in modelling pedestrians as SDOF systems explained earlier) or in the alternation in the GRF (which is better modelled by pendulum models), or both? Experimental evidence of presence (or absence) of the two potential changes observed at the same time would be an ideal approach to answering this question, which will then enable choosing the model that is best suited to reflect the measured phenomena and that can be recommended for design use (as an answer to the last question of direct interest to structural engineering design community).

Again, there is a scope for addressing similar questions in other than pedestrian activity – successful modelling of interaction between a person jumping/bouncing/running and the structure would be valuable for improving reliability of modelling the vandal loading, or crowds in stadia, and therefore understanding extreme human-generated loading conditions to which the structure could be exposed.

5.3. Propagation of uncertainties in VS assessment

This challenge is expected to influence design practice to a large degree. It is clear that dynamic loading and response to vibration within a human population are influenced by both intra- and inter-subject variability. These can be accounted for by utilising probabilistic approach to modelling both components in the VS assessment. In addition, predicting the dynamic properties of the civil engineering structure itself is not a straightforward task given inevitable presence of uncertainties in mechanical properties of the construction materials, dynamic behaviour of connections and supports, the amount of friction and material damping, and similar. Although some work has been conducted on this challenge (Tubino *et al.*, 2020), further research is required to define an effective strategy for inclusion of the uncertainties in the VS assessment.

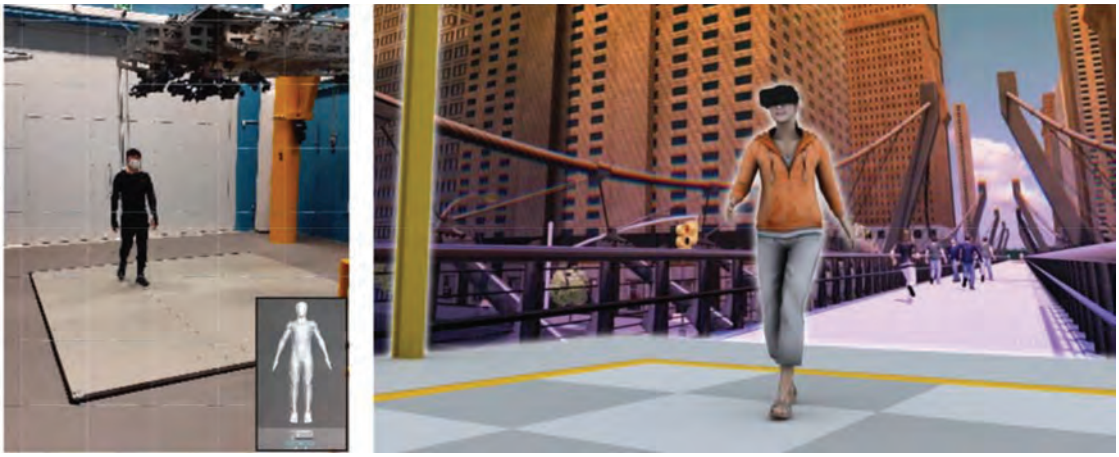


Figure 9. Left: A test sibject in VSim and his avatar. Right: Walking in a virtual reality scene in VSim

6. RECENT RESEARCH DEVELOPMENTS IN THE RESEARCH GROUP

This section introduces selected research endeavours in which the author is or has recently been involved. Key research facilities in use at the University of Exeter are introduced first.

6.1. Experimental facilities

Given the lack of fundamental insight into PSI and WBV challenges as well as verification of various theoretical/modelling developments, sophisticated facilities are required to make a research breakthrough. This was the reason that the University of Exeter have recently invested in two unique, state-of-the-art facilities: a motion simulator VSim, and an ultra-lively Exeter Footbridge (EF). Both facilities are situated at the Exeter Science Park, and will be described here.

The £3.8 million vibration simulation platform, VSim, is a six degree-of-freedom system driven by electric actuators. It is optimised for simulating motion typical of civil engineering structures in the frequency range 0.5-35 Hz. The maximum displacement range for an individual axis is ±21 mm. Nine multi-axis force plates, covering the 3.6 m × 3.6 m area, are built into the floor of the platform. A treadmill can be installed on the top of the plates, as and when needed. Kinematics of the human body in walking is observed using an optical or inertial motion capture system. All raw data streams (acceleration of the platform, human body kinematics and GRF) are synchronised. VSim can generate the vibration signals of chosen vibration frequency, amplitude, and duration. In addition, it is equipped with nine HTC Vive-pro tethered Head-mounted displays for realistic simulation of built environments. VSim provides a unique capability of exposing a test subject to chosen vibration and visual stimuli. There is also an opportunity to incorporate additional sound effects. Figure 9 shows both physical and a virtual environment in VSim.

The EF is a footbridge made of fibre-reinforced-polymer (FRP) composites. This simply supported shallow truss structure is 19.8 m long with 2.1 m wide deck. The two supports can be moved to vary the span between 14.0 m and 19.6 m to achieve a fundamental natural frequency of the vertical vibration mode from 3.4 Hz to 1.8 Hz, with the frequency of the second flexural mode being between 9.8 and 6.7 Hz. A pressure insole system Tekscan is available to measure pressure distribution beneath the feet, a portable force plate for measuring the force generated in one step while walking and inertial measurement units (IMUs) could be used to record kinematics of body motion. The bridge and these measurement devices are shown in Figure 10.

The bridge complements the VSim as it enables investigations of PSI and WBV when vibrations are generated by a pedestrian, rather than imposed by motion platform. The first case involves a two-way interaction while the structure-to-human interaction only takes place in VSim.

An example of vibration generated by a test subject crossing the EF at various pacing rates (from a slow rate at 1.4 Hz to fast at 2.5 Hz) is shown in Figure 11. The test subject excited the bridge up to the vibration level of 6 m/s^2 when walking fast and matching the natural frequency of the bridge of 2.5 Hz. This low-frequency vibration response is larger than on any other pedestrian laboratory structure known to the author, and it is suitable for triggering and studying human reactions and adaptations to lively decks.

6.2. Research developments

Pedestrian's response to WBV is one of the subjects that is expected to benefit substantially from the new facilities. We started this line of work at the University of Warwick on the former Warwick Bridge – a steel-concrete composite bridge of similar span to the new EF, but about 20 times heavier, and with natural frequency at around 2.2 Hz. Three pedestrians, one at a time, were asked to

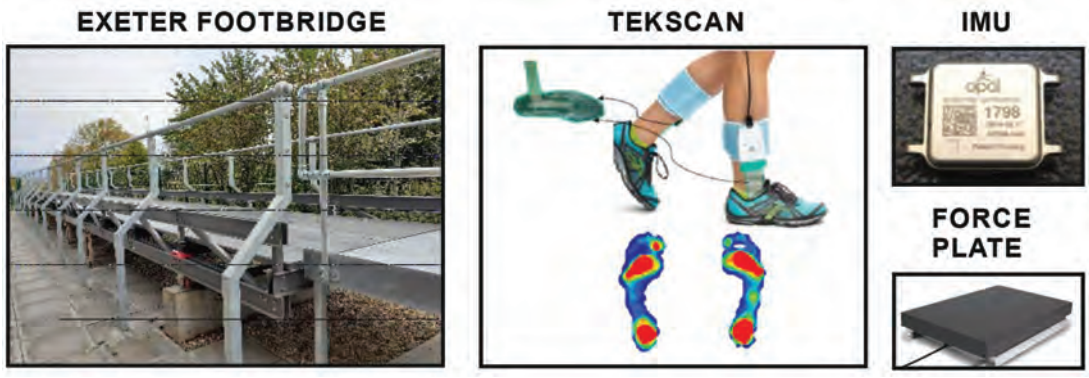


Figure 10. Exeter Footbridge and accompanying measuring equipment (a pressure insole Tekscan, an inertial measurement unit and a force plate)

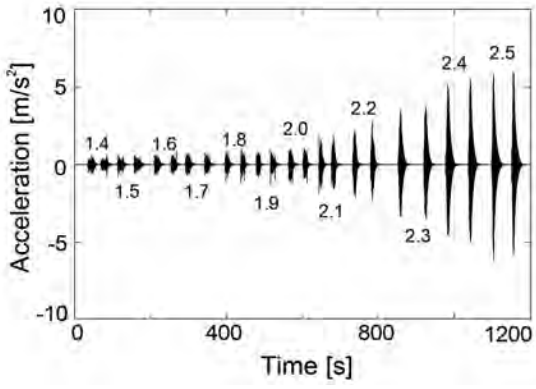


Figure 11. Mid-span acceleration response to a single pedestrian crossing EF at different pacing rates (in Hz) denoted in the figure. Two crossings are recorded at each pacing rate

walk on a treadmill placed at the midspan. The bridge was excited either by the pedestrian themselves or by the pedestrian and a shaker (to further increase the vibration amplitude). The walking was performed at a range of different treadmill belt speeds, from which the pacing rate of pedestrian was determined. To simulate a long vibration exposure, each test accommodated about 400 steps. Figure 12 shows that, when walking at frequency of 1.75 Hz or more, the pedestrians' sensitivity to vibration decreased with increase in the pacing rate. This might be due to two features of the locomotion at fast pacing rates: a decrease in foot's contact time with the structure and a larger BCoM movement which makes the structure movement less noticeable (Dang and Živanović, 2016). For slower than 1.75 Hz walking, the pacing rate is not as influential, perhaps because the contact time is sufficiently long for pedestrian to feel the vibration sensation. This initial study is significant as it shows that that if large structural vibration, say that exceeding the BSI (2008) limit of 2 m/s^2, occurs on a structure having natural frequency of, say 2.2 Hz, then the vibration levels should be interpreted in the following context: not only this vibration will be generated on a small number of occasions as only a small number of people walks this fast, but also these people are less likely to notice the vibrations, which is significant if they are main receivers of the vibrations in question. This example illustrates complexity of details intrinsic to evaluation of comfort of a structure. It also shows that pacing rate is another factor that plays a role in comfort evaluation by pedestrians, in addition to the previously accepted influence of vibration magnitude, frequency and exposure time. Note that in this case, fast walkers were shown to be at least two times less sensitive to vibration than the slow walkers. It is also interesting to note that the slow walkers were about 50 times less sensitive compared with people standing.

Another pilot study we conducted on the EB (that can be traversed in around 30 steps) suggests that the pedestrians could tolerate 3-4 times larger vibrations than those from the previous study. This reduced sensitivity due to shorter vibration exposure time is of direct interest to evaluation of short-span footbridges, mentioned earlier, that are currently required to satisfy the same vibration criteria as longer bridges (Uyttersprot and De Corte, 2021).

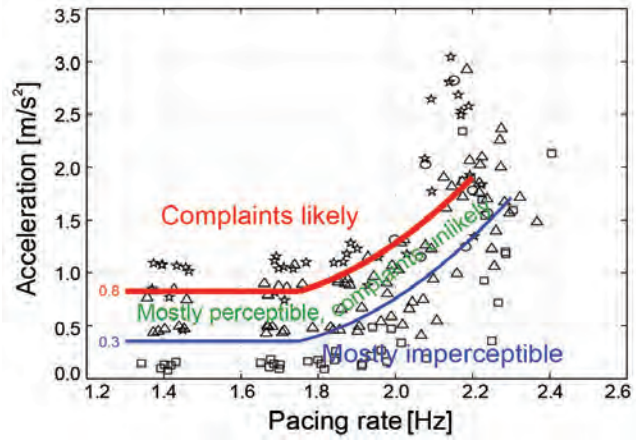


Figure 12. Pedestrian's rating of vibration as a function of pacing rate (Dang and Živanović, 2016)

Another focus of our experimental work is evaluating drivers of PSI in VSim. Human's bipedal locomotion is inherently unstable, and at the same time equipped by

means of controlling balance and adapting to various surface perturbations. There are strong indications that increase in vibration of the footbridge deck increase variability in walking locomotion parameters, such as the pacing rate, the attack angle, orientation of the trunk and step length (Dang and Živanović, 2016). In addition, we observed a reduction in the forcing harmonic responsible for causing resonance of the structure on two laboratory bridges (Ahmadi *et al.*, 2018; Dang and Živanović, 2016). However, only the recent opening of the new facilities provided an opportunity to initiate tests whereby the detailed kinematics of the human walking as well as the GRF and foot pressure distribution can be recorded. The tests that are under way are looking to identify which parameters provide best indicators of human's adaptation to the vibration and under which vibration conditions the adaptation occurs. The evidence of the adaptation will be sought by observing mechanisms for walking/balance control, including foot clearance, knee flexing, ankle rotation and pressure redistribution. In addition, energy expenditure will also be observed to test the hypothesis that the walker's priority on the vibrating surface remains minimisation of energy use.

Importance of these studies can be seen in the context of a recent survey by the UK Institution of Structural Engineers (Pavic, 2019). The survey found that more than 40% of respondents (structural engineers) stated that they encountered limitations in the current design code provisions with respect to VS assessment. While not all limitations might be due to poor assessment of WBV or lack of reliable modelling of PSI, there is no doubt that the contradictory vibration limits and neglecting of the interaction in the codes are significant contributors to low confidence in the current design practice. This situation suppresses innovation and leads to ineffective structural design solutions.

7. CONCLUSIONS

This paper provides an overview of key components that are essential for vibration serviceability assessment of structures, and a brief insight into their historical context.

It can be concluded that the humans using the structures are in the centre of the design process. Whilst we have always aimed to design the structure fit for purpose and acceptable by users, it is only now that this design requirement is evaluated in the context of climate change and requirement for minimising use of structural materials. Design solutions that are not only acceptable by humans but also kind to environment present a technical challenge, as they often require contradictory approaches to design. Technical resolution of this challenge relies on advancing our understanding of the human side of the equation as opposed to traditional approach of compensating for our lack of knowledge by overdesign.

In this context, the vibration serviceability assessment in the future is expected to incorporate sophisticated models for human adaptations and reaction to vibration on both individual and population levels. Key enablers for this research are state-of-the-art experimental facilities, such as those available at the University of Exeter. Some pointers for advancing both fundamental and applied side of research have been provided in this paper. It is hoped that their successful resolution will be followed by a systematic transfer of knowledge to industry and suitable upgrade of vibration serviceability guidelines.

ACKNOWLEDGEMENTS

The author acknowledges the support of numerous colleagues – past and present – who were involved in collaborations and research discussions. Special thanks go to colleagues at Monash, Warwick, Belgrade, Oviedo, Cambridge, Delft, and Exeter Universities, as well as Politecnico di Milano, and to former and current researchers in my group: Hiep, Madi, Justin, Xiaojun, Andre, Nimmy, Sigong and Bintian. Finally, thanks go to the UK Engineering and Physical Sciences Research Council whose financial support was crucial for development of some of the research presented (on projects: EP/M021505/1: Characterising dynamic performance of fibre reinforced polymer structures for resilience and sustainability and EP/I03839X/1: Pedestrian interaction with lively low-frequency structures), as well as EU funding on Horizon 2020 MSCA-IF project vPERFORM (grant ID 898216: Developing advanced vibration performance assessment for new generation of lightweight pedestrian structures using motion platform and virtual reality environments).

REFERENCES

Ahmadi, E., Caprani, C., Živanović, S. and Heidarpour, A. (2018). Vertical ground reaction forces on rigid and vibrating surfaces for vibration serviceability assessment of structures. *Engineering Structures*, *172*, 723–738.

Ahmadi, E., Caprani, C., Živanović, S. and Heidarpour, A. (2019). Assessment of human-structure interaction on a lively lightweight gfrp footbridge. *Engineering Structures*, *199*, 109687.

Brownjohn, J. M., Pavic, A. and Omenzetter, P. (2004). A spectral density approach for modelling continuous vertical forces on pedestrian structures due to walking. *Canadian Journal of Civil Engineering*, *31*(1), 65–77.

BSI. (1978). *Steel, concrete and composite bridges—part 2: Specification for loads; appendix c: Vibration serviceability requirements for foot and cycle track bridges, bs 5400.* UK: British Standards Association, London.

BSI. (2008). *Uk national annex to ec1: Actions on structures - pt. 2: Traffic loads on bridges. na to bs en 1991-2:2003.* UK, London.

Caprani, C. C. and Ahmadi, E. (2016). Formulation of human–structure interaction system models for vertical vibration. *Journal of Sound and Vibration*, *377*, 346–367.

Cavagna, G. A., Thys, H. and Zamboni, A. (1976). The sources of external work in level walking and running. *The Journal of physiology*, *262*(3), 639–657.

Chen, J., Xu, R. and Zhang, M. (2014). Acceleration response spectrum for predicting floor vibration due to occupant walking. *Journal of Sound and Vibration*, *333*(15), 3564–3579.

Dallard, P., Fitzpatrick, A., A. J. Flint, Le Bourva, S., Low, A., Ridsdill-Smith, R. and Willford, M. (2001). The london millennium footbridge. *The Structural Engineer*, *79*(22), 17–33.

Dang, H. V. (2014). *Experimental and numerical modelling of walking locomotion on vertically vibrating low-frequency structures* (PhD Thesis). School of Engineering, University of Warwick, UK.

Dang, H. V. and Živanović, S. (2016). Influence of low-frequency vertical vibration on walking locomotion. *Journal of Structural Engineering*, *142*(12), 04016120.

Dougill, J., JW Wright and Parkhouse, R., JG Harrison. (2006). Human structure interaction during rhythmic bobbing. *Structural Engineer*, *84*(22), 32–39.

Ellis, B., Ji, T. and BRE. (1997). Human-structure interaction in vertical vibrations. *Proceedings of the Institution of Civil Engineers-Structures and Buildings*, *122*(1), 1–9.

Gaspar, C., Caetano, E., Moutinho, C. and da Silva, J. G. S. (2020). Active human-structure interaction during jumping on floors. *Structural Control and Health Monitoring*, *27*(3), e2466.

Geyer, H. (2005). *Simple models of legged locomotion based on compliant limb behavior= grundmodelle pedaler lokomotion basierend auf nachgiebigem beinverhalten* (PhD Thesis). Friedrich-Schiller-Universitat Jena.

Griffin, M. (2012). *Handbook of human vibration*. Academic Press.

Hansen, A., Gard, S. and Childress, D. (2000). The determination of foot/ankle roll-over shape: clinical and research applications. In *Pediatric gait: a new millennium in clinical care and motion analysis technology* (pp. 159–165).

ISO. (2007). *Bases for design of structures—serviceability of buildings and walkways against vibration: Iso 10137*. International Organization for Standardization, Geneva, Switzerland.

IStructE. (2008). *Dynamic performance requirements for permanent grandstands subject to crowd action: Recommendations for management, design and assessment*. Dept. for Communities and Local Government and Dept. for Culture Media and Sport, London, UK.

Jones, C., Reynolds, P. and Pavic, A. (2011). Vibration serviceability of stadia structures subjected to dynamic crowd loads: A literature review. *Journal of Sound and Vibration*, *330*(8), 1531–1566.

Kerr, S. (1998). Human induced loading on staircases phd thesis.

Kim, S. and Park, S. (2011). Leg stiffness increases with speed to modulate gait frequency and propulsion energy. *Journal of biomechanics*, *44*(7), 1253–1258.

Lee, C. R. and Farley, C. T. (1998). Determinants of the center of mass trajectory in human walking and running. *The Journal of experimental biology*, *201*(21), 2935–2944.

Leonard, D. (1966). *Human tolerance levels for bridge vibrations, trrl report no. 34*. Road Research Laboratory.

Lin, B., Zhang, Q., Fan, F. and Shen, S. (2021). Reproducing vertical human walking loads on rigid level surfaces with a damped bipedal inverted pendulum. In *Structures* (Vol. 33, pp. 1789–1801).

Macdonald, J. H. (2009). Lateral excitation of bridges by balancing pedestrians. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *465*(2104), 1055–1073.

McDonald, M. G. and Živanović, S. (2017). Measuring ground reaction force and quantifying variability in jumping and bobbing actions. *Journal of Structural Engineering*, *143*(2), 04016161.

Pavic, A. (2019). Results of istructe 2015 survey of practitioners on vibration serviceability. In *Seced conference*. London, 9-10 September.

Perry, J. and Burnfield, J. M. (2010). Gait analysis. normal and pathological function 2nd ed. *California: Slack*.

Qin, J., Law, S., Yang, Q. and Yang, N. (2013). Pedestrian–bridge dynamic interaction, including human participation. *Journal of Sound and Vibration*, *332*(4), 1107–1124.

Racic, V. and Brownjohn, J. M. W. (2011). Stochastic model of near-periodic vertical loads due to humans walking. *Advanced Engineering Informatics*, *25*(2), 259–275.

Racic, V., Pavic, A. and Brownjohn, J. (2009). Experimental identification and analytical modelling of human walking forces: Literature review. *Journal of Sound and Vibration*, *326*(1-2), 1–49.

Ramírez, D. Z. M., Strike, S. and Lee, R. Y. (2013). Measurement of transmission of vibration through the human spine using skin-mounted inertial sensors. *Medical engineering & physics*, *35*(5), 690–695.

Sachse, R., Pavic, A. and Reynolds, P. (2004). Parametric study of modal properties of damped two-degree-of-freedom crowd–structure dynamic systems. *Journal of Sound and Vibration*, *274*(3-5), 461–480.

Saunders, J., Inman, V. T., Eberhart, H. D. et al. (1953). The major determinants in normal and pathological gait. *The Journal of Bone and Joint surgery. American Volume*, *35*(3), 543–558.

Schiavi, A., Rossi, L. and Ruatta, A. (2016). The perception of vibration in buildings: A historical literature review and some current progress. *Building Acoustics*, *23*(1), 59–70.

Sétra. (2006). *Footbridges, assessment of vibrational behaviour of footbridges under pedestrian loading, the technical department for transport*. Roads and Bridges Engineering and Road Safety, Paris, France.

Shahabpoor, E., Pavic, A. and Racic, V. (2016). Identification of mass–spring–damper model of walking humans. In *Structures* (Vol. 5, pp. 233–246).

Tubino, F., Pagnini, L. and Piccardo, G. (2020). Uncertainty propagation in the serviceability assessment of footbridges. *Structure and Infrastructure Engineering*, *16*(1), 123–137.

Uyttersprot, J. and De Corte, W. (2021). Measured dynamic properties of web-core sandwich panel frp composite footbridges and their relation to pedestrian comfort analysis. *Composite Structures*, *259*, 113236.

Van Nimmen, K., Lombaert, G., Jonkers, I., De Roeck, G. and Van den Broeck, P. (2014). Characterisation of walking loads by 3d inertial motion tracking. *Journal of Sound and Vibration*, *333*(20), 5212–5226.

Van Nimmen, K., Van den Broeck, P., Lombaert, G. and Tubino, F. (2020). Pedestrian-induced vibrations of footbridges: An extended spectral approach. *Journal of Bridge Engineering*, *25*(8), 04020058.

Wang, H., Chen, J. and Brownjohn, J. M. (2017). Parameter identification of pedestrian's spring-mass-damper model by ground reaction force records through a particle filter approach. *Journal of Sound and Vibration*, *411*, 409–421.

Zhang, M., Georgakis, C. T. and Chen, J. (2016). Biomechanically excited smd model of a walking pedestrian. *Journal of Bridge Engineering*, *21*(8), C4016003.

Zheng, X. and Brownjohn, J. M. (2001). Modeling and simulation of human-floor system under vertical vibration. In *Smart structures and materials 2001: Smart structures and integrated systems* (Vol. 4327, pp. 513–520).

Živanović, S., Pavic, A. and Reynolds, P. (2005). Vibration serviceability of footbridges under human-induced excitation: a literature review. *Journal of sound and vibration*, *279*(1-2), 1–74.

Živanović, S., Pavić, A. and Reynolds, P. (2007). Probability-based prediction of multi-mode vibration response to walking excitation. *Engineering structures*, *29*(6), 942–954.

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

624(082)

INTERNATIONAL Conference Civil Engineering 2021 - Achievements and Visions (2021 ; Beograd)
Proceedings of the International Conference Civil Engineering 2021 - Achievements and Visions, celebrating 175th Anniversary of the Faculty of Civil Engineering, University of Belgrade, October 25-26, 2021, Belgrade, Serbia / [organizer University of Belgrade Faculty of Civil Engineering] ; editors V. [Vladan] Kuzmanović, I. [Ivan] Ignjatović. - Belgrade : University, Faculty of Civil Engineering ; Belgrade : Academic Mind, 2021 (Belgrade : Planeta Print). - VIII, 195 str. : ilustr. ; 30 cm

Tekst štampan dvostubačno. - Tiraž 500. - Str. III-V: Preface / Editors = Предговор / уредници. - Biografske beleške o autorima uz svaki rad. - Bibliografija uz svaki rad.

ISBN 978-86-7466-895-5 (AM)

а) Грађевинарство -- Зборници

COBISS.SR-ID 47441417

www.akademska-misao.rs
office@akademska-misao.rs