Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 364-376, 2021
DOI: 10.14513/actatechjaur.00597

The impact of marketing strategy on logistics decisions and the implementation of JIT

Eszter Sós*, Péter Földesi

Széchenyi István University, Department of Logistics and Forwarding
H-9026 Győr, Egyetem tér 1.
*e-mail: sos.eszter@sze.hu

Submitted: 12/02/2021; Accepted: 30/04/2021; Published online: 03/05/2021

Abstract: Marketing has a significant impact on logistics systems and corporate performance, so it is nowadays necessary to define a marketing logistics strategy in the lives of larger companies so that logistics decisions are not distorted by the marketing strategy. According to the current direction of development, smaller and smaller quantities of goods are constantly being moved, with ever shorter deadlines. Therefore, companies prefer to use the Just In Time (JIT) system, which results in a gradual reduction in inventory costs, but increases the burden on the environment. In this article, we analyze the interaction between marketing and logistics and the logistics environment required to implement a JIT system. For research we have used the matrix of Quality Function Deployment (QFD) technique to evaluate the needs of the customer with respect to the JIT system, thus clarifying the logistics strategy applicable to the introduction of the specific product.

Keywords: *logistics; marketing strategy; JIT; QFD*

1. Introduction

The organizational structure of companies has undergone significant changes in recent decades due to the competitive situation in the market. Meeting customer expectations at the highest possible level has come to the fore, and marketing has gained more and more ground. Customer service also requires a well-established logistics system, so it is necessary to define a marketing logistics strategy when developing corporate strategies. By applying it, it is possible to avoid that the

developed marketing concepts distort logistics decisions and to introduce a logistics system for which the company's logistics environment is not yet prepared.

The aim of our article is to describe the importance of marketing-logistics strategy to meet customer expectations and to conduct an examination of the logistics environment required to implement the JIT system, using the QFD technique matrix. The intent of the QFD technique is to meet the highest possible level of consumer needs through the design of manufacture and production design processes developed by engineers, taking into account customer expectations. [1]

The QFD method can be effectively applied in the field of logistics. It can be used to identify the logistics sub-processes that help to meet the primary customer needs.

2. The role of logistics and marketing in the sales system

We will describe the interaction between logistics and marketing through the sales system, as the efficiency of a company's logistics system is reflected in the quality of customer service [2].

Objectives to be defined during the examination of the operation of the sales systems: examination of the market coverage and development of the sales system taking into account the characteristics of the product, as well as ensuring the quality of service and achieving profitability. Of the operational goals listed, logistics ensures the right level of service and partly profitability, while marketing covers the rest [3].

The efficiency of logistics services has an impact through the quality of customer service, as the product is sold to the user together with the contributing services. Therefore, it is important that the logistics concept is completed as soon as the marketing schedule and strategy related to the product is defined. To meet the highest possible level of customer needs, it requires a unified management of logistics and marketing, as proper customer service includes optimal delivery time, quality and flexible delivery, as well as delivery skills and the ability to respond appropriately to suddenly increased demands [4]. Logistics is a means to achieve marketing targets, as marketing stimulates demand and transmits market impulses to production; logistics executes demand and implements production-related flow processes [5].

From the point of view of both logistics and marketing, the right IT background is essential for a company, as this will allow a faster and more accurate information flow. This will reduce uncertainty and thus reduce safety stocks, while keeping the proportion of inaccurate, defective and non-scheduled deliveries to a minimum [6].

The information system that supports the operation of the company must provide data from the inside - about the needs of the company and the possibilities of substitution, and from the outside - about the needs of the market, suppliers, and our reliability and costs.

3. Examining the conditions necessary for the development of a logistics strategy

An important part of the corporate strategy is the logistics strategy, the main content elements of which are the purchasing and inventory strategy, as well as the physical distribution, information flow and waste management logistics strategy, but nowadays we also list the marketing logistics strategy. From the point of view of marketing - logistics, the following general requirements and objectives related to purchasing and sales can be defined during the development of the strategy. Purchasing - supply side requires short delivery times, flexible supplier relationships, and low inventory management costs. On the sales - distribution side, in addition to short delivery times, adequate delivery quality and readiness, the reliability of delivery and the ability to respond flexibly to unforeseen situations are of paramount importance [7].

In many cases, marketing strategy distorts logistics decisions, which is why the Just In Time system is often used even where the corporate environment is not properly prepared yet. For this reason, it is essential to examine the logistics environment, order sizes, the need for JIT, and the environmental burdens it causes before implementing the Just In Time system.

4. The Just In Time system production philosophy

Just In Time is a general philosophy, a principle of production, organization and management that is often and erroneously used as a synonym for out-of-stock production. JIT is a pull system, i.e. the production of the product, the flow of material can only start on a specific order or instruction, and so the material is actually just in time, it is not waiting in the system, as it is already waiting for it. Its operations fight for cost reduction, service and quality improvement through loss reduction, employee involvement, and continuous process improvement [8].

The importance of Just In Time today lies in speeding up and simplifying the flow of material and information, reducing costs, minimizing inventory levels, and helping to meet consumer needs. The JIT system seeks to ensure that the raw materials, semi-finished and finished products required for production are available in just the right place, in the right quality and in the right quantity to serve efficient and smooth production, preferably in such a way that their “destination” is as short

as possible, and be the most cost effective. For the sake of this cost-effectiveness, they do not store only what is absolutely necessary, so stocks of raw materials, semi-finished and finished products are kept to a minimum and more frequent deliveries are used [9].

4.1. Examination of possible application of Just In Time

The appropriate conditions for the applicability of JIT must be examined in all cases, because not all companies are suitable for the introduction of JIT. The first step is to explain the tasks to the management and monitor what obstacles are present in the corporate environment (Table 1).

Table 1. Examination of possible application of Just In Time

Key responsibilities of management when implementing JIT	***Barriers to JIT implementation***
You need to be determined to eliminate losses in the processes.	Lack of commitment from top management.
You need to be committed to innovation.	Lack of plans to obtain support.
A flexible production system needs to be set up with spare capacities.	
A well-motivated and flexible workforce is needed.	Bad attitude of employees, lack of support.
In quality, a "zero error rate" approach should be pursued.	Poor product quality.
Suitable suppliers must be selected.	Reluctant cooperation and willingness of the supplier to cooperate.
A long-term relationship with suppliers needs to be established.	
You need to be prepared for smaller and more frequent deliveries.	
Reliable delivery times are required.	
Establishing a close communication relationship with suppliers.	Weaknesses in the supply chain communication system.

4.2. The basic elements of Just In Time implementation

Management and employees can be made fit to implement JIT through training, but the proper design of physical conditions is also at least as important as improving

the thinking of the workforce at the system level. The adaptation and preparation of the work environment begins with an examination of the logistics environment, which includes an examination of the need for a JIT.

We would like to describe the basic elements needed to build Just In Time (Table 2), the existence of which is worth reviewing before introducing the system.

Table 2. The basic elements of Just In Time implementation

Basic elements needed to design a Just In Time system	
Puller system	The manufacturer starts production of the product to a specific order according to the user's needs.
Constant production volume	Constant flow of material between operating sites, accurate timing of all activities.
Low stock	JIT is created as a result. Advantages: less storage space and tied up capital.
Small items	It reduces inventory costs and space requirements. If there is a quality problem, fewer products need to be repaired or reworked.
Fast, low-cost change-over	Small production batches assume frequent switching, which should be cheap and feasible. Conversion losses can be reduced by having similar products follow each other. For this reason, proper involvement of the workforce is necessary.
Operation order suitable for the purpose	JIT systems are product-oriented, mostly based on material flow layout.
Effective preventive maintenance	Applying a preventive and forward-looking maintenance method is important because a machine failure can ruin the production schedule.
Multidisciplinary, multi-tasking staff	Those working with the JIT system should be able to diagnose problems, eliminate minor faults, and be proficient in work at other workplaces.
Ability to work together to solve problems	A high degree of cooperation between employees is essential for the effective operation of JIT, and special attention must be paid to developing this skill before introducing a JIT system.
Continuous improvement	Directions: inventory reduction, mitigation of transition times and their costs, improvement of quantity and quality, depreciation of losses.

It can be clarified from Table 2 that the examination of the logistics environment covers all areas of the company. I present in Table 3 the supplier expectations outside the company, which play an important role in the implementation of JIT.

Table 3. Requirements from the supplier

Supplier expectations for the development of Just In Time	
Reliable suppliers	In a JIT system, a good supplier-customer relationship is very important. Expectations of suppliers are: consistently high quality, flexibility, frequency, delivery in small batches and quick resolution of any problems encountered.
High quality level	The number of incoming materials without a given quantity, without "retention", so a high level of quality is both a prerequisite and a result of JIT operation. In case of a defective piece, identification and replacement are important.
Strong IT connection, communication between buyer and supplier	The management systems of the two organizations must work closely together.

In the Just In Time system reliability is a basic requirement, as it does not include a security set, in addition to this essential condition for its implementation is an up-to-date knowledge of consumer needs and the construction of the necessary IT network [10].

5. Impact of marketing strategy on the JIT system

Whether the Just In Time system should be served or whether the company has a traditional supplier role has a significant influence on the design side of the marketing logistics strategy.

Examining the management competencies required to implement JIT and the basic elements identified for implementing the system described in the previous chapter is essential when developing a marketing strategy, as a poorly chosen marketing strategy can easily distort logistics decisions and implement a JIT system that is not suitable. When designing the distribution side of a marketing-logistics strategy, it is necessary to examine the characteristics of the logistics environment that will help select the right logistics system for the company. [11]

It is necessary to know what principles and goals should be applied during the development of the distribution logistics system. To do this, the company's

management must determine the criteria on the basis of which the finished product storage system and the unit load training applied to the finished product structure should be selected.

The direct delivery of the product from the finished goods warehouse, the means of warehousing technology and the picking method used also have a significant influence on the development of the logistics environment. In order to determine the range of potential suppliers, it is necessary to know what modes of transport and priorities are used.

In terms of product deliverability, it is necessary to examine for which products it is possible to deliver the finished product directly from the end of the production lines, and for which products it is advisable to use distribution warehouses during the distribution logistics activity. In order to carry out the distribution activity, an IT and information system must be set up that adequately supports its product identification.

It must be determined which management-, leadership- and controlling system must be developed for the distribution activity, and it is also necessary to establish a quality assurance system applied during the distribution logistics process.

Examining the criteria above, the answers are clearly provided by the Just In Time system, as a regulated environment is essential for its proper functioning. But what happens when we don't take into account the specific needs of JIT while developing a marketing strategy?

On the one hand, it distorts logistics decisions, so JIT can be introduced in a company where the company's logistics environment is not suitable for it yet. On the other hand, the determination of order batch sizes is not adapted to the scheduled material supply of the logistics system, so marketing may be successful, but the logistics system (production, delivery, etc.) is not prepared for the order quantity, so the customer has to wait for the product, which in many cases can lead to customer dissatisfaction. In addition, increased demand may necessitate unplanned deliveries, which cause a high environmental burden.

The Just In Time system is only viable with a certain level of stability of supply and demand [12], so efforts should be made to consider the logistics environment from both the supplier and the sales company when making marketing logistics decisions. On the one hand, this is important for the proper and cost-effective operation of the JIT system, and on the other hand, a malfunctioning JIT system generates occasional deliveries, which leads to environmental damage.

6. Environmental loads caused by JIT

Most of the literature deals with the benefits of the Just In Time system, the environmental effects it causes are not described. In 2007, the European Parliament already addressed the issue of freight transport in its report on trade and climate change (2007/2003 (INI)), with particular reference to the negative effects of the Just In Time system on the environment [13].

The Just In Time system is designed to eliminate stock, but due to frequent deliveries, a significant increase in shipping volume is expected. In addition, it is important that the transport of goods is carried out with sufficient flexibility and precision, so that the movement of goods is mostly carried out by road, which in turn causes a significant environmental burden in terms of pollutants, toxic gases and noise [14].

In the European Union the road transport is responsible for 72% of CO^2 emissions from transport, of which 38.1% can be related to freight transport [15], so one of the important logistical factors in terms of environmental impact is the emission of pollutants from material handling activities [16]. Rail freight emits about 65% less CO_2 than road freight transport [17], but rail freight is not always feasible due to track constraints and unreliability to serve the strict schedule of the JIT system.

Due to frequent deliveries, the amount of packaging materials can also increase, as units are usually set up as required by the company and quality control is also done before delivery. Therefore, the parts are delivered to the user company in such a way that their inspection and handling takes as little time as possible, so in order to protect the quality of the goods, more packaging material is used than in preparation for delivery. In addition, many suppliers do not invest energy or time and money in the development and use of multi-way packaging systems, but use cheaper and logistically easier-to-use one-way packaging materials, which become waste immediately upon arrival and therefore have a significant impact on the environment.

7. Examining customer needs for Just in Time using QFD technique

Quality Function Deployment (QFD) is a technique used by service providers to gather requirements, expectations, and customer purchasing decision factors [18]. A practical formalism of transforming consumer needs into technical, quality features is the House of Quality – (HOQ) [19].

In this article we use the matrix that forms the basis of the QFD technique, which gives the “trunk” of the House of Quality (HOQ). We chose this technique because,

with the help of QFD, consumer needs can be compared with logistics processes, thus making it possible to meet the needs of consumers at the highest possible level [20].

In addition to consumer needs, the elements of the 6R principle (The 6 Rights of Logistics) used in logistics were included in the matrix, and the basic elements necessary for the development of the JIT system were contained as technical characteristics.

			Essential elements for implementing JIT system										
			Constant production volume	Keep inventory levels low	Low inventory cost due to frequent deliveries	Fast, low-cost migration	Material flow operation	Effective preventive maintenance	Multitasking staff	High quality level	Ability to work together to solve problems	Reliable suppliers	Continuous improvement	Strong IT connection
Customer expectations	**Product requirements**	Low lead time	++	+						-				+
		Low price	+	-	+					-				
		Short delivery time		+	+						+	++	+	
		Reliable, long-lasting product						+		++			+	
		The stock warehouse should be close			+							++		++
		Stock to suit customer needs	+	-		+								
		Constant supply (even production)	++	-		+	+			+				++
		Complete built-in tracking								++				++
		Great variance	-						++					+
	Requirements of logistics services (6R)	The right product						+		+			+	+
		In the right quantity	+										+	+
		In the right quality				+		+	+	++	+		+	
		At the right time			+		+				+	+	+	+
		To the right place					+				+	+	+	
		For the right cost	+	++	+	+			+			+	+	

Legend	
++	Significant positive effect
+	Moderate positive effect
-	Can cause a negative effect

Figure 1. Examination of customer expectations for JIT using QFD technique

In the intersections of the present matrix, we used the symbols according to the relevance of the given relationship. In the legend, we clarified what meanings we associated with the different symbols. During the analysis according to the whole QFD technique, the symbols are replaced by quantified data, so that the values can be weighted with Fuzzy later. In this way, the customer's needs for the JIT system in relation to the product and logistics services can be evaluated.

The developed system can be used primarily in the commercial field, as it helps to develop the logistics strategy of the products purchased as a result of the marketing strategy. However, it can also be used in an industrial environment for any product that is made to meet specific customer needs (e.g. a custom-made car).

It can be used as a decision tool in all areas where strategic logistics development is required. By using it, the logistics processes become clearer and the goals set by the customer are easier to achieve. Existing logistics strategies (e.g. Push and Pull logistics strategy) can also be reviewed with it.

By applying the OFD technique in the field of logistics, an analysis can be made of the already developed logistics strategies, which also points out the shortcomings of the logistics environment.

8. Conclusion

One of the important logistics problems today is the scheduling of deliveries. Companies strive to keep their inventories low, resulting in lower inventory management costs and tied-up capital.

For this reason, the pursuit of a Just In Time system is popular, but in many cases it is also introduced by companies whose logistics environment is not suitable for using the system. Prior to implementation, it would be necessary to coordinate marketing strategy decisions with the logistics environment, as a poorly chosen marketing strategy decision causes significant negative effects on logistics processes, resulting in the customer not receiving the purchased goods on time. The efficiency of a company’s logistics system is reflected in the quality of customer service, so a successful Just In Time system causes satisfaction for both the supplier and the customer.

In order to determine the appropriate logistics strategy, we applied the procedure and matrix of the QFD technique. During the evaluation of the matrix, we examined the customer needs in terms of the basic conditions required for the development of the JIT system, thus clarifying the applicability of the JIT system when introducing the given product.

References

[1] Y. Akao, QFD: Past, Present, and Future, in: 3rd International Symposium on Quality Function Deployment (ISQFD'97), 1997, Linköping, pp. 1-12.
URL
http://www.las.inpe.br/~perondi/19.10.2009/Akao_1997_QFD_History.pdf

[2] V. E. Bottani, A. Rizzi, Strategic management of logistics service: A fuzzy QFD approach, *International Journal of Production Economics* 103 (2) (2006) pp. 585-599.
doi: *https://doi.org/10.1016/j.ijpe.2005.11.006*

[3] M. S. Akdogan, A. Durak: Logistic and marketing performances of logistics companies: A comparison between Germany and Turkey, *Procedia - Social and Behavioral Sciences* 235 (2016) pp. 576 -586.
doi: *https://doi.org/10.1016/j.sbspro.2016.11.084*

[4] A. E. Ellinger: Improving Marketing/Logistics Cross-Functional Collaboration in the Supply Chain, *Industrial Marketing Management* 29 (1) (2000) pp. 85-96.
doi: *https://doi.org/10.1016/S0019-8501(99)00114-5*

[5] P. V. Hong, T.-T. Nguyen: Factors affecting marketing strategy of logistics business – Case of Vietnam, *The Asian Journal of Shipping and Logistics* 36 (4) (2020) pp. 224-234.
doi: *https://doi.org/10.1016/j.ajsl.2020.03.004*

[6] A. P. Barbosa-Povoa, J. M. Pinto: Process supply chains: Perspectives from academia and industry, *Computers & Chemical Engineering* 132 (2020) 106606.
doi: *https://doi.org/10.1016/j.compchemeng.2019.106606*

[7] J. Salkovska, N. R. V.Danovics: Marketing and Logistics Cooperation Problems in Latvian Companies, *Procedia - Social and Behavioral Sciences* 110 (2014) pp. 390-397.
doi: *https://doi.org/10.1016/j.sbspro.2013.12.883*

[8] T. C.E. Cheng, S. Podolsky: Just-In-Time Manufacturing: An Introduction, Chapman & Hall Second Edition (1996) ISBN: 0-412-73540-7.

[9] K.-H. Lai, T. C.E. Cheng: Just-In-Time Logistics, MPG Books Ltd. (2009).

[10] R. Z. Farahani, S. Rezapour, L. Kardar: Logistics operations and management: Concepts and models, *Elsevier* (2011).
doi: https://doi.org/10.1016/C2010-0-67008-8

[11] U. Y. Alvarado, H. Kotzab: Supply Chain Management: The Integration of Logistics in Marketing, *Industrial Marketing Management* 30 (2) (2001) pp. 183-198.
doi: https://doi.org/10.1016/S0019-8501(00)00142-5

[12] C. H. Pragman: JIT II: A purchasing concept for reducing lead times in time-based competition, *Business Horizons* 39 (4) (1996) pp. 54-58.
doi: https://doi.org/10.1016/S0007-6813(96)90052-X

[13] European Parliament, Trade and climate change 2007/2003(INI), (2007).
URL
https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+REPORT+A6-2007-0409+0+DOC+XML+V0//HU

[14] A. Memari, A. Rahman, A. Rahim, N. Absi, R. Ahmad, A. Hassan: Carbon-capped Distribution Planning: A JIT Perspective, *Computers & Industrial Engineering* 97 (2016) pp. 111-127.
doi: https://doi.org/10.1016/j.cie.2016.04.015

[15] European Environment Agency (2016).
URL www.eea.europa.eu/hu

[16] Z. Chena, B. Bidanda: Sustainable manufacturing production-inventory decision of multiple factories with JIT logistics, component recovery and emission control, *Transportation Research Part E* 128 (2019) pp. 356-383.
doi: https://doi.org/10.1016/j.tre.2019.06.013

[17] Guidelines for Measuring and Managing CO2 Emission from Freight Transport Operations (2020).
URL
https://www.ecta.com/resources/Documents/Best%20Practices%20Guidelines/guideline_for_measuring_and_managing_co2.pdf

[18] M. M. H. Chowdhury, M. A. Quaddus: A multi-phased QFD based optimization approach to sustainable service design, *International Journal of Production Economics* 171 (2) (2016) pp. 165-178.
doi: https://doi.org/10.1016/j.ijpe.2015.09.023

[19] D. R. Kiran: Quality Function Deployment, *Total Quality Management, Key Concepts and Case Studies* (2017) pp. 425-437.
doi: https://doi.org/10.1016/B978-0-12-811035-5.00030-1

[20] A. M. Oddershede, L. E. Quezada, J. E. Valenzuela, P. I. Palominos, H. Lopez-Ospina: Formulation of a Manufacturing Strategy Using the House of Quality, *Procedia Manufacturing* 39 (2019) pp. 843-850.
doi: https://doi.org/10.1016/j.promfg.2020.01.417

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 377-405, 2021
DOI: 10.14513/actatechjaur.00605

Analysis of vehicle-pedestrian and bicyclist conflicts in Győr-Hungary using Swedish conflict technique

A. Kizawi*

Széchenyi István University, Department of Transport Infrastructure and Water Resources Engineering, Egyetem tér 1, 9026 Győr, Hungary
***e-mail: ahmad.kizawi.24@gmail.com**

Submitted: 15/03/2021; Accepted: 26/04/2021; Published online: 18/05/2021

Abstract: Increasing traffic volumes leads to changes in traffic conditions, especially at intersections. These changes affect the ability of Vulnerable Road Users (VRUs) such as pedestrians and bicyclists to cross the road safely, especially at un-signalized crossings (without traffic lights), where many road users are conflicting at the same point of the road and neglecting 'in some cases' the priority issue. Although the emphasis on pedestrian safety has recently increased, there is still a need to analyze the causal factors of VRUs accidents and define their relationship to road design characteristics. This paper presents a study bout vehicle – pedestrian and bicyclist conflicts analysis using Swedish conflict technique at three un-signalized pedestrian crossings in the city of Győr- Hungary, where some pedestrian accidents were happened between 2014 and 2018, and reported based on accidents database in Hungary. The aim is to analyse vehicle-pedestrian and bicyclist interactions, and evaluate the severity of conflicts with the help of the Swedish technique graph. The findings concluded that there is a compatibility between Swedish conflict technique and accident records with regard to conflicts severity and the probability of accidents occurrence.

Keywords: *traffic conflict; Swedish conflict technique; Vulnerable Road Users (VRUs) safety*

1. Introduction

Road accidents is a major cause of injury and death, especially to pedestrians. Not all pedestrians have the same ability to cross the road easily. Many have some restrictions that require special attention on their part and/or modifications of roadway infrastructure and operations to accommodate their needs. [1] The un-signalized intersection that is not controlled by traffic lights is a popular situation in any road network, as it is used by two or more traffic flows, which have different priorities. At un-signalized intersections, the absence of traffic lights may disregard, in some cases, the priority of vulnerable road-users to cross, such as pedestrians and bicyclists, or restrict their ability to pass safely. [2]

The emphasis on the safety of VRUs has recently received more interest. Since the 1960's, several traffic conflict studies have been accomplished aiming to develop new methods to evaluate road safety issues in different road geometric and operating conditions. [3] Traffic Conflict Techniques (TCTs) can be used as alternative or complementary tools for the use of accident records to verify traffic conflicts and estimate accidents probability [4]. Traffic conflict indicators are applicable to assess road safety performance thanks to their ability to capture conflict data, including the severity of conflicts, in an objective way within a shorter time compared to accident records.

One of the most common conflict techniques is the Swedish technique; it is based on observing the traffic conflicts between two road-users, with their evasive actions taken to avoid collision. The main characteristics of this technique are related to several points such as the requirement for a collision course in a conflict; the conflict severity based on the onset of an evasive action; and the distinction between severe and non-severe conflicts. The serious conflicts were found to be an indicator of a breakdown in the interaction – similar to a breakdown preceding an accident. [5]

This technique aims to analyze traffic conflicts on a given road, where the probability of accident occurrence is highly expected. This technical approach provides a better insight into conflict characteristics and helps in estimating the severity of conflicts and providing a possibility to suggest accidents countermeasures to improve the safety of VRUs.

2. Problem statement

Many pedestrians (especially from children and elderly people) find it difficult and unsafe to cross at un-signalized crossings that are not controlled by traffic lights. [6]. At pedestrian crossings that are not controlled by pedestrian traffic lights (un-signalized crosswalks), the priority of the pedestrian is not clear, even with legislation supporting pedestrian priority, and this results in higher levels of accident and death of pedestrians compared to those controlled by pedestrian traffic lights. Moreover, pedestrians often tend to make more sudden decisions and more prone to hectically change their speed and trajectory. Nowadays, it is obvious that the traffic conflict studies at un-signalized intersections are quite necessary. Many researchers have focused on this approach aiming at improving this situation. [7]. An accident analysis has been accomplished in order to determine the main causes of pedestrian related accidents based on the records of WEB-BAL (Hungarian accident database) from year 2018 [8]. Fig. 1 shows the results of the analysis. In 63% of all pedestrian related accidents, drivers are at fault while in 37% pedestrians mistaken. The highest share of causes is yield failure, but speeding and turning failure are also responsible for a high proportion of the pedestrian related crashes.

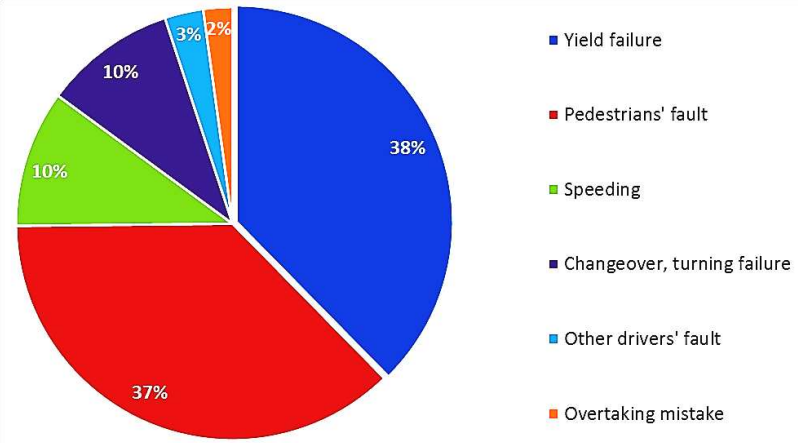


Figure 1. Causes of pedestrian accidents in Hungary, 2018 (based on [8])

Fig. 2 shows the results of the accident type analysis in Hungary 2018. Pedestrian accidents occurring at designated pedestrian crossings in intersection have the highest share among the accident types. (29% of accidents are not classified.)

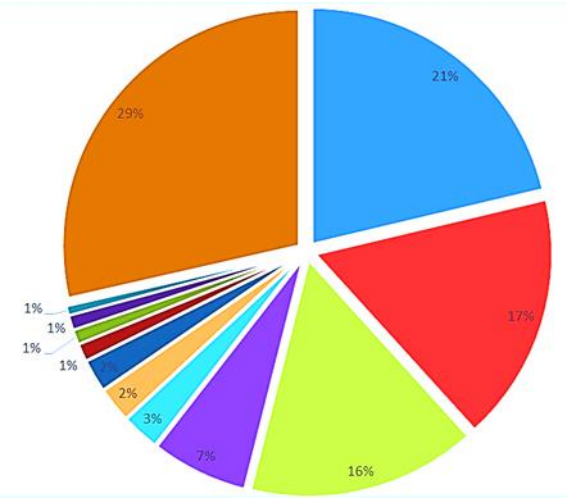


Figure 2. Pedestrian accidents according to accident type in Hungary, 2018 (based on [8])

3. Literature review

The idea of the conflict had been around for many years before it was first carried out systematically in the traffic safety community in the late 1960s. The first practical application of a traffic conflict technique was accomplished in the late 1960s by a team of researchers at General Motors Corporation Perkins and Harris [9]. Since that time, traffic conflict techniques have been developed and used more frequently.

The first study to develop a measure by which accidents can be predicted was designed by Perkins and Harris [9]. More than twenty types of potential traffic conflicts were identified and defined by the occurrence of evasive actions, such as braking or sudden change the path. Since that time, and for many objectives, traffic conflict techniques have been developed in many countries, although in most cases the emphasis is still on the diagnosis of safety problems. [4]

The common governing indicator for all methods addressed of most literatures regarding traffic conflicts is "Time-to-Collision"(TTC), which defined as time remaining to collide for two conflicting road users if they continue on the same trajectory and at their present speed [4].

In the 1970's and 1980's, the traffic conflict research noted significant development at the Technical University of Lund, Sweden. The method of traffic conflict observation is carried out by trained observers focused on situations happen on the road, in which two road users would collide if neither of them made some kind of an evasive action, such as stopping, braking, or changing the lane. The point, in which such evasive action is made, is recorded personally by the observer as the "Time-to–Accident" (TA). The TA value, the speed of conflict, and the distance between the two conflicting road users are used to determine whether the conflict is severe or not using a graph, which has disaggregating lines to categories the different conflicts according to severity of the conflict. According to the country of origin, this method has been called the "Swedish Traffic Conflict Technique" [10]. Reliability tests indicate very strongly that observers record conflicts in a reliable way after some days of training. The results from validation studies indicate that this technique, based on a criterion that the degree of seriousness should be taken into account, shows a strong correlation between conflicts and accidents. [11].

According to the literature review provided by A. Kizawi et al. on the conflict analysis of vehicle-pedestrian interactions, the frequent interactions between vehicles and pedestrians deserve special interest to analyze safety especially at intersections. As an alternative to traffic safety analysis based on historical crash data the use of non-crash events is becoming more popular. Various researchers have developed a number of safety indicators (e.g. Time to Accident (TA), which is

related to one of the most commonly used Surrogate Measures of Safety "Time-To Collision", where TA is calculated at the moment in which an involved road user makes an evasive action). There is a consensus among researches that observable non-crash events can be useful for traffic safety assessment as a substitutional tool in parallel with analysis based on crash data [12].

Almqvist and Hydén [13] [14] applied proximity based technique i.e. Swedish Traffic Conflict Technique (conflict observation using Time to Accident) in the city of Cochabamba in Bolivia with a view to guiding a method to assess the safety problem in developing countries. The study indicated that the technique is useful in its present form for this condition.

Tiwari et al. [15] evaluated conflicts at 14 locations in Delhi, in a heterogeneous traffic environment. The conflict was evaluated using the concept of Time-to-accident (TA). This study recorded seven types of conflicts occurring at mid-block in heterogeneous traffic, such as head-on, rear-end, sideswipe, change direction, fixed object, angle and traverse angle.

4. Conflict observation using Swedish technique

According to the Swedish Traffic Conflict Technique STCT, a collision course is a necessary condition for a traffic conflict. Collision course implies that, unless one of the road users takes an evasive action, a collision will occur. The severity of a conflict can be defined at the moment in which one of the road users takes an evasive action such as hard braking, sudden stopping, or sudden changing the lane. The relevant road user is the name of the road user who firstly takes an evasive action to avoid the collision with the other user. [10]

The conflict severity is based on two indicators: [10]

1. Time-to-Accident (TA): it is the time remaining to a collision when the relevant road user takes an evasive action. Lower TA values indicates that the traffic conflict is nearer to a collision, this means the situation is more dangerous.
2. Conflicting Speed (CS): it is the speed of the relevant road user when taking the evasive action to avoid the collision with the other road user. Higher CS values lead to a more severe conflict.

In practice, when this method is carried out at a given location personally, it would be easy to estimate the distance from the road user to the collision point rather than the time remaining to the collision. In this method, we estimate the distance between the two conflicting road users and the speed of the relevant road user. After that, the conversion table (see Table 1) can be used to calculate time to accident (TA) based on the distance to the collision point and the speed of the relevant road user.

Table 1. Time-to-accident values based on conflict speed and distance (based on [10])

100	95	90	85	80	75	70	65	60	55	50	45	40	35	30	25	20	15	10	5	km/h	***Speed***
27.8	26.4	25.0	23.6	22.2	20.8	19.4	18.1	16.7	15.3	13.9	12.5	11.1	9.7	8.3	6.9	5.6	4.2	2.8	1.4	m/s	
												0.0	0.1	0.1	0.1	0.1	0.1	0.2	0.4	0.5	***Distance [m]***
0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.4	0.7	1	
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.3	0.4	0.5	0.7	1.4	2	
0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.4	0.4	0.5	0.7	1.1	2.2	3	
0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.5	0.6	0.7	1.0	1.4	2.9	4	
0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.4	0.4	0.5	0.5	0.6	0.7	0.9	1.2	1.8	3.6	5	
0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.5	0.5	0.6	0.7	0.9	1.1	1.4	2.2	4.3	6	
0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.5	0.5	0.6	0.6	0.7	0.8	1.0	1.3	1.7	2.5	5.0	7	
0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.6	0.6	0.7	0.8	1.0	1.2	1.4	1.9	2.9	5.8	8	
0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.7	0.8	0.9	1.1	1.3	1.6	2.2	3.2	6.5	9	
0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.9	1.0	1.2	1.4	1.8	2.4	3.6	7.2	10	
0.5	0.6	0.6	0.6	0.7	0.7	0.8	0.8	0.9	1.0	1.1	1.2	1.4	1.5	1.8	2.2	2.7	3.6	5.4		15	
0.7	0.8	0.8	0.8	0.9	1.0	1.0	1.1	1.2	1.3	1.4	1.6	1.8	2.1	2.4	2.9	3.6	4.8	7.2		20	
0.9	0.9	1.0	1.1	1.1	1.2	1.3	1.4	1.5	1.6	1.8	2.0	2.3	2.6	3.0	3.6	4.5	6.0	9.0		25	
1.1	1.1	1.2	1.3	1.4	1.4	1.5	1.7	1.8	2.0	2.2	2.4	2.7	3.1	3.6	4.3	5.4	7.2			30	
1.3	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.1	2.3	2.5	2.8	3.2	3.6	4.2	5.0	6.3	8.4			35	
1.4	1.5	1.6	1.7	1.8	1.9	2.1	2.2	2.4	2.6	2.9	3.2	3.6	4.1	4.8	5.8	7.2	9.6			40	
1.6	1.7	1.8	1.9	2.0	2.2	2.3	2.5	2.7	2.9	3.2	3.6	4.1	4.6	5.4	6.5	8.1				45	
1.8	1.9	2.0	2.1	2.3	2.4	2.6	2.8	3.0	3.3	3.6	4.0	4.5	5.1	6.0	7.2	9.0				50	
2.0	2.1	2.2	2.3	2.5	2.6	2.8	3.0	3.3	3.6	4.0	4.4	5.0	5.7	6.6	7.9	9.9				55	

The graph shown in Fig. 3 can be used to define the severity level of the traffic conflict based on the TA and CS values. [10] According to the graph, the conflicts with severity level above 26 (red line on the graph) are ranked as serious, while the conflicts with severity under 26 are ranked as non-serious conflicts. The serious conflicts have a strong statistical relation with police-reported accidents and even can be converted into expected number of accidents with a reasonable accuracy. [10]

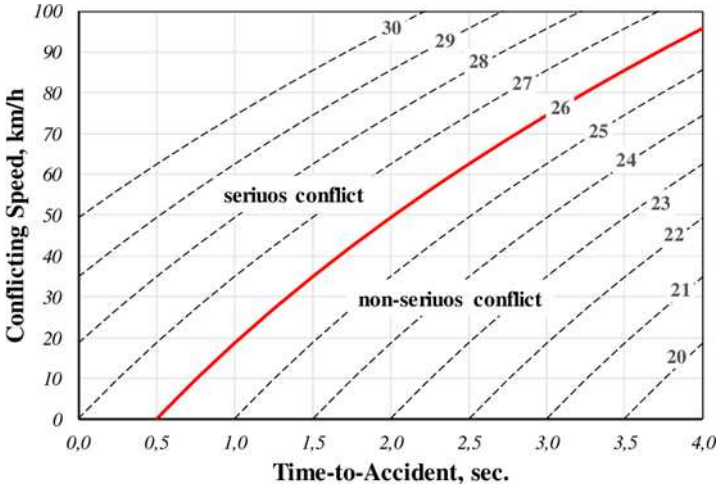


Figure 3. Swedish technique graph [10]

5. Methodology

5.1. Sites selection

The pedestrian accident data that observed in the city of Győr between 2014 and 2018 have been signed on a map using google map service [16] as shown in Fig 4. This map contains locations where accidents were occurred and reported with details; (accident location, accident date, causes of the accident and other relevant information). The accidents were distinguished by colours according to their type as follows:

- Pedestrian accident, crossing in front of / behind a standing vehicle or other obstacle on the right hand side of the road
- Pedestrian accident, vehicle hitting the pedestrian outside the road on the LEFT hand side (e.g. on the sidewalk)
- Pedestrian accident, vehicle hitting the pedestrian outside the road on the RIGHT hand side (e.g. on the sidewalk)

- Pedestrian accident, vehicle hitting the pedestrian at a bus (or other PT) stop
- Pedestrian accident NOT in intersection, in a pedestrian-crossing, passing a standing or moving vehicle
- Pedestrian accident at pedestrian crossing, not in intersection
- Pedestrian accident in intersection, at pedestrian crossing
- Pedestrian accident in intersection, at pedestrian crossing, passing a standing car
- Pedestrian accident not in intersection, walking in the opposite direction to automobile traffic
- Pedestrian accident in intersection, not a pedestrian crossing, passing a standing car
- Pedestrian accident not in intersection
- Irregular pedestrian movement at the designated pedestrian crossing
- Other type of pedestrian accident

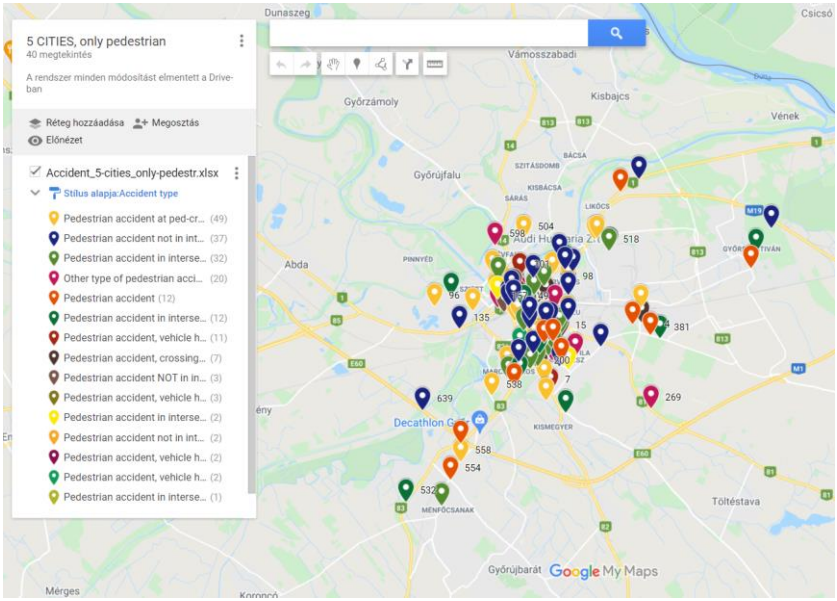


Figure 4. Pedestrian accidents map in Győr city – Hungary (based on [8])

According to the accident map, pedestrian accidents are attributed to different causes at different road characteristics. After trying to figure out a common denominator between many locations by this map to deal with a specific situation three locations could be selected with the same road type (un-signalized intersection

with three legs, with only one pedestrian crossing on the main road), where at least an accident happened at the pedestrian crossing.

5.2. Sites visit

Before starting the conflicts observation, each location was personally visited. Two hours, at least, were spent for the first visit for each location. The aim was to recognize the road users who are frequently present and the prevailing traffic conflicts at each location.

5.3. Sites description

5.3.1. Location (1): Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St

This location is the intersection of the main road (Tihanyi Árpád út 73, 9023) and the secondary road (Kassák Lajos 73 St) (see Figs. 5 and 6). A zebra crossing exists on the main road. This location is one of the most crowded place in the city, where all categories of road-users exist (vehicles, buses, motorcycles, bicyclists, and pedestrians). Many facilities exist near this intersection such as Fekete István Győr Primary School, Petz Aladár County Teaching Hospital, and Győr Pláza shopping Mall. Table 2 shows some details related to this location. The specifications of recorded accidents are listed in Table.3.

Figure 5. Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St. [16]

Table 2. Some details related to location (1) (Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.) (based on [8] and [16])

City	***The main road***	***GPS Coordinates***	***Number of accidents observed according to accidents map***
Győr	Tihanyi Árpád út 73, 9023	47°40'19.8"N 17°39'03.6"E	3

Table 3. The specifications of recorded accidents at location (1) (Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.) (based on [8] and [16])

	Time of accident	***Outcome of accident***	***GPS Coordinates***	***Accident Type***
Accident 1	19/4/2018	Slight injury	47.672166 17.650999	irregular movement at the designated pedestrian crossing
Accident 2	8/1/2016	Slight injury	47.672194 17.650944	irregular movement at the designated pedestrian crossing
Accident 3	13/12/2017	Slight injury	47.672194 17.650944	irregular movement at the designated pedestrian crossing with another vehicle (standing)

Figure 6. Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.

5.3.2. Location (2): Tihanyi Árpád út 78, 9023 way - Jereváni way

This location is the intersection of the main road (Tihanyi Árpád út 78, 9023 way -Jereváni way) and the secondary road (to Győr Pláza shopping Mall) (see Figs. 7 and 8). A zebra crossing exists on the main road. This location is also one of the most crowded place in the city. This location exists very close to location (1). Table 4 shows some details related to this location. The specifications of recorded accidents are listed in Table 5.

Figure 7. Tihanyi Árpád út 78, 9023 way - Jereváni way [16]

Table 4. Some details related to location (2) (Tihanyi Árpád út 78, 9023 way - Jereváni way) (based on [8])

City	***The main road***	***GPS Coordinates***	***Number of accidents observed according to accidents map***
Győr	Tihanyi Árpád út 78, 9023	47°40'11.3"N 17°39'10.1"E	2

Table 5. The specifications of recorded accidents at location (2) (Tihanyi Árpád út 78, 9023 way - Jereváni way) (based on [8] and [16])

	Time of accident	***Outcome of accident***	***GPS***	***Accident Type***
Accident 1	14/11/2015	serious injury	47.669805 17.652805	irregular movement at the designated pedestrian crossing
Accident 2	6/2/2017	Slight injury	47.669916 17.652666	irregular movement at the designated pedestrian crossing

Figure 8. Tihanyi Árpád út 78, 9023 way - Jereváni way

5.3.3. Location (3): Ifjúság krt. 97, 9023- Kodály Zoltán

This location is the intersection of the main road (Ifjúság krt. 97, 9023) and the secondary road (Kodály Zoltán) (see Figs. 9 and 10). A zebra crossing exists on the main road. This site locates close to the previous locations but it is less crowded. The drivers use higher speeds on this main road compared to locations 1 and 2.

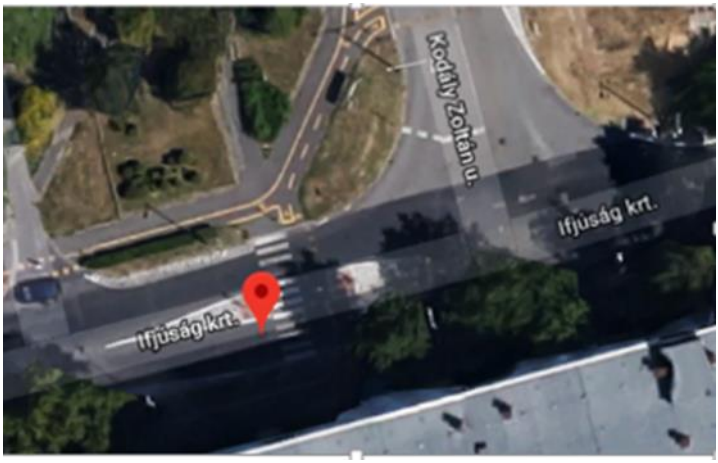


Figure 9. Ifjúság krt. 97, 9023 - Kodály Zoltán [16]

Table 6. Some details related to location (3) (Ifjúság krt. 97, 9023 - Kodály Zoltán) (based on [8] and [16])

City	***The main road***	***GPS Coordinates***	***Number of accidents observed according to accidents map***
Győr	Ifjúság krt. 97, 9023	47°40'16.8"N	Győr

Table 7. The specifications of recorded accidents at location (3) (Ifjúság krt. 97, 9023 - Kodály Zoltán) (based on [8] and [16])

	Time of accident	***Outcome of accident***	***GPS***	***Accident Type***
The reported accident	9/1/2017	serious injury	47.671333 17.654642	irregular movement at the designated pedestrian crossing

Figure 10.Ifjúság krt. 97, 9023- Kodály Zoltán

5.4. Data collection protocol

The aim of study is to observe drivers, pedestrians and bicyclists behaviour at the pedestrian crossing and cycling lanes, and to apply the Swedish traffic conflict technology at each site to evaluate the severity of observed conflicts.

In this work, vehicle-pedestrian and bicyclist conflicts were selected to be observed. In case of bicyclists, it is much more common to swerve rather than brake (as an evasive action). As for pedestrians, they are able to stop or even jump back at literally no time. Moreover, in conflicts including pedestrians or bicyclists, they also very often become the relevant road users, i.e. take the first evasive action. As their speeds are generally lower compared to motor vehicles, the Conflict Speed value "pulls down" the conflict into lower severity. At least two potential solutions for this problem have been suggested: [10]

1. Svensson [17] suggested that the line dividing the serious and not serious conflicts should be moved one or two levels down (i.e. to the severity levels 25 or 24). This would include more conflicts into the final analysis and compensate for the under-scoring of the bicyclist/pedestrian conflicts.
2. Shbeeb [18] suggested that in the situations with pedestrian, it should always be the speed of the motor vehicle used as the Conflicting Speed, regardless to who is taking the evasive action. This is what was adopted in this study.

It was considered that the driver always takes the evasive action (braking). In this case, we estimate the driving speed as a conflicting speed, and the distance (between vehicle and pedestrian in case vehicle-pedestrian conflict observation, and between vehicle and bicyclist in case vehicle-bicyclist conflict observation).

The tasks as an observer on- site were as follows:

- detecting the conflict between the two road-users.
- filling in other relevant information together with a verbal description of the course of events.
- making a sketch of the conflict
- calculating Time-to-Accident (TA) values based on the conflicting speed and the distance between the two conflicting road-users for each conflict and then drop the data (conflicting speed and time-to-accident values for each conflict) on the Swedish conflict graph. This is to be made in both cases, vehicle-pedestrian conflicts and vehicle- bicyclist conflicts.

5.4.1. Equipment required for observation

- Conflict register forms.

No specific form was adopted during observation period. The data were registered on a notebook, such as: the location, observations date and time, the two conflicting road users, the speed of vehicle when taking an evasive action (braking), the distance between the two confliction road users (the distance to the collision point), the conversion table (to convert speed and distance to time TA) shown in table (1), the calculated TA-value, a sketch of the conflict, and verbal description of the course of events,

We can also calculate the time as a distance divided by the speed by an equation in a form of excel-sheet for each location. (Especially for the speed values, which fall ranging between intervals that were not mentioned directly in the converting table).

- A watch.
- A pencil.
- A simple video camera works together with the field work.

5.4.2. Considerations during observation period

According to the literature review related to the Swedish technique, the number of days proposed for the observations and the periods of recording these observations during the day are determined based on the frequency of the expected conflicts at the given location. The expected frequency of conflicts is usually derived from previous experiences. According to e.g. Abdul Manan and Várhelyi [12] at sites with major road safety problems short observation periods can be done as the number of conflicts per time unit is still relatively high there. Observations are usually done in periods of 1 -2 hours with breaks in between for the observer to recover. [10]

Since the three selected sites are often located in the most crowded areas of the city (especially the first and second sites), and they are the same sites where at least an accident occurred in recent years between 2014 and 2018, the survey was conducted first in order to reveal the nature of traffic conflicts between road users personally through on-site visits, so that the measurements taken are objective as possible, and then field measurements and the necessary data (speeds and distances) were estimated. Even though the estimated measurements (in some cases) may not accurately reflect the real situation (the realistic values), but this approach may give at least a good background about the level of severity of the observed conflicts.

Some other considerations during the observation:

- An observational point has selected at each location to offer a clear view over the observational area.
- The observer is good at driving process; speed estimations were not an issue.
- To facilitate estimation of distance and speed, initial measurements by the observer have done at the first arrival at the scene (distance between salient objects or marks can be measured).
- The same number of conflicts were considered for all locations. Once we got an equal number of conflicts at each location, the observation process stops. The field measurements ended within two weeks for the three locations.
- Three types of conflicts were adopted at each location. Six conflicts for each type, as will be illustrated in field measurement section.

5.5. Field measurements

Since many facilities exist near the three selected sites such as Fekete István Győr Primary School, Petz Aladár County Teaching Hospital, and Győr Pláza shopping Mall, the sites survey was comprehensive, several road users (motorists-pedestrians-bicyclists) were frequently present and different kinds of conflicts could be observed. The results are represented in Figs. 11–16 and in Tables 8–10 (see Paragraphs 5.5.1–5.5.3).

5.5.1. Location (1): Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.

Figure 11. Photos taken for some conflicts observed at location (1) – Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.

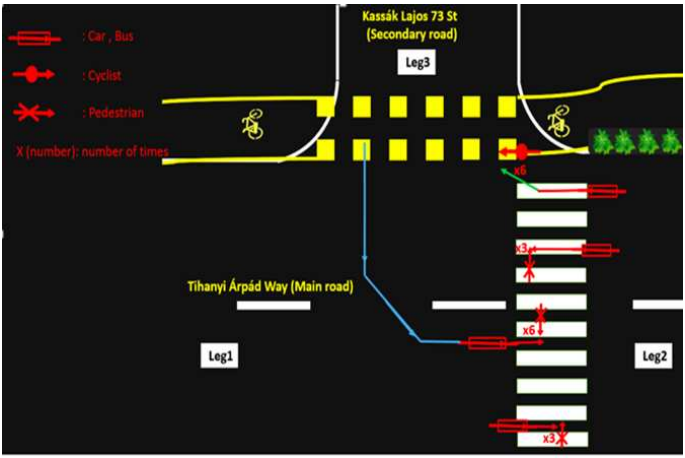


Figure 12. A sketch of the types of conflicts observed at location (1) – Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.

Table 8. Summary table of conflicts observation at location (1) – Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.

Main Road: Tihanyi Árpád Way. Secondary Road: Kassák Lajos 73 St.							
Conflict Type	Road user 1	Road user 2	conflict number	Distance [m]	Conflicting Speed [km/hour]	Conflicting Speed [m/Sec]	Time-To-Accident [Sec]
vehicle straight, pedestrian	vehicle	Pedestrian	1	3.5	30	8.3	0.42
			2	3.5	32	8.9	0.39
			3	3.5	30	8.3	0.42
			4	4	35	9.7	0.41
			5	3	30	8.3	0.36
			6	4	32	8.9	0.45
vehicle turning left, pedestrian	vehicle	Pedestrian	1	3.5	20	5.6	0.63
			2	3	22	6.1	0.49
			3	3.5	18	5	0.7
			4	3.5	20	5.6	0.63
			5	2.5	18	5	0.5
			6	3	20	5.6	0.54
vehicle turning right, bicyclist	vehicle	bicyclist	1	2.5	18	5	0.5
			2	2.5	18	5	0.5
			3	2.5	20	5.6	0.45
			4	3.5	20	5.6	0.63
			5	3	18	5	0.6
			6	3.5	18	5	0.7

5.5.2. Location (2): Tihanyi Árpád út 78, 9023 way - Jereváni way

Figure 13. photos taken for some conflicts observed at location (2) – Tihanyi Árpád út 78, 9023 way - Jereváni way

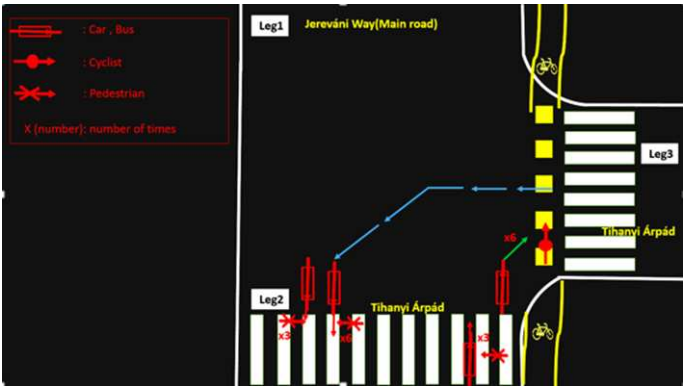


Figure 14. A sketch of the types of conflicts observed at location (2) – Tihanyi Árpád út 78, 9023 way-Jereváni way

Table 9. Summary table of conflicts observation at location (2) – Tihanyi Árpád út 78, 9023 way-Jereváni way

Main Road: Tihanyi Árpád Way-Jereváni Way. Secondary Road: Tihanyi Árpád.							
Conflict Type	**Road user 1**	**Road user 2**	**conflict number**	**Distance [m]**	**Conflicting Speed [km/hour]**	**Conflicting Speed [m/Sec]**	**Time-To-Accident [Sec]**
vehicle straight, pedestrian	vehicle	Pedestrian	1	6	45	12.5	0.48
			2	6	42	11.7	0.51
			3	3.5	32	8.9	0.39
			4	5.5	42	11.7	0.47
			5	4	28	7.8	0.51
			6	4	35	9.7	0.41
vehicle turning left, pedestrian	vehicle	Pedestrian	1	4	28	7.8	0.51
			2	3.5	28	7.8	0.45
			3	4	25	6.9	0.58
			4	2.5	22	6.1	0.41
			5	3.5	22	6.1	0.57
			6	3.5	22	6.1	0.57
vehicle turning right, bicyclist	vehicle	bicyclist	1	2	15	4.2	0.48
			2	2.5	15	4.2	0.6
			3	1.5	12	3.3	0.45
			4	2	15	4.2	0.48
			5	2.5	15	4.2	0.6
			6	2.5	18	5	0.5

5.5.3. Location (3): Ifjúság krt. 97, 9023 - Kodály Zoltán St.

Figure 15. Photos taken for some conflicts observed at location (3) – Ifjúság krt. 97, 9023 - Kodály Zoltán St.

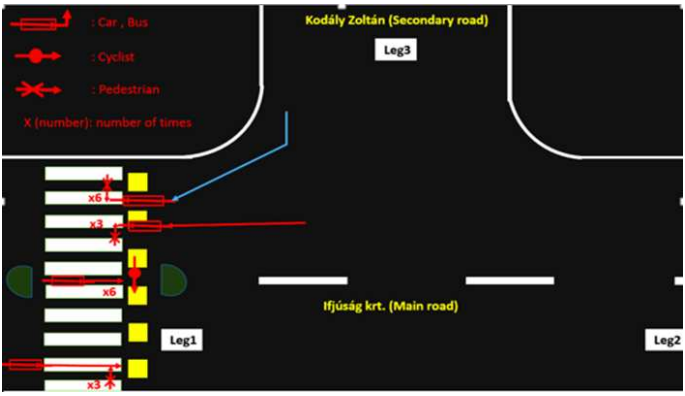


Figure 16. A sketch of the types of conflicts observed at location (3) – Ifjúság krt. 97, 9023 - Kodály Zoltán St.

Table 10. Summary table of conflicts observation at location (3) – Ifjúság krt. 97, 9023 - Kodály Zoltán St.

Main Road: Ifjúság krt. Secondary Road: Kodály Zoltán St.							
Conflict Type	**Road user 1**	**Road user 2**	**conflict number**	**Distance [m]**	**Conflicting Speed [km/hour]**	**Conflicting Speed [m/Sec]**	**Time-To-Accident [Sec]**
vehicle straight, pedestrian	vehicle	Pedestrian	1	6	42	11.7	0.51
			2	5.5	38	10.6	0.52
			3	5	35	9.7	0.51
			4	3.5	32	8.9	0.39
			5	4	32	8.9	0.45
			6	5	38	10.6	0.47
vehicle turning right, pedestrian	vehicle	Pedestrian	1	4.5	20	5.6	0.81
			2	3.5	22	6.1	0.57
			3	4	20	5.6	0.72
			4	4	22	6.1	0.65
			5	3.5	22	6.1	0.57
			6	4	25	6.9	0.58
vehicle straight, bicyclist	vehicle	bicyclist	1	5.5	40	11.1	0.5
			2	4.5	35	9.7	0.46
			3	5	38	10.6	0.47
			4	4	32	8.9	0.45
			5	3.5	28	7.8	0.45
			6	4.5	30	8.3	0.54

6. Results and discussions

The authors summarize their results in Figs. 17–19, see Paragraphs 6.1–6.3. The results are based on the author's literature review [19].

6.1. Location (1): Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.

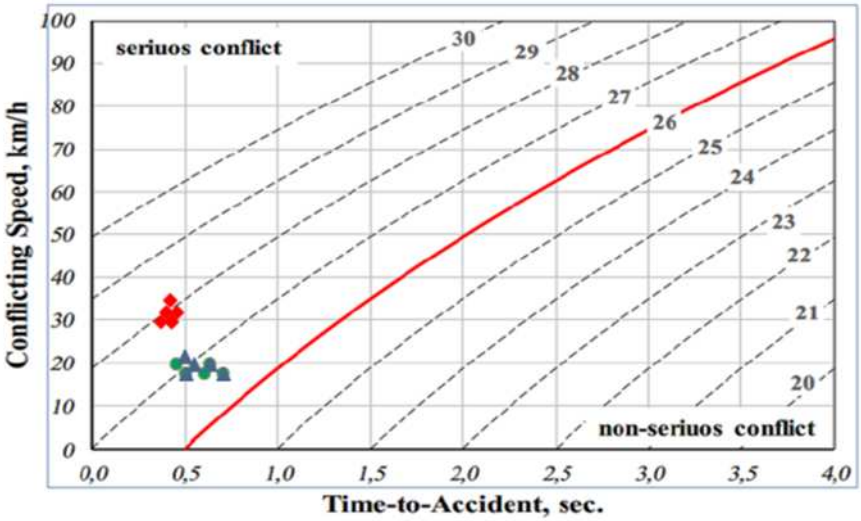


Figure 17. The Swedish graph for location (1) – Tihanyi Árpád út 73, 9023 - Kassák Lajos 73 St.

- The conflicts indicated by a rhombus: "vehicle move straight on the main road, pedestrian at zebra crossing"; Since this location is classified as one of the most crowded streets in the city, this causes a great potential of traffic conflicts occurrence between road users with a high frequency during the day. Pedestrian-vehicle conflicts on the Main road fall around severity curve (No. 28) to be considered as the most serious conflicts compared to other types of conflict at this location.

- The conflicts indicated by a triangle: "vehicle on the secondary road turning left to the main, pedestrian at zebra crossing" and the conflicts indicated by a circle: "vehicle on the main turning right to the secondary, bicyclist at cycling lane"; The speeds at which the drivers began to take the evasive actions are almost identical for both types; the fact here is that when drivers change their direction from main to secondary or from secondary to main), they often try to keep attention to the traffic movements and move with lower speeds. Most conflicts fall within the 26-27 severity levels. They are classified within the serious conflict zone, but less severity than the first type. The values of time-to-accident are a bit higher than first type due to the lower speeds.

6.2. Location (2): Tihanyi Árpád út 78, 9023 way - Jereváni way

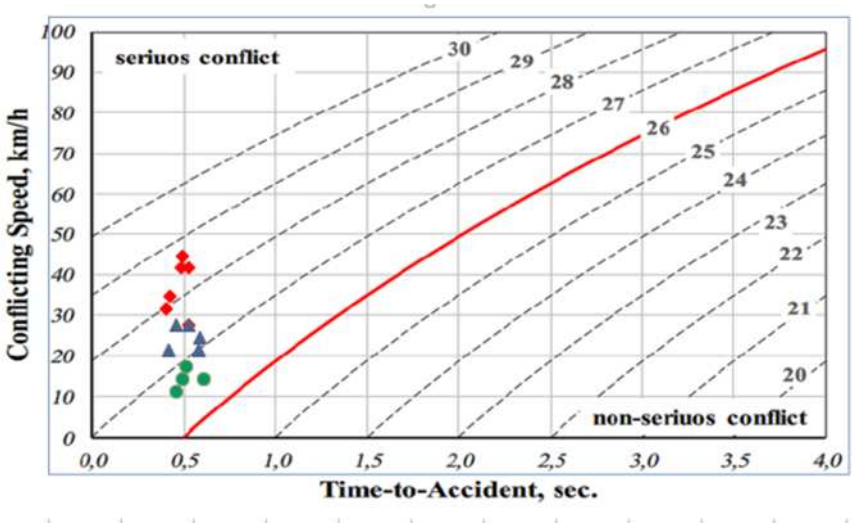


Figure 18. The Swedish graph for location (2) – Tihanyi Árpád út 78, 9023 way - Jereváni way

- The conflicts indicated by a rhombus: "vehicle move straight on the main, pedestrian at zebra crossing"; the speed values at which the drivers began to take the evasive action on the main road (Tihanyi Árpád Way-Jereváni Way) are higher than location (1). The drivers here have more freedom at driving speeds as they move (from Tihanyi Árpád Way) to the wider main road (Jereváni Way). However, they are obligated to slow down suddenly when they arrive the zebra crossing. Since the speeds are relatively higher, most of pedestrian-vehicle conflicts on the main road fall within the 28 – 29 severity levels.
- The conflicts indicated by a triangle: "vehicle on the secondary turning left to the main, pedestrian at zebra crossing"; most conflicts of this type fall within the 27-28 severity levels. The drivers who turn from the secondary road to Tihanyi Árpád way on their left are trying to accelerate a bit when they do not meet any traffic movements approaching to Jereváni Way, at this acceleration they may suddenly brake for a pedestrian at the zebra crossing.
- the conflicts indicated by a circle: "vehicle on the main turning right to the secondary, bicyclist at cycling lane "; the speeds at which the drivers began to take the evasive actions when turning from the main road to the secondary road are low, since the secondary road here is an entrance / exit for Győr Plaza shopping mall, and at this point they can only enter with low speeds. All conflicts

of this type fall within the 26 - 27 severity levels (the area close to the neutral line).

6.3. Location (3): Ifjúság krt. 97, 9023 - Kodály Zoltán St.

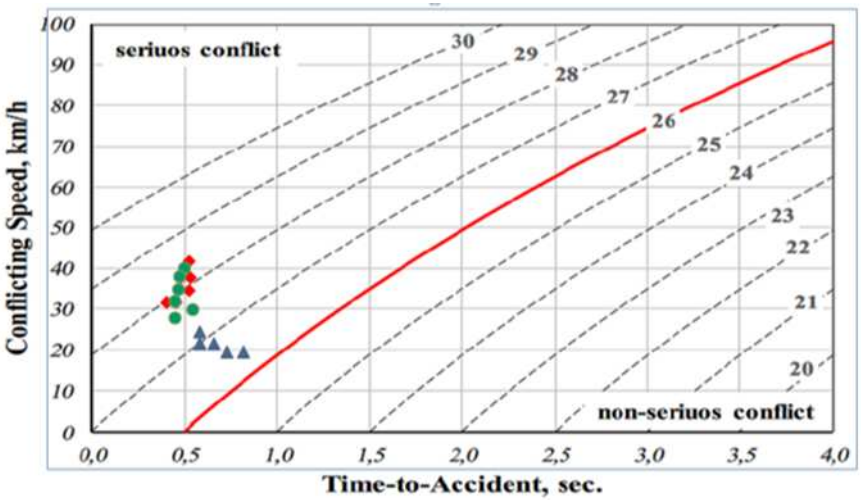


Figure 19. The Swedish graph for location (3) – Ifjúság krt. 97, 9023 – Kodály Zoltán St.

- The conflicts indicated by a rhombus: "vehicle move straight on the main, pedestrian at zebra crossing"; this street is less crowded than the previous locations (1, 2). Even though the drivers are braking mostly at logical distances reaching to 6 meters, the situation is still dangerous since we are on a main road. We note that pedestrian-vehicle conflicts on the main road fall around severity curve (No. 28) and are considered the most dangerous conflicts.
- The conflicts indicated by a triangle: "vehicle on the secondary turning right to the main, pedestrian at zebra crossing"; most conflicts of this type fall within the 26-27 severity levels. This type has less severity than the situation on the main road in the first type.
- The conflicts indicated by a circle: "vehicle move straight on the main, bicyclist at cycling lane "; we note that bicyclist -vehicle conflicts on the main road fall around severity curve (No. 28) and are considered the most dangerous conflicts. Most of time to accident values are less than 0.5 second. This indicates shorter time to collision.

7. Conclusions

Accidents database and the technical approach for conflicts observation can be used as two methods to analyze traffic conflicts, as these two methods complement each other to gain more accurate results about the probability of accidents at a given location.

Looking at the results, and since the Swedish technique has been applied at the same locations where some accidents happened, we can say that there is a compatibility between Swedish technology and accident database regarding the seriousness of conflicts and probability of accidents.

Most pedestrian and bicyclist – vehicle conflicts are the most serious conflicts at all locations. The drivers on the main roads do not stop "in some cases" at zebra crossing for pedestrians as it was noticed during sites visit. Based on the findings, improving pedestrian accident countermeasures is always needed. The safety of vulnerable road users has to be always supported by implementation the traffic calming measures and accidents countermeasures especially at un-signalized intersections.

References

[1] A. Levy, In Canada, In Adoption & Fostering 8 (1984) doi: https://doi.org/10.1177/030857598400800119

[2] Abhishek, M. Mandjes et al., Congestion analysis of unsignalized intersections (2018), ArXiv.

[3] S. M. Sohel Mahmud, L. Ferreira et al., Reviewing traffic conflict techniques for potential application to developing countries. *Journal of Engineering Science and Technology* 13 (6) (2018) pp. 1869–1890.

[4] G.B. Grayson, C. Hyden et al., The Malmo study, a calibration of traffic conflict techniques. A study organised by ICTCT - the International Committee on Traffic Conflict Techniques. Institute for Road Safety Research SWOV, The Netherlands (1984).

[5] C. Hydén, The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis. Lund University, Department of Traffic Planning and Engineering, (1987).

[6] S.Turner, K. Fitzpatrick et al., Motorist yielding to pedestrians at unsignalized intersections findings from a national study on improving pedestrian safety, Journal of the Transportation Research Board, No. 1982, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 1–12.
doi: https://doi.org/10.3141/1982-03

[7] T.Q. Vuong, Traffic Conflict Technique Development for Traffic Safety Evaluation under Mixed Traffic Conditions of Developing Countries. *Journal of Traffic and Transportation Engineering* 5 (4) (2017) pp. 228–235.
doi: https://doi.org/10.17265/2328-2142/2017.04.004

[8] WEB-BAL. Accident database management software (2018).

[9] R.S. Perkins, J.I. Harris, Traffic conflict characteristics – Accident potential at intersections. Highway Research Record No. 224 (1968), pp. 35–43.

[10] S. Almqvist, L. Ekman, *The Swedish traffic conflict technique observer's manual*. (July 2001), pp. 1–4.

[11] P. Zajíc, Traffic Conflicts and Road Transport Safety New Development (2012). Directive of the European Parliament and of the European Council No. 2008/96/EC on road infrastructure safety management. 4(VII).

[12] M. M. Abdul Manan, A. Várhelyi, Motorcyclists' road safety related behavior at access points on primary roads in Malaysia – A case study. *Safety Science* 77, (2015), pp. 80-94.
doi: https://doi.org/10.1016/j.ssci.2015.03.012

[13] S. Almqvist, C. Hyden, Methods for Assessing Traffic Safety in Developing Countries, *Building Issues (LCHS)* 6 (1) (1994) pp. 3–20. [cited 2021-03-15].
URL
http://www.lth.se/fileadmin/hdm/BI_Volume_06_1_1994_Methods_for_Assessing_Traffic_Safety_in_Development_Countries.pdf

[14] C. Hydén, L. Linderholm, The Swedish Traffic-Conflicts Technique. *International Calibration Study of Traffic Conflict Techniques* (1984) pp. 133–139.
doi: https://doi.org/10.1007/978-3-642-82109-7_12

[15] G. Tiwari, D. Mohan, J. Fazio, Conflict analysis for prediction of fatal crash locations in mixed traffic streams. *Accident Analysis & Prevention* 30 (2) (1998) pp. 207–215.

[16] Google Maps (2021) [cited 2021-03-15].
URL *maps.google.com*

[17] Å. Svensson, A method for analysing the traffic process in a safety perspective. Doctoral thesis. University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering, (1998).

[18] L. Shbeeb, Development of a traffic conflicts technique for different environments - a comparative study of pedestrian conflicts in Sweden and Jordan. Doctoral thesis. University of Lund, Lund Institute of Technology, Department of Technology and Society, Traffic Engineering, (2000).

[19] A. Kizawi, A. Borsos, A Literature review on the conflict analysis of vehicle-pedestrian interactions. *Acta Technica Jaurinensis* (2021).
doi: *https://doi.org/10.14513/actatechjaur.00601*

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 406-423, 2021
DOI: 10.14513/actatechjaur.00620

Development of comparative investigation method for timing chain wear analysis using oscillating tribometer

L. Paulovics[1,*], R. Kuti[2], J. Rohde-Brandenburger[1], Cs. Tóth-Nagy[1]

[1]Széchenyi István University, Department of Internal Combustion Engines and Propulsion Technology, Egyetem tér 1., 9026 Győr, Hungary
[2]Széchenyi István University, Department of Mechatronics and Machine Design Egyetem tér 1., 9026 Győr, Hungary
*e-mail: paulovics.laszlo@sze.hu

Submitted: 17/05/2021; Accepted: 03/08/2021; Published online: 30/08/2021

Abstract: This paper presents the development of a test procedure to investigate timing chain components under abrasive stress on a tribometer. Engine developers use cost and time expensive engine dynamometer tests to investigate timing chain life expectancy under different conditions. Tribometer tests are fast and cost effective, but these use standardized specimen material and geometry that greatly differ from the original tribological system of the timing chain. Manufacturing specimens from the original chain material using the original technology is complicated; surface quality and hardness properties cannot be guaranteed. The aim of research was to develop a test method for rapid and cost-effective comparison of engine lubricants, timing chain materials or coatings, as well as to assess the wear resistance of the chain to contaminants. Various uncontaminated and carbon black blended lubricants were compared using standard-based ball-on-disc tribometer tests to tribometer tests using actual timing chain components (bush-on-pin test) of a Diesel engine. Lubricants were ranked in terms of coefficient of friction and wear. Results showed that bush-on-pin tests were comparably suitable for testing lubricants when evaluated against standard ball-on-disc tribometer tests.

Keywords: *timing chain; wear measurement; tribometer, motor oil, soot,*

1. Introduction

The trend of engine development over the last 20 years (reduction of mechanical losses, increase of specific power, start-stop systems, hybridization, etc.) increased the load of timing chains similarly to most engine components. In addition to the increasing load, it is also a challenge that the timing chain is especially sensitive to the quality of the lubricant [1] [2]. Due to low viscosity motor oils, increasing oil change intervals and the soot formation (especially in Diesel engines, but it is already characteristic of gasoline engines as well), under certain conditions chain wear may be higher than planned or permitted [3].

Timing chain wear adds up at each link and can be measured as chain elongation, that can cause various problems during engine operation. Due to the elongation of the chain, valve timing is shifted, resulting in a loss of power and a deterioration of the emission values. In addition, as the chain elongates, the chances of the chain breaking or the chain jumping over the sprocket (one or more teeth) increase. These faults could cause immediate engine damage. It is therefore necessary to improve timing chains continuously, to adapt to changing conditions and new types of motor oils [1] [2].

In the field of engine development, long-term engine tests and radionuclide technology (RNT) wear measurements are used to test chain wear [4] [5], but due to the high cost and time required of these test methods, a compromise must always be made in the measurement program. This means that the number of variations of engine working point and oil quality is limited. In the case of tests performed on the engine dynamometer, the occupancy and operation of the dynamometer and the fuel consumption represent a significant time requirement and cost. In the case of RNT wear measurement, the high activation cost of the components and the management of the radioactive components and waste must also be taken into consideration. It is also a problem that under normal engine load, it can take hundreds of engine working hours to produce measurable wear. In addition, the carbon footprint of these tests is also significant.

Testing chains, chain segments or chain links on a component test bench or tribometer can provide an opportunity to replace expensive engine tests. However, such test methods are still novel for timing chains, at least they are rarely used [6]. Therefore, it is necessary to develop and analyze similar methods for correlation with wear processes in the real environment.

2. Levels of tribological testing

There are several levels of tribological testing available for testing timing chains (like any machine part). According to Czichos and Habig [7] these can be:

I. Field test: test of the original and complete tribological system under original operating conditions (in case of timing chains, it means in-vehicle test)

II. Dynamometer test: examination of the original and complete tribological system on nearly realistic load on a dynamometer. (in case of timing chains, it means vehicle dynamometer test)

III. Subsystem test: examination of subsystems, component groups under nearly realistic conditions. (in case of timing chains, it means engine dynamometer test)

IV. Component test: testing of original or simplified components under conditions similar to the real ones. (in case of timing chains, it means a chain drive test bench)

V. Simplified component test: examination of specimens similar to the real part under conditions similar to the real ones. (in case of timing chains, it means testing of chain segments, links or chain-like specimens)

VI. Model test: a test performed on special test specimens under arbitrarily defined conditions for the purpose of basic research of friction and wear processes. (in case of timing chains, it means simple specimens modeling chain links – but in shape and material very different from it – at a different load from the original)

Moving from field test to model test, the real system is more and more simplified, with the obvious disadvantage of moving away from reality. However, simplifying complex systems – while retaining important parameters from the viewpoint of the examination – offers a number of benefits: cost, time and material requirements, and ecological footprint of the testing can be reduced.

In case of simplified component and model tests, the relevant parameters correspond to reality, but other test parameters are standardized for the purpose of cost-effectiveness, comparability and the simpler and more efficient exploration of causal relationships. In practice, this means that for example the material of the specimens and the quality and temperature of the lubricant correspond to the real operating conditions, but the shape of the specimens, the contact surface, the magnitude of the load are standardized, they do not want to follow the reality. Thus, the effect of irrelevant factors can be ruled out.

There are only few researchers dealing with wear testing of timing chains, and therefore the range of testing technologies used is not very wide. Mostly a test engine, possibly a component test bench, i.e. a chain test bench, is used, but just a few are involved in model testing. The using of a component test bench can seem to

be an obvious solution, but its big disadvantage is that a custom special machine has to be developed for it. Due to the scarcity of the field, it is not profitable to develop such a special purpose machine for sales purposes. Therefore, even worldwide a small number of research institutes or development centers performing chain drive development design and create a chain test bench for their own use, according to their own needs, e.g. the Institute of Machine Elements, Gears, and Transmissions (MEGT) at the University of Kaiserslautern [6] [8] [9] [10].

The advantage of tribometer tests is that they are generally flexible in design to be suitable for testing a variety of specimens. Thus, for existing, commercially available tribometers, only a single specimen clamp needs to be developed and manufactured for chain link testing at minimal cost. However, its applicability is limited, and since it is a highly simplified tribological system compared to the original part, it is absolutely necessary to validate the results obtained with this method. Due to the narrowly used method, there is neither a mature target machine or adapter suitable for holding chain links and chain segments, nor a measurement method or standard, therefore it is necessary to develop such methods individually.

In the following, the method for wear testing of timing chain elements on tribometer is compared to the method of standardized ball-on-disc tribometer test in order to validate the results of the first one. Research was performed in the tribology laboratories at the Department of Internal Combustion Engines and Propulsion Technology at Széchenyi István University.

3. Comparative experimental investigations on tribometer

3.1. Ball-on-disc model tests performed on ball-on-disc tribometer

There are several tests defined by international standards for the qualification and comparative testing of lubricating oils and greases [11] [12] [13]. Of these, the 4-ball-test is most commonly used to test greases [11]. To compare lubricating oils, the ball-on-disc test is also used, in which a ball of a bearing is pressed onto the planar surface of a disc at a given load, reciprocating the ball at a given frequency and stroke [13]. The wear and friction between the two is reduced by the lubricating oil passed between the contact surfaces.

3.1.1. Methodology of ball-on-disc measurements

In the ball-disc assembly (Fig. 1.), the specimens are in a point-like contact or they come in contact with a small contact surface close to it, so the tribological system operates in a mixed lubrication regime. Thus, the effect of additives and contaminants in the oil is definitely shown. The purpose of such tests is therefore to test the system in a mixed lubrication regime.

The tribometer can measure friction online using built-in load cells, while wear can usually be determined after the tribometer test by measuring the size of the wear mark – in case of ball specimens wear scar diameter (WSD).

Figure 1. Ball-on-disc specimens

In the Department of Internal Combustion Engines and Propulsion Technology, such tests are used for various oil qualification, oil comparison tests, and the effects of oil additives and contaminants. These tests are performed on an Optimol SRV®5 tribometer [14]. In the last years, more than 500 new, used, aged and contaminated oil samples have been compared with this test, and the results showed a significant difference between different new (unused, uncontaminated, unaged) oils as well as between differently contaminated and used oils in terms of both friction and wear.

The material of the disc used in the test is 100Cr6, its surface is lapped and its surface roughness Rz is between 0.5-0.65 μm. The other part is a standard 100Cr6 bearing ball with polished surface. Additional parameters for the specimens are shown in Table 1.

Table 1. Parameters of the ball-on-disc specimens

specimen	***material***	***dimensions***	***hardness***	***roughness***
ball	100Cr6	ø10 mm	61.5 HRC	Ra 0.02 μm
disc	100Cr6	ø24x7.9 mm	62 HRC	Ra 0.047 μm

The test parameters are based on ISO 19291: 2016 [13], with the exception that the load is 150 N instead of 300 N and the temperature is 100 °C instead of 120 °C. In addition, instead of the 3ml of oil required by the standard, a peristaltic pump provides continuous oil circulation, so the heat removal is continuous from the contact surface and the operation of the tribological system is stable, approaching the conditions in a real tribological system. The examination parameters are shown in Table 2.

Table 2. Examination parameters of the ball-on-disc test

	stroke	***frequency***	***load***	***time***	***oil flow***	***temp.***
step 1.	1 mm	50 Hz	50 N	30 s	225 ml/h	100 °C
step 2.	1 mm	50 Hz	150 N	2 h	225 ml/h	100 °C

All of these changes are necessary because during standard tests, the system was overloaded with frictional heat in the first half hour, which resulted that the oil lost its lubricity and the surface burned and then was severely worn. This phenomenon can be attributed to the fact that the standard has not been developed specifically for motor oils, that have a much lower viscosity compared to average lubricating oils (e.g. gear oils).

On the contact surfaces, the evaluation of the wear is based on the diameter (on the ball) and width (on the disc) of the wear scar. Measurements were performed using a Keyence VHX-1000 digital microscope. Fig. 2. illustrates the wear scar diameter measured on the ball, that was used for the evaluation.

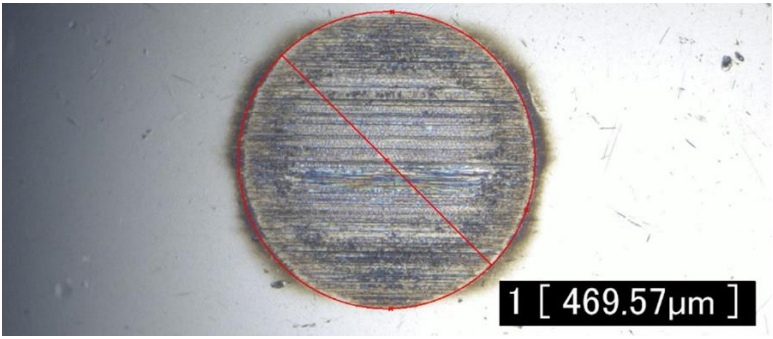


Figure 2. Wear scar on the ball specimen measured by digital microscope

During the test, the coefficient of friction was continuously measured. However, just the mean of the last 10 minutes was considered for the evaluation, because it was always stable enough.

The method presented above was developed for the purpose of comparing and ranking engine lubricants. It was important that the test was not specified for a specific engine part and could be repeated at any time. Because the goal was to create a database, it was important that the specimens were available anytime, in constant quality. Therefore, standard test specimens were used. The 100Cr6 material of these can be considered neutral in terms of all components of an average engine, this material or similar is mostly used as the material of camshafts, valve lifters or rocker arm rollers.

The tested timing chain is from a V6 Diesel engine with a factory oil of Shell Helix Ultra Professional AV-L 0W-30, which was the reference oil. It was compared to two additional new oils from different viscosity grades (Castrol Edge Professional TWS 10W-60, Castrol BOT 920 0W-20) and versions of Shell Helix Ultra Professional AV-L 0W-30 mixed with different concentrations of carbon black. Hereinafter, the tested oils will be referred by their manufacturer and viscosity grade.

Carbon black is primarily used for industrial goals, generally as black color pigment, e.g. coloring tires. It can be used to replace engine soot generated during the combustion process of Diesel engines [15] [16] [17]. This is a commonly used method in tribological research, as it makes possible to control of quantity and quality of soot in the oil. In the case of real engine soot (generated in engine combustion process on test bench or vehicle), the repeatable quality and quantity cannot be ensured. Uy et al. [15], and La Rocca et al. [18] have shown that the quality of soot samples taken from an engine depends strongly on many factors: engine type, engine operation mode, and even the sampling location inside the engine. In the research, Cabot Monarch® 430 carbon black was mixed into the fresh oil samples in concentrations of 0.5 wt%, 1 wt% and 2 wt%, using magnetic stirrer as first step (1000 rpm, 10 min), then ultrasonic homogenizer (70 °C, 15 min) and as last step magnetic stirrer (1000 rpm, 5 min) again.

Due to the mixed lubrication regime, the measurable difference in friction can be expected not from the different internal friction of the oil, but from the realized lubrication regime, how efficient the oil and its additives can form a coherent lubricating film on the surface, how much load the lubricating film can withstand and how its anti-wear effects are deteriorated by contaminants.

3.1.2. Results of ball-on-disc measurements

The measurements were performed at 100 °C, so the warm viscosity of the oils was decisive. At least 3 tribometer tests were performed with each oil sample and the mean values and standard deviation of the coefficient of friction (CoF) were calculated and are shown in Fig. 3.

The coefficient of friction was lowest with Castrol 10W-60 oil, followed by the Castrol 0W-20, which preceded the Shell 0W-30 oil. It should be noted that the listed oils differ significantly not only in their viscosity but also in their additives, so the differences in friction and wear values depend on the additive package at least as much as on the viscosity. It was not the aim in this case to explore the cause of the differences between the oils and to examine the additive packages.

In case of oils mixed with carbon black (CB), it can be observed that the coefficient of friction increased slightly with the concentration of carbon black.

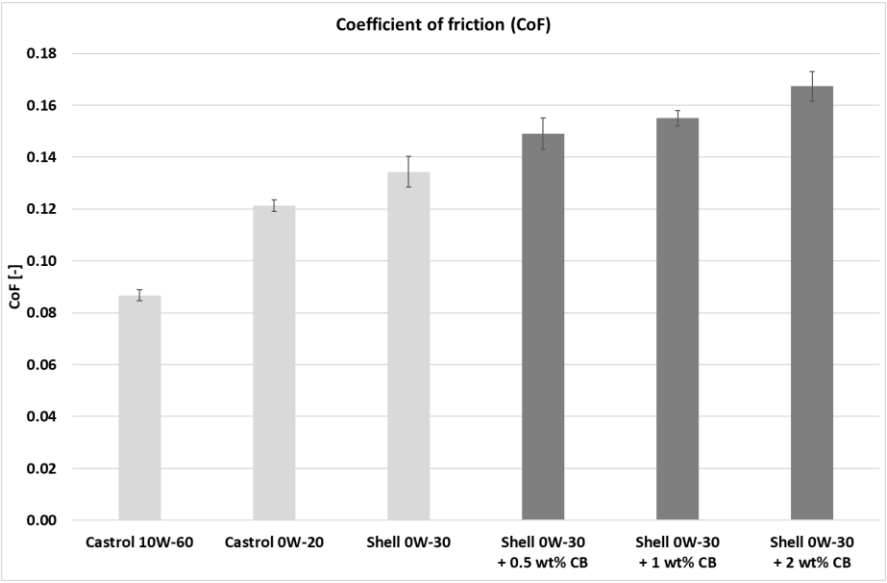


Figure 3. Coefficient of friction values measured with different oils on a ball-on-disc tribometer

In case of the wear scar diameters (Fig. 4.), a similar order has developed to the case of the coefficient of friction. This is logical in light of the fact that the thinner oil film presumably allowed more metallic contact between the surfaces. Therefore, both friction and wear were higher, although anti-wear and other additives also have a big impact on this process. Of the fresh oils, the high viscosity Castrol 10W-60 oil was clearly the best, followed by Castrol 0W-20 and Shell 0W-30 oils.

The effect of soot on the operation of Shell 0W-30 oil is all the more interesting. By adding 0.5 wt% and 1 wt% carbon black, the wear increased slightly, but not in proportion to the carbon black content. This is also related to the increase in friction. Here, it can be stated that the low carbon black content increased the stochasticity of the process. However, with a 2 wt% carbon black content, wear increased sharply. This phenomenon was expected regarding the observations of Ratoi et al. [19], Antusch et al. [16], and Hu at al. [17]. Through various tribological and chemical studies, they have found that carbon black – in addition to its abrasive effect – is strongly bound to metallic surfaces and to the anti-wear additives due to its polar structure. Therefore, the carbon black binds strongly to the metallic surface, isolating it from anti-wear additives, and also binds anti-wear additives, from which it does not get enough on the metallic surface to form a wear-resistant protective layer.

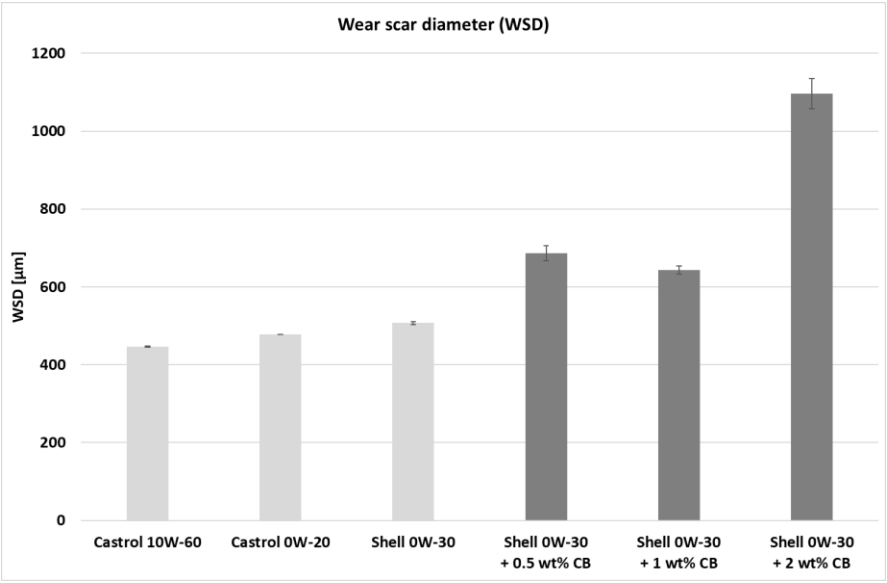


Figure 4. Wear scar diameter values measured with different oils on a ball-on-disc tribometer

Thus, the critical value for the carbon black content in this case were somewhere between 1 wt% and 2 wt% with this oil. This was in line with preliminary expectations, as 2 wt% soot content is typically the critical limit for Diesel engine lubricants according to experiences with MOL LubCheck.

A series of tribological tests were also performed on timing chain elements with the same oil samples. The purpose of this was to validate the specimen test with the model test.

This was necessary because specimen testing is a process under development and therefore results need to be validated. Ball-on-disc model testing is based on a widely used standard, it has been used for a long time in the department, and its results were reliable [20][18]. Therefore, it can be used as a reference in the development process of bush-on-pin specimen tests, performed on tribometer.

3.2. Chain bush-on-pin specimen tests performed with chain elements on tribometer

There is a Plint TE-77 reciprocating tribometer at the department that was used for experimenting with wear testing of timing chains.

3.2.1. Methodology of bush-on-pin measurements

On the department's tribometers (and on the most type of tribometers in general), it is not possible to test the chain joints in the original ("pin-in-bush") configuration and on load similar to the original, the chain links are tested in a "bush-on-pin" configuration instead of the original configuration "pin-in-bush".

In case of a bush chain, the outer surface of the bush (in case of a tooth chain, the back of a moving plate) is pressed onto the surface of a pin fixed in the stationary clamp while reciprocating it at a given stroke and frequency, see Fig. 5.

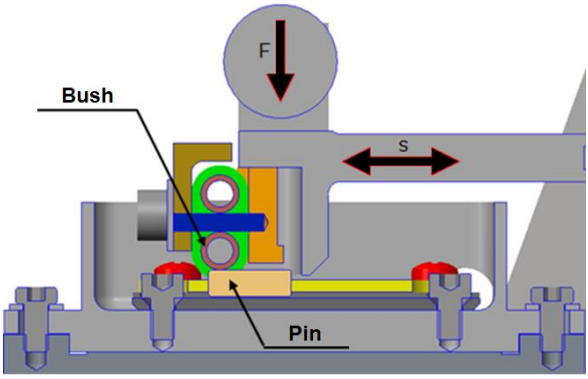


Figure 5. Position of the bush and pin in the reciprocating tribometer

The obvious disadvantage of the bush-on-pin method is that the model moves away from the real system by one step. However, it is still closer to the original tribological system than the ball-on-disc model used to classify lubricants [13], because the original material qualities are used. A further advantage over the real chain (pin-in-bush) assembly is that a point-like contact surface is created instead of a linear one (similar to the ball-on-disc test). This speeds up wear on one hand, and less uncertainty in the results can be expected on the other hand, because it is less sensitive to assembly inaccuracies (e.g. angular misalignment between the bush bore and the pin axis) and the deviation in the production quality (actual size within the tolerance field) of the parts has also much less influence.

An additional advantage is that due to the higher surface pressure (initially the p_{max} is around 2,9 GPa - approximate calculation in the absence of knowledge of the exact material properties) the parts wear faster, so the tests can be shorter.

It has to be mentioned, that hardness and surface quality of the bushes are very similar on their outer and inner surface, so it does not have more influence on the tribological behavior than the point-like contact instead of line-like. The hardness values were 880 to 970 HV on the outer and 820 to 920 HV on the inner surface. Also roughness parameters were of a similar magnitude, Rz values of the outer surface were around 1.3 μm, inner surfaces around 1.0 μm.

The bush-on-pin tests are suitable for relative comparison, like model tests. It may be suitable to compare different oil qualities, taking into consideration the given material quality of the chain, or to compare different chain material qualities and coatings.

The comparison of oils is based on coefficient of friction (CoF) and wear scar diameter (WSD). The advantages of using WSD as wear indicator are, that it is simple and fast to measure, and it is comparable to standardized ball-on-disc tests, where WSD is the standard indicator [13]. Other opportunities are weighing the specimens and 3D scanning the wear scars, but the weighing in this order (tens to hundreds of micrograms) is unreliable (because of the residual oil and wear particles in case of imperfect cleaning), 3D scanning is very time-consuming and hard to compare to standardized tests.

The whole test procedure with the chains (the custom-made chain clamps, the test parameters and the wear measurement procedure) had to be developed in several steps in order to eliminate errors and reduce measurement uncertainty [21] [22].

As a result of the improvements, the uncertainty of the wear measurement (relative standard deviation of the WSD) was reduced from a value of above 50 % to below 10 % in all cases, moreover, in most cases it was below 5 % [23]. However, this required compromises in the measurement program. At the relatively high stroke (first 6 mm, later 4 mm) and high frequency (40 Hz) initially used, due to the construction of the reciprocating tribometer (component elasticity, oscillations caused by clearance of components), during the test, the moving clamp stochastically shifted with the specimen in lateral direction. Therefore, instead of the previously used 4 mm stroke and 40 Hz frequency, in the final version of the test procedure 1 mm stroke and 10 Hz frequency were used. To compensate the large reduction in distance traveled by the specimens and its wear, the 2-hour duration of the tests had to be increased to 22 hours.

With the new parameters, the uncertainty (standard deviation of WSD) became small enough to allow a comparative test on different oil samples.

The Lubrication was provided in a 15 ml oil bath, which completely covered the pin fixed in the stationary clamp. The test parameters are given in Table 3.

Table 3. Final parameters of the bush-on-pin test

stroke	***frequency***	***load***	***time***	***oil vol.***	***temperature***
1 mm	10 Hz	50 N	22 h	15 ml	100 °C

The chain components were from the timing chain of a passenger car's Diesel engine, which is a bush chain by type. The pins were chrome-plated and had a hardness of at least 1500 HV. The surface of the bushes was carbonitridated and their hardness was at least 700 HV.

For the tests, a new timing chain was disassembled, and a pin and bush were installed in the tribometer for each test, as shown in Fig. 5. Due to the large difference in hardness, wear occurred practically only on the bush, the wear of the pins was negligible and hardly measurable.

The coefficient of friction (CoF) was measured during the test, the wear scar diameter (WSD) on the bush was measured with a digital microscope after the test. To evaluate and compare the CoF values, the mean value of the last hour was always considered, because it was always stable enough at the end of the tests. The amount of wear was defined from the average of two diameters of the wear scar, see Fig. 6.

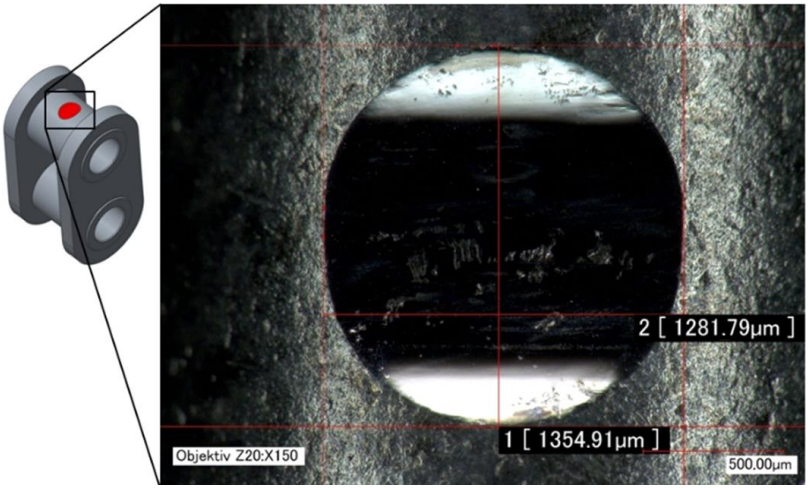


Figure 6. The two main dimensions of the wear scar on the bush

3.2.2. Results of the bush-on-pin measurements

The goal of the series of measurements was to compare different oil samples in terms of friction and wear between the chain elements. If a bush-on-pin tribometer test can detect a small but significant difference between new engine lubricants, it can be considered suitable for comparative tests. The results of the same oil samples obtained by the standard-based ball-on-disc test were used as a reference (see Section 3.1.2).

During the test series, the same lubricants were used as in case of the ball-on-disc test. Three different commercially available motor oils (Castrol 10W-60, Castrol 0W-20, Shell 0W-30) and three samples of contaminated Shell 0W-30 oil with 0.5 wt%, 1 wt% and 2 wt% carbon black (Cabot Monarch® 430).

The comparison of the coefficients of friction measured by the two methods is shown in Fig. 7.

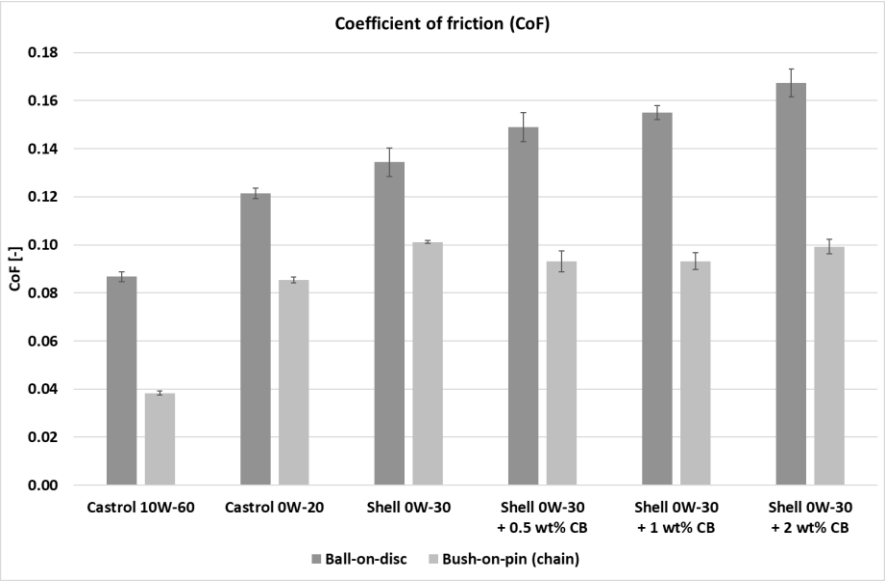


Figure 7. Comparison of coefficients of friction measured by both methods

Based on Fig. 7., it can be concluded that the bush-on-pin method reproduced the order of oils with minimal standard deviation. The standard deviation is even smaller in most cases than that of the ball-on-disc measurements taken as reference.

The best results in both test series came with the Castrol 10W-60 oil, as due to its high viscosity it is best able to separate the contact surfaces. This oil is specifically used for high-performance engines, so the result is in line with preliminary expectations.

Wear results showed a similar trend, see Fig. 8.

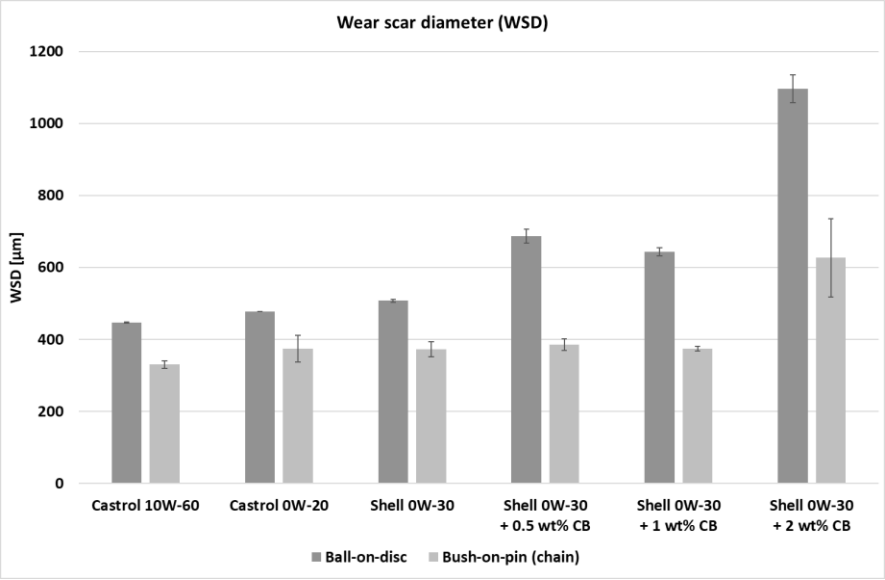


Figure 8. Comparison of wear scar diameters measured by two methods

In case of Castrol 0W-20 and Shell 0W-30 oils, the wear values are clearly higher than with Castrol 10W-60. There is no significant difference that can be detected between the two low-viscosity oils, wear results of these are in the same range in case of bush-on-pin, with higher relative standard deviation values (10,1 % by Castrol 0W-20 and 5,7 % by Shell 0W-30). With bush-on-pin test, small differences cannot be detected so clearly compared to ball-on-disc method. What can be clearly stated is that in the case of Castrol 10W-60 oil the wear value is significantly lower and the standard deviation is also sufficiently small (3,1 %).

With oils mixed with carbon black – similar to ball-on-disc tests – the wear with 1 wt% carbon black was minimally less than with 0.5 wt%, however, the difference is not significant. In fact, it is not significant compared to the wear measured with the unblended sample of this oil. Wear increased significantly only with the addition of 2 wt% carbon black, similar to ball-on-disc measurements.

Based on these results, it can be concluded that with the present test parameters of bush-on-pin test method, the effect of a small difference between the oils (small difference between viscosity classes and additives) can be detected in wear, but not as clearly and definitely as with a ball-on-disc test. However, in the case of a larger difference in viscosity or in additive package, as well as in the case of high carbon black or soot content, a significant difference can be detected.

In turn, the differences in the coefficient of friction between the pure oils can also be detected definitely and with negligible uncertainty by both of the test methods.

4. Conclusions

The results of tribological experiments with chain elements show a similar trend in both coefficient of friction and wear as the ball-on-disc measurements. Minor differences can be attributed to the differences between the two tribological systems, such as geometry, material, method of lubrication, and motion conditions of the specimens.

The bush-on-pin method, just like the ball-on-disc, is suitable for evaluating the effect of oil quality, contamination and other factors on friction, wear and the life expectancy of chains.

Because the bush-on-pin method proves to be reliable when compared to the ball-on-disc method, it can also be used in cases where ball-on-disc testing is not possible. In case of simple oil ranking, the ball-on-disc test is more suitable because it is much simpler and it is a standard-based test. The bush-on-pin method is expected to have a promising potential in evaluating chain materials, manufacturing technology, geometry, surface coating and quality. It would be quite expensive to produce standardized tribometer test specimens with the original material, coating and heat treatment, surface characteristics of the chain. The tribometer specimens would behave differently during heat treatment than the chain elements.

It can be concluded that it is worth using the chain measurements on tribometer to make the development of chain drives more cost-effective and environmentally friendly.

References

[1] T. Fink, V. Hirschmann, Kettentriebe für den Einsatz in modernen Verbrennungsmotoren. *MTZ – Motortechnische Zeitschrift* 62, (2001), pp. 796–806., in German
doi: https://doi.org/10.1007/BF03227087

[2] T. Fink, K. Elwan, Steuerkettentriebe mit innovativen Lösungen zur Steigerung der Leistungsfähigkeit, in: Kay Schintzel (Ed.) *4. Győrer Tribologie Tagung*, Győr, Universitas-Győr Nonprofit Kft. (2016), pp. 47-53., in German

[3] H. Schwarze, L. Brouwer et al. Lubricant degradation and wear behaviour in a spark-ignition engine. *MTZ Worldwide* 69 (2008) pp. 60–67 doi: https://doi.org/10.1007/BF03226922

[4] T. Gergye, M. R. Dreyer et al., Analysis of the Wear Behavior of Combustion Engine Components Using Radionuclide-Technique. in: SAE-China, FISITA (Eds.) *Proceedings of the FISITA 2012 World Automotive Congress* (1) (2012) pp. 171-181. doi: https://doi.org/10.1007/978-3-642-33841-0_13

[5] T. Gergye, M. R. Dreyer, Tribologische Untersuchung einer Steuerkette, in: Dreyer, Mathias Roman; Demmelbauer-Ebner, Wolfgang (szerk.) *2. Győrer Tribologie Tagung*, Győr, Universitas-Győr Nonprofit Kft., (2012) pp. 173-186., in German

[6] A. Becker, D. Meffert, B. Sauer, Friction and wear investigations on single chain joints. *Forschung im Ingenieurwesen* 83, (2019) pp. 53–63. doi: https://doi.org/10.1007/s10010-019-00297-x

[7] H. Czichos, K. Habig, Tribologische Mess- und Prüftechnik, in: Tribologie-Handbuch, 3. Auflage, *Vieweg + Teubner* (2010) ISBN 978-3-8348-0017-6, pp. 193-195, in German

[8] A. Gummer, Cs. Fábián, B. Sauer, Experimental Investigation of Roller Chain Wear, *18th International Colloquium Tribology*, Technische Akademie Esslingen, 10-12 January 2012, Ostfildern

[9] A. Gummer, Cs. Fábián, B. Sauer, Modularer Kettenprüfstand für Verschleiß- und Wirkungsgraduntersuchungen, *52. Tribologie-Fachtagung*, 26.-28.09.2011, Göttingen, in German

[10] D. Sappok, A. Gummer, B. Sauer, Experimental and analytical wear investigations of bush- and roller chain drives. *5th World Tribology Congress, WTC 2013*. (2) (2013) pp. 1204-1207.

[11] Standard Test Method for Wear Preventive Characteristics of Lubricating Grease (Four-Ball Method), ASTM D2266 - 01(2015)

[12] Standard Test Method for Wear Preventive Characteristics of Lubricating Fluid (Four-Ball Method), ASTM D4172 - 20

[13] Lubricants — Determination of tribological quantities for oils and greases — Tribological test in the translatory oscillation apparatus, ISO 19291:2016

[14] Á. D. Tóth, Á. I. Szabó, R. Kuti, Tribological Properties of Nano-Sized ZrO2 Ceramic Particles in Automotive Lubricants, *FME Transactions* 49 (1) (2021) pp. 36-43.
doi: https://doi.org/10.5937/fme2101036T

[15] D. Uy, M. A. Ford et al., Characterization of gasoline soot and comparison to diesel soot, Morphology, chemistry, and wear, in: *Tribology International* 80, (2014) pp. 198-209.
doi: https://doi.org/10.1016/j.triboint.2014.06.009

[16] S. Antusch, M. Dienwiebel et al., On the tribochemical action of engine soot, in: *Wear* 269 (1–2) (2010) pp. 1-12.
doi: https://doi.org/10.1016/j.wear.2010.02.028

[17] E. Hu, X. Hu, et al., The role of soot particles in the tribological behavior of engine lubricating oils, *Wear* 304 (1–2) (2013) pp. 152-161
doi: https://doi.org/10.1016/j.wear.2013.05.002

[18] A. La Rocca, G. Di Liberto et al.: The nanostructure of soot-in-oil particles and agglomerates from an automotive diesel engine, in: *Tribology International* 61 (2013) pp. 80-87
doi: https://doi.org/10.1016/j.triboint.2012.12.002

[19] M. Ratoi, R. Castle et al., The influence of soot and dispersant on ZDDP film thickness and friction. in: *Lubrication Science* 17 (1) (2006) pp. 25–43.
doi: https://doi.org/10.1002/ls.3010170103

[20] Á. D. Tóth, L. Paulovics et al., Methodenentwicklung zur Einstufung von Motorölen anhand tribologischer Eigenschaften, in: *58. Tribologie-Fachtagung 2017: Reibung, Schmierung und Verschleiß, Forschung und*

praktische Anwendungen, Gesellschaft für Tribologie, Göttingen, Aachen, (2017) P8, in German

[21] L. Paulovics, Á. D. Tóth et al., Tribometerversuche an Steuerkettenteilen, in: *58. Tribologie-Fachtagung 2017: Reibung, Schmierung und Verschleiß, Forschung und praktische Anwendungen*, Gesellschaft für Tribologie, Göttingen, Aachen, (2017) P9, in German

[22] L. Paulovics, M. Németh, J. Knaup, Alternative Verschleissmessverfahren für Steuerketten, in: Knaup, Jan (Ed.) *5. Győrer Tribologie- und Effizienztagung*, Győr, Universitas-Győr Nonprofit Kft., (2018) pp. 199-216., in German

[23] Zs. M. Tabakov, L. Paulovics, Development of wear analysis of timing chain elements on a high frequency reciprocating rig, In: A. Dernóczy-Polyák (Ed.) *Kutatási jelentés 2. - Research Report*, Győr, Universitas-Győr Nonprofit Ltd. (2020) pp. 248-252.

Acta Technica Jaurinensis

Vol. 14, No. 4, pp. 424-439, 2021
DOI: 10.14513/actatechjaur.00641

Integration of BIM in architecture and structural engineering education through common projects

K. Ajtayné Károlyfi[1,*], D. Szalai[2], J. Szép[1], T. Horváth[2]

[1]Széchenyi István University, Department of Structural and Geotechnical Engineering, Egyetem tér 1, 9026, Győr, Hungary
[2] Széchenyi István University, Department of Architecture and Building Construction, Egyetem tér 1, 9026, Győr, Hungary
*e-mail: karolyfi.kitti@sze.hu

Submitted: 23/07/2021; Accepted: 24/08/2021; Published online: 30/08/2021

Abstract: Building Information Modeling (BIM) is one of the most significant developments in architecture and civil engineering in recent years, therefore it becomes increasingly important to promote its integration into university education. Currently, several universities worldwide offer BIM courses in architecture and civil engineering programs, while many others are under the process of integrating BIM into their curricula. The goal of this research is the implementation of BIM into higher education (Széchenyi István University, Hungary) by integrating the architectural, structural, and mechanical engineering disciplines. An important advantage of the university is that the architecture and civil engineering programs belong to the same faculty, which allows students to work together on common projects. A new course was performed in the previous semester, in which the students are designing contemporary buildings as case studies. In this paper, the results of the integration process are presented and evaluated based on the trainers' and students' experiences.

Keywords: building information modelling; collaboration; higher education; project-based learning

1. Introduction

The building industry of our accelerated world has demanded the digitization of architectural design and construction organization. Early digital design has begun to apply 3D modeling and then linked more and more information to the elements of the 3D model. In this way, Computer-Aided Design (CAD) has imperceptibly transformed into Building Information Modeling (BIM). The industry felt the effectiveness of the new method and found several other benefits as well. Thus, the newly formulated BIM concept has begun to reform the architectural design and construction. Standardization of the BIM methodology has begun worldwide [1]. In several countries, BIM-based planning is becoming mandatory depending on project size or funding. The Hungarian BIM national standard [2] was released in 2019, in which the application of this design method is highly recommended. Therefore, to prepare the students for the professional challenges of real life, the education must integrate this newly emerged methodology.

The Széchenyi István University (Győr, Hungary) has an important advantage because the architecture and civil engineering programs belong to the same faculty, which gives the possibility for students to work together on common projects. This research aimed to implement the BIM method into the curricula by integrating these two disciplines. In this paper, the integration process is presented and its advantages are evaluated and compared to the traditional building design education.

The university is a quite young educational institution, its ancestor was established in 1968. Civil engineers have been educated since 1974, and architects since 1990 (Fig. 1.). For today, the university has become an important regional higher education institution with nine faculties, and it makes efforts to increase its international recognition.

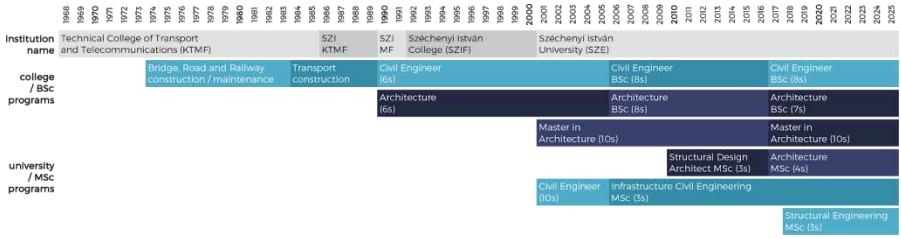


Figure 1. Civil engineer and architecture programs of the Széchenyi István University

Reformed programs started in 2017, as a result of a countrywide review of higher education. In the new curricula of the civil engineering and architecture programs, the education of the BIM method was inserted. Several courses refer to the different

aspects of BIM, and dedicated courses were also introduced such as Structural BIM for civil engineers and BIM management for architects. Both courses belong to the 7th semester, which offered a great opportunity for joint training of the two disciplines. Teaching BIM requires a project approach and intensive teamwork. These techniques were used in both programs, but they appeared most intensively in the course of the Building Structure Design Project. We have been teaching these subjects for more than 10 years with good experiences. The BIM approach has often appeared in students' work, due to their interest [3]. In the new BIM course this approach has become mandatory and the key principle of the design process.

2. Literature review

There is a clear industry demand for engineers with advanced BIM skills, therefore its integration into higher education is crucial. There are three main ways for the implementation: (1) introducing BIM as an optional course or extracurricular activity, (2) introducing new BIM-related courses or implementing it into existing courses, (3) introducing an advanced BIM-related program, or any combination of the above. The implementation of BIM into academic programs is at different levels around the world. In the UK, many universities offer BIM-related courses in their AEC programs and at least 14 institutions have specific education programs due to the decisive government efforts to use BIM in public building projects [4, 5]. In Australia, universities mainly offer BIM courses with special attention to teaching software skills. Similarly in the USA, more than half of the institutions offer courses with embedded BIM content [5]. In Hungary, the integration of BIM into AEC programs has begun in recent years, and already two institutions offer dedicated BIM programs.

The current practice of BIM education is very diverse according to the wide range of disciplines involved. In architecture education the documentation and visualization tools of the BIM is enhanced, while the structural engineering discipline is strongly connected to structural analysis and design software, the management discipline is more involved in the collaborative platforms of BIM and the construction-related subjects are usually dealing with scheduling, quantity statement and cost estimate. [6]. According to an extensive study of the NATSPEC [7], the proper use of the authoring software plays an important role in the current BIM education, while the implementation of open BIM and collaborative working environment is not receiving all the attention that it requires. Wang et al. [8] emphasizes the importance of the connection between BIM and other courses and recommend the adoption of BIM in the final semester project to integrate the previously acquired competencies. Based on the review and categorization of 70 publications from 24 countries made by Badrinath et al. [9], there is a clear trend

towards the development of BIM curricula and educational frameworks, in which multidisciplinarity becomes a key factor.

The effective education of BIM methodology can be achieved by different methods such as self-study, web-based learning, project-based learning, lab tutorials, or lectures [4]. Baradi et al. [10] investigated studies on teaching collaboration within BIM education and identified the best practices. Results show that project and team-based learning is an effective way to teach collaboration skills, where students are assigned a project task to be fulfilled during the course. The learning process is the most successful when the task comes from live projects, where the entire design documentation is not available and it requires both team and independent work from the participants by considering alternative solutions as well. In this methodology, the trainers' role is mainly mentoring and guiding the process instead of teaching professional knowledge, where the involvement of industry mentors can contribute significantly to the development of practical skills [4,11].

3. Research method

The research method includes three main steps: first of all, (1) the current education method of architectural and structural design was explored and evaluated by considering the industrial needs. Secondly, (2) the benefits of the BIM method were identified and the possibilities of implementation into the curricula were outlined by performing an educational pilot project. Finally, (3) the results of the implementation were evaluated by presenting the trainers' experiences and the students' opinions based on a questionnaire.

4. The current state of architectural and structural design education

The design-related courses of the university's architectural program are mainly using the traditional two-dimensional drawing methods. Generally, two main types of design tasks are present: (1) the entire building is designed by the student, or (2) an existing building is processed and redesigned based on the available plans published in a journal or on the Internet. In both cases, creating manual drawings is necessary for the beginning to develop or to understand the house, especially when designing an entirely new building. However, during the detailed design, the software is still used for creating two-dimensional architectural plans and technical specifications. The three-dimensional modeling is only a possibility for the students, which is mainly used for creating low-detailed visualization of the building independently from the plans.

The structural design courses of the civil engineering program have a similar approach: in general, an existing building is processed based on the available architectural plans from different sources. The task begins with performing the preliminary two-dimensional structural plans, then a simplified finite element model is created, which is mainly planar or in some cases three-dimensional. The results of the structural analysis are mainly two-dimensional formwork and reinforcement plans, drawings of the structural joints, and construction plans.

The issues relating to other disciplines are usually discussed with the trainers of the course and consultation with specialists is also provided if necessary. Therefore there is no collaboration between students from different domains. However, the current practice requires strong cooperation between the project participants, which is rooted in the understanding of the significance of other disciplines [12]. With the implementation of the BIM method into higher education, these skills and abilities can be improved.

5. Implementation of BIM into the curricula

The basic idea of BIM is to create a common platform for cooperation among the disciplines [13]. The BIM model is the digital representation of the building, which is working as a shared knowledge resource for all of the project participants during the building's life cycle. In general, the architectural model is the basis of the project, which is used by the different professionals for their work. Each discipline adds specific data to the shared model, which information is directly accessible for the entire design team. Due to the nature of the method, BIM enables strong collaboration between the team members, which contributes to developing interdisciplinary thinking and skills. In 2020 the Structural BIM and BIM management courses were integrated for the civil engineering and architecture students to broaden their perspective and advance their understanding of the other's discipline. This educational pilot project was a simulation of the BIM working method, where the students are creating a common model, which will be the basis of further discipline-specific work. In this way, the civil engineering students gain skills in building construction design, while the architecture students learn how to create a model capable for structural analysis.

The education method was based on a mentoring scheme including short presentations connected to the main steps. The completed subtasks were discussed and evaluated in common lessons involving trainers from each discipline. On the trainer side, two BIM experienced architects (Kitti Ajtayné Károlyfi and Dóra Szalai), a building construction expert architect (Tamás Horváth), a structural and a mechanical engineer (János Szép, László Petrikó) helped the work. Students presented their work in progress plans every two weeks and received feedback and

they had the possibility of weekly consultations. After the 10th week, the education was implemented in online form due to the epidemiological situation, however, the original system was kept.

In the newly introduced lecture, different buildings were selected from a construction-oriented architectural journal [14,15,16] to ensure feasibility and give an interesting architectural and structural challenge to the students. Three groups were formed of four members including at least one architecture student and two or three civil engineering students. Each member was in the last or one before the last semester of their bachelor training.

The first group has chosen the Salt Office Building in Amsterdam (Fig. 2.a) designed by MVRDV [14], which is made of prefabricated reinforced concrete frame structure and framed facade panels with variously segmented glazing. The second building was the Freemen's School Swimming Pool (Fig. 2.b) near London [15] constructed from a large span laminated glued frame structure starting from a basement level with reinforced concrete structure. A smaller project, a single-family house (Fig. 2.c) in Tuttlingen [16] was chosen by the third group. It has an interesting asymmetric form, where the load-bearing structure is based on two reinforced concrete cross-walls, but the building envelope was made of lightweight timber construction.

In the course assignment, student groups had the task to design these buildings. Understanding the building began with the help of draft plans. From there, continuous development was performed until the detailed construction plans.

Figure 2. a) Salt office building in Amsterdam, designed by MVRDV, photo: Ossip van Duivenbode [14], b) School swimming pool near London, designed by Hawkins\Brown, photo: Jack Hobhouse [15], c) Single-family house in Tuttlingen, designed by Yonder, photo: Brigida González [16]

The course description divided the whole design process according to the usual development stages designating both individual and group tasks. The design task was divided into the following parts:

1. Creating the architectural and structural model

2. Improving LOD of the model
3. Structural analysis
4. Design of building structural details
5. Collision detection, quantity statements, and consignments
6. Preparation of the final documentation of the project.

The modeling process was started as a common task for each group member. The architecture students were responsible for the architectural and mechanical engineering model elements, while the civil engineering students for the load-bearing elements. From the 3rd step, the two disciplines had distinct subtasks, in which the feedback and communication have become even more important.

5.1. Creating the architectural and structural model

For the modeling of the buildings, the ArchiCAD 23 software was used. Three projects were created and shared on a BIMcloud server and the students were connected to the projects by using the teamwork function. The minimum level of information content for each object was the use of correct object type, position, material characteristic, and load-bearing function. This phase aimed to generate a model and automatic building permit design documentation at LOD 200, which means that the elements are modeled with approximate size, shape, and location (Fig. 3.).

Figure 3. a) Model of the Salt office building [17], b) Model of the school swimming pool [18], c) Model of the single-family house [19]

5.2. Improving LOD of the model

The second step aimed to further development of the architectural and structural model to reach the LOD 300, in which the elements have the correct size, shape, and location and the main information content is also attached. Civil engineers prepared the model for export. Furthermore, a mechanical engineering model was also created including heating, plumbing, and air-conditioning systems at LOD 200 (Fig. 4.).

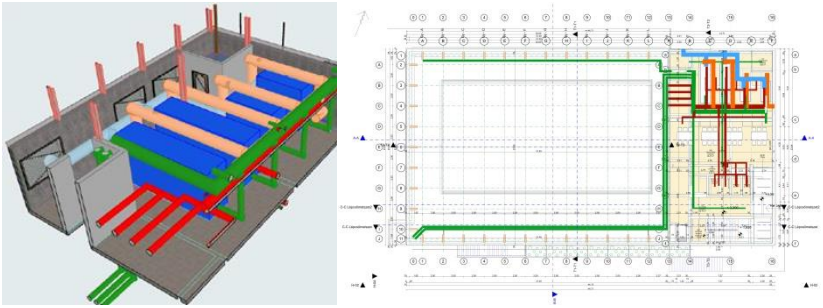


Figure 4. a) MEP model part and b) construction plan of the swimming pool [18]

5.3. Structural analysis

In the third phase, the civil engineering students have exported the structural models from ArchiCAD (Fig. 5.a) using an IFC translator. After the determination of the loads, the AxisVM software was used for the structural analysis, in which the cross-sectional dimensions and deformations of structures were checked (Fig. 5.b) and the reinforced concrete elements were designed in detail. After calculating the required reinforcement area, the used relative steel quantities were imported into ArchiCAD as an IFC attribute for each object. After that, students were able to create reinforcement quantity estimation for the whole building and develop the reinforcement plan of the main elements.

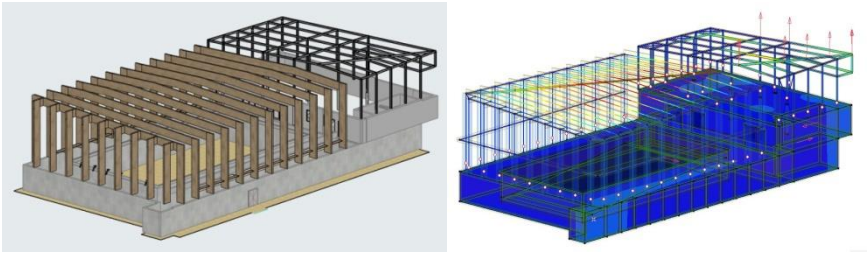


Figure 5. a) Structural model of the swimming pool in ArchiCAD 23, b) Result of the structural analysis in AxisVM [18]

5.4. Design of building structural details

After updating the model with the correct cross-sections based on the structural analysis, the detailed design of the construction joints was started. Architecture students worked on the building structural details (Fig. 6.a), while civil engineering students made calculations and plans for the main structural joints (Fig. 6.b).

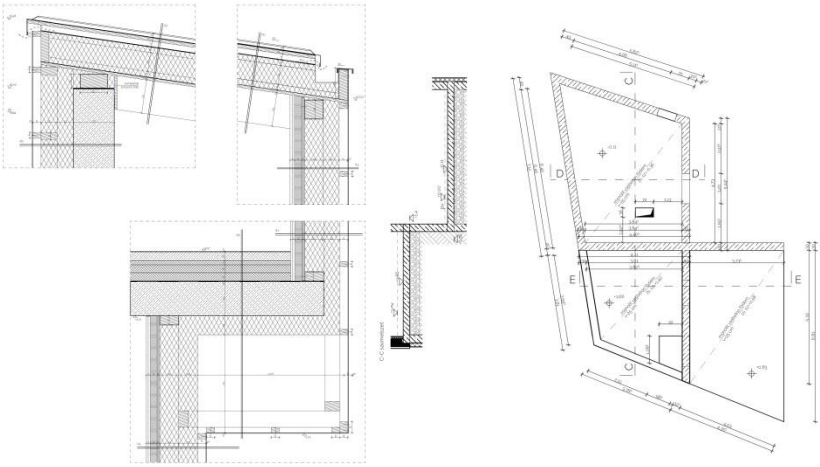


Figure 6. a) Detailed plan of the joints and b) formwork plan of a single-family house [19]

5.5. Collision detection, quantity statements, and consignments

In this phase, students got individual subtasks to investigate the model using different aspects. The collision detection was made between the architectural, structural, and mechanical engineering elements by using the Solibri Office software. Based on the results the detected errors were reported and imported back into ArchiCAD using .bcf format and the collisions were corrected. Furthermore, quantity statements were generated for the main structural materials and consignments of the doors, windows, and curtain walls, using the automatized BIM solutions of ArchiCAD. Fabrication plans were generated for the prefabricated concrete structural elements and facade frames of the office building and the timber frame structure of the swimming pool.

5.6. Preparation of the final documentation of the project

The final documentation of the projects included the completed architectural plans, the structural analysis, the structural drawings, the collision detection, the quantity statements, and the consignments. During the course, several deficiencies were revealed by the examinations, therefore this final documentation was requested to be submitted and presented in the exam period, leaving some weeks to complete the designs. The groups took the opportunity and improved a lot on their projects, so they could present really good quality work at the end of the semester.

6. Discussion

6.1. Opinion survey

To improve the subject in the future, trainers wanted to get feedback from the students based on their experiences, therefore a customized survey was performed in the form of an online questionnaire. The survey was conducted among the students who participated in the course. The questionnaire was received by all students and 10 of 12 students also answered anonymously. The 13-item-long survey focused on the effectiveness of BIM education and the impact of the BIM approach on workflow. The following topics were included in the questionnaire:

- ArchiCAD user skills at the beginning and the end of the course;
- the usefulness of the complex collaborative task;
- the effect of the BIM processing method
- on team communication;
- on the production of design documentation;
- on the structural analysis;
- on the preparation of quantity statements;
- on the design of consignments/element plans;
- on finding and correcting faults;
- the difficulties of applying the BIM working method;
- understanding and experience expansion in BIM during the semester;
- individual ideas for improvement.

According to their first experiences, students find that making the quantity statements, identifying, correcting errors, and designing consignments can be done easier with BIM methods than with the traditional way (Fig. 7.). It is a great achievement because most of the students encountered this processing method this semester firstly, and their software background was also low. Even without special training and practice, this BIM method made the exercises considerably easier.

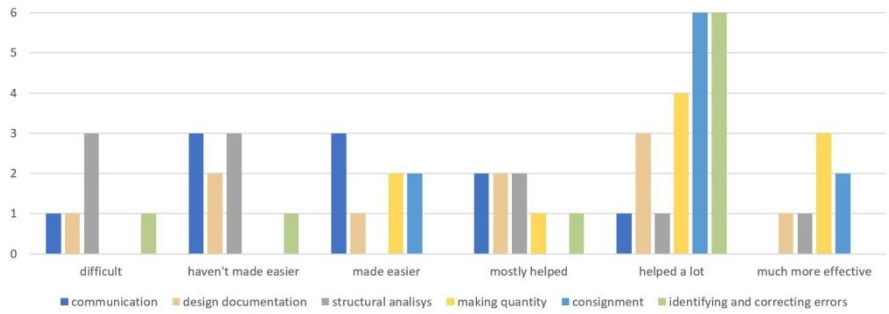


Figure 7. Effect of BIM processing method on the workflow according to the experiences of the students

In the case of interdisciplinary communication, the students' opinions were divided. Some of them thought that this method did not help communication, but others thought that it helped a lot. In this university assignment, it was possible to simulate live project situations, which are often cumbersome and challenging. The pandemic period and the switching to online training forced intensive communication among group members. This problem prepares students also for real life, where working in distance is becoming increasingly necessary.

In the case of preparing the design documentation, the opinion was also controversial. Probably, those who worked on a larger-scale building may have found this job easier. In their case, there was a lot of repetitive work and the BIM method shows a return relatively quickly due to the scale. [20] On the other hand, those who found this method more difficult, probably used fast, non-transparent processing methods in the past (e.g. fill-in corrections on sections, floor plans, documentation hierarchical neglect, etc.). For them, it was more difficult to apply the controlled design processing.

Almost all civil engineers found the BIM-based structural analysis more complicated than the usual method. The biggest challenge was the direct use of the architect model to make the structural analysis. In their previous tasks, they did not depend on the work of others, and they were able to proceed at their own pace and discretion. However, now they also had to deal with the model to be imported, which was a completely new job for them. Importing the BIM model back and forth required close cooperation between the two disciplines to transfer the required data.

Students generally reported that their CAD/BIM skills improved significantly during the course (Fig. 8.). The real-life-based assignment, also the contemporary buildings, and the complex design task were welcomed. This was the first time at the university when architects and civil engineers could work together on one

common project. All students evaluated the course positively, they also found it useful to see processes like real planning.

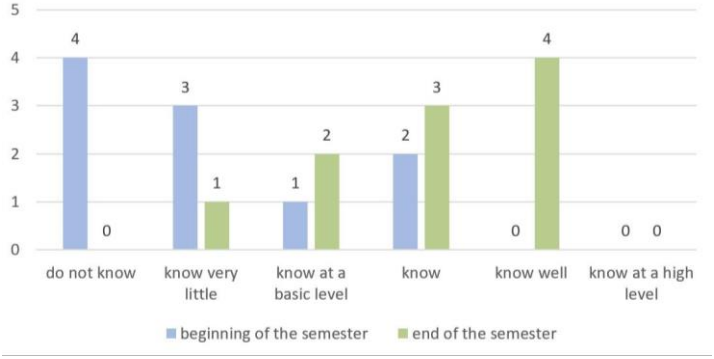


Figure 8. Students' CAD/BIM knowledge at the beginning and end of the semester

6.2. Trainers' experiences

However, this new course has some antecedents at the university, organizing a common course for architects and civil engineers was a challenging task and needed a serious collaboration of the departments and trainers. The pioneering course can be concluded as a successful training, but all trainers see the possibilities of improvements, which can be implemented in the next few years. Students attending the course were senior students, but they arrived with different knowledge, experiences, and software skills. One of the biggest problems was the deficiencies in the software skills. The assignment expected a quite good level of ArchiCAD use, but this was not given in the case of all students. ArchiCAD was educated for architects in a compulsory course, but civil engineers can meet this software only in an optional course. These differences should be considered in the future. The modification of the curricula is difficult, therefore the problem may be solved by a course-integrated intensive ArchiCAD training. Experiences of the BIM course can be also considered as feedback for the ArchiCAD courses, where the importance of accurate, rule-following, professional modeling, and design workflow should be emphasized. Practiced software users can also have problems if they are accustomed to improper practice.

In this very first course, the proportions of the distribution of students were not ideal, civil engineers were overrepresented in the groups. In an ideal situation, the members of the two disciplines are balanced. Design of building construction projects employs fewer civil engineers than architects. Hopefully, the ratio will change in the future when more architects will reach the last semester of the new

BSc training. It was also experienced both from the students' and trainers' side, that BIM is an ideal concept that can help in the design of the built environment, but theory needs a lot of practical solutions. The integration of the BIM aspect into the design is an ongoing process, which requires and holds significant research results soon. These results must be implemented in practical solutions and built into the software. Users, such as designers, must have the intent to learn these new techniques and properly use them. Important learning was that BIM competencies had to be defined by the trainers, and after that, the aims and methods of education can be adapted to the competencies [21].

BIM development levels can be different in different projects or even inside in the case of different subtasks. These levels should be chosen carefully, according to the available software and hardware endowments, and last but not least, according to the available intellectual capacity and labor force.

An important advantage of the course arrangement was the continuous feedback from the development of the projects and the contribution of the participants. The modeling was a common task at the beginning for all team members, which resulted in a higher contribution of the architecture students. After the individual tasks were defined and the students were accountable for their assignments, the contribution of the team members became more balanced. Therefore, these roles and responsibilities should be defined more precisely in the future at the beginning of the course.

7. Conclusions

It can be concluded, that the cooperation between the architect, structural and mechanical engineer can be simulated in the university environment, therefore the implementation of BIM into the curricula was successful. Comparing to the traditional design method, with the help of the BIM students get a better understanding of the building and the aspects of the connecting disciplines, though not without its challenges. While in the case of the traditional method the revision of changes in the plans is the most time-consuming task, with the BIM method this work can be made automatically based on the model. Here, the greatest challenge is to create an appropriate model according to the project's purposes and to find the proper way of data exchange between the disciplines. Once this task has been achieved, the BIM method provides many possibilities for further use; the introduced course covers automatic documentation, collision detection, quantity statements, and consignments, therefore it has many development opportunities. In the future, it is also aimed to create cost estimation, use the model in construction management and analyze the efficiency of the different solutions. These applications have even greater importance in the industry, therefore their integration into higher education improves the students' competencies and prepare them for the current challenges of

the industry. It can be concluded, that students' interests were raised in a developing area that will certainly have an impact on their work. Some of them started working on a BIM-related topic as a thesis work, which gives them more freedom to explore the common points between disciplines.

References

[1] K. Zima, E. Plebankiewicz, D. Wieczorek, A SWOT analysis of the use of BIM Technology in the Polish construction industry, *Buildings* 10 (1):16 (2020)
doi: https://doi.org/10.3390/buildings10010016

[2] MSZ EN ISO 19650:2019 Organization and digitization of information about buildings and civil engineering works, including building information modeling (BIM). Information management using building information modelling

[3] T. Horváth, Project-based education of building construction design for architects, *Hazai és külföldi modelek a projektoktatásban*, Óbuda University, Budapest, Hungary (2017) pp. 134-142, in Hungarian
URL https://www.researchgate.net/publication/317193175

[4] A. Shibani, S. B. Souliman, BIM Implementation Strategies in Higher Education (Case Study Coventry University) pp. 406-420.
URL http://www.ieomsociety.org/ieomuk/papers/92.pdf

[5] A. Abbas, Z. U. Din, R. Farooqui, Integration of BIM in construction management education: an overview of Pakistani Engineering universities, International Conference on Sustainable Design, Engineering and Construction, *Procedia Engineering* 145 (2016) pp. 151-157.
doi: https://doi.org/10.1016/j.proeng.2016.04.034

[6] X. Zhang, R. Jin, T. Yang, Review of BIM Adoption in the Higher Education of AEC Disciplines, *Journal of Professional Issues in Engineering Education and Practice* 146 (3):06020001 (2020)
doi: https://doi.org/10.1061/(ASCE)EI.2643-9115.0000018

[7] K. Rooney, BIM Education – Global Summary 2015 Update Report, Sydney, Australia, *NATSPEC Construction Information*, 2015.
URL https://icis.org/wp-content/uploads/2017/11/ICIS2015-BIM_Education_Global_2015_Update_Report_V2.0.pdf

[8] L. Wang, X. Yan et al., Incorporating BIM in the Final Semester Undergraduate Project of Construction Management-A Case Study in Fuzhou University, *KSCE Journal of Civil Engineering* (2020) in PressDetail, 5 (2019).
doi: https://doi.org/10.1007/s12205-020-1971-4

[9] A. C. Badrinath, Y. Chang, P. Hshieh, A review of tertiary BIM education for advanced engineering communication with visualization, *Visualization in Engineering* 4:9 (2016).
doi: https://doi.org/10.1186/s40327-016-0038-6

[10] K. Baradi, M. Oraee et al., Teaching collaboration in tertiary BIM education: A review and analysis, *The 42nd Australasian Universities Building Education Association (AUBEA) Conference Proceedings*, Singapore, 2018.
URL https://minerva-access.unimelb.edu.au/bitstream/handle/11343/230682/Oraee.pdf?sequence=2&isAllowed=y

[11] K. Chen, W. Lu, J. Wang, University–industry collaboration for BIM education: Lessons learned from a case study, *Industry and Higher Education* 34 (6) (2020) pp. 401-409.
doi: https://doi.org/10.1177/0950422220908799

[12] C. Merschbrock, M. R. Hosseini et al., Collaborative Role of Sociotechnical Components in BIM-Based Construction Networks in Two Hospitals, *Journal of Management in Engineering* 34 (4):05018006 (2018).
doi: https://doi.org/10.1061/(asce)me.1943-5479.0000605

[13] U. Poerschke, R. J. Holland et al., BIM collaboration across six disciplines (2010) *Proc. of the Int. Conf. on Computing in Civil and Building Engineering*, Nottingham University Press.
URL https://www.researchgate.net/publication/280087526

[14] MVRDV, Office building in Amsterdam, *Detail* 5 (2019) pp. 56-59.

[15] Hawkins/Brown, School swimming pool near London, *Detail* 7/8 (2019) pp. 48-53.

[16] Yonder – Architektur und Design, Single-family house in Tuttlingen, *Detail* 10 (2019) pp. 54-59.

[17] E. Szabó, E. Schram et al., Construction plans for the MVRDV: Office building in Amsterdam, 2021.

[18] Zs. Kalmár, P. F. Módy et al., Construction plans for the Hawkins/Brown: School swimming pool near London, 2021.

[19] Sz. Dugár, M. Bak et al., Construction plans for the Yonder: Single-family house in Tuttlingen, 2021.

[20] A. J. Garcia, S. Mollaoglu, M. Syal, Implementation of BIM in Small Home-Building Businesses *Practice Periodical on Structural Design and Construction* 23 (2): 04018007 (2018).
doi: https://doi.org/10.1061/(ASCE)SC.1943-5576.0000362

[21] S. Kolarić, T. Mandičák et al., BIM training in construction management educational practices in Croatia and Slovakia, *Creative Construction Conference*, Ljubljana, Slovenia, 2018, pp. 1002-1009.
doi: https://doi.org/10.3311/CCC2018-130

Acta Technica Jaurinensis

Vol. 14, No. 4, pp. 440-454, 2021
DOI: 10.14513/actatechjaur.00599

Analysis of driver behavior in Amman using Manchester Driver Behavior Questionnaire

K. Jadaan[1], N. Albeetar[2,*], D. Abuhalimeh[3], Y. Naji[4]

[1]University of Jordan, Department of Civil Engineering
Amman, 11942, Jordan
[2] Bitar Consultants, Department of Civil Engineering
Amman, 11196, Jordan
[3]Collaboration Management & Control Solutions, Department of Civil Engineering
Amman, 11180, Jordan
[4]Arab Business Corporation, Department of Civil Engineering
Amman, 11942, Jordan
[*]e-mail: bitarnoor96@gmail.com

Submitted: 23/02/2021; Accepted: 21/04/2021; Published online: 30/04/2021

Abstract: A key component in combating traffic accidents is to study the contributory factors behind them, among these factors, the driver behavior stands out as the main causative factor. One of the most effective tools used worldwide in measuring self-reported driving components is the Manchester Driver Behavior Questionnaire (DBQ), it investigates the relationship between the driver and accidents involvement, throughout the analysis of both sociodemographic characteristics of drivers, and the risky driving components practiced such as; violations, errors and lapses. The present study investigates the factor structure of the DBQ and examines the relationships between the driver behavior factors and accident involvement. A survey questionnaire including the DBQ and background information was filled by a randomly selected sample of drivers in Amman, the capital of Jordan and the Statistical Package for Social Sciences (SPSS) software was used for data analysis. Driver behavior differed according to the gender, educational level and driving experience of the respondents. The results reflected the lifestyle, way of thinking and the general attitude of the driver and its relationship with traffic safety.

Keywords: driver behavior; Manchester Driver Behavior Questionnaire; traffic safety; Jordan

1. Introduction

Road Traffic Accidents (RTA) may cause severe physical losses and disabilities, hold back economic growth especially in developing countries and affect the country's Gross Domestic Product (GDP) negatively. In fact, RTAs cost most countries about 2-3% of their GDP. Traffic accidents are increasingly being recognized as a major cause of death and a growing health problem, Significant part of the road traffic accidents are originated in inappropriately selected speeds or specifically to speeding [1]. Accident issues can be recognized: approximately 1.3 million crashes annually lead to nearly 40,000 fatalities and more than 1 million injuries [2].

Manchester DBQ (Driver Behavior Questionnaire) was used to measure the aberrant driving behaviors leading to accidents [3]. Hence, driver behavior is regarded by researchers as the main contributor to RTAs. Furthermore, Human factors in driving can be seen as being composed of two separate components, driving style and driving skills [4]. The effect of specific aspects of driver characteristics and behavior on accident occurrence were investigated.

In Jordan, RTA are considered as a serious problem with an estimated yearly cost of about $500 million (about 2.5% of the country's GDP). This cost has increased 7.4% in the past five years [4]. According to the Jordan Traffic Institute (JTI), RTAs resulted in 571 deaths and 16203 injuries in 2018, and driver is the main contributor to the causation of RTAs (about 98% of total contributors). As such, a recommended road safety strategy was developed that focused on the role of driver characteristics and behavior [5]. A most recent study investigated the driver behavior relative to drivers' gender, while the current study measures the accident involvement in accordance with different driver characteristics studied in the questionnaire [6].

Manchester Driver Behavior Questionnaire (DBQ) is one of the commonly used tools in traffic psychology for measuring self-reported driving style and investigating the relationship between driving behavior and accident involvement. [7]

DBQ contains three subscales to capture different aspects of driver behavior; violations, errors and lapses [8]. Violations are defined as the behaviors which endanger traffic safety [9], such as driving under the influence of alcohol or drugs and tailgating the car in front. In particular, violations have been reported to be associated with active loss-of-control as well as with speeding and parking offences [10]. Errors are breaches of the rules, such as not noticing pedestrians on road and

braking too hard, they seemed to be the main predictor of accident involvement among elderly drivers [11]. Furthermore, lapses are a set of problems related to the lack of attention and memory defects that cause embarrassment [9], lapses of concentration, for example; forgetting where you parked the car and switching on the wipers when you meant to switch on the lights.

The current study aims to explore specific aspects of driving behavior in Jordan and to investigate the reoccurrence of committing violations, errors, and lapses of the DBQ among Jordanian drivers. The results could be used as a reference in further studies to develop a country specific "Jordanian DBQ" which will help in controlling the driver aberrant behaviors that lead to accidents.

2. Methodology

2.1. Sample

A survey was conducted in Amman city, the capital of Jordan, through direct interviews with a total of randomly selected 200 drivers aged 22 years and above. Only 134 drivers (94 males and 40 females) agreed to participate in the survey and were asked to fill the DBQ.

2.2. Measures

The DBQ contains three subscales to capture different aspects of driver behavior; violations, errors and lapses. DBQ with extended violations was used to measure aberrant driver behaviors. The DBQ includes 10 items of ordinary violations, 8 items of lapses, and 8 items of errors It has 26 behaviors on a six-point scale (0 = never, 1= hardly ever, 2 = occasionally, 3 = quite often, 4 = frequently, and 5 = nearly all the time) The research assistants asked the participants to indicate how often they have committed every behavior.

2.3. Demographic variables

The participants were also asked to indicate their demographic information, age, gender, marital status, educational level, occupation, place of living, housing conditions, and other data related to their driving experience, such as the type of car, whether or not they use the seatbelt and why, their usual driving speed, number of miles their cars recorded and the history of accidents and injuries.

2.4. Statistical analysis

The Statistical Package for Social Sciences "IBM SPSS" was used to analyze the collected data. Descriptive statistics tools and central tendency measures, mean,

median, mode and distribution measures (standard deviation) were also used to get further useful information.

3. Results

3.1. Socio-demographic characteristics and accidents involvement

Analysis of the socio-demographic characteristics of the participants showed that their age distribution ranged between 22 and 74 years with a mean age of 27.5 years. Most of the participants were males aged between 22-29 years (69 of the respondents). Moreover, 49% of participants had 2-5 years of driving experience, with a mean mileage of 45,678.65 km.

Seat belt usage spots light on the driving attitude; 55% of drivers admitted that they don't use seatbelt, and 79.5% of them reported that it causes discomfort and restricts their movement.

When relating seatbelt usage with gender, 75% of female drivers reported that they always use the seatbelt. On the other hand, only 33% of male drivers use seatbelt. Hence, this result assures the finding of the paper "Role of gender and driver behavior in road traffic crashes" [5], which reported that typically, male participants were less likely to wear seat belts.

Table 1 shows age groups of the participants, and whether or not they had an accident. 51% of the participants below 30 years of age were involved in an accident, while the percentages of drivers who experienced a car accident in the age groups 30-39 years, 40-49 years, and above 50 years were 60%, 83%, and 100% respectively. These results indicate that accident involvement increases with age of drivers.

Table 1. Age group and accident involvement

Age group	Accident involvement	Not involved
<30 years	51% (52)	49% (51)
30-39	60% (12)	40% (8)
40-49	83% (5)	17% (1)
>=50	100% (5)	0% (0)

Table 2 shows that 47% of female drivers were involved in traffic accidents, whereas the corresponding percentage of males was 55%.

Table 2. Gender and accident involvement

Gender	Accident involvement	Not involved
Female	47% (19)	53% (21)
Male	55% (52)	45% (42)

Comparing these results with the results found in ”Role of gender and driver behavior in road traffic crashes” [5], female drivers were less likely to be involved in traffic accidents.

Table 3 shows the relationship between the educational level of participating drivers and accident involvement. Surprisingly, the highest group involved in accidents was the one of drivers holding university degree (most educated, 54%) compared with 46% of people with secondary school qualification. for drivers with an intermediate school educational level the number of drivers in the sample was not sufficient to derive conclusive result. These findings show that educational level is not related to the skills the driver needs to avoid being involved in an accident.

Table 3. Educational level and accident involvement

Educational level	Accident involvement	Not involved
University	54% (64)	46% (55)
Secondary	46% (6)	54% (7)
Intermediate	50% (1)	50% (1)

The effect of driving experience on accident involvement is represented in Table 4, It can be seen that drivers who have been driving for a longer period of time appear to commit higher cumulative number of accidents mainly due to their longer exposure to accident risk, thus increasing their probability of accident involvement.

Table 4. Driving experience and Accident involvement

Driving experience	Accident involvement	Not involved
Below 2 years	20% (3)	80% (12)
2-5 years	53% (35)	47% (31)
5-10 years	54% (19)	46% (16)

More than 10 years	74% (14)	26% (5)

The results that link seatbelt usage and accident involvement aren't directly correlated, due to the fact that seatbelt usage reduces traffic fatalities and serious injuries, yet it doesn't prevent the accident from happening. As shown in Table 5, 22% of drivers who don't use seatbelt were seriously injured compared to 9% of those who use it.

Table 5. Seat belt usage and injuries

Seatbelt use		Accident involvement	Were injured
Yes	45%	Yes (55%)	9%
		No (45%)	-
No	55%	Yes (51%)	22%
		No (49%)	-

3.2. Causes of accidents

The main contributing cause of accidents reported by drivers who were previously involved in an accident is careless driving (46.5%), followed by excessive speeding (19.7 %), alcohol and drugs (7%). Furthermore, 25.4% of the participants reported their involvement in traffic violations as an accident cause (Fig.1)

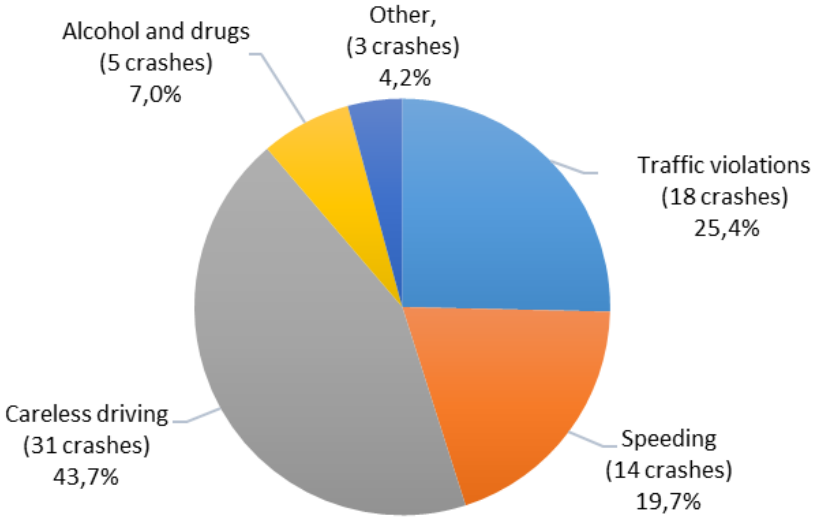


Figure 1. Causes of crashes among the surveyed drivers in Jordan

3.3. Violations, errors and lapses

Traffic violations occur when drivers violate laws that regulate vehicle operation on streets and highways. These may vary depending on the country and area.

In Jordan the most common traffic violation is speeding, with 34% of people reporting this, when 30% of drivers reported that they disregard the speed limits late at night and early in the morning, and they also tend to disregard speed limits on main roads. Moreover, 29% of drivers become impatient with slow drivers in the outer lane and overtake on the inside lane (right lane). On the other hand, running a red light and getting involved in unofficial races with other drivers are the least violations committed, with the percentage of 7% and 11% respectively of the recorded sample.

After analyzing the violations measured in the DBQ, it turned out that they were divided into two main categories, the first one was violations of traffic regulations, and the other one was the violations related to other road users. Drivers in Amman tend to violate traffic regulations more than violations related to other road users

The errors, are typically misjudgments and failures of observation that may be hazardous to others. Such as attempting to overtake someone that you hadn't noticed to be signaling a left turn, failing to check the rear mirror before pulling out or changing lanes and under estimating the speed of an oncoming vehicle when overtaking, these errors were found to be the most common in Jordan.

By the descriptive analysis of the collected data about lapses in driving, and calculating the measures of central tendency, the most common answer for the frequency of people realizing that they have no clear recollection of the road they have been travelling along was that they occasionally realized that, which shows that this is not a very common lapse. Also, drivers said that they occasionally forget the place they have parked their car in due to lack of concentration.

Table 6 and 7 show the DBQ mean scores and standard deviations for each of the individual items related to violations, errors and lapses among Jordanian drivers. In these tables the mean represents the following: 0 – never, 1 – hardly ever, 2 – occasionally, 3 – quite often, 4 – frequently, 5 – nearly all the time.

Table 6. Means and Standard deviations of items of Driver Behavior Questionnaire (DBQ) part 1

Variables	***Mean***	***St. deviation***
Violations		
Disregard the speed limits late at night or early in the morning	2.18	1.76
Disregard the speed limits on a motorway	2.13	1.67
Become impatient with a slow driver in the outer lane and overtake on the inside (right) lane	2.13	1.50
Sound your horn to indicate your annoyance to another road user	2.02	1.45
Stay in a motorway lane that you know will be closed ahead until the last minute before forcing your way into the other lane	1.94	1.67
Drive especially close to the car in front as a signal to its driver to go faster or get out of the way	1.91	1.69
Have an aversion to a particular class of road user and indicate your hostility by whatever means you can	1.59	1.48
Angered by another driver's behavior, you give chase with the intention of giving him/her a piece of your mind	1.13	1.17
Get involved with unofficial 'races' with other drivers	0.86	1.13
Crossing a red light	0.75	0.99
Errors		
Attempt to overtake someone that you hadn't noticed to be signaling a left turn	1.40	1.32
Fail to check your rear-view mirror before pulling out or changing lanes, etc.	1.22	1.34
Underestimate the speed of an oncoming vehicle when overtaking	1.14	0.98
Miss 'Give Way' signs and narrowly avoid colliding with traffic having right of way	1.06	1.13
Apply sudden brakes on a slippery road, or steer wrong way in a skid	0.96	0.87
Queuing to turn right onto a main road, you pay close attention to the mainstream of traffic that you nearly hit the car in front	0.92	1.14
Fail to notice that pedestrians are crossing when turning into a side street from a main road	0.79	0.94
On turning right nearly hit a two-wheeler who has come up on your inside	0.69	0.88

Table 7. Means and Standard deviations of items of Driver Behavior Questionnaire (DBQ) part 2

Variables	***Mean***	***St. deviation***
Lapses		
Attempt to drive away from the traffic lights	3.47	1.47
Realize you have no clear recollection of the road along which you have been travelling	1.75	1.34
Forget where you left your car in the car park	1.50	1.31
Get into the wrong lane when approaching a roundabout or a junction	1.43	1.09
Intending to drive to destination A and, you 'wake up' to find yourself in destination B, because the latter is your more usual destination	1.43	1.15
Hit something when reversing that you had not previously seen	1.31	0.99
Switch on one thing, such as headlights, when you meant to switch on something else, such as wipers	1.21	1.10
Misread the signs and exit from the roundabout on the wrong road	1.10	0.99

3.4. Strong and weak driving components

The participants were asked to describe the strength of the components of their driving style. Table 8 shows the results where 77.6% of drivers claim that they are careful drivers, while 81.3% of them can react fast in critical situations, and 79.1% show consideration for other road users, 87.4% have full control over the vehicle in normal conditions.

Secondly, when it came to the weak components, it was found that 42.6 % of the respondents lose their patience when driving behind a slow car, and 29.1% of the recorded sample can't tolerate other driver errors calmly. This indicates that drivers are usually short-tempered. Yet 22.4% of drivers face difficulties in driving within the speed limits.

Moreover, controlling the car through a skid, fluent lane changing in heavy traffic, and reverse parking in narrow gap scored a weak performance, with a percentage of 20.2% of the recorded sample.

Table 8. Strong and weak components in driving style

	Definitely Strong	***Strong***	***Neither weak nor strong***	***Weak***	***Definitely Weak***
Performance in a critical situation	21.6%	41%	26.1%	7.5%	3.7%
Driving behind a slow car without getting impatient	6%	25.4%	26.1%	23.9%	18.7%
Managing the car through a skid	11.2%	38.1%	30.6%	14.9%	5.2%
Predicting traffic situations ahead	16.4%	53%	19.4%	8.2%	3%
Driving carefully	24.6%	53%	14.9%	6%	1.5%
Knowing how to act in particular traffic situations	14.9%	53%	24.6%	7.5%	0%
Fluent lane changing in heavy traffic	14.9%	44%	20.8%	16.4%	3.7%
Fast reactions	23.1%	58.2%	14.9%	3%	0.7%
Showing consideration for other road users	25.4%	53.7%	12.7%	6%	2.2%
Staying calm in irritating situations	7.5%	23.9%	33.6%	20.9%	14.2%
Controlling the vehicle	29.9%	57.5%	9.7%	2.2%	0.7%
Avoiding competition in traffic	35.8%	34.3%	11.9%	7.5%	10.4%
Keeping a sufficient following distance	20.9%	38.1%	23.8%	10.4%	6.7%
Overtaking	23.9%	39.6%	23.9%	9%	3.7%
Relinquishing legitimate rights when necessary	12.7%	41.8%	25.4%	16.4%	3.7%
Confronting to the speed limits	15.7%	37.3%	24.6%	18.7%	3.7%
Avoiding unnecessary risks	27.6%	44%	14.1%	9.7%	4.5%
Tolerating other drivers' errors calmly	6.7%	32.8%	31.4%	20.1%	9%
Reverse parking in a narrow gap	27.6%	37.3%	14.9%	7.5%	12.7%

4. Road safety campaigns awareness in Jordan

People awareness of road safety campaigns in Jordan was investigated during the data collection phase. Analysis of the responses revealed the followings:

People who are aware of road safety campaigns constitute only 14.9% of drivers with only 12.7% of them believe that the campaigns are effective.

1) The most popular mediums for road safety campaigns are the internet and radio as reported by 23.9% of the responses followed by internet and TV (17.9%).
2) The most effective traffic enforcement tools are presented in figure 2,
    - Speed cameras (as reported by 36%)
    - Police presence (25%)
    - Traffic light cameras (21%)
    - Others (18%)

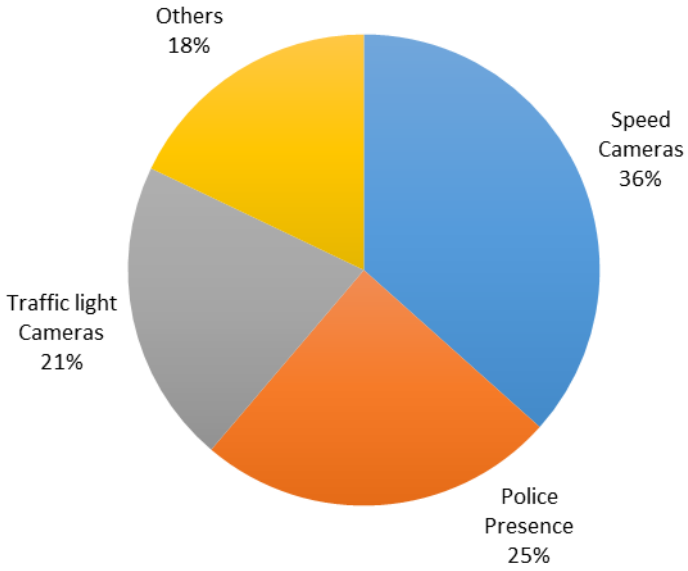


Figure 2. Most effective traffic enforcement tools

3) Drivers who think that speeding fines must be raised constitute 29.9% of the total participants whereas 70.1% think that speeding fines are high enough.
4) Drivers who think that there should be an increased police presence at main junctions constitute (56%) of drivers.
5) Drivers who think that more road safety strategies should be implemented (85.8%).
6) Drivers who think that high powered cars should have speed restricting devices (71.6%).
7) Drivers who think that should be a special hotline to report aggressive drivers (90.3%).
8) Drivers who think that mobile radars should be used during peak volume traffic (58.2%).
9) Drivers who think all high peak roundabouts should be equipped with traffic lights (55.2%). A remarkable decrease in traffic congestion at two of the high peak roundabouts in Amman city was detected after this solution was implemented in 2019.

5. Drivers' suggestions

The participants reported some suggestions that they perceive effective in reducing the magnitude of road safety problem in Jordan. These include: regular maintenance of roads, filling holes in streets and regularly renewing street markings, and providing proper street lighting.

Another field that the drivers feel requires improvement is the use of traffic calming techniques. They suggested that more humps should be installed to reduce drivers' speed, while providing road signs and adding more posted speed signs will force drivers to abide with traffic laws and regulations. Effective enforcement of traffic law was also perceived as another intervention with potential for accident reduction in Jordan.

6. Methodological limitations

Although several studies have reported that self-reports of driving may correspond well to actual driving behavior, the collected data used for the purpose of this study were based mostly on drivers' self- reported behavior and no observations were made. It is also possible that some respondents could mislead their answers about positive and aggressive driving. Furthermore, the measurements of accidents and

injuries involvement were based on a self-report of all accidents. Therefore, some respondents may have underestimated the number of accidents in which they had been involved.

7. Conclusion

The following conclusions were drawn based on the results of data analysis and evaluation conducted in this study.

- Jordanian drivers who have longer driving experience tend to be involved in higher cumulative number of traffic accidents through their driving years; the reason behind this is their longer years of exposure, therefore it is very likely that they have been involved in more accidents.
- It was proven that careless driving is the most common self-reported reason behind traffic accidents with 46.5% of the participating drivers supporting this view.
- The most common traffic violation practiced is speeding, 34% of drivers disregard the speed limits especially late at night and early in the morning. (hence that there are no motorways in Amman).
- The most common errors reported by drivers were found to be; failing to check rear mirror before pulling out or changing lanes, and under-estimating the speed of an oncoming vehicle in overtaking.
- Regarding lapses, 27.6% of drivers showed that they have no clear recollection of the road they have been travelling along, and this is not a highly common lapse. However, 26.1% usually forget where they parked their car in due to lack of concentration.
- The strong components of the driving style of Jordanian drivers were found to be their claim being careful drivers (77.60%), while 81.30% of them reported that they can react fast in critical situations.
- Internet and radio media forms are the most effective in delivering traffic awareness campaigns to drivers.
- Traffic enforcement tools such as speed cameras and police presence, are the most effective in regulating driver's aberrant behavior.

8. Recommendations

Considering the burden and impact of traffic accidents in Amman, more road safety campaigns should be launched, initiated, and supported, yet strategic road

safety management mechanisms should be developed and implemented to control and reduce the growing accidents rate.

Driver behavior questionnaire is an effective tool in this process since the results found in this study can be used as a reference in further studies. However, it is highly recommended to conduct this survey online, since face to face interviews consumed lots of time, and this is guaranteed to reach a larger sample of drivers in order to result in more representative results, identify outliers and provide a smaller margin of error.

References

[1] Zs. Sándor, A. Monostori, Traffic Analysis of Specific Motorways with Different Usage Characteristics in Hungary with the Method of Section Control, *Acta Technica Jaurinensis* 13 (3) (2020) pp. 211–245. doi: https://doi.org/10.14513/actatechjaur.v13.n3.554

[2] A. Kizawi, A. Borsos, A literature review on theconflict analysisof vehicle-pedestrian interactions, *Acta Technica Jaurinensis* (2021). doi: https://doi.org/10.14513/actatechjaur.00601

[3] A. Bener, T. Özkan, et al., The Impact of Aggressive Behaviour, Sleeping, and Fatigue on Road Traffic Crashes as Comparison between Minibus/Van/Pick-up and Commercial Taxi Drivers, *Journal of Traffic and Transportation Engineering* 5 (2017) pp. 21-31. doi: https://doi.org//10.17265/2328-2142/2017.01.003

[4] A. Bener, D. Crundall, Role of gender and driver behaviour in road traffic crashes, *International Journal of Crashworthiness* 13 (3) (2008) pp. 331-336. doi: https://doi.org//10.1080/13588260801942684

[5] K. Jadaan, A. Bener, E. Braizat, Formulation of road safety strategy for developing countries with special reference to Jordan, *Advances in Transportation Studies an International Journal* 40 (2016) pp. 101-111. doi: https://doi.org/10.4399/97888548970078

[6] A. Bener, K. Jadaan, et al., The effect of aggressive driver behavior, violation and error on vehicle crashes involvement in Jordan, *International Journal of Crashworthiness* 25 (3) (2019) pp. 276-283. doi: https://doi.org//10.1080/13588265.2019.1583422

[7] J.C.F. de Winter, D. Dodou, The driver behaviour questionnaire as a predictor of accidents: A meta-analysis, *Journal Safety Research* 41 (6) (2010) pp. 463-470.
doi: https://doi.org//10.1016/j.jsr.2010.10.007

[8] N. Zhao, B. Mehler, et al., An investigation of the relationship between the driving behavior questionnaire and objective measures of highway driving, *Transportation Research Part F: Traffic Psychology and Behaviour* 15 (6) (2012) pp. 676-685.
doi: https://doi.org// 10.1016/j.trf.2012.08.001

[9] A. Kashani, M. Ravasani, E. Ayazi, Analysis of Drivers' Behavior using Manchester Driver Behavior Questionnaire Based on Roadside Interview in Iran, *International Journal of transportation Engineering* 4 (1) (2016) pp. 61-74.
doi: https://doi.org// 10.22119/IJTE.2016.33497

[10] A. Bener, E. Yildirim, et al., The Driver Behavior Questionnaire as an accident predictor in Cross-cultural Countries in Qatar and Turkey: Global Public Health Problem, *British Journal of Medicine and Medical Research* 15 (7) (2016) pp. 1–9.
doi: https://doi.org//10.9734/BJMMR/2016/25719

[11] D. Parker, L. McDonald, et al., Elderly drivers and their accidents: The aging driver questionnaire, *Accident Analysis and Prevention* 32 (6) (2000) pp. 751–759.
doi: https://doi.org//10.1016/S0001-4575(99)00125-6

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 455-476, 2021
DOI: 10.14513/actatechjaur.00615

Analysis of Embankment Supported by Rigid Inclusions Using Plaxis 3D

R.Alsirawan[1,*]

[1] Széchenyi István University, Department of Structural and Geotechnical Engineering.
Egyetem tér 1, 9026 Győr, Hungary
*E-mail: Rashad.seirawan@gmail.com

Submitted: 07/05/2021; Accepted: 16/08/2021; Published online: 27/08/2021

Abstract: A rigid inclusion-supported embankment is used to overcome the problems of soft soils. This system is considered complex due to the various interactions between its elements, namely the embankment body, load transfer platform, geogrid layers, piles, and soft soils. The load transfer mechanism is based on the phenomenon of soil arching, the tension in the geogrid layers, support of the soft soils, and friction between piles and soft soil. In this paper, the first part highlights the behaviour of a rigid inclusion-supported embankment validated by field measurements, and the contribution of rigid inclusions technology to the reduction of settlement and creep settlement. In addition, the effect of geogrid in improving the load efficiency and reducing the settlements is presented. In the second part, a comparison is made between many analytical design methods and a three-dimensional finite element analysis method. The results show the inconsistencies between the analytical methods in calculating the load efficiency and the tension in the geogrid.

Keywords: rigid inclusion-supported embankment; load efficiency; 3D finite element analysis

1. Introduction

Due to the unfavourable properties of soft soil, different technologies are used to support embankments on this type of soil to avoid excessive settlement and loss of

bearing capacity. Many advantages motivate engineers to use rigid inclusions technology under embankments, buildings, and industrial facilities. Rigid inclusions fulfil the requirements of bearing capacity and settlements, rapid construction, and competitive cost.

The typical rigid inclusions system consists of rigid piles, with relatively small diameters, and load transfer platform LTP, located over the piles, which includes generally one or more layers of geosynthetics [1]. The vertical load is divided into three parts, Fig. 1:

1. Part A: this part is transferred directly to the rigid piles via the soil arch.
2. Part B: this part is transferred by geosynthetic layers to the rigid piles.
3. Part C: this part is transferred to the soft soil [2].

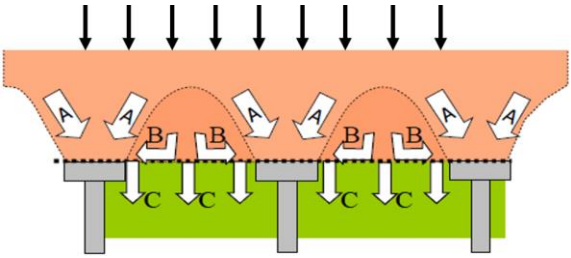


Figure. 1. Load transfer mechanism in embankment supported by rigid inclusions [2]

Han and Gabr [3] conducted a numerical study using the program Flac to investigate the interactions between 5s, soils, and geosynthetic reinforcement. The analysis showed that using the geosynthetic layers within the LTP and increasing the stiffness of the pile material leads to a reduction in the settlement at the embankment base. This study also found that the load efficiency (defined as the ratio of load acting on the pile head to the total vertical load resulting from the embankment weight and surcharge in a unit cell) increases as the elastic modulus of the pile increases and the height of the embankment increases. As for geosynthetic stress, this value increases as the stiffness of the geosynthetic, the elastic modulus of the pile, and the height of the embankment increase. A numerical study coupling the finite element method and the discrete element method described the load transfer mechanism. In this study, it was found that the load efficiency increases with high strength parameters of LTP soil, Chevalier et al. [4]. Abusharar et al. [5] proposed an analytical method of design. According to this method, the load efficiency increases with the height of the embankment at large cover ratios (defined as a ratio of the pile head (cap if any) to the LTP area), it also increases at high soil elastic modulus and small soft soil

thicknesses. The geosynthetic tension increases linearly with the embankment height and decreases with the soil elastic modulus. Pham [6] presented a new analytical design method, the parametric study demonstrated that the load efficiency increases with the consolidation and low values of subgrade reaction modulus of subsoils, this study also showed that the tension decreases with the increase of the consolidation degree of subsoil, friction angle of LTP fill, and cover ratio. Fischer et al. [7] [8] found through an experimental tests that the performance and effectiveness of the geosynthetic reinforcement are affected by the ratio of the geosynthetic aperture size and the soil particle size, which in turn affects on the geosynthetic tension. Zhuang et al. [9] developed an analytical method and found that the tension increases linearly with the embankment height, pile spacing, and geogrid stiffness. van Eekelen et al. [10] performed a series of model tests in the laboratory, which showed that the use of a gravelly platform (∅=49°) instead of a sandy platform (∅=40.88°) increases the load efficiency. It increases also as a percentage whenever the uniform loads increase. The behaviour of geotextile and geogrid is similar in load transfer; therefore, the load efficiency and the tension are not affected by the type of reinforcement. These experiments led to the conclusion that the tension decrease in two cases, namely, when the friction angle of the LTP fill (gravelly platform) is high and when the stiffness of the reinforcement material is high. Girout et al. [11] performed thirty-three geotechnical centrifuge tests. The findings were, the presence of the reinforcement enhances the load efficiency, and the load efficiency depends on the stiffness and the positions of the reinforcement layers within the LTP.

This paper briefly describes the behaviour of rigid inclusions and the advantages of using this technology to reduce the settlements at the embankment surface in various cases. A comparison is also made between many analytical design methods and a numerical method using a validated model of an embankment supported by rigid inclusions.

2. Background

The design of rigid inclusions is considered complex due to the various interactions between the elements, which include the piles, geosynthetic layers, load transfer platform, and pile caps (if any) [1]. Over the past decades, many researchers have attempted to understand the interactions between the elements and the mechanism of load transfer. To achieve this end, many laboratory models, centrifuge tests, and full-scale tests have been conducted.

Various generations of methods are used to design rigid inclusions under embankments. All these methods are based on the theory of soil arching, which is based on a trap door experiment proposed by Terzaghi [12]. The first generation of methods includes Guido et al. [13], Hewlett and Randolph [14], Low et al. [15],

Kempfert [16], Abusharar et al. [5], and others [17]. The differences between geosynthetic tension and load efficiency calculated by these methods are significant [17][18]. The differences can be attributed to the inadequacy of the tests or models and the inability to consider the effect of all parameters in this complex system. The finite element method is considered to be more comprehensive and can avoid the disadvantages of the previous methods.

BS8006 [19] and EBGEO [20] have proposed new design guidelines in both United Kingdom and Germany. Ellis and Aslam [21][22], van Eekelen et al. [10][23], Zhuang [9], Cui [24], Pham [6] worked on the analysis and design of this technology in the last decade. However, none of the proposed design methods provides realistic solutions in all case studies.

To calculate the geosynthetic tension and load efficiency, many design methods from different generations are chosen in this paper, as follows:

2.1. Terzaghi's design method (1943)

The theory of arching was used as the basis of this method, Russell and Pierpoint [25] expanded this method. The load efficiency is calculated as follows:

$$E = 1 - \frac{S_{3D}(s^2 - a^2)}{s^2} \qquad (1)$$

where: S_{3D} is the stress reduction ratio and calculated according to equation (2), a is the pile cap width (m), s is the pile spacing (m).

$$S_{3D} = \frac{(s^2 - a^2)}{4HaK_0 tan\emptyset}\left(1 - e^{\frac{-4HaK_0 tan\emptyset}{(s^2 - a^2)}}\right) \qquad (2)$$

where: H is the embankment height (m), K_0 is the earth pressure coefficient at rest, $\emptyset$ is the friction angle of the embankment fill.

The tension in the geosynthetic layer is given by the following equation (kN/m):

$$T = \frac{S_{3D}\gamma H(s^2 - a^2)}{4a}\sqrt{1 + \frac{1}{6\varepsilon}} \qquad (3)$$

where: γ is the LTP soil unit weight (kN/m^3), ε is the initial strain and equal to 5%.

2.2. Guido et al. design method (1987)

The load efficiency is calculated by equation (1) and S_{3D} is calculated according to the following equation [13]:

$$S_{3D} = \frac{s-a}{3\sqrt{2}\,H} \qquad (4)$$

The tension in the geosynthetic layer can be calculated using equation (3)

2.3. Hewlett and Randolph design method (1988)

The load efficiency is calculated according to equation (1) and The tension in the geosynthetic layer can be calculated according to equation (3).

The stress reduction ratio S_{3D} is calculated using equations (5) and (6), the higher value is used in the design [14],

$$S_{3D} = (1-\frac{a}{s})^{2(K_P-1)}\left(1-\frac{2(K_P-1)\,s}{\sqrt{2}H\,(2K_P-3)}\right)+\frac{2(K_P-1)(s-a)}{\sqrt{2}H\,(2K_P-3)} \qquad (5)$$

$$S_{3D} = \frac{1}{\left(\frac{2K_P}{1+K_P}\right)\left[\left(1-\frac{a}{s}\right)^{(1-K_P)}-\left(1-\frac{a}{s}\right)\left(1+\frac{a}{s}K_P\right)\right]+(1-\frac{a^2}{s^2})} \qquad (6)$$

where: K_P is the passive earth pressure coefficient and given by:

$$K_P = \frac{1+\sin\emptyset}{1-\sin\emptyset} \qquad (7)$$

2.4. Abusharar et al. design method (2009)

Equation (7) is used to calculate the load efficiency [5]:

$$E = 1-\frac{(s^2-a^2)\,\sigma_s}{s^2\,\gamma\,H} \qquad (8)$$

where: σ_s is the vertical stress acting on the geosynthetic layer (kN/m^2).

The geosynthetic tension is given by:

$$T = 4\beta^2 J + 0.25\,(s-a)\,\lambda(\sigma_s\,tan\emptyset + \frac{tE_c}{D}tan\,\emptyset_c) \qquad (9)$$

where: β is the sag ratio and given as follows $\beta = t/(s-a)$, t is the maximum settlement of soft soil midway between rigid pile caps when geosynthetic is used (m), J is the geosynthetic tensile stiffness (kN/m), λ is a factor that depends on the type of the geosynthetic and ranges between (0.7-0.9), $E_c, \emptyset_c, D$ are the elastic modulus (kN/m^2), friction angle, and depth of the soft soil (m) .

2.5. BS8006 design method (2010)

According to BS8006, two equations are proposed to calculate the load efficiency, equation (9) is used for partial arching, and equation (10) is used for full arching [19].

$$E = 1 - \frac{1}{s^2}\left[s^2 - a^2\left(\frac{P_c}{\gamma H}\right)\right] \qquad (10)$$

$$E = 1 - \frac{1.4\,(s-a)}{s^2\,H}\left[s^2 - a^2\left(\frac{P_c}{\gamma H}\right)\right] \qquad (11)$$

$$\frac{P_c}{\gamma H} = \left[\frac{C_c\,a}{H}\right]^2 \qquad (12)$$

where: P_c is the vertical stress acting on the pile caps ($\mathrm{kN/m^2}$), C_c is the arching coefficient, for end-bearing piles ($C_c = 1.95\,H/a - 0.18$, for frictional piles ($C_c = 1.5\,H/a - 0.07$).

The geosynthetic tension is given by:

$$T = 0.5\,W\,\frac{s-a}{a}\,(1 + \frac{1}{6\varepsilon})^{0.5} \qquad (13)$$

For partial arching $[0.7(s-a) \le H \le 1.4(s-a)]$,

$$W = \frac{\gamma\,H}{2(s-a)}\left[s^2 - a^2\left(\frac{P_c}{\gamma H}\right)\right] \qquad (14)$$

For full arching $[H > 1.4(s-a)]$,

$$W = 0{,}7\gamma\left[s^2 - a^2\left(\frac{P_c}{\gamma H}\right)\right] \qquad (15)$$

where: W is the load acting on geosynthetic (kN).

Equations (13), (14) are proposed by van Eekelen et al. [26] for use in the British Standard.

2.6. Tuan A. Pham design method (2020)

The load efficiency and geosynthetic tension are obtained from equations (15), (16) [6]:

$$E = \frac{P_{ca} + P_{cm}}{(\gamma H + q)s^2} \qquad (16)$$

$$T = \frac{8}{3}\left(\frac{8y}{3(s-a)}\right)^2 J + (\alpha_P \tan \emptyset_p \, \sigma_s^a + \alpha_s \tan \emptyset_s \, \sigma_{up} + 0.1 c_s)(s-a) \qquad (17)$$

where: P_{ca}, P_{cm} are the loads on the pile (cap) via soil arch and through the geosynthetic layer (kN) , y is the maximum deflection of the geosynthetic (m), α_P, α_s are the interaction coefficients between the soils and geosynthetic layer, $\emptyset_p, \emptyset_s$ are the friction angles of the soils at the top and bottom of the geosynthetic layer respectively, σ_s^a is the vertical stress acting on the soft soil (kPa), σ_{up} is the upward counter pressure from the soft soils (kPa), c_s is the total cohesion of the soils at the top and bottom of the geosynthetic (kPa).

3. Case study

A full-scale model of an embankment supported by piles and two layers of uniaxial geogrids was performed in the Virvée swamp (France) within the frame of the new South Europe Atlantic high-speed line project [27]. The geometry information of the supported embankment and the geotechnical profile can be seen from the cross-section in Fig. 2. A working platform with a thickness of 1.0 m was constructed before the construction stages to support the movement of the equipment. The pile has a width of 0.274 m and a length of 12.7 m. The precast concrete piles were driven in 60 working days, followed by the construction of the LTP with a thickness of 0.7 m, this LTP includes two layers of geogrid at (0.2-0.4) m above the pile head. Finally, the embankment was constructed in two stages (represents the embankment body and traffic load), each 1.9 m thick.

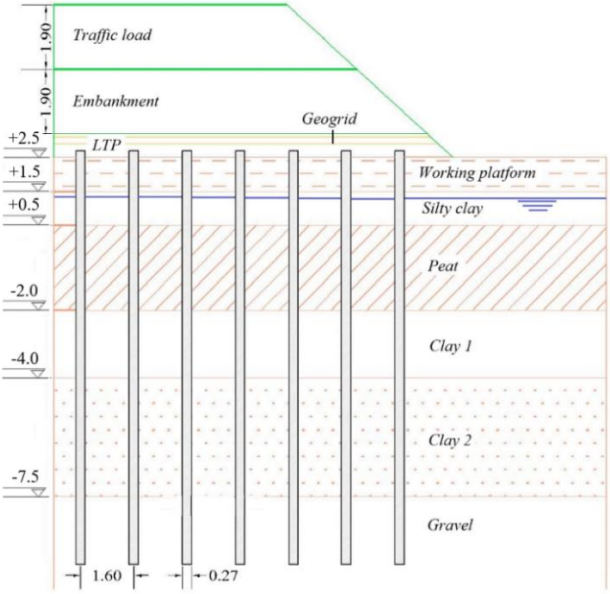


Figure 2. Cross-section of the rigid inclusion-supported embankment

3.1. 3D Finite Element modelling

Plaxis 3D CONNECT Edition V20 program is used to analyse the behaviour of the rigid inclusion-supported embankment, the dimensions of the model are 45 m in the x-direction, 4.8 m in the y-direction, as well as 25 m in the z-direction. Fig. 3 shows the typical FE mesh of this model.

For the validation of the model, six rings fixed along a magnet extensometer were used to measure the settlements in the soft soil layers during the construction stages and consolidation period, four earth pressure cells (EPCs) were fixed inside and over the LTP to measure the stresses in these points, three settlement sensors were installed to measure the vertical displacement of the pile head, and settlements at two representative measurement points inside the LTP [27].

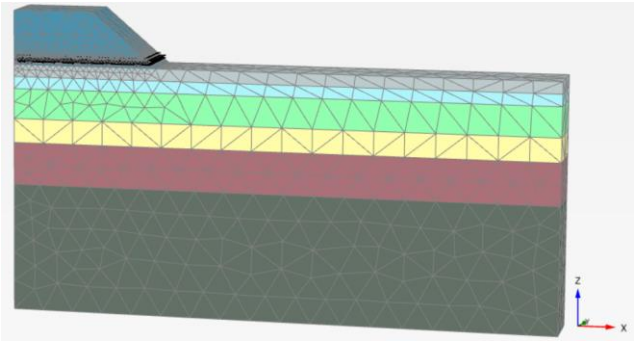


Figure 3. FE mesh of the supported embankment

The piles were modelled as embedded beam elements and the geogrid was modelled as elastoplastic material. The properties of the pile material and geogrid are listed in Table (1). Hardening soil model (HS) was used to simulate the behaviour of gravel, working platform, and embankment soils. Table (2) summarize the soil properties.

Table 1. Properties of pile material and geogrid

	Unit weight $\gamma\ (kN/m^3)$	Young modulus $E\ (GPa)$	Poisson ratio (v)	Stiffness (kN/m)
Pile	24	20	0.2	-
Geogrid	-	-	0.2	13000

The soft soil creep model (SSC) was used to simulate the behaviour of soft soils (silty clay, peat, and clay). The analyses for the gravel, working platform and embankment were modelled as drained condition and undrained condition for soft soils. The parameters of soft soils are tabulated in Table (3).

Table 2. Parameters of the embankment, working platform, and gravel soils

	Embankment	Working platform	Gravel
γ_{unsat} (kN/m^3)	21	21	19
γ_{sat} (kN/m^3)	21	21	20
φ°	35	35	35
Ψ°	5	5	5
c (kPa)	5.0	5.0	10.0
E_{50}^{ref} (kN/m^2)	16000	12860	63000
E_{oed}^{ref} (kN/m^2)	16000	12860	63000
E_{ur}^{ref} (kN/m^2)	48000	38580	189000
m	0.5	0.5	0.5

k (m/day)	0.864	0.864	1.00

Table 3. Soft soil parameters

Layer	Silty clay	Peat	Clay 1	Clay 2
γ_{unsat} (kN/m^3)	11.56	6.2	13.0	13.5
γ_{sat} (kN/m^3)	15.0	10.6	14.0	14.5
φ°	29	29	29	29
Ψ°	0	0	0	0
c (kPa)	4	4	4	4
λ^*	0.0928	0.2560	0.2016	0.1895
κ^*	0.0232	0.0845	0.0537	0.0475
μ^*	0.0027	0.0010	0.0019	0.0018
k_x, k_y (m/day)	8.64E-4	5.55E-4	6.25E-4	5.55E-4
k_z (m/day)	8.64E-4	6.75E-4	6.25E-4	5.55E-4
OCR	8.40	7.85	3.23	1.45

4. Results and discussion

The construction of the embankment over soft soil layers with undesirable characteristics yields excessive settlements and high excess pore pressure which is considered unsuitable for time-bound construction projects. In this section, the advantages of the rigid inclusions to overcome these difficulties are presented. Add to that, the behaviour of this system (stress distribution, soil-inclusion interactions), and the inconsistencies in some design methods are discussed also.

4.1. Embankment surface settlements

One of the objectives of using rigid inclusions technology is to reduce the settlement at the embankment surface, especially for embankments over soft soil layers, which are characterized by their high compressibility. Fig. 4 shows the maximum settlements at the embankment surface at different heights of the embankment and in different cases (unsupported embankment, embankment supported with piles, embankment supported with piles, and one layer of geogrid, and embankment supported with piles and two layers of geogrid). A comparison between the three improvement methods shows that the placement of one layer of geogrid reduces the settlements by 23% compared to embankments supported only with piles, while the placement of two layers reduces the settlements by 20%.

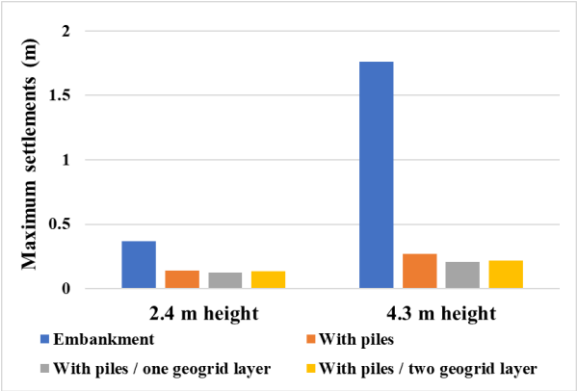


Figure 4. Maximum settlements at the embankment surface

Differential settlements at the embankment surface should be studied as maximum settlements. Rigid inclusions help to reduce differential settlements to acceptable levels, especially on high embankments. Differential settlements can be reduced also by adding one or two layers of geogrid within the LTP, making these settlements almost non-existent, as shown in Fig. 5.

4.2. Excess pore pressure

Fig. 6 shows the evolution of the excess pore pressure. Concerning the unsupported embankment, which is constructed in two stages to avoid soil collapse, the expected time of full consolidation is indefinite. When rigid inclusions technology is used, the dissipation of excess pore pressure is faster and the expected time of full consolidation is about 400 days, which is considered suitable for time-bound construction projects.

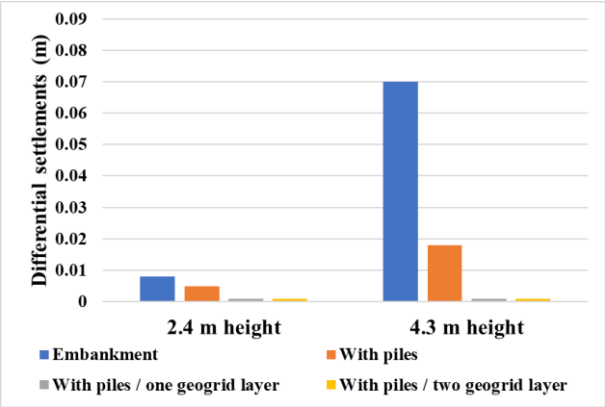


Figure 5. Differential settlements at the embankment surface

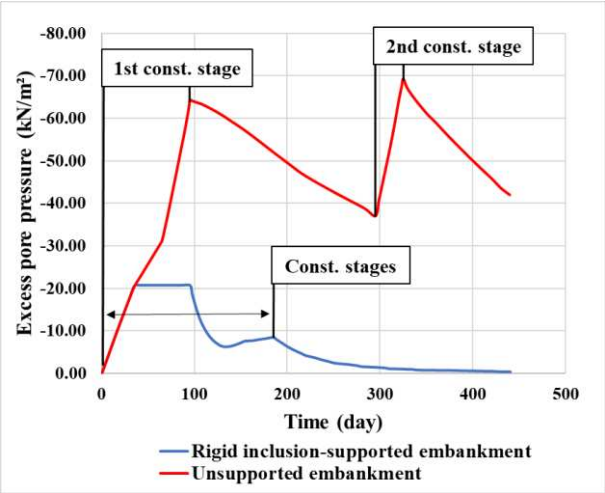


Figure 6. Excess pore pressure

4.3. Creep (secondary settlement)

Creep (secondary settlement) is one of the problems of soft soil. Because of their high compressibility, these soils suffer more from the secondary settlement than other soil types. Although the creep phenomenon is noticeable only after a long period, it is necessary to consider its effect. The use of rigid inclusions can ensure the stability of the embankment in the long term. After 25 years, the secondary

settlement of the unsupported embankment reaches 0.54 m, while it does not exceed 0.012 m for the supported embankment by rigid inclusions, as shown in Fig. 7.

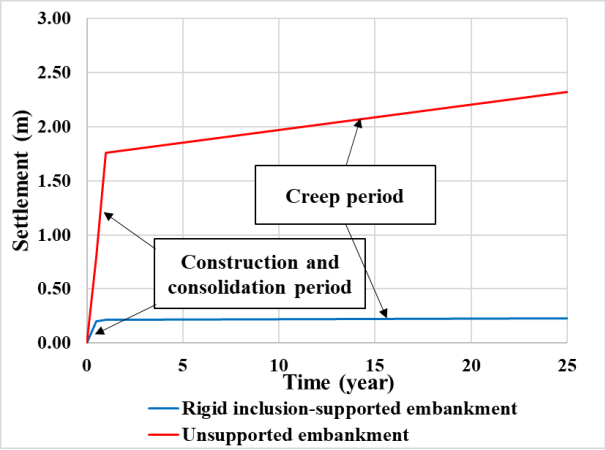


Figure 7. Creep (secondary settlement)

4.4. Vertical and Shear stresses distribution

The embankment settles more over the soft soil than over the piles due to the difference in stiffness. This movement causes the soil arch to form over the piles to carry the main load resulting from the weight of the embankment and the uniform surcharge load. This in turn results in a reduction in the settlement of the soft soil. Fig. 8 illustrates the soil arch in the embankment body. (Plaxis 2D program was used to clarify this phenomenon).

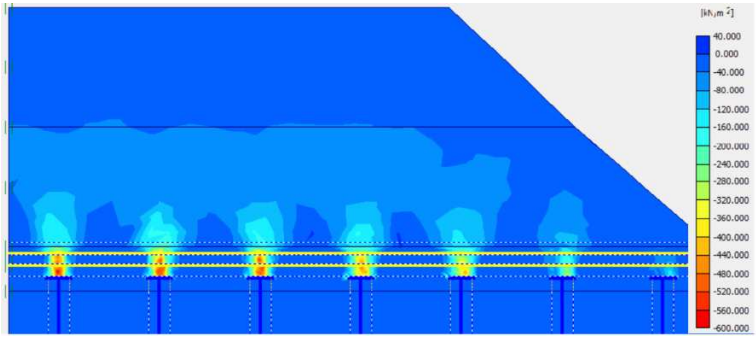


Figure 8. Soil arch in the embankment body

The vertical stress in the midway between two piles in the embankment body increases to a certain depth and then decreases to a depth near the base of the embankment. The vertical stress then increases again due to the soil weight under the soil arch, where the depth of decrease in vertical stress represents the height of the soil arch. Fig.9 shows the distribution of vertical stress during the different stages of construction.

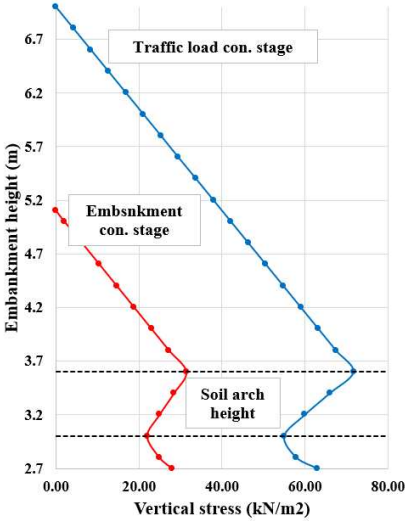


Figure 9. Vertical stress distribution

The shear stresses contribute to the determination of the shear planes which in turn represent the external and internal surfaces of the soil arch, Fig. 10 illustrates the shear stresses over the pile head. (Plaxis 2D program was used to clarify this phenomenon).

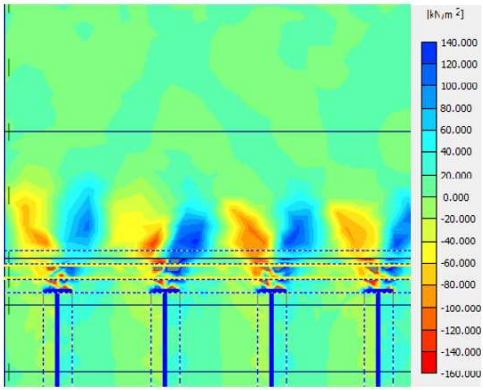


Figure 10. Shear stress distribution

4.5. Soil-pile interaction

Along the upper part of the soft soil and the working platform (if any), soil settlement is greater than the vertical displacement of the pile, which in turn leads to produce negative skin friction. Positive skin friction occurs when the vertical displacement of the rigid inclusion is greater than the soil settlement.

The settlements of the soft soil and the settlements along the pile axis are equal at three neutral lines, which can be characterized as follows:

1. The first neutral line: this line is located inside the embankment body and represents the upper limit of the soil arch.
2. The second neutral line: this line is located inside the soft soil layers that are penetrated by the pile, the loads acting on the pile are composed of the load at the pile head and the load from negative skin friction. These loads increase up to the second neutral line and then begin to decrease due to resisting forces, which include the positive skin friction force and tip resistance.
3. The third neutral line: this line is located within the gravel layer, Fig. 7a shows the settlement distribution at the end of the consolidation period in a unit cell and Fig. 7b shows the axial force profile in the pile.

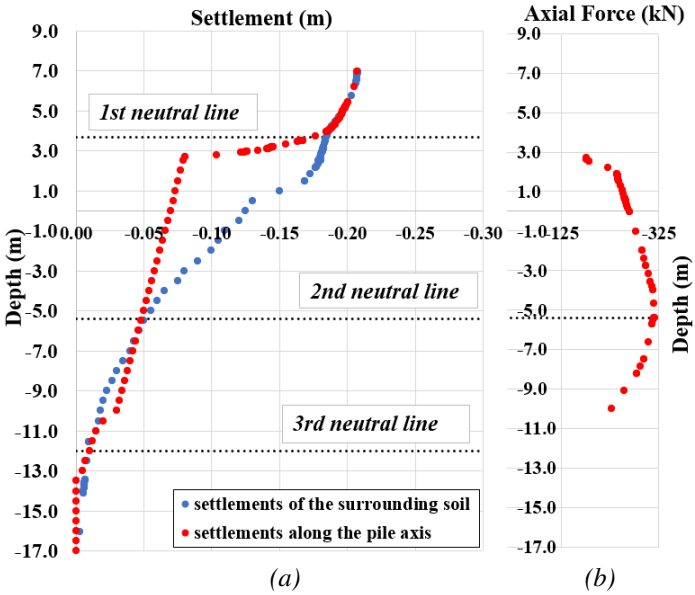


Figure 11. The settlements (a) and the axial force in the unit cell (b)

4.6. Load efficiency

Mathematically, the load efficiency can be expressed by:

$$E = \frac{\sigma_p}{(\gamma H + q)} \qquad (18)$$

where σ_p is the vertical stress applied on the pile head.

Fig. 12 shows the effect of using geogrid layers on the load efficiency at different heights of the embankment. For this case study, the load efficiency is constant in the two height cases studied when this technology is used without geogrid layers. The use of one geogrid layer increases the load efficiency by about 42% and the use of two geogrid layers increases the load efficiency by 44%, the geogrid layers contribute to improving the load transfer to the piles.

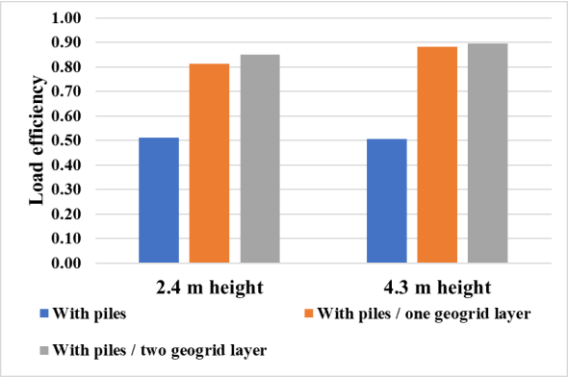


Figure 12. Load efficiency of the pile.

4.7. Comparison of design methods

4.7.1. Comparison using the load efficiency

The load efficiency for each design method is calculated for the present case study. The results are compared with those predicted by the 3D analysis. Out of the six design methods, Guido's method yields a slightly higher value but close to the prediction of the numerical analysis, and Tuan A. Pham's method yields a relatively lower value but can be considered close to the prediction of the numerical analysis. Terzaghi, Hewlett and Randolph, Abusharar, and BS8006 methods give low values for the load efficiency as shown in Fig. 13.

4.7.2. Comparison using the geogrid tension

The methods of Guido and Tuan A. Pham and relatively Abusharar yield values of geogrid tension in good agreement with the results of the 3D analysis. The methods of Terzaghi, Hewlett and Randolph, and BS8006 tend to overpredict the geogrid tension with different proportions as shown in Fig. 14.

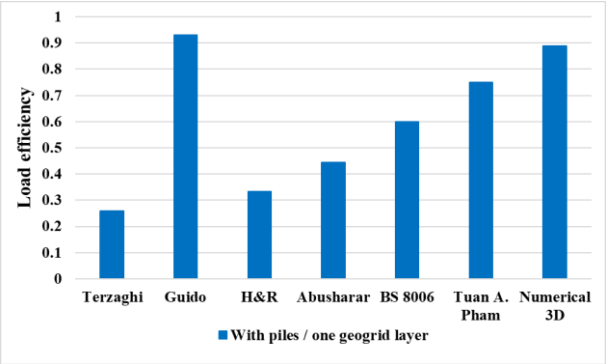


Figure 13. Pile load efficiency with different methods.

Guido's method depends on the clear pile spacing, and embankment height in the calculations, while Tuan A. Pham's method takes into account the properties of embankment fill, soft soil support, stiffness and deflections of geosynthetic layer, friction between soil and geosynthetic layer, embankment height, and clear pile spacing. On this basis, the Tuan A. Pham method can be relied upon as it is the most comprehensive not only with the Guido method but also with other analytical methods. The only disadvantage of the Tuan A. Pham method is the complexity of the calculations.

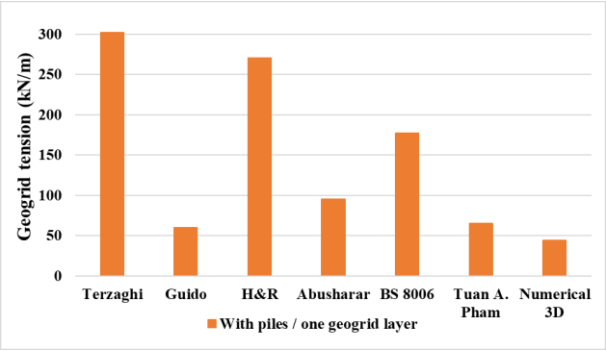


Figure 14. Geogrid tension with different methods.

5. Conclusions

In this paper, a numerical analysis of a validated model of a rigid inclusion-supported embankment has been performed using Plaxis 3D program. The results are as follows:

1. The main part of the load is directly transferred to the pile heads (caps) through soil arches. This phenomenon can be explained by the difference in stiffness between piles and surrounding soil. The shear stresses contribute to form these arches in the embankment body.
2. Three neutral lines can be found in this system where the settlements of the soft soils and the displacement of the piles are the same. the first one is in the embankment body and represents the upper limit of the soil arch. the second one is in the soft soil layers and the third one is in the firm layer where the piles rest.
3. The use of rigid inclusions helps to reduce the maximum and differential settlements at the embankment surface to allowable limits compared to an unsupported embankment.
4. The insertion of the geogrid layers within the load transfer platform increases the load efficiency by 42% in the case of one layer and by 44% in the case of two layers.
5. The use of rigid inclusions reduces the excess pore pressure dissipation period which is considered suitable for time-bound construction projects.
6. Rigid inclusions reduce the secondary settlement to be imperceptible compared to 0.54 m settlement of the unsupported embankment after 25 years of operation, this phenomenon is considered one of the soft soils' problems.
7. Six analytical design methods are used in this paper and compared with the results of numerical analysis. These methods give different results because they consider only some of all the parameters in this complex system and use different calculation methods. The methods of Guido and Tuan A. Pham provide values of load efficiency that are close to the prediction of numerical analysis. The methods of Terzaghi, Hewlett and Randolph, Abusharar, and BS8006 provide low values of load efficiency and tend to overestimate the geogrid tension.
8. Tuan A. Pham design method is showing the closest results to the 3D FEM calculations

References

[1] IREX. Recommendations for the design, construction and control of rigid inclusion ground improvements. *Project National ASIRI*. Presses des Ponts, (2012).

[2] S.J.M. van Eekelen, A. Bezuijen, Model experiments on geosynthetic reinforced piled embankments, 3D test series, *EuroFuge Conference*, Delft University, Netherlands, (2012).

doi: https://doi.org/10.4233/uuid:1e82ac26-c5db-4247-9619-e482e8916390.

[3] J. Han, A. Gabr, Numerical Analysis of Geosynthetic-Reinforced and Pile-Supported Earth Platforms over Soft Soil, *Journal of Geotechnical and Geoenviromental engineering* 128 (1) (2002) pp. 44-53. doi: https://doi.org/10.1061/(ASCE)1090-0241(2002)128:1(44).

[4] B. Chevalier, P. Villard, G. Combe, Investigation of Load-Transfer Mechanisms in Geotechnical Earth Structures with Thin Fill Platforms Reinforced by Rigid Inclusions, *International Journal of Geomechanic* 11 (3) (2011) pp. 239-250. doi: https://doi.org/0.1061/(ASCE)GM.1943-5622.0000083.

[5] S.W. Abusharar, J.J. Zheng et al. A simplified method for analysis of a piled embankment reinforced with geosynthetics, *Geotextiles and Geomembranes* 27 (1) (2009) pp. 39-52. doi: https://doi.org/0.1016/j.geotexmem.2008.05.002.

[6] T. A. Pham, Analysis of geosynthetic-reinforced pile-supported embankment with soil-structure interaction models, *Computers and Geotechnics* 121 (1) (2020) pp. 1-17. doi: https://doi.org/10.1016/j.compgeo.2020.103438.

[7] S. Fischer, T. Szatmari, Investigation of the geogrid-granular soil combination layer with laboratory multi-level shear box test, *Eurogeo 6 conference,* Ljubljana, Slovenia, 2016, pp. 439-449.

[8] T. Szatmari, S. Fischer, A. Trombitas, Investigation of geogrid-granular soil combination layer, *Geosyntetika conference*, Zilina, Slovakia, 2017.

[9] Y. Zhuang, K. Y. Wang, H. L. Liu, A simplified model to analyze the reinforced piled embankments, *Geotextiles and Geomembranes* 42 (2) (2014) pp. 154-165. doi: https://doi.org/10.1016/j.geotexmem.2014.01.002.

[10] S.J.M van Eekelen, A.Bezuijen et al., Model experiments on piled embankments. Part I, *Geotextiles and Geomembranes* 32 (2012) pp. 69-81. doi: https://doi.org/0.1016/j.geotexmem.2011.11.002.

[11] R. Girout, M. Blanc et al., Geosynthetic reinforcement of pile-supported embankments, *Geosynthetics International* 25 (1) (2018) pp. 37-49. doi: https://doi.org/10.1680/jgein.17.00032.

[12] K. Terzaghi, Theoretical Soil Mechanics, *John Wiley & Sons, Inc*, New York, USA, 1943, 526 p.

[13] V. A. Guido, J. D. Kneuppel, M. A. Sweeney, Plate loading tests on geogrid reinforced earth slabs. *Geosynthetic Conference 87*, New Orleans, 1987, pp. 216-225.

[14] W. J. Hewlett, M. F. Randolph, Analysis of piled embankments, *Ground Engineering* 22 (3) (1988) pp. 12-18.

[15] B. K. Low, S. K. Tang, V. Choa, Arching in piled embankments, *Journal of Geotechnical Engineering* 120 (11) (1994) pp. 1917-1938.

[16] H. G. Kempfert, C. Gobel et al., German recommendations for the reinforced embankments on pile-similar elements, *4th Asian Regional Conference on Geosynthetics*, Shanghai, China, 2008, pp. 697-702. doi: https://doi.org/10.1007/978-3-540-69313-0_128.

[17] R. Alsirawan, Review of Geosynthetic-Reinforced Pile-Supported (GRPS) embankments - parametric study and design methods, *Acta Technica Jaurinensis* 14 (1) (2020) pp. 36-59. doi: https://doi.org/10.14513/actatechjaur.00566.

[18] P. Ariyarathne, D. S. Liyanapathirana, Review of existing design methods for geosynthetic-reinforced pile-supported embankments, *Soils and Foundations* 55 (1) (2015) pp. 17-34. doi: https://doi.org/10.1016/j.sandf.2014.12.002.

[19] Code of Practice for Strengthened/reinforced Soils and Other Fills, BS8006-1:2010 (2010).

[20] Recommendations for Design and Analysis of Earth Structures using Geosynthetic Reinforcements, EBGEO:2010 (2010).

[21] E. Ellis, R. Aslam, Arching in piled embankments. comparison of centrifuge tests and predictive methods, Part 1 of 2, *Ground Engineering* 42 (6) (2009) pp. 34-38.

[22] E. Ellis, R. Aslam, Arching in piled embankments. Comparison of centrifuge tests and predictive methods, Part 2 of 2, *Ground Engineering* 42 (6) (2009) pp. 28-31.

[23] S.J.M van Eekelen, A.Bezuijen, A.F.van To, Model experiments on piled embankments. Part II, *Geotextiles and Geomembranes* 32 (1) (2012) pp. 82-94.
doi: https://doi.org/10.1016/j.geotexmem.2011.11.003.

[24] X. Cui, Y. Zhuang et al., An analytical method to calculate the settlement of reinforced piled embankment considering three dimensional deformed geogrid, *European Journal of Environmental and Civil Engineering* (2020) doi: https://doi.org/0.1080/19648189.2020.1810130.

[25] D. Russell, N. Pierpoint, An assessment of design methods for piled embankments, *Ground Engineering* 30 (11) (1997) pp. 39-44. URL https://trid.trb.org/view/476724.

[26] S.J.M.van Eekelen, A.Bezuijen, A.F van Tol, Analysis and modification of the British Standard BS8006 for the design of piled embankments, *Geotextiles and Geomembranes* 29 (3) (2011) pp. 345-359. doi: https://doi.org/0.1016/j.geotexmem.2011.02.001.

[27] L. Briançon, B. Simon, Pile-supported embankment over soft soil for a high-speed line, *Geosynthetics International* 24 (3) (2017) pp. 293-305.
doi: https://doi.org/10.1680/jgein.17.00002.

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 477-487, 2021
DOI: 10.14513/actatechjaur.00593

Methods and simulation to reduce fuel consumption in driving cycles for category N1 motor vehicles

A.A. Kolin*, S.E. Silantyev, P.S.Rogov, M.E. Gnenik

**Nizhny Novgorod State Technical University n.a. R.E. Alexeev,
Nizhny Novgorod, 24 Minin street, 603950, Russia.
*e-mail: kolinaa@nntu.ru**

Submitted: 27/01/2021; Accepted: 14/04/2021; Published online: 19/05/2021

Abstract: The paper presents the results of using the simulation model estimating the fuel consumption of a light commercial vehicle in road traffic cycles; virtual tests are performed. The impact analysis of the motor vehicle design parameters on fuel consumption in NEDC and WLTC cycles is conducted. Numerical values of average fuel consumption are obtained for variation of the main parameters of the structure in NEDC and WLTC cycles. Energy distribution is shown during the motion of category N1 light commercial vehicle.

Keywords: fuel consumption in cycles; simulation model; light commercial vehicle

1. Introduction

Fuel consumption is an important criterion that determines the attractiveness of a car for a costumer. All of the world's leading automakers are committed to reduce fuel consumption [1]. The continuous rise in fuel prices and deterioration of the environment associated with higher intensity traffic lead to creation and implementation of single-approach methods that are used to determine and estimate fuel consumption, economic and environmental properties based on standard driving cycles.

The most common driving cycles include the New European Driving Cycle (NEDC) [2], the American Driving Cycle (FTP-75) and the Japanese Driving Cycle (JC-08).

Additionally, the World Harmonized Driving Cycle (WLTC) [3] was introduced in 2017. It is based on the world statistical study of driving modes and features high accelerations and lack of steady-motion intervals.

The purpose of the study is to build a simulation model quantifying the effect of the design parameters on the fuel efficiency indicator of N1 category light commercial vehicles in NEDC and WLTC driving cycles. The model building environment is matlab/simulink.

It is to be noted that a number of papers are devoted to simulation of a motor vehicle driving cycle [4] [5] [6]. Main methods of improving the fuel efficiency of vehicles are specified in earlier studies [7] [8] [9] [10] [11]. These include the engine parameters optimization, reduction of aerodynamic drag and tire rolling resistance. Also, a number of works are devoted to the study and search for optimal transmission ratios [12] [13] [14] [15] [16] [17]. Everything stated above should have an impact on the result of the vehicle motion process simulation in driving cycles.

A distinctive feature of this work is a detailed study of the dynamics and fuel efficiency of a light commercial vehicle and obtaining of the quantitative influence of each of the motion resistance forces on the total motion energy in the NEDC and WLTC cycles.

2. Methods

The light commercial vehicle motion on NEDC and WLTC cycles is considered in the paper. NEDC cycle consists of one urban driving cycle or 4 simple urban cycles of 195 seconds, and one suburban driving cycle of 400 seconds.

Driving cycle is performed using a detailed operational map included in the standard [1]. The map details the gears of the gearbox to be used for every section of the vehicle motion as well as its accelerations. Application of the standard: transport vehicles of category M1, M2, N1 and N2 with the reference mass not exceeding 2 610 kg. The vehicle driving diagrams based on NEDC cycle are shown in Fig. 1.

As shown in Fig. 2, WLTC cycle is separated by short stops into four phases:

Low-speed phase when the vehicle accelerates to maximum 56.5 km/h; medium-speed phase (76.6 km/h), high-speed phase (97.4 km/h) and extra-high speed phase (131.6 km/h).

In each phase of the cycle the driving is performed based the operational map according to [2] that details basic requirements for the gear shift timing depending on the power-to-weight ratio class of a vehicle.

Extra high-speed phase was not included in the study which is acceptable for the vehicle class under review.

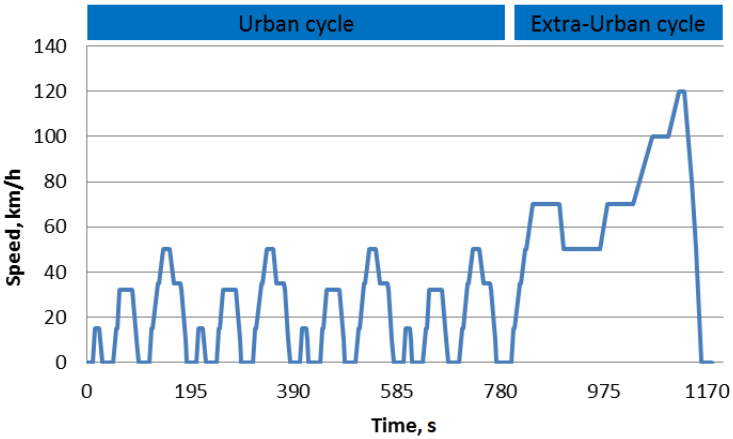


Figure 1. Vehicle motion diagram based on NEDC driving cycle

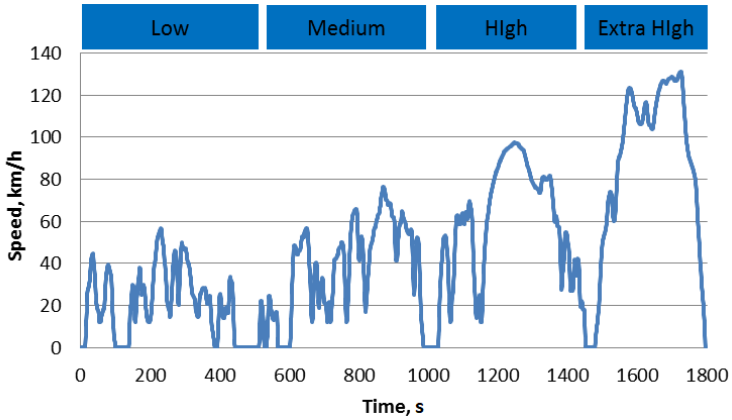


Figure 2. Vehicle motion diagram based on WLTC driving cycle

The simulation model is used for simulation of the vehicle motion. The mathematical formulation of the vehicle motion is based on the following approaches [18].

The traction force is computed by formula:

$$F_T = \frac{T_e u_{tr} \eta_{tr}}{r_w}, \quad (1)$$

where T_e – internal combustion engine (ICE) torque, u_{tr} – transmission gear ratio, η_{tr} – transmission efficiency, r_w – wheel rolling radius.

T_e is normally a function of the accelerator pedal position d(%) and ICE speed ω_e(rad/s), or $T_e = f(d,\omega_e)$. T_e is represented by data array $T_{e\,i,j}$, i =1…17, j =1…11 in the mathematical model. In this particular case, for acceleration of the vehicle at full throttle j = 11 and $Te = f(\omega_e)$.

The rolling resistance force is determined as [18, p. 22]:

$$F_f = fm_a g;\ f = f_0 + k_f V^2 \quad (2)$$

where f – coefficient of rolling resistance; g – gravity acceleration.f_0 – rolling resistance coefficient at a speed close to 0 km / h; k_f – coefficient that takes into account the influence of speed; V – vehicle speed.

Air resistance is determined as follows:

$$F_w = 0{,}5 C_x \rho_{air} A_v V^2, \quad (3)$$

where C_x – aerodynamic drag coefficient; ρ_{air} – air density; A_v – transverse projection area of the vehicle.

The rotational inertia coefficient is determined as:

$$\delta = 1 + \frac{I_e u_{tr}^2 \eta_{tr}}{m_a r_к^2} + \frac{I_{ds} u_{fd}^2 \eta_{fd}}{m_a r_w^2} + \frac{\sum I_w}{m_a r_w^2}, \quad (4)$$

where I_e – ICE moment of inertia, I_{ds} – drive shaft moment of inertia, u_{fd} – final drive gear ratio, η_{fd} – final drive efficiency, I_w – moment of inertia of a wheel with half-axle (if any). Moments of inertia are determined from design documentation.

When in a steady-speed driving mode, the vehicle speeds up without acceleration. ICE torque and the instantaneous fuel consumption are the unknowns. The unknowns are computed by the following formula:

$$F_т = F_f + F_w + \delta m_a a_a, \quad (5)$$

when ICE torque T_e is calculated, given (1) and the engaged transmission gear that is determined with account for (6). The specific fuel consumption in the model is represented by data array$g_{e\,i,j}$, and is a function of ICE torque and data array $g_e = f(T_e,\omega_e)$. Therefore, the current specific fuel consumption BSFC is determined according to (6) using the Te and we values obtained with the (1), (5), (7).

$$BSFC = \frac{m_f}{P};\ P = T_e \omega_e \quad (6)$$

where m_f – fuel consumption rate in grams per second (g/s); P - power (W); ω_e – engine speed (rad/s)

Gear shifting is performed as required in [2] [3]. After gear shifting the engine's starting speed of rotation is determined subject to correlation of the vehicle speed and ICE speed of rotation:

$$V = \frac{\omega_e r_w}{u_{tr}}, \tag{7}$$

The acceleration of the vehicle is computed using the following formula:

$$a_a = \frac{V_{j+1} - V_j}{3{,}6(t_{j+1} - t_j)}, \tag{8}$$

where a_a (m/s) – acceleration of the vehicle; V_j, V_{j+1} (km/h) – initial and final speeds over the selected interval, respectively; t_j, t_{j+1} (s) – start and end times, respectively.

Input data for the simulation, such as the mechanical characteristics of the ICE and the data on its fuel consumption was obtained by performing bench tests. The value of the air drag coefficient was obtained using numerical simulation. Other input data were taken from the technical documentation of the automaker and component manufacturers. To be confident in the reliability of the results, obtained in this formulation, the developed model is then verified. Model initial data in Table 1.

3. Results of the study

Simulation model must reliably react to a change of input data i.e. the reduction of the aerodynamic drag coefficient contributes to lower fuel consumption while the higher aerodynamic drag coefficient results in higher fuel consumption. It can be stated the verification of the model was carried out successfully only if the response of the model is adequate and proportional to the change in the initial data. The simulation results for a light commercial vehicle are specified in Table 2.

It is to be noted that the results obtained are indicative of the indisputable impact of these parameters on the average fuel consumption. An increase of the drive shaft efficiency by 5% brings down the average fuel consumption rate by 0.36%, for NEDC and WLTC cycles, and vice versa.

Table 1. Model initial data.

Parameter	***Value***
Transmission	
Gear ratio	1 – 3.786
	2 – 2.188
	3 – 1.304
	4 – 1
	5 – 0.794
	6 – 0.643
Differential	
Final drive ratio	4.3
Car	
Vehicle kerb weight, kg	2548
Vehicle draft coefficient	0.356
Vehicle frontal area, m^2	5.375
Tire	
Tire size	185/75R16

Increasing the final drive and the gearbox efficiency by 1% in NEDC cycle brings down the average fuel rate by 0.6%, and 0.48%, respectively. For WLTC cycle the change of these parameters results in the reduction of the average fuel consumption rate by 0.72% and 0.66%, respectively.

If the aerodynamic drag coefficient is reduced by 5%, the average fuel rate becomes less by 0.84% for NEDC cycle, and 0.88% for WLTC cycle.

The rolling resistance factor will have less impact as compared to other indices reviewed, since its variation to the extent of 5% causes the average fuel rate to change within the limits of 0.4% for NEDC cycle, and 0.66% for WLTC cycle.

In general, the obtained results show an adequate response of the model to changes in the initial data and, thus, it can be stated that the model was verified successfully.

This study is the initial stage of a comprehensive search and development of practical recommendations for reducing the fuel consumption rate of a motor vehicle in the physical world. As the resulting simulation model allows a qualitative assessment of the vehicle motion parameters, it is decided to evaluate the main areas of further activities. Specifically, to analyse the distribution of the energy consumed on the movement of a light commercial vehicle.

Table 2. Impact of specified parameters for NEDC and WLTC cycles.

Index	***Index variation, %***	***Driving cycle***			
		NEDC		***WLTC***	
		Fuel rate, l/100 km	***Change of fuel rate, l/100 km***	***Fuel rate, l/100 km***	***Change of fuel rate, l/100 km***
Aerodynamic drag coefficient	5	16.87	0.14	13.79	0.12
	0	16.73	0	13.67	0
	-5	16.59	-0.14	13.55	-0.12
Rolling resistance factor	5	16.80	0.07	13.76	0.09
	0	16.73	0	13.67	0
	-5	16.67	-0.07	13.58	-0.09
Ring and pinion set efficiency	1	16.63	-0.1	13.57	-0.1
	0	16.73	0	13.67	0
	-1	16.83	0.1	13.77	0.1
Drive shaft efficiency	0.5	16.67	-0.06	13.62	-0.05
	0	16.73	0	13.67	0
	-0.5	16.79	0.06	13.72	0.05
Gearbox efficiency	1	16.65	-0.08	13.58	-0.09
	0	16.73	0	13.67	0
	-1	16.81	0.08	13.76	0.09

The motion resistance energy values were determined by integrating the corresponding motion resistance forces over the distance of the driving cycle. The energy expended for the light commercial vehicle motion is specified in Table 3.

It is to be noted that the energy of mechanical losses in transmission can be as high as 15% of the total energy expended on the transport vehicle motion, as specified in table 3. The energy of aerodynamic drag accounts for 30 to 40% of the total energy used for cycles NEDC and WLTC, respectively. It is also of major importance, and will become the main area of focus in terms of reduction of the fuel consumption rate.

Table 3. Energy consumption of the vehicle motion for NEDC and WLTC cycles.

	Distribution of energy	***Driving cycle***	
		NEDC	***WLTC***
Total resistance of motion energy	Aerodynamic drag energy, %	40.2	30.7
	Rolling resistance energy, %	26.3	26.8
	Speed up resistance energy, %	18.8	28.4
Energy of mechanical losses in transmission	Half-axles and hubs, %	5.9	5.1
	Gearbox, %	4.4	4.5
	Final drive, %	3.3	3.3
	Drive shaft, %	1.1	1.1

4. Conclusions

According to Tables 2 and 3, the following results are obtained.

Increasing the drive shaft efficiency by 5% results in 0.36% reduction of the average fuel consumption rate for a light commercial vehicle in NEDC and WLTC cycles, and vice versa.

Increasing the final drive and the gearbox efficiency by 1% in NEDC cycle brings down the average fuel rate by 0.6%, and 0.48%, respectively. Changing these parameters in WLTC cycle results in the reduction of the average fuel consumption rate by 0.72% and 0.66%, respectively.

If the aerodynamic drag coefficient is reduced by 5%, the average fuel rate also becomes less by 0.84% for NEDC cycle, and 0.88% for WLTC cycle.

The rolling resistance factor will have less impact as compared to other indices reviewed, since its variation to the extent of 5% causes the average fuel rate to change within the limits of 0.4% for NEDC cycle, and 0.66% for WLTC cycle.

It is to be noted that according to table 3 the energy of mechanical losses in transmission can be as high as 15% of the total energy used for the transport vehicle motion. The aerodynamic drag energy can be as high as 30 to 40% of the total energy used for cycles NEDC and WLTC, respectively.

Given the simulation results, the highest potential for reducing the average fuel rate of a light commercial vehicle is associated with the improvement of its aerodynamic shape. Higher transmission efficiency can have a substantial impact on

the average fuel consumption rate, however, the range of improvement of this index is limited. The rolling resistance factor has the least impact on the parameter of interest. All the parameters listed above can be optimized together that can help to achieve the desired fuel efficiency index of a vehicle.

Acknowledgement

The studies are conducted under the sponsorship of the Ministry of Education and Science of the Russian Federation as part of the project High-end production of the product range of GAZelle Next vehicle fitted with new electronic systems architecture under Agreement No. 075-11-2019-027 dated 29.11.2019 (Decree No.218 dated April 09, 2010 of the Government of the Russian Federation).

References

[1] U. Tietge, P. Mock, N. Zacharof, V. Franco, Real-world Fuel Consumption of Popular European passenger car models, ICCT working paper 2015-8 (2015) URL *https://theicct.org/sites/default/files/publications/ICCT_Real-worldFC-EUcars_28122015.pdf*

[2] Regulation No. 83. Revision 5.Uniform provisions concerning the approval of vehicles with regard to the emission of pollutants according to engine fuel requirements (2015).
URL
https://unece.org/fileadmin/DAM/trans/main/wp29/wp29regs/R083r5e.pdf

[3] Global technical regulation No. 15. Worldwide harmonized Light vehicles Test Procedure (2014).
URL
https://unece.org/fileadmin/DAM/trans/main/wp29/wp29r-1998agr-rules/ECE-TRANS-180a15e.pdf

[4] S.V. Gusakov, D.V. Markov, D.V. Mikhryachev, Experiment-calculated method to adjust a drive cycle for vehicle motion phase in urban conditions, Proceedings of Higher Educational Institutions. *Machine Building* 5 (2012) pp. 23–30, in Russian.

[5] A.V. Manyashin, S.A. Manyashin, Issledovanie gorodskih ezdovyh ciklov dvizheniya, Transportnye i transportno-tekhnologicheskie sistemy. Materialy

Mezhdunarodnoj nauchno-tekhnicheskoj konferencii, Tyumenskij industrial'nyj universitet, Tyumen', 2010, pp. 113–114, in Russian.

[6] S.I. Antipov, Yu.V. Dementiev, Sovremennye ispytatel'nye ezdovye cikly i ih aktual'nost' pri sozdanii algoritma raboty sistemy upravleniya avtomobilya s keu, Izvestiya volgogradskogo gosudarstvennogo tekhnicheskogo universiteta. *Seriya: nazemnye transportnye sistemy* 10 (113) pp. 8-11, in Russian.

[7] A.N. Evgrafov, N.S. Kuzovkov, Aerodynamic of automobile, *Trudy NAMI* 245 (2010) pp. 57-70, in Russian.

[8] D.Kh.Valeev, V.S. Karabtsev, Methods for fuel consumption reduction in commercial vehicles, *Mechanics of machines, mechanisms and materials* 4 (29) (2014) pp. 33-39, in Russian.

[9] V.A. Petrushov, Some milestones in the development of nami works on aerodynamics and rolling resistance of vehicles and road trains, *Trudy NAMI* 4 (279) (2019) pp. 11-21, in Russian.

[10] Transportation Research Board and National Research Council, Technologies and Approaches to Reducing the Fuel Consumption of Medium- and Heavy-Duty Vehicles, The National Academies Press, Washington, 2010. doi: https://doi.org/10.17226/12845

[11] M. S. Lyu, Optimization on vehicle fuel consumption in a highway bus using vehicle simulation, *International Journal of Automotive Technology* 7 (7) (2006) pp. 841-846.

[12] N.V. Savenkov, Method of selection of gear ratios transmission of the n1 category vehicle based on the driving cycle, *Bulletin of the Kuzbass State Technical University* 2 (114) (2016) pp. 64-71, in Russian.

[13] P. Bera, D. Wedrychowicz, The influence of number and values of ratios in stepped gearbox on mileage fuel consumption in NEDC test and real traffic, Scientific Conference on Automotive Vehicles and Combustion Engines (KONMOT 2016), *IOP Conf. Series: Materials Science and Engineering* 148 (2016). doi: https://doi.org/10.1088/1757-899x/148/1/012001

[14] W. Wang, C. Chen, Sport utility vehicle refinement by engine upgrading and gear ratio optimization, International Journal of Advanced Science and Technology 54 (2013) pp. 61-76.

[15] C. J. Oglieve, M. Mohammadpour, H. Rahnejat, Optimization of the vehicle transmission and the gear-shifting strategy for the minimum fuel consumption and the minimum nitrogen oxide emissions, *Proc IMechE Part D: J Automobile Engineering* 231 (7) (2017) pp. 883-899. doi: https://doi.org/10.1177/0954407017702985

[16] İ. Küçükoğlu, Differential Evolution Algorithm for Gear Ratio Optimization of Vehicles, *Int. Journal of Engineering Research and Application* 6 (10) (2016) pp. 29-33.

[17] D. Robinette, D. Wehrwein, Automatic Transmission Gear Ratio Optimization and Monte Carlo Simulation of Fuel Consumption with Parasitic Loss Uncertainty, *SAE Int. J. Commer. Veh.* 8 (1) (2015) pp. 45-62. doi: https://doi.org/10.4271/2015-01-1145

[18] V.N. Kravets, Teoriya avtomobilya (Automobile theory), Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhny Novgorod, 2013, in Russian.

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 488-507, 2021
DOI: 10.14513/actatechjaur.00606

Modelling the ZalaZONE Proving Ground: a benchmark of State-of-the-art Automotive Simulators PreScan, IPG CarMaker, and VTD Vires

K. Gangel[1], Z. Hamar[1], A. Háry[2], Á. Horváth[1], G. Jandó[1], B. Könyves[3], D. Panker[1], K. Pintér[4], M. Pataki[1,6], M. Szalai[6], Zs. Szalay[1,5], T. Tettamanti[6,*], V. Tihanyi[1,5], B. Tóth[9], B. Varga[6], Zs. J. Viharos[7,8]

[1]Automotive Proving Ground Zala Ltd., ZalaZONE Tér 1, H-8900 Zalaegerszeg, Hungary
[2]ZalaZONE Industry Park Ltd. Dr. Michelberger Pál út 3, H-8900 Zalaegerszeg, Hungary
[3]Continental Automotive Hungary Ltd., Házgyári u. 6-8, H-8200 Veszprém, Hungary
[4]Bay Zoltán Alkalmazott Kutatási Közhasznú Nonprofit Ltd., Kondorfa u. 1, H-1116 Budapest, Hungary
[5]Budapest University of Technology and Economics, Department of Automotive Technologies, Stoczek J. u. 6, H-1111Budapest, Hungary
[6]Budapest University of Technology and Economics, Department of Control for Transportation and Vehicle Systems, Stoczek J. u. 2, H-1111 Budapest, Hungary
[7]Institute for Computer Science and Control, Kende u. 13-17, Budapest H-1111, Hungary
[8]John von Neumann University, Izsáki str. 10, H-6000 Kecskemét, Hungary
[9]TÜV Rheinland - KTI Kft., Than Károly u. 3-5, H-1119 Budapest, Hungary
[*]e-mail: tettamanti@mail.bme.hu

Submitted: 17/03/2021; Accepted: 13/08/2021; Published online: 12/11/2021

Abstract: In our days, simulation based development is a core element of vehicle engineering, especially considering highly automated or fully autonomous vehicles. Accordingly, the paper presents a benchmark of

three different automotive simulators: PreScan, IPG CarMaker, and VTD Vires. The three software tools were applied for the same goal, namely, modelling the ZalaZONE Proving Ground of Hungary for vehicle testing. The paper aims to highlight the experiences while creating the virtual models by presenting and comparing the relevant software features and providing suggestions for scientific or practical application.

Keywords: *automotive proving ground; simulation software; automotive testing and validation; autonomous vehicles; virtual testing*

1. Introduction

Autonomous driving became one of the most researched topics of today's vehicle industry. The interest of the society is growing continuously due to the various beneficial effects like the vision of zero accidents, the increasing mobility, and the decreasing pollution, which can change our everyday life as well [1]. However, self-driving cars have to perform well in those difficult traffic situations where the most developed vehicles mounted with the newest advanced driver assistance systems currently fail. To cover the arising wide range of driving situations, the number of test scenarios and tested kilometres must be increased also continuously [2]. It is already a standard practice in the automotive industry to perform a part of the test kilometres needed for development through simulations. Several virtual testing software tools are available and the number of public datasets is also increasing to create even more realistic scenarios [3]. The software mentioned in the title, *IPG CarMaker, VTD Vires*, and *PreScan* are such best practice virtual testing software. These tools are already well known in the automotive industry, providing a variety of realistic and high-precision models for simulations to guarantee real-world testing. The paper reports the applications and comparisons of the experiences on all the mentioned, three software tools for proving ground test simulation.

There are different levels of simulations determined for testing some or all features of vehicles with self-driving capabilities. Typically, numerous scientific research deals with the simulation of vehicle functions on various levels. The lowest level of simulation means testing of generic models and the highest level is the simulation of (almost) all vehicle functions. However, the issue of simulation focusing specifically on proving ground tests has not yet investigated, it is one of the essential contributions of the presented paper.

A related work in [4] provides opportunities for multi-level testing, from Software-in-the-Loop (SiL) tests to Vehicle-in-the-Loop (ViL) procedures. Many forward-looking vehicle functions and ADAS (Advanced Driver Assistance System) systems

have already being tested through simulations, e.g., this study tests a new feature in Active Safety Light in ViL using *VTD Vires* [5]. ViL testing is also used in the work of [6] by extending the test environment with augmented reality technology to simulate virtual interference in a real-world environment. The contribution in the article [7] reports also on simulations of a Cooperative Collision Warning function. In the case of ADAS systems, the simulation-driven validation process may also play an important role [8], but some companies are also addressing the possibility of virtual validation, e.g., Siemens with PreScan [9]. The article [10] deals with the "mixed virtual and real environment" testing capabilities, in which the virtual test environment also plays a vital role, and a "proof of concept" has been implemented for that concept in [11].

The authors of this article are all experienced in the use of simulation software and are all related to the R&D works at the ZalaZONE Automotive Proving Ground. 2. ZalaZONE is a unique proving ground for autonomous vehicle technologies constructed near Zalaegerszeg, Hungary (see Section 2. for more details). One of the main goals at ZalaZONE is to facilitate the testing of future vehicles with highly automated and self-driving functions by providing state-of-the-art procedures and technologies to support the execution of vehicle test simulations. Therefore, creating an accurate, virtual model of the complete ZalaZONE Proving Ground was realized in each of the three simulation software mentioned above, which also provides the opportunity to perform tests in this virtual environment, and later in a unique framework created by combining virtual and real-world environments [12]. The paper's goal, in general, is to draw a detailed picture of experiences while creating the ZalaZONE virtual models in these different software tools. It is important to emphasize that this article aims not to prioritize or evaluate any of the software tools but to present and compare their relevant features and serve with advices during their scientific and practical application as well.

The structure of the paper is as follows. Section 2 briefly introduces the ZalaZONE Proving Ground. Section 3 presents the applied simulation software tools, in the next, fourth section, the experiences about the different software environments are provided with particular emphasis on the modelling process and the construction of the test track modules. Conclusions, acknowledgements and references close the paper.

2. ZalaZONE, a unique proving ground for autonomous vehicle technologies

The ZalaZONE proving ground is located in Zalaegerszeg, Hungary, Europe. The traditional test track features focusing on driving and driving stability are implemented together with the research and development infrastructure elements for

future vehicle validation. The proving ground provides not only dynamics tests for conventional vehicles, but it also allows validation tests for autonomous vehicles and electric vehicles as well [13]. The construction works of the ZalaZONE Proving Ground started in 2017 and will be fully completed in 2020 (Fig. 1). ZalaZONE (beyond the classic vehicle dynamics testing) enables to carry out handling and control scenarios, test cutting-edge automotive technologies (such as electric vehicles and self-driving vehicles) as well as communication technologies [14]. The 250 ha area (Fig 2) incorporates the following test features:

- standard vehicle dynamics testing and validation,
- fully integrated autonomous vehicle testing and validation,
- environment preparation (obstacles, traffic signs, traffic control, other vehicles, vulnerable road users),
- complex driving and traffic situations,
- smart city area,
- testing and validation, from prototype testing to mass production.

ZalaZONE Proving Ground enables complex test cycles on either classical or interconnected elements for automated/autonomous vehicles. The main elements of the ZalaZONE test track are listed briefly below.

- Dynamic Platform:
  The Dynamic Platform is a special, continuous, asphalt paved element that provides the ability to perform high-speed tests under safe conditions. The dynamic element is a circular surface with a diameter of 300 m, provided with multi-layer asphalt pavement. To the west side of the Dynamic Platform, a 740 m long acceleration track is connected for safe acceleration during tests.

- Braking Platform:
  The braking platform is a track element designed for testing ABS, ASR, and ESP systems with special pavement elements and a built-in irrigation system. This module offers six different lanes with different frictions providing different braking conditions. Moreover, the Braking Platform has the possibility of wetting by lanes and has corresponding drainage. With an acceleration lane of almost 700 m and a brake surface of 200 m, the Braking Platform makes also possible to test long distance brake tests.

Figure 1. Visualization image of the Automotive proving ground dedicated to autonomous vehicle testing and validation at Zalaegerszeg city, Hungary, Europe (GPS: 46.889771, 16.836956)

- Handling Course:
  The Handling Course is a test track element where it is possible to examine vehicle behavior, vehicle handling, and test different technical settings under controlled traffic conditions. The ZalaZONE test track contains an outer and an inner handling course section, for higher and lower speeds respectively. The total length of the high speed track is 2030 m with the pavement width of 12 m. Here, the length of straight sections and the arc radius allow 120 km/h to be reached. On the low speed handling track similar tests can be performed up to 60 km/h. Its overall length is 1330 m with the pavement width of 6 m.

- Internal Road Network:
  The primary role of the internal road network is to provide access to the various modules. At the same time, the internal road network of the test track is designed to meet road standards, allowing road tests and providing a connection between the elements such that the test vehicles can move from one element to the next without turning. This has an important role in the multi-element, longer cycle, special tests, even for autonomous vehicles.

- High Speed Oval:
  The term high-speed oval means a closed course which, without significant deceleration, allows continuous driving in a given direction even at high speed. The proving ground has a high-speed oval track which has

curved bends with a neutral speed. The High Speed Oval track enables the vehicle turning with released steering wheel at 200 km/h due to the geometry of the pavement.

- Smart City Zone:
  One of the uniqueness of the ZalaZONE test track is its built-in urban environment, called the Smart City Zone, which allows testing automated or fully autonomous vehicles in complex urban situations. The Smart City Zone includes different types of roads, intersections, and other features that are common in real-world urban area, e.g. bus stops, facades, traffic signs, or traffic lights. In addition to the listed items, bike paths, roundabouts, tram rails, low-adhesion surfaces, rainfall sections, and parking garages help the testing procedures.

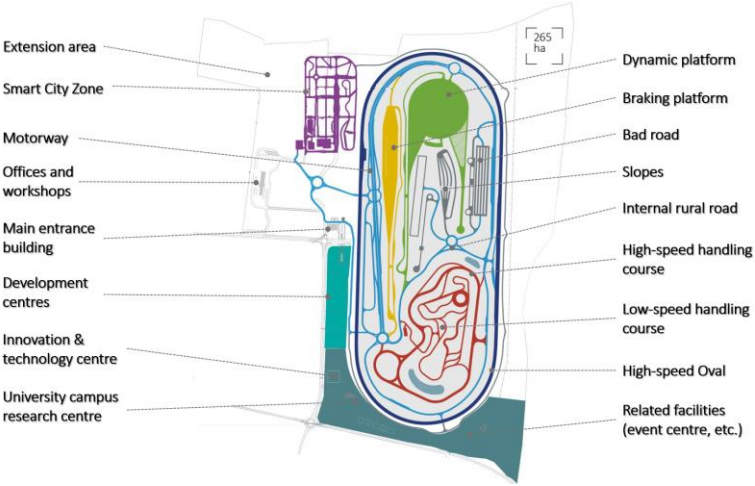


Figure 2. Main elements of ZalaZONE Proving Ground (AD in figure means Automated Driving)

3. State-of-the-art simulation tools at ZalaZONE

In ZalaZONE, currently, three state-of-the-art vehicle simulation software tools are in daily use. In the next paragraphs, three short summaries are provided based on the working experiences of the ZalaZONE virtual environment creation.

3.1. IPG CarMaker (Version 8.1) – IPG Automotive GmbH

IPG CarMaker is an automotive simulation software of IPG Automotive GmbH (*https://ipg-automotive.com/company/*) specializing in passenger cars. The vehicle models available in it are characterized by sophisticated, all-round adjustments, and their vehicle dynamics models are efficient and accurate. Various tests can be performed using pre-set scenarios according to a measurement cycle set by a standardization authority. However, other methods, defined by the researcher for his own research purposes can also be used. It is possible to perform conventional vehicle dynamics tests even in the presence of the Advanced Driver Assistance System (ADAS). *So-called "third-party" software (for example, MATLAB/Simulink, AVL Cruise, GT Suite, etc.) can also provide interfaces to refine the vehicle model and extend the modelling features that is also a very important functionality.*

The environment building function is eligible to build simpler road network for standard tests, but it is possible to create more complex test environments by making some compromises. The so-called Scenario / Road module is used to create test environments. Newer versions of the software support the more and more popular and recognized OpenDRIVE ("de facto standard") format [15], which is widely used by the automotive industry to describe road networks.

The trajectories of the created road networks can be linked to real locations based on geographical latitude, longitude, and altitude, which can be simultaneously checked with an external map software and can even track vehicle movements during simulations.

3.2. Prescan (version 2019) – Siemens PLM

PreScan is another vehicle simulator software developed by Siemens PLM (*https://www.plm.automation.siemens.com/global/en/products/simcenter/prescan.html*). *PreScan* is primarily desgined to test and simulate various sensors. In the secondary aspect, the vehicle dynamics used within the program is noteworthy. There are several sensor types of sensor models in the software and it is possible to use them as a real sensor, or an "idealized" one. Sensor settings, as well as vehicle parameterization and use of their functions, are available through the *Matlab/SIMULINK* software package.

The software is straightforward and user-friendly in the creation of an environment; it does not use a separate module; the central GUI is suitable for environment building. It has a relatively limited toolbar for creating road sections but contains unique elements (such as roundabouts) that are not available in the other software. A wide range of additional items can fit into the environment (e.g., trees, buildings, boards). *Experience from modeling the ZalaZONE test track has shown*

that PreScan allows extremely easy track construction, even with a limited set of tools. By being compatible with *Matlab/SIMULINK*, there is plenty of Add-On available, allowing for a much more extensive range of testing and scenario building. Another advantage of the program is the simple trajectory drawing method. It is sufficient for a logically well-built road network to click on the connection points to draw the route.

PreScan can also work with other software (e.g. *IPG CarMaker*) using *MATLAB/Simulink* software package. It also allows for testing of built-in ADAS functions. It provides simulation capabilities in Model-in-the-Loop (MiL), SiL, and Hardware-in-the-Loop (HiL) environments with the ability to automate tests.

Another benefit of the software package is that its help guide is exceptionally user-friendly, describes many features with the use of the existing tools (e.g. Auto Emergency Braking (AEB) functions, creation of various ADAS scenarios).

3.3. Vires Virtual Test Drive (version 2.2) – Hexagon, Vires Simulationtechnologie GmbH

VTD Vires is also one of the most well-known automotive simulation software (`https://vires.mscsoftware.com/`). The software provides the ability to perform HiL, SiL, DiL (Driver-in-the-Loop), ViL tests. The software consists of two distinct parts: RoadDesigner and Scenario Editor. The program has the advantage of supporting OpenDrive, OpenCRG, and OpenScenario file formats, which are native formats of VTD Vires. It can likely be a strong basis for vehicle simulation software in the future [16]. It is possible to interface external software and modules to simulations such as Matlab/SIMULINK or other third-party software.

The RoadDesigner module allows the creation of different, well detailed virtual environments. With this tool, the virtual test environments can be more realistic, and it is easier to create a digital twin of a real test area based on the *OpenDrive* standard. With the Scenario Editor, one can create simulations of varying degrees of details to test basic vehicle and ADAS functionalities. Different sensor models can also be attached to the vehicle to extend the types of possible test cases with environment sensing simulations. Some of the Scenario parameters, such as weather, time, and time of day, can be changed in real-time when running simulations.

4. Benchmark of automotive simulation software programs via modelling proving ground modules

The main goal of the paper is to analyse and compare the three simulation software from a road modelling perspective. It is emphasized that rating of any of the investigated software is not planned because each has its advantages from a different

point of views, i.e., in this paper, only the comprehensive experiences related to ZalaZONE Proving Ground modelling are provided

4.1. Environment design basics in the various software

In order to introduce the modelling capabilities of the software, a 'dummy' environment was created that contains all the typical road network elements. Hence, the created model includes a roundabout, straight sections with slopes, junction, turn, and a multilane section with entrance and exit lanes. The created modules are shown in Fig. 3 (a – *IPG CarMaker*, b – *PreScan*, c – *VTD Vires*).

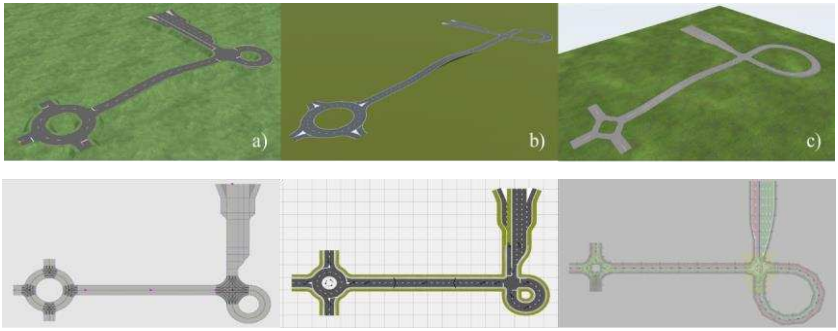


Figure 3. The dummy road-network in different software (a – CarMaker, b – PreScan, c – VTD)

In *IPG CarMaker*, one can find under the Parameters menu the Scenario/Road editor, which enables creating a new test environment. Road type can be selected easily from the menu, and then road drawing is straightforward. One can connect new road segments to the existing area, then parameterize them, and so on. There are more groups for the user like Road, Accessories, Scenery, Traffic, but the Road objects are the most important for the modelling. The Road objects group contains the following: Straight, Turn, Junction, Ram, Clothoid, Point list, Connect, and File. The last one provides the opportunity to import other file formats or existing road network segments.

Additional elements like Road marking, Traffic sign, Bridge, and Tunnel were also relevant for our work. After placing the Road objects, many more parameters can be set, e.g., the lanes' width or their lateral shift separately. More lanes can be also added to existing roads, and other features can be assigned to them afterward, like Elevation, Slope, and Camber profiles. The surfaces can be modified with the following elements: Beam, Cone, CRG, Friction, Mesh, Wave, or Lateral profile. Besides, additional elements can be added to each road segment, like a guard rail, lane markings, or textures for the different surfaces. Connecting two separate

existing road sections, connect command has to be used, or the two roads must end at the same coordinates. In this case, there is the possibility of establishing the connection, even using a junction. The software does not include roundabouts, so these are built using junctions and turns. The dummy example road network can be built as described above. After the roundabout building, the straight road sections can be connected to it, and then the straight line elevation can be added using the Elevation profile. A junction can then be added to the existing road, and its two exits can be connected using turn roads. An expanding highway-type road can be connected to the last exit, which is turned into 2x2 lanes in each direction. The final steps are the placement of the guardrails and the road markings (e.g. dashed and continuous lines, double closing lines, perpendicular white lines for a give-way situation, etc.). The example model is illustrated in Fig. 3 – a).

In *PreScan*, it is possible to access the builder elements from the Library Elements menu within the GUI. Each element can be selected and moved to the drawing area with a Drag & Drop method. Each item that is dropped in the Build Area gets a unique ID. *PreScan* contains several groups of objects (Environment, Infrastructure, Actors, Sensors, etc.), but the infrastructure and environment elements are the most important for the modelling.

The range of elements that can be used to road network building is quite broad, including the following elements: Bend Road, Curved Road (Bézier), Entrance Lane Road, Exit Lane Road, Flexible Road, Lane Adapter Road, Pedestrian Crossing, Ramp, Roundabout, Spiral Road (Clothoid), Straight Road, X Crossing, Y Crossing). Besides, there are several object groups (Road Markings, Buildings, Abstract Objects, Traffic Signs, Animated Elements, Reflectors & Botts' dots, Generated Content, Other). However, from these, the Traffic Sign and Roadsign were the most relevant. Most objects can be parameterized in *PreScan*, and many other parameters can be set for path elements. In the case of a road section, different parameters can be adjusted including, the number of lanes, the curve radius, the width of the different paintings, the quality of the bench, if necessary, assigned to the road a sidewalk also can be added, which can also be parameterized. The speed limit for the road and the detectability by the built-in sensors also can be defined.

In the intersections (X Crossing, Y Crossing) and roundabouts, the properties, orientation, and parameters of the safety island can be set. Besides, guard rails, walls, or other texture can be added to each segment. In *PreScan*, the connection of the different roadblocks is simply by dragging the roads' connection points. It is also possible to connect the elements, but it is also possible to connect the lanes separately in some cases. In *PreScan*, slopes can be created using the Ramp element. The angle of the elevation can be defined by the length and height of the elevated section. One of *PreScan's* advantages is that it includes a predefined Roundabout element, so

there is no need to build it separately. Lane Adapter is available to expand the lanes, but highway acceleration and exit lanes can also be used as separate elements. The dummy example road network can be built as described above.

After the roundabout is dropped in, the straight road sections can be connected to it, followed by the up and down ramp sections in the appropriate direction. To this, a junction can be added, and bend roads can connect its two exits.

To the last exit, a highway-type road with acceleration and exits lanes can be connected. The example file is illustrated in Fig. 2.

In *VTD Vires,* the first step of road and environment modelling is drawing the centreline as the road's base. The length and orientation of this road can be modified at this phase. This is not identified as a road network, but the sketch geometry can be turned into a path. *VTD Vires* is working with the *OpenDRIVE* file format since the path axis will be in the middle of the road. We can place lanes with any width on any side from this axis, which can be various types: driving, none driving, shoulder, bike, sidewalk, border, parking, etc. These different types perform different logical functions (For example, in a bike lane, a car cannot drive in traffic simulations.) The roads are connected using junctions. The roads' ends have to be selected for the connection, and then every possible driving path has to be defined separately.

VTD Vires does not contain a roundabout element, so this should be created using T-junctions and small bend roads. Several small junctions can be connected with a larger one, which gives the final logical connection. Each lane can be assigned different materials, road markings, and additional environmental elements, such as a guardrail. Based on the above, the example road network can be easily built; this is illustrated in Fig. 3 – c).

4.2. Created proving ground modules

Every module was created based on the blueprints (the original technical drawings). In *VTD Vires* and *PreScan* the blueprints could be used as underlays. A scaled blueprint of the test track was inserted on the right coordinate, and the roads were drawn on it. The underlay feature is only available from *IPG CarMaker* version 9.1, but the model was created in *IPG CarMaker* version 8.1.1. Therefore, the model is based on the accurate measurement of the blueprint.

In general, it is not possible to create a road network by hand that matches the blueprints completely, it serves as a basic layer, but with some compromise, the road network on the underlay can be well approximated. A general problem with modelling the test track accurately stems from the limitation of the different software. For example, cornering radii, special intersections, and unique test track elements, they require several overlapping layers. If the right logical connection for

traffic simulation is required, application of additional layers is necessary. *These layers are usually defined vertically on the same level, which can sometimes cause visualization issues, but this compromise is necessary to get a geometrically and logically correct road network.* Thus, when preparing the test track, we tried to keep both geometrical and logical correctness, but in the case of certain elements, this could only be achieved with special solutions. The modules created in each software are listed below:

- High-speed Oval
- Dynamic Platform
- Braking Platform
- High-Speed Handling Course
- Low-Speed Handling Course
- Smart City Zone
- Motorway
- ADAS Surface
- Rural/Connection Roads

In this section, the Dynamic Platform and the Smart City Zone are described more verbosely as these test track elements have some specialties compared to conventional roads or race tracks. The Dynamic platform is a large flat surface with a few entry or exit connectors, and no traffic rules or routes prevail. The Smart City Zone on the other hand, is a typical "urban" environment but contains many unique or "tricky" intersections, roundabouts, and road sections.

4.3. Modelling of the Smart City Zone

In *IPG CarMaker*, the road network is based on the original blueprint of the test track. The model (Fig. 4 upper left) matches the real network with some minor deviations from the blueprint. Some differences are in the southern (figure 4 down middle) and northern section with a small roundabout and narrow T-junctions. These areas can affect the driving behaviour and sensor detection results, causing disparities in the simulations. To solve this problem, these sections are built with different road-width, and curve radius compared to the original blueprint. The software does not contain the special traffic signs of the ZalaZONE Proving Ground or Hungarian traffic signs. Some road signs could be found in the Japanese pack, but mainly third-party programs were used for the creation of the unique signs as shown in Figure 4 lower left. After the right scaling, these could be imported from image files. The road marking placement was simpler because every necessary road marking could be found in IPG's road designer as demonstrated in Figure 4 down right. Although the blueprint contains an elevation profile (Figure 4 upper right), it could not be accurately modelled.

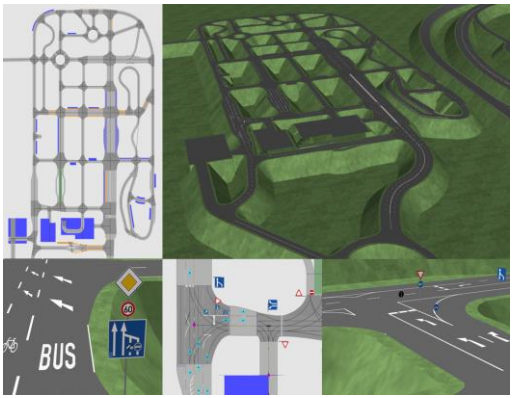


Figure 4. ZalaZONE Smart City Zone in IPG CarMaker

The built model in *PreScan* is logically correct, but there are geometrical differences from reality, mostly at junctions, where the program limits the parameters of the exit radius. Parking places are also planned in several locations, which could be created with the help of simple asphalt or concrete underlays. The most serious difficulties were in building the intersection on the western side shown in Figure 5, inside the Smart City module. The road markings are mostly included in the software, but the road signs set in the toolbox are incomplete for the given purpose. The track elements were built in 2D; hence, the elevation profile is not defined.

In *VTD Vires* the roads are built very close to reality because the model is created using underlays (Fig. 6). The *OpenDRIVE* format is currently under development (standardization), hence some track sections cannot be correctly defined. The parking module is currently unavailable, so for the traffic simulations in the parking areas an invisible underlay must be created with normal roads. The appearance of the parking area can be built by increasing the border width and setting the material (pavement). While the large roundabout was built without difficulties, the small one is still under development due to the limitations of *OpenDRIVE*. The construction is logically correct, mostly. Scenarios can be run on 90% of the smart city zone (merging lanes are not quite right). The biggest problems are the creation of multi-lane junctions and roundabouts. The road markings are placed as planned, but some of the signs are incomplete. There is no elevation profile defined on the track elements.

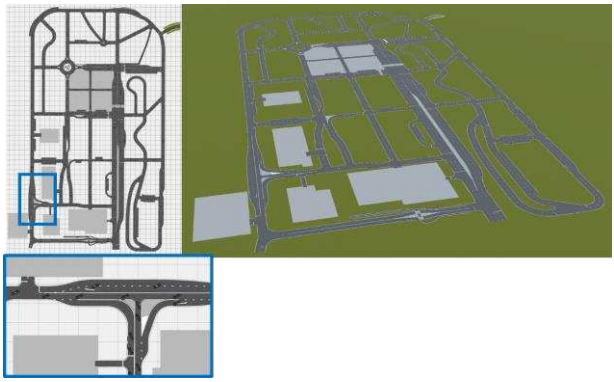


Figure 5. Smart City Zone in PreScan's GUI (upper left and right) and visualization tool with the most difficult junction (lower left)

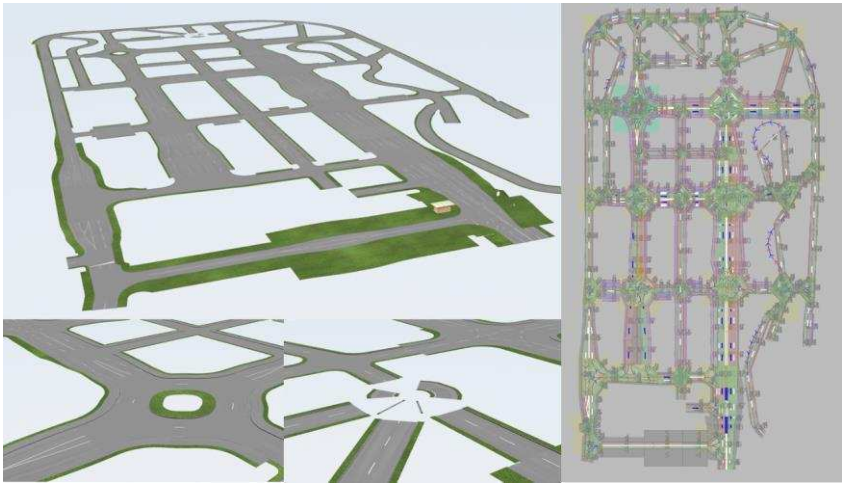


Figure 6. ZalaZONE Smart City Zone in VTD Vires

4.4. Modelling of the Dynamic Platform

The creation of the Dynamic platform (Fig. 7 – a)) is a special task because the surface itself is also unique. Moreover, in *IPG CarMaker* the environment builder tool is not able to generate this full circle as a single road network element. First, an outer arc was built with the necessary junctions. Then, the internal lanes were created. These lanes are connected in one junction in the centre of the full circle and starting from here, and they become wider until they reach the outer arc, where they

are connected with its junctions. They seem like circular-sectors. The triangle-shaped connector parts of the acceleration lanes were built with similar methods. The ending sectors of these lanes have to be equal with the radius of the curves. (Fig. 7 – b)) In this case, the layers of the acceleration triangles cover the original circular sectors, but this is necessary for the right road network connection in order for the vehicles will be able to drive on the full dynamic platform (Fig. 7 – c) d)).

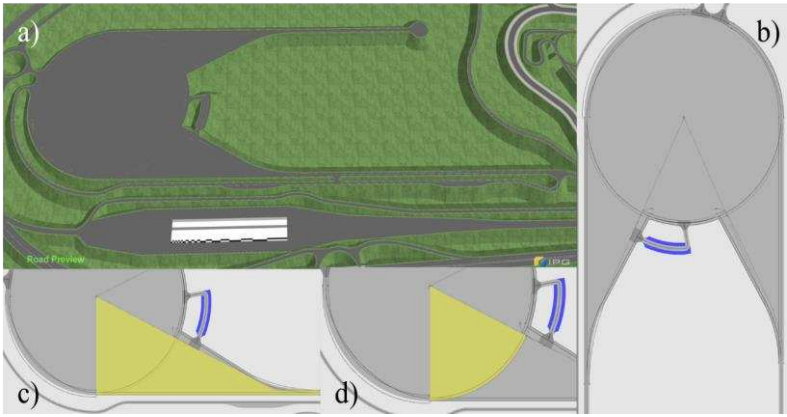


Figure 7. Dynamic platform in IPG CarMaker

PreScan contains the roundabout object, which made the building of the dynamic platform simpler (**Hiba! A hivatkozási forrás nem található.**). For the full circle, two roundabouts are needed because one roundabout is able to handle only ten different lanes. The outer roundabout has all of the exit lanes, which are necessary for the junctions of the platform (like the connectors of the turn-back lane and the parking area). In this software, only just the existing acceleration lane was created. In this case, a similar method was applied as in *IPG CarMaker*, because the acceleration triangle is an overlay that partially covers the outer roundabout. *PreScan* is not sensitive to the missing road connections, but one connected lane was defined between the acceleration lane and the turn-back lane for traffic simulation purposes.

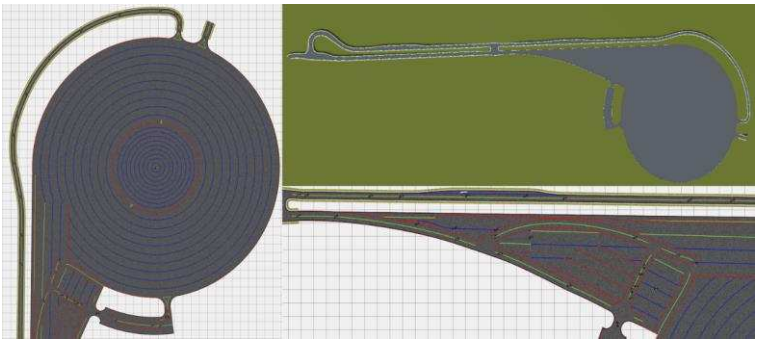


Figure 8. The PreScan model of the Dynamic Platform in the GUI and in the visualization tool

In *Vires VTD*, two different layers were created (Fig. 9). The first layer (Fig. 9 down right) contains all the junctions and necessary road connectors and provides the right road-network logic for the traffic simulation, but this layer is hidden in the visualization. The second layer (Fig. 9 left side) is responsible for the right visual appearance. This could be drawn easily in this software. However, this object itself is not identified as a road object or road-network. The final result is shown by Fig. 9 (upper right).

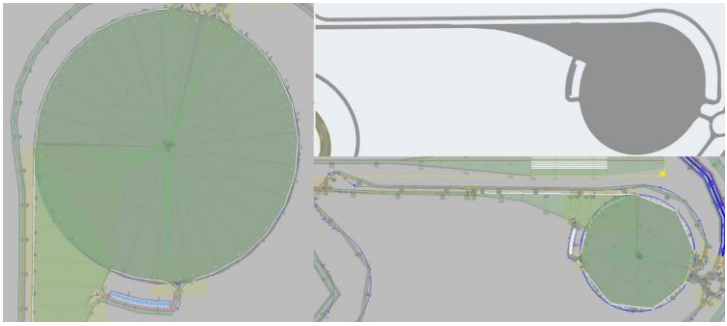


Figure 9. VTD model of the Dynamic platform in ROD editor and in the simulations

5. Conclusion

In this paper, state-of-the-art simulation software tools were presented, in which the ZalaZONE proving ground models have been built up for scientific research and practical vehicle pre-testing purposes. *Through this work presented, advantages, different features, even some limitations and novel, proposed model building techniques of the different simulation tools were identified.* It is important to note

that all of the considered software could have advantages or disadvantages. It ust be noted that the presented work did not aim to prioritize any of the software tools but to present the relevant features and provide useful suggestions concerning scientific or practical application. The strengths and weaknesses of the tools basically depend on the specific usage. *It was clearly observable that these tools are not developed for the creation of road networks with such large size and complexity as the ZalaZONE Proving Ground has*. Some software can handle this easier, or provide the opportunity to save the whole model in separate files.

It can be noted in general that the creation of a road-network that can match the original plans geometrically in the whole area is only possible by using multilayers that cover each other. It does not provide the right logical connections in the road-network; hence usually, an additional layer with the right connections is also necessary for the software. Moreover, the models of the investigated tools also have difficulties with special elevation attributes, e.g., banked curves, slopes, etc.

When the construction of ZalaZONE is finished, a laser scanning about the whole area is be carried out to provide a more accurate digital copy of the real environment. Based on this future data set, *OpenDRIVE* network is also planned to be generated, which then will be imported smoothly into any modelling software. (via the modelling the ZalaZONE Proving Ground of Hungary)

As a continuous result of the research of proving ground modelling, an open access repository has been started on GitHub where the models are freely available in different file formats for simulation software *IPG CarMaker, PreScan, VTD Vires*: `https://github.com/BMEAutomatedDrive` [16], [17]. The more, this repository contains two additional models created in *SUMO* and *Unity 3D* which were not investigated in the paper as they are not specific automotive programs. *SUMO* (Simulation of Urban MObility) [18] is a free and open microscopic road traffic simulation suite. *Unity 3D* [19] is a graphical game engine for realistic 3D visualization. As another future work the modelling works of ZalaZONE is planned within Matlab/Simulink (Automated Driving Toolbox) being a generic tool for vehicle engineers.

Acknowledgments

The research reported in this paper was supported by the Higher Education Excellence Program of the Ministry of Human Capacities in the frame of Artificial Intelligence research area of Budapest University of Technology and Economics (BME FIKP-MI/FM). The project has also been supported by the European Union, co-financed by the European Social Fund. EFOP-3.6.2-16-2017-00002.

References

[1] T. Tettamanti, I. Varga, Zs. Szalay, Impacts of Autonomous Cars from a Traffic Engineering Perspective, *Periodica Polytechnica Transportation Engineering* 44 (4) (2016), pp. 244-250.
doi: https://doi.org/10.3311/PPtr.9464

[2] P. Koopman and M. Wagner, Challenges in Autonomous Vehicle Testing and Validation. *SAE Int*. J. Trans. Safety 4 (1) (2016), pp. 15-24.

[3] Y. Kang, H. Yin, C. Berger, Test Your Self-Driving Algorithm: An Overview of Publicly Available Driving Datasets and Virtual Testing Environments, *IEEE Transactions on Intelligent Vehicles,* 4 (2) (2019), pp. 171-185.
doi: https://doi.org/10.1109/TIV.2018.2886678

[4] S. Riedmaier, J. Nesensohn, et al., Validation of X-in-the-Loop Approaches for Virtual Homologation of Automated Driving Functions. In: Proceedings of the 11th Graz Symposium Virtual Vehicle (2018).

[5] Y. Laschinsky, K. von Neumann-Cosel, et al., Evaluation of an Active Safety Light using Virtual Test Drive within Vehicle in the Loop, *IEEE International Conference on Industrial Technology*, Vina del Mar, (2010), pp. 1119-1112.
doi: https://doi.org/10.1109/ICIT.2010.5472583

[6] R. Kallweit, P. Prescher, M. Butenuth, Vehicle-in-the-loop: augmenting real-world driving tests with virtual scenarios in order to enhance validation of active safety systems. *25th International Technical Conference on the Enhanced Safety of Vehicles (ESV),* Detroit (2017).

[7] D. Gruyer, S. Demmel, et al., Simulation architecture for the design of Cooperative Collision Warning systems. 2012 *15th International IEEE Conference on Intelligent Transportation Systems*, (2012), pp. 697-703.
doi: https://doi.org/10.1109/ITSC.2012.6338733

[8] A. Belbachir, J-C. Smal, J-M. Blosseville, D. Gruyer, Simulation-Driven Validation of Advanced Driving-Assistance Systems, *Procedia - Social and Behavioral Sciences*, 48, (2012), 1205-1214, ISSN 1877-0428.
doi: https://doi.org/10.1016/j.sbspro.2012.06.1096

[9] M. Worm, R. van der Made, Verifying ADAS and autonomous driving performance, *Siemens PLM Software White paper* (2019).

[10] H. Németh, A. Háry, et al., Proving Ground Test Scenarios in Mixed Virtual and Real Environment for Highly Automated Driving. In: Proff H. (eds) Mobilität in Zeiten der Veränderung. *Springer Gabler*, Wiesbaden (2019), pp. 199-210.
doi: https://doi.org/10.1007/978-3-658-26107-8_15

[11] Zs. Szalay, M. Szalai, et al., Proof of concept for Scenario-in-the-Loop (SciL) testing for autonomous vehicle technology, Graz: *IEEE International Conference on Connected Vehicles and Expo (ICCVE)*, (2019), pp. 1-5.
doi: https://doi.org/10.1109/ICCVE45908.2019.8965086

[12] Zs. Szalay, Z. Hamar, P. Simon, A Multi-layer Autonomous Vehicle and Simulation Validation Ecosystem Axis: ZalaZONE. In: Strand M., Dillmann R., Menegatti E., Ghidoni S. (eds) Intelligent Autonomous Systems 15. IAS 2018. Advances in Intelligent Systems and Computing, Vol 867, *Springer*, Cham, 2019.
doi: https://doi.org/10.1007/978-3-030-01370-7_74

[13] Zs. Szalay, T. Tettamanti, et al., Development of a Test Track for Driverless Cars: Vehicle Design, Track Configuration, and Liability Considerations, *Periodica Polytechnica Transportation Engineering*, 46 (1) (2018), pp. 29-35.
doi: https://doi.org/10.3311/PPtr.10753

[14] Zs. Szalay, Á. Nyerges, Z. Hamar, M. Hesz, Technical Specification Methodology for an Automotive Proving Ground Dedicated to Connected and Automated Vehicles, *Periodica Polytechnica Transportation Engineering*, 45 (3) (2017), pp. 168-174.
doi: https://doi.org/10.3311/PPtr.10708

[15] M. Dupuis and H. Grezlikowski, OpenDRIVE® - an open standard for the description of roads in driving simulations. *In Driving Simulation Conference* (2006), pp. 25–36, Paris, France.

[16] BME AutomatedDrive github repository
URL https://github.com/BMEAutomatedDrive

[17] BMEAutomatedDrive homepage
URL https://www.automateddrive.bme.hu/

[18] P. A. Lopez et al., Microscopic Traffic Simulation using SUMO, 2018 *21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI, (2018), pp. 2575-2582.
doi: https://doi.org/10.1109/ITSC.2018.8569938

[19] M. Szalai, B. Varga, T. Tettamanti and V. Tihanyi, Mixed reality test environment for autonomous cars using Unity 3D and SUMO, 2020 *IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, Herlany, Slovakia, (2020), pp. 73-78.
doi: https://doi.org/10.1109/SAMI48414.2020.9108745

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 508-520, 2021
DOI: 10.14513/actatechjaur.00618

Influence of frequency resolution in case of frequency response function measurement in structural dynamics

Z. Gazdagh[1,*], B. Vehovszky[2]

[1]Audi Hungaria ZRt, Whole Vehicle Engineering,
Győr, Audi Hungária út 1. H-9027, Hungary

[2]Széchenyi István University, Department of Whole Vehicle Engineering,
Győr, Egyetem tér 1. H-9026, Hungary

*e-mail: zoltan.gazdagh@audi.hu

Submitted: 10/05/2021; Accepted: 18/08/2021; Published online: 16/09/2021

Abstract: Frequency resolution is an essential parameter in acoustical testing, even if we are using numerical or experimental method, for example when determining frequency response function (FRF) of a dynamic mechanical system, or executing modal analysis based on the FRFs. Finer resolution leads to more accurate results, at the expense of longer calculation/measurement process and larger data size. This parameter is generally set based on rules of thumb, prior practice or with big margin for safety. This results in waste time and data storage if the required frequency resolution is overestimated, or even significant errors in the results, if it is underestimated. Present paper offers a direct, method for the conscious determination of optimal frequency resolution. It is based fully on theoretical considerations, and investigates amplitude and phase distortion at resonances as target parameters. Beside defining the steps of the process, it is tested on a real structure, and the results are presented as well, proving the applicability and the appropriateness of the method. With this method, development engineers get a practical tool for adjusting the parameters of dynamic measurements and simulations.

Keywords: *transfer function; frequency resolution; dynamic analysis; engineering acoustics*

1. Introduction

During prototype testing, dynamic measurements are vital, some of them are mandatory to validate virtual results in concept phase. The low frequency measurements are often referred to vehicle comfort, the mid frequency measurements are used for validating the global vehicle dynamics and used in acoustic assessments [1]. The transfer function (or frequency response function, FRF) measurement is one of the most frequently used method in dynamic assessment of vehicle parts or structures. The typical input parameters, that help the engineers to understand local and global dynamics, are transfer functions, auto- and cross-power functions as well as the coherence curves. These parameters are dependent on each other as well. They are the input for trimmed body experimental modal analysis, but they can be input parameters for the evaluation of equivalent torsional stiffness of the chassis as well. In general, transient or stationary excitation types are used for frequency response function measurements in automotive industry. Measurement parameter settings mostly depend on prior results and experience, however inappropriate frequency resolution may lead to imprecise damping calculation and magnitude distortion, when these values are determined from the FRFs with the widely used half-power bandwidth method [2]. The current study focuses on the understanding and quantification of these effects.

Several parameters, such as structural damping, frequency resolution, or the frequency range of the modal base, play an important role in the accuracy of excited vibration analysis. In case of simulation, the frequency response functions (FRFs) of the investigated structure are often synthetized. This simulation based FRF synthesis is highly dependent on the used modal range or usage of lower and upper compensation of modes - when residual compensation is not used, the upper frequency limit of the modal range should be at least twice of the upper limit of the frequency range used for the FRF synthesis. Another dominant parameter is the frequency resolution. It must be chosen sufficiently small to minimize the amplitude distortion of the response functions, but this parameter determines the length of the FFT blocks during data processing as well. The better the frequency resolution, the longer the acquisition time (1 Hz resolution requires 1 second of data, but 0.1 Hz needs 10 seconds during a single acquisition without averaging), moreover, the larger the recorded data.

Equation (1) can be found in the user guide of MSc Nastran finite element solver [2]. The equation is applied for a single resonance peak and splits the half power bandwidth to n discrete points. This equation provides guidance for a possible determination of the correct frequency resolution in case of $\xi < 0.1$.

$$\Delta\omega = \frac{2 \cdot \xi \cdot \omega_r}{n-1} \tag{1}$$

In the context of (1), ξ denotes the damping ratio, ω_r denotes the resonant frequency, n is the number of discrete points in half power bandwidth and $\Delta\omega$ is the frequency resolution. However, when using this equation, the number of points (n) is up to the user, and the method does not take into account the level of amplitude distortion and frequency shift. In the following part, a newly developed, supplementary method is presented that takes into consideration the above effects as well, which are influenced by acquisition parameter settings.

2. Considerations on amplitude distortion

Let's focus on a single mass-spring-damper system, where the frequency transfer function can be written in the following form [3]:

$$H_d(j\omega) = \frac{\frac{1}{m\omega^2}}{\left[\left(\frac{\omega_r}{\omega}\right)^2 - 1\right] + 2j\xi \cdot \left(\frac{\omega_r}{\omega}\right)} \tag{2}$$

This relationship describes the receptance (H_d, i.e. the displacement response divided by the single exciting input) when system is excited by a unit force at a frequency of ω, its mass is m and damping is ξ.

In acoustic measurements, however, the velocity response function (the so-called mobility equation) of the structure is generally required, which can be obtained by multiplying the displacement response function (2) by $j\omega = \frac{\omega}{-j}$ (the relation $j = -\frac{1}{j}$ can easily be understood as $\frac{j}{j} = 1$ so $\frac{1}{j} = \frac{1}{j} \cdot \frac{j}{j} = \frac{j}{-1} = -j$). Aiming further simplification, by multiplying the nominator and the denominator equally by $\left(\frac{\omega}{\omega_r}\right)$, the resulted transfer function (velocity response or mobility equation) can be written as:

$$H_v(j\omega) = \frac{\frac{1}{m\omega_r}}{j\left(\frac{\omega}{\omega_r}\right) - j\left(\frac{\omega_r}{\omega}\right) + 2\xi} \tag{3}$$

In order to quantify the amount of distortion caused by the chosen frequency resolution, lets perturb the transfer function equation (3) with $\pm\frac{\Delta\omega}{2}$ frequency shift (referring to the $\Delta\omega$ frequency resolution), which represents the distance between

the theoretical and the actually acquired frequency value. As a result, the new transfer function can be written as:

$$H_v\left(j\left(\omega \pm \frac{\Delta\omega}{2}\right)\right) = \frac{\frac{1}{m\omega_r}}{j\left(\frac{\omega \pm \frac{\Delta\omega}{2}}{\omega_r}\right) - j\left(\frac{\omega_r}{\omega \pm \frac{\Delta\omega}{2}}\right) + 2\xi} \tag{4}$$

We are looking for the worst case, so when the actual frequency (ω) is located in the middle of two acquired, neighbouring frequency values ($\omega - \frac{\Delta\omega}{2}$ and $\omega + \frac{\Delta\omega}{2}$, see Fig. 1 for the case when $\omega = \omega_r$). The amplitude distortion $\varepsilon_H(j\omega)$ for such case can be given by dividing the acquired amplitude (4) by the theoretical one (3), resulting:

$$\varepsilon_H(j\omega) = \left|\frac{H_v\left(j\left(\omega \pm \frac{\Delta\omega}{2}\right)\right)}{H_v(j\omega)}\right| = \left|\frac{j\left(\frac{\omega}{\omega_r}\right) - j\left(\frac{\omega_r}{\omega}\right) + 2\xi}{j\left(\frac{\omega \pm \frac{\Delta\omega}{2}}{\omega_r}\right) - j\left(\frac{\omega_r}{\omega \pm \frac{\Delta\omega}{2}}\right) + 2\xi}\right| \tag{5}$$

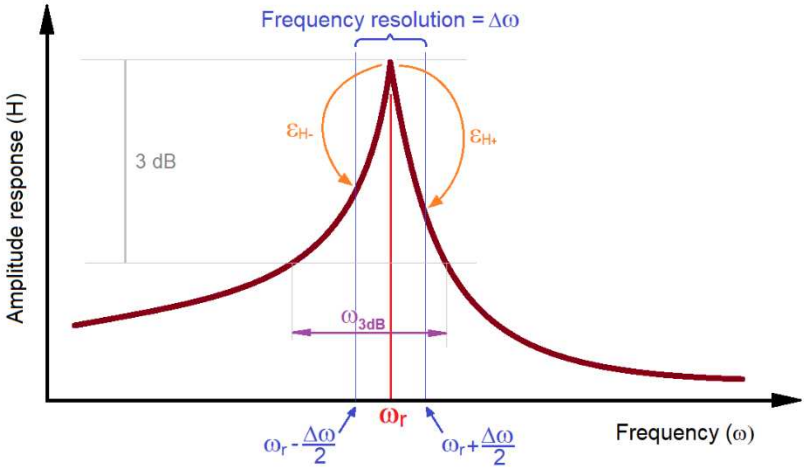


Figure 1. Resonance peak and frequency resolution – worst case

Consider the amplitude distortion at resonance frequency ω_r in the following. Substituting $\omega = \omega_r$ in equation (5) and highlighting $\frac{\pm\Delta\omega}{2\omega_r}$, after performing the simplifications, the obtained amplitude distortion is:

$$\varepsilon_H(j\omega|\omega = \omega_r) = \frac{1}{\left| j\left(\frac{1}{2\xi}\cdot\frac{\pm\Delta\omega}{2\omega_r}\cdot\frac{2+\frac{\pm\Delta\omega}{2\omega_r}}{1+\frac{\pm\Delta\omega}{2\omega_r}}\right)+1\right|} \tag{6}$$

Being a complex function, as a result of the absolute value calculation, the following equation is obtained:

$$\varepsilon_H(j\omega|\omega = \omega_r) = \frac{1}{\sqrt{\left(\frac{1}{2\xi}\cdot\frac{\pm\Delta\omega}{2\omega_r}\cdot\frac{2+\frac{\pm\Delta\omega}{2\omega_r}}{1+\frac{\pm\Delta\omega}{2\omega_r}}\right)^2+1}} \tag{7}$$

Using equation (7) the level of maximal amplitude distortion ε_H can be determined at resonance, which is the function of ξ (damping ratio), ω_r (resonance frequency) and $\pm\frac{\Delta\omega}{2}$ frequency resolution. Note: this value is the theoretical maximum distortion, when the resonance frequency (ω_r) falls exactly in the middle of two neighbouring, acquired frequency values (see Fig. 1.) The best case is when the acquired frequency hits exactly the resonance frequency, in this case there is no distortion, $\varepsilon_H = 1$.

In order to represent the practical applicability of equation (7), the ε_H amplitude distortion is plotted as a function of $\frac{\Delta\omega}{2\omega_r}$ in Fig. 2. Curves are plotted for different damping ratios to visualize the influence of damping. Lower damping shows higher risk of distortion.

It is worth mentioning that the amplitude distortion in plus and minus direction is not equal as resonance curves are not symmetric (Fig. 1.), however, the difference between ε_{H+} and ε_{H-} is negligibly small, thus Fig. 2. illustrates the average of the two values.

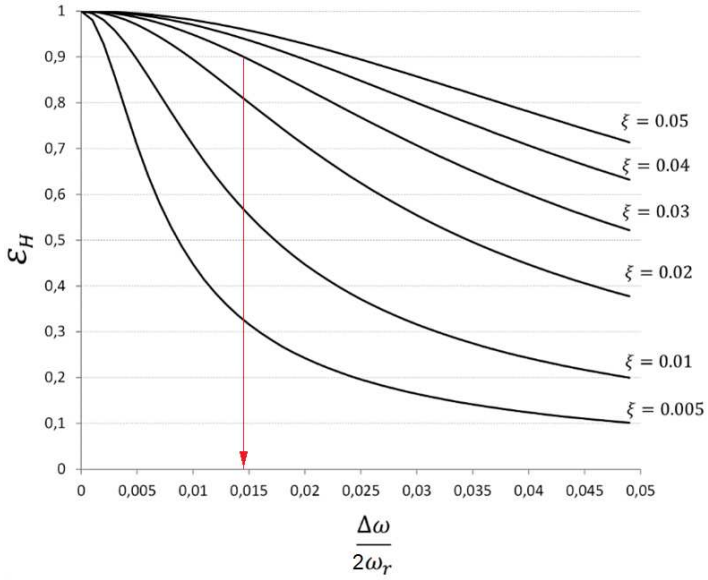


Figure 2. ε_H amplitude distortion as a function of $\frac{\Delta\omega}{2\omega_r}$

3. Optimal parameters for FRF measurements

Let us consider the practical application of the above findings: Using the curves of Fig. 2 and setting the permissible distortion to be 90% (meaning $\varepsilon_H \geq 0.9$) and estimating the damping ratio to be 3% ($\xi = 0.03$), the maximum allowable $\frac{\Delta\omega}{2\omega_r}$ value is 0.014, meaning $\Delta\omega = 2.8\ Hz$ maximum allowable frequency resolution beside $\omega_r = 100\ Hz$ resonance frequency. By decreasing ω_r to 50 Hz, maximum allowable frequency resolution decreases to 1.4 Hz. In short: the method gives us direct values for frequency resolution, assuming a maximum damping ratio and giving the lowest frequency limit of interest.

In the following, let's examine the amplitude distortion ε_H in relation with the half-power bandwidth range. For this, first take a look at the correlation of damping ratio and half-power bandwidth. As it is known, the damping ratio at resonance can be calculated by the following equation [4]:

$$\xi = \frac{\omega_2 - \omega_1}{\omega_r \left(tan\frac{\theta_1}{2} + tan\frac{\theta_2}{2} \right)} \tag{8}$$

where θ_1 and θ_2 are the angles between ω_r and ω_1, and ω_r and ω_2, respectively; ω_1 and ω_2 being the lower and upper frequency limits or the half power (3 dB) frequency band. Assuming $\xi \leq 0.1$, the relative phase angles at both ends of the 3 dB bandwidth (θ_1 and θ_2) are close to 90°, so the equation can be simplified to the following form [5]:

$$\xi = \frac{\omega_{3dB}}{2\omega_r} \tag{9}$$

ω_{3dB} being the half-power bandwidth ($\omega_2 - \omega_1$).

As a general rule, in case of ξ above 0.1, the calculation of 3 dB bandwidth using equation (9) is inaccurate, and only equation (8) can be used for damping estimation, while below 0.1 equation (9) gives adequate accuracy, meaning that the half-power bandwidth is assumed to be proportional to the damping ratio.

For further consideration, substitute the rearranged format of formula (1) $\left(\frac{\Delta\omega}{2\cdot\xi\cdot\omega_r} = \frac{1}{n-1}\right)$ in equation (7). The result is as follows:

$$\varepsilon_H(j\omega|\omega = \omega_r) = \frac{1}{\sqrt{\left(\pm\frac{1}{2(n-1)}\cdot\frac{2 \pm \frac{\xi}{n-1}}{1 \pm \frac{\xi}{n-1}}\right)^2 + 1}} \tag{10}$$

Equation (10) determines the amplitude distortion ε_H as a function of n (the number of discrete frequency values in the half-power bandwidth). Fig. 3. illustrates it graphically, by giving the amplitude distortion for $\omega_r - \frac{\Delta\omega}{2}$ and $\omega_r + \frac{\Delta\omega}{2}$ for $\xi = 0.1$ case (grey) and the average of the two (black). The difference of the two curves tends firmly to zero when damping ratio is decreased. The inset diagram shows the average curves only for different damping values. One can conclude that the amplitude distortion does not depend directly on the resonance frequency (see equation (10)), moreover, the dependence on damping is negligible too when $\xi \leq 0.1$.

Note: Fig. 2 and 3 both demonstrates the amplitude distortion, but the former focuses on the frequency resolution (which is important from practical point of view), while the latter examines the number of discrete frequency points within the half-power bandwidth, which is more useful for analytical considerations. The correlation of the two is given by equation (1).

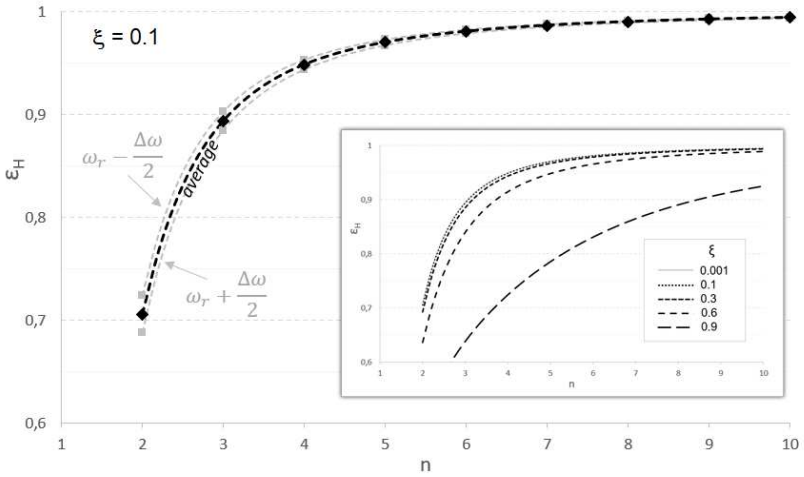


Figure 3. ε_H amplitude distortion as a function of number of discrete points in the half power bandwidth of a resonance at a constant damping ratio (ξ = 0.1 for the base diagram and for different ξ values for the inset)

By using the detailed correlations, the maximum error of amplitude distortion due to frequency discretization can be simply controlled, which gives us useful suggestions when setting acquisition parameters during dynamic measurements. This method can be used in simulations as well to reduce measurement-based error in hybrid simulation models. Such models are mentioned in [6] and [9]. Furthermore – as damping is generally calculated using the half-power bandwidth principle (equation (8) and (9)) into which the amplitude distortion also introduces error – the accuracy of damping calculation can also be kept under control.

The next examples show some further consequences of the appropriateness of frequency resolution value. Fig. 4. shows measured FRF curves in Nyquist diagram, from which the importance of frequency resolution is obvious: Even though measuring with coarser resolution is faster, results may be practically unusable.

One good method to check the adequacy of frequency resolution is plotting the FRF curves in Nyquist diagram [7]. In order to validate the number of discrete points in half power bandwidth, an experimental test was carried out and the FRFs were plotted in Nyquist diagrams: Fig 4 shows the coarser, while Fig. 5 shows finer frequency resolution cases.

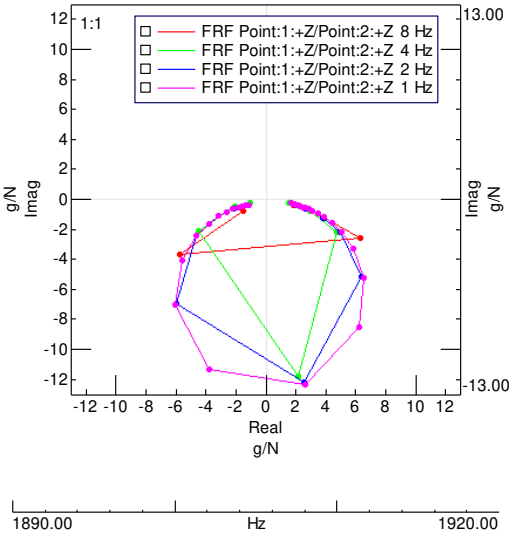


Figure 4. Measured FRFs with coarse frequency resolution

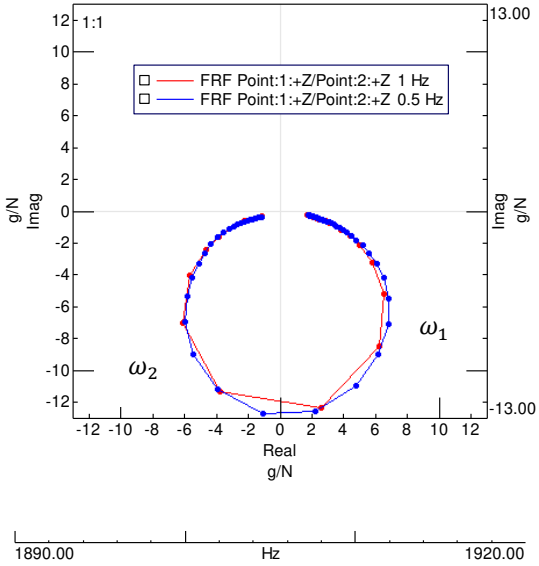


Figure 5. Nyquist diagram of FRFs with higher resolution around resonance

The applied measurement method was the so-called impact testing [8, 12], with an average of 5 impact measurements. During this transient type testing, exponential window was set to be 100%, but on force window 10% exponent was applied to remove possible noise from force signal after impact. For FRF estimation, the H1 estimation method was set, which is the most commonly used estimator [11], assuming all the noise to be on the output. The impact hammer was a low weight, modally tuned hammer with a steel tip, to ensure the high frequency energy input. The impact duration was around 0.3 milliseconds.

In Fig. 5 two cases with finer frequency resolution are compared for the same experiment (red curve with 1 Hz and blue curve with 0.5 Hz). The two ends of the 3 dB bandwidth are marked with $\omega_1$and ω_2. In case of the red curve n = 5 and blue FRF gives n = 9. Amplitude distortions were calculated as $\varepsilon_{H,1Hz} = 0.92$ and $\varepsilon_{H,0.5Hz} = 0.976$.

The damping ratio and resonance amplitude were also calculated (Table 1.) One can see that coarser resolution result not only in larger amplitude distortion but inaccurate damping value as well.

Table 1. Damping ratio and FRF amplitude of resonance frequency

Frequency Resolution	***Damping ratio***	***FRF amplitude***
1 Hz	0.09 %	12.63 g/N
0.5 Hz	0.1 %	12.76 g/N

The same FRFs are plotted in Bode diagram, see Fig 6. Small change in frequency resolution between the two FRF measurement lead already to a frequency shift of the resonance peak and produces different damping ratio when using the 3 dB bandwidth equation (8).

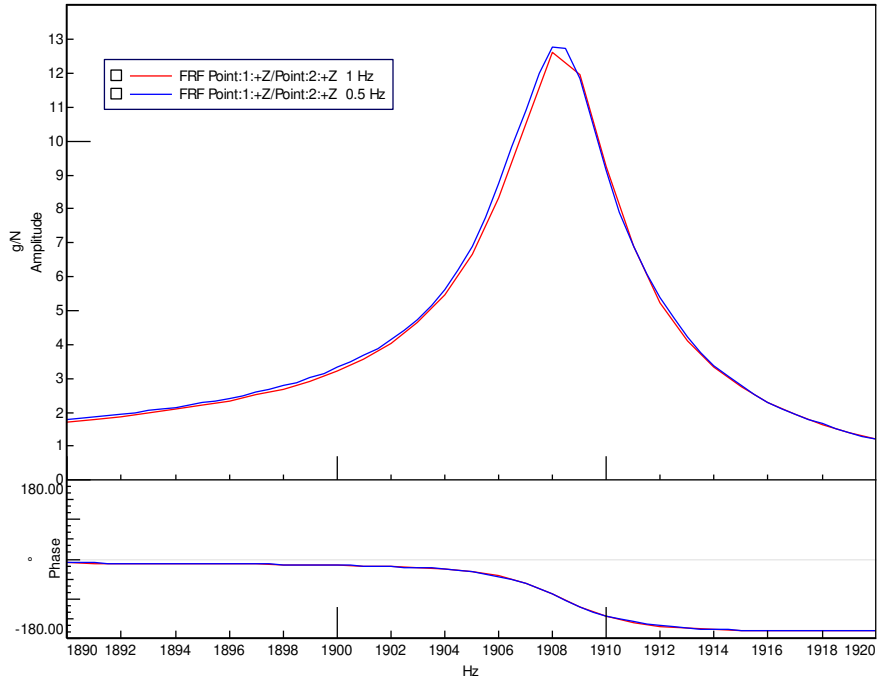


Figure 6. Bode Plot of FRFs near resonance

4. Conclusion

The method presented in this article offers a solution for the correct choice of the frequency resolution as a measurement setup parameter, taking into account the potential amplitude distortion of the FRFs. The level of distortion is determined as a function of frequency resolution, damping ratio and resonance frequency. Using this equation, one can determine the minimum frequency resolution which is required to keep the maximal amplitude distortion in a given range.

The number of discrete points within the half-power bandwidth is also evaluated, which makes easy to work with, as resonance frequency in this case is not a variable, and neither damping has significant influence (if lower than 0.1). By defining the appropriate criteria for measurement accuracy, the described functions and graphs provide good guidance for the selection of suitable parameters, so that amplitude of the frequency response function can be measured with pre-defined accuracy, and the error of damping value can be controlled as well. The application area covers the field of modal analysis [10] and other FRF based measurements.

A potential continuation of the work is the analysis of the effect of exponential windowing used during the impact testing, which also causes pseudo-damping effect [13]. It would be useful to form similar direct suggestions for the optimal parameter selection.

References

[1] M. F. Treszkai, D. Feszty, Review of Hybrid Finite Element – Statistical Energy Analysis methods in vehicle NVH predictions, *Acta Technica Jaurinensis* 12 (4) (2019) pp. 347–370.
doi: https://doi.org/10.14513/actatechjaur.v12.n4.513

[2] MD Nastran Dynamic Analysis User's Guide, 2010.

[3] M. R. Hatch, Vibration simulation using MATLAB and ANSYS, *Chapman & Hall/CRC*, Washington D.C., 2000.

[4] W. Heylen, S. Lammens, P. Sas, Modal Analysis Theory and Testing, *KU Leuven Academic Press*, Belgium, 1999.

[5] M. Geradin, D. Rixen, Mechanical vibrations: theory and application to structural dynamics, *Wiley*, New York, 2015.

[6] P. Peeters, S. Manzato et al., Reducing the impact of measurement errors in FRF-based substructure decoupling using modal model, *KU Leuven Academic Press*, Belgium, 2018.

[7] M. F. Treszkai, D. Sipos, D. Feszty, Damping Determination by Half-Power Bandwidth Method for a Slightly Damped Rectangular Steel Plate in the Mid-Frequency Range, *Acta Technica Jaurinensis* 13 (3) (2020) pp. 177–196.
doi: https://doi.org/10.14513/actatechjaur.v13.n3.545

[8] The Fundamentals of Modal Testing – Application Note 243-3, *Agilent Technologies*, 2000.
URL
https://www.modalshop.com/techlibrary/Fundamentals%20of%20Modal%20Testing.pdf

[9] Z. Gazdagh, Z. Jamil, J. M. Allport, A model answer: hybrid vibration analysis, *Cummins Turbo Technologies HTi Magazine* 13 (2009) pp. 3-4.

URL *https://pure.hud.ac.uk/en/clippings/a-model-answer-hybrid-vibration-analysis*

[10] A. Schweighardt, B. Vehovszky et al., NVH and Modal Analysis in Vehicle Industry, *Perner's Contacts* Special Issue 2 (2019) pp. 258-266.

[11] Z. Mao, M. Todd, Statistical modeling of frequency response function estimation for uncertainty quantification, *Mechanical Systems and Signal Processing*, 38 (2) (2013) pp. 333-345.
doi: *https://doi.org/10.1016/j.ymssp.2013.01.021*

[12] D. L. Brown, R. J. Allemang, A. W. Phillips, Forty Years of Use and Abuse of Impact Testing: A Practical Guide to Making Good FRF Measurements, Experimental Techniques*, Rotating Machinery and Acoustics* 8 (2015) pp. 221-241.
doi: *https://doi.org/10.1007/978-3-319-15236-3_21*

[13] D. Tajiri, M. Matsubara, S. Kawamura, Identification of modal parameters in a lightly damped system based on impact vibration testing: Application of exponential window and removal of its effect, *IOP Conf. Series: Journal of Physics* 1264 (2019) 012011.
doi: https://doi.org/10.1088/1742-6596/1264/1/012011

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 521-552, 2021
DOI: 10.14513/actatechjaur.00611

Review of the effect of sand on the behavior of expansive clayey soils

A. N. Alnmr[1,*], R. P. Ray[1]

[1]Széchenyi István University, Department of Structural and Geotechnical Engineering, Egyetem tér 1, 9026 Győr, Hungary
[*]e-mail: ammar888999@hotmail.com

Submitted: 19/04/2021; Accepted: 02/05/2021; Published online: 13/05/2021

Abstract: Clayey soils often showed undesirable engineering behavior such as low bearing capacity, swelling and shrinkage characteristics. However, chemical improvement, thermal improvement and improvement by additives like lime, cement and sand offer an efficient technique to overcome the problems resulting from Expansive soils. This paper presents a review of the swelling behavior of sand-clay mixtures as well as the effect of sand on the physical and mechanical characteristics of expansive soils. Results highlight the importance of sand in improving the behavior of expansive soils. Finally, the most important general conclusions about the behavior of expansive soils and suggestions for future researches are highlighted.

Keywords: *expansive soils; additive material; sand-clay mixtures; mechanical characteristics; physical properties*

1. Introduction

Expansive clayey soils spread over large areas of land within the world. The total value losses resulting from the destruction of structures built on expansive soils are estimated to be several billions of dollars worldwide so expansive soils are considered a high-cost problem when establishing structures on them. For example, the annual estimated loss of roads, buildings, airports and other structures in China, Britain and America exceeds $1 billion, £150 million and $15 billion, respectively [1] [2] [3] [4]. According to Nelson and Miller [5] the financial loss resulting from the devastating effects of expansive soils would be greater than the loss caused by earthquakes or floods.

Besides, clayey soils cover 630 thousand hectares, which is approximately 6-9% of Hungary's land, and another 1.7 million hectares are covered with clayey loamy soils. Its significant expanse and unique properties led to early scientific recognition in Hungary as well, and thus expansive soils can be found within the Hungarian soil science literature from its inception, from József Szabó's scientific work [6].

The foundation soil layer plays an important role in creating safe and cost-effective constructions. In case the site is unsuitable or the engineering works in which backfill plays an essential role, such as dams, roads and railways, requires large quantities of soil of certain specifications and bringing them to the site from long distances may be uneconomical, which call for improving the behavior of site soil itself.

There are many ways to improve the expansive soil behavior and choose between them depends on economic, practical and environmental standards. Common methods are chemical improvement, thermal improvement and improvement by additives such as lime, cement and sand.

Many researchers have conducted many laboratory and field experiments that confirmed that expansive soils with certain additives of other materials such as sand, silt, lime, fly ash, etc. can improve their behavior and used efficiently in the above engineering works[7] [8] [9].

Sand is a natural granular material and due to its high bearing capacity, it can be used as a filler material to improve the behavior of expansive soils. Therefore, sand can be used in various percentages by mixing it with expansive soil and thus altering the properties of plasticity, compaction, strength and permeability (Hydraulic conductivity, k) of mixtures [8] [9] [10] [11]. It is also well known that expansive soils comprise of different proportions of coarse fractions, which vary as function of soil, and therefore, studying the effect of sand on the behavior of the expansive soil has a high degree of importance.

2. Literature review

The swelling properties of expansive soils are due to several factors related to the initial soil conditions such as initial dry density and initial moisture content, the granularity and the content of clay so many correlations have been suggested depending on them for prediction of swelling potential in addition to various index properties such as liquid limit, plasticity index, shrinkage index, activity etc. However, the first factor that conditions swelling of expansive soil and affects these physical, chemical and mechanical properties is the mineralogical composition and the amount and type of minerals present in expansive soil. It is known for instance, that the phenomenon of swelling depends not only on the content of clay but also on

the surface activity of the clay itself which will be more intense as the specific area of the particles will be high [12].

At the beginning, it is necessary to understand the process and behavior of swelling to enable us to interpret the results related to the effect of sand on the behavior of expansive soils. Then, the effect of sand on the most important parameters affecting the behavior of expansive soils will be reviewed.

2.1 Swelling process

The swelling-pressure behavior of expansive soil are often described in three steps: before, during, and after absorption of the water.

Fig. 1 shows the three steps. In the first step, expansive soil is formed of a mix of swelling clay minerals (such as montmorillonite), void, other non-swelling minerals and sand particles. Through this step, air and free water occupy the voids [13].

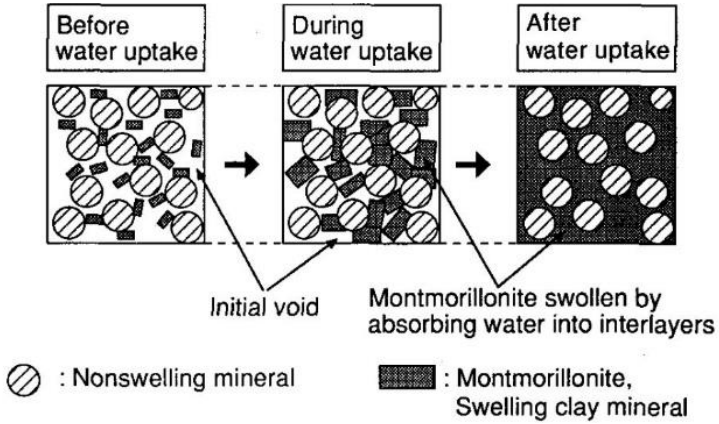


Figure 1. Mechanism on the swelling pressure of expansive soil from [13]

During water absorption, swelling clay mineral adsorbs water into interlayers and swells and occupies the void in the expansive soil. Therefore, expansive soil increases in size and swelling pressure occurs. Anyway, the interlayer and double layer forces are liable for the swelling behavior of the bentonite.

Finally, after the water is totally absorbed, and therefore the entire voids are filled with water, the volume of expansive soil cannot increase further. then, the swell pressure of the expansive soil can be measured

2.2. Swelling behavior

The mechanism of the swelling has been extensively discussed in the literature [14] [15] [16] [17]. The swelling behavior of sand-clay mixtures is caused by two mechanisms known as crystalline swelling that relates to the degree of hydration and osmotic swelling is related to the cation in the interlayers. Also, the relative humidity plays major role in the swelling behavior.

2.2.1. Crystalline swelling

Crystalline swelling is that the first mechanism that occurs when montmorillonite absorbs water, that is, it is an outcome of short-term hydration. This mechanism results in an increase in the space between the interlayers of swelling minerals such as montmorillonite. As a result, an increase in the volume of the expansive soil is gained and as a result, the swelling pressure appears. Crystalline swelling is mainly caused by the energy related to the initial hydration of interchangeable interlayer cations and hydrogen bonds or the effects of charged surface dipole attraction associated with the solid-liquid interactions that occur in the immediate vicinity of the clay particle surfaces [15] [16]. Fig. 2 shows the conceptual model of the sequential crystalline swelling process of montmorillonite.

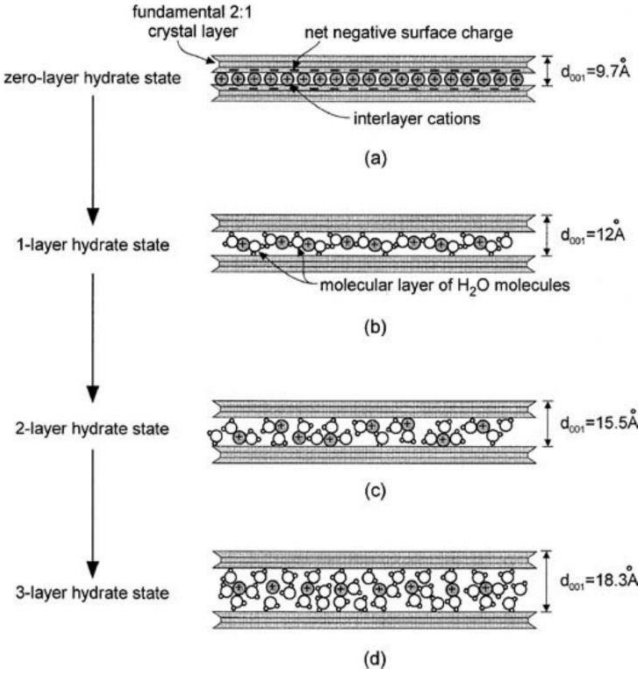


Figure 2. Mechanism of crystalline swelling of montmorillonite [16]

It has not been confirmed but it is first supposed that the surface of the clay is completely enveloped with a layer of water molecules. Then the second molecular layer is covered and so on. In sum, the water molecules are distributed layer by layer. For sodium montmorillonite, the water molecules are enveloped from zero to four layers [18]. And according to foster [12] the cation of clay play an important role in The dissociation where the important relationship between clay composition and swelling is very satisfactorily explained by the concept of cation dissociation. When calcium is the accompanying cations, the degree of dissociation is not large enough to cause a significant degree of swelling; Whereas when sodium is the predominant bound cation, the degree of dissociation is much greater and sodium montmorillonite has swelling properties.

As shown in Fig.3, during the absorption of water, the negative dipoles of the water molecules are oriented towards the cation. Consequently, the electrostatic interaction between the layers and the interlayer cations decreases [15].

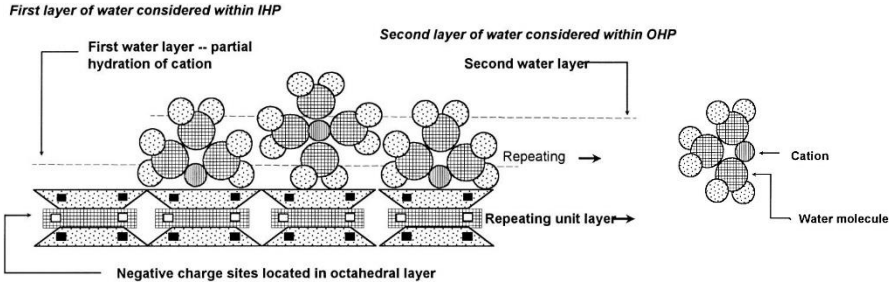


Figure 3. Arrangement of water on the clay layer [19]

The outcome is an increment in the space between the layers of swell clay minerals and an increase in the volume of it. The swelling process can absolutely arrive at a separation of the layers because of the forces that developed as a result of absorbing water. In addition, if there are organic compounds, the distance between layers of montmorillonite can be diminished since the organic compound makes the surface hydrophobic [15] [16].

2.2.2. Osmotic swelling

Osmotic swelling is the second mechanism resulting from the adsorption of water by swelling clay minerals. This mechanism creates a persistently increasing distance between layers of the swelling clay mineral. The difference in concentration between ions near the surface of swelling clay mineral layers and in the porous water is the main cause of osmotic swelling. This mechanism is resulted in because of diffuse multiple-layer interactions and van der Waals gravity.

Swelling clay mineral comprises of a lattice of Al_2O_3 and SiO_2 units; A net negative charge creates on the clay surface when divalent minerals, for example, magnesium are substitute in the lattice for aluminum or silicon [20] . Within the presence of an aqueous electrolyte solution, negatively charged mineral surfaces pull cations and polar water, composing a double-layer system (Fig. 4). Interference between two double layers makes a driving force that pushes the clay platelets away from each other; It additionally causes an increase in the cation concentration between the platelets and consequently, free water must be drawn into the system to reestablish balance. For example, Clay layers are negatively charged, and it prompts a repellent force between layers. However, cations near the surface of the layer counterbalance this negative charge. Replaceable cations are not too strong on the external surfaces. Consequently, they will in general diffuse from zones of high concentration towards zones of low concentration in the bulk solution. The outcome is a diffuse set of ions around a clay particle. This arrangement is called the diffuse electric double layer.

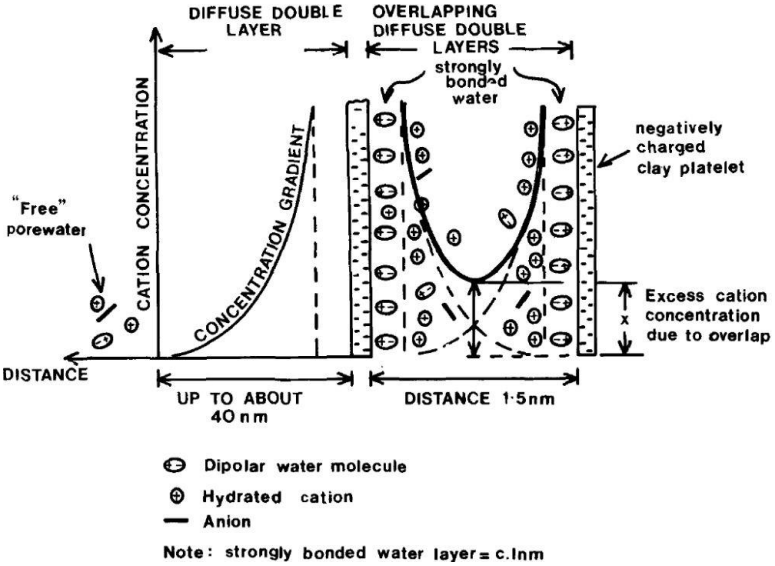


Figure 4. A double-layer osmotic swelling model of two clay mineral platelets [14]

Fig. 5 shows the diffuse double layer. It is defined as a group of negative charge set on an outer surface of clay and a cloud of positive ions placed on the outer surface.

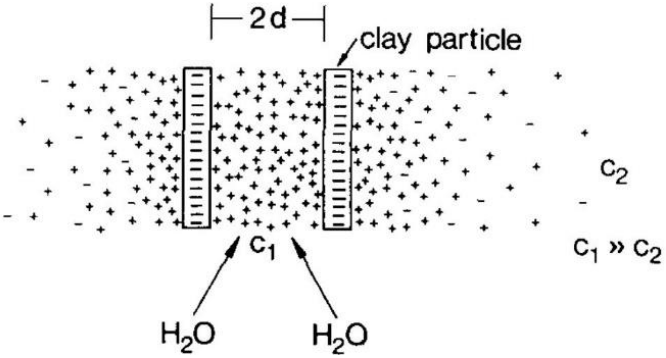


Figure 5. Diffuse double layer from [15]

To form the diffuse double layer, cations close to the surface balance the negative charge and then the exchangeable cations tend to diffuse. As a result, a repellent force appears between the two overlapping layers of the clay particles. This repellent force varies with the balance and radius of the counterions in the double layer and the concentration of electrolytes in the pore water [18].

From this literature review on the process and swelling behavior, we note that the existence of coarse materials (sand) in expansive soil has a significant effect on swelling. Whereas, Increasing the hydrophobic sand ratio reduces the hydrophilic clay ratio and thus reduces the efficacy of crystalline and osmotic swelling and improves the behavior of the expansive soil.

Knowing the swelling behavior is very important, as it enables to give a preliminary explanation about the effect of sand on the behavior of the expansive soil, but knowing the swelling behavior alone is not sufficient and it is necessary to know the effect of more than one parameter together, such as density and initial moisture with different sand ratios on the physical and mechanical properties of the expansive soil. This is will be highlighted in the following sections.

3. Effect of sand on physical characteristics

3.1. Consistency (Aterberg) Limits

There are many studies that show the positive effect of sand on the consistency of the expansive soil [7] [10] [21].

Louafi and Baher [10] conducted an experimental study to examine the effect of sand on soil consistency. Fig. 6 shows the granular gradient curves for sand and

clay (bentonite) used in the study. the physical characteristics of bentonite are showed in Table 1.

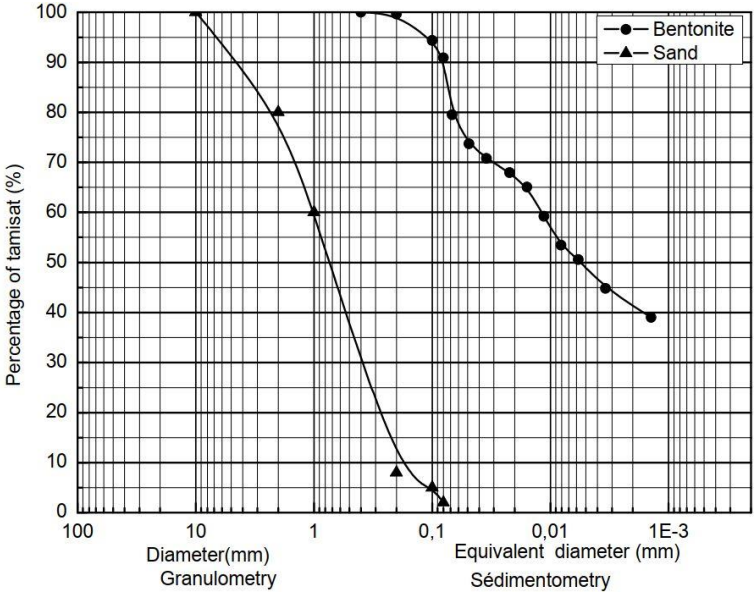


Figure 6. The granular gradient curves of sand and bentonite [10]

It is noticed from Fig. 6 that the sand used is medium poorly graded sand and the bentonite used contains in its composition 65% clay particles of less than 2 μm in diameter and 10% coarse (sand) with a diameter greater than 0.075 mm (75 μm).

Table 1. Physical characteristics of bentonite [10]

Property	***Value***
Specific gravity	*2.6*
Consistency properties	
Liquid limit (%)	*187*
Plastic limit (%)	*47*
Plastic limit (%)	*140*
Optimum moisture content (OMC) (%)	*25*
Maximum dry unit weight (γ_{dmax}) (kN/m^3)	*1.6*

Louafi and Baher [10] conducted a series of tests to study the evolution of limits of consistency using two different diameter sizes of sand ([0.1 mm < d < 0.2 mm]

and [0.2 mm < d < 0.4 mm]) to study its effect on the consistency of expansive clay soils. To do this, the soil is mixed with different percentages of sand (10-20-30-40-50-60-70%) and then limits of consistency is defined using distilled water.

Fig. 7 and 8 show the change of both the liquid limit and the plasticity index in terms of the percentage of fine and medium sand. The results show that adding sand to clay (bentonite) considerably reduces both the liquid limit and the plasticity index. The greater decrease is for sand of larger diameters, however the difference in result between the two fine and medium sand is small. This reduction in the liquid limit and the plasticity index is explained by the decrease in the content of fines that contribute to the plasticity with the increase in the percentage of sand. Thus, the number of bonds between the clay particles decreases with the increase in the bonds of the clay-sand particles or the bonds of the sand and sand particles.

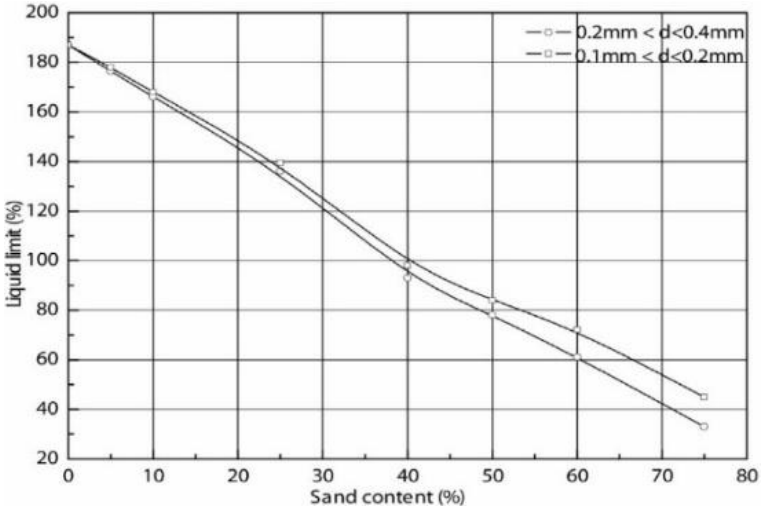
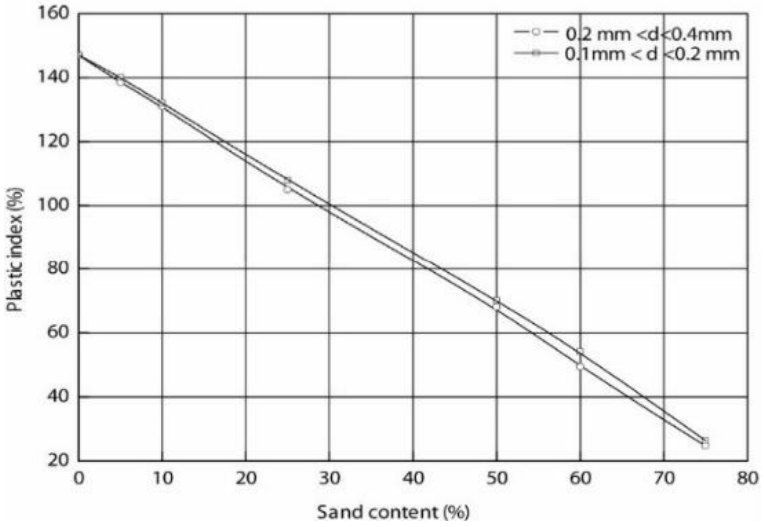


Figure 7. Change of the liquid limit in terms of the sand content [10]

As for the difference in the liquid limit and the plasticity index related to the effect of the size of the grains of added sand. This is probably due to the total specific surface area of water absorption of the sample which depends on the size of the sand grains. This specific surface is more important when the sand particles are finer, permitting the mixtures to absorb more water and consequently increasing of liquid limit.

Roy [7] also used three gradients of sand: fine, medium and coarse sand with percentages from 0 to 15 % in his experimental study. It was also found that increasing the percentage of sand reduces the values of the Liquid and plastic limits,

however, the effect of sand gradation was not clear because of the fact that the percentage of added sand is small, not more than 15% and the effect of the gradient on the consistency is small as noted in Table 2.



(b)

Figure 8. Change of the plastic index in terms of the sand content [10]

Table 2. LL, PL and PI of clay sand mixtures [7]

Type of Sand Mixed	***% of Sand***	***LL (%)***	***PL (%)***	***PI (%)***
Fine Sand	0	40.00	24.90	15.10
	5	35.85	24.50	11.35
	10	35.15	24.35	10.80
	15	34.80	24.10	10.70
Medium Sand	0	40.00	24.90	15.10
	5	36.15	24.35	11.80
	10	36.00	24.30	11.70
	15	35.60	24.05	11.55
Coarse Sand	0	40.00	24.90	15.10
	5	36.25	24.30	11.95
	10	36.05	25.14	10.91
	15	35.85	24.65	11.20

Whereas in the study of Srikanth and Mishra [22], they use fine (FS) and medium (MS) sand, but with percentages of sand of more than 50%, and with two different types of clay (B1 low plasticity and B2 High plasticity) and they have obtained the same previous results. However, the effect of the size of the sand particles becomes more evident with an increase in the plasticity of the clay used, as shown in Fig. 9.

Regarding the shrinkage limit, Srikanth & Mishra [22] [23] and Yang et al [24] found that the shrinkage limit increases with an increasing percentage of sand. As shown in Fig. 10, which shows the results of the shrinkage limit of fine clay-sand and medium clay-sand mixtures. For an identical clay content in the mixture, fine clay-sand mixtures showed higher shrinkage limits.

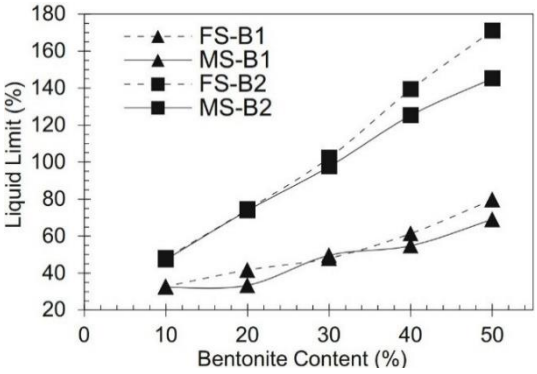


Figure 9. Variation of the liquid limit for clay-sand mixtures with fine sand and medium sand according to the content of bentonite [22]

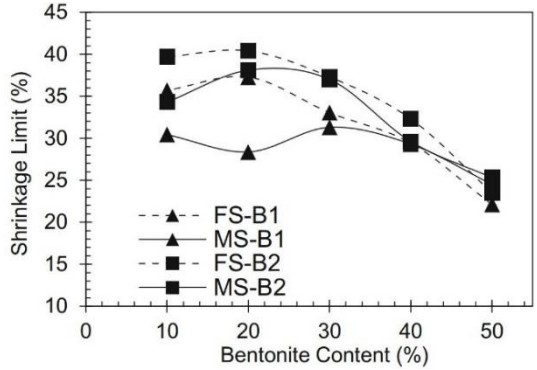


Figure 10. Variation in shrinkage limit of fine sand-clay(bentonite) and medium sand- clay mixtures with bentonite content [22]

3.2. Optimum moisture and maximum dry density

Roy [7] made a visible comparison of the effects of mixing different types of sand, for instance, fine, medium, coarse and with variable proportions in soil on maximum dry density (MDD) and optimum moisture content (OMC), Fig. 11 shows the experimental results in terms of maximum dry density (MDD) and Fig. 12 shows the experimental results in terms of optimum moisture content (OMC) for various percentages of sand.

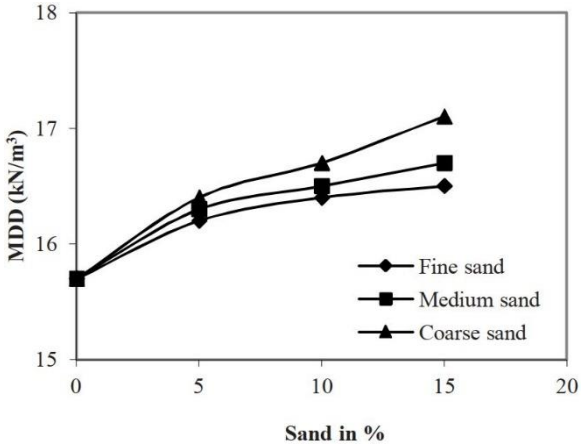


Figure 11. Variation of maximum dry density of clay-sand mixtures [7]

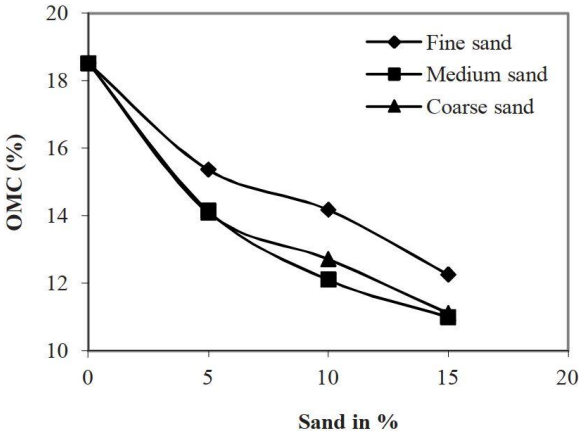


Figure 12. Variation of optimum moisture content of clay-sand mixture [7]

From Fig.11, it is observed that when mixing expansive soil with fine sand from 0 to 15%, MDD increases from 15.70 kN / m3 to 16.50 kN / m3. But for the same expansive soil when medium sand is mixed in increasing percentages, the MDD increases from 15.70 kN / m3 to 16.70 kN / m3. In the case of mixing coarse sand, the effect on MDD increases even more. Thus, increasing the size of sand grains results in an increase in the maximum dry density.

The result of mixing sand in numerous percentages on the OMC of expansive soils is shown in Fig. 12. It is noted that adding all kinds of sand in increasing percentages decreases progressively the OMC values compared to the tested soil and this value decreases progressively from 18.52% to 11.00% when the percentage of medium-grained sand increases from 0 to15%. this is often a sign of reduced water demand to achieve the required compaction in the field when the percentage of sand increases in expansive soil [7].

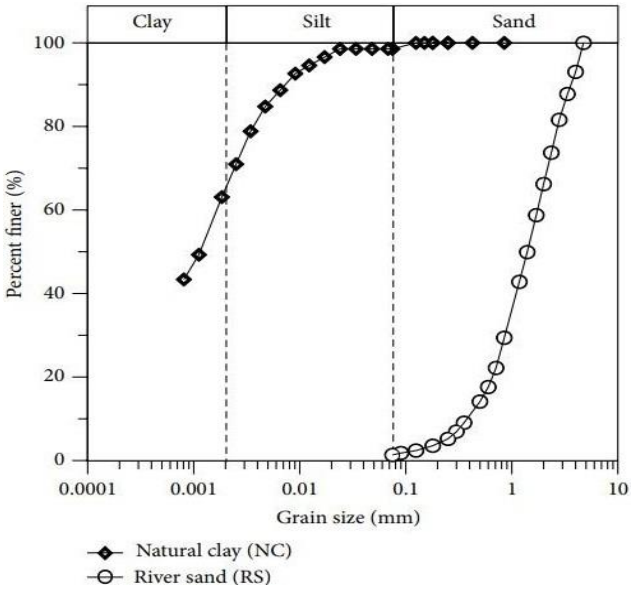
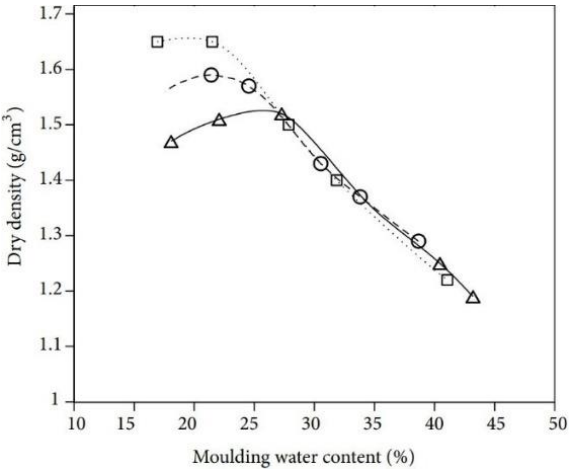


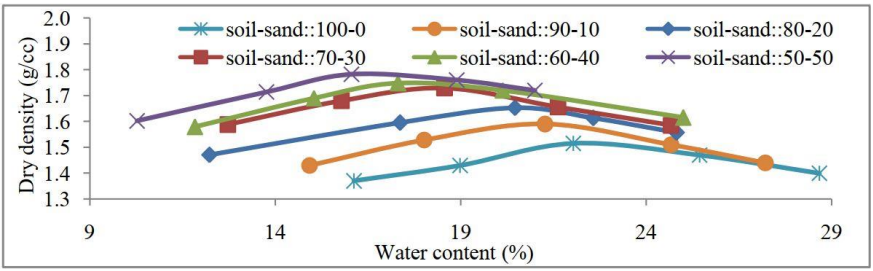
Figure 13. Grain size distribution of investigated sand and clay [25]

Khan et al [25] conducted standard proctor experiments for three clay-sand mixtures with percentages of sand of (0-20-40)% to work out the maximum dry density and optimum moisture content for each percentage of mixing. Clay-sand mixtures were prepared based on the dry weights of the materials. Fig. 13 presents the Grain size distribution of sand and clay materials. It is observed that there is a continuous increase in the maximum dry density Fig. 14-(a). The reason is Due to

the decrease in the volume of pores in the soil and the decrease of absorption as a result to the replacement of a soft part that can hold a large amount of water with a coarse part with little ability to retain water and this is confirmed by the optimum moisture values of the mixtures wherever it is noted the optimum moisture values decrease by increasing the percentage of added sand.



(a)



(b)

Figure 14.Compaction characteristics of clay-sand mixtures: (a) Khan et al [25] - (b) Gupta and Sharma [26]

Where: NC represents the Nature clay without any added sand (0% of added sand); CS-I represents the clay-sand mixture with added sand of 20%; CS-II represents the clay-sand mixture with added sand of 40%.

Gupta and Sharma [26] also conjointly performed standard Proctor experiments. Fig. 13-(b) shows the curves of the water content and the dry density of the expansive soil for different percentages of sand ranging from 10% to 50%. It is observed that

the maximum dry density (MDD) of the clay-sand mixtures increases with the increase in the percentage of sand. While the opposite happens with optimum moisture content (OMC), for the same reason mentioned previously.

Srikanth and Mishra [23] conducted standard Proctor experiments on expansive soil with different percentages of sand (50-90%) for two grain size of sand (fine sand FS and medium sand MS) and determined the optimum moisture content (OMC) and maximum dry density (MDD) for all added sand percentages as shown in Table 3.

Table 3. OMC and MDD of various sand-clay mixture [23]

Sand: bentonite proportions	***Bentonite-1 (B1)***		***Bentonite-2 (B2)***	
	OMC (%)	MDD (kg/m3)	OMC (%)	MDD (kg/m3)
Medium sand				
50:50	22.5	1565	12.9	1598
60:40	19.0	1628	11.6	1624
70:30	16.0	1662	17.5	1657
80:20	15.0	1710	18.5	1651
90:10	18.5	1670	17.9	1611
Fine sand				
50:50	22.5	1585	20.9	1555
60:40	20.5	1595	20.2	1582
70:30	18.0	1655	18.8	1607
80:20	18.4	1701	20.5	1612
90:10	17.2	1641	22.4	1579

The data from the standard Proctor compaction test in Table 3 shows that for the same percentage of sand, the FS-bentonite and MS-bentonite mixture showed a different value for OMC and MDD indicating the likelihood that sand particle size affected the compaction properties of the mixtures. Mixtures with MS resulted in relatively higher MDD and lower OMC values for both types of bentonite (Clay) that may be attributed to the effective filling of the clay particle within the spaces formed between the sand particles. The data in Table 3 shows that mixtures with percentage of sand of 80% exhibited the highest density among all mixtures. Once the voids formed between the individual sand particles are completely filled by clay particles, with another addition of sand more than 80%, there will be voids between the grains of sand that are not filled with clay particles, but rather with water and air, i.e. the clay particles will not occupy all the spaces between the sand grains, and this

is the reason for the decrease in MDD. Similar results have been reported by Yang et al [24], K. Raw &G. Raw [27] and Atemimi [28].

It is obvious that sand improves the maximum dry density of expansive soils, however increasing the dry density of expansive soils as a result of increasing the percentage of sand will cause more swell pressure and swell amplitude. This calls for studying the effect of sand content in expansive soils on its mechanical characteristics.

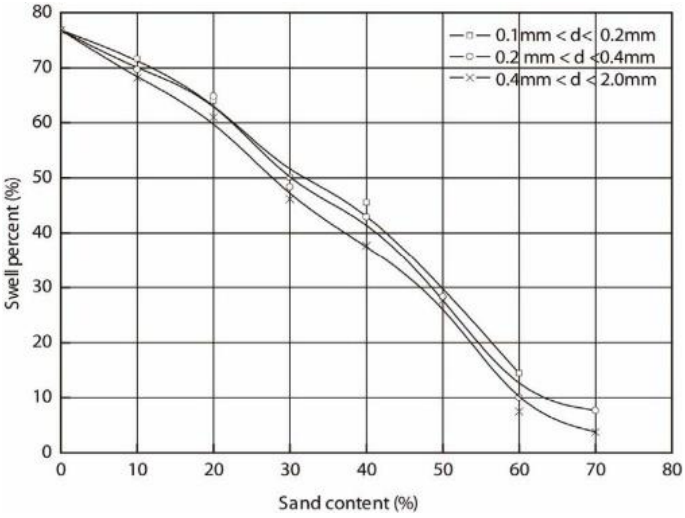
4. Effect of sand on mechanical characteristics

4.1. Swelling Potential and Swelling Pressure

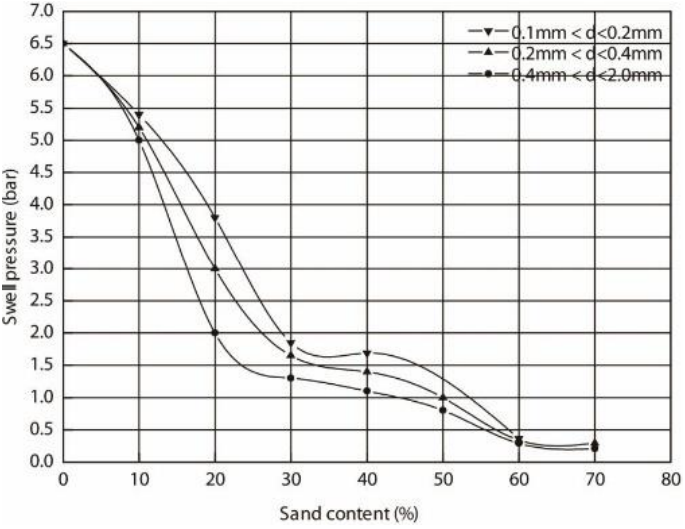
Louafi and Baher [10] studied the effect of added sand and its grain size on the swelling potential and swelling pressure by conducting free-swell experiments of different grain sizes of sand and different added sand percentages from 0 to 70% for each type of grain size. In all experiments, both the dry density of 15kPa and the initial moisture of 26% were fixed.

Samples were prepared at laboratory temperature by hand mixing and pressed statically into oedometric cell in three layers, each layer with a thickness of 5 mm, to ensure homogeneous samples were obtained, using a hand press. the swell was allowable under a seating surcharge of 3 kPa by immersion. After obtaining the final heave (ΔH), the sample was consolidated under increasing vertical loads until the initial thickness was reached.

The curves of Fig. 14 show that the swelling pressure and swelling potential of expansive soils decrease with an increase in the percentage of added sand. Coarse sand with a larger grain size seems to produce a greater reduction in both the swelling potential and the swelling pressure. This difference expresses the impact of sand grains on swelling. These above results can be explained by the fact that the clay content decreases by adding sand, which is usually an inert material. Indeed, increasing the percentage of sand reduces the swelling potential and swelling pressure; This is on the one hand. On the opposite hand, much of this swelling is absorbed by the voids between the sand grains. Once the sand fraction is coarser, these voids become larger. consequently, there is a decrease in the swelling potential and swelling pressure caused by soil swelling.



a)



b)

Figure 15.Variation of swelling potential and swelling pressure with percentage of sand for clay-sand mixture prepared at same dry density and initial moisture [10]

Srikanth and Mishra [23] conducted free-swell experiments by standard consolidometer according to ASTM D 2435. To highlight the importance of water content on the behavior of the mixtures, each mixture was pressed at 5% dry OMC and 5% wet OMC.

Samples were formed with dimensions of 60 mm in diameter and 15 mm in height in a way that achieves uniformity with density and moisture. All samples were formed according to their maximum dry density, with three different initial moistures. The sample was formed in a consolidation cell with a seating pressure of 4.9 kPa, then submerged in water and left until complete swelling. Then the samples were consolidated, gradually increasing the stress from 4.9 up to 784.5 kPa.

Tables 4 and 5 show the swelling pressure and swelling potential data exhibited by various mixtures of sand and bentonite (clay). From the data shown within the tables 4 and 5, we note that the swelling pressure and swelling potential decrease with an increase in the percentage of sand no matter what the initial compression condition. The data showed that, for any given percentage of the sand in the mixture, the samples compacted on the dry side of OMC exhibited higher swelling potential followed by those compacted in OMC.

Table 4. Swelling Potential results of sand clay mixture [23]

Sand–bentonite proportions	***Swelling potential (%)***		
	MDD-OMC	***MDD-5 % dry of OMC***	***MDD-5 % wet of OMC***
Medium sand (MS)			
50:50	18.7	35.0	17.0
60:40	16.2	8.0	7.8
70:30	4.6	18.5	1.6
80:20	1.0	6.3	0
90:10	0	0	0
Fine sand (FS)			
50:50	18.4	24.1	10.2
60:40	13.3	11.5	22.7
70:30	2.0	5.2	4.1
80:20	1.4	7.4	1.1
90:10	0.2	0.2	0.2

It was also found from tables 4 and 5 that sand with FS exhibits relatively higher swelling pressure and swelling potential for any given percentage of the sand in the mixture. Regardless of the size of the sand grains, mixtures containing bentonite less

than 20% showed a general decrease in the observed swelling. Whereas, the mixture with high bentonite contents (i.e. 50 and 40%) showed very high values of swelling pressure and swelling potential. Comparison of the mixtures shows that for the same percentage of sand and bentonite, the bentonite containing FS showed a higher value for swelling potential and swelling pressure compared to the bentonite with MS. Bentonite effectively fills the smaller spaces between individual FS particles compared to the relatively larger void spaces between MS particles. Once the spaces between the FS particles are filled in with the clay, along with a further increase in the volume of the clay, it will begin pushing out the sand particles leading to a higher value for swelling potential and swelling pressure [23]

Table 5. Swelling Pressure results of sand clay mixture [23]

Sand–bentonite proportions	***Swelling pressure (kPa)***		
	MDD-OMC	***MDD-5 % dry of OMC***	***MDD-5 % wet of OMC***
Medium sand (MS)			
50:50	340.5	388.4	335.3
60:40	300.1	219.1	242.2
70:30	134.4	215.8	50.4
80:20	23.5	79.7	21.1
90:10	10.2	0	0
Fine sand (FS)			
50:50	372.2	431.7	343.3
60:40	344.4	258.9	332.3
70:30	79.5	101.0	71.5
80:20	52.6	94.7	21.6
90:10	7.8	10.1	5.0

Phanikumar et al [9] conducted one-dimensional swell-consolidation tests, in order to study the swelling pressure and potential swelling of clay-sand mixtures. The dry density (γ_d) and initial water content (w_i) of the samples were fixed at 12kN/m3 and 0%, respectively.

Clay-sand mixtures were thoroughly mixed with a spatula and statically compacted into the consolidometer ring (diameter 60 mm, thickness 20 mm) in four layers of 5 mm thick each. An initial token surcharge of 5 kPa has been applied. Samples were allowed to undergo free swelling by immersion until the swelling had stabilized. To assess the equilibrium heave (ΔH), ΔH (Y axis), and log time in

minutes (X-axis) plots were made and when the curves became a straight line, equilibrium heave was understood to have been reached Fig. 15.

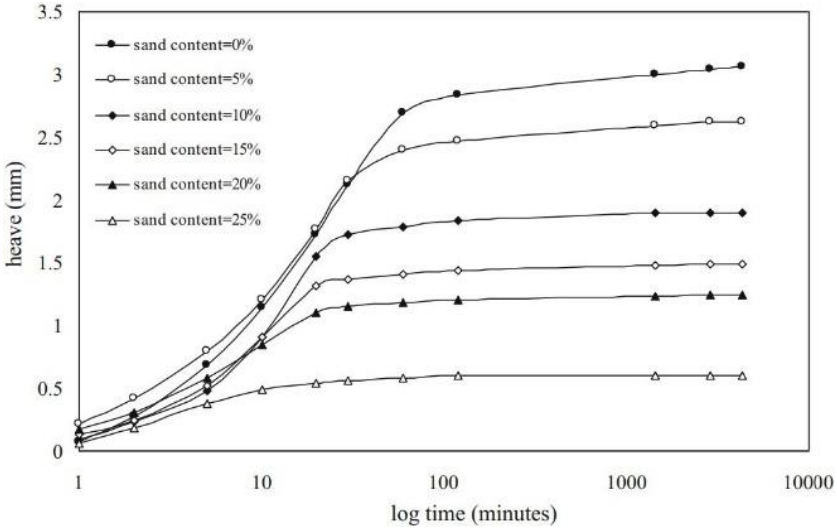


Figure 16. Rate of heave for clay-sand mixture [9]

Fig. 15 shows the heave ratio of clay-sand mixtures having completely different sand percentages. The equilibrium for all mixtures was obtained at the end of 3 days or 4,320 minutes of immersion besides that the equilibrium time becomes shorter as the percentage of sand in the mixture increases. The height of the unmixed clay sample (0% fine sand content) was the highest at 3 mm, giving a swelling potential of 15%. Heave decreases as the percentage of fine sand in the mixture increases. It reaches 0.60 mm at a percentage of sand of 25% which corresponds to the swelling potential (S%) 3%.

Fig. 16 shows the change of swelling potential with an increase in the percentage of sand. A noticeable decrease in swelling potential with increasing sand percentage is observed. Swelling potential (S%) decreases from 15 to 3% at 25% sand content [9].

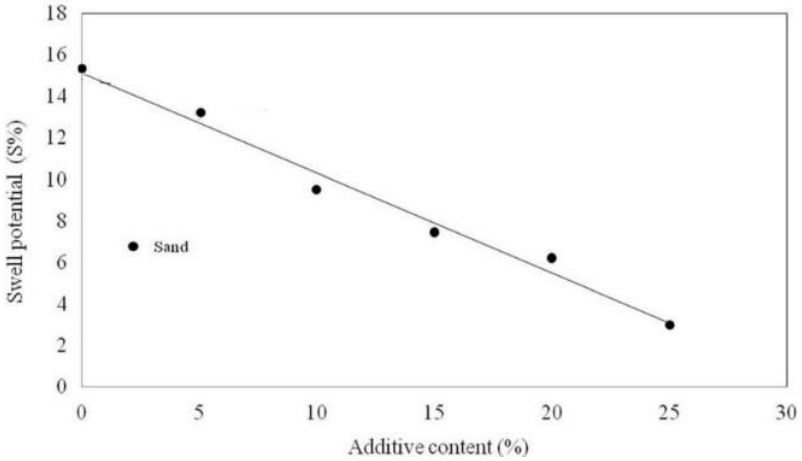


Figure 17. Changing of swell potential with Percentage of sand [9]

Fig. 17 shows the variation of the swell pressure (kPa) as the percentage of sand in the mixture increases. The swelling pressure decreased from 350 to 54 kPa once the percentage of sand in the mixture increased from 0 to 25%. Because the non-expansive material such as sand replaces the expansive clay material in the mixture, the swelling potential decreases, and therefore the swelling pressure also decreases [9].

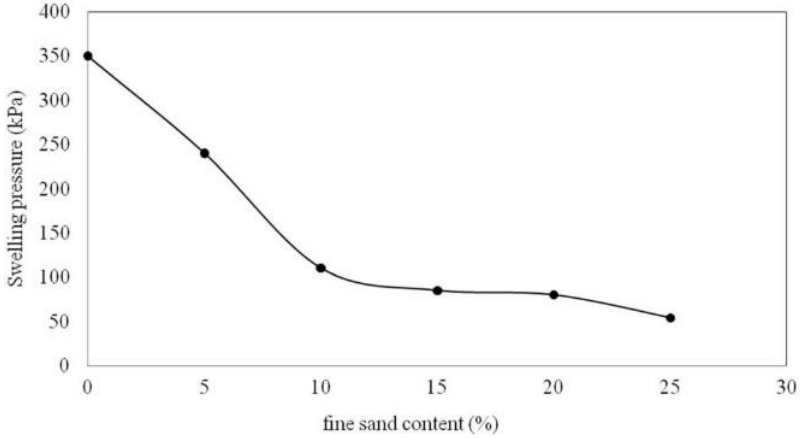


Figure 18. Changing of pressure swell with Percentage of sand [9]

It is evident from previous studies on the swelling potential and swelling pressure [9] [10] [23] [28] [27] [29] [24] that increasing sand in expansive soils improves both of swelling potential and swelling pressure, but most of studies are conducted

on one density and moisture for all percentages of added sand. However, it is better to take more than one density and moisture for each added sand ratio separately, in order to better study the effect of the initial density and the initial moisture and achieve a more comprehensive understanding of the effect of the percentage of sand in the mixture especially the increase in the percentage of sand in the expansive soil causes an increase in its dry density, as indicated by the results of compaction tests.

4.2. Strength characteristics

Khan et al [25] performed unconfined compressive experiments of cohesive soils according to the ASTM Standard Test method (D2166-13). Samples 50 mm in diameter and 110 mm in height were used. The height-to-diameter ratio is equal to 2.2 and it falls within the range (2.0 to 2.5) specified by the ASTM standard. The strain was applied at a rate of 0.5 mm/min and the test was stopped when the load decreased with increased strain or until 15% strain was reached.

Fig. 18 displays the compressive strength (half the peak axial pressure) according to the dry density (Fig. 19 (a)) and water content (Fig. 19 (b)). The compressive strength increases with the increase in the dry density of each material. NC (natural clay) shows the highest increase in compressive strength followed by CS-I (20% sand content) and CS-II (40% sand content). In contrast, an increase in water content shows a reverse trend. NC shows the highest increase in compressive strength followed by CS-I and then CS-II with a decrease in water content [25].

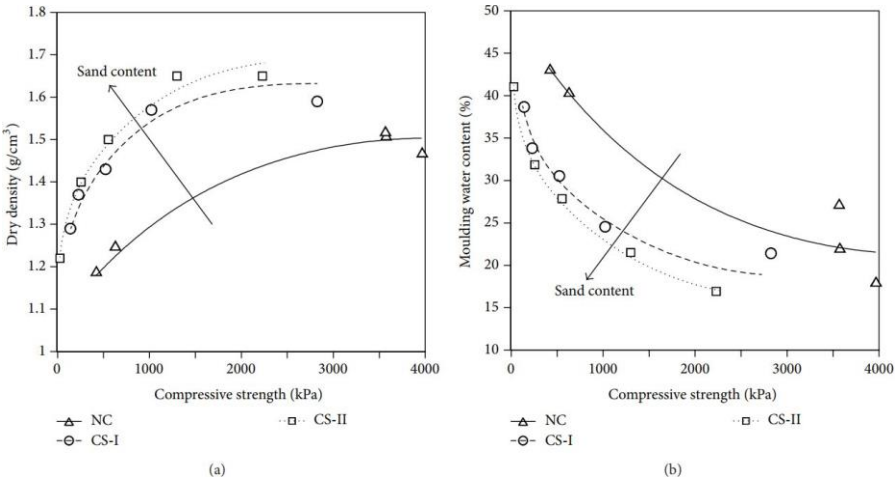


Figure 19. Compressive strength versus dry density and water content according to sand content [25]

Despite the lowest maximum dry density of the clay, however, it showed the highest compressive strength. The higher degree of heterogeneity within the clay-sand mixture compared to the clay resulted in lower strength values as a result of the failure plane had to pass through the weakest zone within the sample [30].

Unconfined compression tests were also performed on statically reconstituted samples according to IS: 2720-part 10 by Nagaraj [8]. Cylindrical specimens were 76 mm high and 38 mm in diameter, and therefore have a height to diameter ratio of 2. Samples were prepared using a static compactor at the optimum moisture for each obtained from the proctor test. All samples were sheared at a strain rate of 0.25mm/min.

Fig. 19 is a plot of unconfined compression strength versus the percentage of sand for clay-sand series. It can be seen from the Fig. 20 that, in general, there is an increase in the strength of unconfined compression with an increase in the percentage of sand up to 40 or 50%, and an extra increase in the percentage of sand resulted in a decrease in strength. Once the percentage of fine particles (i.e. clay) is more, coarse particles (i.e. sand) float within the fine particle matrix, and thus, fines dominate the mechanical behavior of the clay-sand mixture. Sand gradation has shown a significant influence on the strength of the clay-sand mixture in any percentage of sand. Clay-medium sand mixtures have shown to have the maximum strength, followed by clay-fine sand and clay-coarse sand mixtures [8].

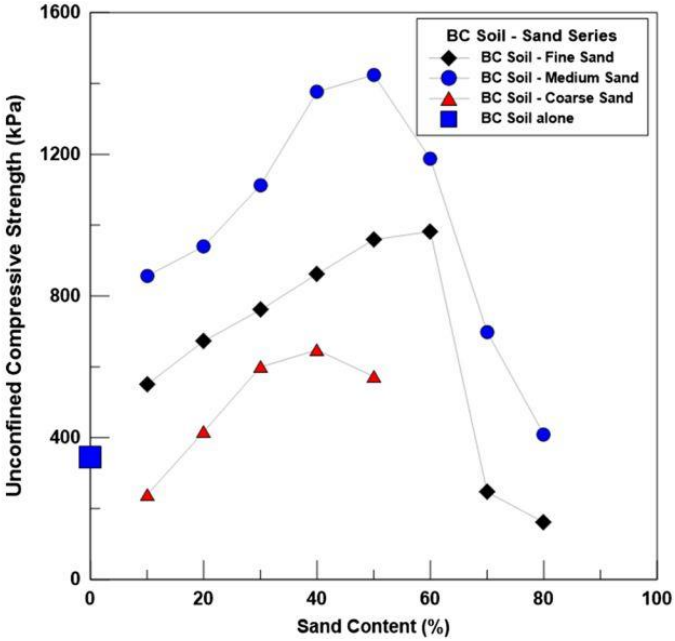


Figure 20. Unconfined compression strength versus percentage of sand with different graded sand [8]

Nagaraj [8] did not mention in his paper the dry density values of the mixtures that he worked on in the study of their strength. Whether as he adopts a constant dry density for all mixtures, or he adopts the maximum dry density corresponding to each mixture which is more likely. Also, its result is opposite to that reached by Khan et al [25], as in Khan's study, the unconfined resistance decreases with an increase in the percentage of sand up to 40%, while it increases in Nagaraj [8] study. Although the Nagaraj result is more logical, as Khan [25] mentioned the presence of heterogeneity sometimes in the mixtures in his research besides that he used only three percentages of sand (0-20-40)%, with a big difference between three percentages. This calls for a more comprehensive study of the effect of moisture and density with the proportion of sand in expansive soils for a more comprehensive understanding of the behavior.

The effect of sand at percentages of 0-10-20-30-40-50% on the strength of expansive soil was studied by Yang et al [24]. The samples were formed according to the maximum dry density and optimum moisture. And Yang et al [24] reached a conclusion that increasing the percentage of sand leads to a decrease in the unconfined compression strength and cohesion, and the sand percentage has a slight

effect on the angle of internal friction. However, the expansive soil that he used contains in its composition a high percentage of sand, about 27%, and thus the results that Yang et al [24] reached do not show the effect of sand on strength at low percentages within the composition of the expansive soil.

Unconfined compression tests also conducted by K. Raw & G. Raw [27] on soil samples compacted at optimum moisture and maximum dry density to determine the unconfined compression strength (UCS). Two types of expansive soils were used: Soil-A and Soil-B. Soil-A and Soil-B have in their composition about 15% and 3% sand, respectively. Soil-B has higher plasticity compared to Soil-A.

K. Raw & G. Raw [27] used in their study four percentages of added sand (Coarse Fraction) 0-25- 40-70%. Fig. 20 shows the variance of unconfined compression strength with sand for two soils. Unconfined compression strength for two soils were 140.06 and 83.39 kPa. From the Fig. 20 it can be seen that unconfined compression strength decreases with the increasing percentage of sand for both soils. However, there is a relatively steep drop in strength up to a certain percentage of sand equal to about 25% for soil-A and 40% for soil-B. Percentage of sand greater than 40% has a little effect on unconfined strength in both soils. This study also used only four percentages of sand with a big difference between each percentage and the next. Therefore, the results do not exhibit the effect of low percentages of sand on unconfined compression strength.

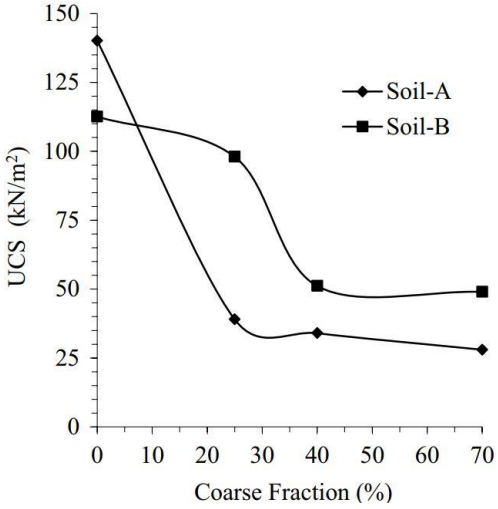


Figure 21. Variation of unconfined compression strength (UCS) with sand (coarse fraction) [27]

Roy [7] has determined soil strength characteristics in terms of CBR with added sand in varying percentages (0-5-10-15)%, and CBR tests were performed in a non-submerged condition in OMC (optimum moisture content) consistent with IS 2720 (part) and the result is presented in Fig. 21.

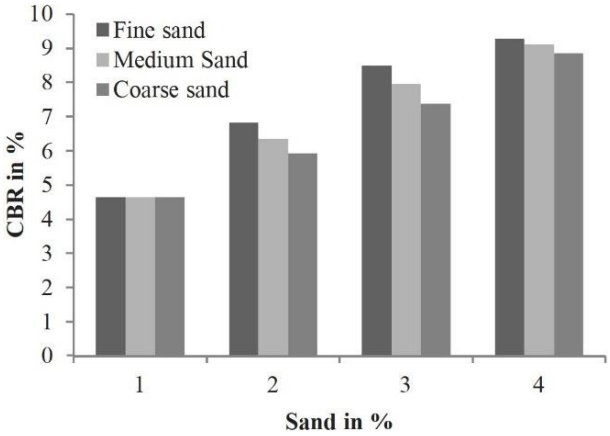


Figure 22. Variation of CBR (Unsoaked) with variable percentage of different sand [7]

Fig. 21 clarifies that the CBR value of the tested soil without any added sand in unsoaked condition is 4.65%. But for mixing different types of sand with their increased percentage, this value has been amended to much higher values. It has additionally been observed that the addition of fine sand offers the best strength, whereas coarse sand offers the lowest strength. For fine sand, the CBR value increases from 4.65% to 9.28% when the percentage of sand increases from 0 to 15%. This improvement might be caused by achieving better grain size distribution for the soil mixture sample. an identical trend was additionally reported by Ravi Shankar [21] and Yang et al [24].

5. Discussion

The results obtained from the review demonstrated that Sand is an effective solution to overcome the problems resulting from the expansive soil such as low strength and big swell amplitude and swell pressure, as it is improved the behavior of expansive soil noticeably. Increasing the hydrophobic sand ratio reduces the hydrophilic clay ratio and thus reduces the efficacy of crystalline and osmotic swelling and improves the behavior of the expansive soil. and this improvement becomes best when the percentage of sand in the expansive soil exceeds 20% as the

absorption forces after this percentage are reduced greatly, and this is due to improving the behavior of expansive soil.

Different results have been observed in the reference studies, and sometimes the results are contradictory. The reason is the heterogeneity in the tested samples, Insufficient percentage of sand used to illustrate the behavior, or the lack of clarification of the densities and moistures that were worked on, this is a must, given the great importance of their influence on the behavior.

Finally, although most studies have shown the good effect of sand on the behavior of expansive soils, more comprehensive studies of the parameters should be worked on to clarify a comprehensive behavior of the expansive soil according to the percentage of sand.

6. Conclusions

The aim of this paper is to conduct an overview of the effect of sand on the behavior of expansive soils, and the results can be summarized as follows:

1. The swelling behavior of clay-sand mixtures is caused by two important mechanisms known as crystalline swelling that relates to the degree of hydration and osmotic swelling is related to the cation in the interlayers.
2. The results got showed the viability of sand in the stabilization of expansive soils and improvement of its behavior.
3. The addition of sand decreases all Atterberg limits. Except for the shrinkage limit. The decrease is slightly affected by the grain size of sand.
4. In compaction characteristics (Proctor tests), Maximum dry density (MDD) values show a significant increase with the addition of any types of sands utilized. On the other hand, OMC values show a gradual decrease with an increase in the values of sand percentages added, independent of the sand utilized.
5. Heave (mm), swell potential (S%) and swelling pressure (ps) decreased significantly with an increase in the percentage of sand in the mixtures. The rate of heave of the mixtures as well decreased with an increase in the percentage of sand in the mixtures.
6. The unconfined compressive strength will increase with a rise in density. Conversely, the compressive strength decreases with a rise in water content of the fabric.
7. Use of fine or medium sand fractions is additionally useful to enhance the undrained strength of clay-sand mixtures.
8. The CBR values are shown a significant amount with addition of any type of sand. Anyway, a better outcome is noticed for mixing of fine sand, so an increase of the percentage of sand in expansive soil might be distinguished as cost-

effective due to reducing the thickness of various layers of flexible pavement because of better improvement in the CBR value.

7. Suggestions for future researches

This research has presented the groundwork for the effecting of sand on the physical and mechanical properties of expansive soil. There is much work that remains to be completed to gain a complete understanding of the effect of sand on the behavior of the expansive soil. Due to the large number of parameters affecting the behavior of the expansive soil, which must be carefully and systematically studied. There are a number of research works that can be presented as suggested below.

1. In recent years, most of the studies were done with expansive soil without separating the sand existing within, and therefore the studies did not take the true percentage of sand on the behavior. One of the aspects to consider for further work is separating the sand from the expansive soil and then adding the desired percentage of sand precisely.
2. It would also be interesting to use more than one initial density and initial moisture for each percentage of added sand to study its effect on behavior.
3. Another aspect to consider in future work is to adopt enough percentages of added sand. It is advised to adopt the following percentages: 0-10-20-30-40-50% at the beginning and increase them if needed.

References

[1] L.D. Jones, I. Jefferson. Expansive soils: In: Burland, J., (ed.) ICE manual of geotechnical engineering. Volume 1, Geotechnical engineering principles, problematic soils and site investigation *ICE Publishing*, London, UK, 2012, pp. 413–441.
URL `http://nora.nerc.ac.uk/id/eprint/17002`

[2] A. Aqeel, Investigation of expansive soils in Obhor Sabkha, Jeddah-Saudi Arabia, *Arabian Journal of Geosciences* 9 (4) (2016) pp. 1–14.
doi: `https://doi.org/10.1007/s12517-016-2341-x`

[3] L. Chen, X. Zhao, L. Lu, The impact of initial placement conditions on the shear strength of expansive soil, *Journal of Computational and Theoretical Nanoscience* 13 (4) (2016) pp. 2419–2425.
doi: `https://doi.org/10.1166/jctn.2016.4596`

[4] A.R. Goodarzi, H.R. Akbari, M. Salimi, Enhanced stabilization of highly expansive clays by mixing cement and silica fume, *Applied Clay Science* 132–133 (2016) pp. 675–684.
doi: https://doi.org/10.1016/j.clay.2016.08.023

[5] J. Nelson, D.J. Miller, Expansive soils: problems and practice in foundation and pavement engineering, *John Wiley & Sons*,1997.
URL http://nora.nerc.ac.uk/id/eprint/17002

[6] M. Fuchs, Development of the classification of high swelling clay content soils of Hungary based on diagnostic approach, Ph.D. thesis, *Szent István University*, Hungary, Gödöllő, 2012.
URL
http://archivum.szie.hu/JaDoX_Portlets/documents/document_13382_section_13836.pdf

[7] T.K. Roy, Influence of sand on strength characteristics of cohesive soil for using as subgrade of road, *Procedia-Social and Behavioral Sciences* 104 (2013) pp. 218–224.
doi: https://doi.org/10.1016/j.sbspro.2013.11.114

[8] H.B. Nagaraj, Influence of gradation and proportion of sand on stress-strain behavior of clay-sand mixtures, *International Journal of Geo-Engineering* 7 (1) (2016) pp. 1–10.
doi: https://doi.org/10.1186/s40703-016-0033-8

[9] B.R. Phanikumar, S. Dembla, A. Yatindra, Swelling Behaviour of an Expansive Clay Blended With Fine Sand and Fly Ash, *Geotechnical and Geological Engineering* 39 (1) (2021) pp. 583–591.
doi: https://doi.org/10.1007/s10706-020-01480-6

[10] L. Bahia, B. Ramdane, Sand: an additive for stabilzation of swelling clay soils, *International Journal of Geosciences* 3 (4) (2012) pp. 719–725.
doi: https://doi.org/10.4236/ijg.2012.34072

[11] Y. Deng, Z. Wu et al., Sand fraction effect on hydro-mechanical behavior of sand-clay mixture, Applied Clay Science 135 (2017) pp. 355–361.
doi: https://doi.org/10.1016/j.clay.2016.10.017

[12] M. D. Foster, The Relation Between Composition and Swelling in Clays, *Clays and Clay Minerals* 3 (1) (1954) pp. 205–220.
doi: https://doi.org/10.1346/ccmn.1954.0030117

[13] H. Komine, N. Ogata, Prediction for swelling characteristics of compacted bentonite, *Canadian Geotechnical Journal* 33 (1) (1996) pp. 11–22.
doi: https://doi.org/10.1139/t96-021

[14] R.K. Taylor, T.J. Smith, The engineering geology of clay minerals: swelling, shrinking and mudrock breakdown, *Clay Minerals* 21 (3) (1986) pp. 235–260.
doi: https://doi.org/10.1180/claymin.1986.021.3.01

[15] F.T. Madsen, M. Müller-Vonmoos, The swelling behavior of clay, *Applied Clay Science* 4 (2) (1989) pp.143-156.
doi: https://doi.org/10.1016/0169-1317(89)90005-7

[16] L. David Suits, T. Sheahan, W. Likos, Measurement of Crystalline Swelling in Expansive Clay, *Geotechnical Testing Journal* 27 (6) (2004) pp. 540–546.
doi: https://doi.org/10.1520/GTJ11857

[17] E.I. Stavridakis, A solution to the problem of predicting the suitability of silty–clayey materials for cement-stabilization, *Geotechnical & Geological Engineering* 24 (2) (2006) pp. 379–398.
doi: https://doi.org/10.1007/s10706-004-7934-6

[18] L. Liu, I. Neretnieks, Interaction between colloidal particles. Literature review. SKB Technical Report TR-10-26, Stockholm, Sweden, 2010 [cited 2021-05-01].
URL
https://inis.iaea.org/collection/NCLCollectionStore/_Public/41/103/41103173.pdf?r=1

[19] R.N. Yong, Soil suction and soil-water potentials in swelling clays in engineered clay barriers, *Engineering geology* 54 (1–2) (1999) pp. 3–13.
doi: https://doi.org/10.1016/S0013-7952(99)00056-3

[20] J.A. Greathouse, S.E. Feller, D.A. McQuarrie, The modified Gouy-Chapman theory: comparisons between electrical double layer models of clay swelling, *Langmuir* 10 (7) (1994) pp. 2125–2130. doi: https://doi.org/10.1021/la00019a018

[21] A.U. Shankar, A. Chandrasekhar, P.H. Bhat, Experimental investigations on lithomargic clay stabilized with sand and coir, *Indian Highways* 40 (2) (2012) pp. 21–31.

[22] V. Srikanth, A.K. Mishra, Atterberg limits of sand-bentonite mixes and the influence of sand composition, *Geotechnical Characterisation and Geoenvironmental Engineering* 16 (2019) pp. 139–145. doi: https://doi.org/10.1007/978-981-13-0899-4_17

[23] V. Srikanth, A.K. Mishra, A laboratory study on the geotechnical characteristics of sand–bentonite mixtures and the role of particle size of sand, *International Journal of Geosynthetics and Ground Engineering* 2 (1) (2016) pp. 1–10. doi: https://doi.org10.1007/s40891-015-0043-1

[24] J. Yang, X. Li et al., Expansive soil improvement with weathered sand, in ICTE 2013 - *Proceedings of the 4th International Conference on Transportation Engineering* (2013) pp. 2082–2087. doi: https://doi.org/10.1061/9780784413159.303

[25] F.S. Khan, S. Azam et al., Compressive strength of compacted clay-sand mixes, *Advances in Materials Science and Engineering* (2014) pp. 1–6. doi: https://doi.org/10.1155/2014/921815

[26] C. Gupta, R.K. Sharma, Influence of marble dust, fly ash and beas sand on sub grade characteristics of expansive soil, *Journal of Mechanical and Civil Engineering* 13 (2014) pp. 13–18.

[27] K. Mallikarjuna Rao, G.V.R. Subba Rao, Influence of Coarse Fraction on Characteristics of Expansive Soil–Sand Mixtures, *International Journal of Geosynthetics and Ground Engineering* 4 (19) (2018) pp. 1–7. doi: https://doi.org/10.1007/s40891-018-0136-8

[28] Y.K. Atemimi, Effect of the grain size of sand on expansive soil, *Key Engineering Materials* 857 (2020) pp. 367–373.

doi:
https://doi.org/10.4028/www.scientific.net/KEM.857.367

[29] S.B. Ikizler, M. Aytekin, M. Vekli, Reductions in swelling pressure of expansive soil stabilized using EPS geofoam and sand, *Geosynthetics International* 16 (3) (2009) pp. 216–221.
doi: https://doi.org/10.1680/gein.2009.16.3.216

[30] C.E. Mullins, K.P. Panayiotopoulos, The strength of unsaturated mixtures of sand and kaolin and the concept of effective stress, *Journal of Soil Science* 35 (3) (1984) pp. 459–468.
doi: https://doi.org/10.1111/j.1365-2389.1984.tb00303.x

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 553-576, 2021
DOI: 10.14513/actatechjaur.00591

Possibilities of porous-structure representation – an overview

H. Szűcs[1,2*], B. Vehovszky[1]

[1]Széchenyi István University, Department of Whole Vehicle Engineering, Egyetem tér 1, 9026 Győr, Hungary
[2]Audi Hungaria Zrt., Engine Development, Audi Hungária út 1, 9027 Győr, Hungary
***e-mail: szucsherman@outlook.hu**

Submitted: 26/01/2021; Accepted: 13/04/2021; Published online: 21/04/2021

Abstract: Porous media can be found in all areas of scientific life, such as medicine, civil engineering, material science, fluid dynamics. Computing has achieved high efficiency and computational capacity – so far. However, three-dimensional Computational Fluid Dynamics (CFD) simulations of microstructure remain significant challenges. Pore-scale simulations can help understand the physical processes and determine macroscopic parameters such as the high-frequency limit of dynamic tortuosity, viscous, and thermal characteristic lengths. Independent of whether the computational problem is two or three-dimensional, the geometry as input parameter must be prepared. For this reason, geometry representation methods play a crucial role in the analysis at the pore-scale, especially in numerical simulations. In this article, an insight into microstructures' visualization capabilities is provided essentially for CFD simulations.

Keywords: *Pore-scale simulation; Micro-structure, Porous-structure, Microscopy*

1. Introduction

Researchers and engineers are mainly concerned with fluid flow and transport phenomena on a scale that is much larger than atomic; the pore-scale is commonly adopted in practice [1]. Fluid flows at pore-scale are detailed by the equation found on the law of Darcy [2] [3], together with the transport of mass by the advection-dispersion equation [4], in which bulk averaged fluxes satisfy the two equations [4] [5]. A correct and thorough understanding of larger-scale flow and transport processes requires knowledge of micro-scale processes, highly dependent on input porous media' geometrical details [6-9]. For example, simple stagnant zones can influence the transport phenomena and dispersive mixing [4]. Today, multiple flow and transport problems [4] in porous material necessary to be studied with the help of different numerical approaches, which allows the analysis of various processes and phenomena [2] in the microstructure. Moreover, their impacts on a larger scale (i.e., fluid flow behavior at the macro-scale) also must be studied using these methods. Therefore, the so-called "linking" between specific scales plays more and more significant role in the industrial field despite the reachable computational power, which heavily limits the spread of multi-scale methods [10] [11].

The natural pore materials have unique geometries, which are incredibly elaborate and enormously complicated. Therefore, solving the governing equations (Navier-Stokes or Stokes in case of Creeping flow) for actual pore materials is absolutely a challenging and exciting computational fluid dynamics (CFD) simulation task currently. Hence different numerical approaches are suited to handle real porous geometries. Therefore, it is needed to develop the simulation methods for pore-scale analysis [4]. With the continuous improvement of various visualization techniques, which can also imagine the internal structure [12] [13], and micro-fluidics experiments, more pore-scale physics and higher spatial resolution can be obtained [14-18]. The area of fluid flow and transport in porous materials [2] (micro-structures) has been completely changed by our ability to create images, in three dimensions, of porous media with modern microscopes at several resolution ranges [6] [9] [19]. Recent modern advances of various imaging techniques, which can represent the material's specific geometry with the complex inside-structure, can help in numerical simulation (e.g., CFD-simulations) with negligible loss in pore topologies and porous media [20-22]. It should be noted that the other imaging methods are also used for simulation, which cannot present an image of the inside structure of the material.

2. Theory of Porous Media

The pore-scale is characterized by the solid grains or body (black parts) of the porous material as well as the pore-spaces, which contain fluid (blue parts); Fig. 1. shows an example. If the pores are closed and not connected with each other, representing the whole solid by a macroscopic material law is ordinary practical. In a way, the material attributions are 'smeared' over the solid frame and its pores. However, a too 'coarse' definition might affect neglecting any probably critical effects [4].

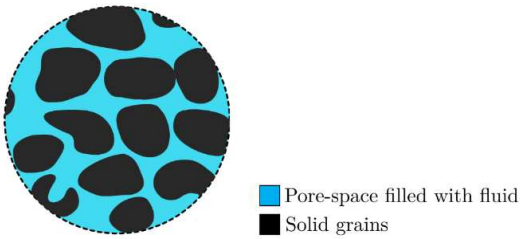


Figure 1. Schematic illustration of a pore-scale porous medium [4] [23]

Microscopic pores and cracks can have a significant influence on macroscopic behavior [20]. At all points of the present scale, either a fluid (marked with blue) or a solid phase (marked with black) has definite phase-boundaries. The fluid placed in the void spaces could be specified with several elemental quantities like density (ρ), kinetic (ν) and dynamic viscosity (η). Since the void boundaries' features are applicable, it is feasible to directly taking the particular void space geometries and merging with the correct boundary conditions in order to detail the flows of fluids – solution of the conservations of mass, momentum, and energy – in the pore space [4]. Fig. 2. shows two examples of porous media in the field of material science [1][20].

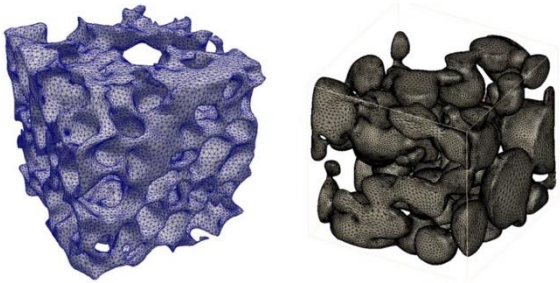


Figure 2. Two numerical models of porous media in the field of engineering [1] [20]

The porous medium simulation of fluid flow and transport could be divided into two types separated by scale. Simulating the fluid-transport inside the individual pores constituting the porous material's pore space is the first type. In this category, methods solve the so-called "governing equations of motion" inside the pore space where pores' wall treatment is a no-slip boundary. This first type is denoted as pore-scale simulations. The second type treats larger scales where the definition of averaged effective-properties such as porosity and permeability are possible. However, these are up-scaled properties from the smaller scales of porous media, and the physics and flow that take place on the pore-scale ultimately determine how larger-scale flows must be treated [1].

Generally, the porous material contains a solid skeleton or matrix within a massive number of microscopic voids/pores. The tiny porous inside the structure are commonly attached. For this reason, different flow and transport processes can happen inside. Natural substances are good examples, such as rocks, soils, sandstone, and artificial or industrial materials. Porous media is commonly applied in the field of engineering or sciences like biophysics, geosciences, or material science. Thus, it is incredibly critical in the research and application of fluid transport and flow through them. The flow and transport processes are affected by the pore structure, the matrix's physical characteristics, and the fluid-molecules (in the pore-space) [4].

These processes are treated as quite complicated due to the complexity of micro-structures. The pores mostly have irregular surfaces; therefore, fluid flow through them is genuinely problematic. Several pores make dead ends and have a severe effect on the flow and transport behavior. In single-phase flow through porous materials, just one fluid phase is flowing over the pores, e.g., water. Multi-phase flow or two-phase flow is in the case where multiple phases are flowing over the void spaces. Multi-phase flow can be present in unsaturated zones. Such as a petroleum recovery of hydrocarbons and various industrial contexts [4].

3. Porous-Structure Representation

New perspectives on the study of microstructures are being opened by modern three-dimensional imaging methods/tools capable of reconstructing porous materials' structure [6] [9] [19]. Different modern imaging techniques are able to provide high-fidelity visualization [24] and characterization of 2-D/3-D porous micro-structures [25] such as pore topologies with negligible loss, i.e., affordable computational resources. The digital micro-structure can be applied for a variety of simulation or diagnostic purposes [6] [9].

3.1. Two or Three-Dimensional Methods without Internal Structure

The commonly used techniques to provide 2-D (surface) or 3-D (without internal structure) visualizations of micro-structures are summarized in Table 1 [9] [14]. The methods are the Scanning Tunneling Microscopes (STM) [26], the Scanning Electron Microscopes (SEM) [27-28] (it has Backscattered Electron (BSE) mode [28]), the Optical Microscopes (OM) [14], the Scanning Probe Microscopes (SPM) [26] (Subtype: Atomic Force Microscope (AFM) [29]), and Confocal Laser Scanning Microscopes (CLSM) [30].

Table 1. 2-D (surface) and 3-D (without internal structure) geometry representation methods [6] [7] [9] [12-14] [19] [25-36] [38] [39]

Method	***Range***	***Resolution***	***Feature***
STM	500 × 500 nm	0.1 nm – 30 nm	• STM represents the surface, which makes it possible to analyze a considerable number of characteristics, e.g., surface defects, roughness, and inspecting the molecule's feature such as conformation and size;
SEM	10 × 10 ×10 μm	1 nm – 10 mm	• SEM is an efficient tool for characterization of particle size at the surface, but it does not fit for pore size; • Backscattered electrons produced at each point are collected, with practical electronics to imaging on the computer;
OM	50 × 50 μm	0.1 μm – 1.0 mm	• OM is the primary technique used for the analysis of the surface profile. • It is required for quantitative studies as it provides the base of image analysis systems illustrations/visual documentation;
SPM	10 × 10 μm	0.1 nm – 70 μm	• SPM has excellent resolution (spatial) and can imaging the structure and the surface topography also; • AFM is an effective technique of SPM for single-cell characterization;
CLSM	400 × 400 × 0.8 μm	180 nm – 1.5 mm	• CLSM has a unique ability to provide 3D images of the porous media with submicron resolution; • Although CLSM is greatly limited in depth (z-dimension);

Scanning Tunneling Microscopy (STM) is a type of electron microscopy. It is capable of producing an image of a sample's surface at the molecular scale [26]. The STM's fundamental is based on quantum tunneling theory, Fig. 3 shows the schematic representation. Unlike SPM techniques (such as AFM), this technique normally not requires contact with the specimen's surface, although some volts or millivolts are used between the specimen and the tip, while the current of tunneling is measured [40]. Just after the conducting tip is appropriately positioned, several angstroms from a semiconducting or metallic surface, a bias voltage is used among the surface. The probe tip makes it possible to tunnel electrons via the gap. When the tip scanning over the surface, the fluctuation in the tunneling current is registered. Therefore, it can create topographical images of the sample's surface. [26] The STM's regular depth and lateral resolutions are 0.1 and 0.01 nm. STM is a well-known technique, which is surface-sensitive, and unlike TEM, it requires a definitive clean surface, so it has a low success rate [26] [40] [41].

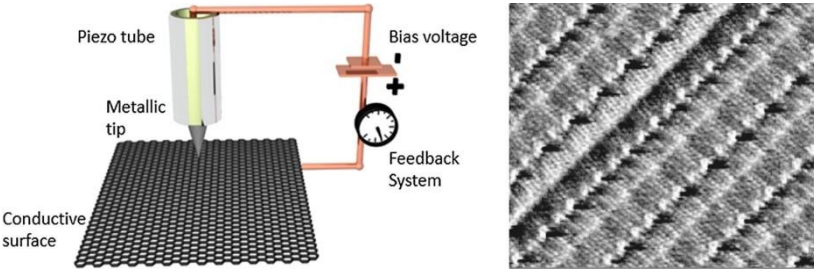


Figure 3. Schematic representation of an STM device and an STM-image [41]

Scanning Electron Microscope (SEM) is able to produce much information like surface (2D) structures, topographies, crystallines, and chemical compositions [28]. Present SEM-techniques are considered to fall within the general class of electron microscopes, which create 2D-images of different samples. In terms of its operation, illustrated in Fig. 4., the surface is scanned with a focused electron beam [42]. The electrons – used for scanning – have interaction with the sample's (investigated material) atoms. This process provides numerous data about the composition of the whole sample and the surface's topography (also see Fig. 4) [27]. The scanning process of the beam of electrons follows raster-scan-patterns, plus the beam's position relates to the intensity of the detected signal, which can produce two-dimensional images [38]. Backscattered Electron (BSE) operation mode has been commonly used to characterize phases with diverse chemical bases in various materials [28]. Moreover, BSE is generally used to study porous-structure substance in which differences in atomic number are expected [27]. Thus, backscattered images can be a proper method to characterize metallic phases [42].

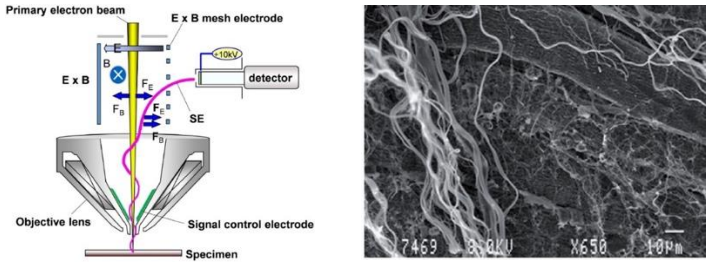


Figure 4. Schematic illustration of SEM signal detection and an SEM image [43-44]

Optical Microscopy possesses a wide field of view, ensuring a high throughput screening and the possibility to collect vast amounts of data (see Fig. 5) for post-processing and machine learning [35]. In the case of optical microscopy techniques, the preparation of the sample is not necessary [45]. These methods have low applicability in 3D-CAD modeling because it requires comprehensive data manipulation. For example, a high resolution of the sample requires to be split into fragile slides [14]. This method is profoundly labor-intensive and resulting 2D images, which need to be digitally processed to build the 3D images [46]. If the sample is inhomogeneous, this time-consuming process can provide the best solution to get a 3D image for CFD simulations. This method is hugely computation and memory demanding; therefore, it is incredibly challenging nowadays for computers to reach a level capable of identifying cells by optical characteristics [25].

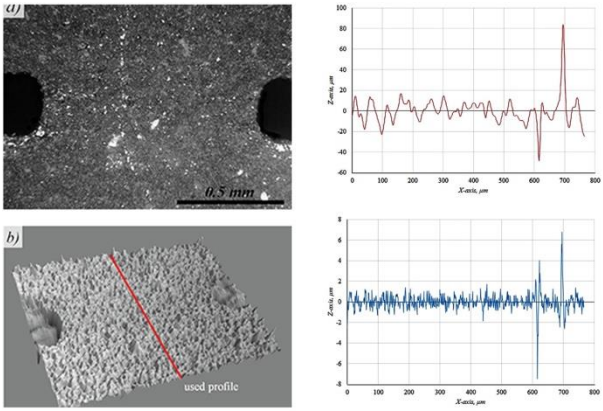


Figure 5. Fully focused optical image (a) and 3D rendering of surface morphology (surface profilometry) (b), the waviness profile (c) and roughness profile (d) taken from line profile across the image [47]

The scanning probe microscopes (SPM) can give several pieces of information about the material sample at the atomic scale. One of the SPMs is the atomic force microscopy (AFM) with extremely high resolution. Its application's main reasons are the following: the powerful piezoelectric actuation for the precise tip/sample driving; applying sharp tip; controlling with rapid feedback for tip-sample interaction defined by the user. The control of the movement of the tip/sample is possible at the picometer scale [26]. With this method, surface topology and intermolecular forces can be directly measured with a resolution of ≈ 0.1 nm vertically. AFM scans the surface of the sample with a sharp probe and collects information about the surface properties. This probe is assembled with a cantilever, substrate, and a sharp tip. AFM can provide both 2D and 3D images (similar to OM), which is represented in Fig. 6 [39].

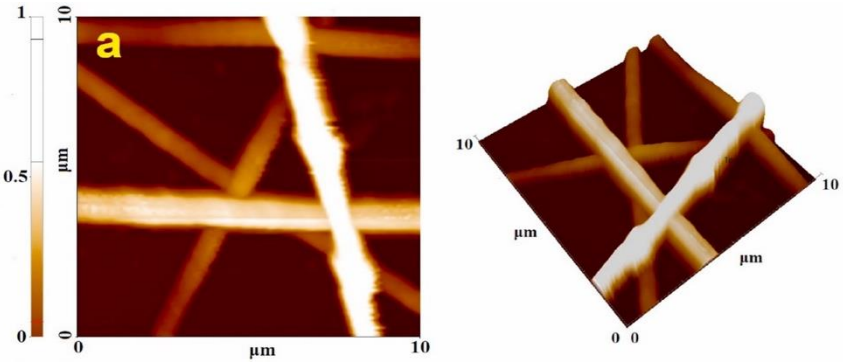


Figure 6. AFM 2D and 3D topography images [48]

Confocal Laser Scanning Microscopy (CLSM) is an effective device to visualize the porous materials [30] [49]. This method's resolution is relatively lower than electron microscopy (e.g., SEM), but its most significant advantage is that remarkably fewer sample preparations are necessary, which could be intensely time-consuming [50]. Moreover, CLSMs make possible the 3D imaging of porous media (without internal structure) by collecting the fluorescence signals from the sample's different planes and assembling the 2D-images (planar) into a 3D. Fig. 7 shows the process of imaging [51]. For this reason, CLSM is suitable for three-dimensional "live" imagine, allowing researchers to study both the biological and atomic processes [50]. Notably, that this image is the surface of the media. The maximum resolution is specified by the light's diffraction limit, like in the case of OM [46] [52]. The compensation of light collection's losses by raising the exciting laser light is not an appropriate solution because more laser power leads to more phototoxicity and photobleaching [25].

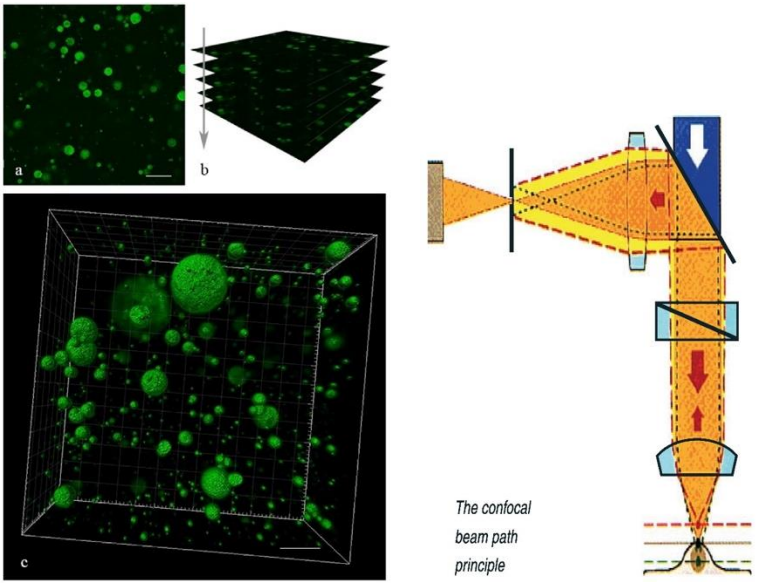


Figure 7. (a) 2D, (b) Z-stack, and (c) 3D CLSM images, and the principle of confocal (d) [53] [54]

3.2. Three-Dimensional Methods with Internal Structure

Three-dimensional microstructures or pore materials, including the whole internal structure, can be imagined with numerous techniques like Nuclear Magnetic Resonance (NMR) (one of NMR application: Magnetic Resonance Imaging (MRI) [34]), Focused Ion Beam technique (FIB) [7], Transmission Electron Microscopy (TEM) [12], X-ray Micro-Computed Tomography (μ-CT or XCT) [13]. Besides, these mentioned microscopy techniques are commonly combined, mentioning some effective methods of them: The Focused Ion-Beam combined with Scanning Electron Microscopes (FIB-SEM) [28] [37] and the Serial Block-Face combined with Scanning Electron Microscopes (SBF-SEM) [28]. These modern microscopy imaging techniques are able to represent a porous medium's complex geometry with internal structure into a 2-D or 3-D image, which are listed and summarized in Table 2 [6] [9].

Table 2. 3-D (with internal structure) geometry representation methods [6] [7] [9] [12-14] [19] [26-36] [38] [39] [55]

Method	***Range***	***Resolution***	***Feature***
NMR	10 × 10 ×10 cm	100 µm – 1.5 cm	• Effective, non-invasive method, which is suitable for studying spatial resolved fluid saturation; • MRI is an application of NMR, which gives 3-D geometry with internal structure • However, it is less likely available to researchers and expensive;
FIB	5 × 5 ×5 µm	2 nm – 5 nm	• FIB operates in principle like the SEM technique; • FIB uses an ion (not electron) beam instead to create secondary electrons and ions for imaging;
TEM	10 × 10 ×10 µm	0.5 nm – 50 µm	• Suitable for an accurate representation of pore size and distribution with high resolution, it highlights the particle's center • Other 3-D imaging methods are required for a complete examination of porous structures;
Micro-CT	10 × 10 ×10 mm	5 µm – 150 µm	• XCT is capable of various in situ time-resolved investigations during imaging; • 3-D imagine within the inside structure is possible, but the sample size is limited and not suitable for nano-structures;
FIB-SEM	45 × 45 ×45 µm	4 nm – 16 nm	• Common FIB technique (using electrons and ions) in combination with simple SEM (using an ion column) • One of the most potent methods for 3-D imaging of porous microstructures;
SBF-SEM	80 × 80 ×80 µm	10 nm – 100 nm	• SBF-SEM technique is a useful tool for characterizing porous media with relatively high resolution;

Nuclear Magnetic Resonance (also known as NMR) is a widely used spectroscopic method for observation of local magnetic fields, which formed around the atomic nuclei. The investigated materials are put into a magnetic field. The signal of NMR is generated in the following way: radio waves excite the nuclei sample to nuclear magnetic resonance; this can be detected using a simple, sensitive radio receiver [34]. Examining a simple atom in a molecule, we find that the intramolecular magnetic field around it modifies the frequency of resonance. Consequently, it gives a report of the specific functional group of the molecule as well as its electric structure. This formed field is quite special and individual and specifies the compounds; therefore, these techniques are widely applied in chemistry for identifying monomolecular organic compounds [31]. Magnetic Resonance Imaging (MRI) techniques are found in the theory that with the main magnetic field, hydrogen (H_2) protons align themselves. Added a second, orthogonally magnetic field to the main field, which is oscillating at radio frequencies, then protons are pushed out of alignment. When these protons turn back to align with the main magnetic field, they are emitting a measurable signal at radio frequencies, which can be detected. Various material's protons have dissimilar realigning speeds. This attribution is suited to analyze different materials that specify the actual media. MRI-method is widely applied, especially in medical science or the food industry; an example is given in Fig. 8. The main reasons for its attractiveness are that MRIs are applying non-ionizing frequencies, and they are not destructive. As disadvantage can be mentioned that the imaging process is time-consuming, and the resolution of image is restricted [7].

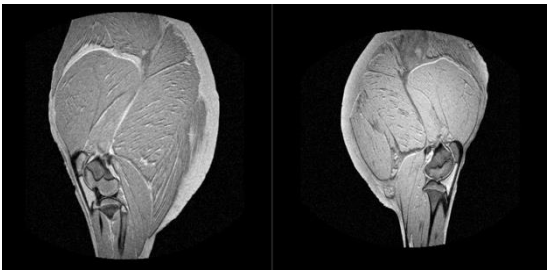


Figure 8. MRI image of different hams [31]

The fundamental of Focused Ion Beam (FIB) is genuinely like SEM [7], but it uses a focused beam of ions (see Fig. 9), not an electron beam. The diameter of the beam is down to about five nm. Using the FIB technique, internal structures can be imaged [33]. The equipment's primary (main) ion beam must be working at low currents for imaging of different materials. High currents give other alternative usages of FIB, like cutting, milling (see Fig. 9), drilling, or structuring at the atomic-scale, where critical material ablations are observable [7].

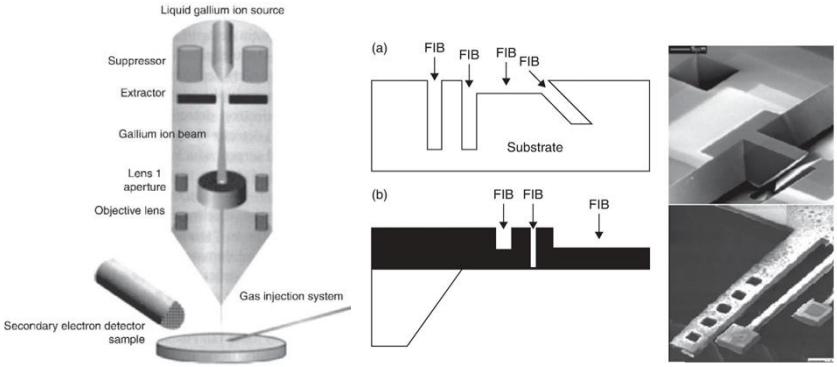


Figure 9. Schematic illustration of focused ion beam system and two examples for FIB milling: milling pattern (a) and milling structure (b) [56]

Transmission electron microscopy (TEM) is widely used and a viral technique in the characterization of nanomaterials [12]. For images of nanomaterials, the same spatial resolution as the atomic level can be achieved by using TEM [57]. The present device uses the same fundamental principles as Optical Microscopy (OM), but it utilizes not light but electrons. Fig. 10 shows a schema of the TEM technique with two different images. The wavelength of light is shorter than in the case of electrons. Consequently, TEM can offer high-resolution images compared to optical microscopy. Generally, TEM techniques allow researchers to investigate objects in the order of 10^{-10} m, reaching atomic levels [7].

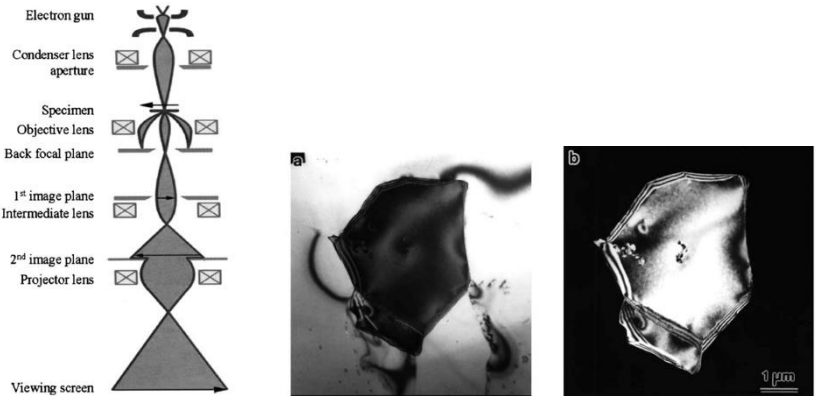


Figure 10. Schematic optical diagram of TEM (left), imagines with different objective aperture positions: bright- and dark-fields microscopy images (center and right, respectively) [58] [59]

X-ray Computed Tomography – also known μ-CT or XCT – makes it possible to enhance the applied resolution (approx. 7-8 orders of magnitude regarding the geometry volume), which strongly relates to the application [60]. This microscope's required resolution depends on the porous material's actual characteristic (especially shape and type) [7]. Micro-CTs are effective tools in the research of evaluating 3D-structures of the different porous scaffolds. Consequently, XCT can quantify important parameters, e.g., average pore size, pore interconnectivity, and porosity [19]. As both animal studies and cell culture are costly and time-consuming, advanced scaffold's optimal micro-structure needs to be verified in the early phase of development [61]. Computed Tomography imaging is an effective and powerful technique that is not destructive and utilizes thermal neutrons, but a nuclear reactor is needed [60]. A micro-CT scanner's direct output is commonly a grey-scale image (see Fig. 11) that is post-processed by filtering and applying segmentation techniques to identify discrete material phases in the image. Fig. 11 shows the whole process of 3D reconstruction of an XCT image [7] [19].

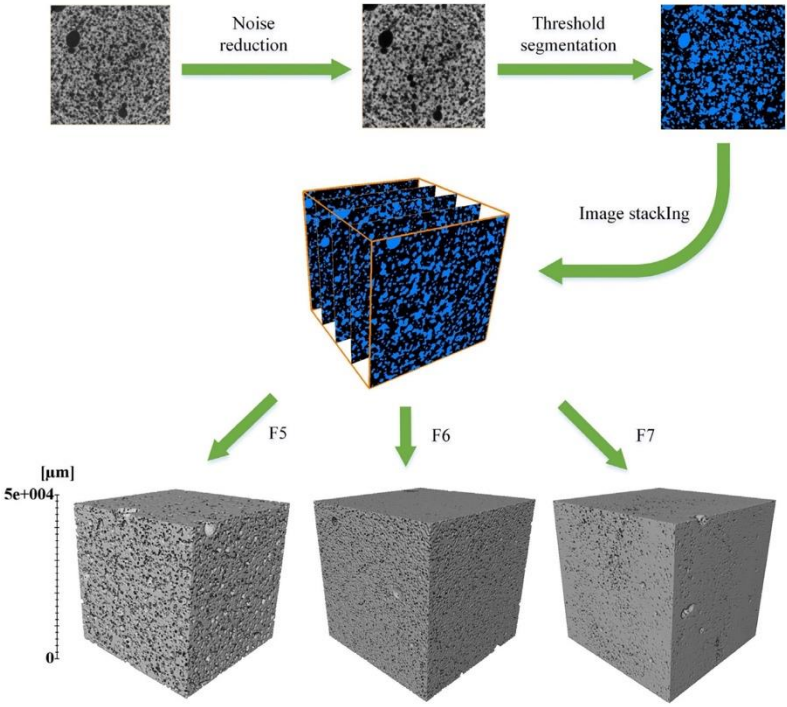


Figure 11. Process of XCT image 3D reconstruction [62]

Focused Ion-Beam – Scanning Electron Microscopies (FIB-SEM) are a combination of an ordinary SEM microscopy applying ion column (i.e., gallium) and the specific beams – namely, electrons and ions (FIB) – focused on one concrete point [32-33]. The latter purpose is to remove the unnecessary material and preparation of flat 2-D surfaces. A simple FIB can be realized based on SEM, which can be used to identify the location of homogeneous portions (in terms of X-ray attenuation and surface attributes) and, thus, the specific volumes for the process milling [32]. With the FIB-SEM methods, internal structures can also be mapped; an example is given for its images in Fig. 12 [33].

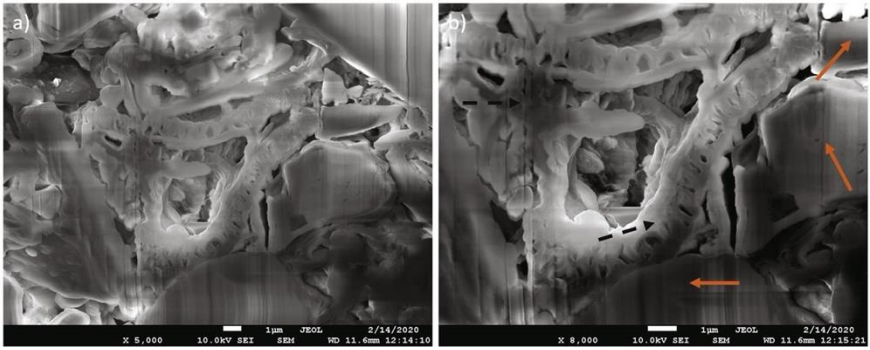


Figure 12. The imagine of FIB-SEM of an inner structure, using secondary (a) and back-scattering electrons (b) [32]

Serial Block-Face Scanning Electron Microscopes (SBF-SEM) are new and innovative techniques developed by Horstmann et al. [36] at the Max Planck Institute. Their commercialized microscopy is available as an accessory to several SEM models. Additionally, the microscope is operating with automatic serial sectioning and scanning of the sample embedded in SEM. Using the SBF-SEM technique, a three-dimensional image can be reconstructed from an aligned density map of the sample by choosing longer-term, extended processing of the sample (from 1 day to 30 days). The SEM's image is similar to a conventional TEM, except that the resolution of the SEM is lower than the TEM at higher magnification ranges. At lower magnifications, the SEM has sufficient resolution to the competition with TEM used in the same range, illustrated in Fig. 13 [36].

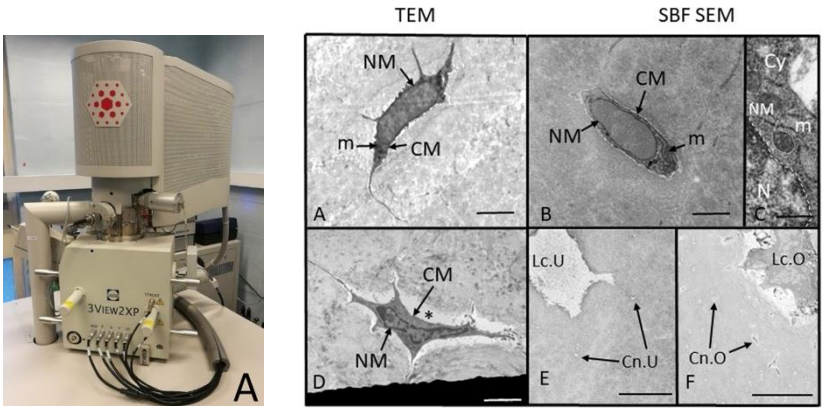


Figure 13.Modern SBF-SEM system and TEM and SBF SEM images comparison [63]

4. Summary

Geometry representation methods [25] [38] are especially critical because the two- or three-dimensional images are input parameters of numerical simulations. In complex cases, there are two different ways to produce the 3D geometry [38] for simulations as input: 3D imagines with the whole internal structure and reconstruction from numerous, separate 2D surfaces [6]. In this context, pore-scale simulations can improve the study and research process [64], as well as the comprehension of the physical processes at the pore-level and allow the determination of several macroscopic parameters. However, these parameters – obtained by simulation – can be applied in a continuity model, which could simply be determined from costly plus too consumptive of time field measurements or laboratory experiments. Besides, it is essential to note that it is not possible to implement experimental setups for various practical issues. Theoretical research in fluid dynamics supports simulation approaches' continuous improvement at different scales such as micro and continuum-scales. In addition to that, the practical application of these theoretical methods, especially in automotive engineering, requires in-depth technical knowledge of fluid behavior. Furthermore, it is necessary to extend them, involving multi-phase flows, fluid flow interactions with other physical phenomena – also known as multi-physics problems.

5. Outlook and Future Work

With the continuous development and spread of fluid flow simulations, the representation method of different materials is becoming increasingly important, especially in the field of microstructures. In the future, there will be a particular need to develop representation methods that can provide high-quality input with wide range for simulations without any further, time-consuming post-processing. The availability of such applicable methods would be of great benefit in material research.

In our subsequent research, we will scan a sample of porous material using a 3D representative method that can provide adequate input for CFD simulations. The sample size and the equipment availability also need to be considered to select the suitable/applicable method. At this time, the sample is still not determined. After scanning the material, the resulting geometry must be prepared for mesh generation. This future research aims to investigate the flow properties and characteristics of microstructures in a numerical simulation environment, e.g., viscous and thermal characteristic length, the high-frequency limit of the dynamic tortuosity, etc. Finally, the results will be compared to laboratory measurements and experiments to validate the simulation.

Acknowledgement

The publishing of this paper was supported by the Department of Whole Vehicle Engineering of Széchenyi István University.

References

[1] T. Ramstad, C. F. Berg, K. Thompson, Pore-Scale Simulations of Single- and Two-Phase Flow in Porous Media: Approaches and Applications, *Transport in Porous Media* 130 (2019) pp. 77–104.
doi: https://doi.org/10.1007/s11242-019-01289-9

[2] N. Kovalchuk, C. Hadjistassou, Laws and Principles Governing Fluid Flow through Porous Media, *The European Physical Journal E* 42 (2019) pp. 1–6.
doi: https://doi.org/10.1140/epje/i2019-11819-6

[3] W. R. Zimmerman, Fluid Flow In Porous Media, in: W. R. Zimmerman (Ed.), Imperial College Lectures in Petroleum Engineering, 5th Edition, *World Scientific Publishing Company,* New Jersey, 2018, pp. 1–201.
doi: https://doi.org/10.1142/q0146

[4] M. Oostrom, Y. Mehmani et al., Pore-Scale and Continuum Simulations of Solute Transport Micromodel benchmark experiments, *Computational Geoscience* 20 (2014) pp. 857–897.
doi: https://doi.org/10.1007/s10596-014-9424-0

[5] A. Fick, About Diffusion, *Annals of Physics* 170 (1855) pp. 59–86, in German.
doi: https://doi.org/10.1002/andp.18551700105

[6] J. Fu, S. Cui et al., Statistical Characterization and Reconstruction of Heterogeneous Microstructures Using Deep Neural-Network, *Computer Methods in Applied Mechanics & Engineering* 373 (2021) pp. 1–38.
doi: https://doi.org/10.1016/j.cma.2020.113516

[7] M. Matrecano, Porous Media Characterization by Micro-Tomographic Image Processing, *Ph.D. thesis*, Colorado School of Mines (2014).
doi: http://dx.doi.org/10.6092/UNINA/FEDOA/8518

[8] G. Pawar, Modelling, and Simulation of the Pore-Scale Multiphase Fluid Transport in Shale Reservoirs: A Molecular Dynamics Simulation Approach, *Ph.D. thesis*, The University of Utah (2018).
URL: https://collections.lib.utah.edu/ark:/87278/s6rv3x25

[9] H. Xu, F. Usseglio-Virett et al., Microstructure Reconstruction of Battery Polymer Separators by Fusing 2D and 3D Image Data for Transport Property Analysis, *Journal of Power Sources* 480 (2020) pp. 1–9.
doi: https://doi.org/10.1016/j.jpowsour.2020.229101

[10] N. Abdussamie, Navier-Stokes Solutions for Flow and Transport in Realistic Porous Media, in: COMSOL (Ed.), *Proceedings of the COMSOL Conference*, Boston, 2010, pp. 1–5.
URL: https://www.comsol.com/paper/download/101193/abdussamie_paper.pdf

[11] Q. Sheng, Pore-to-Continuum Multiscale Modeling of Two-Phase Flow through Porous Media, *Ph.D. thesis*, Louisiana State University – LSU (2013).
URL: https://core.ac.uk/download/pdf/217390152.pdf

[12] J. Alvarez, G. Saudino et al., 3D Analysis of Ordered Porous Polymeric Particles using Complementary Electron Microscopy Methods, *Scientific Reports* 9 (2019) pp. 1–10.
doi: https://doi.org/10.1038/s41598-019-50338-2

[13] J. Wanek, C. Papageorgopoulou, F. Rühli, Fundamentals of Paleoimaging Techniques: Bridging the Gap Between Physicists and Paleopathologists, in: L. A. Grauer (Ed.), A Companion to Paleopathology, 1st Edition, *Blackwell Publishing Ltd*, Hoboken, 2011, pp. 1–43.
doi: https://doi.org/10.1002/9781444345940.ch18

[14] S. Wieghold, L. Nienhaus, Probing Semiconductor Properties with Optical Scanning Tunneling Microscopy, *Joule* 4 (2020) pp. 1–15.
doi: https://doi.org/10.1016/j.joule.2020.02.003

[15] X. Yin, Pore-Scale Mechanisms of Two-Phase Flow through Porous Materials – Volume of Fluid Method and Pore-Network Modeling, *Ph.D. thesis*, Utrecht University Repository (2018).
URL:
https://dspace.library.uu.nl/bitstream/1874/361289/1/Yin.pdf

[16] I. Kozma, I. Zsoldos, CT-based tests and finite element simulation for failure analysis of syntactic foams, *Engineering Failure Analysis* 104 (2019) pp. 371-378
doi: https://doi.org/10.1016/j.engfailanal.2019.06.003

[17] I. Kozma, I. Zsoldos, G. Dorogi, S. Papp, Computer tomography based reconstruction of metal matrix syntactic foams, *Periodica Polytechnica Mechanical Engineering* 58 (2), 2014, pp. 87-91.
doi: https://doi.org/10.3311/PPme.7337

[18] I. Kozma, I. Fekete, I. Zsoldos, Failure Analysis of Aluminum – Ceramic Composites, *Materials Science Forum* 885, 2017, pp. 286–291.
doi: https://doi.org/10.4028/www.scientific.net/MSF.885.286

[19] M. Shams, Modelling Two-phase Flow at the Micro-Scale Using a Volume-of-Fluid Method, *Ph.D. thesis*, Imperial College London (2018).
doi: https://doi.org/10.25560/62652

[20] A. T. Vuong, A Computational Approach to Coupled Poroelastic Media Problems, *Ph.D. thesis*, Technische Universität München – TUM (2016).
URL: http://mediatum.ub.tum.de/?id=1341399

[21] K. Wang, B. Xu, Current Status and Perspectives, in: X. Guo (Ed.), Molecular-Scale Electronics, 1st Edition, *Springer Press*, Cham, 2019, pp. 1–43.
doi: https://doi.org/10.1007/978-3-030-03305-7

[22] P. Kowalczky, A. P. Gauden, et al., Atomic-Scale Molecular Models of Oxidized Activated Carbon Fibre Nanoregions: Examining the Effects of Oxygen Functionalities on Wet Formaldehyde Adsorption, *Carbon* 165 (2020) pp. 67–81.
doi: https://doi.org/10.1016/j.carbon.2020.04.025

[23] T. Zhu, Unsteady Porous-Media Flows, *Ph.D. thesis*, Technische Universität München – TUM (2017).
URL: https://mediatum.ub.tum.de/doc/1279870/1279870.pdf

[24] J. Fu, R. H. Thomas, C. Li, Tortuosity of Porous Media: Image Analysis and Physical Simulation, *Earth-Science Reviews* 212 (2021) pp. 1–98.
doi: https://doi.org/10.1016/j.earscirev.2020.103439

[25] V. D. Chapman, H. Du et al., Optical Super-Resolution Microscopy in Polymer Science, *Progress in Polymer Science* 111 (2020) pp. 1–71.
doi: https://doi.org/10.1016/j.progpolymsci.2020.101312

[26] E. Widiatmoko, M. Abdullah, K. Khair, A Method to Measure Pore Size Distribution of Porous Materials Using Scanning Electron Microscopy Images, *American Institute of Physics (AIP) Conference Proceedings* 1284 (2010) pp. 23–27.
doi: https://doi.org/10.1063/1.3515554

[27] A. Borel, A. Ollé et al., Scanning Electron and Optical Light Microscopy: Two Complementary Approaches for the Understanding and Interpretation of Usewear and Residues on Stone Tools, *Journal of Archaeological Science* 48 (2014) pp. 46–59.
doi: https://doi.org/10.1016/j.jas.2013.06.031

[28] G. Zou, J. She et al., Two-Dimensional SEM Image-Based Analysis of Coal Porosity and its Pore Structure, *International Journal of Coal Science & Technology* 7 (2020) pp. 350– 361.
doi: https://doi.org/10.1007/s40789-020-00301-8

[29] C. C. Moura, A. Miranda et al., Correlative Fluorescence and Atomic Force Microscopy to Advance the Bio-Physical Characterisation of Co-Culture of Living Cells, *Biochemical and Biophysical Research Communications* 529 (2020) pp. 392–397.
doi: https://doi.org/10.1016/j.bbrc.2020.06.037

[30] S. M. Shah, J. P. Crawshaw, E. S. Boek, Three-Dimensional Imaging of Porous Media Using Confocal Laser Scanning Microscopy, *Journal of Microscopy* 265 (2016) pp. 1–11.
doi: https://doi.org/10.1111/jmi.12496

[31] T. Antequera, D. Caballera et al., Evaluation of Fresh Meat Quality by Hyperspectral Imaging (HSI), Nuclear Magnetic Resonance (NMR) and Magnetic Resonance Imaging (MRI): A Review, *Meat Science* 172 (2021) pp. 1–12.
doi: https://doi.org/10.1016/j.meatsci.2020.108340

[32] V. K. Gerke, V. E. Korostilev et al., Going Submicron in the Precise Analysis of Soil Structure: A FIBSEM Imaging Study at Nanoscale, *Geoderma* 383 (2021) pp. 1–12.
doi: https://doi.org/10.1016/j.geoderma.2020.114739

[33] C. Kizilyaprak, J. Daraspe, B. M. Humbel, Focused Ion Beam Scanning Electron Microscopy in Biology, *Journal of Microscopy* 254 (2014) pp. 109–114.
doi: https://doi.org/10.1111/jmi.12127

[34] C. E. Muir, V. O. Petrov et al., Measuring Miscible Fluid Displacement in Porous Media with Magnetic Resonance Imaging, *Water Resources Research* 50 (2014) pp. 1859–1868.
doi: https://doi.org/10.1002/2013WR013534

[35] J. M. Noel, J. F. Lemineur, Optical Microscopy to Study Single Nanoparticles Electrochemistry: From Reaction to Motion, *Current Opinion*

in Electrochemistry 25 (2021) pp. 1–13.
doi: https://doi.org/10.1016/j.coelec.2020.100647

[36] A. M. Parades, MICROSCOPY Scanning Electron Microscopy, in: A. B. Carl, L. T. Mary (Eds.), Encyclopedia of Food Microbiology, 2nd Edition, *Academic Press*, New York, 2014, pp. 693–701.
doi: https://doi.org/10.1016/B978-0-12-384730-0.00215-9

[37] M. Röding, C. Fager et al., Three-Dimensional Reconstruction of Porous Polymer Films from FIB-SEM Nanotomography Data Using Random Forests, *Journal of Microscopy* 281 (2021) pp. 76–86.
doi: https://doi.org/10.1111/jmi.12950

[38] K. D. Veron-Parry, Scanning Electron Microscopy: An Introduction, *III-Vs Review* 13 (4) (2000) pp. 40–44.
doi: https://doi.org/10.1016/S0961-1290(00)80006-X

[39] H. Zhang, J. Huang et al., Atomic Force Microscopy for Two-Dimensional Materials: A Tutorial Review, *Optics Communications* 406 (2018) pp. 3–17.
doi: https://doi.org/10.1016/j.optcom.2017.05.015

[40] S. Yesilkir-Baydar, N. O. Oztel et al., Evaluation Techniques, in: M. Razavi, A. Thakor (Eds.), Nanobiomaterials Science, Development and Evaluation, 1st Edition, *Woodhead Publishing*, New York, 2017, pp. 211–232.
doi: https://doi.org/10.1016/B978-0-08-100963-5.00011-2

[41] D. den Boer, A. A. W. J. Elemans, Triggering chemical reactions by Scanning Tunneling Microscopy: From atoms to polymers, *European Polymer Journal* 83 (2016) pp. 390-406.
doi: https://doi.org/10.1016/j.eurpolymj.2016.03.002

[42] C. M. M. Rodrigues, M. Militzer, Application of the Rolling Ball Algorithm to Measure Phase Volume Fraction from Backscattered Electron Images, *Materials Characterization* 163 (2020) pp. 1–7.
doi: https://doi.org/10.1016/j.matchar.2020.110273

[43] Y. Hashimoto, S. Takeuchi et al., Voltage contrast imaging with energy filtered signal in a field-emission scanning electron microscope,

Ultramicroscopy 209 (2020) pp. 1-22.
doi: https://doi.org/10.1016/j.ultramic.2019.112889

[44] T. Kanemaru, K. Hirata et al., A fluorescence scanning electron microscope, *Materials today* 12 (1) (2010) pp 18-23.
doi: https://doi.org/10.1016/S1369-7021(10)70141-3

[45] W. Chrzanowski, F. Dehghani, Standardised Chemical Analysis and Testing of Biomaterials, in: V. Salih (Ed.), Standardisation in Cell and Tissue Engineering, 1st Edition, *Woodhead Publishing*, New York, 2013, pp. 166–197.
doi: https://doi.org/10.1533/9780857098726.2.166

[46] T. Xu, J. I. Rodrigez-Devora et al., Bioprinting for Constructing Microvascular Systems for Organs, in: R. Narayan (Ed.), Rapid Prototyping of Biomaterials, 1st Edition, *Woodhead Publishing*, New York, 2014, pp. 201–220.
doi: https://doi.org/10.1533/9780857097217.201

[47] E. S. Statnik, A. I. Salimon, A. M. Korsunsky, On the application of digital optical microscopy in the study of materials structure and deformation, M*aterials today: PROCEEDINGS* 33 (2020) pp 1917-1923.
doi: https://doi.org/10.1016/j.matpr.2020.05.600

[48] Y. E. Bulbul, T. Uzunoglu et al., Investigation of nanomechanical and morphological properties of silane-modified halloysite clay nanotubes reinforced polycaprolactone bio-composite nanofibers by atomic force microscopy, *Polymer Testing* 92 (2020) pp. 1-11.
doi: https://doi.org/10.1016/j.polymertesting.2020.106877

[49] M. Potter, A. Li et al., Artificial Cells as a Novel Approach to Gene Therapy, in: S. Prakash (Ed.), Artificial Cells, Cell Engineering and Therapy, 2nd Edition, *Woodhead Publishing*, New York, 2007, pp. 236–291.
doi: https://doi.org/10.1533/9781845693077.3.236

[50] A. Canette, R. Briandet, MICROSCOPY Confocal Laser Scanning Microscopy, in: A. C. Batt, L. M. Tortorello (Eds.), Encyclopedia of Food Microbiology, 2nd Edition, *Academic Press*, New York, 2014, pp. 676–683.
doi: https://doi.org/10.1016/B978-0-12-384730-0.00214-7

[51] P. Prabhakaran, D. T. Kim, K. S. Lee, Polymer Photonics, in: K. Matyjaszewski, M. Möller (Eds.), Polymer Science: A Comprehensive Reference, 2nd Edition, *Elsevier Press*, New York, 2012, pp. 211–260.
doi: https://doi.org/10.1016/B978-0-444-53349-4.00207-7

[52] L. A. Trinh, E. S. Fraser, Chapter 21 – Imaging the Cell and Molecular Dynamics of Craniofacial Development – Challenges and New Opportunities in Imaging Developmental Tissue Patterning, in: Y. Chai (Ed.), Current Topics in Developmental Biology. 2nd Edition, *Elsevier Press*, New York, 2015, pp. 599–629.
doi: https://doi.org/10.1016/bs.ctdb.2015.09.002

[53] Z. Földes-Papp, U. Demel, G. P. Tilz, Laser scanning confocal fluorescence microscopy: an overview, *International Immunopharmacology* 3 (2003) pp. 1715-1729.
doi: https://doi.org/10.1016/S1567-5769(03)00140-1

[54] P. Mhaske, L. Condict et al., Phase volume quantification of agarose-ghee gels using 3D confocal laser scanning microscopy and blending law analysis: A comparison, *LWT* 129 (2020) pp. 1-9.
doi: https://doi.org/10.1016/j.lwt.2020.109567

[55] P. Parlanti, F. Brun et al., Size and Specimen-Dependent Strategy for X-Ray Micro-CT and TEM Correlative Analysis of Nervous System Samples, *Scientific Reports* 7 (2017) pp. 1–12.
doi: https://doi.org/10.1038/s41598-017-02998-1

[56] M. Utlaut, Focused ion beams for nano-machining and imaging, in: M. Feldman (Ed.), Nanolithography, 1st Edition, *Woodhead Publishing Limited*, Cambridge, 2014, pp. 116-157.
doi: https://doi.org/10.1533/9780857098757.116

[57] S. P. Kumar, G. K. Pavithra, M. Naushad, Characterization Techniques for Nanomaterials, in: S. Thomas, M. H. E. Sakho et al. (Eds.), Nanomaterials for Solar Cell Applications, 2nd Edition, *Elsevier Press*, New York, 2019, pp. 97–124.
doi: https://doi.org/10.1016/B978-0-12-813337-8.00004-7

[58] Z. L. Wang, J. L. Lee, Electron Microscopy Techniques for Imaging and Analysis of Nanoparticles, in: R. Kohli, K. L Mittal (Eds.), Development in

Surface Contamination and Cleaning, 2nd Edition, *Elsevier Inc.*, Amsterdam, 2016, pp. 395-443.
doi: https://doi.org/10.1016/B978-0-323-29960-2.00009-5

[59] H. Saka, Transmission Electron Microscopy, in: E. Yasuda, M. Inagaki et al. (Eds.), Carbon Alloys, 1st Edition, *Elsevier Ltd.*, Oxford, 2003, pp. 223-238.
doi: https://doi.org/10.1016/B978-008044163-4/50014-0

[60] Y. Gao, Micro-CT Imaging of Multiphase Flow at Steady State, *Ph.D. thesis*, Imperial College London – ICL (2019).
URL: https://doi.org/10.25560/76496

[61] A. Palmroth, S. Pitkanen et al., Evaluation of Scaffold Microstructure and Comparison of Cell Seeding Methods Using Micro-Computed Tomography-Based Tools, *Journal of the Royal Society Interface* 17 (2020) pp. 1–12.
doi: https://doi.org/10.1098/rsif.2020.0102

[62] Y. Guo, X. Chen et al., Analysis of foamed concrete pore structure of railway roadbed based on X-ray computed tomography, *Construction and Building Materials* online (2020) pp. 1-11.
doi: https://doi.org/10.1016/j.conbuildmat.2020.121773

[63] P. Goggin, E. M. L. Ho et al., Development of protocols for the first serial block-face scanning electron microscopy (SBF SEM) studies of bone tissue," *Bone* 131 (2020) pp. 1-47.
doi: https://doi.org/10.1016/j.bone.2019.115107

[64] M. J. Carcione, Wave Fields in Real Media, in: M. J. Carcione (Ed.), Wave Propagation in Anisotropic, Anelastic, Porous and Electromagnetic Medium, 3rd Edition, *Springer Press*, Berlin, 2015, pp. 560–690,
doi: https://doi.org/10.1016/C2013-0-18893-9

Acta Technica Jaurinensis

Vol. 14, No. 4, pp. 577-598, 2021
DOI: 10.14513/actatechjaur.00579

Physical Internet: A solution for the Supply Chain disruptions during the COVID-19 pandemic

B. Safwen*, P. Németh
Széchenyi István University, Department of Logistics and Forwarding, Egyetem tér 1, 9026 Győr, Hungary
***e-mail: saf.benneila@gmail.com**

Submitted: 29/11/2020; Accepted: 13/05/2021; Published online: 16/07/2021

Abstract: In this paper, we will specify our research on the impact of the COVID-19 on Supply Chain Management and show up the critical issues and how the digital solution like the Physical Internet (PI, π) could solve the disruptions of production or delivery of a Supply Chain. The Physical Internet is still an emerging phenomenon in which it is intended to replace the current logistics model by encompassing new technologies such as Artificial Intelligence, Big Data, the Internet of Things (IoT), and others. This article aims to compose conceptual research to describe Supply Chain Management problems during the COVID-19 pandemic and represent the Physical Internet as a solution for this disruption based on various journal articles, papers, websites, and managers' experiences. Further, this study helps increase the understanding of scholars and practitioners on how the novel PI paradigm can solve COVID-19-induced Supply Chain disruptions. Furthermore, the Physical Internet and other modern technologies in the business world are necessary and recommended, as these current issues now require quick decisions and up-to-date knowledge.

Keywords: Supply Chain Management; Physical Internet; COVID-19; Supply Chain disruption; pandemic

1. Introduction

Since 2000, numerous studies and analyses have been conducted on Supply Chain disruptions and have shown great potential. However, researchers and experts like in [1], [2], and [3] have concentrated on disturbances of the Supply Chain, usually defined as disruptive events that interrupt the structure of the Supply Chain.

According to [4], the Supply Chain could be affected by various factors such as natural disasters, civil disputes, financial crises, crime, or breakdown in transport infrastructure. Moreover, the Supply Chain was interrupted by COVID-19 [5]. COVID–19 was declared a global pandemic by the World Health Organisation [6] in March 2020 and described it in terms of infectious coronavirus diseases. According to the most recent OECD Economic Outlook, the COVID-19 pandemic has led to significant health, employment, and welfare harm [7]. Practically every country in the world was contaminated with COVID-19. This generation has encountered nothing that can be compared to this epidemic.

During their stages, the Supply Chain has faced significant disruption due to this pandemic and the resulting global crisis for health, especially in the range of connectivity, knowledge, and data sharing between its participants [8]. Upstream, though hedging and panic buying triggered equivalent downstream volatility, supply and demand equilibrium was further impacted and disrupted in many countries around the world by the travel restrictions and closures. The latest short-term pandemic predictions and also their effect on the Supply Chain have proved to be critical for management and strategy due to these disruptions. It is crucial to focus on tactical and strategic planning to increase customer satisfaction and avoid trouble in the Supply Chain. Precise predictions of emerging conditions help handle excess demand across the Supply Chain more effectively. A common understanding and current reality suggest that the acceleration and growth of COVID-19 in countries are inducing reforms in the immediate real needs and behavior of clients (healthcare [9] and food [10]). These types of changes put enormous pressure on the Supply Chains. For example, the entire Supply Chain, including eggs, flour, and grains, is affected when customers start to worry about purchasing dry noodles. However, a pandemic has some unique effects on Supply Chains. Compared to natural events, the pandemic is not confined or limited to a single area. Sequentially or simultaneously, the various Supply Chain elements are affected: production, supply centers, consumers, and logistics may be interrupted by time overlap windows. A new notion of the Physical Internet was suggested by Benoit Montreuil in 2006 [11] as an instrument to improve the world's importance regarding sustainable logistics. It is designed to interconnect worldwide logistics networks through temporary and intermediate storage shipping and delivery, as the Digital Internet has done with computer networks that allow data transmission in packets in standardized formats through heterogeneous equipment. In the discipline of logistics, these systems already exist because every company possesses its network, but they are far from getting interconnected; the idea of the Physical Internet is to make these systems compatible and much more open. The following three questions guide the research objectives of this study:

What are the SCM problems during this pandemic?

What are the advantages of using the Physical Internet in the COVID-19 Supply Chain?

What is the capability of the Physical Internet that could solve these issues in SCM during COVID-19?

1. Methodology

Our conceptual research is based on various journal articles, papers, websites, and managers' experiences. Currently, the topic is still in its infancy, ruling out the opportunity for a systematic review of the literature. The study's main objective is to highlight the Physical Internet's critical role in solving disruptions in the Supply Chain. First, we will present the different impacts of COVID-19, such as socioeconomic, macroeconomic, demand, and supply impacts. In the other part, we will explain the SCM challenges in resolving these disruptions by referring to transparency, estimation of stocks, improvement of customer and employee relations, estimation of risks, and cash management and networking, and how these methods are not enough. In the next section of this study, we will interpret the Physical Internet's ability to resolve SCM disruptions using an intelligent container that offers online access and identifications and environmentally friendly, that complies with international standards. It can be used in various transport modes—finishing our article with a summarized conclusion.

2. Impact of COVID-19 on SCM

"The effect of the COVID-19 initially identified in Wuhan, Hubei Province, China, in late 2019" [12] is illustrated in this section. But the effects of this virus on SCM have already been brought to the attention of academics (Choi 2020 [13]; Govindan et al. 2020 [8]; Lin et al. 2020 [14]; Sarkis et al. 2020 [15]), as well as industrial specialists (Business Insider 2020 [16]; Deloitte 2020 [17]; Fortune 2020 [18]).

2.1. Socioeconomic impacts

"The COVID-19 pandemic is far more than a health crisis: it is affecting societies and economies at their core" [19]. COVID-19 has a widespread human impact with unknown duration and unknown effects. There is a long-term public fear, and risks are invisible. Furthermore, there is a significant drop in international trade. Coronavirus causes a $47B drop in world exports [20]. Moreover, the overall scope impacts both goods and services. On the one hand, the movement of commodities is likely to be significantly interrupted. That could be because there was an insufficient workforce to transport the goods physically. The government acts such as port limits

on the handling of goods could theoretically impact them. On the other hand, for companies delivering services, workers' inability to reach and leave the regions concerned may be an essential obstacle to the supply of resources [21]. Nonetheless, workforce capacity has been reduced globally. Also, travel was restricted due to long breaks and extended quarantine times or work permit difficulties. "Without urgent socio-eco-nomic responses, global suffering will escalate, jeopardizing lives and livelihoods for years to come. Immediate development responses in this crisis must be undertaken with an eye to the future. Development trajectories in the long-term will be affected by the choices countries make now and the support they receive." [19].

2.2. Macroeconomic impacts

Due to COVID-19, according to Stephan [22], a recession likelihood in global GDP growth or has likely already begun. However, there is a wide range of economic impact timelines and various estimates about the recovery from V to U to other potential shapes over different time horizons. Therefore, there are very uncertain expectations. That means a heavy attachment to the global financial system and a heavy reliance on cooperative government interventions to sustain economies. The macroeconomic influence on world food production and food safety has already been significantly impacted by the COVID 19 Pandemic [23]. That should be considered by politicians when adopting steps to counter the full effects of the latest pandemic. According to Arslan, Drehmann, & Hofmann [24], the dissemination mechanisms of the COVID-19 shock into the economy, the interplay of policy decisions, epidemics, and political consensus need to be better understood.

2.3. Demand and Supply impacts

COVID-19 has a potential impact on consumer demand. We realize that specific product categories for personal protection, hygiene, medical supplies, and even some necessary foods are out of stock in some countries during this pandemic. This phenomenon is known as stockpiling. However, Tinglong Dai, from the Johns Hopkins University Carey business strategy and operations management and business analytics specialist, describes stockpiling in 2020 as "a spike in demand in the short term, and an immediate effect is you start seeing many empty shelves"[25]. Additionally, online shopping has increased due to social distancing, self-quarantining, and consumer avoidance of physical stores and crowded meeting places.

One of the most significant impacts of COVID-19 is many manufacturers' production delays or stoppages due to labor shortages because of quarantine and some cities' lockdown. That's why we have lacked many products. According to

[26], responding to the COVID-19 pandemic, the European Commission prioritizes citizens' welfare. That implies that necessary items such as emergency equipment and other essential goods are carried out. From a logistical point of view, the labor shortage caused congestion at airports and seaports. However, alerts about limited export loads and additional delays in port operations due to the lack of shipping operators and staff in the port have been released from the Clecat European Freight Forwarding Association and the Polish Chamber of Forwarding and Logistics [27]. As a result of the coronavirus, the eastern countries understand that the European economic situation is growing and improving.

COVID-19 is the crucial reason behind the Demand and Supply shock [28]. But other reasons make the situation even worst, such as the lack of horizontal collaboration, the inadequate response to technology trends, and inaccurate planning and forecasting [29]. Organizations should be ready for natural or environmental occurrences.

3. SCM challenge during COVID-19

This section describes SCM's various challenges during COVID-19, such as transparency, inventory estimation, relationships with customers and employees, risk estimation, cash managing, and networking.

3.1. Transparency

Keeping the Supply Chain sustainable needs transparency, which "involves communicating with key stakeholders about the firms' current activities and incorporates stakeholder feedback for Supply Chain improvement" [30]. We could create transparency in Supply Chains at multiple levels by listing critical components, determining the supply's origin, and identifying alternative sources [31]. Knowing the lineage of suppliers is very important in our current situation as it can detect whether or not the suppliers are in severely affected regions. And to go through transparency, Nancy Jalbert, management consulting at CPA [32], notes that it is essential to develop good relationships with vendors by cash flow management so that orders can be paid for on arrival. Second, decide who the leading suppliers are and how they are affected by the situation. Third, identify and notify replacement suppliers as soon as possible, if necessary. Finally, agree with the payment terms of the manufacturers.

3.2. Inventory Estimation

The COVID-19 pandemic has taught different lessons for all facets of the Supply Chain. However, during the COVID-19 pandemic, organizations that don't have

good inventory management processes and don't have access to accurate numbers will likely struggle to meet demand in specific categories or have oversupply articles that are suddenly less of interest to customers. That's why the accuracy and reliability of an organization's inventory levels have become extremely important. Moreover, companies must assess the stock available along the Supply Chain to provide replacement parts and components to be used as a bridge for continuing production and consumer distribution. Furthermore, to better track inventory levels, Nancy Jalbert [32] offers some great solutions to follow, for example, calculating sales variation to allow for efficient estimating of demand; identifying critical inventory levels and constructing inventories where necessary; designing quarantine areas for specific products and utilizing detailed list and refill roles of management systems.

3.3. Strengthen relationships with customers and employees

The pandemic of COVID-19 pressured corporations to preserve and establish connections with customers as their environment was changed. This global crisis is really about customer moments that matter. By putting your customers' interests first, maybe now is the time for your company's brand to take leadership. [17] suggests some ideas to help companies manage their customers in the current situation. Besides, businesses need to be brand loyal, as customers will see how their enterprises fare in such difficult times. Suppliers should strengthen their relationship with customers through communication, especially in a difficult situation. Customers will appreciate the supplier's assistance and gain an audience from a simple gesture. Not only with customers but also with employees, especially those in direct communication with your clients, has a significant effect on your business. Therefore, they are representative of your products, and they should be helpful and take care of hygiene responsibility to reduce the transmission of the virus. Moreover, to ensure employees' safety, it is essential to optimize the production and distribution capacity by working from home. It's also necessary to provide protective equipment to the workers and educate them on the virus's risks. These measures would help managers take into account the existing and expected capacity levels of personnel and machines.

3.4. Risk estimation

The estimation of the risk is a step to avoid them. According to The Global Fund [33], there are many risks to be taken into account, specifically related to logistics like capacity estimation, a summary of high-risk items or product types, and a description of existing buffer inventory and locations. Also, conduct scenario preparation to consider the consequences of Supply Chain disruption. Therefore, collect the appropriate demand for some months to create the necessary supply. At the same time, we must avoid these risks by exploring options to develop alternative

products, suppliers, and additional sourcing options. Moreover, change the mode of transport and take advantage of other existing logistics capabilities to reduce replenishment times. Finally, strengthen the market management mechanism to correct the excessive demand to minimize the bullfighting impact.

3.5. Manage cash and networking

COVID-19 has a crucial impact on the management of cash and network capital. However, according to [31], companies can perform stress testing to consider the economic effect of Supply Chain problems by planning for processes such as challenging key assumptions, identifying rapid mitigation actions, and developing an action plan for building simulation with different scenarios to understand short-term needs. With the recession unfolding, small production channels, sluggish deliveries, and shrunk profits contribute to the strain of earnings and liquidity even further. Companies are used to planning ambitious projects, and a great deal of realism is now required to unlock resources. Businesses would have to test their capital requirements weekly and monthly with the complete internal forecasting capability available.

4. The need for Physical Internet

The World Economic Forum WEF [34] highlighted the need to reengineer and adapt SCs to its potential trade problems. Therefore, we suggest Physical Internet as a workable solution to the logistical disruption of the COVID-19 pandemic. PI is an innovative concept of logistics network interconnection, which can be considered an extension of horizontal collaboration. The Physical Internet depends on the global interconnection of logistics. This concept aims to create an open and interconnected global logistics system that is economically, environmentally, and socially efficient and sustainable.

At this stage, we present a systematic literature review. First of all, we frame research questions that are mentioned in the introduction. In the second step, we identify the most relevant literature, such as journal articles, papers, websites, that mention the words "Physical Internet" and "COVID-19" in their titles or keywords. After that, we analysed this literature and arranged our summarised knowledge on this topic. In the last step, we provide valuable recommendations for organizations about the various benefits and solutions of the Physical Internet to face Supply Chain challenges during COVID-19.

4.1. Physical Internet: Definition and concept

Benoit Montreuil was the first researcher who introduced Physical Internet terms on 15th June 2006 and was inspired by an Economist magazine article title [35]. According to Benoit Montreuil, Physical Internet is "the way in which physical objects are handled, moved, stored, produced, delivered, and used, with a focus on global logistical efficiency and sustainability" [36]. Notice that physical goods' movement close to the data flow can be structured in the Digital Internet. Between the Physical Internet and the Digital Internet, there is an analogical similarity [36]. However, on Digital Internet, there is a flow of information from a host computer to another host computer. And the Physical Internet is a vision related to managing the physical flow in the logistic process from an origin point to the desired destination via a set of systems and procedures. This metaphor has generated interest within the university and has attracted substantial business coverage.

The novel term Physical Internet is grabbing the attention of a lot of practitioners and academics. Montreuil defines PI as "a global logistical system based on the interconnection of logistical networks through a standardized set of collaboration protocols, modular containers, and smart interfaces for increased efficiency and sustainability." [37]. However, the Physical Internet concept presented as a solution and response to the Global Logistics Sustainability Grand Challenge [36] founded on physical, digital, and operational interconnectivity, consisting of PI-hubs run by encapsulation, interfaces, and protocols through various services.

The PI is based on standard containers easily transported by different PI hubs through various transport means (planes, trucks, and cars) by several operators. Hence, this logistical structure can be shared with all customers, including manufacturers, suppliers, and consignees, including complete horizontal and vertical logistical cooperation through uniform interconnections, standard encapsulation, standard intelligent interfaces, and standard coordination protocols to optimize the loading and unloading. Open hubs and open logistics material are part of the network, enabling a global Logistics Web. The Physical Internet aims to maximize the transport of goods in terms of costs and speed, creating an efficient and sustainable logistics network similar to the way data packages flow in the Digital Internet in a reliable and resilient manner.

4.2. Physical Internet as a solution during COVID-19

The Physical Internet uses and extends to logistics the ideas of the digital Internet. PI is a network of more flexible, agile, and resilient logistics networks globally, open and interconnected. PI uses many collaborative protocols and standard innovative interfaces to transmit and receive physical products [36]. These goods are stored in standardized modules to increase global logistics efficiency and sustainability [38].

The purpose of the Physical Internet is to make intelligent containers accessible online and exploit them. The omnipresent use of PI containers allows any logistics services provider to manage and store the goods of any organization while it does not process and store its items. PI aims to maximize the capability of intelligent PI-containers connecting to the Internet of Things and their integrated smart objects to enhance customer efficiency and the overall performance of logistics networks (Fig. 1.) [39].

The Physical Internet may solve several Supply Chain Management issues stemming from the COVID-19 pandemic. During this time, the planet experienced significant disruptions in transportation and distribution. The PI-nodes allow us to ensure that the intelligent containers are loaded, unloaded, composed, decomposed, sent, moved, and traced quickly. Note that PI-nodes should generally be linked to the logistics activities to encourage cooperation and knowledge sharing among various players and companies. The run-on Blockchain technology can anonymously and transparently transfer data between operators. The proper order of these transfers may be confirmed by the possession of the moved items in the whole shipping process [40]. The higher point density in a PI system enables multiple tours at the same service level to be shortened and means of transportation to be improved [41]. Moreover, the π-nodes provide higher-scale economies. However, the bigger the distribution network's size, the more participating parties are involved [42].

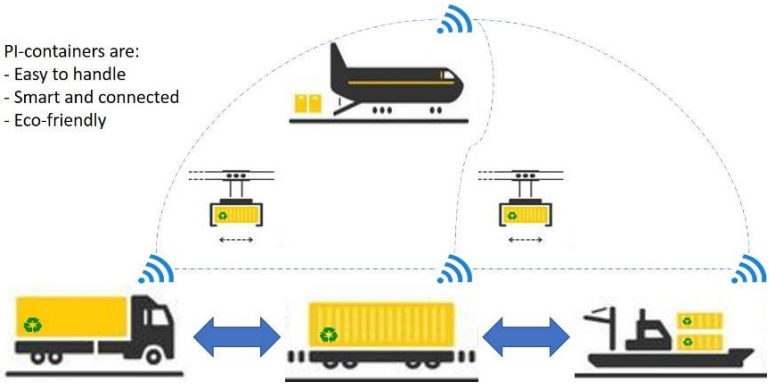


Figure 1. An integrated transport system through Physical Internet containers

As a result, the transport network can be faster and cheaper by consolidating shipments due to the cross-docking of π-hubs (Fig. 2.). Hence, the loan consolidation of goods from different parties is faster, easier, and more efficient [43]. The processes between procurement and shipment can be long and complex, but transportation is where a business comes into direct contact with its customers across

all other operations. The distribution point represents the image of the entire organization. If a business delivers its goods late, the client takes a very unfavourable opinion of this company and therefore may not re-use the services.

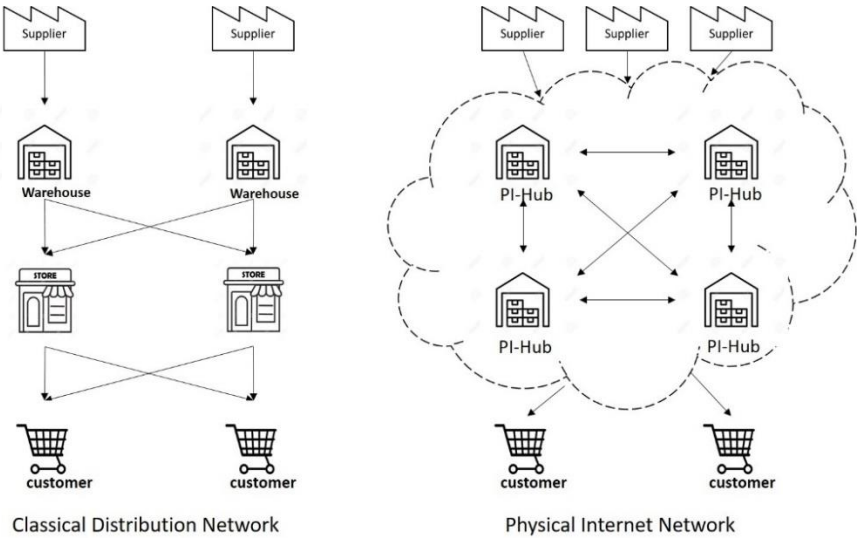


Figure 2. Difference between Classical Distribution Network and Physical Internet Network

The COVID-19 pandemic has also created significant stock problems worldwide, which led us to minimize an average overall expense, decrease an average stock, and decrease remaining service levels using a PI-inventory management model [44]. Moreover, it offers more options for repositioning and replenishment between hubs and also more opportunities for multi-sourcing. Furthermore, it provides different ways to adjust inventory positions rapidly (e.g., in case of demand variations) [45]. Effective transport maintains a seamless operation of the Supply Chain. Stock can be held lean by efficient operation of the vehicle and can be carried rapidly in and out of a store. This increases warehouse turnover, decreases lead time, and saves storage capital.

The inventory has a direct impact on the manufacturing system. However, in the PI manufacturing system, all resources are interconnected, providing better control, focus on processes, "quick response, balanced production, highly effective throughput, low consumption, and scientific decision-making in a manufacturing system" [46]. After all, "inventory models using PI respond to SC facility disruptions and their resilience." [47]

We can't discuss the transportation and inventory process without referring to π-containers in which they have a potential effect on the transport and the storage of goods. Hence, similar to the Digital Internet packets, there are containers created for the Physical Internet called π-containers. They have physical content and structure rather than being purely informational. From a functional perspective, the main objective of π-containers is to facilitate their handling, storage, transport, sealing, locking, loading, unloading, construction, and dissociation. These PI-containers are standardized containers with the global norm. The sizes, duties, and fittings are particularly modularized and standardized worldwide [48]. Besides, unit loads are constructed from modular containers loaded into trucks to optimize size and weight [49]. From a marketing perspective, π-containers should be easy to panel with publicity, and precise product information supports for business marketing. The enterprise can advertise in big and flashing image sizes without confronting any constraints by using PI-container as a billboard. Businesses can reach a large audience with an economical marketing step that can generate a permanent impact and achieve great results.

From an environmental perspective, π-containers must have a limited footprint when away from service, dismantle, and assemble on demand. They should be as environmentally friendly as possible, in range with toughness principles [50]. The efficient utilization of PI-containers will reduce space and transport many goods with lower fuel consumption. Also, PI-containers have a longer life cycle than pallets and cartoons and can be reused and recycled.

From an informational perspective, each container is attached with a smart tag to act as the agent that it represents. It supports the coverage, protection, routing, tracking, traceability, and security of containers across the Physical Internet [39]. PI provides real-time visibility using radio frequency identification and leads to better decision making, increased quality, and efficiency through advanced information sharing and models. Also, PI offers closed-loop visibility, methods, development, and expense tracking capability. [51]. Visibility in the PI gives the possibility to access some reliable information about the movement of PI-containers in the logistics network in real-time, leading to better analysis and improvement of Supply Chain processes.

For collecting this big data, we run to Cloud computing usage and sharing. Nonetheless, in this decade, the cloud has driven innovation for many organizations and companies. New, fast, private, safe, and efficient, all of these characteristics distinguish the cloud from traditional data treatment methods. The PI is a supporter of cloud development and storage to ensure the growth and sharing or support from virtually present specialists [52]. Besides, the cloud can provide the ability to pool resource use, save time and costs, reduce risks, and enhance interaction promptly,

and contribute to economic and environmental sustainability and socially by offering the passive capacity to other participants in an open logistics network. Among the many incentives for using the cloud, the cost of shipping and distribution through warehouse PI hubs closer to end-users is reduced [53].

Additionally, by applying the Physical Internet worldwide, we will have intelligent city logistics that offer operators a variety of easily understandable information, facilitating the execution of the urban freight transport plans. However, logistics and transit services can be used (public and private) for developed metropolitan areas. And these cities become smart logistics cities by reducing transports fuel consumption and CO2 emissions in cities and reducing traffic jams [54].

To sum up, "Physical Internet is a solution for the logistics disruptions during the COVID-19 pandemic" [55] (Table 1 and 2). However, "Physical Internet is the logistics of the future" [56].

Table 1. The SCM challenges and the main technological opportunities of PI Part 1.

Supply Chain challenges during COVID-19	How can the Physical Internet overcome these challenges?	Benefits	Technologies
Transparency	The integration systems and collaborative mechanisms of the PI will lead to improving transparency along with the Supply Chain network [57].	Competitive advantage Data transparency and safety Predictability Agility Flexibility	Blockchain [40] Cloud computing [53]
Inventory estimation	The real-time traceability of PI-containers will improve the efficiency of the inventory estimation [58]	Real-time identification Better service to consumers Protection from demand fluctuations Shorter lead times	Blockchain [40] RFID [58] Bare code [58] Cloud computing [53]

Table 2. The SCM challenges and the main technological opportunities of PI Part 2.

Supply Chain challenges during COVID-19	How can the Physical Internet overcome these challenges?	Benefits	Technologies
Strengthen relationships between consumers and employees	The availability of the goods in the markets under challenging events will increase trust between customers and organizations [59], and technology will also facilitate consumers' purchases in a secure manner [60]. The use of technology will facilitate the work from home, and employees will feel safer and more comfortable [61]	Trust Loyalty Safety Low the risk of getting infected	Internet of Things [60] Artificial Intelligence & machine learning [61]
Risk estimation	The adoption of the Internet of Things will guide the managers to select the appropriate risk estimation technique [62]	Reduce future uncertainties Training and enhancement Improve decision making Proper results forecast	Internet of Things [62]
Manage cash and networking	For effective cash flow control, it's essential to use the new generation of technologies to solve recent problems [63]	Develop better cash-flow forecasts Reduce the risk of loss Faster access to cash and data Increase operational efficiency Cost savings	ERP software Blockchain Artificial Intelligence [63]

5. Conclusion

After the COVID-19 crisis started, shipping and logistics companies were led by sustaining shelves and vital supply lines. However, the year began with global demand expansion already behind existing trends, and now the sector faces

worldwide disruption. No corporation was resistant to the recession, with a significant impact on the transportation and logistics sector. The Supply Chain is disrupted, and executives should make rapid and operational decisions to secure and help their workforce and guarantee vital company processes to fulfill society's current and long-term needs.

The Supply Chain network can benefit enormously from digital technologies in terms of transparency, visibility, cost savings, reliability, and resiliency. The stakeholders of the PI network, such as suppliers, consumers, and logistics service providers, must implement and improve emerging technology to enable the whole network to be agile with continuous delivery, which has a significant effect on reducing inventory risk mitigation. We can ensure a digital future through the Physical Internet to serve individuals and organizations' logistics needs. However, providing an online and maintained connection enables the physical movements of goods to be transported to other locations. Therefore, it avoids shortages and conflicts of the whole content system thanks to an increased and timely flow of information—moreover, shorter lead times due to PI interconnections and increases digitalization's effect on inventory management. Mainly, data coordination and PI visibility qualities are enabled by Blockchain, IoT, RFID in the simulation, and activation of the reorganization plan.

Consequently, businesses become data-driven organizations aiming to maximize data importance and uses it as an essential strategic instrument for forecasting and reacting to future events. Not only to anticipate future events but also to create a new organizational model, processes, and update technical capabilities to enhance flexibility, achieve speed, adaptiveness, and stability.

References

[1] Craighead, C. W., Blackhurst, J. et al., The severity of supply chain disruptions: design characteristics and mitigation capabilities. *Decision Sciences* 38 (1) (2007) pp. 131-156.
doi: https://doi.org/10.1111/j.1540-5915.2007.00151.x

[2] Ivanov, D., Das, A., Coronavirus (COVID-19/SARS-CoV-2) and supply chain resilience: A research note. *International Journal of Integrated Supply Management* 13 (1) (2020) pp. 90-102.
doi: https://doi.org/10.1504/IJISM.2020.107780

[3] Queiroz, M. M., Ivanov, D. et al., Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a

structured literature review. *Annals of Operations Research* (2020) pp. 1-38. doi: https://doi.org/10.1007/s10479-020-03685-7

[4] Gurning, S., Cahoon, S., Analysis of multi-mitigation scenarios on maritime disruptions. *Maritime Policy & Management* 38 (3) (2011) pp. 251-268. doi: https://doi.org/10.1080/03088839.2011.572701

[5] Nikolopoulos, K., Punia, S. et al., Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *European journal of operational research* 290 (1) (2020) pp. 99-115. doi: https://doi.org/10.1016/j.ejor.2020.08.001

[6] The world health organization [cited 2020-03-11] URL https://www.who.int/

[7] OECD. Food Supply Chains and COVID-19: Impacts and Policy Lessons. Comparing Crises: Great Lockdown versus Great Recession, pp. 1-11. [cited 2020-06-02] URL http://www.oecd.org/coronavirus/policy-responses/food-supply-chains-and-covid-19-impacts-and-policy-lessons-71b57aea/

[8] Govindan, K., Mina, H., Alavi, B., A decision support system for demand management in healthcare supply chains considering the epidemic outbreaks: A case study of coronavirus disease 2019 (COVID-19). *Transportation Research Part E: Logistics and Transportation Review* 138 (2020) 101967. doi: https://doi.org/10.1016/j.tre.2020.101967

[9] UNAIDS, The impact of the COVID-19 response on the supply chain, availability and cost of generic antiretroviral medicines for HIV in low- and middle-income countries (1) (2020). URL https://www.unaids.org/sites/default/files/media_asset/covid19-supply-chain-availability-cost-generic-arv_en.pdf

[10] Singh, S., Kumar, R. et al., Impact of COVID-19 on logistics systems and disruptions in food supply chain. *International Journal of Production Research* (2020) pp. 1-16. doi: https://doi.org/10.1080/00207543.2020.1792000

[11] Montreuil, B., Meller, R. D., Ballot, E., Physical internet foundations. *IFAC Proceedings Volumes* 45 (6) (2012) pp. 26-30.
doi: *https://doi.org/10.3182/20120523-3-RO-2023.00444*

[12] Ian Twinn, Qureshi, N. et al., The Impact of COVID-19 on Logistics. International Finance Corporation (IFC), (2020-06-02) pp. 1–5.
URL *https://www.ifc.org/wps/wcm/connect/2d6ec419-41df-46c9-8b7b-96384cd36ab3/IFC-Covid19-Logistics-final_web.pdf?MOD=AJPERES&CVID=naqOED5*

[13] Choi, T. M., Innovative "bring-service-near-your-home" operations under Corona-virus (COVID-19/SARS-CoV-2) outbreak: Can logistics become the messiah?. *Transportation Research Part E: Logistics and Transportation Review* 140 (2020) 101961.
doi: *https://doi.org/10.1016/j.tre.2020.101961*

[14] Lin, Q., Zhao, S. et al., A conceptual model for the outbreak of Coronavirus disease 2019 (COVID-19) in Wuhan, China with individual reaction and governmental action. *International Journal of Infectious Diseases* 93 (2020) pp. 211-216.
doi: *https://doi.org/10.1016/j.ijid.2020.02.058.*

[15] Sarkis, J., Cohen, M. J. et al., A brave new world: lessons from the COVID-19 pandemic for transitioning to sustainable supply and production. *Resources, Conservation, and Recycling* 159 (2020) 104894.
doi: *https://doi.org/10.1016/j.resconrec.2020.104894.*

[16] Schmalz, F., The coronavirus outbreak is disrupting supply chains around the world—Here's how companies can adjust and prepare. Business Insider (2020-03-30).
URL *https://www.businessinsider.com/covid-19-disrupting-global-supply-chains-how-companies-can-react-2020-3*

[17] Main, A., Stephen, A. et al., COVID-19 Maintaining customer loyalty and trust during times of uncertainty. Deloitte (2020-03-30).
URL
https://www2.deloitte.com/content/dam/Deloitte/ie/Documents/covid19/gx-coronavirus-customer-loyalty.pdf

[18] Sherman, E., 94% of the Fortune 1000 are seeing coronavirus supply chain disruptions: Report. Fortune (2020-03-30).

URL *https://fortune.com/2020/02/21/fortune-1000-coronavirus-china-supply-chain-impact/*

[19] United Nations, A UN framework for the immediate socioeconomic response to COVID-19 (2020-04-19).
URL *https://unsdg.un.org/sites/default/files/2020-04/UN-Framework-for-the-immediate-socio-economic-response-to-COVID-19.pdf*

[20] FOX Business, 2020 Coronavirus causes $47B drop in world exports: UN (2020-03-04)
URL *https://www.foxbusiness.com/markets/coronavirus-causes-47-billion-drop-in-world-exports-un*

[21] Francis, C., Managing the impact of the coronavirus on supply chains. Pinsent Masons (2020-03-13).
URL *https://www.pinsentmasons.com/out-law/analysis/coronavirus-supply-chains*

[22] Gay, S. H., Adenäuer, M., Frezal, C., Potential macroeconomic impact of the Covid-19 pandemic on food demand. *EuroChoices* 19 (3) (2020) pp. 40-41.
doi: *https://doi.org/10.1111/1746-692X.12271*

[23] Guan, D., Wang, D. et al., Global supply-chain effects of COVID-19 control measures. *Nature Human Behaviour* 4 (2020) pp. 577-587.
URL *https://doi.org/10.1038/s41562-020-0896-8*

[24] Arslan, A., Drehmann, M., Hofmann, B., Central bank bond purchases in emerging market economies BIS Bulletin. (20) (2020) p. 9.
URL *https://www.bis.org/publ/bisbull20.pdf*

[25] Tim Parsons, How Coronavirus will affect the global supply chain (2020-03-06).
URL *https://hub.jhu.edu/2020/03/06/covid-19-coronavirus-impacts-global-supply-chain/*

[26] The European Union ec.europa.eu (2020-09-13).
URL *https://ec.europa.eu/transport/coronavirus-response_en*

[27] Teleroute. Effects of the coronavirus COVID-19 on the transport sector (2020-03-12).
URL *https://teleroute.com/en-en/blog/article/effects-of-the-coronavirus-covid-19-on-the-transport-sector/*

[28] del Rio-Chanona, R. M., Mealy, P. et al., Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective. *Oxford Review of Economic Policy* 36 (Supplement_1) (2020) S94-S137.
doi: *https://doi.org/10.1093/oxrep/graa033*

[29] Ostdick, N., 5 Primary Causes of Supply Chain Disruptions (2017-08-17)
URL *https://blog.flexis.com/5-primary-causes-of-supply-chain-disruptions*

[30] Duan, Y., Aloysius, J. A., Supply chain transparency and willingness-to-pay for refurbished products. *The International Journal of Logistics Management* 30 (3) (2019) pp. 797-820.
doi: *https://doi.org/10.1108/IJLM-01-2019-0025*

[31] Alicke, K., Azcue, X., Barriball, E., Open interactive popup Supply-chain recovery in coronavirus times—plan for now and the future. McKinsey & Company (2020-03-18).
URL *https://www.mckinsey.com/business-functions/operations/our-insights/supply-chain-recovery-in-coronavirus-times-plan-for-now-and-the-future*

[32] Jalbert, N., Coronavirus: Effective supply chain management during a crisis. Raymond Chabot Grant Thornton (2020-03-24).
URL *https://www.rcgt.com/en/insights/coronavirus-effective-supply-chain-management-during-crisis/*

[33] The Global Fund, COVID-19 Impact on Supply Chain Logistics: Assessment and Recommendations (2020).
URL *https://www.theglobalfund.org/media/9440/psm_covid-19impactonsupplychainlogistics_report_en.pdf*

[34] Betti, F., Ni, J., How China can rebuild global supply chain resilience after COVID-19. World Economic Forum (2020-04-05).
URL *https://www.weforum.org/agenda/2020/03/coronavirus-and-global-supply-chains/*

[35] The economist, The physical Internet (2006-06-17).
URL https://www.economist.com/special-report/2006/06/17/the-physical-internet

[36] Montreuil, B., Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logistics Research* 3 (2-3) (2011) pp. 71-87.
doi: https://doi.org/10.1007/s12159-011-0045-x

[37] Ballot, E., Montreuil, B., Meller, D.R., "The network of logistics networks", Paris, France: Direction de L'information Legale et Administrative (2014) p. 23.
URL
http://www.ladocumentationfrancaise.fr/catalogue/9782110098658/index.shtm

[38] Montreuil, B., Ballot, E., Fontane, F., An open logistics interconnection model for the Physical Internet. *IFAC Proceedings Volumes* 45 (6) (2012) pp. 327-332.
doi: https://doi.org/10.3182/20120523-3-RO-2023.00385

[39] Sallez, Y., Pan, S. et al., On the activeness of intelligent Physical Internet containers. *Computers in Industry* 81 (2016) pp. 96-104.
doi: https://doi.org/10.1016/j.compind.2015.12.006

[40] Meyer, T., Kuhn, M., Hartmann, E., Blockchain technology enabling the Physical Internet: A synergetic application framework. *Computers & Industrial Engineering* 136 (2019) pp. 5-17.
doi: https://doi.org/10.1016/j.cie.2019.07.006

[41] Sarraj, R., Ballot, E. et al., Analogies between Internet network and logistics service networks: challenges involved in the interconnection. *Journal of Intelligent Manufacturing* 25 (6) (2014) pp. 1207-1219.
doi: https://doi.org/10.1007/s10845-012-0697-7

[42] Pan, S., Nigrelli, M. et al., Performance assessment of distributed inventory in Physical Internet. Proceedings of International Conference on Computers and Industrial Engineering, CIE (2013) pp. 178–192.
URL https://hal.archives-ouvertes.fr/hal-00876280

[43] Venkatadri, U., Krishna, K. S., Ülkü, M. A., On Physical Internet logistics: modeling the impact of consolidation on transportation and inventory costs.

IEEE Transactions on Automation Science and Engineering 13 (4) (2016) pp. 1517-1527.
doi: https://doi.org/10.1109/TASE.2016.2590823

[44] Yang, Y., Pan, S., Ballot, E., A model to take advantage of Physical Internet for vendor inventory management. *IFAC-PapersOnLine* 48 (3) (2015) pp. 1990-1995.
doi: https://doi.org/10.1016/j.ifacol.2015.06.380

[45] Pan, S., Nigrelli, M. et al., Perspectives of inventory control models in the Physical Internet: A simulation study. *Computers & Industrial Engineering* 84 (2015) pp. 122-132.
doi:https://doi.org/10.1016/j.cie.2014.11.027

[46] Wang, J. Q., Fan, G. Q. et al., Research on initiative scheduling mode for a physical internet-based manufacturing system. *The International Journal of Advanced Manufacturing Technology* 84 (1-4) (2016) pp. 47-58.
doi: https://doi.org/10.1007/s00170-015-7915-3

[47] Manopiniwes, W., Irohara, T., Stochastic optimisation model for integrated decisions on relief supply chains: preparedness for disaster response. *International Journal of Production Research* 55 (4) (2017) pp. 979-996.
doi: https://doi.org/10.1080/00207543.2016.1223379

[48] Gontara, S., Boufaied, A., Korbaa, O., Routing the PI-Containers in the Physical Internet using the PI-BGP Protocol. In 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA). IEEE Xplore (2018) pp. 1-8.
doi: https://doi.org/10.1109/AICCSA.2018.8612885

[49] Lin, Y. H., Meller, R. D. et al., A decomposition-based approach for the selection of standardized modular containers. *International Journal of Production Research 52* (15) (2014) pp. 4660-4672.
doi: https://doi.org/10.1080/00207543.2014.883468

[50] Zhang, Y., Liu, S. et al. Smart box-enabled product–service system for cloud logistics. *International Journal of Production Research* 54 (22) (2016) pp. 6693-6706.
doi: https://doi.org/10.1080/00207543.2015.1134840

[51] Zhong, R. Y., Peng, Y. et al., Towards Physical Internet-enabled prefabricated housing construction in Hong Kong. *IFAC-PapersOnLine* 48 (3) (2015) pp. 1079-1086.
doi: https://doi.org/10.1016/j.ifacol.2015.06.227

[52] Montreuil, B., Meller, R. D., Ballot, E., Physical Internet Foundations Bureaux de Montréal : Bureaux de Québec. Cirrelt, 10 (2012).
URL https://www.cirrelt.ca/DocumentsTravail/CIRRELT-FSA-2012-56.pdf

[53] Darvish, M., Larrain, H., Coelho, L.C., A dynamic multi-plant lot-sizing and distribution problem. *International Journal of Production Research* 54 (22) (2016) pp. 6707-6717.
doi: https://doi.org/10.1080/00207543.2016.1154623

[54] Crainic, T. G., Montreuil, B., Physical internet enabled hyperconnected city logistics. *Transportation Research Procedia* 12 (Supplement C) (2016) pp. 383-398.
doi: https://doi.org/10.1016/j.trpro.2016.02.074

[55] Youssef A., Physical Internet: A solution for the logistics disruptions during the COVID-19 pandemic (2020-08-20).
URL https://medium.com/@youssefag99b/physical-internet-a-solution-for-the-logistics-disruptions-during-the-covid-19-pandemic-2ee49d3757ae

[56] Matusiewicz, M. Logistics of the Future—Physical Internet and Its Practicality. *Transportation Journal* 59 (2) (2020) pp. 200-214.
doi: https://doi.org/10.5325/transportationj.59.2.0200

[57] Plasch, M., Pfoser, S. et al., Why Collaborate in a Physical Internet Network? - Motives and Success Factors. *Journal of Business Logistics* 42 (1) (2021) pp. 120-143.
doi: https://doi.org/10.1111/jbl.12260

[58] Herrera, M. M., Orjuela-Castro, J., An Appraisal of Traceability Systems for Food Supply Chains in Colombia. *International Journal on Food System Dynamics* 12 (1) (2021) pp. 37-50.
doi: http://dx.doi.org/10.18461/ijfsd.v12i1.74

[59] Jeżewska-Zychowicz, M., Plichta, M., Królak, M., Consumers' Fears Regarding Food Availability and Purchasing Behaviors during the COVID-19 Pandemic: The Importance of Trust and Perceived Stress. *Nutrients* 12 (9) (2020) 2852.
doi: https://doi.org/10.3390/nu12092852

[60] Ho-Sam-Sooi, N., Pieters, W., Kroesen, M., Investigating the effect of security and privacy on IoT device purchase behaviour. *Computers & Security* 102 (2021) 102132.
doi: https://doi.org/10.1016/j.cose.2020.102132

[61] Dubey, A. D., Tripathi, S., Analysing the sentiments towards work-from-home experience during covid-19 pandemic. *Journal of Innovation Management* 8 (1) (2020) pp. 13-19.
doi: https://doi.org/10.24840/2183-0606_008.001_0003

[62] Atlam, H., Alenezi, A. et al., An overview of risk estimation techniques in risk-based access control for the internet of things. In Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security – IoTBDS (2017) pp. 254-260.
doi: https://doi.org/10.5220/0006292602540260

[63] NatWest, Tech solutions for managing working capital (2019-05-23).
URL https://natwestbusinesshub.com/articles/tech-solutions-for-managing-working-capital

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 599-611, 2021
DOI: 10.14513/actatechjaur.00601

A literature review on the conflict analysis of vehicle-pedestrian interactions

A. Kizawi[1,*], A. Borsos[1]

[1] Széchenyi István University, Department of Transport Infrastructure and Water Resources Engineering
Egyetem tér 1., 9026 Győr, Hungary
*e-mail: ahmad.kizawi.24@gmail.com

Submitted: 05/03/2021; Accepted: 23/03/2021; Published online: 08/04/2021

Abstract: An alternative to traffic safety analysis based on historical crash data the use of non-crash events is becoming more popular thanks to the rapid improvement in video-based vehicle trajectory processing. By means of Surrogate Measures of Safety (SMoS) in traffic conflict studies, the most critical elements on the road network can be identified and the probability of accidents can be proactively determined. This paper aims to summarize the state-of-the-art research regarding the analysis of pedestrian-vehicle interactions at unsignalized crossings, to synthetize the previous studies using Surrogate Measures of Safety (SMoS), and to identify the research gaps.

Keywords: *traffic conflict; Surrogate Measures of Safety (SMoS); road safety; pedestrian*

1. Introduction

The consequences of road accidents represent a real problem worldwide. Despite of the implementation many different effective traffic measures in the EU, the number of road accidents remains unacceptably high. The paper by Fülep and Óberling [1] provided the causal factors of road accidents with the appropriate recommendations. Since the statistical database are easily available (from Hungary and Europe), accident issues can be recognized: approximately 1.3 million crashes annually lead to nearly 40,000 fatalities and more than 1 million injuries.

Several comprehensive studies in the European Union discuss road accidents from both analytical and statistical perspectives. The paper by Holló [26] provided an overview about road safety in the EU and Hungary including some statistical studies. Although there was no significant increase in the number of motor vehicles between 1987 and 1990, the number of road accident deaths in the EU increased significantly. However, from 2001 until 2009, overall the number of fatal victims of road accidents lessened by 36% in the EU member states thanks to the planned development by the European Union transport policy. [26]

The most sensible method for road safety assessment is scrutinizing the accidents using data records. However, this kind of methods has some restrictions especially for Vulnerable Road Users (VRUs) such as bicyclists and pedestrians. These limitations are related to underreporting problems, and missing accident data. From an ethical point of view, it makes no sense to wait for road accidents to occur to evaluate safety which we primarily want to avoid. [2]

Various researchers have developed a number of safety indicators aiming to overcome the shortcomings of safety analyses based on accidents (underreporting, quality issues and rare nature). By using Surrogate Measures of Safety (SMoS), the probability of accident at a location can be proactively determined before it would happen. There is a consensus among researches that observable non-crash situations can be useful for traffic safety assessment as a substitutional tool in parallel with analysis based on crash data [3], [4].

This paper aims to synthetize the previous studies using Surrogate Measures of Safety (SMoS) and to illustrate the most important indicators that can be used in traffic studies to evaluate the traffic conflicts seriousness.

2. Literature review

2.1. Pedestrian-vehicle interactions

Pedestrians and vehicles share the road in the complex internal traffic of the city resulting in interactions between vehicles and pedestrians affecting each other's movement. While vehicle trajectories are more predictable, there are several issues regarding pedestrian movements since pedestrians often tend to make more sudden decisions and more prone to hectically change their speed and trajectory. For pedestrian and vehicle interactions at intersections, the traffic conflict between the pedestrian and the vehicle is an important factor that influences some intersection characteristics such as capacity and safety. Conflicts between pedestrians and vehicles are growing in developing countries due to the ever-increasing traffic. The frequent interactions between vehicles and pedestrians deserve special interest to analyze safety at intersections. [5]

The priority of the pedestrians at the unsignalized crosswalks is not entirely clear. This leads to higher levels of accident of pedestrians compared to signalized crosswalks. Pedestrian and Vehicle Interactions (PVI) have been receiving more attention among governmental agencies and decision-makers aiming to improve road accident countermeasures that decrease the rate of accidents and severe conflicts [6].

The study by Dey and Terken [7] attempted to define the importance of communication between drivers and pedestrians. Their study was carried out at a city in the Netherlands, at an intersection where approximately a hundred observations about pedestrian actions were observed at the pedestrian crossing after video recordings and field measurements for pedestrian's movements. At the pedestrian crossing, five behaviors were detected from pedestrians while interacting with approaching vehicles. Most of pedestrians (61%) stepped on the pedestrian crossing without waiting for the driver's approval, while in some cases (20.7%), pedestrians acted by changing their path based on a vehicle's approach. In other cases (12.6%), pedestrians wait until an oncoming vehicle decelerates, and then they stepped on the pedestrian crossing. In 1.8% of the cases, a pedestrian tended to give the driver the priority for passing first. In 3.6% of the cases, pedestrians wait for an approaching car to completely stop, and get the driver's approval to cross.

2.2. Traffic conflict analysis

The first study to evolve a measure by which road accidents can be predicted was designed by Perkins and Harris [8]. Their study could be useful to gain a better insight into causal factors regarding traffic safety issues. In their study, they identified different potential accident situations, which they classified as traffic conflicts. They clarified more than twenty types of traffic conflicts that occur between road users, which were essentially identified by the occurrence of evasive actions, such as swerving, stopping, and braking.

In the terms of traffic conflicts based on evasive actions, the study by Johnsson et al. [9] described different surrogate measures of safety in the literature. Few indicators concentrated on aspects that can be useful when studying Vulnerable Road Users (VRUs) issues. Their paper evaluated different safety indicators depending on the capability of these indicators to take into account both injury risk and collision risk (talking the evasive actions into account). The findings in their paper showed that several indicators have focused on braking as an important indicator to define the critical traffic situations and do not consider other kinds of evasive actions such as running or swerving.

The supposition that there is a relationship between the seriousness and the frequency of traffic events is the basic idea behind the use of non-crash situations to

investigate road safety [10]. Hyden [11] clarified the so-called "safety pyramid" where the upper part represents the most severe events, which can be considered as the rarest events in traffic, that we generally define them as "accidents". Immediately below the accidents come traffic conflicts, which are grouped as severe, slight or potential conflicts according to their dangerousness. Below the conflicts come the majority of traffic encounters that are characterized as natural events. Fig. 1 shows the Safety pyramid with the severity levels of traffic events.

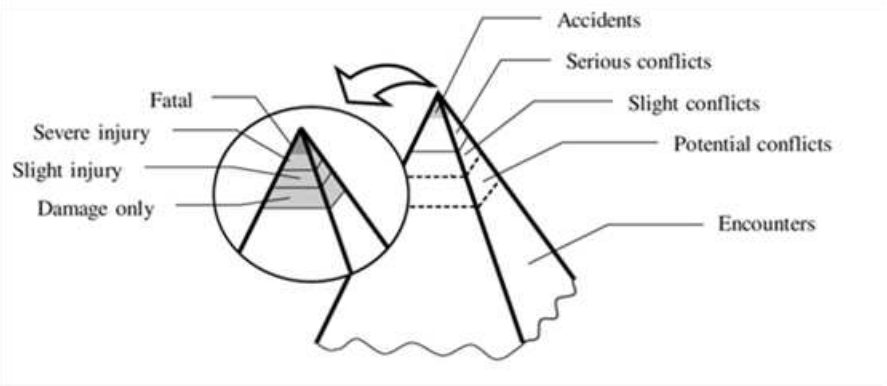


Figure 1. "Safety pyramid" [11]

2.3. Surrogate Measures of Safety (SMoS)

The term "surrogate" reveals that the indicators do not depend on crash database. Instead, they are meant to be complementary tools of historical records analysis. Several different indicators were proposed and developed by various researchers throughout the 1970s and the 1980s. Recently, several traffic safety indicators have been proposed and developed including those related to VRUs. Since there are many issues related to VRUs safety analysis such as underreporting problems of VRU crashes, there is a growing awareness in transport modes including VRUs [9].

Several papers have described, summarized and compared a group of safety indicators. For instance, Laureshyn et al. [4], [12] provided an overview about nearness to-collision and severity indicators. Zheng et al. [13] clarified the temporal and spatial proximity characteristics in traffic situations. An observed situation is considered a traffic conflict or not depending on closeness in distance and (or) time of the concerned road users.

Ceunynck [3] presented the previous studies on the application of safety indicators and looked into the frequency of use. He compiled indicators into groups using the Time-To-Collision (TTC), the Post Encroachment Time (PET), and the deceleration

families, plus two extra groups for other and unspecified indicators. According to the study by Ceunynck [3], indicators originating from the (TTC) family are frequently used, followed by those from the (PET) family. There are a few indicators that do not belong to these families and are classified as "other" type, as shown in Fig. 2.

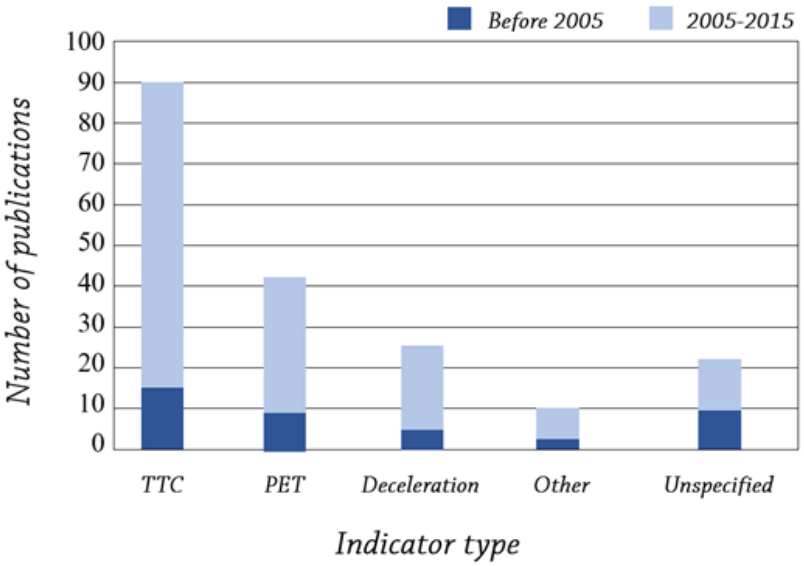


Figure 2. Frequency of using surrogate safety indicators [3]

The "unspecified" type belongs to papers that present incomplete details to determine the techniques that have been adopted.

The most commonly used Surrogate Measures of Safety (SMoSs) for pedestrian conflict analysis include but not limited to the indicators described in the following subsections. [14] These will be illustrated using figures showing vehicle-vehicle interactions; however, their calculation is identical to vehicle pedestrian interactions.

2.3.1. Time-To-Collision (TTC)

This indicator has been suggested by Hayward [15]. It is the time to collision in seconds when the two vehicles continue their trajectory at the same angle and at the same speeds without any kind of evasive behaviors. The minimum value is 0 second (which means collision). When TTC has a small value, there is a high risk of collision.

Fig. 3 illustrates TTC for right-angle collision. In this case, TTC is calculated by the following equation (1), (2).

$$TTC = \frac{d_2}{v_2}, if\ \frac{d_1}{v_1} < \frac{d_2}{v_2} < \frac{d_1+l_1+w_2}{v_1}, \quad (1)$$

$$TTC = \frac{d_1}{v_1}, if\ \frac{d_2}{v_2} < \frac{d_1}{v_1} < \frac{d_2+l_2+w_1}{v_2}, \quad (2)$$

Where v_1 and v_2 are vehicles speeds; d_1 and d_2 are distances from the front of the vehicle to the conflict area; l_1, l_2, and w_1, w_2 are lengths and widths of vehicles, respectively.

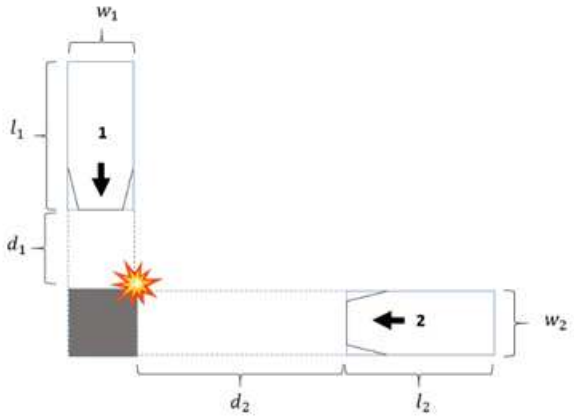


Figure 3. Illustration of TTC [17]

There are two important indicators that can be used based on TTC: Time to Accident (TA), which is the TTC at the moment in which an involved road user acts by making an evasive action, and TTC min, which is the minimum TTC value calculated in a conflict. For differentiating between a serious and a not serious event, these two indicators usually use a threshold value.

Jiang [16] conducted a study about the vehicle–pedestrian conflict by field observations, video recordings, and TTC calculations. According to the paper, 3.0 and 1.0 s can be recommended for urban areas as the average TTC and TTCmin, respectively.

2.3.2. Post Encroachment Time (PET)

It is the difference between times when a car enters a conflict point until another one arrives to this conflict point [17], [18]. When PET has a small value, this indicates a higher dangerous of vehicle- pedestrian collisions.

To measure PET in the case of pedestrian-vehicle conflict, we only need to measure the passing times at the conflict point for both conflicting road users, while to measure TTC, we need to determine the time remaining to the conflict point at each time instant. Fig. 4 displays the procedure for calculating PET.

$$\mathbf{PET = t2 - t1}$$

Figure 4. Illustration of PET [17]

Several papers [17] [19] [20] found that PET and TTC indicators are considered as the most precise indicators for road safety assessment at intersections due to the simplicity of mensuration, consistency over time, and its relationship to other indicators.

2.3.3. Delta V

This indicator reveals the velocity vector change experienced by the involved road user during an accident. It is very susceptible to the vulnerability of the road user, since a light body will bounce back when colliding with a massive one, while the speed of the massive body will not change. This is an important measure in accidents studies between, for example, a truck and a bicyclist. [3]

For an inelastic collision, Delta V values are calculated for each of the two conflicting road users using the following equation (3),(4). The highest value is used to describe severity. Fig. 5 illustrates the Delta V calculation based on the principle of momentum conservation.

$$\Delta v_1 = \frac{m_2}{m_1 + m_2} \cdot \sqrt{v_1^2 + v_2^2 - 2v_1 v_2 cos\alpha}, \quad (3)$$

$$\Delta v_2 = \frac{m_1}{m_1 + m_2} \cdot \sqrt{v_1^2 + v_2^2 - 2v_1 v_2 cos\alpha}, \quad (4)$$

Where: v_1,v_2 the speeds of the involved road users 1 and 2 respectively; m_1, m_2 their masses; α: the approach angle.

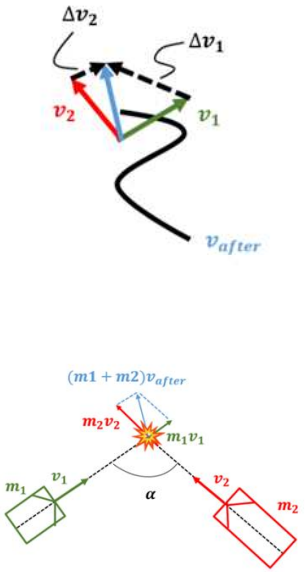


Figure 5. Calculation of Delta V [3]

2.4. Using extreme value theory (EVT) in traffic conflicts

Extreme Value Theory (EVT) is useful for estimating crash probabilities using Surrogate Measures of Safety (SMoS). To evaluate the benefits of active safety techniques, the Weibull type of generalized extreme value (GEV) distribution was applied to traffic events [21]. The EVT method was also used and validated in [2]. For estimating the frequency of right-angle collisions at intersections, the GEV distribution was further used and the findings concluded that there was a favorable relationship between the estimated accidents and detected accidents. Tarko [22] clarified the use of the generalized Pareto distribution (GPD) as a complementary approach of GEV distribution in the EVT method, especially for estimating the severity of crashes based on traffic conflicts.

Some studies [23] used the EVT method in order to evaluate the head-on conflicts associated with crossing maneuvers situations in rural areas on two-lane roadways. Other researches [24] adopted the EVT method for estimating the road departure accident recurrence based on the index of time to road edge passing. It was found plausible accident estimates compared to the recorded accident database.

A recent study [25] aimed to anticipate the near-accident situations that can happen at signalized intersections between vehicles and pedestrians. With TTC and PET indicators obtained from field measurements and video recording, EVT method was applied to fit the distributions of PET and TTCmin values and to extract the threshold values. A Gated Recurrent Unit (GRU) neural network was used to for anticipating the most severe situations after the sequent data was obtained from pedestrians' and vehicles' movements, with an accuracy of 0.878 and the AUC (Area Under the Curve) value of 0.865.The model presented in the study is considered as a useful model to alert drivers of the potential severe traffic situations with the pedestrians.

3. Research gap and methods

As for pedestrian-vehicle interactions, there are still some gaps related to the use of Surrogate Measures of Safety (SMoS) especially since the vulnerable road users (VRUs) tend to make sudden changes in their trajectories. The analysis of such interactions using Extreme Value Theory is also very scarce in the currently available literature.

The research related to the current paper focuses on analyzing pedestrian-vehicle interactions at unsignalized crossings using Surrogate Measures of Safety (SMoS) and Extreme Value Theory (EVT). An important characteristic of the EVT is that it enables the researcher to model the stochastic behavior of unusually large or small processes. This behavior is typically unobservable within a plausible data collection time period due to its rareness. [27] The assumption behind the EVT method is that

the underlying stochastic behavior of the process being modeled is sufficiently smooth to enable extrapolations to unobserved levels [28]. EVT involves estimating extreme events probability over an extended period of time given very short and limited historical data. [27]

According to Tarko et al. [29] The Extreme Value Method offers an important advantage over the traffic conflict technique that the risk of crash given the surrogate event is estimated for any conditions based on the observed variability of crash proximity without using crash data. The crash proximity measure precisely defines the surrogate event. An unsignalized intersection in the city of Győr, Hungary has been selected where several different pedestrian-vehicle interactions will be video recorded. Video recordings will be used to collect SMoS with the help of the T-Analyst Software. EVT approach will be used to determine the probability of crashes and the ability of SMoSs to assess the severity of pedestrian-vehicle interactions. After the in-depth analysis, recommendations will be suggested for improving the levels of safety.

4. Conclusions

In this paper, we gave an overview about using Surrogate Measures of Safety (SMoS) and the latest research results with regard to the analysis of pedestrian-vehicle conflicts. The most important indicators that can be used to evaluate the seriousness of such conflicts such as TTC, PET and Delta-V were introduced. Extreme Value Theory has been recently used by a few researchers to estimate crash probabilities using SMoS; the most relevant literature sources in relation to that were summarized. It is concluded that there is an existing research gap in the conflict analysis of pedestrian-vehicle conflicts mainly because of the less predictable behavior of pedestrian movements. The use of SMoSs coupled with EVT to such conflicts also offers research opportunities.

References

[1] T. Fülep, J. Óberling, Traffic Safety based on Accident Statistics Concerning Road Vehicles and Infrastructure, *Acta Technica Jaurinensis* 5 (3) (2012) pp. 197–205.
doi: https://doi.org/10.14513/actatechjaur.v12.n1.489

[2] P. Songchitruksa, A. P. Tarko., The extreme value theory approach to safety estimation, *Accident Analysis & Prevention* 38 (4) (2006) pp. 811–822.
doi: https://doi.org/10.1016/j.aap.2006.02.003

[3] D. Ceunynck, Defining and applying surrogate safety measures and behavioural indicators through site-based observations *Ph.D. thesis*, Lund University (2017). [cited 2021-03-05]
URL
https://portal.research.lu.se/portal/files/30184385/170823Dissertation_TimDeCeunynck_final_inclcover.pdf

[4] Laureshyn A., Svensson Å., Hydén, C., Evaluation of traffic safety, based on micro-level behavioral data: Theoretical framework and first implementation. *Accident Analysis & Prevention* 42 (6) (2010) pp. 1637–1646.
doi: *https://doi.org/10.1016/j.aap.2010.03.021*

[5] A. Kumar, M. Paul, I. Ghosh, Analysis of pedestrian conflict with right turning vehicles at signalized intersections in India, Proactive safety assessment and improvements at intersections, *Journal of Transportation Engineering Part (A) Systems* 145 (6) (2018) pp. pp. 04019018-01–04019018-12.
doi: *https://doi.org/10.1061/JTEPBS.0000239*

[6] H. Amado, S. Ferreira et al., Pedestrian-vehicle interaction at unsignalized crosswalks, A systematic review. *Sustainability* 12 (7) (2020) pp. 1–23.
doi: *https://doi.org/10.3390/su12072805*

[7] D. Dey, J. Terken, Pedestrian interaction with vehicles, Roles of explicit and implicit communication, in: *Automotive UI 2017 - 9th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, Proceedings*, Association for Computing Machinery, Oldenburg, 2017, pp. 109–113.
doi: *https://doi.org/10.1145/3122986.3123009*

[8] R. S. Perkins, J. I. Harris, Traffic conflict characteristics-Accident potential at intersections, *Highway Research Record* 225 (1968) pp. 35–43.

[9] C. Johnsson, A. Laureshyn, T. De Ceunynck, In search of surrogate safety indicators for vulnerable road users, a review of surrogate safety indicators, *Transport Reviews* 38 (6) (2018) pp. 765–785.
doi: *https://doi.org/10.1080/01441647.2018.1442888*

[10] H. C. Chin, S. T. Quek, Measurement of traffic conflicts, *Safety Science* 26 (3) (1997) pp. 169–185.

[11] C. Hydén, The development of a method for traffic safety evaluation, the Swedish traffic conflict technique, *Doctoral thesis*, Lund University, Department of Traffic Planning and Engineering (1987).

[12] A. Laureshyn, T. De Ceunynck et al., In search of the severity dimension of traffic events, Extended Delta-V as a traffic conflict indicator, *Accident Analysis & Prevention* 98 (2017) pp. 46–56.
doi: https://doi.org/10.1016/j.aap.2016.09.026

[13] L. Zheng, K. Ismail, X. Meng, Traffic conflict techniques for road safety analysis, Open questions and some insights, *Canadian Journal of Civil Engineering* 41 (7) (2014) pp. 633–641.
doi: https://doi.org/10.1139/cjce-2013-0558

[14] P. Chen, W. Zeng et al., Surrogate Safety Analysis of Pedestrian-Vehicle Conflict at Intersections Using Unmanned Aerial Vehicle Videos. *Journal of Advanced Transportation*, (2017).
doi: https://doi.org/10.1155/2017/5202150

[15] J. C. Hayward, Near-miss determination through use of a scale of danger, *Highway Research Record* 384 (1972) pp. 24-34.

[16] X. Jiang, Intercultural Analyses of Time-to Collision in Vehicle Pedestrian Conflict on an Urban Midblock Crosswalk, *IEEE transactions on intelligent transportation systems* 16 (2) (2015) pp. 1–6.
doi: https://doi.org/10.1109/tits.2014.2345555

[17] B. L. Allen, B. T. Shin, D. J. Cooper, Analysis of traffic conflicts and collision, *Transportation Research Record* 667(1978) pp. 67–74.

[18] P. J. Cooper, Experience with traffic conflicts in Canada with emphasis on "post encroachment time" techniques, In Proceedings of the *NATO Advanced Research Workshop on International Calibration Study of Traffic Conflict Technique* (1983).

[19] D. Gettman, L. Head, Surrogate Safety Measures from Traffic Simulation Models, Final Report, Publication No. FHWARD-03-050, Federal Highway Administration, Washington, DC, USA. (2003).

[20] D. Gettman, L. Sayed et al., Surrogate Safety Assessment Model and Validation," FHWA Report, Publication No.: FHWA-HRT-08-051, Federal Highway Administration, Washington, DC, USA. (2008).

[21] K. Campbell, H. C. Joksch, P. E. Green, A bridging analysis for estimating the benefis of active safety technologies, UMTRI-96-18, Final Report. University of Michigan, Transportation Research Institute (1996).

[22] A. P. Tarko, Use of crash surrogates and exceedance statistics to estimate road safety, *Accident Analysis & Prevention* 45 (2012) pp. 230–240. doi: https://doi.org/10.1016/j.aap.2011.07.008

[23] H. Farah, C. L. Azevdo, Safety analysis of passing maneuvers using extreme value theory, *IATSS Research* 41 (1) (2017) pp. 12–21. doi: https://doi.org/10.1016/j.iatssr.2016.07.001

[24] T. J. Gordon, A Multivariate Analysis of Crash and Naturalistic Driving Data in Relation to Highway Factors Report No. S2-S01C-RW-1. Transportation Research Board, Washington, DC (2013).

[25] S. Zhang, M. Abdel-Aty et al., Modeling pedestrians' near-accident events at signalized intersections using gated recurrent unit (GRU), *Accident Analysis & Prevention* 148 (2020) 105844. doi: *https://doi.org/10.1016/j.aap.2020.105844*

[26] P. Holló, Road Safety Situation of Hungary Reflected by National and International Objectives, *Acta Technica Jaurinensis* 4 (2) (2011) pp. 2–7.

[27] P. Songchitruksa, A. P. Tarko, The extreme value theory approach to safety estimation. *Accident Analysis and Prevention* (2006), 38:811–822.

[28] S. Coles, *An Introduction to Statistical Modeling of Extreme Values, (2001)*

[29] A. Tarko, G. Davis, N. Saunier, T. Sayed, S. Washington, White Paper SURROGATE MEASURES OF SAFETY ANB20(3) Subcommittee on Surrogate Measures of Safety ANB20 Committee on Safety Data Evaluation and Analysis (2009).

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 612-630, 2021
DOI: 10.14513/actatechjaur.00634

Asphalt layers within railway tracks' substructure

Al Qatamin Jehad[1,*], Al Amir Ibrahim[1]

[1] Széchenyi István University, Department of Transport Infrastructure and Water Resources Engineering
Egyetem tér 1., 9026 Győr, Hungary
*e-mail: Qataminjihad@yahoo.com

Submitted: 02/06/2021; Accepted: 04/11/2021; Published online: 24/11/2021

Abstract: For both heavy cargo Rail lines and High-Speed Lines, the railway industry continues to develop railway track design technologies. Within the past years, using an asphalt layer as a part of the railway track became so common as a support layer, it helps to reduce noise and vibration levels, reduce the thickness of the cross-sectional layers, reduce the lifetime maintenance, and many other great advantages. This paper will discuss the application of asphalt layers in railways, functionality, application fields, and types of asphalted trackbeds. Design requirements and parameters of asphalt layers, the mix design, installation process, and international experiences will be mentioned. Asphalt layers can improve the overall performance of the trackbeds and their behavior under repeated stresses of trains, the effect of asphalt layers on stress and strain ratios will be briefly discussed.

Keywords: asphalt underlayment, asphalt mix design, stress and strain ratios, asphalt trackbeds performance.

1. Introduction

Ancient Railway track designs simply consisted of two parallel rails fixed on wooden sleepers or ties, as in figure 1, those sleepers were just put on the natural ground. by the time people recognized that they should improve the performance of that kind of trackbed so they thought that it could be more efficient if they underlay and surround the cross ties or sleepers with a natural stone aggregate (ballast), they noticed that this stone aggregate layer restrained the horizontal and vertical

movement and reduced the displacement as well. And that was the first ballasted track beds.

Figure 1. The old Railway of Alhijaz 1902 [11]

With the evolution and the development of trains and the need for heavier loads and more speeds, also the need for higher numbers of trains, railways had to be more efficient with larger rail profiles and more ties, to provide good load distribution and to reduce the effect of these loads on the subgrade materials, a sub-ballast granular layer was provided with a specific thickness.

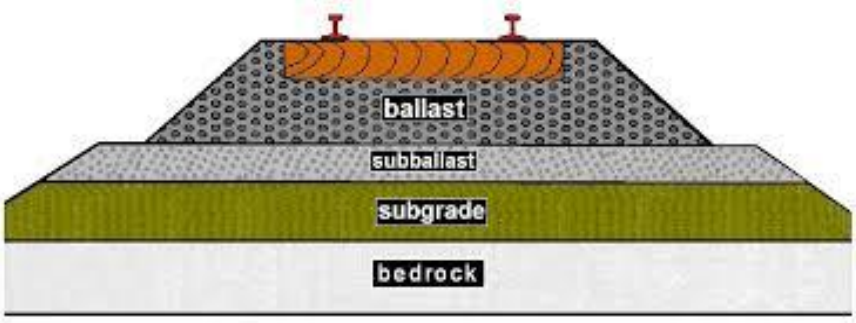


Figure 2. Typical all-granular trackbeds cross-section. [1].

This layer has a smaller maximum aggregate diameter than the ballast, and it has a well-graded granular composition. This layer acted as a waterproof layer due to its low void ratio and low permeability [1]; so, it drives the water away from the track. As a result, better trackbeds with mainly all granular layers have better performance and more bearing capacity. This type of trackbed requires a drainage system to prevent water from seeping down into the sub-ballast and the subgrade to avoid the undesirable effect of water on those layers. Figure 2 shows the typical all-granular trackbeds cross-section.

Later on, the trackbeds need to be more stable and more load distribution efficient, and hence it reduces subgrade stresses and prevents softening and deformation of the subgrade layers. So, the asphalt underlayment-as showed in figure 3- appeared to be the perfect solution. This asphalt layer is quite similar in Mix Design to the one used in highway construction. [2].

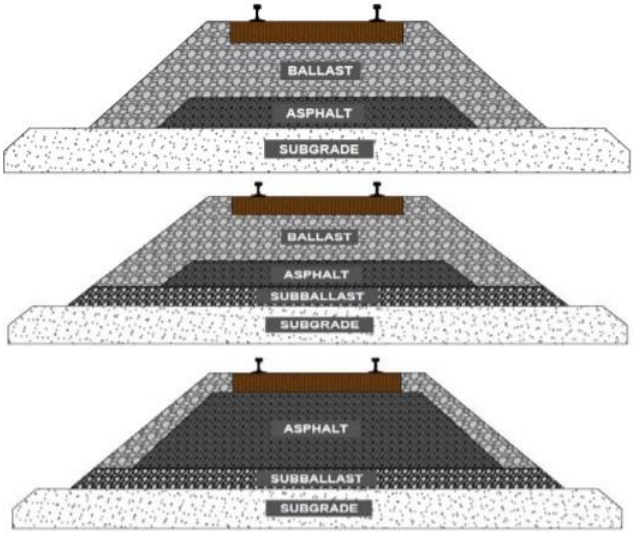


Figure 3. Asphalt underlayment [3]

Asphalt

Asphalt is mainly a mixture of aggregate and bitumen. See figure 4. Bitumen is the resulting material of the distillation of crude oil. Aggregate in asphalt should be well-graded and combines particle sizes from fine to coarse aggregate, the properties of asphalt mixture can differ in accordance to the construction

requirements; these properties can be controlled by varying the proportions of the mixture's compositions, it could be rigid and stable or flexible.

Some projects require exact specifications for the asphalt; such as heavy-duty and lower temperatures; this could be done by adding some special additives or using polymer modified bitumen.

Figure 4. Core samples show the composition materials of asphalt.[12]

Asphalt pavement should meet the basic requirements in terms of use:

1- Strength: asphalt pavement should be strong enough to resist the stresses coming from the loads above to avoid cracking and failure.

2- Durability: durability could be affected by many factors such as temperature, humidity, excess loads, and the performance of each component.

3- Stability: stability is the term that describes the consistency of performance level due to the change of the affecting factors, such as temperature and the existence of water.

2. The functionality of Asphalt in railways

Because of the accelerated growth in traffic volumes, train numbers, and heavy loadings, it was necessary to customize the railways and provide better characteristics of railways that make them more supportive, stronger, and longer-lasting. Because of its good properties, hot mixed asphalt has to provide trackbeds with the desired features that are needed in railway structures. Asphalt to be added as a layer to be a medium-modulus, flexible, low-voids, fatigue-resistant layer that can withstand heavy tensile strains without cracking. [10]. The addition of an asphalt layer to the trackbeds -see figure 5- can increase the bearing capacity and protect the

substructure layers beneath besides its perfect separation function, [9] it also can do a great job in terms of separation; it prevents the fine-graded materials from the sub-ballast or subgrade to seep up to the ballast. The asphalt layer to be used as a supplementary layer within the railway trackbed could play a major role as a load distributor layer which will consequently reduce the settlement while the trackbed is on loading [14]. Thanks to the asphalt waterproofing ability to replace the use of additional geomembrane. The functionality of asphalt in railways could be counted as follows:

1- A strong-supportive layer below the ballast distributes the loads uniformly to the substructures.

2- A waterproofing layer prevents the water seepage and keeps the necessary consistency of load-bearing capacity in the trackbed structure.

3- Ease the water's movement toward the side ditches as it is an impermeable layer with a low void ratio.

4- Prevent the fine-graded materials from pumping up and mix with the Ballast.

5- A Resilient layer tends to reduce noise and vibration.

6- Longer-lasting trackbeds and lower maintenance costs.

Figure 5. Asphalt layer under ballast [3]

3. Asphalt applications in railways

The properties of asphalt offer an ideal solution for many problems and applications in railways construction; it improves both the stability and durability of the track beds and consequently reduces the maintenance needs; it could also be a perfect solution when a lesser construction height of superstructure is a must especially in tunnels and bridges.

3.1 Types of asphalt trackbeds

Asphalt could be involved in the construction design of tracked within two basic patterns; either hot mix asphalt as a sub-ballast layer or hot mix asphalt overlayment layer designs could be adopted in trackbeds.

Asphalt as a sub-ballast layer could also be divided into two designs; the so-called "asphalt underlayment" and "asphalt combination" trackbeds, both of them consider the ballast layer as a part of the structural design. [1]. The ballast layer will play a crucial role as it offers a cover to protect the asphalt from the sunlight, air, and water, and it also keeps a stable environment and temperature above the asphalt, which leads to a longer asphalt lifetime. [4].

3.1.1 Asphalt underlayment

This type of trackbed has almost the same cross-section as the all-granular cross-section but the substitution of the sub-ballast layer with the asphalt layer. So, the asphalt layer is placed directly on the subgrade. The typical cross-section is shown in figure 6.

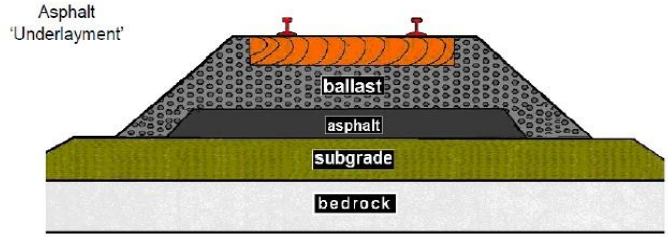


Figure 6. Asphalt underlayment trackbed cross-section without sub-ballast layer. [1]

3.1.2 Asphalt combination

The asphalt layer is an additional layer placed between the ballast, and the sub-ballast layer, the thickness of the asphalt layer is lesser than the one used in asphalt underlayment trackbeds; this is because of the existence of the granular sub-ballast layer below. Figure 7 shows this design.

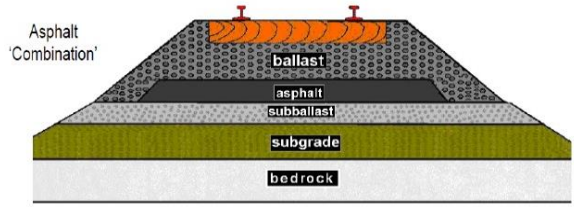


Figure 7. Asphalt combination trackbed cross-section with sub-ballast layer. [1]

3.1.3 Ballast-less trackbeds

This type of trackbed has the sleepers placed directly on the asphalt layer. So, there is no ballast layer placed on the top of the structure. The asphalt layer and the granular sub-ballast layer should be relatively thick to compensate for the deleted ballast layer, cribbing rock should be used to restrain sleepers from both lateral and longitudinal movement. as showed in figure 8.

Normal paving machines are being used for such trackbeds, and it is possible to reach a level of accuracy up to ±2 millimeters, resulting in a perfect level and stiff surface for the sleepers to be fixed over it, trackbed and sleepers are filled with elastic synthetic material or a granular ballast could be used between sleepers as an extra stability option. [2].

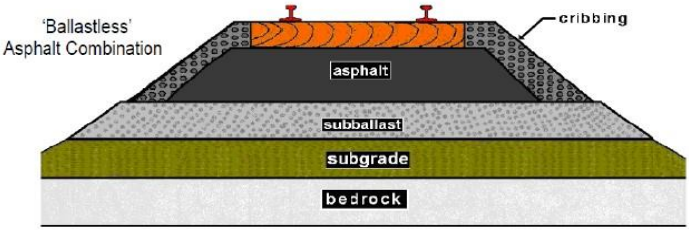


Figure 8. Asphalt ballast-less trackbeds cross-section. [1]

A horizontal anchoring system should be used for the rail track while constructing the ballast-less trackbed design; to prevent the transverse movement of rails. This kind of design allows for lowering the base of the trackbed as it eliminates the ballast layer so, this is an option to consider for tunnels.

The most important feature of this design is elasticity, particularly; when polymer-modified asphalt is being used. Also, it is considered a time-saving system because the construction could take place immediately after the asphalt cools down,

maintenance and small corrections could be done without demolishing and reconstruction. In the 1990s, the first successful ballastless track was built in Germany. Since then, this design has started to appear in many other countries. [5]

3.2 Application fields

Hot mix asphalt underlayment tended to be used where there is heavy rail traffic, and one or more of the following conditions exist.

1- When a strong, and durable substructure below ballast is difficult to be set up and maintained.

2- When a good drainage surface for water is difficult to be set up and maintained.

3- When lowering the groundwater table to prevent the weakening of the trackbed is difficult.

4- When stresses are expected to be relatively high. Such as joints and approaches of bridges and tunnels.

Moreover, Hot-Mix-Asphalt could be used for the maintenance and rehabilitation of the tracks as it provides in some cases a cheaper solution to maintain an acceptable operating performance in locations that have a high conventional maintenance cost. Railway ballast is a non-cohesive, granular crushed rock that has a poor capacity to bear vibrations caused by trains [17]. So, the existence of asphalt underlayment can improve the vibration absorption and will reduce the maintenance needs.

In some specific track works, such as crossings and turnouts, the trackbed tends to break down faster. So; more maintenance cycles are needed; The existence of the hot-mix-asphalt layer has proven that it improves performance and reduces maintenance costs in those areas. [6].

4. Asphalt trackbed parameters

In the early 1980s, asphalt trackbed design properties have developed and changed following the design requirements.

4.1 Asphalt layer thickness and width

The width of the asphalt layer for underlayment trackbeds designs is around 3.6 - 3.7m; As in figure 9. This width provides 0.5m as an asphalted space beyond the sleepers. [1].

Some special trackwork requires a wider asphalt layer, some crossings are provided with 7.5 - 30 m asphalt width beyond the sleepers. Asphalt underlayment

thickness depends on the quality of subgrade support, traffic loading, and type of installation. [1].

In average support conditions, a layer of 125 to 150 mm thickness is adequate, but for the poor support conditions and high impact areas, a minimum 200-300 mm asphalt layer thickness is required. [1].

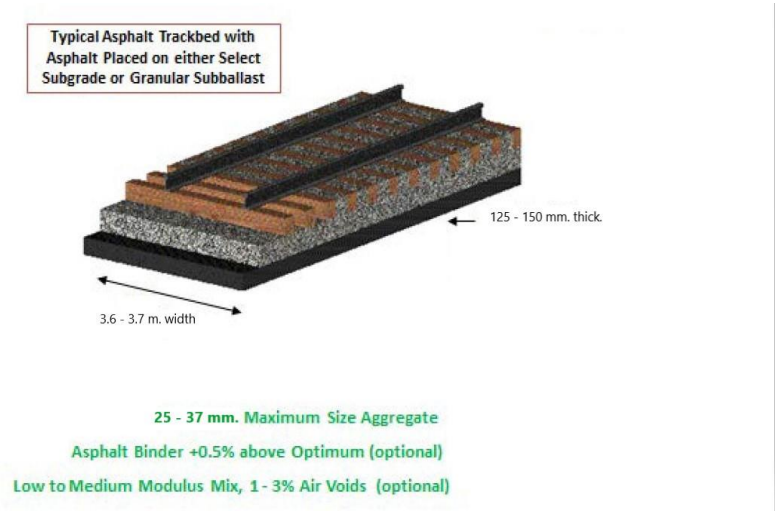


Figure 9. Typical asphalt trackbed parameter [3]

4.2 Asphalt mixture

When the asphalt is designed to be a part of the structural system of the trackbed, it is recommended that the mix should be a low modulus (plastic) mix, with a design air-void percentage of 1 - 3%, and in place air-void percentage of less than 5%. Furthermore, the maximum aggregate size is almost the same as the highway pavement mix; 25 to 37 mm is perfect for such a pavement. (Table 1 shows a suggested aggregate gradation for both underlayment and overlayment).

Binder content should be increased by 0.5% above the optimum binder content used for highway pavements to make the mix easier to be compacted. The binder used for trackbeds should be AC-10 or AC-20 viscosity-grade. The viscosity of

asphalt is usually measured at 60 °C. This bitumen grade is usually used for road construction and asphalt mixtures with superior properties.

Table 1. A suggested aggregate gradation suitable for both underlayment and overlayment [7]

Sieve size	Percent passing
1-1/2 in. (37.5 mm)	100
1- in. (25.0 mm)	90-100
3/4 in. (19.0 mm)	---
1/2 in. (12.5 mm)	70-90
3/8 in. (9.5mm)	---
No. 4	40-65
No. 10	25-45
No. 40	10-26
No. 80	6-18
No. 200	3-8
Percent AC-10, 20 OR 30* Asphalt Cement	4-8
*Based upon total weight of mixture	

5. Asphalt installation

Installation equipment needed for the new construction of trackbeds is almost the same as the equipment needed for highway construction. Asphalt is placed with conventional paving machines and compacted with vibratory rollers.

For an old trackbeds rehabilitation, the track should be removed first, and then excavations take place to the perfect depth as per the specification of the design, asphalt mix to be placed and paved as normal highway paving procedure. See figure 10. To provide positive drainage for the water away from the trackbed; the top of the hot mix asphalt should slightly be crowned or sided sloped at the same time, it should be leveled with the shoulders or a little higher. A proper drainage system necessitates a well-formed substructure crown, appropriate water spillage material, and relatively clean ballast. A protective layer that can guarantee drainage and bearing capability is required for a longer lifetime. [13].

If the asphalt layer is only 150 mm thick, it could be placed all at once as one lift, and if the asphalt layer has a higher thickness than 150 mm, it should be placed as lifts of 100 mm thick for each. The asphalt mix is best compacted when the temperature of the mixture is between 100 and 150 °C, and this is to achieve the perfect air voids ratio and the desired compaction degree; which is 95% of the maximum density or higher. [6].

Even though the construction of the asphalt protection layer is costly, the lifetime cost is more economical than alternative protection layers in the long run. This is due to the slower aging that the asphalt layer provides which leads to lower cost maintenance [15].

Stress and strain ratios in highways and railways

Depending on the base type, either full-depth asphalt, underlayment, or overlayment, for both highways and asphalted railways, and the thickness of hot mix asphalt layer, stress, and strain ratios will vary in different ways. The effect of hot mix asphalt and its thickness is showed in figure 11. [8].

Figure 10. Rail/highway crossing rehabilitation. [1].

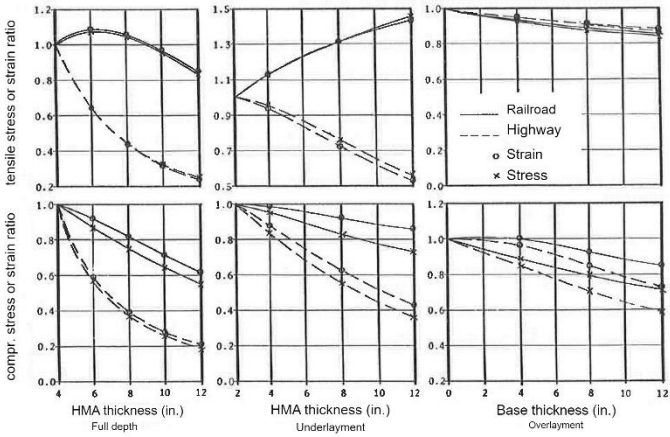


Figure 11. Effects of thickness on stress and strain ratios. [8].

The contact pressure for highways is the tire contact pressure, which is different in the railways where the maximum contact pressure is between sleepers and the underneath layers. From figure 11, many points could be summarized to explain the effect of the hot mix asphalt layer on the stress and strain for both compressive and tensile states. [8].

1- In full-depth asphalt, tensile strain, and tensile stress, both are affected in the same way by increasing the asphalt layer's thickness as the curves of stress and strain are close to each other.

2- Increasing the thickness of hot mix Asphalt in highways will decrease the tensile strain, but for railways, it is not necessary. Increasing hot mix asphalt underlayment thickness will increase the tensile strength for the railroad, this is because of the concentration of loads near the wheels due to the increasing contact pressure. (Contact pressure increases by the increase of hot mix asphalt thickness).

3- To reduce the strain in the railroad trackbed; it is more effective to use a thicker base course rather than using a thicker asphalt layer.

4- The thickness of hot mix asphalt has a different effect on compressive strength and compressive stress, as there is a large space between the curves of stress and strain. So, it is preferable to use compressive stress for the design as it is more sensitive to thickness changes.

5- The effect of thickness on compressive stress in highways is much greater than its effect on railway trackbed, and this is due to the load distribution difference; in highway, the tire loads are distributed on a small area, but in railways trackbed, the loads are distributed through rails and sleepers on a large area.

By using hot mix asphalt underlayment design in railways, the thickness of asphalt has a significant effect on compressive strength as the increased thickness will decrease the compressive stress.

6. Performance of asphalted trackbeds

The asphalted trackbed was being under monitoring to record the overall performance and the operating system evaluation.

1. Trackbed moisture

That existence hot mix asphalt will maintain the moisture level in the layers beneath the asphalt at or very close to the optimum moisture content; due to the Asphalt waterproofing role, which consequently leads to provide a consistent load-carrying capability for the layers below, also preventing pumping of subgrade materials to the ballast.

2. Long term settlement

The use of asphalt in railways improves track performance in terms of long-term settlement even in the existence of low stiffness subgrade, and this is because of its high stiffness related to the weak subgrade stiffness and the ability of asphalt to redistribute the loads.

3. Track geometry

The geometric parameters; (alignment, gage, and elevation) were periodically monitored in railways, asphalted track lines were being geometrically tested at six months intervals; those tests didn't show any changes in geometric parameters.

The tested track lines were historically in need of high maintenance costs before rehabilitation, due to weak support and drainage problems.

4. Long-term operating costs

By keeping the trackbed quality insured, the operating costs will be reduced, as the operating efficiency of the train movement is improved. Maintaining a proper track geometry will increase the speed and safety of operating systems of trains, it also reduces train resistance and fuel consumption. Less resistance means lesser wear and repair for rolling stocks.

7. Failure modes in asphalt

Asphalt roadway pavements are prone to the following breakdown modes:

1) High-temperature rutting.

2) Low-temperature cracking and fatigue.

3) Stripping/raveling on wet surfaces due to the suction of high wheel pressures.

4) Progressive fatigue cracking owing to insufficient subgrade support, which is often exacerbated by excessive moisture and poor drainage.

These circumstances do not exist on railway trackbeds made of asphalt. Temperatures, for instance, aren't high enough to encourage rutting. In the insulated trackbed environment, however, the temperatures are not low enough to cause low temperature cracking and lower fatigue life, and neither does the asphalt binder weather or harden excessively, which would have a detrimental impact on cracking and fatigue life. Because there is no rubber suction action in the trackbed environment, the potential to strip/ravel is largely avoided. [16].

8. International applications and experiences

Many countries adopted the asphalt track bit systems as a response to the high demand for stronger and durable rail lines.

9.1 Asphalt railways applications in Italy

From the early 1970s, asphalt mixes have been used in Railway Construction in Italy; more than 1200 km of asphalted trackbeds are used now. The experience showed great results in terms of high-stress elimination, protecting the embankments, eliminate Ballast fouling, and distributing the static and dynamic loads. The Italian asphalted Railway systems adopt the asphalt underlayment systems; where the cross-section consists of a super-compacted sublayer over the embankment, asphalt sub-ballast, ballast, sleepers, and rails. as shown in figure 12.

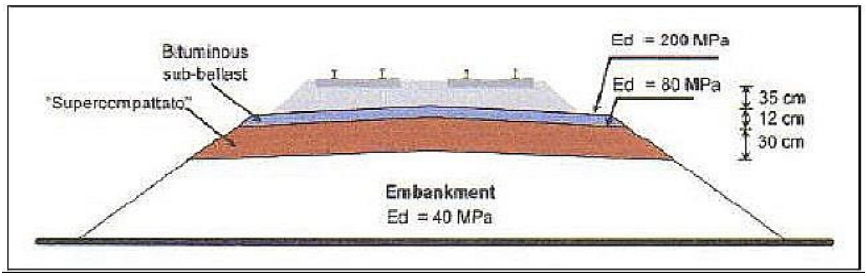


Figure 12. Italian high-speed railways cross-sectional profile. [1].

The bearing capacity of the embankment must be 40 MPa over it there, will be a super compacted layer with a bearing capacity of 80 Mpa, and a minimum thickness of 30 cm, consists of a sand /gravel mixture crowned with a slope of 3.5%.

The asphalt mixture for the sub-ballast has a maximum aggregate size of 25mm. The thickness of the layer is 120mm, and it covers the whole track cross-section. The asphalt layer must have a minimum of 200 Mpa; To be able to carry the repeated loads which come from the movement of the trains. The Italian experience showed that the asphalt layer is very effective in improving the overall performance of trackbeds. [1]. The following advantages were proven:

- Reduced vibrations and Noises.
- A reduced overall thickness of the trackbed.
- Lesser maintenance costs.
- Better water drainage protects the trackbed.

9.2 Asphalt railways applications in France

In France, the traditional railways' cross-section consists of a ballast layer of 300 mm thickness sets on a 200 mm thick sub-ballast layer, while a 500 mm thick layer of limestone carries the upper layers. In 2005 the French national railways brought the new concept of asphalted railways, where the 500mm thick layer is replaced by two layers; the upper one is an asphalt layer with 140 mm thickness, and the lower one is called the adjustment layer with 200 mm thickness, which means that the thickness is reduced by 140 mm. To protect the adjustment layer against the construction equipment; a surface addressing is applied, another surface dressing is applied over the asphalt layer too. [5]. figure 13 shows the cross-sectional profile for each traditional and asphalted railway in France.

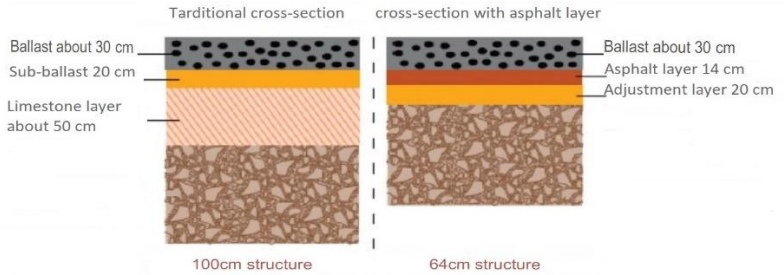


Figure 13. Traditional and asphalted sub-ballast cross-sections in France [5]

9.3 Asphalt railways applications in Japan

For several years, asphalted lines in Japan have been commonly used on high-Speed and normal railway lines in railway ballast tracks. The first goal was to support the ballast effectively and to eliminate track irregularities using asphalt track beds, the load level in the subgrade is also decreased to stop subgrade deformations. [5] The performance-based design process has three separate standard performance designs:

• Rank I performance: ballastless roadbed or asphalt roadbed.

• Rank II performance: Asphalt roadbed for the ballasted track.

• Rank III Performance: Crushed stone roadbed for the ballasted track which is the typical design using granular layers only.

The Performance Rank I is a ballast-less slab track with a concrete or asphalt supporting where sleepers are directly fastened to the slab. It is designed as the highest quality track in Japan. It is designed and tested for track settlement, concrete reinforcement base breakage, fatigue, cracking, and thermal stresses. For the Performance Rank I asphalt ballast-less track, the typical dimensions are:

• Slab width: 2220 mm

• Concrete slab thickness: 190 mm

• Asphalt-concrete base thickness: 150 mm

• Well graded crushed stone layer thickness: 150 mm

A ballasted track with a 50 mm thick asphalt layer is the **Performance Rank II design**. As shown in figure (14). Because of the ability of asphalt to spread loads and promote drainage, this style has been used for over 30 years in Japan. The design requirements for the performance rank II design are track settlement and fatigue damage to the asphalt. the typical dimensions are:

- Ballast thickness under the sleeper: 250-300 mm.
- Asphalt layer thickness: 50 mm.
- Well-graded crushed stone layer thickness: 150-600 mm.

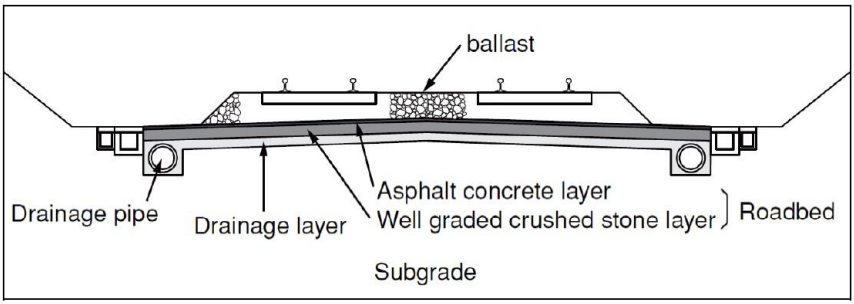


Figure 14. Performance Rank II design [5]

9.3 Asphalt railways applications in Germany

In Germany, Deutsche Bundesbahn AG (DB AG) concluded that the rail web (rails and sleepers) with a ballast bedding type of construction had reached a level that could hardly be improved as a classical construction method for railroad tracks.

Asphalt will replace the load-bearing ballast. For the first time in Germany, this construction method was used about 35 years ago, with a base level in asphalt. Since then, DB AG has accepted several systems of the asphalt design concept. They are the following, in detail [5]:

- ATD Asphalt base course with a rail web
- SATO Concrete sleeper or Y-steel sleeper with double base
- Walter System Walterbau
- Getrac German Track Corporation Asphalt.

10. Conclusion

The asphalt layer is being used worldwide now in combination with granular layers to add the preferable properties of asphalt to the trackbed; the main purpose of asphalt in railways is to provide an additional supportive, waterproofing layer to bear the repeated stresses coming from the movement of trains, also to be a load distributor layers to increase the lifetime of the tracked.

When asphalt underlayment is used with a conventional ballasted track, the track stiffness increases, and total settlement decreases. This is directly impacting the cost of track maintenance. Furthermore, as compared to a typical granular sub-ballast, ballasted rails with asphalt layers will require less maintenance.

The amount of stress and strain beneath an asphalt layer in a railway track is mostly determined by the thickness of the layer; however, substituting granular

materials with asphalt pavement can reduce the total height of the ballasted railway track. This is a critical consideration in situations where the total height of the trackbed must be critically controlled, such as tunnels and bridges.

When it comes to ballastless railway tracks, asphalt is an excellent choice. It offers a good degree of stiffness balance throughout the track system, which is critical for overall stability and, at the same time, response to the train's dynamic loads.

Many countries adopted this technology and tested it during the last years it was clear that the presence of asphalt either as an underlayment or instead of ballast carried the railway industry to a new level.

Using asphalt in railway trackbed increases maintenance cycles' gaps. However, using new asphalt technologies such as self-healing asphalt and nanoparticles in asphalt will promote the performance of the railway trackbed in general and transfer it to a new level.

More studies are being made to reach the best combination and form of using the asphalted trackbeds.

Acknowledgment

The publishing of this paper was supported by Professor Szabolcs Fischer.

References

[1] Jerry G. Rose, Reginald R. Souleyrette, *Asphalt Railway Trackbeds: Recent Designs, Applications, and Performances*, AREMA 2015.

[2] Coenraad Esveld. *Modern Railway Track.* Second ed., Zaltbommel, Mrt Productions, 2001.

[3] Jerry G. Rose, *Asphalt Underlayment Railway Trackbeds: Designs, Applications, and Long-Term Performance Evaluations*, NURail Project ID: NURail 2017-UKY-R12b, 2017.

[4] Jerry G. Rose, Henry M. Lees, Jr., *Long-Term Assessment of Asphalt Trackbed Component Materials' Properties and Performance*, 2008.

[5] European Asphalt Pavement Association (EAPA), Asphalt in Railway Tracks, EAPA Position paper, June 2014.

[6] Jerry G. Rose, E. Ray Brown, and Monica L. Osborne, *Asphalt Trackbed Technology Development the First 20 Years,* Transportation Research Record 1713, Paper No. 00-0207, 2000.

[7] Jerry G. Rose and M. J. Hensley, *Performance of hot-mix-asphalt railway trackbeds*, Transportation research record 1300, 1991.

[8] Yang H. Huang, Jerry G. Rose, and Charles J. Khoury, *Hot-mix asphalt railroad trackbeds,* Transportation research record 1095, 1986.

[9] B. Eller, Sz. Fischer, *Tutorial on the emergence of local substructure failures in the railway track structure and their renewal with existing and new methodologies*, Acta Technica Jaurinensis, Vol. 14, No. 1, pp. 80-103, 2021.

[10] Jerry G. Rose, Lindsey Sebastian Bryson, *Hot Mix Asphalt Railway Trackbeds: Trackbed Materials, Performance Evaluations, and Significant Implications*, The International Conference on Perpetual Pavements, 2009.

[11] Moritz, B. (1916) *Bau der Ḥiğâzbahn Schienenlegung bei Tebûk*. Saudi Arabia, 1916. Berlin: Dietrich Riemer. [Photograph] Retrieved from the Library of Congress, `https://www.loc.gov/item/96513602/`

[12] Gilmore & Associates, Inc, *Infrastructure management plan, Narrative report,* prepared for Borough of Conshohocken, 2018.

[13] Eller, B. and Fischer, S., 2019. *Review of the modern ballasted railway tracks' substructure and further investigations. Science and Transport Progress.* Bulletin of Dnipropetrovsk National University of Railway Transport, 0(6(84), pp.72-85.

[14] Fischer, S., 2021. *Investigation of effect of water content on railway granular supplementary layers.* Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, (3), pp.64-68.

[15] Fischer, Szabolcs & Eller, Balázs & Kada, Zoltán & Németh, Attila. (2015). Railway construction. Győr: Universitas-Győr Nonprofit Kft., pp.54-55.

[16] Sehgal, V., Sehgal, L., Garg, A. and Buttlar, W., 2017. *Hot Mix Asphalt in Ballasted Railway Track: International Experience and Inferences*. [online] Iricen.gov.in. Available at: <https://www.iricen.gov.in/iricen/ipwe_seminar/2017/36.pdf> [Accessed 4 October 2021].

[17] Sysyn, M., Nabochenko, O., Kovalchuk, V., Przybyłowicz, M. and Fischer, S., 2021. Investigation of interlocking effect of crushed stone ballast during dynamic loading. *Reports in Mechanical Engineering*, 2(1), pp.65-76.

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 631-649, 2021
DOI: 10.14513/actatechjaur.00626

Methods to detect and measure scour around bridge foundations

M. Al-Jubouri[1*], R. P. Ray[1]

[1]Széchenyi István University, Structural and Geotechnical Engineering Department
Egyetem tér 1, 9026 Győr, Hungary
*e-mail: muhanad.kh.99.oo@gmail.com

Submitted: 19/05/2021; Accepted: 11/09/2021; Published online: 18/10/2021

Abstract: Bridges are indispensable structures vital to the operation of road and rail transportation networks. Crossing rivers and artificial waterways, however, presents a risk to their foundations due to scour actions. Scour is the number one cause for bridge failures and may occur beneath any bridge, large or small, with supports located within the waterway. This paper provides a summary of present scour detection and measurement equipment and associated assessment methodologies. In this regard, particular emphasis is placed on structural health monitoring better to evaluate the presence and influence of potential scour. A Sensitivity Analysis on a newly introduced monitoring system is also assumed. Furthermore, much research has been undertaken to create a technology that can instantly identify and detect bridge scour, improving survey reliability through prior inspection and prompt intervention. This research will explore and evaluate bridge scour detection methods employed and suggest a possible path for developing the detection system to identify scour depth effectively and efficiently. Finally, our key aim is to minimize human effort in identifying and bridge scour by using a quick, easy-to-use, cost-effective process, resulting in fewer injuries and economic savings.

Keywords: *water action; Bridge scour; detection methods*

1. Introduction and motivation

Bridges are critical engineering components within transportation networks (road, rail, pipeline, and waterway) that provide a means of crossing waterways, valleys (viaducts), and other networks (road/rail overpass). As a result, they may be exposed

to a variety of hazards, both natural and man-made. They are at the confluence of natural barriers and human networks by the very nature of their location and function. Any damage or loss of function could result in lives lost, as well as an economic and social disaster. Engineers who study bridge damage and collapse cite scour as the number one cause for loss of operational capacity. Bridge failures due to scour have been reported by ([1], [2], and [3]), with 6000 bridge failures in the USA and 140 UK rail bridge failures during 1846-2013. A database of over 600,000 US bridges was studied and compared to 1,700 collapses due to hydraulic action by [4]. They found a wide divergence of causes and evaluated many rating systems to predict performance. Scour is also identified as the bridge management risk most likely to be impacted by climate change; approximately 1 in 20 in the UK are expected to be at high risk by 2080 [5]. Given the potential impact, mitigation measures are equally daunting with costs averaging 25 to 76 million euros annually for 2040–2100 [6].

Within risk and resilience, scour is one of many natural hazards that bridges are designed to resist. Others would include wind (hurricane, tornado), flooding (high water level, high velocity, and debris), earthquake (shaking, liquefaction), and temperature extremes (thermal stresses, expansion, and contraction). Man-made hazards would include extreme traffic and vehicle loads, collisions, and accidents from vehicles as well as ships and barges. Depending on what is permitted on the bridge, there is also the possibility of fire and explosions. The recent focus on force protection for critical infrastructure highlights the potential of terrorist attacks [7].

Further assessment of risk involves quantifying the severity and recurrence of these hazards and estimating the structural resilience of the bridge system to them. Finally, exposure levels and response to an event must be evaluated. So, the types and depth of bridge scour are an integral part of risk assessment for every bridge crossing a waterway [8].

2. Scour along waterways

Scour is the removal and transport of sediment from around hydraulic structures. As a natural process, sediments are constantly eroded, transported, and deposited along waterways. However, certain hydraulic conditions will erode large volumes of material from around bridge foundations that weaken their support capacity and compromise the safety of the superstructure.

Three primary terms are used to categorize scour, specifically general scour, contraction scours, and local scour. General scour includes the natural processes of aggradation and degradation of streambeds due to changing hydraulic parameters such as variations in flow discharge or sediment amount [9]. It drives the natural

evolution of a waterway and results in the constant re-routing of the river channel when there are no natural or artificial obstacles [10].

Contraction scour occurs due to sudden channel geometry changes that constrict flow, causing an increase in water velocity. For example, bridge piers, shoreline facilities, hydraulic control structures, even ship traffic, can alter the channel cross-section and induce higher flow velocities. The increased speeds lead to higher sediment bed shear stresses, and if the stresses exceed the threshold level of the bed material, sediments mobilize and scour initiated [11].

The term "local scour" is given to describe erosion that occurs around hydraulic structures such as bridge piers and abutments (Fig. 1). An interacting set of unsteady flow features lifts and transports sediment from the pier foundation. They comprise flow impact with the pier face, creating a downward flow and an upward flow with a roller vortex; flow converging, constricting, then diverging; the creation, movement, and collapse of large scale turbulence in the foundation of the pier junction (horseshoe vortex, Fig. 1); a detaching shear layer where the pier cross-section curves back past its widest point; and, wake vortices generated through the pier's wake. The features evolve as scour develops. For example, when the water flow faces an obstruction such as a bridge pier, downward flow is convinced at the upstream end, causing localized corrosion around this construction [12]. The combined actions of the three forms of scour described may lead to significant losses in soil from around foundation elements.

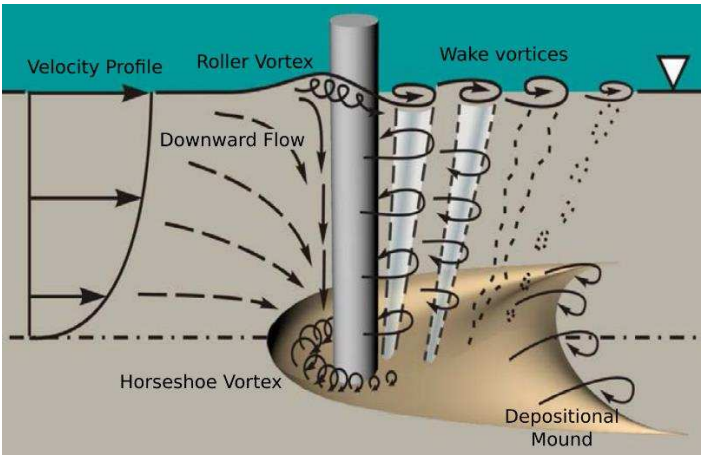


Figure 1. Local scour around the cylindrical pier in deep water, adapted from [13].

3. Damage from scour

Examples of scour causing serious bridge failures as the 1987 Schoharie Creek Bridge (Fig. 2a), part of Interstate-90 New York USA, where ten lives were lost. At the collapse, peak flow was 1,840 m3/s with a (70- 100) year return period. The foundations of the four bridge piers were large spread footings 25 m long, 5.5 m wide, and 1.5 m deep without piles. The footings were set 1.5 meters into the streambed on highly dense ice contact layered glacial that the designers deemed non-erodible. Flume investigations of the stratified drift, on the other hand, revealed that some material would be eroded at 1.5 m/s, with substantial rates at 2.4 m/s. A 1:50 scale 3-D physical model study indicated a prototype flow velocity of 3.3 m/s along the pier that failed. Additional modeling estimated 4.6 m of maximum scour depth. The prototype pier (pier 3) had a scour depth of 4.3 meters at the breakdown time. In 1995, the Arroyo Pasajero Bridge (Fig. 2b) on Interstate-5, California (USA), collapsed, killing seven people. The stream system is temporary (usually dry), with a sand bed having planar topography. While discharges were hard to quantify, the USGS using slope-area methods, determined that the 1995 discharge ranged 462–1,141 m3/s with the best estimate of 773 m^3/s and recurrence 75 years [14].

Factors contributing to the I-5 bridge failure were:

- Regional subsidence causing an increase in channel slope (higher velocities)
- The original design changed by placing a solid web wall between columns to repair damage from an earlier flood. The wall had an angle of attack from 15 to 26 degrees, blocking thru-flow and potentially increasing local pier scour depth by a factor of 3.6 – 4.4
- There was a 33 percent increase in drainage area upstream from land-use change and a channel to link two streams.
- Long-term degradation of 3 m since the bridge was built
- Channel width reduction of 90 to 120m to a bridge width of 37 m.

(a) (b)

Figure 2. (a)Schoharie Creek Bridge [15] and (b) Arroyo Pasajero [14]

4. Methods to monitor bridge scour

Bridge designers can reduce scour through both hydraulic and structural countermeasures [15]. Hydraulic approaches may reduce abrupt flow expansion or contraction caused by poor streamlining, such as blunt pier faces. Such changes often lead to the generation of the vortices responsible for scour. Maintaining larger bridge openings at the design stage and streamlining pier geometries reduce vortex generation. Another critical factor is to keep openings clear by removing debris that often obstructs flow. However, deposits of the natural canal and upstream erosion may frequently modify the flow angle, and rationalized abutments pose similar difficulties. Structural features can be included during the design stage to locate spread footing below maximum scour depth and add rip-rap to the base of piers. This assumes the maximum depths are specific and hydraulic conditions will remain relatively constant over the bridge's lifetime. Another approach is based on the observational method where scour is monitored over time, and remediation works are implemented as needed [16]. When implemented correctly, this can be an efficient and economical method.

As part of a general bridge asset management scheme, visual inspection is the most common monitoring technique [17]. They often incorporate the employment of divers to check foundation elements and estimate their depth of scour by use of basic instruments with a structural examination [18]. Two important disadvantages of this method are that inspections during flooding are impossible, and scour holes tend to be filled in as floodwaters subside. Since scour holes may refill after flooding, misinterpretation of the inspection may hide the true extent of the scour problem.

Scour depth may be measured or monitored in a variety of ways. They include:

- In-place devices that sense the presence (or absence) of soil. They may be penetration rods, floats with embedded switches, conductivity or dissolved oxygen sensors, or a variety of other approaches. They are often used when frequent or regular monitoring is required [15].
- Remote/portable sensors, which scan a greater area than sensors in (1) using sonar, geophysics, or similar concepts.
- Response analysis of the bridge, single pier, or abutment. Instead of measuring scour depth directly, they measure the effect of scour on the vibration response of the bridge or its components.

All the measuring and monitoring methods may be adopted into an overall program of operational health monitoring for the bridge. Each of these approaches has its advantages, and they are discussed further in the following sections

4.3. Scour monitoring using in-place (fixed) instrumentation

Most instrumentation fixed to the structure consists of sounding rods, driven rods with sensors, fathometers (sonar), buried transmitters, or tilt and vibration sensors. Table 1, presented at the end, lists a summary of all methods discussed here. Sonar devices can be mounted on the upstream face of the bridge pier and measure the distance to the stream bed. Measurements can be recorded continuously over any period to determine the depth of scour and refill during a high-flow event (Fig. 3a). Additional data logging and transmission equipment, as well as a solar panel or standard line power supply, would complete the package. Magnetic sliding collars (Fig. 3b) are rods attached to the face of a pier or abutment and driven into the streambed. A magnetic sensor collar is lowered to the streambed. If the streambed corrodes, the collar recession the rod into the scour hole, and the depth indicates the amount of scour. Magnets in the collar come into proximity with switches within the rod that close as the magnet slides into the scour hole. The data logger senses the switch condition and determines the level of the collar and scour activity. Magnetic sliding collars can only be used to monitor the maximum scour depth. Buried float-out devices can be active or inert buried sensors (Fig 3c) with a radio transmitter that may be active or awaiting activation. The signal, when engaged, is detected by a nearby receiver. The float-out is buried in a horizontal orientation and either transmits this state or does not transmit at all. When scour occurs to the depth of the device, it floats, changing its direction to vertical, and either activates (if previously dormant) or transmits a different signal. The receiver listens to the new signal and knows which device is sending it. Scour depth is based on the burial depth of the newly activated signal. The float-out sensor is easy to install in dry or nearly dry conditions, under armor stone and rip-rap, or at various pre-determined depths. Due

to low power requirements, the float-out sensor can remain buried for many years and activate when released by scour activity [20].

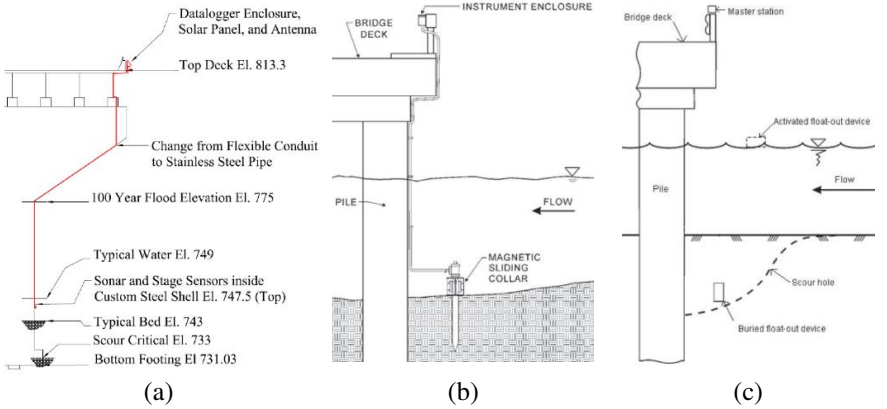
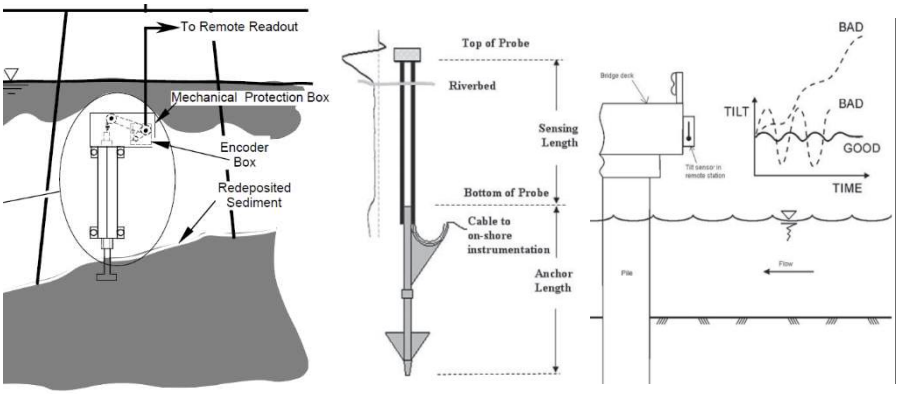


Figure 3. Fixed devices (a) sonar, (b) magnetic collars, (c) float-out device, modified from [21]

Other fixed systems may use sounding rods, time-domain reflectometry, or other sensors such as dissolved oxygen to report the scour depth. Sounding rods are placed on the surface of the river sediment (Fig. 4a). They have an enlarged base that will not penetrate the surface but move downward as the sediments erode beneath. An encoder box measures the downward movement and relays the information to the data collection/transmission system. Time-domain reflectometry (TDR, Fig. 4b) uses two wires to sense the dielectric environment outside the sensor. Sediments and water have different dielectrics, and the TDR device will detect the interface. Only the sensing length is used to measure depth. A tilt meter is sensitive enough to measure variations in abutment tilt (Fig. 4c) that may be a result of daily temperature changes, expected changes in water levels, or regular operations. Once those behaviours are catalogued, deviations from that behaviour would trigger an alarm that warns engineers to check the bridge for dangerous conditions. With some further calibration and analysis, the cause of the unusual behaviour may be determined. Vibration monitors are worked similarly; however, this approach uses the principle of calculating the critical occurrence of the rod fixed in the streambed. The opposite relationship between fundamental frequency and the sensor of the rod length is applied to monitor scour depth. It uses structural vibration sensors, such as accelerometers or fiber-optic (FBG) devices, to be used as the scour sensor's dynamic sensing feature. However, this approach is yet to be thoroughly tested. It is continuing research, and studies are being carried out on it [22].

All of the in-place instruments are limited to a small area of interest immediately surrounding the probe. They give little information about the areal extent of scour or specific hydraulic conditions (aside from depth of scour).



(a) (b) (c)

Figure 4. In-place measurements (a) recording sounding rod, (b) TDR probe, (c) tilt meter [9, 21]

4.2 Scour monitoring using portable instrumentation

Schall and Price [23] presented an extensive discussion on the theory and practice of using portable monitoring equipment. Being written in 2004, much of the technology is less expensive, but many of the concepts, design processes, and practical problems with deployment and handling are still the same. Portable instruments include four components:

- The measuring instrument (sonar, geophysical, etc.)
- The system to deploy the instrument (crane, paddleboard, remote-controlled boat, drone).
- Method to record horizontal (x, y) position of the instrument. It may also include some vertical (z) references as well.
- Data storage and transmission.

With newer technologies, components 3 and 4 have become much more economical and efficient. RTK GNSS modules, weighing 35gm, can record location information with centimetre accuracy for less than 500 euro [21]. Data storage options can be on-device with SD cards holding 128GB+ of data or transmitted to a receiver station or smartphone. With higher accuracy come different challenges: one must know the

various locations of the RTK module and the sensor. If the sensor is moving, then any time lag between sensor trigger and position report will impact the sensor's reported vs. actual location.

Sonar systems use a variety of transponders and methods. Generally, the signal is a narrow beam to reduce the possibility of signal averaging over a wider area. Monitoring concepts for structural structures have been through a steady growth phase over the last decade. As a result, they play a role in the design of new and emerging architectures. This chapter briefly reviews a selection of the more often used scour detection and measuring techniques. In terms of the instruments used, they are divided into portable instrumentation methods and integrated instrumentation methods. Mobile instrumentation methods include Physical probing, fathometers (Sonar), and Geophysical information, while fixed or integrated instrumentation methods include buried RF sensors, sonar, and other similar devices [24] as shown in Fig. 5.

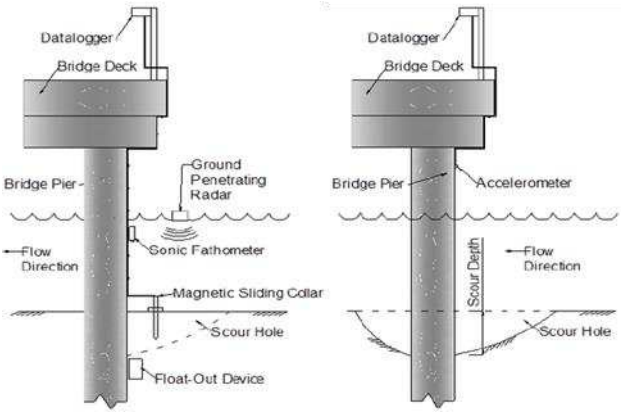


Figure 5. Portable (mobile) system measurement [26].

4.2.1. Fathometers and Manned ships

Fathometers or acoustic depth sounders are commonly used for portable scour measurements. In addition, precision survey-grade hydrographical survey fathometers and fish finders are used. Transducers are fixed to a pole, hand-line, tethered buoy, or boom while measurements are taken from the bridge. Kneeboards and pontoon-style floats are examples of tethered float platforms. In fast-moving, swirling water, the size of the float is critical for stabilization. In addition, a bridge inspection truck can deploy floating or non-floating structures. When the bridge is well over the river, this is especially helpful. For instance, bridges more than 15

meters above the sea are usually inaccessible from the bridge surface unless this method is used.

[25] Created an articulated arm to position a sonar transducer. The machine was trailer-mounted and could be used on bridge decks ranging from 5 to 15 meters above the water's level. An onboard computer measured the transducer's orientation about a known location on the bridge deck based on the boom's angle and the space between the boom pivot and the transducer.

Manned ships are used as a scour measurement platform. They usually need enough distance below the bridge and near the launch facilities. During floods, it may be a challenge as the river level can exceed or submerge the bridge's low chord, and boat ramps may be submerged. A fathometer is typically used for depth measurements, and GPS devices are usually used for location. The construction of an unscrewed prototype boat was prompted by concerns about safety, launching, and approval. A compact flat-bottomed jon yacht, an 8-hp outboard engine, a fathometer, and GPS with remote controls are applied. It was put to the test through six floods with great results [25].

The benefit of GPS over conventional land-based intrusion detection eliminates the need for line-of-sight between control points. GPS can be used at night and in inclement weather, particularly useful for scour tracking during floods. The downside of GPS is that it cannot be used in parts where the overhead obstructions, such as tree canopy or bridge decks. Though, GPS measurements up to the bridge face have been accurate without going under the bridge.

4.2.2. Sonar and Geophysical Techniques

The interfaces between different resources with different physical characteristics are determined by geophysical instruments based on wave spread and reproduction measurements. Sonar and geophysical techniques vary in that geophysical approaches have sub-bottom detail, while sonar can only detect the water-soil interface and not reach the sediment layer. The key differences between various geophysical techniques are the kinds of signals emitted and the physical property modifications that produce reflections. Like sonar, a seismic instrument uses acoustic waves but at a lower range (2-16 kHz). Seismic sounds, like sonar, are susceptible to being dispersed by air bubbles and high sediment concentrations [27]. The best use of geophysical technology is to assess scour depth in infilling areas during a flood under lower flow conditions. The equipment's expense and complexity and the data's interpretation restrict widespread usage and use as a portable scour monitoring system. More developed, minor price GPR devices with electronic data dispensation capability have been recognized, these issues have

lessened. However, expense and difficulty and the need for borehole details and reliable bridge plan details to calibrate and interpret the results can also restrict GPR.

4.2.3. Diving Probing and sound rods

Diving is a basic scour screening system in which a trained bridge inspector conducts a manual inspection of the bridge underwater. This system will capture scour data from various sites, and the water clarity does not affect the data collection process. However, the downside of this approach is that it can be costly, making it more ideal for worst-case cases. It also has a strong potential for risk. Furthermore, owing to the subjectivity of the regulators, the data produced from such visual inspections may have a high degree of uncertainty [28]. A sound rod is used for the bridge inspector to manually test a bridge by placing a rod or weight on the streambed to determine the sediment depth. The bottom of the rod must be wide to keep the rod from entering the streambed due to the rod's weight and friction induced by running water. When the riverbed is sand, sounding rods appear to go into the streambed, affecting their accuracy. Diving probing has many advantages, including the fact that it is not compromised by air entrainment or heavy sediment levels, and it can be used in fast, shallow water. The critical problem with this approach is the inaccuracy of the data samples gathered and the possible risks inherent with this method. Besides, this approach can be costly and does not have the capability of automatic warnings.

4.3. Applying other new methods for scour detection

The soil-structure interface mechanism is complicated during scour. However, the material displacement under (or around) the base during scour induces increased tension and decreased residual soil stiffness. Although the vibration of the structure depends on the device's rigidity, the observation of variations in vibration rates is a possible tool for detecting damage and monitoring health. The framework's average frequency related to the bridge piers could be found by applying spectral analysis approaches, for example, the FFTs accelerometers equestrian and frequency domain fragmentation on the bridges. The underwater instruments have often been used to calculate the progression of the scour depths over time. However, few investigations have been accompanied to understand the impact of the scour on the bridge system's reaction. Some of the instruments designed to test the bridge structure's reaction to the scour including tilt-meters which measure the comparative variation of the structural feature and, as such, also for distinguishing differential settlements that may happen as a consequence of the scour process. The only significant drawback of the system is that it does not straight designate the scour's extent. Devices capable of specifically assessing structural distress are expected to help engineers undertake the necessary repair arrangements for critical structures before the collapse.

Furthermore, Accelerometers are used to compute the structural reaction, especially in a change in boundary conditions [19].

There has been a spectral analysis reaction, and an elastic reaction spectrum of the seismic guidelines was created for this method of analysis that characterizes the activities of the earthquake [29]. In addition, a nonlinear study of the complex past of time was conducted to detect the potential behind non-elastic conditions. The principal findings are that the continuum analysis of reactions is an estimated method for assessing the optimum importance of inner forces and the nonlinearity of the pier segment reduces the inner forces and movement compared to linear analysis. This method could be considered a possible tool for detecting bridge damage and monitoring health, and distinguishing differential settlements that may happen due to the scouring process.

An Internet of Things (IoT) and Artificial Intelligence (AI) combination scoring system is created and deployed to get real-time measurements in scour depth. These vibration-based MEMS devices are packaged in a water-resistant steel ball in a barrel cage to survive extreme flooding. First, the fluvial water-level fluctuations surrounding the jetty are made using the Mask R-CNN profound learning model with real-time CCTV pictures (Fig. 6). Then, the scour-depth development is simulated with the hydrodynamic model and the sediment transport formula using the specified local scour formulas [30]. The overall performance of the hydrodynamic modelling is reasonable, based on the technique stated above, for whole scour-depth development. Moreover, it can anticipate scour-depth changes in bridge failure for early warning.

Figure 6. Fixing of the scour observing scheme, sensors, and the wireless station [30].

5. Summary

The author summarises the main results in Table 1 and Table 2 that contain the relevant methods, their tools, and experiments, as well as the properties of the methods based on international literature.

6. Conclusion

All methods presented in scour detection are costly in monetary costs, and more human resources are required. Therefore, instead of investing so much money on new gadgets, sensors can be used that are already integrated into our smartphones. A primary mobile phone, which anybody can use, can identify scour beneath bridges foundation (piers and abutments). The key aim is to minimize human effort in identifying scour by using a quick, easy-to-use, cost-effective process, resulting in minor injury due to bridge collapse.

Conversely, traditional procedures mostly use underwater tools to detect bridge scour depths and shape, which are also difficult to distinguish in instrument deployments and facilities. Recently, the method of Vibration-based destruction monitoring has been discovered to overcome specific difficulties by investigating the natural frequency range of a bridge or bridge part. Additionally, the advancement of these fixed and portable scour measurement devices, along with GPS, remotely operated ships, instrumented vehicles, and knowledge of the need to calculate and monitor bridge spacing, have greatly enhanced the scour database, methods for forecasting spacing depths, bridge scour and bridge protection.

Table 1. The portable bridge scour instrumentation detection methods

	Tools and experiments	***Properties***
Diving probing and sound rods	• A basic scour screening system in which a trained bridge inspector conducts a manual inspection of the bridge underwater. • A sound rod is used for the bridge inspector to manually test a bridge by placing a rod or weight on the streambed to determine the sediment depth.	• Capture scour data from a variety of sites, and the water clarity does not affect the data collection process. • It can be costly, making it more ideal for worst-case cases. • It is a strong potential for risk. • Owing to the subjectivity of the regulators, the data produced from such visual inspections may have a high degree of uncertainty. • It is not compromised by air entrainment or heavy sediment levels, and it can be used in fast, shallow water. • The critical problem with this approach is the inaccuracy of the data samples gathered and the possible risks inherent with this method.
Fathometers	• Acoustic depth sounders are commonly used for portable scour measurements. • Precision survey-grade hydrographical survey fathometers and fish finders are used. • Transducers are fixed to a pole, hand-line, tethered buoy, or boom. • Kneeboards and pontoon-style floats are examples of tethered float platforms • Articulated arm to position a sonar transducer.. • Manned ships are also used as a scour measurement platform.	• A bridge inspection truck can deploy floating or non-floating structures. • The advantage of GPS over traditional land-based vulnerability scanning is that it reduces the need for line-of-sight among control points. • The GPS may be used at night and in inclement weather, which can be especially useful for flood monitoring. • The downside of GPS is that it cannot be used in parts where overhead obstructions, such as tree canopy or bridge decks. • GPS measurements up to the bridge face have been accurate without going under the bridge.
Geophysical data	• The interfaces between different resources with different physical characteristics are determined by geophysical instruments based on wave spread and reproduction measurements. • Sonar and geophysical techniques vary in that geophysical approaches have sub-bottom detail, while sonar can only detect the water-soil interface and not reach the sediment layer • More developed, minor price GPR devices with electronic data dispensation capability have been renowned.	• Differences between various geophysical techniques. Like sonar, a seismic instrument uses acoustic waves at a lower range (2-16 kHz). • The best use of geophysical technology is to assess scour depth in infilling areas during a flood under lower flow conditions • The equipment's expense and complexity and the data's interpretation restrict widespread usage and use as a portable scour monitoring system. • Expense, difficulty, and the need for borehole details and reliable bridge plan details to calibrate and interpret the results can also restrict GPR.

Table 2. The fixed bridge scour instrumentation detection methods

Methods	***Tools and experiments***	***Properties***
Sounding rods	• The binding of the rod in its supporting enclosure (pipe), • The producer (Cayuga Industries) and the USGS on coarse-bed streams • long probes installed in the stream bed at the point of interest.	• A substantial depth of scouring, in sand-bed streams, sand deposited between the rod and its supporting section would also tie it up • In sand-based non-cohesive bed materials, the rod would penetrate a considerable and indeterminate volume into the bed • It could be costly and timewasting. • The precision of the measurement is affected by varying temperatures in the channel, with relative errors of 5% reported in some studies
Driven Rods (Buried)	• Both sensors and instruments supported by a vertical support member such as a shaft, rail, or column mounted vertically in the bed • The gravity sensor is placed in the stream bed near the guided rod scheme. • A remote sensing component is typically used to detect changes in the depth of the gravity sensor.	• This system offers a reasonably easy method for tracking the scour depth's advancement. • There is a range of drawbacks as Scour depths may only be noticed in the instrument's immediate vicinity, meaning that several devices could be needed to catch the scour's actual (global) effect. • The element uses a gravity device that remains at the deepest depth of the scour throughout each flood incidence. This means that it will have to be reset, which could be costly and timewasting, and would not include details on the scour hole's refilling
Electrical conductivity devices	• Use the variations in the different media's electrical conductivity to determine the water-sediment interface's direction using two probes	• As the material between the probes varies, the capacity to draw a current would also change. • This phenomenon can be used to show the presence and extent of the scour.
Sonar-based sensors	• permanent devices that are typically mounted on the pier and abutment. • Sonar emits pulse waves and processes the trip portable time of a pulse from the riverbed.	• Both scour, and accumulation of sediments can be measured using sonar sensors • The measurements are influenced by heavy sediment and turbulent flow in the sea. • High-end sonars with a large depth capacity and high resolution can be costly
Time Domain Reflect-meter (TDR)	• Time Domain Reflect-meter (TDR) based approaches use a similar sonar approach to measure scour.	• Determining the portion of the conduits submerged in the stream bed. • TDRs are capable of repeatable measurements and are very durable. • involve much fuel, making them costly.
Vibration-based method	• Calculating by the essential occurrence of the rod fixed in the streambed. • It uses structural vibration sensors, such as accelerometers or fiber-optic (FBG) devices.	• To monitor scour depth, the opposite relationship between essential frequency and the sensor of the rod length is applied. • This approach is yet to be thoroughly tested. and studies are being carried out on it.

Acknowledgment

Professor Csaba Koren helped and supported the publishing of this paper in Széchenyi István University as part of the New solutions in the design of road and railway infrastructure course submission.

References

[1] K. Faulkner, J. Brownjohn, Y. Wang et al., Tracking bridge tilt behaviour using sensor fusion techniques, *Journal of Civil Structural Health Monitoring* 10 (4) (2020) pp. 543-555.
doi: https://doi.org/10.1007/s13349-020-00400-9

[2] L. J. Prendergast, K. Gavin, A review of bridge scour monitoring techniques, *Journal of Rock Mechanics and Geotechnical Engineering* 6 (2) (2014) pp. 138-149.
doi: https://doi.org/10.1016/j.jrmge.2014.01.007

[3] Z. Van Leeuwen, R. Lamb, Flood and scour related failure incidents at railway assets between 1846 and 2013, *Railway Safety and Standards Board*, 2014 [cited 2021-06-29].
URL https://www.jbatrust.org/how-we-help/publications-resources/risk-analysis/flood-and-scour-related-failure-incidents-at-railway-assets/

[4] C. Montalvo, W. Cook, T. Keeney, Retrospective analysis of hydraulic bridge collapse, *Journal of Performance of Constructed Facilities* 34 (1) (2020): 04019111.
doi: https://doi.org/10.1061/(ASCE)CF.1943-5509.0001378

[5] R. J. Dawson, D. Thompson et al., A systems framework for national assessment of climate risks to infrastructure, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (2018): 20170298.
doi: https://doi.org/10.1098/rsta.2017.0298

[6] F. Nemry, H. Demirel, Impacts of Climate Change on Transport: A focus on road and rail transport infrastructures, *European Commission, Joint Research Centre (JRC), Institute for Prospective Technological Studies (IPTS)*, 2012.

[7] E. B. Williamson, G. David, Risk management and design of critical bridges for terrorist attacks, *Journal of Bridge Engineering* 10(1) (2005) pp. 96–106. doi: https://doi.org/10.1061/(ASCE)1084-0702(2005)10:1(96)

[8] J. L. Briaud, P. Gardoni, C. Yao, Statistical, risk, and reliability analyses of bridge scour, *Journal of Geotechnical and Geoenvironmental Engineering* 140 (2) (2014) . doi: https://doi.org/10.1061/(ASCE)GT.1943-5606.0000989

[9] M.C. Forde, D. M. McCann et al., Radar measurement of bridge scour, I*ndependent Non-destructive Testing and Evaluation International* 32 (8) (1999) pp. 481-492. doi: https://doi.org/10.1016/S0963-8695(99)00026-2

[10] F. Federico, G. Silvagni, F. Volpi, Scour vulnerability of river bridge piers, *Journal of Geotechnical and Geoenvironmental Engineering* 129 (10) (2003) pp. 890-899. doi: https://doi.org/10.1061/(ASCE)1090-0241(2003)129:10(890)

[11] J. L. Briaud, F. C. Ting et al., SRICOS: Prediction of scour rate in cohesive soils at bridge piers, *Journal of Geotechnical and Geoenvironmental Engineering* 125(4) (1999) pp. 237-246. doi: https://doi.org/10.1061/(ASCE)1090-0241(1999)125:4(237)

[12] D. Ryan, G. A. Hamill et al., The hydraulics and resulting bed scour within the vicinity of Submerged Single Span Arch Bridges, *In Civil Engineering Research in Ireland,* 2014 [cited 2021-06-19]. URL https://pureadmin.qub.ac.uk/ws/portalfiles/portal/13538682/The_hydraulics_and_resulting_bed_scour_within_the_vicinity_of_Submerged_Single_Span_Arch_Bridges.pdf

[13] R. Ettema , B.W. Melville, G. Constantinescu, Evaluation of bridge scour research: Pier scour processes and predictions, *National Cooperative Highway Research Program Transportation Research Board of the National Academies*, Washington, DC, USA, 2011.

[14] E.V. Richardson, Photo collection, Library Archive, Colorado State University (1995) [cited 2021-05-11].
URL https://mountainscholar.org/handle/10217/170449

[15] P. F. Lagasse, L. W. Zevenbergen et al., Bridge scour and stream instability countermeasures: experience, selection, and design guidance: Volume 2. No. FHWA-NHI-09-112, *National Highway Institute,* USA, 2009.

[16] J. L. Briaud, S. Hurlebaus et al., Realtime monitoring of bridge scour using remote monitoring technology*,* No. Report 0-6060-1. *Texas Transportation Institute,* USA, 2011.

[17] L. J. Prendergast, K.Gavin, A review of bridge scour monitoring techniques, *Journal of Rock Mechanics and Geotechnical Engineering* 6(2) (2014) pp. 138-149.
doi: https://doi.org/10.1016/j.jrmge.2014.01.007

[18] R. R. Avent, M. Alawady, Bridge scour and substructure deterioration: Case study, *Journal of Bridge Engineering* 10(3) (2005) pp. 247-254.
doi: https://doi.org/10.1061/(ASCE)1084-0702(2005)10:3(247)

[19] S. Foti, D. Sabia, Influence of foundation scour on the dynamic response of an existing bridge, *Journal of bridge engineering* 16(2) (2011) pp. 295-304.
doi: https://doi.org/10.1061/(ASCE)BE.1943-5592.0000146

[20] B. E. Hunt, Monitoring scour critical bridges: a synthesis of highway practice, *The National Cooperative Highway Research Program Synthesis Report,* 2009.

[21] L. J. Zabilansky, Ice force and scour instrumentation for the White River, Vermont, *Cold Regions Research and Engineering Lab Hanover N,* US Army Corps of Engineers, USA, 1996.

[22] A. Zarafshan, A. Iranmanesh, F. Ansari, Vibration-based method and sensor for monitoring of bridge scour, *Journal of bridge engineering* 17(6) (2012) pp. 829-838.
doi: https://doi.org/10.1061/(ASCE)BE.1943-5592.0000362

[23] J. D. Schall, G. R. Price, Portable scour monitoring equipment (Vol. 515), *Transportation Research Board,* 2004.
URL http://www.trb.org/Publications/Blurbs/154328.aspx

[24] N. Anderson, A. Ismael, T. Thitimakorn, Ground-penetrating radar: a tool for monitoring bridge scour, *Environmental & Engineering Geoscience* 13(1) (2007) pp. 1-10.
doi: https://doi.org/10.2113/gseegeosci.13.1.1

[25] D. S. Mueller, M. N. Landers, Portable instrumentation for real-time measurement of scour at bridges (No. FHWA-RD-99-085), *Federal Highway Administration*, United States, 2000.

[26] L.J. Prendergast, K. Gavin, Monitoring of scour critical bridges using changes in the natural frequency of vibration of foundation piles-A field investigation, *In Transport Research Arena,* Paris, France, 2014.

[27] N. E. Yankielun, L. Zabilansky, Laboratory investigation of time-domain reflectometry system for monitoring bridge scour, *Journal of Hydraulic engineering* 125(12) (1999) pp. 1279-1284.
doi: https://doi.org/10.1061/(ASCE)0733-9429(1999)125:12(1279)

[28] W. Zheng, Instrumentation and computational modeling for evaluation of bridges substructures across waterways (No. FHWA/MS-DOT-RD-13-229). *Department of Transportation,* Mississippi, USA, 2013.

[29] M. Brinissat, R. Kuti, Z. Louhibi, Dynamic seismic analysis of bridge using response spectrum and time history methods, *Acta Technica Jaurinensis* (2021) pp. 171-185.
doi: https://doi.org/10.14513/actatechjaur.00595

[30] Y. B. Lin, F. Z. Lee et al., The artificial intelligence of things sensing system of real-time bridge scour monitoring for early warning during floods, *Sensors* 21(14) (2021) pp. 1-18.
doi: https://doi.org/10.3390/s21144942

Acta
Technica
Jaurinensis

Vol. 14, No. 4, pp. 651-680, 2021
DOI: 10.14513/actatechjaur.00625

Material Models to Study the Effect of Fines in Sandy Soils Based on Experimental and Numerical Results

A. Alnmr[1,*]

[1] Széchenyi István University, Department of Structural and Geotechnical Engineering
Egyetem tér 1, 9026 Győr, Hungary
*e-mail: ammar888999@hotmail.com

Submitted: 18/05/2021; Accepted: 24/11/2021; Published online: 24/11/2021

Abstract: Choosing and calibrating a robust and accurate soil material model (constitutive model) is the first important step in geotechnical numerical modelling. A less accurate model leads to poor results and more difficulty estimating true behaviour in the field. Subsequent design work is compromised and may lead to dangerous and costly mistakes. In this research, laboratory experimental results were used as a basis to evaluate several soil material models offered in Plaxis2D software. The deciding feature of the soil model was how well it could represent effects of percentage of fine material within sandy soils to simulate its behaviour. Results indicate that the Hardening Soil (HS) model works well when the percentage of fine (soft) materials is less than 10%. Above that level, the Soft Soil model (SS) becomes the most suitable. Finally, some important conclusions about this research and recommendations for future research are highlighted.

Keywords: sandy soil, Sand, fine material, consolidation, constitutive model

1. Introduction

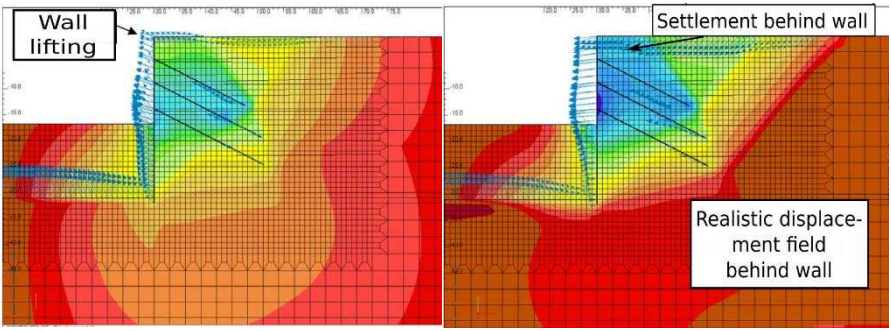
In numerical calculations, the relationships between stresses and strains in a given substance are represented by a constitutive model, which consists of mathematical expressions that model the behaviour of the soil [1]. The ideal constitutive model

would faithfully represent soil behaviour in the field, require a small number of parameters gathered from field and laboratory tests, and require minimum computational effort. While no single ideal model exists, a more practical approach is to choose a particular model that performs well under a more limited set of conditions. Such an approach has led to the development of a set of 10-15 material models commonly used in finite element software. Finding the best model for a set of conditions is not always a simple exercise and often requires some trial and error, as well as comparison between different model types.

Finite element analysis has become the de-facto numerical tool to investigate a wide array of geotechnical engineering problems, but the quality of any numerical prediction directly depends on the chosen material model and its required input parameters. The best models would not only predict behaviour previously measured in the field or lab, but also yield reasonable results for a wide variety of design alternatives, loading conditions, and variations in soil lithology and groundwater conditions. These predictions should be useful for considering both serviceability and limit states [2] [1].

While simpler models such as elasto-plastic with Mohr-Coulomb failure condition may be helpful in predicting limit states and estimating movement, many problems require more complex soil behaviour. This becomes especially true when there may be sequences of loading and unloading, construction stages, consolidation, and creep effects that impose very complex loading, drainage, and deformation conditions on a site [2]. Depending on these stages, the soil may compress or expand, its stiffness may vary due to changes in confinement, creep, or shear strain. In fact, the behaviour of the soil is elastic only at a level of very small strains. At slightly higher strain levels, the soil will exhibit a nonlinear stress-strain behaviour.[3] [1].

In contrast to the Mohr-Coulomb (MC) model, the Hardening Soil (HS) and Soft Soil (SS) models [Plaxis Ref] allow for stress dependency of stiffness as well as reduction in stiffness due to shear strain and more sophisticated treatment of dilatancy and yield. The HS and SS models offer a large choice of defining parameters, based on either laboratory testing or back calculated from field experience. The differences between the HS model and the simple MC model can be illustrated by a FE model by Obrzud [5] solving a benchmark excavation problem in Berlin sand from [Schweiger]. Figure 1 shows the displacement fields for an MC model (1a) and HS model (1b). The MC results show an over estimation of elastic rebound and a more general displacement field behind the tie-back wall. The HS results highlight a much lower rebound, more localized movement within the tieback zone and greater vertical settlement at the ground surface directly behind the wall. These estimates were consistent with observations and monitoring results.



(a) (b)

Figure 1. Comparison of numerical predictions of horizontal displacements of excavation wall in Berlin sand (a) MC model. (b) HS model [4]

While the qualitative effect percentage of silt content in sand has been known since Terzaghi, quantifying that effect has proved more elusive. Thevanayagam [5] presented a structured approach by considering the silty sand as a balanced matrix consisting of two submatrices: the coarser-grain matrix, and the finer-grain matrix, and analyzing how they coexist within the overall framework of particle-particle contact and transfer of forces through the soil. He and other authors Thev et al [6] , Lade [1], Sibley and Polito [7], viewed the behavior as bridging between sand silt where a fairly narrow transition zone existed between the end states. The location of a threshold fines content would dictate when the transition occurred. These authors also studied various engineering properties such as compressibility, shear strength, cyclic stability and pore pressure generation.

It is worth noting that sand does not exist without other components, but it contains a percentage of fine materials within it, and according to it, the sandy soils has been divided into zones which are zone1, transitional zone and zone2 [8]. Zone 1 is the zone in which fine materials partially fill the voids between the sand grains, meaning that the voids between the sand grains are not completely filled and the sand grains are in contact with each other, while transitional zone which is the zone that gives the lowest values of compressibility. It was found that the percentage of fine material in transitional zone for poor-graded sandy soils ranges between 10-15%, which is the soil used in this research [8]. transitional zone varies according to the grain gradation [8] [9], particle shape and size of sands [10]. In well-graded sand, the percentage of fine material in transitional zone may range between 15-30% [9]. At transitional zone fine material fills all the voids between sand grains and sand grains

remain in contact with each other, and thus it plays a major role in the behavior of sandy soils. While zone 2 is the zone where fine materials fill the spaces between the sand grains and separate them from each other, which means that the sand grains are not in contact with each other and the fine materials control the behavior in this region,

Adding various amounts of silt to the host sand causes a noticeable strain-softening response in the behavior even at high relative densities, this difference is represented by the reduced strain-softening behavior and the shift of the steady-state line. However, silty sands present a more tendency to flow contrasted with clean (pure) sand [11] [12]. And as indicated by Porcino et al [13], the addition of fines fundamentally changes the undrained monotonic stress-strain response of sand when tested at a constant void ratio. Up to fines content located at transitional zone, the behavior of the sand–silt blend turns out to be more contractive and strain softening gets more noticeable with increasing fines content, obliging a decrease in the peak and steady-state strengths. Thus, increasing the percentage of fine materials in the sandy soils will affect the choice of the constitutive soil model that simulates its actual behavior. The current study will seek to find out the optimal constitutive soil model according to the percentage of fine materials in sandy soils.

In this paper, both the HS and SS models will be adopted as being advanced models besides that it is easy to get their parameters from the direct shear and consolidation tests. The HS model has proven efficient in simulating soil behavior and has been used in a variety of research [14][15][16][17][18] [19]. The SS model is one of the models capable of simulating the behavior of soft soils with high efficiency. It is a model with better features than the MC model, but it has fewer features than the HS model (Table 1), but it is a specially developed material model for soft soils based on the MCC (Modified Cam Clay) model [20], According to Likitlersuang et al [21], analysis of SS model and HS model with soil parameters determined from laboratory and in situ tests provided better agreement with sidewall movements and field observations of surface settlement.

The magnitude of soil deformations in the HS model can be modelled more precisely by combining three different stiffness parameters taken at a specific reference stress Fig. 2:

1. The triaxial loading stiffness (E50)
2. The triaxial unloading-reloading stiffness (Eur)
3. The oedometer loading modulus (Eoed)

The set of parameters entered into the HS model allows the user to distinguish between loading and unloading - reloading stiffnesses for which a typical ratio is around $E_{ur} / E = 3\text{-}10$ as the ratio for compression indices $Cc / Cs = 0.1\text{-}0.4$ measured in consolidation tests [4] .

Table **1**. *Key features of the MC, SS and HS constitutive models* [20]

	Constitutive model		
Model feature	*Mohr-Coulomb Model*	*Soft-Soil Model*	*Hardening-Soil Model*
Non-linear stiffness	X*	X	X
Stress-dependent stiffness		X	X
Different stiffness for loading/unloading		X	X
Associated flow	X	Cap	Cap
Non-associated flow	X	MC	Cone, MC
Stress history effect		X	X
Volumetric hardening		X	X
Deviatoric hardening			X

* Only bi-linear MC: Mohr Coulomb failure surface Cap: Cap yield surface in SS and HS
Cone: Deviatoric hardening conical yield surface in HS

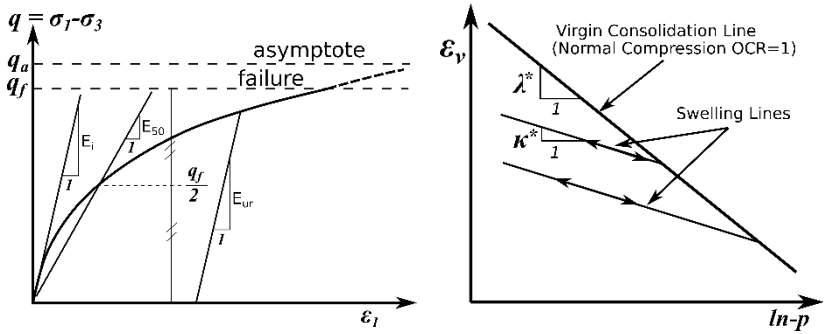


Figure 2. Definitions of moduli for triaxial (left) and oedometer (right) tests [4]

Despite the mathematical complexity of the HS model, its parameters can be determined by conventional soil experiments or can be guessed by following the geotechnical evidence [4]. The HS model represents soil deformations realistically, and for more details about this model go back to [4] [22].

Fig. 2 also shows how to obtain the parameters: λ^* (modified compression index that determines the compressibility of soil in loading) and K^* (modified swelling index that determines the compressibility of soil in unloading-reloading).

The importance of this research lies in facilitating the task for the researcher to propose the most appropriate model for sandy soil according to the percentage of fine materials within it in order to reach more accurate and reliable results. Whereas,

engineers looking for realistic predictions of soil response should be aware that by applying the elastic linear model in finite element analysis and not being bound by the limits and constraints of the model used, the soil movements may be mistaken in the guesswork, which in turn affects the amount of forces calculated to support the structural elements.

The research mainly aims to conduct an experimental numerical study of sandy soils mixed with different percentages of fine materials (Silt) in order to determine the optimum constitutive model that simulates the behavior of sandy soils.

2. Research materials and methods

In this research, the analytical experimental methodology was followed, whereby homogeneous samples of sandy soil consisted of different percentages of fine materials mixed with poorly-graded sand, were formed in order to conduct laboratory consolidation experiments on them. Finally, consolidation experiments were modelled using Plaxis2D software, using two constitutive models SS and HS.

The fine marine sand was brought from Sanawbar Jableh site in Lattakia, it was wet sieving on the N200 sieve to ensure that it is free from fine materials and obtaining completely pure sand.

While the silty soil was brought from the youth housing site in Latakia. The fine material used in this research was separated from the course material by wet sieving on the N200 sieve. Wet sieving was used because the silty soils agglomerate in their dry state therefore separating fine material with dry sieving will be difficult and will cause a lot of dust.

2.1. The laboratory work

Mixtures of sand- fine material were prepared based on dry weight. The fine marine sand was mixed with different percentages of fine materials 5, 10, 15, 20, 25, 30, 35, 45%. Laboratory experiments were performed on mixtures in the laboratories of the Faculty of Civil Engineering at Tishreen University in Lattakia, as follows:

1. Grain size distribution experiments were carried out according to ASTM D6913 [24] for sand and ASTM D7928-17 [25] for fine material, the granular gradient curves are shown in Fig. 3. We note from Fig. 3 that fine material consists of 38% of clay and 62% of silt. Table 2 shows the values of the coefficient of uniformity (Cu) and the coefficient of curvature (C_z) of the mixtures.

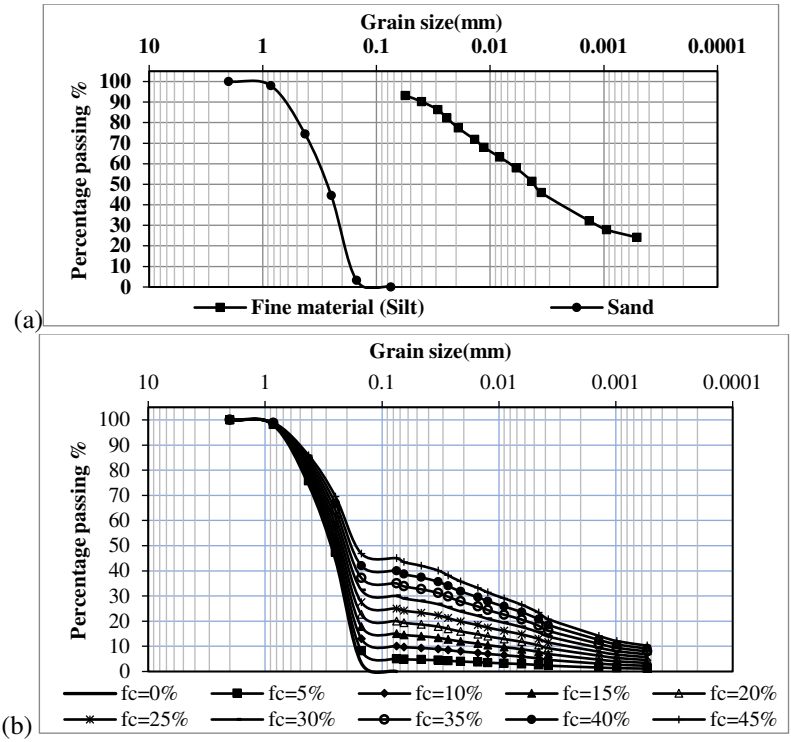


Figure 3. Grain Size Distribution experiments (a)- for sand and silt (b)- for mixtures

2. Specific gravity weight (G) tests were conducted according to ASTM D854-14 [26], and the results are presented in Table 2.

Table 2. Coefficient of uniformity (C_u), coefficient of curvature (C_z) and Specific gravity weight (G) for mixtures

Percentage of added fine materials[%]	***Specific gravity weight (G)***	***coefficient of uniformity (C_u)***	***coefficient of curvature (C_z)***
0	2.650	1.72	0.79
5	2.652	1.88	0.83
10	2.654	4.00	1.60
15	2.656	18.71	7.21
20	2.658	59.83	24.73
25	2.659	74.29	35.60
30	2.661	119.05	10.71
35	2.663	167.83	9.47
45	2.666	363.64	1.27

3. Atterberg limits experiments were conducted for fine material (according to ASTM D4318-17e1 [27]), and it was found that the plasticity index was 26% and the liquid limit was 57%, so it was classified as MH (High Plasticity Silt) according to the Unified Soil Classification System [28].

4. Maximum and minimum void ratio - Experiments were conducted according to the specification (ASTM D4254-00) [29], where the dry soil in its loosened state was filled in a mold size of 2825cm^3 three times and then weighed to calculate the maximum void ratio (e_{max}) and the average results were adopted for each percentage of fine material. After the mixture has been prepared into the mold it was placed on the shaking table and the table was left to vibrate for at least 2 minutes until the height of the soil in the mold became stable and unchanged. After that, the new height of soil in the mold was measured and the minimum void ratio (e_{min}) was calculated. Fig. 5 shows the change of the maximum and minimum void ratio with an increase in the percentage of fine materials.

   Fig. 4- (b) shows the results of e_{max} and e_{min} for different studies which used different types of sands, where the results of the experiment show the existence of three distinct zones [8] [9]. At the transition zone, the fine materials fill the entire voids between the grains of sand and thus the void ratios are at the lowest values and thus lowest compressibility [8] [9]. This transition zone depends on grain size distribution, packing density, size and shape of the sand and it falls within the range of 10-30% of added fine materials as shown in Fig 5- (b), as a result, the fine materials have an important role in controlling sandy soil behavior.

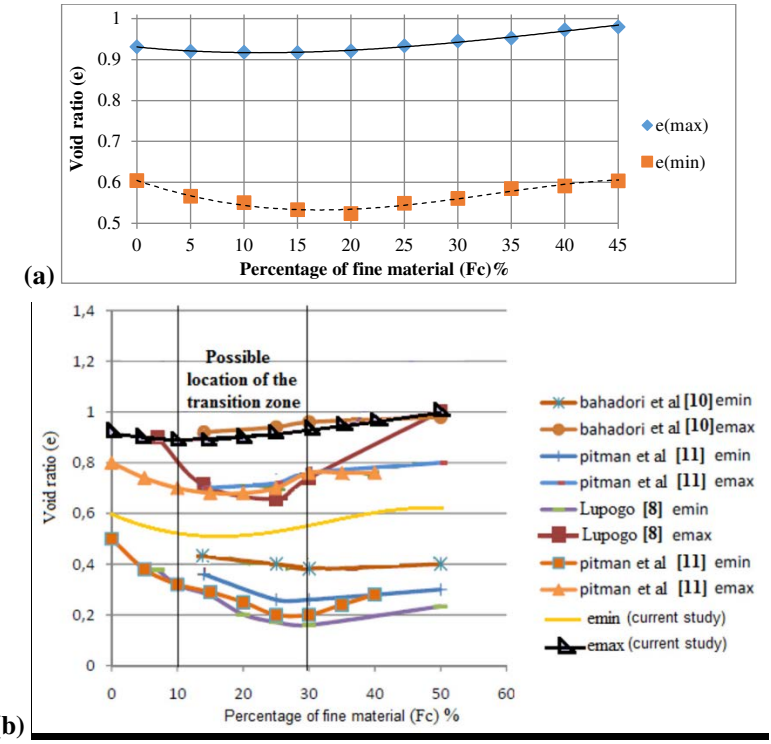


Figure 4. Curves of the relationship between the void ratio and the percentage of fine materials used in this study,(a)- current study (b-) comparison with other studies (adapted from Lupogo [30] *)*

.

5. The oedometer tests were conducted according to ASTM D2435 / D2435M – 11 [31], where, at first, mixtures of sand - silt were prepared on the basis of dry weight, and for each percentage of silt the required weight of silt and sand were found and mixed manually until it became homogeneous, thereafter the required water percentage was added then it was left for 24 hours in a tightly closed plastic bag in an isolated place to achieve homogeneous moisture in the sample, after the expiration of this period it was mixed manually again for a period of 15 minutes, thus achieving the desired satisfactory homogeneity of the samples as possible. Fig. 5 shows some pictures of laboratory samples with different percentages of silt. The vertical stresses were applied according to the following

sequence: 0.25, 0.50, 1, 2, 4, 6, 8, 6, 4, 6, 8, 12 kg/cm^2. Fig. 6 shows the consolidation test device.

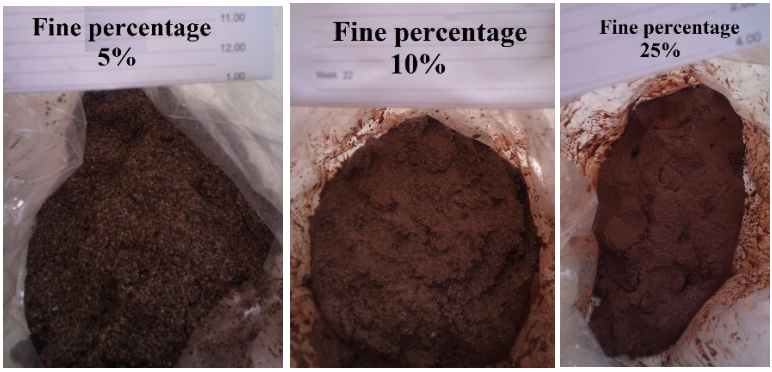


Figure 5. Pictures of some laboratory samples with different percentages of silt

Figure 6. Consolidation test device

6. Direct shear experiments were also conducted for all percentages of silt adopted in this research. Fig.7 shows the direct shear test device. Table 3 shows the physical properties of samples used in the direct shear and consolidation tests. The mixtures were tested with a relative density of 50% and initial moisture of 24% close to the saturation moisture.

Figure 7. Direct shear device

Table 3 . The samples used in the direct shear and consolidation experiments.

Sample symbol	Percentage of fine materials (Fc) %	e_{max}	e_{min}	Moisture [%]	Dry density [gr/cm^3]	Void ratio [e]	Relative density %
S1	**0**	0.931	0.604	24	1.499	0.768	50
S2	**5**	0.920	0.566	24	1.521	0.743	50
S3	**10**	0.918	0.550	24	1.531	0.734	50
S4	**15**	0.917	0.533	24	1.540	0.725	50
S5	**20**	0.921	0.524	24	1.543	0.722	50
S6	**25**	0.933	0.549	24	1.527	0.741	50
S7	**30**	0.945	0.560	24	1.518	0.753	50
S8	**35**	0.952	0.585	24	1.506	0.769	50
S9	**45**	0.979	0.604	24	1.488	0.791	50

Plaxis 2D finite element software was utilized to simulate the consolidation phenomenon in the oedometer test. An axisymmetric model having dimensions

similar to the dimensions of a ring of oedometer test used in this study. A very fine mesh was used for greater precision during matrix solving. Total fixities for the bottom bound of the model and horizontal fixities for the side borders of the model are used. The vertical distributed load (P) was applied over the entire length of the top surface to represent the phases of the load during the loading and reloading process. The default value for this load is 1 kPa, and it was changed during the deactivation stages of the calculation phases**.** Fig. 8 shows the model prepared in Plaxis and the mesh of finite elements, the distributed load, the boundary conditions, and geometric dimensions. To simulate the loading steps that were used in the laboratory, the vertical load was determined in the calculation program as different phases. A point was identified on the top surface to track the load-displacement behavior. In order to represent the loading and reloading stage in the calculation program in Plaxis, the calculation phases allow the required (applied) load to change during the load activation before the phase is updated. This means that for the next stage, the applied load can easily be changed (increased or decreased), as the method of modelling the loading stages in this paper is similar to the method that followed by Aldefae [19].

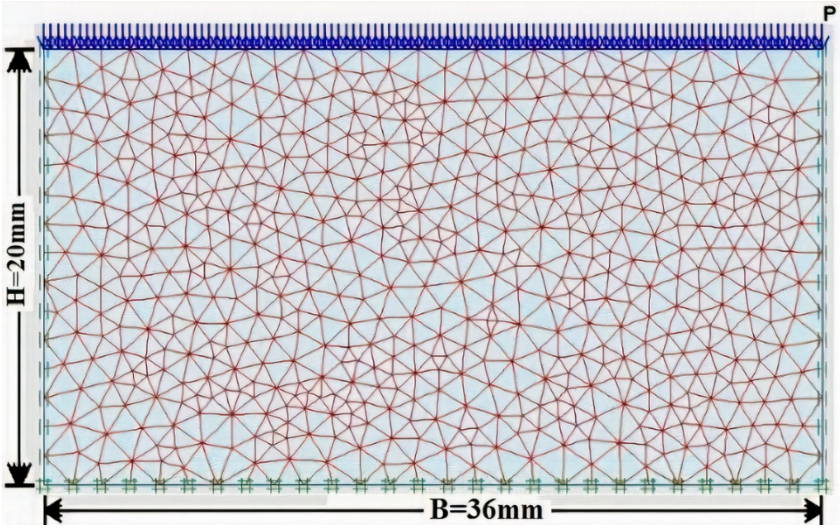


Figure 8. Geometry and finite element mesh for one-dimension consolidation problem

Fig. 9 shows the used phases during calculation processing as it shown the calculation type is Plastic analysis because we consider that the behaviour of soil as drained and the attention is given for final settlement of each loading step only, since the behaviour of sandy soil is drained, as it is known. Every loading step starts after finishing the previous one, exactly as happening in consolidation (oedometer) test.

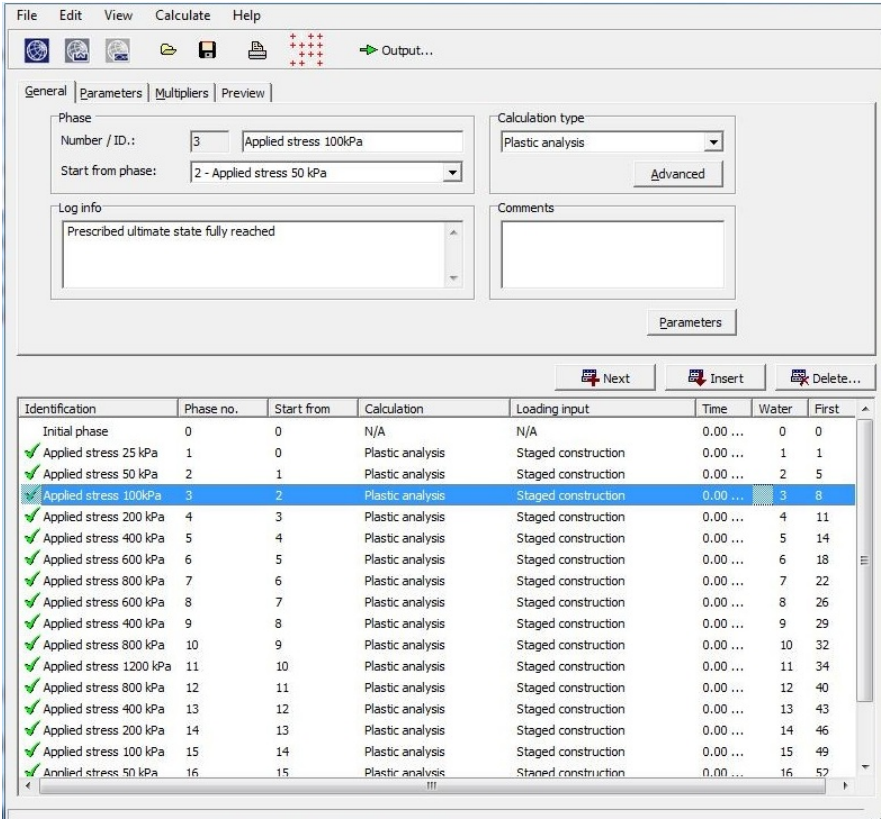


Figure 3. Calculation phases

3. Results and discussion

3.1. Experimental results

Fig. 10 shows the relationship between the void ratio and the vertical effective stress of the tested samples according to the percentages of added fine materials. It appears from the figure that the inclination of the compression line increases with the increase in the content of added fine material.

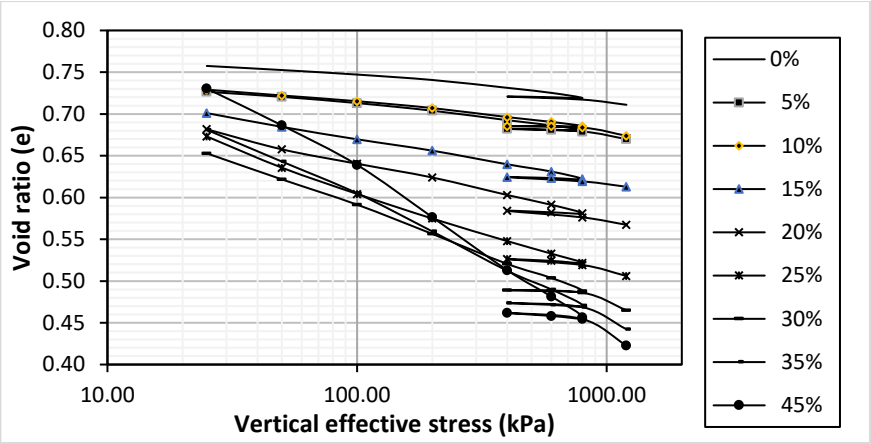


Figure 10. Curves of the relationship between the void ratio and the vertical effective stress of the tested samples according to the percentages of added fine materials

The values of the compression index (Cc) in this study for different effective stresses ranged between 0.018-0.211. And the values of the measured sand-silt compression index are in good agreement with the results of Mesri and Vardhanabhuti [32] & Monkul and Ozden [33].

the compression index (Cc) is the slope of the linear portion of the curve after pre-consolidation stress, it is calculated using the equation (1).

$$C_c = \frac{e_0 - e_1}{\log\left(\frac{\sigma'_1}{\sigma'_0}\right)} \quad (1)$$

Fig. 11 shows compression index values at each monotonic loading increment for every specimen. The horizontal axis represents the average vertical stress between load increments. There is a definite transition in Cc between low and high fine percentages. The (0, 5, 10%) specimens show identical behaviour until the highest stress level. This may be due to fine fractions carrying some of the internal stress at this level. The fines are more compressible compared to 0% where grain interlocking may be stabilizing. The next three specimens (15, 20, 25%) show progressive softening with fine content. The softening reduces at higher stress perhaps due to the sand matrix being more dominant in certain zones. Finally, the last three (30, 35, 45%) show nearly identical behaviour with overall softening increasing directly proportional to fines content. Therefore, there is a definite "sand zone" and a definite "fines zone" with a less well-defined "transition zone". This evolution of behaviour with fines content has been observed by previous researchers [30] [33] [34].

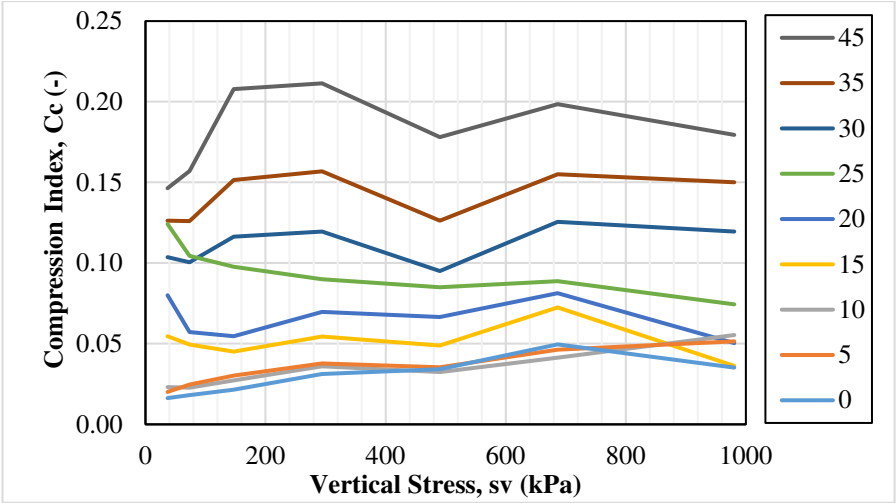


Figure 11. The relationship between the compression index and vertical stress for specimens with increasing percentage of fine materials.

Monkul and Ozden [33] presented an important concept about the transitional percentage of added fine material after which the separation between sand grains occurs. According to this concept, they assumed that the separation of coarse grains may occur when the intergranular void ratio (e_s) of the mixture exceeds the maximum void ratio of the remaining sand matrix.

The intergranular void ratio concept initially assumes only the sand matrix is carrying load. That assumption means for load carrying purposes, the fines matrix is first considered as voids since they do not contribute to carrying load. As percentage of fines increases, this assumption is no longer valid and the impact of fines on the compressibility of sand silt mixture can be quantified. The intergranular void ratio can be written as shown in the equation (2) [33].

$$e_s = \frac{e+\frac{G*f_c}{G_f*100}}{\frac{G}{G_s}*\left(1-\frac{f_c}{100}\right)} \tag{2}$$

Where: G_s *and* G_f are the specific gravity of sand and fine material, respectively. G is the specific gravity of the soil mixture itself.

It is shown from Fig. 12 that the line of the maximum void ratio of sand (e_{max}=0.931) intersects the curves of the relationship between es and percentage of added fine material according to different levels of stresses at the percentage of fine material between 11 to 17%, and this indicates that the contact between the coarse

grains under the applied stresses occurs when the content of fine materials is between 11 to 17% and this proves that the percentage of fine materials that achieve the lowest compressibility increases with increasing the applied stress.

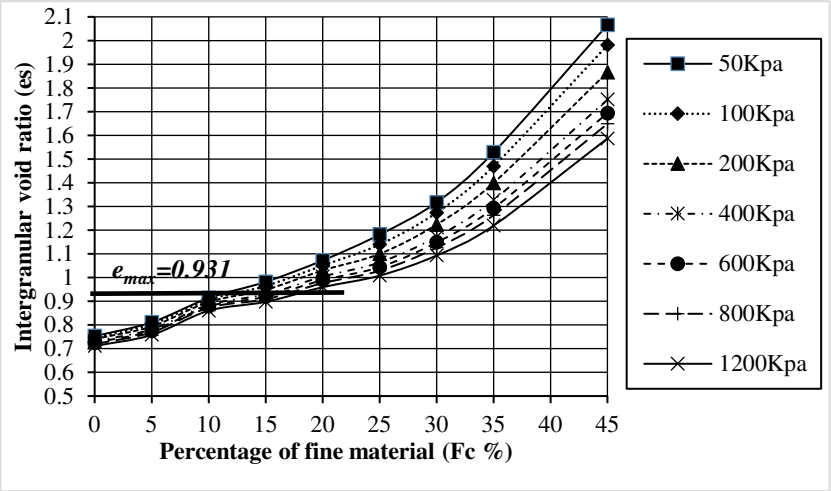


Figure 12. Relationship between intergranular void ratio and percentage of fine material

3.2. Numerical study

3.2.1. Method of fitting

The best fitting of the curve of the numerical modeling of consolidation test using SS model with the curve of laboratory consolidation test was achieved as follows:

1- The parameter $\boldsymbol{\lambda}^*$ was calculated for each percentage of added fine materials at three different ranges of stress, (50-100), (100-200) and (200-400)kPa to take uncertainty into consideration, while the parameter $\mathbf{K}^*$ was calculated from the reloading curve of an unloading-reloading cycle for example, for the percentage of added fine materials 35%, the values of $\boldsymbol{\lambda}^*$ and $\mathbf{K}^*$ are shown in Table 4.

Table 4. Parameters of λ^* and K^* for percentage of added fine materials 35%.

Range of stress (kPa)	$\mathbf{K}^*$	$\boldsymbol{\lambda}^*$
50-100	0.003	0.0273
100-200	0.003	0.034
200-400	0.003	0.038

2- For each range of stress, the curve of numerical modeling using SS model was created as shown in Fig. 13.

Finally, the parameters $\boldsymbol{\lambda}^*$ and $\mathbf{K}^*$ calculated from the considered range of stress, which achieved the best fitting of the curve of the laboratory test, were adopted. As it is seen from Fig. 13 the range of stress (50-100 kPa) achieves the best simulation of the curve of the laboratory test which is also proved in Table 5 (the last row), which shows the sum of the squared error of the three ranges of stress, and we note that this sum is the smallest for the range (50-100), so its parameters are adopted for the percentage of added fine materials 35%. The same methodology was applied to other percentages of added fine materials.

Table 5. The squared error of the three ranges of stress

б(kPa)	$\mathbf{T_{35}}$	$\mathbf{PS_{35,1}}$	$\mathbf{PS_{35,2}}$	$\mathbf{PS_{35,3}}$	$\mathbf{(PS_{35,1}-T_{35})^2}$	$\mathbf{(PS_{35,2}-T_{35})^2}$	$\mathbf{(PS_{35,3}-T_{35})^2}$
0	0	0	0	0	0	0	0
25	0.050	0.034	0.029	0.033	2.3E-04	4.1E-04	2.8E-04
50	0.071	0.078	0.071	0.082	4.2E-05	9.6E-08	1.1E-04
100	0.092	0.103	0.101	0.115	1.1E-04	6.9E-05	5.2E-04
200	0.118	0.126	0.127	0.145	5.3E-05	8.4E-05	7.4E-04
400	0.145	0.147	0.154	0.175	4.2E-06	7.4E-05	8.9E-04
600	0.158	0.160	0.169	0.192	3.9E-06	1.3E-04	1.2E-03
800	0.168	0.169	0.180	0.204	6.3E-10	1.3E-04	1.3E-03
600	0.168	0.168	0.179	0.203	2.3E-07	1.2E-04	1.2E-03
400	0.167	0.167	0.178	0.202	9.0E-08	1.2E-04	1.3E-03
800	0.170	0.169	0.180	0.204	3.0E-07	1.1E-04	1.2E-03
1200	0.185	0.182	0.195	0.222	9.0E-06	1.1E-04	1.4E-03
800	0.184	0.180	0.194	0.220	1.3E-05	1.0E-04	1.3E-03
400	0.182	0.178	0.192	0.218	1.2E-05	1.1E-04	1.4E-03
200	0.179	0.177	0.191	0.217	3.9E-06	1.4E-04	1.5E-03
100	0.176	0.175	0.189	0.215	5.6E-07	1.8E-04	1.6E-03
50	0.172	0.174	0.187	0.214	1.1E-06	2.3E-04	1.7E-03
25	0.168	0.173	0.186	0.213	1.6E-05	3.3E-04	2.0E-03
				Σ	0.0005	0.0024	0.0195

Where:

T_{35} is refer to the results of Laboratory test of 35% of added fine material.

$PS_{35,1}$ refers to the results of the numerical analysis by Plaxis using Soft Soil model of 35% of added fine material for the range of stress (50-100) kPa .

$PS_{35,2}$ refers to the results of the numerical analysis by Plaxis using Soft Soil model of 35% of added fine material for the range of stress (100-200)kPa.

$PS_{35,3}$ refers to the results of the numerical analysis by Plaxis using Soft Soil model of 35% of added fine material for the range of stress (200-400)kPa

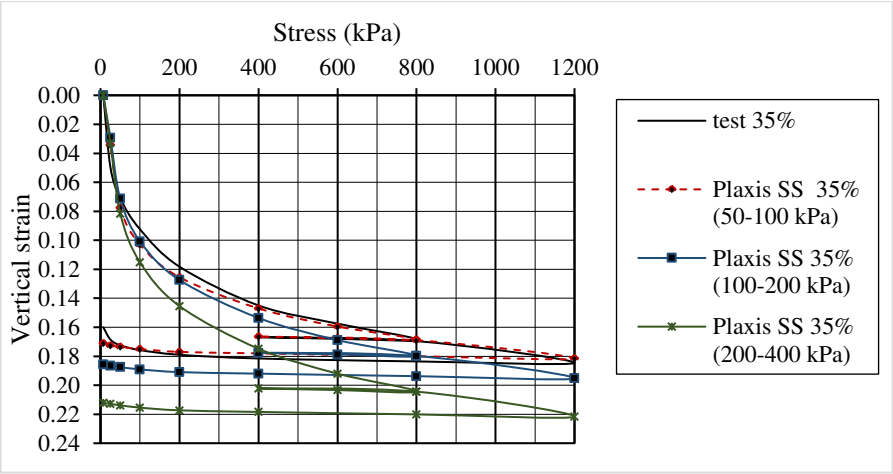


Figure 13. The curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using SS model to determine the best value of $\boldsymbol{\lambda}^*$ for the percentage of added fine material 35%

While, achieving the best fitting between the curves of stress- strain relationship for both laboratory one-dimension consolidation test and numerical modeling using HS model required more time and effort and it was done as follows:

1- The most appropriate reference stress (P_{ref}) that achieved the best fitting was determined. The corresponding stiffness parameters were calculated from curve of the relationship between stress and strain of consolidation test for each reference stress 50, 100, 200, 400 kPa where E_{oed}^{ref} calculated depending on loading curve, while the parameter E_{oed}^{ref} was calculated from the unloading curve for each reference stress, without changing the parameter m in this step, for example, the stiffness parameters corresponding to each reference stress 50-100-200-400 kPa for the percentage of added fine material 35%, shown in Table 6.

Fig. 14 displays the curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using HS model for various reference stresses P_{ref} for the percentage of added fine material 35%.

Table 6. Stiffness parameters of HS model for percentage of added fine materials 35% for various reference stresses Pref.

P_{ref}	m	E_{ur}^{ref} [kPa]	$E_{50}^{ref} = E_{oed}^{ref}$ [kPa]
50	0.7	10300	1750
100	0.7	23000	3200
200	0.7	52000	5700
400	0.7	98500	10200

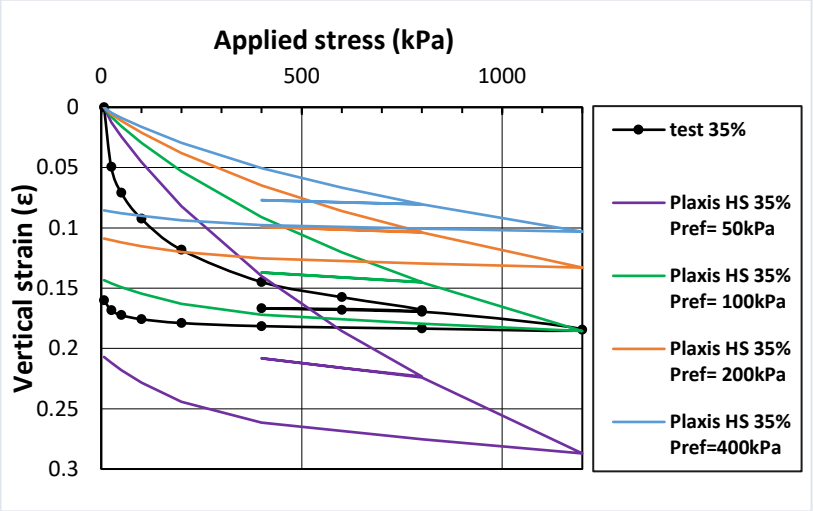


Figure 14. The curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using HS model at various reference stresses Pref for the percentage of added fine material 35%

2- The second step was determining the parameter m. It is noted from the Fig. 14 that the reference stresses P_{ref} that achieve the best fit are P_{ref} = 100kPa, so the reference stress P_{ref} will be fixed and change the value of m within the range from 0.5 to 0.9 and adopt the value of the parameter m that achieves the best fit for the laboratory one-dimension consolidation test.

Fig. 15 shows the numerical modeling curves of the HS model at different values of the parameter m for the percentage of added fine material 35%.

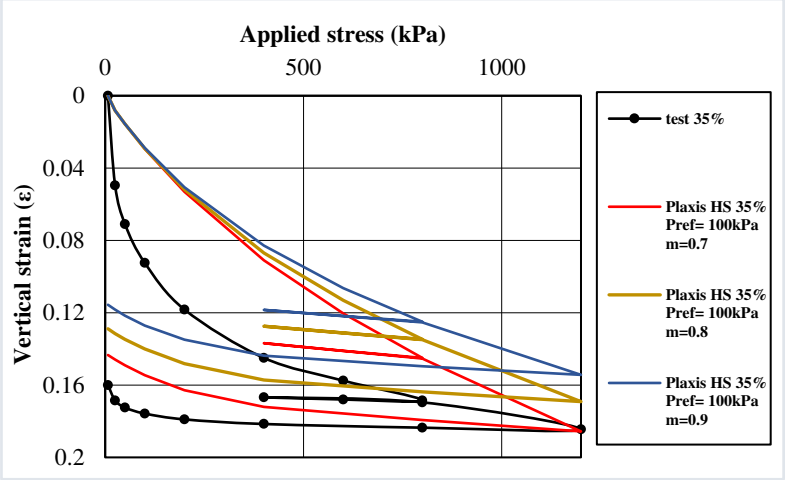


Figure 15. The curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using HS model at various values of the parameter m for the percentage of added fine material 35%

It is obvious from Fig. 15 that HS model does not simulate the laboratory one-dimension consolidation test for 35% of added fine material. The same methodology was applied to all percentages of added fine materials to determine the best parameters for HS model.

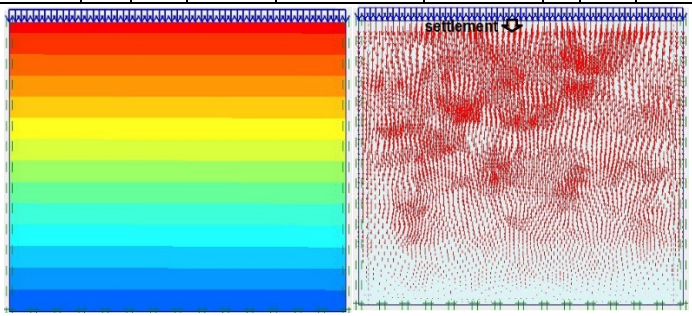
3.2.2 Numerical results

The above methodology was followed to determine the best parameters for both models HS and SS, based on conducted laboratory experiments. Table 7 shows the best parameters of both the HS and SS models which achieve the best simulation of the one-dimension consolidation experiments, which were extracted from the laboratory experiments conducted for each percentage of fine material.

The results of both the laboratory tests of the one-dimensional consolidation test and finite element with SS and HS models are plotted in a stress-strain relationship. Fig. 16 shows the vertical displacement at the end phase of the numerical test in both the rainbow shading Fig. 17 (a) and in the shape of the arrows Fig.16 (b). The maximum vertical settlement is at the top of the model while zero settlement is at the base of the model and this is identical to what should be in the consolidation test.

Table 7. Summary of parameters used for FE model

SS (SoftSoil)		***HS (HardeningSoil)***								
K*	**λ***	**P_{ref}**	**m**	E_{ur}^{ref} **[kPa]**	$E_{50}^{ref} = E_{oed}^{ref}$ **[kPa]**	γ_{sat} **[kN/m³]**	**Ψ [⁰]**	**ϕ [⁰]**	**C [kPa]**	**Fc [%]**
0.0023	0.0047	400	0.65	246100	45800	19.33	9	39	5	0
0.00184	0.006	400	0.7	271100	33200	19.48	8	38	15	5
0.00141	0.0055	400	0.82	296200	34000	19.54	6	36	23	10
0.0019	0.01	200	0.6	85000	12500	19.6	3	33	35	15
0.003	0.0141	100	0.6	51500	6900	19.62	0	25	44	20
0.002	0.0205	100	0.67	41000	4300	19.53	0	24.1	47	25
0.003	0.0273	100	0.7	23000	3200	19.4	0	21	52	35
0.0035	0.0295	100	0.9	12500	2400	19.3	0	17	55	45



(a) (b)

Figure 16. Vertical displacement of the model (a)- shadings and (b)- arrows.

Fig. 17 shows the curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using HS model. There is an almost identical agreement between the two curves when the percentage of added fine material ranges between 0 up to 10%, and after the percentage of 10% of added fine material the two curves move away from each other. Thus, the HS model gives the best results in sandy soils in which the percentage of fine materials does not exceed 10%, meaning that HS model is suitable for sandy soils located in zone 1.

Fig. 18 shows the curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using SS model as well. But on contrary to the result from Fig. 18, there is a great and almost identical agreement between the two curves when the percentage of added fine material exceeds 10%, whilst if the percentage of added fine material is less than 10% the two curves move away from each other. Thus, the SS model gives the best results in sandy soils in

which the percentage of fine materials exceeds 10%, meaning that SS model is suitable for sandy soils located in the Transition zone and zone 2.

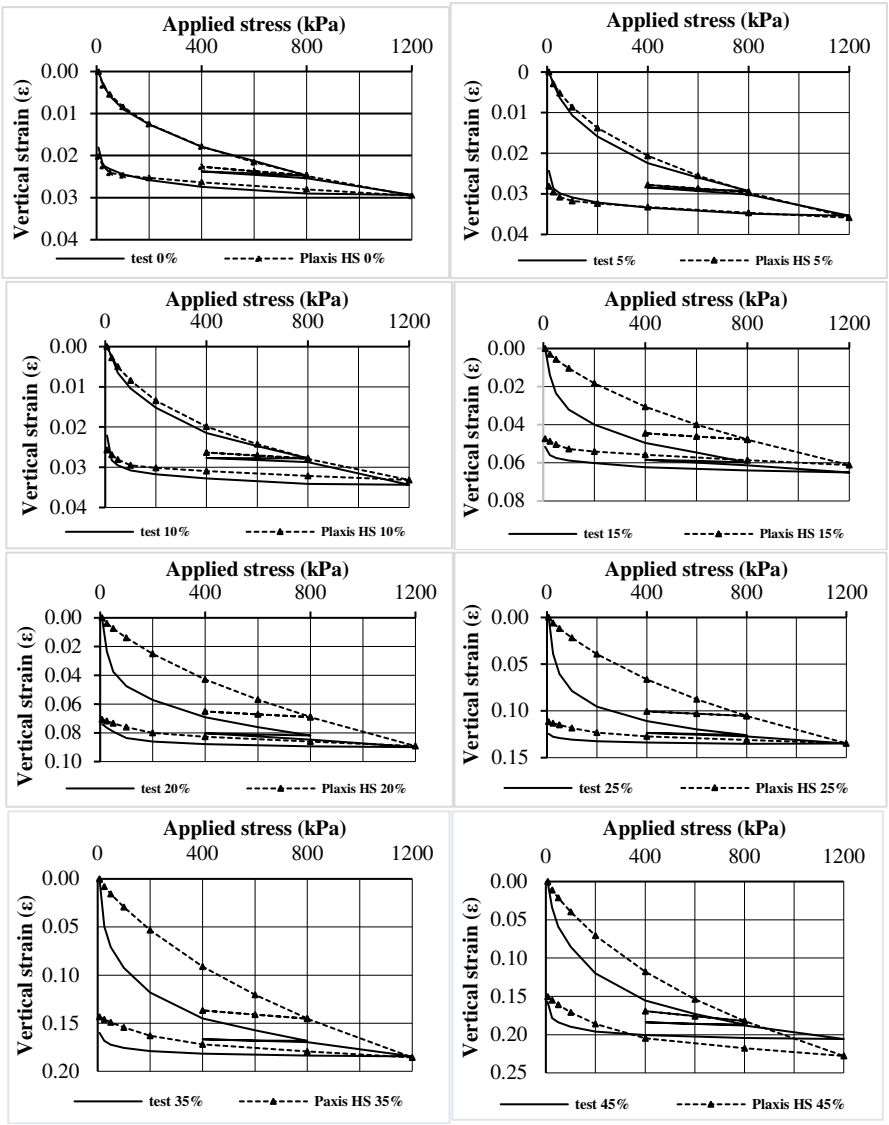


Figure 17. The curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using HS model

Where**: test 0%, test 5% etc** mean laboratory one-dimension consolidation test at 0%, 5% etc of added fine material respectively, **Plaxis HS 0%, 5% etc** and **Plaxis SS 0%, 5% etc** means numerical modeling by Plaxis software using HS model and SS model at **0% 5% etc** of added fine material respectively.

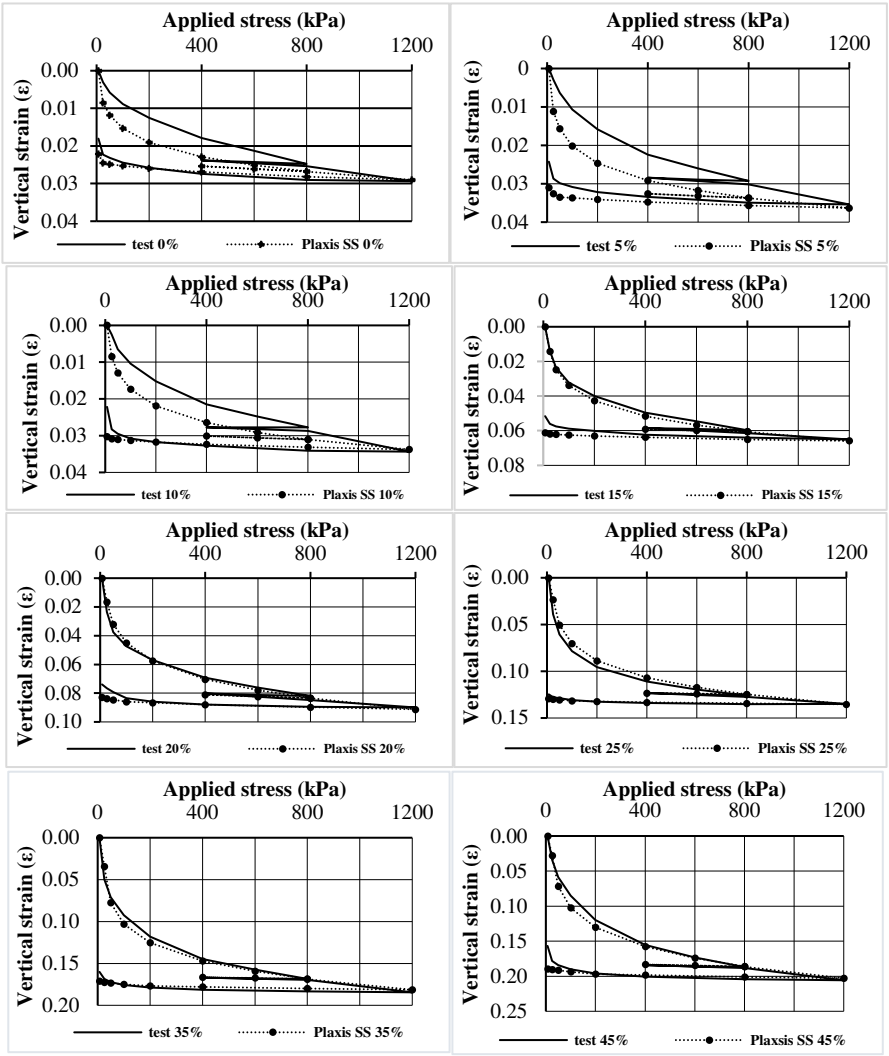


Figure 18. The curves of the stress-strain relationship for both laboratory one-dimension consolidation test and numerical modeling using SS model

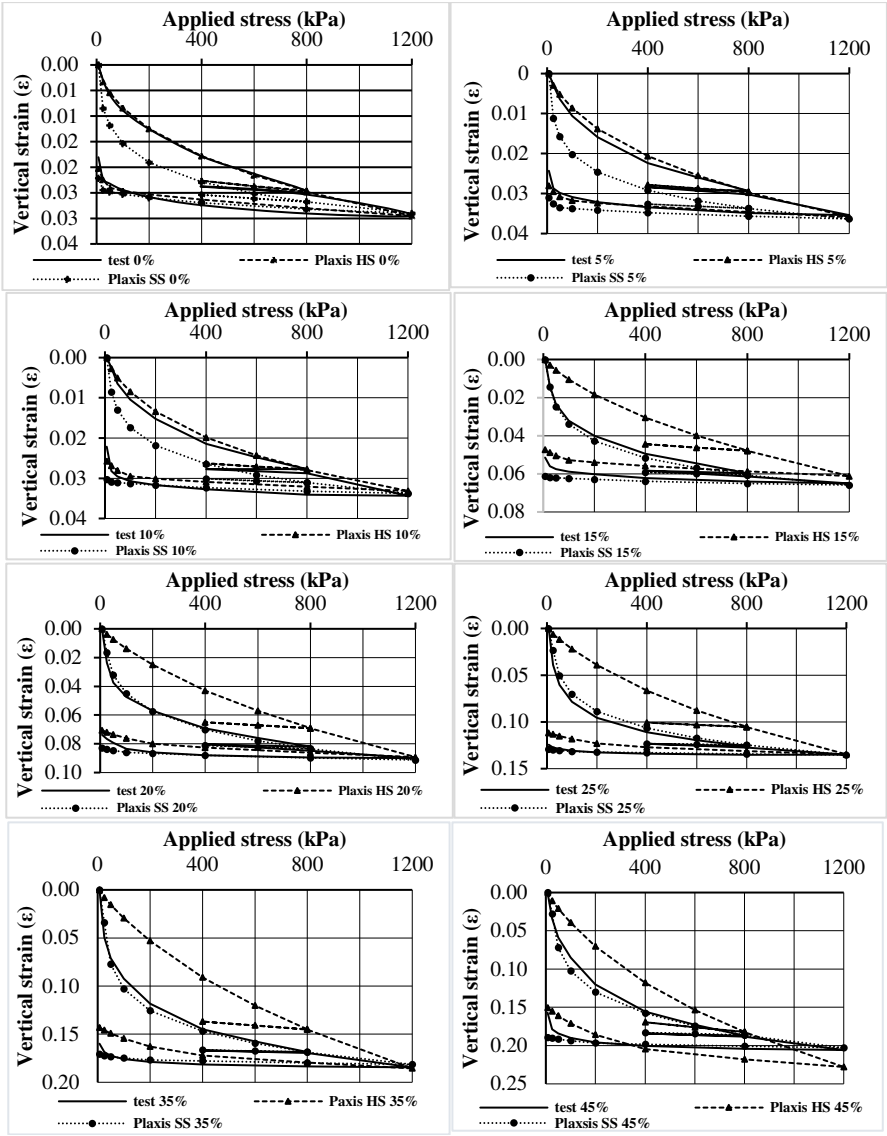


Figure 4. The curves of the stress-strain relationship for laboratory one-dimension consolidation test, numerical modeling using HS model, and numerical modeling using SS model

Fig. 19 collects the three curves together (laboratory one-dimension consolidation test, numerical modeling using HS model, and numerical modeling using SS mode) to clarify the difference between them. It is noted that the HS model does not agree with the SS model and SS model curve is always more concave than HS model curve.

4. Conclusions and recommendations

4.1. Conclusions

Based on the results of the laboratory tests, it was found that fine materials mixed with soft poorly-graded sand for a certain percentage of fine material(silt) greater than 10-15% increase the compressibility of the sand mixture and it is proven by conducting numerous consolidation tests on prepared samples reconstituted in the laboratory. In order to study the consolidation behavior of sand when the percentage of fine materials increase, and through this research, the following conclusions and final recommendations are reached:

1. The compressibility of sandy soils depends on the percentage of fine materials that fill the voids between their sand grains. The percentage of fine materials that achieve the lowest compressibility is in the range of 10-15%, which is affected by the level of applied stress.
2. The percentage of fine materials that achieve the lowest compressibility increases with increasing the applied stress.
3. The minimum values of e_{max} and e_{min} are related to the grain size distribution of the sand and the packing density. At transition zone, e_{max} and e_{min} have the minimum values due to the fact that the voids between the sand particles are completely filled with the fine materials. In zone 2, fine materials separate sand particles and make them move away from each other and void ratios start to increase again.
4. The transitional zone is affected by the grain distribution, size and shape of sand. For poorly graded sand, the transitional zone defined based on compressibility, is on a small range of fine material about (10-15%), while the transitional zone is on a wider range of fine material for other types of sand and fine material according to previous studies.
5. The HS model is most appropriate when the percentage of soft materials added is up to 10% (zone 1), then the SS model becomes the most suitable when the percentage of fine materials exceeding 10% (transition zone and zone 2).

4.2. Recommendations

1. Study the effect of relative density of the sandy soil, as in this research a relative density of 50% has been adopted considering sandy soil as medium dense.

Therefore, it is preferable to use other relative densities to study their effect on the material model selected.

2. Repeating the tests on other types of soil mixtures by experimenting with adding different fine materials (other than silt) to the sand used, such as clay, as well as using other types of sand in the tests (other than soft poorly-graded sands) such as coarse poorly-graded sands or well-graded sand with a varied grain gradient, with the aim of establishing a useful data bank that is not limited to a specific type of mixtures only, and thus determining the transition zones of various sandy soils.

Acknowledgement

The author would like to express thanks to the faculty of civil engineering at Tishreen University, Latakia, Syria, and especially for the department and the laboratory of geotechnical engineering.

References

[1] P. V. Lade, Overview of Constitutive Models for Soils, *Geo-Frontiers Congress,* Austin, Texas, United States, 2005, pp. 1–34. doi: *https://doi.org/10.1061/40771(169)1.*

[2] D. M. Potts, L. Zdravkovic, T. I. Addenbrooke, K. G. Higgins, and N. Kovacevic, Finite element analysis in geotechnical engineering, *Thomas Telford*, London, 2001, 440 pages. URL *https://www.icevirtuallibrary.com/doi/pdf/10.1680/feaigea.27831.fm.*

[3] T. I. Addenbrooke, D. M. Potts, and A. M. Puzrin, The influence of pre-failure soil stiffness on the numerical analysis of tunnel construction, *Geotechnique*, 47 (3) (1997) pp. 693–712. doi: *https://doi.org/10.1680/geot.1997.47.3.693.*

[4] R. F. Obrzud, On the use of the Hardening Soil Small Strain model in geotechnical practice, *Technical report Z_Soil.PC 100701, Zace Services Ltd*, Lausanne, 2010, 17pages URL *http://www.geomod.ch/pdf/zsday-hard.pdf.*

[5] S. Thevanayagam, Effect of Fines and Confining Stress on Undrained Shear Strength of Silty Sands, *Journal of Geotechnical and Geoenvironmental Engineering*, 124 (6) (1998) pp. 479–491. doi: *https://doi.org/10.1061/(ASCE)1090-0241(1998)124:6(479).*

[6] S. Thevanayagam, T. Shenthan, and T. Kanagalingam, Role of intergranular contacts on mechanisms causing liquefaction and slope failures in silty sands, In Final report, USGS Award No. 01HQGR0032 and 99HQGR0021. U.S. Geological Survey, Department of the Interior, Reston, Va, 2003, 396 pages.

[7] E. L. D. Sibley and C. P. Polito, Insights on Threshold Fines Content, *Geovirtual 2020 Resilience and Innovation Sep 14-16*, 2020, 7 pages. URL *https:// www.geovirtual2020.ca/wp-content/files/406.pdf.*

[8] A. N. Alnmr, M. Omran Alzawi, and S. Abdallah, Study the Effect of the Percentage of Fines on the Compressibility Behavior of Sandy Soils, *Tishreen University Journal for Research and Scientific Studies*, 4 (2) (2017) pp. 2079–3001. *(in Arabic)* URL *http://www.journal.tishreen.edu.sy/index.php/engscnc/article/view/3677/0.*

[9] K. Lupogo, Effect of fines mineralogy on the oedometric compressional behavior of sandy soils, *Journal of Civil Engineering and Construction Technology*, 4 (7) (2013) pp. 232–238. doi: *https://doi.org/10.5897/JCECT12.025.*

[10] D. Sarkar, M. Goudarzy, D. König, and T. Wichtmann, Influence of particle shape and size on the threshold fines content and the limit index void ratios of sands containing non-plastic fines, *Soils and Foundations*, 60 (3) (2020) pp. 621–633. doi: *https://doi.org/10.1016/J.SANDF.2020.02.006.*

[11] H. Bahadori, A. Ghalandarzadeh, and I. Towhata, Effect of non plastic silt on the anisotropic behavior of sand, *Soils and Foundation*, 48 (4) (2008) pp. 531–545. doi: *https://doi.org/10.3208/SANDF.48.531.*

[12] T. D. Pitman, P. K. Robertson, and D. C. Sego, Influence of fines on the collapse of loose sands, *Canadian Geotechnical Journal*, 31 (5) (1994) pp. 728–739. doi: *https://doi.org/10.1139/T94-084.*

[13] D. D. Porcino, V. Diano, T. Triantafyllidis, and T. Wichtmann, Predicting undrained static response of sand with non-plastic fines in terms of equivalent granular state parameter, *Acta Geotechnica* 15 (4) (2019) pp. 867–882. doi: *https://doi.org/10.1007/S11440-019-00770-5.*

[14] S. Likitlersuang, C. Chheng, and S. Keawsawasvong, Structural modelling in finite element analysis of deep excavation, *Journal of GeoEngineering* 14 (3) (2019) pp. 121–128.
doi: *https://doi.org/10.6310/jog.201909_14(3).1.*

[15] G. Song and E. X. Song, Selection of soil constitutive models for numerical simulation of foundation pit excavation, *Gongcheng Lixue/Engineering Mechanics* 31 (5) (2014) pp. 86–94.
doi: *https://doi.org/10.6052/j.issn.1000-4750.2012.08.0583.*

[16] K. H. Law, K. H. Geotechnical, and S. Bhd, 3D finite element analysis of a deep excavation considering the effect of anisotropic wall stiffness Impact, *19th Southeast Asian Geotechnical Conference & 2nd AGSSEA Conference (19SEAGC & 2AGSSEA)* , 2016, 6 pages.
URL *https://www.researchgate.net/publication/327867642.*

[17] W. Al-Ani, D. Wanatowski, and S. H. Chan, Numerical Analysis of Piled Embankments on Soft Soils, *Geo-Shanghai*, May 2014, pp. 30–39,
doi: *https://doi.org/10.1061/9780784413401.003.*

[18] B. C. B. Hsiung, K. H. Yang, W. Aila, and L. Ge, Evaluation of the wall deflections of a deep excavation in Central Jakarta using three-dimensional modeling, *Tunnelling and Underground Space Technology* 72 (2018) pp. 84–96.
doi: *https://doi.org/10.1016/j.tust.2017.11.013.*

[19] A. Hafudh and H. Aldefae, "Prediction of One-Dimensional Compression Test using Finite Elements Model," *International Journal of Engineering Research & Technology (IJERT)* 5 (05) (2016) pp. 359–363.
URL *http://www.ijert.org.*

[20] M. Karstunen and A. Amavasai, Soft soil modelling and parameter determination, Chalmers University of Technology, Department of Architecture and Civil Engineering Gothenburg, Sweden, 2017, 78 pages.
URL *https://research.chalmers.se/publication/522789/file/522789_Fulltext.pdf.*

[21] S. Likitlersuang, C. Surarak, D. Wanatowski, E. Oh, and A. Balasubramaniam, Finite element analysis of a deep excavation: A case study from the Bangkok MRT, *Soils and Foundations* 53 (5) (2013) pp. 756–773.
doi: *https://doi.org/10.1016/j.sandf.2013.08.013*.

[22] A. Truty and Z. Services, Hardening Soil model with small strain stiffness, Lausanne, Technical Report 080901, Zace Services Ltd, 2008, 42 pages. URL *http://www.zsoil.com/zsoil_day/2008/Truty_HS-model.pdf.*

[23] R. B. Brinkgreve, W. Broere, and D. Waterman, Plaxis 2D software manual version 8, Delft University of Technology and Plaxis Inc, Delft, The Netherlands, 2006.
URL *http://www.plaxis.nl.*

[24] ASTM International- Standard test method for particle-size analysis of soils,. ASTM D6913/D6913M-17, (2017)
doi: *https://doi.org/10.1520/D6913_D6913M-17.*

[25] ASTM International-Standard Test Method for Particle-Size Distribution (Gradation) of Fine-Grained Soils Using the Sedimentation (Hydrometer) Analysis, ASTM D7928-17 (2017).
doi: *https://doi.org/ 10.1520/D7928-17.*

[26] ASTM International- Standard Test Methods for Specifc Gravity of Soil Solids by Water Pycnometer, . D854-14 (2014).
doi: *https://doi.org/10.1520/D0854-14.*

[27] ASTM International- Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils, ASTM D4318-17e1 (2017).
doi: *https://doi.org/10.1520/D4318 -17E01.*

[28] ASTM International- Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System), ASTM D2487-17e1 (2017).
doi: *https://doi.org/10.1520/D2487-17E01.*

[29] ASTM International- Standard Test Methods for Minimum Index Density and Unit Weight of Soils and Calculation of Relative Density, ASTM D4254-00 (2000).
doi: *https://doi.org/10.1520/D4254-00.*

[30] K. Lupogo, Effects of fines on mechanical behaviour of sandy soils, master thesis, 2009, 136 pages.
URL *https://repository.tudelft.nl/islandora/object/uuid%3Af4b5b6d0-8b00-44da-8011-cd664bfa96bd.*

[31] ASTM International- Standard Test Methods for One-Dimensional Consolidation Properties of Soils Using Incremental Loading, ASTM D2435 / D2435M - 11 (2020).
doi: *https://doi.org/10.1520/D2435_D2435M-11R20.*

[32] G. M. Mesri and B. V. Vardhanabhuti, Compression of granular materials, *Canadian Geotechnical Journal* 46 (4) (2009) pp. 369–392.
doi: *https://doi.org/10.1139/T08-123.*

[33] M. Murat Monkul and G. Ozden, Compressional behavior of clayey sand and transition fines content, *Engineering Geology* 89 (3–4) (2006) pp. 195–205.
doi: *https://doi.org/10.1016/j.enggeo.2006.10.001.*

[34] S. Thevanayagam and S. Mohan, Intergranular state variables and stress–strain behaviour of silty sands, *Geotechnique* 50 (1) (2015) pp. 1–23.
doi: *https://doi.org/10.1680/GEOT.2000.50.1.1.*