Acta
Technica
Jaurinensis

Vol. 12, No. 2, pp. 82-100, 2019
DOI: 10.14513/actatechjaur.v12.n2.491
Available online at acta.sze.hu

Selection of Industrial Packaging System in Closed-loop Automotive Supply Chains

K. Vöröskői[1], P. Böröcz[1]

[1]Széchenyi István University, Department of Logistics and Forwarding
Egyetem tér 1., 9026 Győr, Hungary
E-mail: voroskoi.kata@sze.hu

Abstract: The corporate decisions in the field of packaging aim to find and determine the optimal packaging functions and expenses at the same time. One-way packaging is also used in the practice, but returnable packaging is more common in the outbound flow of the automotive engine manufacturer companies. In addition, it can happen that packaging devices originally designed for one way are used multiple times, because the technical condition of the packaging ensures the repeated usage and the other way around. Transport distance is an important aspect while choosing the right packaging system, but according to our investigation the ratio between the purchasing cost of a new device and the cost of backwards transportation should also influence this decision. Therefore, the scope of this paper is to give a model of the cost structure of returnable industrial packaging operating in a closed loop system. The total cost will be especially determined by the return ratio, cost of return transport and planned number of uses.

Keywords: returnable packaging; closed-loop; packaging cost

1. Introduction

The packaging design has a strategic impact on the efficiency of the supply chain (SC). Packaging links the entire supply chain and coordinates all participants in the process to give a flexible and effective response to customer needs in order to maximize satisfaction at optimal cost [1].

It has been found that paying limited attention to packaging can cause higher costs in the physical distribution. Furthermore, researchers argue that packaging should

not only be considered from the cost point of view, but focus should be put on its role as a value-added function in the SC [2].

Nowadays packaging designers are increasingly focusing on balancing the need for product protection, material use efficiency and the packaging material's impact on the environment along the whole supply chain [3].

Sustainability (green supply chain and environmental aspects) is closely related to packaging decisions in supply chain management [4]. Svanes' decision support methodology for sustainable packaging design rather focuses on packaging optimisation than packaging minimisation. It offers a tool box, but one-way vs. returnable decision is not mentioned. Instead of traditional cost calculations total distribution costs of packed product are considered. It includes the cost of packaging materials, cost of packing process, cost of transport from the producer via the wholesaler to the retailer, costs of handling by users along the distribution chain and cost of product loss [5].

According to Dominic et al. a sustainable packaging development model needs three variables: technical, supply chain and environmental design. This also considers the whole supply chain from the beginning till the end, reverse processes included. However, reuse only appears in context of waste handling and reduction. Packaging cost is merely one of the numerous indicators as part of the business perspective. Because of the discussion of corrugated box packaging in this model, development of a returnable system is not addressed here. [4]

Other returnable packaging management models compare returnable packaging management strategies, the focus is on transportation cost and inventory holding cost, furthermore the division of these among the SC stakeholders. At the same time, they neglect for example procurement or management costs [6].

Therefore, in the field of logistic packaging (industrial transportation, or even consumer packaging) the companies make decisions in order to determine the optimal packaging expenses. This decision-making situation practically means a choice between the one-way and reusable (disposable or returnable) packaging systems [7]. It is significant in the decision making process that which costs and environmental effects the disposable packaging has. The construction of returnable packaging is usually more complex. Numerous factors appear and each of them should be examined separately and then all together. As returnable packaging is planned for multiple uses, packaging material should be usually stronger and the whole design has to handle excessive forces.

2. Theoretical framework for automotive packaging cost evaluation

2.1. Disposable vs. returnable packaging

The best packaging solutions are those that can - beside the optimal cost levels – maximize the use of packaging space so that all the products can easily be packed and stacked and at the same time reduce packaging waste [8]. Besides the support of logistics handling and containment, product and environmental protection is also one of the most important functions of packaging [9].

The following two types of packaging systems are mostly used in automotive industry: one-way and returnable. One-way packaging is only suitable for one use. Reusable containers are loaded with products and shipped to the destination, then the empty container is sent back to the same supplier, refilled with products and this cycle is repeated over and over again in a closed-loop. (Fig 1) In some cases it is an open-loop system, when reusable packaging is collected at a centralised return handling centre, where it is cleaned, stocked, and distributed for refilling. [10]

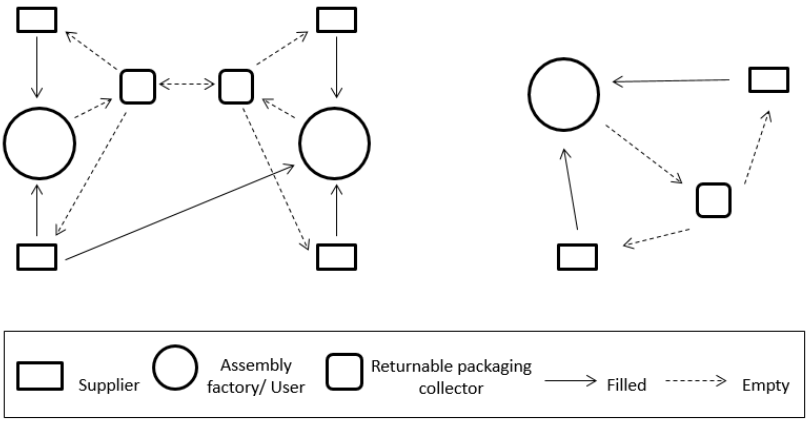


Figure 1. Open-loop (left) and closed-loop (right) returnable packaging system

The main problem with one-way packaging is the waste created after the usage, while relative production costs are lower. On the other hand, transportation and maintaining costs are a relevant issue in the case of returnable packaging [11].

In the automotive industry the primary function of packaging is the protection of products and parts optimised by the total costs of logistics [12]. This way, even if

shipping is performed on land or sea, by rail, trucks, vessels or multi-mode shipping, the distribution environment and logistics costs together define the possible form of packaging, and then determine the final solution from disposable to returnable packaging and systems [13].

Returnable packaging has been frequently used, for example in the US automotive industry, in order to reduce waste, costs, transport damages and to enable JIT deliveries [14]. Standardised shipment materials are usually used as returnable packaging in the automotive supply chain (ASC), like the EUR/-EPAL pallets, racks, containers and specialty bins for certain types of parts [15].

The elimination of waste caused by disposable packaging is one of the main environmental factors to consider when choosing a returnable packaging system. Furthermore, returnable packages may contribute to a better workflow, because they are often easier to handle [16].

Legislation has also forced companies to rethink their packaging operations [17]. The green actions inclues green procurement, green manufacturing, green packaging, recycling and waste management to make the supply chain management more cost efficient and environmental friendly [18]. LCA (Life Cycle Assessment) is a commonly used method for defining and evaluating the total environmental load associated with a product, process or activity (or in this case packaging system), by identifying and quantifying energy and materials consumed and waste released into the environment. [19] [20]

Twede and Clark examined, which types of supply chains facilitate reusable packaging systems. They state that returnable packaging is not appropriate for every product or logistical system. The supply chains with short lead times and shipping distances, efficient sorting, cleaning and tracking systems and industry consortia for standardization are mainly suitable [21].

2.2. Returnable packaging system costs

Compared to disposable packages, returnable packages, because of their longer lifetime, can reduce the needed amount of packages, which can be more environmentally friendly. However, a returnable packaging system may have higher costs of procurement, transportation, and other costs caused by cleaning, repairing, storage and management, etc. [6] Therefore, introduction of a returnable packaging system does not always mean cost reduction. Although it may lead to significant savings in logistics cost in some cases, not all companies find this solution cost effective [22].

These decisions in the practice generally include only the purchase cost savings amortized over the reusable containers' lifetime and often do not include logistics

costs. Nevertheless, the overall cost impact of such decisions remains unclear. A number of other factors affect the system cost, including the type of packaging used, transportation characteristics, handling, labour, and disposal costs. That is why the use of reusable containers does not always result in lower overall costs relative to disposable packaging because firm sourcing strategies and operations differ. [16]

3. System elements of packaging cost

In this section the total cost components and other factors of disposable and returnable packaging systems will be presented.

Table 1. Markings

Markings
CT = total cost
(CT_D = total cost of disposable packaging, CT_R = total cost of returnable packaging)
P = price of the packaging device
(P_D = price of disposable packaging, P_R = price of returnable packaging)
Q = quantity
T_R = cost of return transport
U = number of uses
S = cost of storage
R = repair cost
C = cleaning cost
A = administration cost
W = cost of disposal/waste

3.1. Packaging material cost

Regarding returnable packaging systems, several studies suggest that they improve the impact of packaging on the environment, because they can reduce the amount of packaging material [12]. Within the environmental criterion of sustainable packaging design, six indicators are defined, two of them is related to packaging material:

- Gross material intensity, which covers the total amount of packaging materials used in the packaging system including reuse of five main types (fibre, plastic, glass, metal and wood). This parameter takes reuse into account.

- Net material intensity describes the mass of packaging materials that are not being recycled, so it can be also seen as an estimate for the total amount of packaging waste generated from a logistics system. [5]

This cost can directly affect a company's returnable packaging system investment. Because of the improved product protection reusable containers can be even five to ten times more expensive in some cases than the disposable version that they replace [22].

In our case packaging material refers to purchasing price of the packaging devices, (all factors related to the production of the packaging are included in this component).

3.2. Transport cost

The transport costs can be divided into two parts: transport to the place of use and return transport. The cost for the return transport of empty containers can be mainly affected by the following factors:

- transport distance from the producer to the packaging collector where all the empty packages are gathered.
- possibility to take advantage of unbalance in transport flow, because of the use of joint loading when returning the packages.
- if the packaging is collapsible, it helps to maximize the amount of empty packaging for return transport. [16]

Transport is used in most parts of the ASC, but we examine the route between the engine producer and the buyer (the OEM). Since in our case study the cost of the first transport is the same for disposable or returnable packaging solutions (the destination is the same), only reverse transportation of the empty returnable packaging will be considered.

It was mentioned earlier that transport distance influences transport cost. Here we also have to mention the cycle time, which refers to the time a closed-loop is completed by the returnable packaging. In general, shorter cycle time may lead to lower initial investment cost, because less devices have to be purchased. Also shorter transportation distances usually result shorter cycle time. Nevertheless, to consider transport distance alone is not enough, complexity of the route and the supply chain has to be taken into consideration. Moreover, return transport opportunities and costs are even more important.

3.3. Administration, cleaning, repair and disposal (waste) cost

Administration cost influences economics for both disposable and returnable packaging, but for returnable packaging also includes the management of the system, like related labour cost and cost of the information system. Cleaning and repair costs only appear in the case of reusable packaging in order to maintain the right condition of the packaging device. The use of returnable containers minimizes the disposal cost.

3.4. Number of uses

One-way packaging is only used once, but in the case of returnable ones the number of uses is a major issue. If we plan certain number of uses and the packaging devices can be only used less, because it is damaged due to for example inappropriate use, it will increase the total cost. It can also happen that the number of uses is beyond the originally planned number, in this case the total system cost can be decreased.

If disposable and returnable packaging should be compared in terms of usage ratio, it is logical to choose cheaper disposable packaging devices or more expensive returnable ones, which can be used as many times as possible. In this case a solution has to be found, where the goal of development is to maximize the number of uses. Nevertheless, the unexpected loss of returnable packaging may lead to great economic consequences.

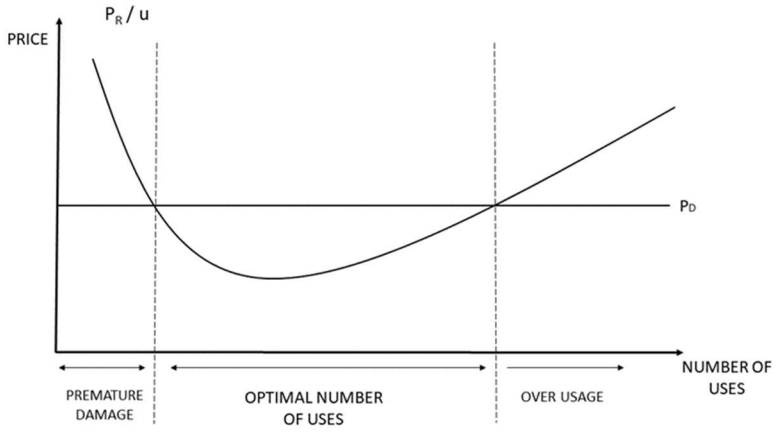


Figure 2. Usage ratio

Fig 2 shows us how the number of uses influences the price of the packaging, compared to the price of the disposable solution. Three cases can occur:

- Premature damage: it can happen because the technical and mechanical properties of the chosen packaging are not appropriate or in case of an unexpected event (like for example damage because of inappropriate material handling or loss of the packaging device as described earlier).
- Optimal number of uses: the cost is optimal, when the planned number of uses is reached, even if the initial price of the returnable packaging is higher than the disposable version.
- Over usage: after the packaging exceeds the optimal number of uses, it can happen that the maintenance of the appropriate technical conditions of the packaging (repair cost) becomes more expensive and because of that the price is higher. Risk of damage can be also significantly higher in this phase.

3.5. Packaging quantity

It means the total number of packaging devices in use in a returnable packaging system or a number of disposable containers need to be purchased. It is strongly affected by the number of uses of a particular packaging device and the return ratio (described later) as well.

Mollenkopf et al conclude that packaging quantity increases should favour the use of reusable containers. In their cost model for reusable packaging they consider average daily volume and packaging quantity two of the most important factors [22]. Daily volume can also significantly influence the initial investment cost, as well as cycle time.

4. An evaluation model for packaging cost structure

4.1. Cost of disposable packaging

The determination of the final expenses (total cost) is effected by the following components in case of disposable (one-way) packaging:

- purchase price of the packaging device (packaging material cost)
- cost of storage

- administration cost
- cost of disposal (waste).

$$\sum CT_D = P_D\, Q + A_D Q + S_D Q + W_D Q \quad (1)$$

where the total cost of disposable packaging (CT_D) is calculated from the price of disposable packaging (P_D), unit cost of administration (A), storage (S) and waste (W) multiplied by the quantity (Q).

4.2. Analysis of returnable packaging cost

The cost structure of a returnable packaging system is more complex compared to the disposable one.

In case of industrial, returnable packaging (operating in a closed-loop) the companies make decisions mainly based on the cost.

The calculation of the final expenses (total cost) is effected by the following components/elements:

- purchase price of the returnable packaging device (packaging material cost)
- cost of return transport
- cost of storage, repair and cleaning
- administration cost
- cost of disposal (waste)
- quantity of the returnable packaging devices needed to maintain the operation of the system
- number of uses

  return ratio

$$\sum CT_R = P_R\, Q + P_R Q\,(1 - RR) + T_R(U-1)Q\,RR + AQ + (S + R + C)\,Q\,RR + WQ, \quad (1)$$

where the total cost of a returnable packaging system (CT_R) is calculated from the price of the returnable packaging material (P_R) squared by the quantity (Q) initially needed for the system the operate smoothly, the price of not returning packaging devices which have to be purchased again, the cost of return transport (except for the last route at the end of the life cycle of the packaging device), furthermore the unit costs of the administration (A), storage (S), repair (R), cleaning (C) and waste (W). Return ratio (RR) is explained the next section.

4.3. Model criteria

1) Relation between packaging (purchase) price and transport cost

There are two possible cases:

- the cost of a new packaging device is higher than the cost of transporting back the same packaging: $P_R \geq T_R$
- the cost of a new packaging device is lower than the cost of the reverse transportation: $P_R \leq T_R$

In our model we assume that in case of a closed loop returnable packaging system the cost of purchasing a new packaging device should be higher than the cost of back transportation, otherwise it is not economical to transport it back, because we rather buy a new one.

2) Introduction of return ratio

The return ratio (RR) represents the percentage of the packaging devices returning back from the whole pool, it has a value between 0% and 100%. We assume that RR can never reach 100%, because it would mean that all packaging devices come back and it is not possible in the practice.

The return ratio (Fig 3) is expressed as the ratio of the transport cost (T_R) and the price of a new packaging device (P_R):

$$RR = 1 - \frac{{T_R}^{\epsilon}}{{P_R}^{\epsilon}} = 1 - \left(\frac{T_R}{P_R}\right)^{\varepsilon}, \qquad (2)$$

We apply an elasticity modulus (Ɛ) in order to model real life circumstances better, because the relationship between RR and T/P is non-linear. These conditions can be for example the lack of operation conditions, i.e. bad road or IT infrastructure or the low motivation level of the stakeholders of the SC, for example willingness to send back the empty packaging devices. Ɛ should be between 0 and 1: $0 < \varepsilon < 1$.

Because $P_R \geq T_R$ and RR should never exceed 100%, the ratio of T/P should be between 0 and 1.

$$0 \leq \frac{T_R}{P_R} \leq 1 \text{ and also } \left(\frac{T_R}{P_R}\right)^{\varepsilon} \leq 1 \qquad (4)$$

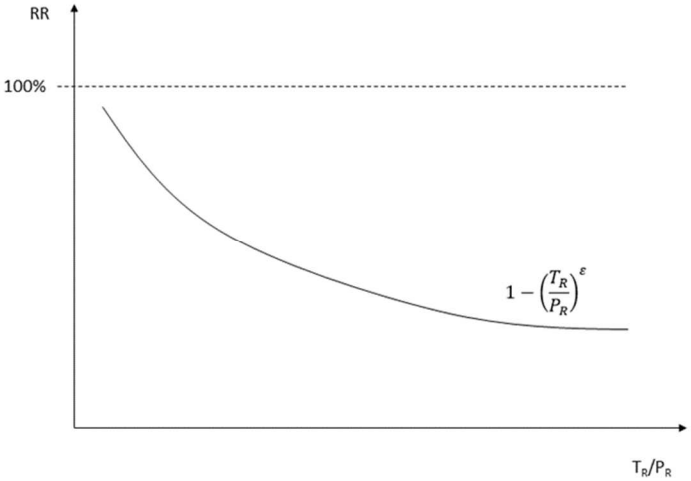


Figure 3. The effect of T_R/P_R on the return ratio

If we substitute RR with the above formula, we get the following equation:

$$\sum C_T = P_R Q + P_R Q\left(1-\frac{{T_R}^{\varepsilon}}{{P_R}^{\varepsilon}}\right) + T_R(U-1)Q\frac{{T_R}^{\varepsilon}}{{P_R}^{\varepsilon}} + AQ + (S+R+C)\,Q\,\frac{{T_R}^{\varepsilon}}{{P_R}^{\varepsilon}} + WQ \tag{5}$$

We are looking for the minimum value of the equation (3).

$$\frac{d\,C_T}{d\,T_R} = -\varepsilon\,\frac{P_R}{P_R^{\varepsilon}}\,Q\,T_R^{\varepsilon-1} + \frac{\varepsilon\,T_R\,(U-1)Q\,T_R^{\varepsilon-1}}{P_R^{\varepsilon}} + \varepsilon\,\frac{(S+R+C)Q\,T_R^{\varepsilon-1}}{P_R^{\varepsilon}} = 0 \tag{6}$$

Using the presented analytical way, we could find an optimum point.

5. Case study

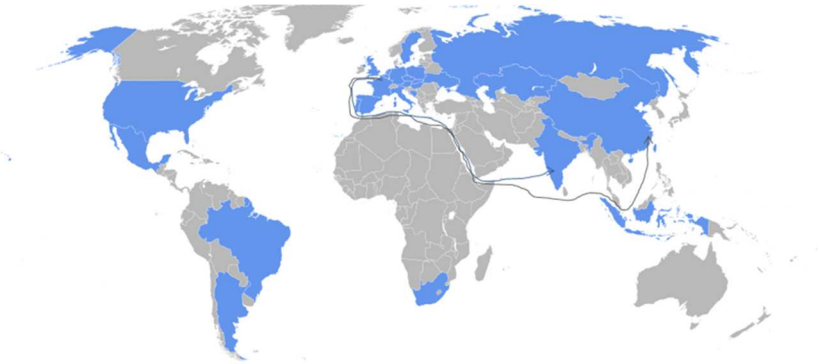


Figure 4. Transport routes to India and China

In the case overseas CKD (completely knocked down) transport will be examined. The same returnable packaging is used for automotive engine (CKD) transport from Europe to two different destinations in India and China (Fig 4). The packaging comes back from China, but it will be discarded in India. In both cases it is transported by multimodal transport, namely road, rail and sea. More than 85% of the transport distance is carried out by maritime transport and merely around 3% (India) and 1% (China) is by road.

The finished CKD engines are sensitive products therefore special racks are mainly used to store and transport them (Fig 5). These ensure safe and reliable transport and storage. The column is usually collapsible in order to save place while returning back as empty transportation. The posts are supposed to keep the engine in place, but these can be also collapsed. The returnable packaging system also contains disposable components, namely VCI (anti corrosion) foil.

Figure 5. Examples of industrial CKD packaging [24] [25] [26]

In this case, the engine company plays the role of system integrator, then the OEM (Original Equipment Manufacturer) complies with the vehicle assembling factory. As it is shown in the case study example, returnable packaging is more common on the outbound flow of the engine companies.

5.1. Cost calculations

In the following section we present examples for how certain cost components included in the returnable packaging system influence the returnable packaging decision. Calculations are based on empirical data and the above described theoretical framework. The effect of variable transport cost and packaging purchase price will be highlighted in the results besides different planned number of uses.

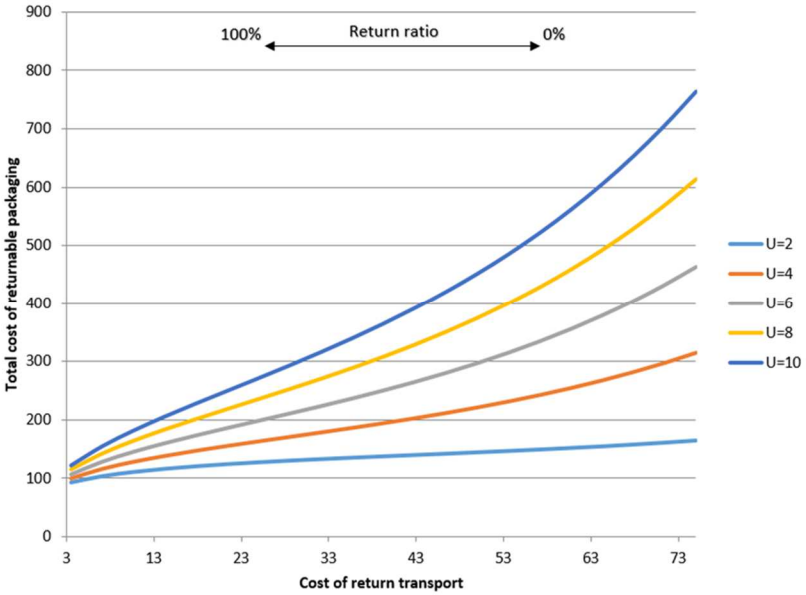


Figure 7. Total cost in case of variable transport cost and different planned usage number

In the first case (Fig 7) the price of the packaging device is fixed, €75. As it is one of the model criteria, we suppose that the return transport cost is less than the purchase price of the packaging. This is why the transport cost is only examined between €1 and €75.

The unit costs of storage, repair, cleaning administration and disposal (waste) are also fixed in the model according to the followings:

- Storage €1,50
- Repair (average) €2,50
- Cleaning €0,20
- Admin €0,01
- Waste €10,00

The graph shows the optimal cost besides different planned number of uses (U=2; 4; 6; 8; 10) taking the changes of the return ratio into consideration at the same time.

The curves present the line until it is worth it to choose returnable packaging (above the line disposable packaging should be preferred).

The model gives us a decision map in terms of how to find the appropriate technical solution besides variable transport cost, if all other data is known.

The following figure (Fig 8) illustrates an example where two solutions with the same logistics performance, but different planned usage (U=2 and U=10) are compared. This means that the packaging device planned for two uses has to be bought 5 times. It also has to be highlighted that because of the formula the return transport cost is only included 5 times in the total cost of this case.

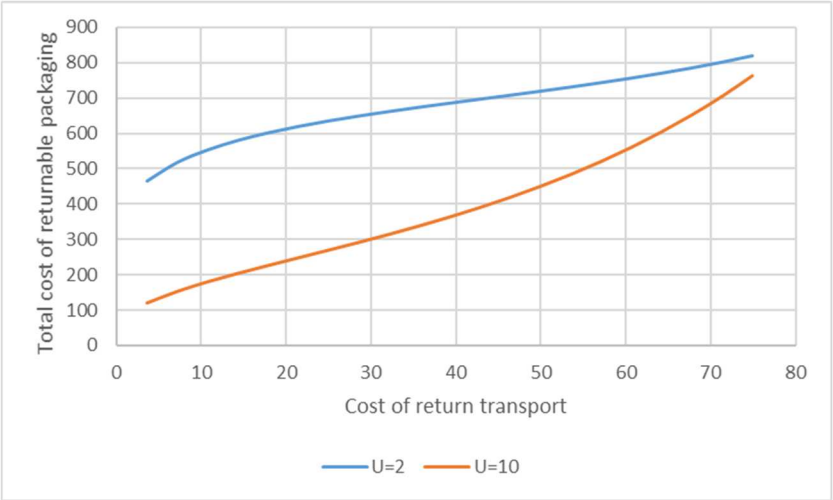


Figure 8. Comparison of total cost besides the same logistical performance

Nevertheless, if the same price is assumed for returnable packaging devices, value of the tied-up capital will be less for the solutions which can be used more times. It leads us to the conclusion that a packaging system needs to be developed which allows as many uses as possible besides acceptable price and also the stakeholders of a closed loop system need to be encouraged to return the packaging. Environmental impact is not considered in our model.

6. Conclusion

Automotive industry is one of the world's most significant economic sectors by its revenue. Engine transportation is carried out between different facilities all over the world. Besides waste reduction and optimal cost level, the best packaging solutions

for automotive packaging are those that can maximize the use of packaging space so that all the products can easily be packed and stacked. Appropriate protection of the sensitive and relatively expensive products (engines) is also crucial, especially during overseas transport where the external forces are multiplied. However, managing returnable packaging systems requires more than just inverse transportation. Many other cost factors should be considered in the process.

In the case study, we assume that the price of a technically suitable packaging device is constant, that is why we optimize the model for the ways and costs of return (back) transport, in order to define the appropriate alternative so that it is economically worth it to introduce a returnable packaging system. We optimize according to this criterion in our model. Although the result is situation specific, the formula could be generalized across other industrial settings for closed-loop returnable packaging systems.

Acknowledgement

The authors would like to thank to EFOP-3.6.1-16-2016-00017 1 'Internationalisation, initiatives to establish a new source of researchers and graduates, and development of knowledge and technological transfer as instruments of intelligent specialisations at Széchenyi István University' for the support of the research.

References

[1] H. M. Gamez-Alban, , O.C. Soto-Cardona, , C. Mejia-Argueta, A. T. Sarmiento, A cost-efficient method to optimize package size in emerging markets, European Journal of Operational Research, 241 (3) (2015) pp. 917–926.
doi: 10.1016/j.ejor.2014.09.020

[2] F.T.S. Chan, H.K. Chan, K.L. Choy, A systematic approach to manufacturing packaging logistics, International Journal of Advanced Manufacturing Technology 29 (9-10) (2005) pp. 1088-1101.
doi: 10.1007/s00170-005-2609-x

[3] C.A.S. Dominic, Towards a conceptual sustainable packaging development model: A corrugated box case study, Packaging Technology and Science, 28 (2015) pp. 397-413.
doi: 10.1002/pts.2113

[4] R. Accorsi, A. Cascini, S. Cholette, R, Manzini, C. Mora, Economic and environmental assessment of reusable plastic containers: A food catering supply chain case study, International Journal of Production Economics, 152 (2014) pp. 88–101.
doi: 10.1016/j.ijpe.2013.12.014

[5] E. Svanes, M. Vold, H. Møller et al., Sustainable packaging design: a holistic methodology for packaging design. Packaging Technology and Science, 23 (3) (2010) pp. 161–175.
doi: 10.1002/pts.887

[6] Q. Zhang, A. Segerstedt, Y. C. Tsao, B. Liu, Returnable packaging management in automotive parts logistics: Dedicated mode and shared mode, International Journal of Production Economics, 168 (2015) pp. 234–244.
doi: 10.1016/j.ijpe.2015.07.002

[7] K. Vöröskői, P. Böröcz, Framework for the Packaging Supply Chain of an Automotive Engine Company, Acta Technica Jaurinensis, 9 (3) (2016) pp. 191-203.
doi: 10.14513/actatechjaur.v9.n3.409

[8] What Are The Best Packaging Solutions For Automotive Packaging (2013) [cited 2016-10-05]
URL http://www.mjspackaging.com/blog/what-are-the-best-packaging-solutions-for-automotive-packaging/

[9] P. Böröcz, Á. Mojzes: The importance of packaging in logistics. Transpack, 8 (2) (2008) pp. 28-32. in Hungarian

[10] K. L. Yam: The Wiley encyclopedia of packaging technology. Wiley, USA, 2009

[11] P. Böröcz, P. Földesi: The application of the game theory onto the analysis of the decision theory of logistic packagings, Acta Technica Jaurinensis, 1 (2) (2008) pp. 259-268.

[12] H. Palsson, et al., Selection of Packaging Systems in Supply Chains from a Sustainability Perspective: The Case of Volvo, Packaging Technology and Science, 26 (2013) pp. 289–310.
doi: 10.1002/pts.1979

[13] P. Böröcz, S.P. Singh.: Measurement and Analysis of Vibration Levels in Rail Transport in Central Europe. Packaging Technology and Science 30 (2016) pp. 361-371.
doi: 10.1002/pts.2225

[14] C. E. Witt: Are reusable containers worth the cost? Material Handling Management, 55 (7) (2000) p.75.

[15] N. Boysen, S. Emde, M. Hoeck, M. Kauderer: Part logistics in the automotive industry: Decision problems, literature review and research agenda. European Journal of Operational Research, 242 (1) (2015) pp. 107-120.
doi: 10.1016/j.ejor.2014.09.065

[16] C. Dominic, et al.: Förpackningslogistik. 2000: Packforsk, Kista

[17] European Commission: Packaging and Packaging Waste [cited 2017-12-09] URL http://ec.europa.eu/environment/waste/packaging/index_en.htm

[18] S. G. Azevedo, H. Carvalho, V. C. Machado, The influence of green practices on supply chain performance: A case study approach, Transportation Research Part E: Logistics and Transportation Review, 47 (6) (2011) pp. 850–871.
doi: 10.1016/j.tre.2011.05.017

[19] A. Conte, G.M. Cappelletti, G.M. Nicoletti, C. Russo, M.A. Del Nobile: Environmental implications of food loss probability in packaging design, Food Research International, 78 (2015) pp. 11–17.
doi: 10.1016/j.foodres.2015.11.015

[20] S. Ross, D. Evans, The environmental effect of reusing and recycling a plastic-based packaging system, Journal of Cleaner Production, 11 (5) (2003) pp. 561–571.
doi: 10.1016/j.foodres.2015.11.015

[21] D. Twede, R. Clarke, Supply chain issues in reusable packaging. Journal of Marketing Channels, 12 (1) (2004) pp. 7–26.
doi: 10.1300/J049v12n01_02

[22] D. Mollenkopf, D. Closs, D. Twede, S. Lee, G. Burgess, Assessing the viability of reusable packaging: A relative cost approach. Journal of Business

Logistics, 26 (1) (2005) pp. 169–197.
doi: 10.1002/j.2158-1592.2005.tb00198.x

[23] Industrial packaging [cited 2017-12-09]
URLhttp://www.dssmith.com/contentassets/448ba5a06897471d921313fcc4377680/industrial_packaging.jpg?w=650&h=340&mode=max&scale=both&format=jpg&quality=75

[24] CKD/SKD packaging [cited 2017-12-09]
URL http://flexol.in/site/wp-content/uploads/2015/09/ckdskd-4.jpg

[25] Stillage for transport of engine [cited 2016-10-03]
URL http://china-rollcontainer.com/product-6-1-engine-transport-stillage.html/138065

Acta
Technica
Jaurinensis

Vol. 12, No. 2, pp. 101-116, 2019
DOI: 10.14513/actatechjaur.v12.n2.496
Available online at acta.sze.hu

Direct Instantaneous Torque Control of the Switched Reluctance Motor for Electric Vehicles Applications Using Fuzzy Logic Control

R. Abdel-Fadil[1,2], L. Számel[2]

[1]Budapest University of Technology and Economics, Electric Power Engineering
18 Egry J. str. Budapest, 1111, Hungary
e-mail: Reyad.abdelfadil@aswu.edu.eg

[2]Aswan University, Electrical Engineering Department
Aswan, 81542, Egypt

Abstract: This paper presents direct torque control of Switched Reluctance Motor (SRM) using Fuzzy Logic Control (FLC) for electric vehicles applications. The PD-FLC is proposed for SRM torque control, to keep the torque of the motor shaft in tracking the reference torque with high accuracy. With the help of FLC techniques, the SRM torque ripples can be reduced compared to traditional control techniques. In this study, the nonlinear 6/4 SRM model is simulated with the symmetrical converter, and the converter controller is programmed using C-language. The proposed method is tested at different load and variable speed conditions, and the obtained results confirm that the FLC direct torque control can be used for torque control to improve the motor performance and reduce the torque ripples compared to other techniques such as direct instantaneous torque control.

Keywords: switched reluctance motor, electric vehicles, fuzzy logic control, control techniques, torque control, and torque ripple

1. Introduction

In the present, the energy sources, environmental pollution, and noise are considered the main problems facing many countries due to using conventional vehicles [1]. In contrast, the Electric Vehicles (EVs) provide a good solution for

pollution and noise problems and reduce petroleum usage as well. Therefore, the using of the EVs in transportation became essential consequently, the EVs proportion in the commercial market gradually increasing [2][3]. Recently, researchers and industrial companies are working to solve the operational problems of EVs, save energy, and achieve the best performance at an appropriate cost. To improve the overall performance of the EVs many different components must be optimized, the electric motors and it's driving circuits are considered the main components and the most important parts in the EVs therefore, the performance of the vehicles depends mainly on electric motors and driving circuits [4][5].

There are many types of electric motor that can be used for EVs applications each type has advantages and drawbacks. The Switched Reluctance Motors (SRMs) are considered one of these types that can be applied to EVs, due to their several advantages such as low-cost manufacturing, simple construction and material composition, high starting torque, rugged construction, high speed ranges, ability to operate at high temperature, higher reliability, and low inertia. Although the SRM has many benefits, still face some drawbacks and challenges which have to overcome such as torque ripples due to doubly salient structure and pulse excitation, and complex control [6][7]. To overcome the problems of the SRM and improve the EVs overall performance there are two main ways, the first one by improving the mechanical design of the electric machine and the second way by selecting suitable control strategies and optimal control techniques. Three main strategies can be used for SRM control these strategies are speed control, current control, and torque control [8]. The main challenge which has to be solved by control techniques is the torque ripples problem, this problem is considered very complicated, and it is not easy to solve because it is affected by many factors [9].

Two types of torque control strategies can be used for SRM: the first strategy is indirect torque control, which uses the complex algorithms or distribution function to obtain the reference current. After that, the current controller is used to control phase torque. The second strategy is the Direct Torque Control (DTC) which uses the torque controller and simple control scheme to reduce the torque ripple [10]. In recent years, the Direct Instantaneous Torque Control (DITC) algorithm has much progress and development to overcome the problems of the indirect torque control methods [11]. The main feature of this method is the instantaneous torque is considered a control variable directly, and the conversion from torque to current and closed-loop currents control became not essential. Also, DITC able to avoid the torque error immediately with a good dynamic response and reducing the torque ripple [12]. Because of the simple hysteresis switch rule is used in the traditional DITC method, just one phase state can be used depending on the error in the torque at each sampling cycle. Two different methods can be used to make sure that the torque ripple within an acceptable range: the first method is the sampling time

reduction but this solution maybe increases the hardware cost, and the second method is the PWM method which can be used to control of the average voltage of phase winding in sampling time [10].

Recently, the applications of artificial intelligence methods have been used to translate human knowledge into a form comprehensible by computers. Advanced control based on artificial intelligence techniques is commonly defined as intelligent control. The intelligent control methods considered one of the applicable technique that can solve the torque ripple issue in the SRM, because of the advantages of intelligent control such as nonlinear control, self-learning, and it is adaptive capacity. The idea of intelligent control is to use off-line or online learning and optimization. Fuzzy Logic Control (FLC) is a technique to make machines more intelligent, it is defined as a mathematical tool to deal with uncertainty and imprecision [13], it was introduced by Lotfi Zadeh in 1965 [14], and it is one of the artificial intelligence methods which suitable for torque control of SRM. FLC system consisted of three main blocks as shown in Fig. 1 [15]. The first one is fuzzification its main function converts the input data to fuzzy sets values. The second block is decision making logic depend on the knowledge base, this part determines how the logic operations are achieved, and together with the knowledge base can regulate the outputs of each fuzzy set. The last one is the defuzzification block, which converted fuzzy values to output data [16], [17].

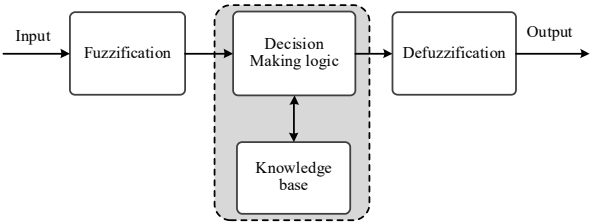


Figure 1. The fuzzy logic system block diagram

2. Direct instantaneous torque control

The online availability of the total instantaneous torque is considered fundamental requirement for the DITC strategy, and due to the nonlinear characteristics of the SRM it is not easy to obtain the instantaneous torque by straightforward analytical equations as in other motors (DC or rotating field) but the instantaneous torque can be estimated only by stored characteristics of the motor [18]. Two methods can be used to estimate the instantaneous torque, torque estimation as a function of phase current and rotor position, or estimate the torque as a function of phase current and

phase flux linkage [10]. Fig. 2 show the overall blocks diagram of the current-position based DITC.

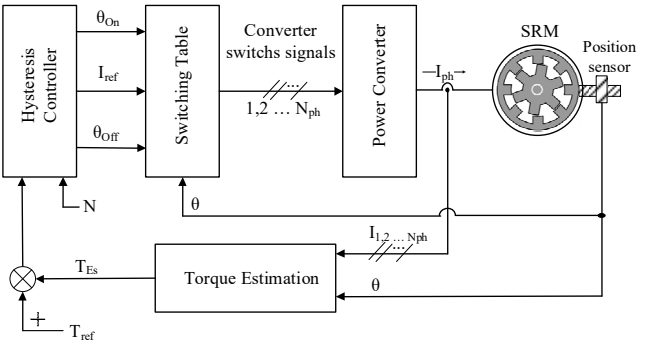


Figure 2. Direct instantaneous torque control

The PWM-DITC consisted of a combination of the traditional DITC and PWM as shown in Fig. 3 [19]. Depending on the phase currents (Iph) and rotor position (θ) the lookup table is used to implement the torque estimation block. The torque control block generating the duty cycle for all activated motor phases depending on the error between the estimated torque and command torque, then the PWM block generate the switching signals for the power converter. By using the PWM method the average phase voltage can be regulated to control the currents variety in single sampling time, the sampling time can be extended and reduce the torque ripple comparing to the traditional DITC. The significant drawback of this method the switching frequency is increased comparing to DITC therefore, the losses due to switching frequency and EMC noise are increased.

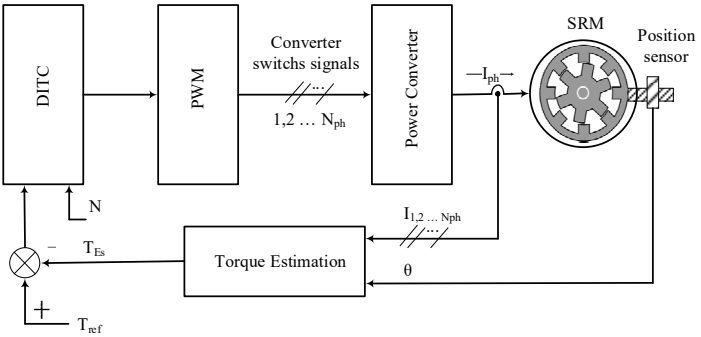


Figure 3. Direct instantaneous torque control with PWM

3. Controller design

To programming the FLC, there are main steps have to be performed: Firstly, classify the inputs, their ranges/limits, and label them. The second step is to classify the outputs. Thirdly, make the degree of the membership function for every input and output. After that, structure the system rule-based and determine how the action will be performed by select optimum rule-based. Finally, combined the rules and defuzzify the output. In this study, PD-FLC is proposed for DITC to enhance motor performance and reducing the torque ripples compare with the DITC method. The inputs of the FLC are the motor shift torque error (e) and the change in this error (Δe). The output of the FLC is a converter modulation index, which used to generate the optimal gating signals after comparing it with a carrier wave in the PWM block as described in Fig. 4 the FLC design steps will be discussed in the following sections.

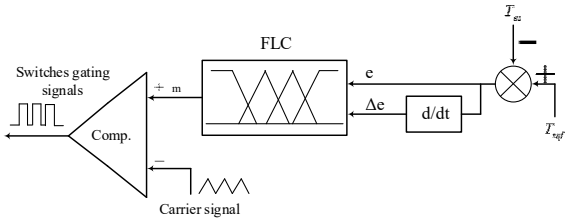


Figure 4. PD-Fuzzy logic direct torque control

3.1. Fuzzification

The FLC uses linguistic variables instead of numerical variables, and the fuzzification process converts these numerical variables (real numbers) to linguistic variables (fuzzy sets). Consequently, the motor torque error signal values can be assigned as follow: Negative Very Big (NVB), Negative Big (NB), Negative Medium (NM), Negative Small (NS), Zero (ZE), Positive Small (PS), Positive Medium (PM), Positive Big (PB), Positive Very Big (PVB). The shape of the membership function for the first input is shown in Fig. 5. Similarly, the second input for the fuzzy system (change of error) converted from numerical value to a linguistic variable according to the membership function shown in Fig. 6.

3.2. Rule Evaluator

The membership functions variables were defined off-line, and its values are selected according to the behavior of the system observed during simulations. The decision-making logic (rules base table) which used in this work are listed in Table

1, and it can be built using the experiences and by observing the performance of the controller. Generally, to evaluate between two inputs (A and B) the basic fuzzy set operations can use one of three rules AND (∩), OR (∪) or NOT (~). In this work, AND (∩) intersection is used which is presented in equation (1) [20].

$$\mu A \cap B = min\,[\mu A(X), \mu B(X)] \quad (1)$$

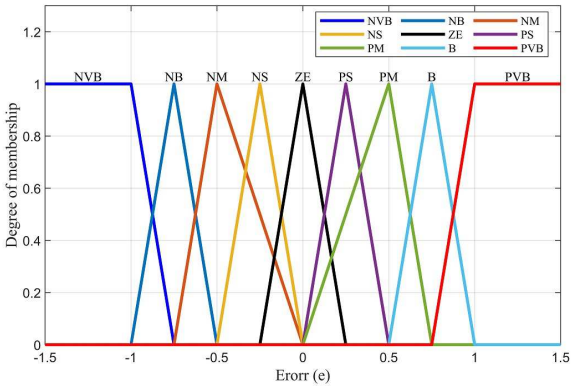


Figure 5. Membership functions representing the error signal

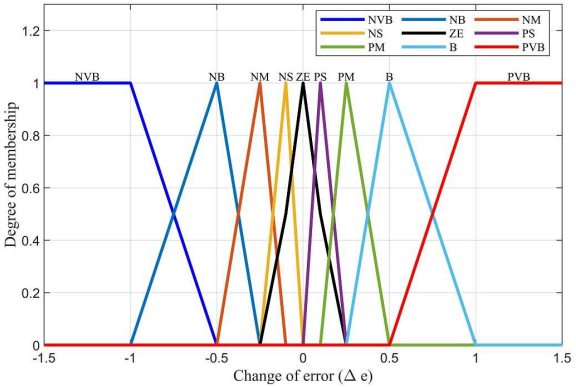


Figure 6. Membership functions representing the change of error

Table 1. If-Then rule base for fuzzy logic control

e / *Δ e*	***NVB***	***NB***	***NM***	***NS***	***ZE***	***PS***	***PM***	***PB***	***PVB***
NVB	EL	EL	EL	EL	EL	VL	L	UM	M
NB	EL	EL	EL	EL	VL	L	UM	M	AM
NM	EL	EL	EL	VL	L	UM	M	AM	H
NS	EL	EL	VL	L	UM	M	AM	H	VH
ZE	EL	VL	L	UM	M	AM	H	VH	EH
PS	VL	L	UM	M	AM	H	VH	EH	EH
PM	L	UM	M	AM	H	VH	EH	EH	EH
PB	UM	M	AM	H	VH	EH	EH	EH	EH
PVB	M	AM	H	VH	EH	EH	EH	EH	EH

3.3. Defuzzification

The function of the defuzzification step is convert-back linguistic variables of the output, which generated from fuzzy logic rules to real numbers. The membership function shape for defuzzification process which used in this study is shown in Fig. 7. In this case, the output signal which represents the modulation index (m) can be assigned as follow: Extremely Low (EL), Very Low (VL), Low (L), Under Medium (UM), Medium (M), Above Medium (AM), High (H), Very High (VH), Extremely High (EH).

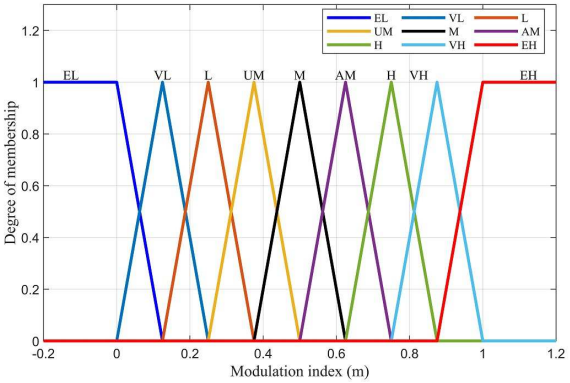


Figure 7. Output Membership functions

Three different methods can be used for membership defuzzification, Center Of Area (COA), Bisector, or Middle Of Maximum (MOM). In this study, the COA method is used, which is presented in equation (2), because it considered the most popular method [20] [21].

$$U(n) = \frac{\sum_{j=1}^{n} \mu(uj)\omega j}{\sum_{j=1}^{n} \mu(uj)}, \quad (2)$$

where μ(uj) the membership function of the jth fuzzy set of input variable uj, and ωj the jth output fuzzy, and n is the number of fuzzy membership functions.

4. Simulation results and discussions

The SRM torque control with the PD-FLC control technique is shown in Fig. 8. The controller needs a reference torque signal to be followed by the motor torque, for this reason, the motor speed is taken as the feedback signal and compared to the desired speed and determine the error values of the motor speed, this error is applied to the PI controller to generate the demanded reference torque. With the help of the phase current and rotor position, the motor shaft torque can be estimated, the most general expression for the instantaneous torque equation of the of SRM can be expressed in equation (3), the estimated torque compared to reference torque, the torque error and the change in this error are used as input variables for FLC. According to the controller design, the controller determines the degree of the modulation index of the power converter then the PWM block generates the switch-state for all active switches of the power converter.

$$T_j = \frac{dW_C(\theta,i)}{d\theta}\Big|_{i_j=const} = \frac{dW_s(\theta,i)}{d\theta}\Big|_{\psi_j=const} \quad (3)$$

where T_j is the phase torque, W_C is the co-energy, and W_S is the stored field energy.

The control technique proposed in this work is conducted on a 60 kW SRM using PSIM software [22], and the controller is programmed using a C-code capability in this software. The simulation parameters of the SRM are given in Table 2. [8] [23]. The first step of the simulation work is applying the DITC using Hysteresis Torque Control (HTC) for SRM and obtain the machine performance (motor speed profile, motor torque, and torque ripples). The simulation carried out through three steps: the first one is applying the conventional DITC using hysteresis control and observe the machine performance (motor speed profile, phase current, motor torque, and torque ripples). After that, in the second step, the DITC is applying but using PD-FLC instead of hysteresis control and observe the performance of the SRM. Finally, comparing the obtained results for the motor performance in step one and two are compared.

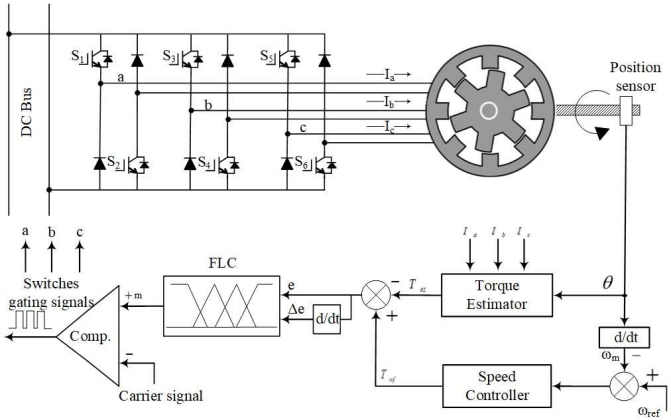


Figure 8. SRM torque control block diagram using PD-FLC

Table 2. SRM simulation parameters

Parameter	***Value***	***Parameter***	***Value***
Power	60 kW	Speed	1000 rpm
DC link voltage	220 V	Load torque	10 - 20 Nm
Maximum current	450 A	Inertia	0.05 kg m^2
Stator resistance	0.05 ohm	No. of rotor pole	4
Unaligned inductance	0.67 mH	No. of stator pole	6
Aligned inductance	23.62 mH		

To verify the effectiveness of the performance for the proposed controller, it is tested at two different study cases: the first one is investigating the controller robustness in following the reference torque, in this case, the reference speed is *1000 rpm*, and the load torque is changed from *10 Nm* to *20 Nm* at *0.3s*. The motor speed profile in case of load changed is shown in Fig. 9 the motor speed reaches to steady state after *0.05s* and tracks the reference speed until the load torque changed at *0.3s*, then a small drop in the motor speed occurs, this drop is not static and decreasing, but it needs more than 0.5s to reach the desired speed as shown in the zoom of the speed profile.

The SRM torque performance is studied using hysteresis control with sampling time $Ts=1\mu s$, and the hysteresis band was the smallest possible value with this sampling time ($\Delta T=\pm 0.5$). PD-FLC also is applied to the motor torque controller and the sampling time is constant in two control methods. The torque of the motor is

shown in Fig. 10 and it's clear that the motor torque was doubled at *Ts=0.3s*, to investigate the controller robustness in track the reference torque. Also, this figure shows the comparison between the two used techniques at different load conditions, where the black color represents a motor torque using hysteresis control (HTC) and this torque is in red color in case of fuzzy control (FLC). This comparison shows that the torque ripples during the phase conduction period at two different load conditions with FLC are smaller than in case of hysteresis control. There is no doubt that the motor torque ripples still needed to reducing particularly in the phase commutation period, this period basically depends on the optimal turn on and turn off angle according to motor design parameters.

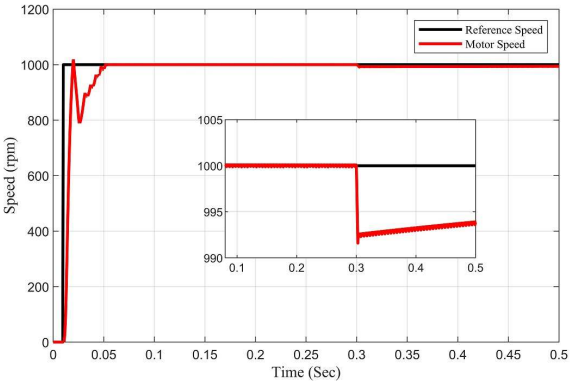


Figure 9. SRM speed profile in case of load torque changed condition

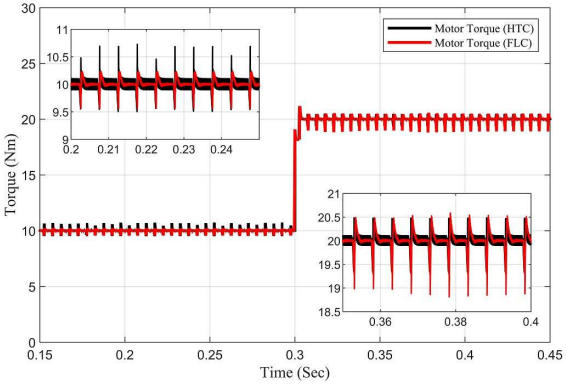


Figure 10. The motor torque with constant speed and load torque changed at 0.3s

The comparison between the motor phase current using FLC, and in the case of hysteresis control in case one is shown in Fig. 11. The obtained results show the difference between the two techniques, for phase (a) as an example the red color represents a phase current using hysteresis control, and this current is in blue color in case of FLC. It is noted that the current ripples during the phase conduction period using FLC are smaller than in the case of hysteresis control. This result shows the ability of the fuzzy logic technique to reduce the motor phase current ripples comparing to traditional methods.

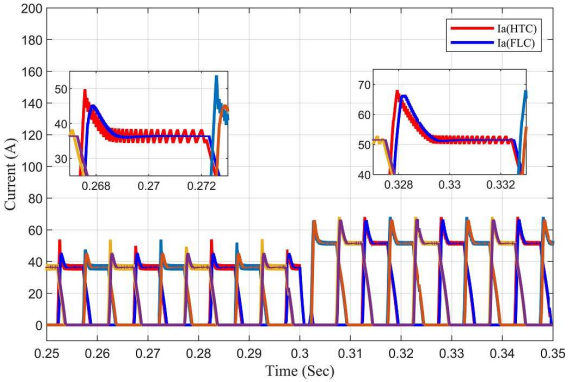


Figure 11. The phase current signals and zoom in case of load torque changed

The second study case tests the tracking performances of the controller, the motor speed profile, in this case, shows in Fig. 12 which load torque is *10 Nm* and the reference speed changed from *800 rpm* to *1200 rpm*, and the results show that the motor speed is tracking the reference speed satisfactorily. The motor torque in case of tracking performance is shown in Fig. 13 by zooming the motor torque performance curve it can be noted that the torque ripples in case of FLC are smaller than hysteresis control case at different speed values. When the motor torque increase to reach *20 Nm* with a variable reference speed, also the results demonstrate that the motor torque performance in case of FLC method for direct torque control is better than the conventional DITC methods. The obtained result at *20 Nm* load torque is shown in Fig. 14.

The previous results confirm that the proposed torque controller (FLC-DITC) has many advantages comparing with traditional DITC technique not only because of the general advantage of FLC as we mentioned in the introduction section but also this technique makes the motor torque tracked the reference signal with the smallest value of torque ripples at different loading conditions and variable speeds. Table 3.

reviews the torque ripples percentage by using FLC and hysteresis control, in case of reference speed *1000 rpm* and the load torque changed from *10 Nm* to *20 Nm*, and Table 4. reviews the torque ripples percentage, when, the reference speed changed from *800 rpm* to *1200 rpm* and the load torque *TL=10 Nm.* These tables summarize and demonstrate the effectiveness of the fuzzy logic DITC method compared to the traditional DITC method.

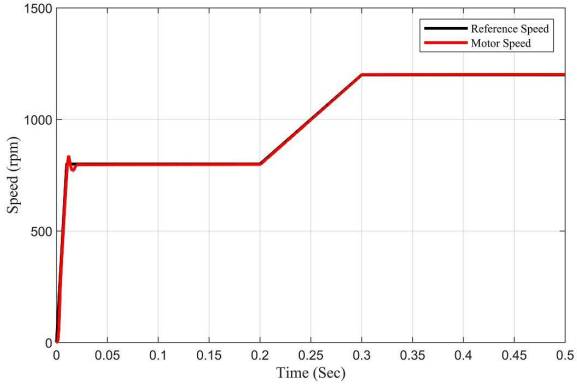


Figure 12. SRM speed profile with constant torque and reference speed changed

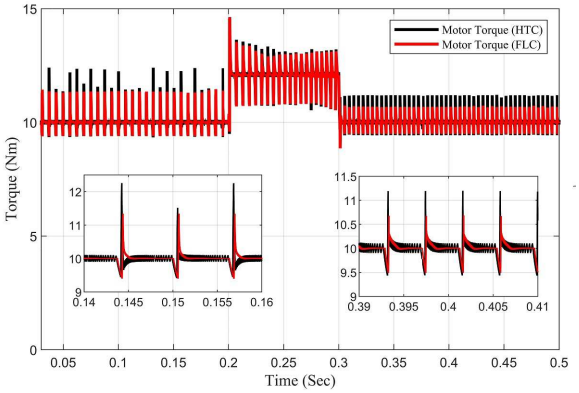


Figure 13. The motor torque in case of tracking performance (TL = 10 Nm)

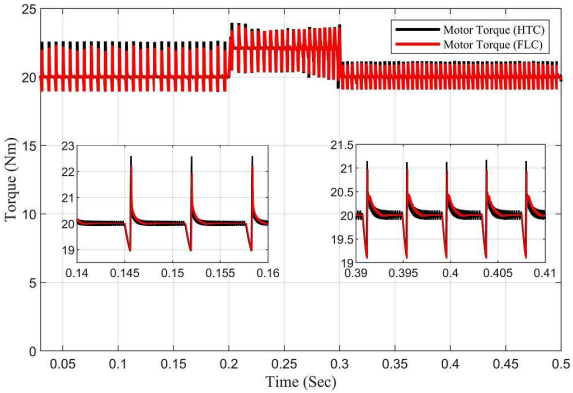


Figure 14. The motor torque in case of tracking performance (TL = 20 Nm)

Table 3. SRM torque ripples in case 1(Speed =1000 rpm)

Load Torque / Torque ripples	10 Nm	20 Nm
Fuzzy Logic DITC	6.9%	5.8%
Conventional DITC	12.3%	8.85%

Table 4. SRM torque ripples in case 2 (Torque = 10 Nm)

Motor Speed / Torque ripples	800 rpm	1200 rpm
Fuzzy Logic DITC	20.7%	12.5%
Conventional DITC	30.65%	17.8%

5. Conclusions

In this paper, fuzzy logic direct torque control of switched reluctance motor for electric vehicles applications was introduced. The FLC is one of artificial intelligence control techniques, and it is suitable for torque control of the SRM because it has many advantages such as nonlinear control, self-learning, and its adaptive capacity. In this work, the FLC programmed by C- code and the simulation test performed with 60 kW SRM. The controller was tested at different loading conditions to investigate the controller robustness, and with variable reference speeds to examine the tracking performance of the proposed controller. The obtained results show the effectiveness of the proposed technique (PD-FLC) to reduce the SRM torque ripples, whether in case of torque load changed or in case of motor

speed changed. With the help of fuzzy logic DITC, the SRM torque tracking the reference signal with smaller values of ripples compared to traditional DITC techniques.

References

[1] A. Y. Anekunu, S. P. Chowdhury, S. Chowdhury, A review of research and development on switched reluctance motor for electric vehicles, 2013 IEEE Power Energy Soc. Gen. Meet., Canada, 2013, pp. 1–5.
doi: 10.1109/PESMG.2013.6672494

[2] C. He, C. Hao, W. Qianlong, X. Shaohui, Y. Shunyao, Design and control of switched reluctance motor drive for electric vehicles, 14th Int. Conf. Control. Autom. Robot. Vision, Phuket, Thailand, 2016, pp. 1–6.
doi: 10.1109/ICARCV.2016.7838783

[3] M. Ruba, D. Fodorean, Motor-drive solution for light electric vehicles based on a switched reluctance machine", 20th IEEE Int. Conf. Autom. Qual. Testing, Cluj-Napoca, Romania, 2016, pp. 1–6.
doi: 10.1109/AQTR.2016.7501372

[4] A. C. Sijini, E. Fantin, L. P. Ranjit, Switched Reluctance Motor for Hybrid Electric Vehicle, Middle-East Journal of Scientific Research 24 (3) (2016) pp. 734–739.

[5] J. Zhu, K. Wai, E. Cheng, X. Xue, Y. Zou, Design of a New Enhanced Torque In-Wheel Switched Reluctance, IEEE Transactions on Magnetics 53 (11) (2017) pp. 1–4.
doi: 10.1109/TMAG.2017.2703849

[6] T. J. E. Miller, Electronic Control of Switched Reluctance Machines, 1st Edition, Newnes, Oxford, 2001.

[7] R. Abdel-Fadil, L. Számel, Enhancement of the Switched Reluctance Motor Performance for Electric Vehicles Applications Using Predictive Current Control, 2018 International IEEE Conference and Workshop in Óbuda on Electrical and Power Engineering, Budapest, Hungary, 2018, pp. 195-200.
doi: 10.1109/CANDO-EPE.2018.8601168

[8] N. Saha, S. Panda, Speed control with torque ripple reduction of switched reluctance motor by Hybrid Many Optimizing Liaison Gravitational Search

technique, Eng. Sci. Technol. an Int. J. 20 (3) (2017) pp. 909–921.
doi: 10.1016/j.jestch.2016.11.018

[9] S. J. Evangeline, S. Suresh Kumar, Minimization of Torque Ripple in Switched Reluctance Motor Drive – A Review, In: Lee J. (eds) Advanced Electrical and Electronics Engineering. Lecture Notes in Electrical Engineering, Springer, Berlin, Heidelberg, 2011, pp. 287–294.

[10] J. Ahn, Torque Control Strategy for High Performance SR Drive, J. Electr. Eng. Technol. 3 (4) (2008) pp. 538–545.

[11] O. Ellabban, H. Abu-Rub, Torque Control Strategies for a High Performance Switched Reluctance Motor Drive System, 7th IEEE GCC Conference and Exhibition, Doha, Qatar, 2013, pp. 257–262.
doi: 10.1109/IEEEGCC.2013.6705786

[12] D. a Shahakar, V. M. Jape, Direct Instantaneous Torque Control of Switched Reluctance Motors, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering 2 (1) (2013) pp. 574–579.

[13] J. Y. M. Cheung, Fuzzy logic control of refrigerant flow, In UKACC International Conference on Control, Exeter, UK, 1996, pp. 125–130.
doi: 10.1049/cp:19960538

[14] L.A.Zadeh, Fuzzy sets, Information and Control 8 (3) (1965) pp. 338-353.

[15] R. Abdel-Fadil, L. Számel, Fuzzy Logic Current Control of Switched Reluctance Motor for Electric Vehicles Applications, International Journal of Engineering and Information Systems (IJEAIS) 2 (4) (2018), pp. 19–28.

[16] Y. Bai, D. Wang, Fundamentals of Fuzzy Logic Control — Fuzzy Sets, Fuzzy Rules and Defuzzifications, In: Bai Y., Zhuang H., Wang D. (eds) Advanced Fuzzy Logic Technologies in Industrial Applications, Springer, London, 2006, pp. 17–36.

[17] R. Abdel-Fadil, A. Eid, M. Abdel-Salam, Fuzzy logic control of modern aircraft electrical power system during transient and steady-state operating conditions, in 2014 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), Mumbai, India, 2014, pp. 1–6.
doi: 10.1109/PEDES.2014.7042125

[18] R. B. Inderka, R. W. A. A. De Doncker. DITC-direct instantaneous torque control of switched reluctance drives, IEEE Transactions on Industrial Electronics 39 (4) (2003) pp. 1046-1051.
doi: 10.1109/TIA.2003.814578

[19] C. R. Neuhaus, N. H. Fuengwarodsakul, R. W. de Doncker, Predictive PWM-based Direct Instantaneous Torque Control of Switched Reluctance Drives, In IEEE 37th Power Electronics Specialists Conference, Jeju, South Korea, 2006, pp. 1-7.
doi: 10.1109/pesc.2006.1712264

[20] A. I. Al-Odienat, A. A. Al-Lawama, A. I. Al-Odienat, A. A. Al-Lawama, The Advantages of PID Fuzzy Controllers Over the Conventional Types, American Journal of Applied Sciences 5 (6) (2008) pp. 653–658.
doi: 10.3844/ajassp.2008.653.658

[21] R. Manikandan, R. Arulmozhiyal, Position control of DC servo drive using fuzzy logic controller, International Conference on Advances in Electrical Engineering, Vellore, India, 2014, pp. 1–5.
doi: 10.1109/ICAEE.2014.6838474

[22] PSIM | Software for Power Electronics Simulation [cited 2019-02-11]
URL https://powersimtech.com/products/psim/

[23] C. Li, G. Wang, Y. Fan, Y. Li, Adaptive RBF Neural Network Controller Design for SRM Drives, 35th Chinese Control Conference, Chengdu, China, 2016, pp. 6092-6097.
doi: 10.1109/ChiCC.2016.7554313

Acta
Technica
Jaurinensis

Vol. 12, No. 2, pp. 117-129, 2019
DOI: 10.14513/actatechjaur.v12.n2.497
Available online at acta.sze.hu

Artificial Neural Networks in Medicine

Z. T. Kocsis[1]

[1] Széchenyi István University, Department of Information Technology
Egyetem Tér 1, 9028, Győr, Hungary
e-mail: kocsis.zoltan@ga.sze.hu

Abstract: In recent years, Information Technology has been developed in a way that applications based on Artificial Intelligence have emerged. This development has resulted in machines being able to perform increasingly complex learning processes. The use of Information Technology, including Artificial Intelligence is becoming more and more widespread in all fields of life. Some common examples are face recognition in smartphones, or the programming of washing machines. As you may think, Artificial Intelligence can also be used in medicine. In this study I am presenting the relationship between machine learning and neural networks and their possible use in medicine.

Keywords: neural networks; medicine; surgery; patient; learning; healthcare

1. Introduction

People are able to learn from their acts, to interpret any mistakes they may have made, and thus to improve a series of repeated actions. However, people are not able to perform a series of actions in succession several times in a row exactly the same way. You may correct the previously committed mistake; however, another error might surface in the process. Nowadays, artificial intelligence systems developed for 'learning' that are capable of solving a pre-programmed problem with the help of a computer have become more and more common.

Artificial intelligence must have a problem-solving algorithm that helps it adapt to the problem. Artificial Intelligence (AI) is still being researched to enable computers to perform tasks that require human thinking, that is, self-learning.

In health care, it is almost impossible to find an area where human thinking is not needed. For example, to set up diagnose and a possible cure based on a patient's complaints human problem solving is necessary.

There are great opportunities for artificial intelligence during surgery. However, performing a surgery is not enough it must also be planned very carefully. There is a great need for the physician's experience and knowledge in both planning and implementing. However, surgery does not end well in all cases. Complications may occur during the operation, and a poorly performed surgery can lead to serious health damage, which also need to be treated by a physician.

The use of AI is already widespread in health care, used for example for collecting data about certain parts of the body, and for comparing surgical results and methods, and also for making suggestions based on these.

AI has several branches. In health care, machine learning is the one that can be used the best. In the following chapters, I am going to explain machine learning, and illustrate some specific examples of cases in which it could be applied in the healing process.

2. Machine learning in medicine

From a scientific point of view, machine learning is nothing more than making the machines suitable for making decisions with the help of the available data and situations using a learning algorithm. The biggest problem is how to create an algorithm that can learn the most effectively. There are two types of machine learning that are distinguished: supervised learning and unsupervised learning. They can be used well in health care [6].

2.1. Supervised learning

In supervised learning (Figure 1) we know the expected goal, that is, we want to teach the algorithm for a specific case. A good example fort this is the handwriting recognition software. We know what a word means, even if it is not written by the same person. The aim here is to recognize the different ways of writing and to identify words. A well-known example is the face recognition application found in mobile phones. Human faces are different but have the same features that AI can recognize. Supervised Learning focuses on grading, that is, selecting the pre-programmed known ones, classifying them, and finding the best of the selected items.

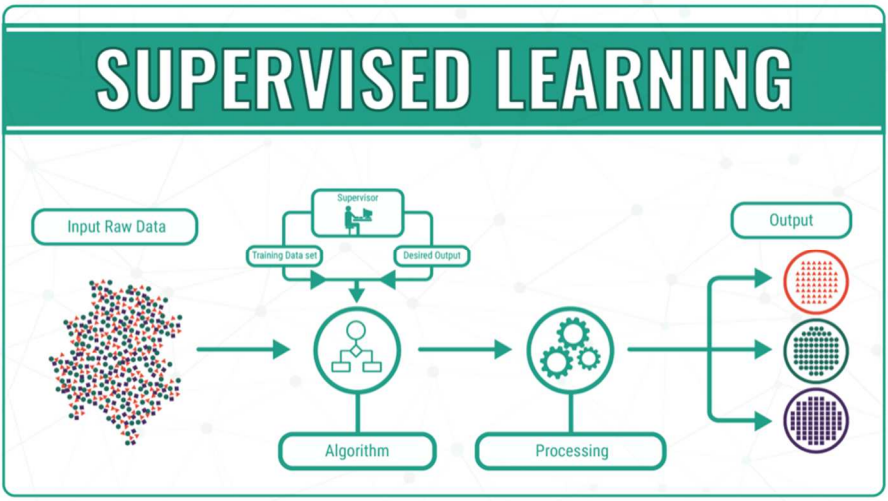
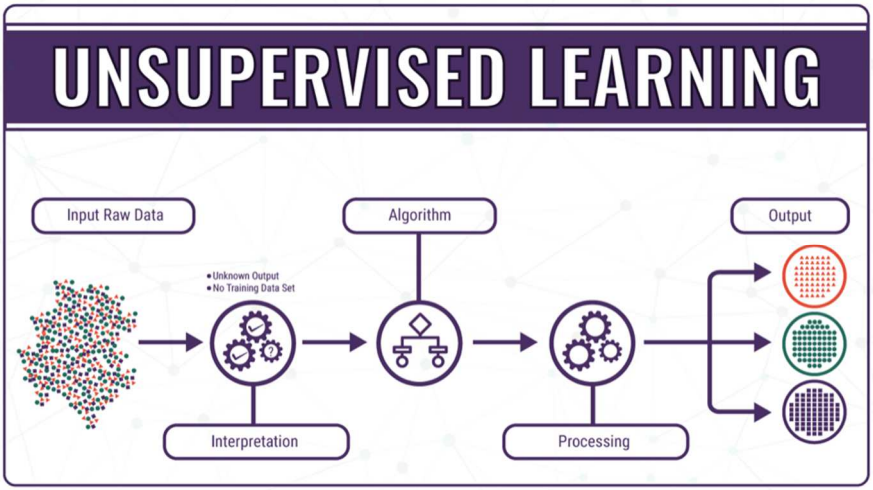


Figure 1. Supervised learning [15]

There are also cases in medicine that can be suitable for supervised learning. [1] In radiology, e.g. this method can be used to analyse a chest radiograph. The doctor knows which deformations he has to look for in the X-ray, so the computer can be programmed to look for these deformations as well (e.g., to recognize bone fractures). Supervised learning can be used by cardiologists to estimate risks. It is able to approach the risks posed by doctors in heart disease, but it also raises new risks that doctors would not notice. There are many other examples in medicine, but it can be seen that supervised learning might be widely used.

2.2. Unsupervised learning

Unsupervised or not supervised learning is a much more difficult task. Here we do not know what kind of outcome we expect, that is, we do not have anything specific. The algorithm tries to find the naturally occurring patterns or groups in the data. The results obtained are much more difficult to evaluate than in supervised learning because we do not know exactly what we want to see as the end result. We can only guess if the new results are useful in some way or not.



Fiugre 2. Unsupervised learning [15]

Unsupervised learning in health care is best used in precision medicine. The essence of precision medicine is to examine the subject as a whole. The patient's genetic characteristics, organ and emotional determinants and medical history are all taken into account. Thus, it is possible to redefine each therapeutic treatment. Another area of application may be for heart patients in the recognition of myocardial inflammation.

3. Machine learning in healthcare

3.1. Diagnosis with machine learning

In actual clinical practice, machine learning usage is limited. I would like to highlight two literature references. In the first case, cardiovascular diseases were tried to be diagnosed, while in the other they tried to diagnose cancer by machine learning. [1] [2] [7] [11]. The diagnosis was attempted by a classification method, where two groups were created. The task was to develop a model that can distinguish between the two classes. However, this is not easy because there are many symptoms that occur not only in patients with heart disease or cancer. The diagnosis was resolved with the help of a neural network.

The artificial neural network models some properties of the biological neural network taken from nature. This is an information tool capable of distributed

operation. The neurons in it represent the same operational elements and are arranged in an orderly manner. It has a learning algorithm that defines a learning process based on a sample and the way the information is processed. It must also have an algorithm that makes it possible to process the information it has acquired.

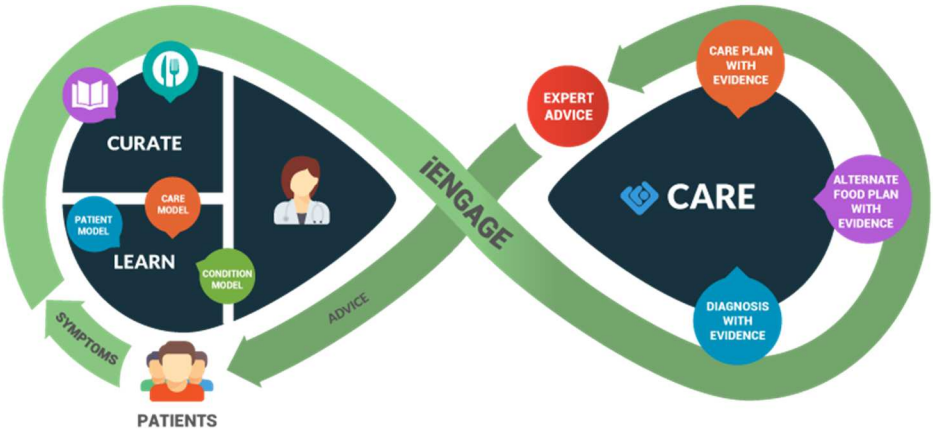


Figure 3. Machine learning in medicine [16]

Another example can be found in stroke research. Stroke is the fourth most common cause of death among adults. Stroke is a brain disorder that occurs when the blood supply to our brain deteriorates to an extent that results in the destruction of brain cells there. Early detection of this disease is an extremely difficult task, but it is essential for successful recovery. Machine learning contributes greatly to the timely detection of this disease. Between 2011 and 2014, a medical team in New York called for the Deep Learning machine learning method to detect the disease. [3][8] The essence of Deep Learning is that the machine can interpret an ever increasing amount of data quickly and accurately. It differs from traditional Artificial Intelligence (AI) by processing data step by step and the classification is performed by the algorithm as well.

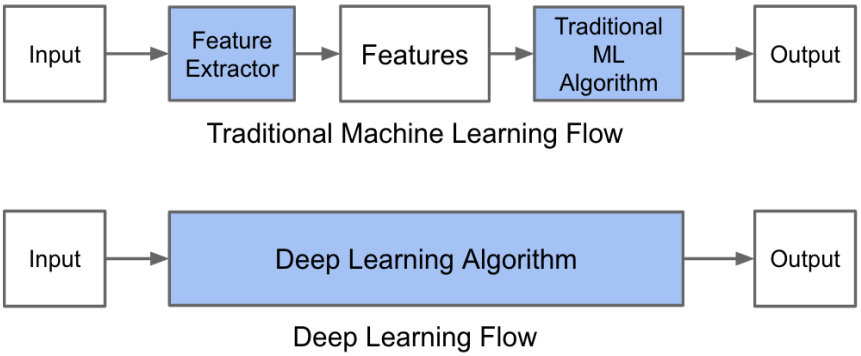


Figure 4. Machine learning vs. Deep Learning [17]

A total of 114 randomly selected patients were included in the study. Patients were given CT scans to identify the stroke. Based on these CT scan images a medical team made a diagnosis. The doctors have divided the images into two groups. One group became the data set from which the neural network could learn and the other group was the test set. With the help of these, a Deep Learning neural network capable of working in a 3-dimensional image was created based on the opinion of doctors on CT scans. The developed model was tested with the help of the set of test kits, by comparing the diagnosis established by the algorithm to the diagnosis of the medical team. Finally, the system has created a computer generated heat map to predict the possibility of infarction (Figure 5).

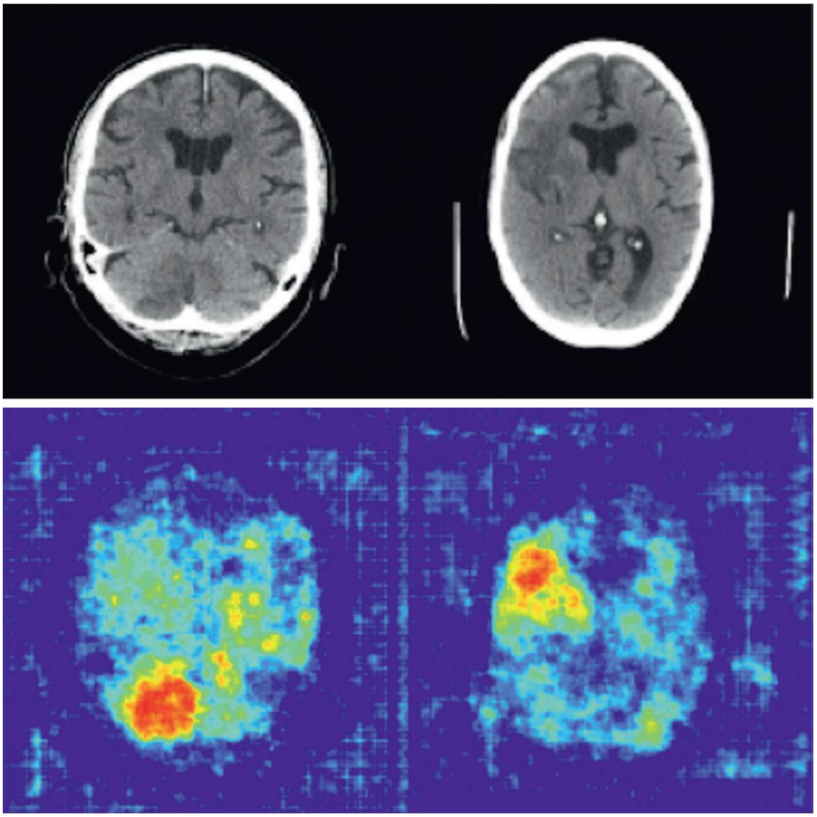


Figure 5. Computer generated heat map [3]

3.2. Machine learning during surgery

During surgery, doctors not only need to pay attention to the outcome of the operation, but also to the movement of the tools they use within the body. The movement of surgical instruments is monitored by the doctors' sight, and it is their responsibility to set their direction and position of these tools. By the automation of the traceability of devices, better surgical results can be achieved [9] [10].

There is already a so-called machine vision, which means that with the help of a camera the machine "sees" and processes the incoming images with the help of a software. In medicine, machine vision is an effective tool for tracking surgical instruments. Since they are controlled by software, altering them is unnecessary [4].

Machine vision tracking systems have been used for minimally invasive surgeries (MIS). The tracking of the tools was used both in 2 and 3 dimensions. The essence of these operations is that the intervention can be done with minimal cutting or possibly without cutting. Since the size of the cut is minimal, the accuracy of in-patient navigation is extremely important. Figure 6 shows a device for surgery.

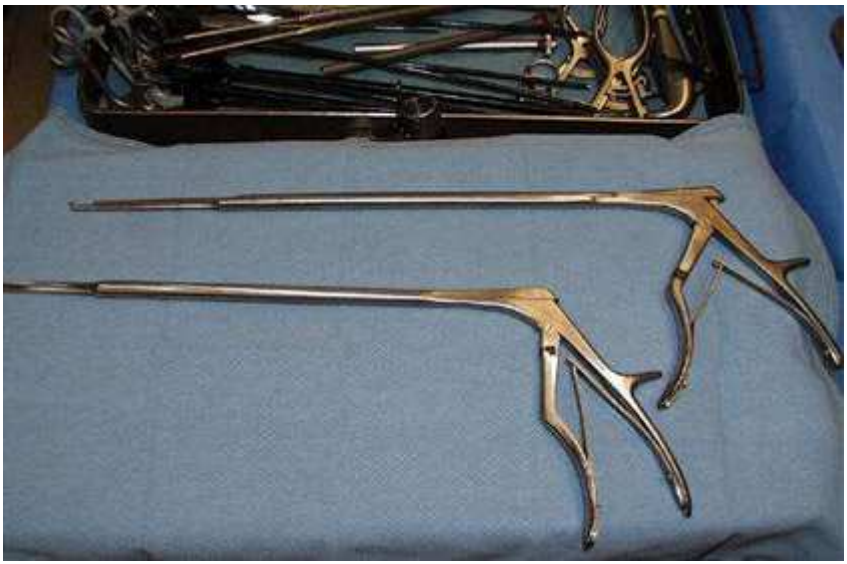


Figure 6. Minimal invasive surgery tool [12]

During surgery, the doctor only sees an image coming from an endoscope, which makes it difficult to coordinate the device and feel the depth. Nowadays, the data needed to track the device is acquired by electromagnetic, optical and vision-based techniques. Electromagnetic tracking techniques are costly and have a complicated tracking algorithm. In contrast, vision-based systems operate much more efficiently and less costly. Software fitted to the camera can measure the position of the camera and monitor the device accordingly, so there is no need to change the device on the surgical procedure or the device used for surgery. During surgery, the position of the end of the device should be measured, which is not an easy task due to continuous movement. Tracking the tool was solved by machine vision and a neural network specifically developed for the task. This created a system that monitored the device with good results and immediately informed the physician performing the surgery (Figure 7).

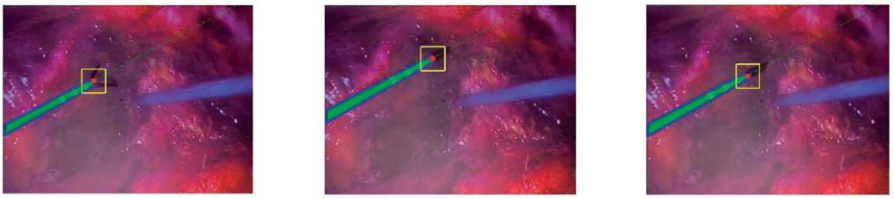


Figure 7. Tracking detection [4]

3.3. Deep Learning and Computer Tomography

So far we have mainly looked at examples of machine learning in only 2 dimensions. However, CT images can provide 3-dimensional images that contain useful information about human anatomy that can be used to diagnose and support surgery and therapies. [5]

Recognizing and separating human body parts play a critical role in the interpretation of CT images. In the traditional sense, they all have to be done by radiologists, but these are time-consuming tasks and require great care. These could be automated, that is, using automated image recognition software, which would, on the one hand, take the burden off the doctors' shoulder and on the other hand be faster and more reliable.

This task is challenging in several respects, because medical imaging procedures never give back the 100% exact image of body parts, so it is not possible to create a precise mathematical model for identifying body parts. Another disadvantage of CT is that it is not possible to set a contrast on a recording as a conventional camera, so there will be blurred boundaries between the individual parts of the body. CT images can be made from every part of the body, but every shot requires different settings, so if you want to recognize each part of the body with a specific method, you should make a special method for each option.

Convolutional neural network (CNN) was used to implement CT image segmentation and solve various problems. [5] [13] [14] In all cases, CT images are made in 2 dimensions, which represent the patient in horizontal or vertical sections, slices. The radiologist should be able to interpret these images and, if necessary, create a 3-dimensional image in the head to determine the specific problems.

However, the diagnostic software is able to combine the slices of CT scans both horizontally and vertically into a 3-dimensional image. From the 3-dimensional images thus obtained, a neural network of problems can already determine the individual parts of the body (Figure 8).

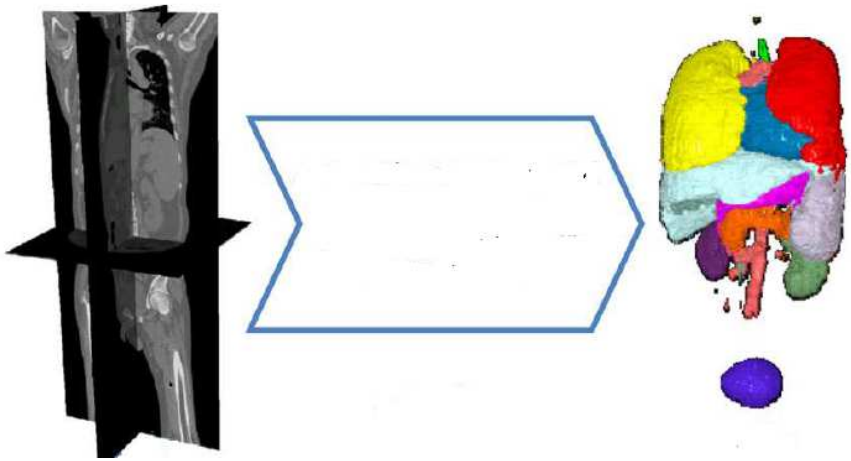


Figure 8. 2D to 3D CT image [5]

Figure 8 shows how the 3-dimensional image and how the individual body parts are reunited from the individual slices of the CT images (vertical and horizontal). With the help of colours, each part of the body can easily be recognized. However, the neural network can be taught for one kind of problem, so you don't have to recognize every part of the body at the same time. The teacher data set must be specified for the given task, so it will only recognize that part of the body.

4. Conclusions

Based on these new procedures we can say that machine learning is a great help for doctors, both in diagnostics and in surgery.

The method also greatly facilitates the processing of CT images, since even an experienced physician may not notice something in the image being examined. Unlike a physician, a software developed for a particular problem is able to focus specifically on the problem, so it does not have to deal with other distractions. If you use a neural network to do this, it will fix itself after every task that it has solved, i.e. it will be more accurate and reliable.

Apart from diagnostics, machine learning can be used during surgery too. In the case presented, the aim was to follow the surgical device and to make it more accurate. In contrast to the tools used so far, the use of the neural network did not require any modification of the surgical intervention or the device. I think neural networks can be used to determine a specific surgical area and intervention scenario.

In connection with my research, I also investigate what kind of mechanical learning or neural networking has already been used in spinal surgery and how I could use this information in reaching my goal

References

[1] Rahul C. Deo, MD, PhD, Machine Learning in Medicine, Circulation 132 (2015) pp. 1920-1930.
doi:10.1161/circulationaha.115.001593

[2] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, Deep Learning for Identifying Metastatic Breast Cancer, arXiv preprint (2016) pp. 1-6.
arXiv:1606.05718

[3] Ashley N. Beecy, Qi Chang, Khalil Anchouche, Lohendran Baskaran, Kimberly Elmore, Kranthi Kolli, Hao Wang, Subhi Al'Aref, Jessica M. Peña, Ashley Knight-Greenfield, Praneil Patel, Peng Sun, Tong Zhang, Hooman Kamel, Ajay Gupta and James K. Min, A Novel Deep Learning Approach for Automated Diagnosis of Acute Ischemic Infarction on Computed Tomography, JACC: Cardiovascular Imaging 11 (11) (2018) pp. 1723-1725.
doi: 10.1016/j.jcmg.2018.03.012

[4] Zhao, Z., Voros, S., Weng, Y., Chang, F., and Li, R., Tracking-by-detection of surgical instruments in minimally invasive surgery via the convolutional neural network deep learning-based method. Computer Assisted Surgery, 22 (sup1) (2017) pp. 26-35.
doi:10.1080/24699322.2017.1378777

[5] Zhou, X., Takayama, R., Wang, S., Hara, T., and Fujita, H., Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. Medical Physics, 44 (10) (2017) pp. 5221-5233.
doi:10.1002/mp.12480

[6] Obermeyer, Z., and Emanuel, E. J., Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. New England Journal of Medicine, 375 (13) (2016) pp. 1216-1219.
doi:10.1056/nejmp1606181

[7] Cabitza, F., Rasoini, R., and Gensini, G. F., Unintended Consequences of Machine Learning in Medicine. JAMA, 318 (6) (2017) pp. 517-518. doi:10.1001/jama.2017.7797

[8] Leung, M. K. K., Delong, A., Alipanahi, B., & Frey, B. J., Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets. Proceedings of the IEEE, 104 (1) (2016) pp. 176-197. doi:10.1109/jproc.2015.2494198

[9] Smith, L. N., Smith, M. L., Fletcher, M. E., & Henderson, A. J., A 3D machine vision method for non-invasive assessment of respiratory function. The International Journal of Medical Robotics and Computer Assisted Surgery, 12 (2) (2015) pp. 179-188. doi:10.1002/rcs.1669

[10] Din, N., Smith, P., Emeriewen, K., Sharma, A., Jones, S., Wawrzynski, J., Saleh, G. M., Man versus Machine: Software Training for Surgeons—An Objective Evaluation of Human and Computer-Based Training Tools for Cataract Surgical Performance. Journal of Ophthalmology 2016 (3548039) (2016) pp. 1-7. doi:10.1155/2016/3548039

[11] Würfl, T., Ghesu, F. C., Christlein, V., and Maier, A., Deep Learning Computed Tomography. Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2016. Lecture Notes in Computer Science, Springer, Cham, 9902 (2016) pp. 432-440. doi:10.1007/978-3-319-46726-9_50

[12] George D. Picetti, Basic Introduction to Minimally Invasive Spine Surgery [cited 2019.01.22] URL: https://www.spineuniverse.com/treatments/surgery/minimally-invasive/basic-introduction-minimally-invasive-spine-surgery

[13] Van Ginneken, B., Setio, A. A. A., Jacobs, C., and Ciompi, F., Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI) (2015) pp. 286-289. doi:10.1109/isbi.2015.7163869

[14] Li, W., Cao, P., Zhao, D., and Wang, J., Pulmonary Nodule Classification with Deep Convolutional Neural Networks on Computed Tomography Images. Computational and Mathematical Methods in Medicine 2016 (6215085) (2016) pp. 1-7.
doi:10.1155/2016/6215085

[15] Ronald van Loon, Machine Learning Explained: Understanding Supervised, Unsupervised, and Reinforcement Learning [cited 2019.01.22]
URL: https://datafloq.com/read/machine-learning-explained-understanding-learning/4478

[16] COGNUB Decision Solutions, Healthcare with Machine Learning [cited 2019.01.22]
URL: http://www.cognub.com/index.php/health-care-2/

[17] Adil Moujahid, A Practical Introduction to Deep Learning with Caffe and Python [cited 2019.01.22]
URL: http://adilmoujahid.com/posts/2016/06/introduction-deep-learning-python-caffe/

Acta
Technica
Jaurinensis

Vol.12., No.2., pp. 130–147, 2019
DOI: 10.14513/actatechjaur.v12.n2.499
Available online at acta.sze.hu

State Space Based Linear Controller Design for the Inverted Pendulum

M. Kuczmann

Széchenyi István University, Department of Automation
Egyetem tér 1, H-9026, Győr, Hungary
E-mail: kuczmann@sze.hu

Abstract: In a previous survey paper the detailed PID controller design to stabilize the inclination angle as well as the horizontal movement of an inverted pendulum system has been presented. In this paper the linear controller design based on the state space representation is shown step by step. Pendulum model is based on Euler-Lagrange modeling, and the non-linear state space model is linearized in the unstable upward position, finally pole placement by Ackermann formula and Bass–Gura equation, moreover linear quadratic optimal control are presented. The pendulum has been inserted into a virtual reality laboratory, which is suitable to use in model based control teaching.

Keywords: *inverted pendulum, model based control, pole placement, optimal control*

1. Introduction

This paper presents a comprehensive study of controller design for an inverted pendulum mounted on a cart which can only move horizontally. Design is based on the state space representation of the plant. PID controller design with some drawbacks has been shown in the latter paper [1].

The inverted pendulum is an unstable system that must be stabilized by the pushing-pulling force $F = F(t)$ acting on the cart by an electric motor (Fig. 1), i.e. to reach

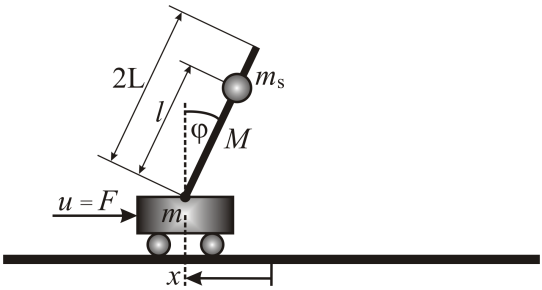


Figure 1. The inverted pendulum model.

the inclination angle φ being zero. The pendulum simply falls over if the cart is not moved to balance it. By the presented methods, the horizontal movement can also be easily stabilized.

The studied plant is a popular example commonly found in control system textbooks and research literature [1–8]. The dynamics of the system are nonlinear as presented in the paper based on the above mentioned literature, but controller design is based on the linearized system.

The aim of this survey paper is to show the Euler-Lagrange modeling of the inverted pendulum system, then the linear state feedback controller design step by step. Pole placement technique and linear quadratic regulator are presented. The mentioned formulations are deeply studied, and the current paper can be used in teaching of model based control.

The latter paper [1] concluded that, two controllers are necessary to design for stabilizing the inverted pendulum. The design of two dependent controllers is tedious in some cases, however the state feedback controllers can solve this problem in an easy way. This is shown in this paper.

The real operating device has not built in this research, however the virtual reality based implementation has been performed which is applicable to understand the steps of control design.

2. Dynamic model of the pendulum

To set up the dynamic model of pendulum, the Euler-Lagrange equation is applied,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial K}{\partial \dot{q}_i} - \frac{\partial K}{\partial q_i} + \frac{\partial P}{\partial \dot{q}_i} = \tau_i, \tag{1}$$

where K is the kinetic energy, P is the potential energy, q_i and τ_i are the generalized coordinates and the generalized torque (force), respectively. In the case of pendulum $i = 1{,}2$, i.e. $q_1 = x$ and $q_2 = \varphi$, moreover $\tau_1 = F$ and $\tau_2 = 0$.

The kinetic energy of the system is as follows:

$$K = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}Mv_M^2 + \frac{1}{2}\Theta\dot{\varphi}^2 + \frac{1}{2}m_{\mathrm{s}}v_{\mathrm{s}}^2, \tag{2}$$

with the mass of cart, m, the mass of rod, M, and the mass of the sphere m_{s}. The inertial moment of the rod belonging to the center of mass is $\Theta = \frac{1}{3}ML^2$ (the length of the rod is $2L$). The sphere can be moved along the rod, the distance from the shaft is l. The velocity of the center of mass of the rod as well as the sphere are v_M and v_{s}, respectively [1],

$$v_M^2 = \dot{x}^2 + 2LC_\varphi\dot{x}\dot{\varphi} + L^2\dot{\varphi}^2, \quad v_{\mathrm{s}}^2 = \dot{x}^2 + 2lC_\varphi\dot{x}\dot{\varphi} + l^2\dot{\varphi}^2. \tag{3}$$

The potential energy of the system is

$$P = MgLC_\varphi + m_{\mathrm{s}}glC_\varphi, \tag{4}$$

where g is the gravitational acceleration. For simplicity, $S_\varphi = \sin\varphi$ and $C_\varphi = \cos\varphi$ notations are used in the paper.

After obtaining the terms in (1), and doing some manipulations, the following differential equations can be got (see [1] for a similar pendulum):

$$\begin{aligned}
&(m + M + m_{\mathrm{s}})\ddot{x} + (ML + m_{\mathrm{s}}l)C_\varphi\ddot{\varphi} - (ML + m_{\mathrm{s}}l)S_\varphi\dot{\varphi}^2 = F,\\
&(ML + m_{\mathrm{s}}l)C_\varphi\ddot{x} + (ML^2 + \Theta + m_{\mathrm{s}}l^2)\ddot{\varphi} - (ML + m_{\mathrm{s}}l)gS_\varphi = 0.
\end{aligned} \tag{5}$$

From these equations, the following second order derivatives can be yielded:

$$\ddot{\varphi} = \frac{-\theta_{\mathrm{R}}^2 S_\varphi C_\varphi \dot{\varphi}^2 + M_\Sigma \theta_{\mathrm{R}} g S_\varphi - F \theta_{\mathrm{R}} C_\varphi}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2 C_\varphi^2}, \tag{6}$$

and

$$\ddot{x} = \frac{\theta_{\mathrm{R}} \Theta_{\mathrm{R}} S_\varphi \dot{\varphi}^2 - \theta_{\mathrm{R}}^2 g S_\varphi C_\varphi + F \Theta_{\mathrm{R}}}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2 C_\varphi^2}. \tag{7}$$

Here the following notations have been introduced: $M_\Sigma = m + M + m_{\mathrm{s}}$, $\theta_{\mathrm{R}} = ML + m_{\mathrm{s}} l$, $\Theta_{\mathrm{R}} = \frac{4}{3} ML^2 + m_{\mathrm{s}} l^2$.

The two second order differential equations can be rewritten as four first order differential equations by introducing state variables: $x_1 = x$, $x_2 = \dot{x}$, $x_3 = \varphi$, $x_4 = \dot{\varphi}$, i.e. $x_2 = \dot{x}_1$ and $x_4 = \dot{x}_3$. Finally, the state space representation of the dynamic modell is the following:

$$\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \frac{\theta_{\mathrm{R}} \Theta_{\mathrm{R}} S_{x_3} x_4^2 - \theta_{\mathrm{R}}^2 g S_{x_3} C_{x_3} + F \Theta_{\mathrm{R}}}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2 C_{x_3}^2}, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= \frac{-\theta_{\mathrm{R}}^2 S_{x_3} C_{x_3} x_4^2 + M_\Sigma \theta_{\mathrm{R}} g S_{x_3} - F \theta_{\mathrm{R}} C_{x_3}}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2 C_{x_3}^2}.
\end{aligned} \tag{8}$$

This nonlinear system can be linearized in the unstable upright position, when $\varphi = 0$ and $\dot{\varphi} = 0$, i.e. the approximations $S_\varphi \cong \varphi$ and $C_\varphi \cong 1$ can be applied.

At the end, the linearized system can be modeled by the state space equations

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u, \tag{9}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{-\theta_{\mathrm{R}}^2 g}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{M_\Sigma \Theta_{\mathrm{R}} g}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2} & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ \frac{\Theta_{\mathrm{R}}}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2} \\ 0 \\ -\frac{\theta_{\mathrm{R}}}{M_\Sigma \Theta_{\mathrm{R}} - \theta_{\mathrm{R}}^2} \end{bmatrix}, \tag{10}$$

and $\mathbf{x} = [x_1\, x_2\, x_3\, x_4]^{\mathrm{T}}$ is the vector of the state variables.

For simplicity, the following notations are used:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & q & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & p^2 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ \beta \\ 0 \\ \alpha \end{bmatrix}. \tag{11}$$

The following data are used in the paper to represent the results: $m = 5$, $M = 10$, $m_{\mathrm{s}} = 50$, $2L = 20$, $l = 20$, in a coherent unit system, i.e. $\alpha = -0.0062$, $p = 1.9925$, $\beta = 0.1208$, $q = -67.1893$.

3. Pole placement control

The typical closed loop block diagrams of the very basic state feedback controller system is shown in Fig. 2. The block diagrams are equivalent, however both will be supplemented. Scalar-valued signals are represented by dashed line, the others are vector-valued.

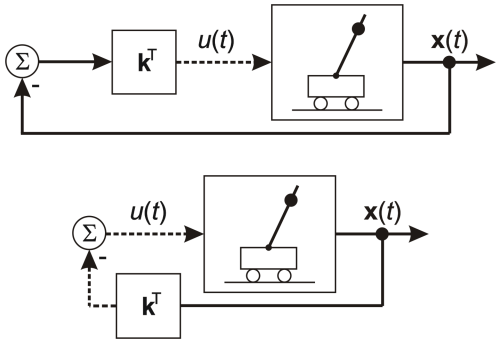


Figure 2. Block diagrams of the state feedback system.

The pendulum can be represented by the nonlinear ordinary differential equations (8), called the nonlinear model, while the linearized model is given by (9) and (10). It

is highlighted that the controller design is based on the linearized model, however the real life pendulum is nonlinear in nature. Linear and nonlinear models are compared in the paper.

The control signal is defined as the weighted sum of the state variables, i.e.

$$u = -\mathbf{k}^{\mathrm{T}}\mathbf{x} = -k_1 x_1 - k_2 x_2 - \cdots - k_N x_N. \tag{12}$$

In this application $N = 4$, and $\mathbf{k}^{\mathrm{T}} = [k_1\, k_2\, k_3\, k_4]$ contains the unknown parameters of the controller that must be designed.

The eigenvalues λ of the system matrix $\mathbf{A}$ can be computed by the characteristic equation

$$\varphi(\lambda) = |\lambda\mathbf{I} - \mathbf{A}| = \lambda^N + a_1\lambda^{N-1} + a_2\lambda^{N-2} + \cdots + a_{N-1}\lambda + a_N, \tag{13}$$

where $\mathbf{I}$ is the identity matrix, and the coefficients $a_1, a_2, \cdots, a_N$ are known building up the vector $\mathbf{a}^{\mathrm{T}} = [a_1\, a_2\, \cdots\, a_N]$. The transient behavior of the plant is depending on the eigenvalues, which can be modified by the negative feedback.

The state space representation of the closed loop system is as follows:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u = (\mathbf{A} - \mathbf{b}\mathbf{k}^{\mathrm{T}})\mathbf{x}, \tag{14}$$

i.e. the system matrix $\mathbf{A}$ of the open loop system is modified by the term $-\mathbf{b}\mathbf{k}^{\mathrm{T}}$.

The eigenvalues of the closed loop system are determined by

$$\varphi_{\mathrm{cl}} = |\lambda\mathbf{I} - (\mathbf{A} - \mathbf{b}\mathbf{k}^{\mathrm{T}})| = \lambda^N + p_1\lambda^{N-1} + p_2\lambda^{N-2} + \cdots + p_{N-1}\lambda + p_N. \tag{15}$$

It is important to recognize, that the coefficients $p_1, p_2, \cdots, p_N$ are not known, because $\mathbf{k}^{\mathrm{T}}$ is unknown yet. The coefficients can be written as a vector $\mathbf{p}^{\mathrm{T}} = [p_1\, p_2\, \cdots\, p_N]$.

The first step of the design is to determine the desired value of the coefficients in $\mathbf{p}^{\mathrm{T}}$ (i.e. the eigenvalues of the closed loop system), then to find the appropriate feedback gains.

The Bass–Gura equation and the Ackermann formula are well known in the literature to design the gain vector $\mathbf{k}^{\mathrm{T}}$ [2–6].

Next, the controllability of the plant must be checked by analyzing the controllability matrix

$$\mathbf{M}_{\mathrm{c}} = [\mathbf{b}\ \mathbf{A}\mathbf{b}\ \mathbf{A}^2\mathbf{b}\ \cdots\ \mathbf{A}^{N-1}\mathbf{b}]. \tag{16}$$

Here $\mathbf{M}_{\mathrm{c}}$ can be computed by using the notations of (11),

$$\mathbf{M}_{\mathrm{c}} = \begin{bmatrix} 0 & \beta & 0 & q\alpha \\ \beta & 0 & q\alpha & 0 \\ 0 & \alpha & 0 & p^2\alpha \\ \alpha & 0 & p^2\alpha & 0 \end{bmatrix}. \tag{17}$$

It is easy to see, that the matrix has four independent columns, i.e. the rank of $\mathbf{M}_{\mathrm{c}}$ is maximum. It means that, the system is controllable and state feedback methods can be applied. The inverse of $\mathbf{M}_{\mathrm{c}}$ appears in the design process.

The Bass–Gura equation looks like

$$\mathbf{k}^{\mathrm{T}} = \left(\mathbf{p}^{\mathrm{T}} - \mathbf{a}^{\mathrm{T}}\right) \tau(\mathbf{a})^{-1} \mathbf{M}_{\mathrm{c}}^{-1}, \tag{18}$$

where $\tau(\mathbf{a})$ is the following Toeplitz matrix

$$\tau(\mathbf{a}) = \begin{bmatrix} 1 & a_1 & a_2 & a_3 & \cdots & a_{N-1} \\ 0 & 1 & a_1 & a_2 & \cdots & a_{N-2} \\ \vdots & & \ddots & \ddots & \cdots & \vdots \\ \vdots & & & \ddots & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{19}$$

The Ackermann formula has the form

$$\mathbf{k}^{\mathrm{T}} = [0\,0\,0 \cdots 0\,1]\mathbf{M}_{\mathrm{c}}^{-1} \varphi_{\mathrm{cl}}(\mathbf{A}). \tag{20}$$

Of course the row vector is $[0\,0\,0\,1]$ in this situation, moreover

$$\varphi_{\mathrm{cl}}(\mathbf{A}) = \mathbf{A}^N + p_1\mathbf{A}^{N-1} + p_2\mathbf{A}^{N-2} + \cdots + p_{N-1}\mathbf{A} + p_N\mathbf{I}. \tag{21}$$

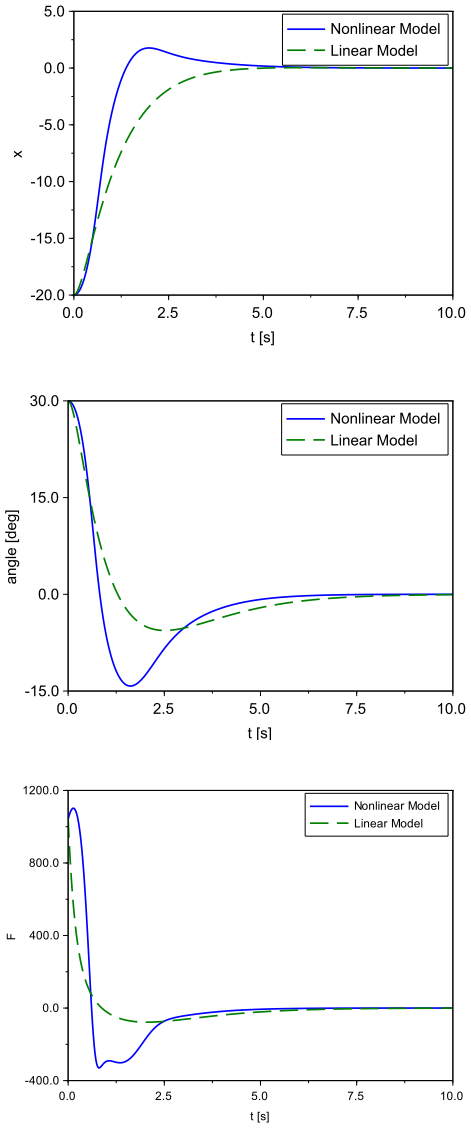


Figure 3. Stabilization of the horizontal movement (top) and the inclination angle (middle) by the control signal (bottom).

Stabilization of the horizontal movement as well as the inclination angle can be performed easily by the pole placement methods as it is shown in Fig. 3, where the control force is also plotted. A significant difference between the linear and nonlinear model behavior can be seen, because of the large initial values. If the initial values are much smaller, then the models output are very close to each other.

In this illustration $\mathbf{p}^{\mathrm{T}} = [8\ 18\ 16\ 5]$ has been applied, i.e. the desired eigenvalues of the closed loop system are $\lambda_1 = -5$, $\lambda_{2,3,4} = -1$. The plant itself is unstable with the following eigenvalues: $\lambda_{1,2} = 0$, $\lambda_{3,4} = \pm 1.9925$, i.e. $\mathbf{a}^{\mathrm{T}} = [0\ -3.97\ 0\ 0]$. The resulting feedback vector is $\mathbf{k}^{\mathrm{T}} = [-81.86\ -261.95\ -5116.12\ -6365.03]$.

Reference tracking can be realized by simple feedforward branches as shown in Fig. 4. It is noted that, there is no feedback from the output of the plant resulting in sensitive and not robust controller, however the pendulum is stabilized.

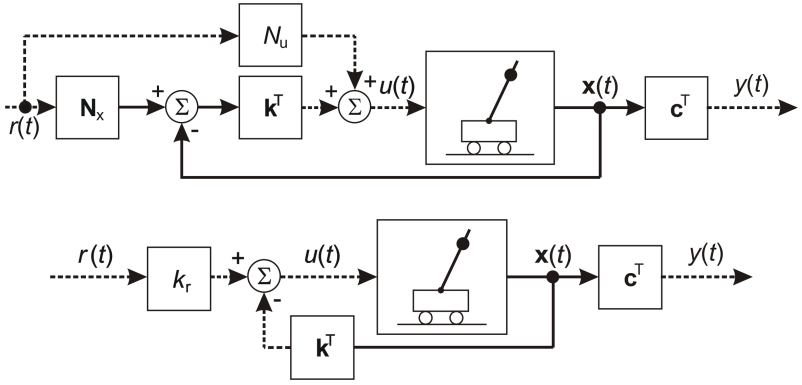


Figure 4. The state feedback system with reference trancking.

The control signal is as follows

$$u = \mathbf{k}^{\mathrm{T}} \left(\mathbf{N}_{\mathrm{x}} r - \mathbf{x} \right) + N_{\mathrm{u}} r, \tag{22}$$

or

$$u = -\mathbf{k}^{\mathrm{T}} \mathbf{x} + k_{\mathrm{r}} r, \tag{23}$$

in the first and second setup, respectively.

The gains in the first realization can be calculated as

$$\begin{bmatrix} \mathbf{N}_{\mathrm{x}} \\ N_{\mathrm{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^{\mathrm{T}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}, \tag{24}$$

and the gain

$$k_{\mathrm{r}} = \frac{-1}{\mathbf{c}^{\mathrm{T}} \left(\mathbf{A} - \mathbf{b}\mathbf{k}^{\mathrm{T}} \right)^{-1} \mathbf{b}} \tag{25}$$

is applied in the second block diagram.

Fig. 5 shows a simple reference tracking problem by the pendulum. The cart is at $x = -10$ at the initial state, where $\varphi = 30°$. The reference signal is jumping from -10 to $+10$, then, after $10\,\mathrm{s}$ it is jumping back to the original state. The cart is following the reference signal, moreover the pendulum is stabilized. Controller gains are the same as above, feedforward gains are set as mentioned in (24) and (25).

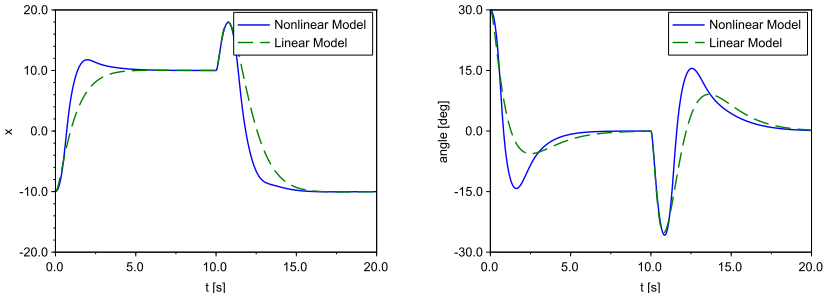


Figure 5. Reference signal tracking.

Reference tracking properties can be improved by a feedback including an integrator as shown in Fig. 6. The variable $z = z(t)$ is a new state variable, and the augmented

system representation is as follows:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{c}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ z \end{bmatrix} + \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} u, \tag{26}$$

where the control signal is

$$u = -\begin{bmatrix} \mathbf{k}^{\mathrm{T}} & k_{\mathrm{I}} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ z \end{bmatrix}. \tag{27}$$

Here $k_{\mathrm{I}} = N_{\mathrm{u}}$ and $k_{\mathrm{I}} = k_{\mathrm{r}}$, respectively (see notations in Fig. 6). The augmented vector $\begin{bmatrix} \mathbf{k}^{\mathrm{T}} & k_{\mathrm{I}} \end{bmatrix}$ can be calculated by the Bass–Gura or the Ackermann formula.

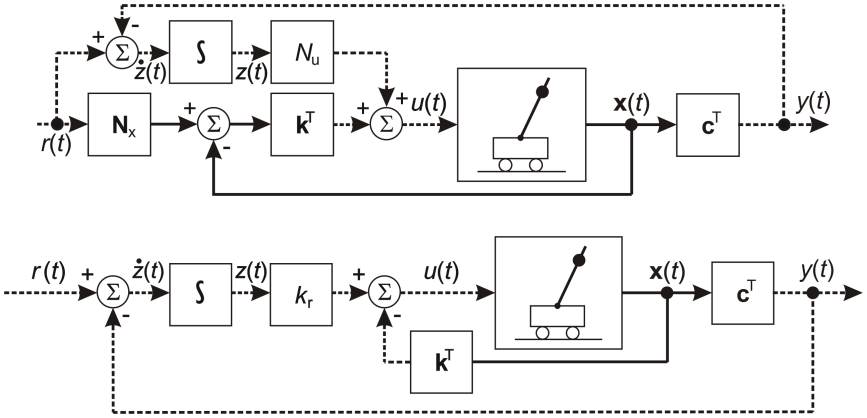


Figure 6. The state feedback system with integrator and reference trancking.

The mentioned pole placement techniques require to measure all the states. The horizontal movement and speed, moreover the inclination angle and angular speed must be measured in this case. It is easy to realize, however, in many cases it is not possible to pick up all the states, that is why state observers are designed. State observers estimate the states from measurements of the input and output of the real system as it is shown in Fig. 7.

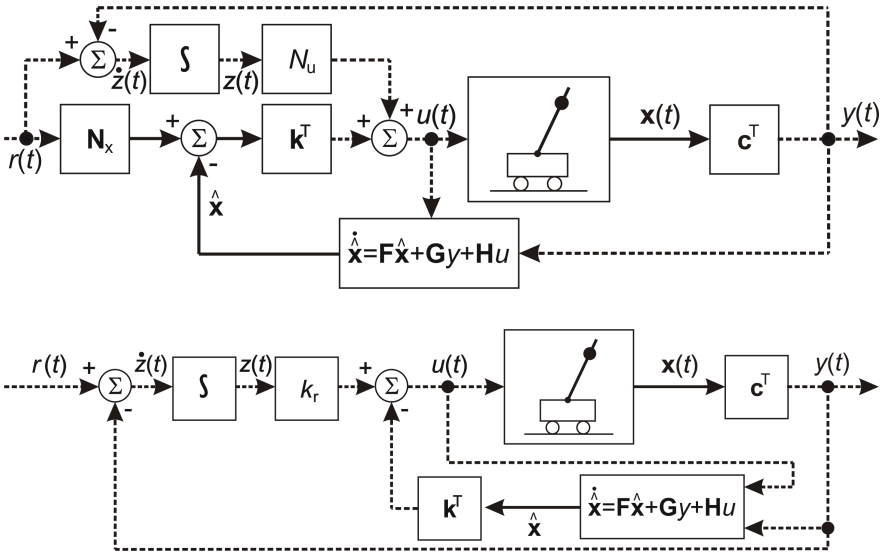


Figure 7. Block diagram of the state feedback system with observer.

The pole placement and observer design are dual to each other. The eigenvalues of the system augmented by the observer are defined by

$$\varphi_{\mathrm{o}} = |\lambda \mathbf{I} - (\mathbf{A} - \mathbf{G}\mathbf{c}^{\mathrm{T}})| = \lambda^N + f_1\lambda^{N-1} + f_2\lambda^{N-2} + \cdots + f_{N-1}\lambda + p_N. \quad (28)$$

The coefficients $f_1, f_2, \cdots, f_N$ can be written as a vector $\mathbf{f}^{\mathrm{T}} = [f_1\ f_2\ \cdots\ f_N]$.

The first step of the observer design is to determine the desired value of these coefficients (i.e. the eigenvalues), then to find the appropriate observer gain $\mathbf{G}$. The Bass–Gura equation and the Ackermann formula can also be applied.

Next, the observability of the plant must be checked, i.e. the observability matrix

$$\mathbf{M}_{\mathrm{o}} = \begin{bmatrix} \mathbf{c}^{\mathrm{T}} \\ \mathbf{c}^{\mathrm{T}}\mathbf{A} \\ \mathbf{c}^{\mathrm{T}}\mathbf{A}^{2} \\ \vdots \\ \mathbf{c}^{\mathrm{T}}\mathbf{A}^{N-1} \end{bmatrix} \tag{29}$$

must have full rank. It can only be reached if the horizontal movement of the cart is measured, i.e.

$$\mathbf{c}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \quad D = 0, \tag{30}$$

otherwise the rank of the observability matrix is less then four. The full rank observability matrix is according to the notations of (11) is

$$\mathbf{M}_{\mathrm{o}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & q & 0 \\ 0 & 0 & 0 & q \end{bmatrix}. \tag{31}$$

Duality means a simple change of the matrices: $\mathbf{A} \to \mathbf{A}^{\mathrm{T}}$, $\mathbf{b} \to \mathbf{b}^{\mathrm{T}}$, $\mathbf{M}_{\mathrm{c}} \to \mathbf{M}_{\mathrm{o}}^{\mathrm{T}}$, $\mathbf{k} \to \mathbf{G}$, with which the Bass–Gura equation has the form

$$\mathbf{G}^{\mathrm{T}} = \left(\mathbf{f}^{\mathrm{T}} - \mathbf{a}^{\mathrm{T}}\right) \tau(\mathbf{a})^{-1} \mathbf{M}_{\mathrm{o}}^{-\mathrm{T}}, \tag{32}$$

while the Ackermann equation is given as

$$\mathbf{G}^{\mathrm{T}} = [0\,0\,0 \cdots 0\,1] \mathbf{M}_{\mathrm{o}}^{-\mathrm{T}} \varphi_{\mathrm{o}}(\mathbf{A}^{\mathrm{T}}), \tag{33}$$

then the other terms of the observer are determined by

$$\mathbf{F} = \mathbf{A} - \mathbf{G}\mathbf{c}^{\mathrm{T}}, \quad \mathbf{H} = \mathbf{b}. \tag{34}$$

Fig. 8 and Fig. 9 show comparisons between observer estimated and the measured signals in the case of small initial value and a larger one, respectively. The first one is very close to the linearized model, resulting very good observer performance. The linear observer is not advantageous when the real life problem is far from the stationary point. Observer eigenvalues have been set to $\lambda_{1,2,3,4} = -2$ in this example.

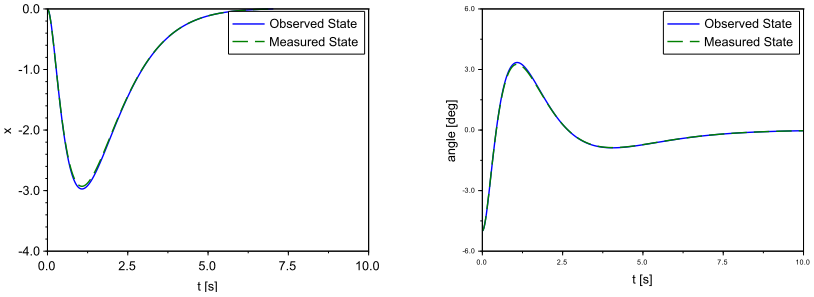


Figure 8. Comparison of observed and measured state variables x and φ, $\varphi(0) = -5°$.

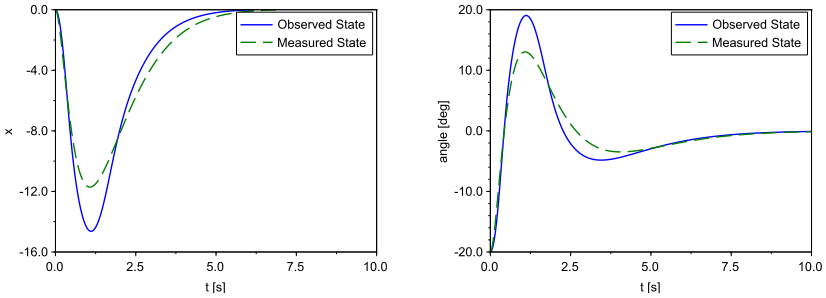


Figure 9. Comparison of observed and measured state variables x and φ, $\varphi(0) = -20°$.

4. Linear quadratic optimal control

Linear quadratic regulator (LQR) minimizes the following functional:

$$J(\mathbf{x},u) = \frac{1}{2}\int_0^\infty \left[\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + ru^2\right] \mathrm{d}t. \tag{35}$$

The first term in the integral ensures stability ($\mathbf{x} \rightarrow \mathbf{0}$), while the second term minimizes the supply effort for the control. Matrix $\mathbf{Q}$ is symmetric and positive semi-definite, moreover $r > 0$. These design parameters determine the relative importance of error in $\mathbf{x}$ and the supplied energy.

The system is defined by (9), with the initial condition $\mathbf{x}(t_0)$.

The optimal control law minimizing the functional (35) is a state feedback control, given by

$$u = -\frac{1}{r}\mathbf{b}^{\mathrm{T}}\mathbf{P}\mathbf{x} = -\mathbf{k}^{\mathrm{T}}\mathbf{x}, \tag{36}$$

where the symmetric positive definite matrix $\mathbf{P}$ is the solution of the control algebraic Riccati equation

$$\mathbf{P}\mathbf{A} + \mathbf{A}^{\mathrm{T}}\mathbf{P} - \frac{1}{r}\mathbf{P}\mathbf{b}\mathbf{b}^{\mathrm{T}}\mathbf{P} + \mathbf{Q} = \mathbf{0}. \tag{37}$$

Riccati equation can be solved by numerical techniques. The following gains have been obtained when $\mathbf{Q} = 5\mathbf{I}$ and $r = 1$: $\mathbf{k}^{\mathrm{T}} = [-2.24\ -19.96\ -1500.66\ -1072.67]$. The value of Q_{11}, Q_{22}, Q_{33} and Q_{44} has effect on the state variable x, $\dot{x}$, φ, $\dot{\varphi}$, respectively. Control time can be decreased by increasing these values, however control signal is increased. Fig. 10 illustrates the effect of Q_{11}.

All the other supplementary blocks are designed in the same way presented is Section 3.

5. Implementation

Controller design and analysis have been realized firstly in Scilab [9]. Then the virtual reality based implementation has been performed in the frame of MaxWhere as a freely available virtual laboratory [10, 11]. A separate future paper is planned

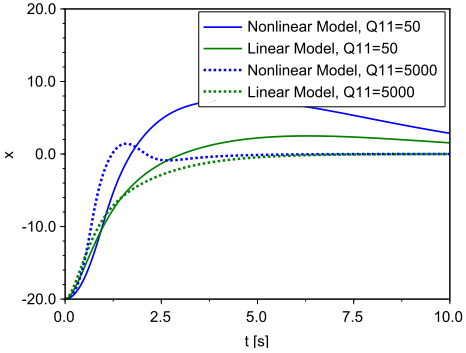


Figure 10. Effect of Q_{11} on the behavior of horizontal movement.

to show the virtual lab. A snapshot about the lab can be seen in Fig. 11, where the controller settings and oscilloscopes showing some signals (e.g. the angle of pendulum, the horizontal position, the acting force versus the time) can be seen among other information like the theoretical background.

6. Conclusion and future work

State feedback controller design for the problem of inverted pendulum has been shown in detail in this paper. Next, nonlinear techniques [12–14] and model predictive control [15] are planned to study.

Acknowledgement

Special thank goes to Tamás Budai to include the mentioned techniques in MaxWhere.

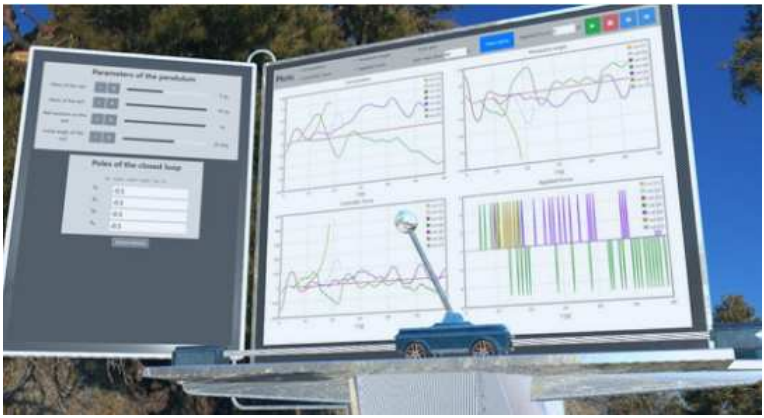


Figure 11. The inverted pendulum model in the virtual laboratory of MaxWhere.

References

[1] M. Kuczmann, Comprehensive Survey of PID Controller Design for the Inverted Pendulum, Acta Technica Jaurinensis 12 (1) (2019) pp. 55–81. doi:10.14513/actatechjaur.v12.n1.492.

[2] L. Keviczky, R. Bars, J. Hetthéssy, C. Bányász, Control Engineering, Springer, 2019.

[3] B. Lantos, Control Systems, Theory and Design I., Academic Press, Budapest, Hungary, 2009, in Hungarian.

[4] K. Ogata, Modern Control Engineering, Prentice-Hall, 1997.

[5] K. J. Aström, R. M. Murray, Feedback Systems: An introduction for scientists and engineers, Princeton University Press, 2008.

[6] P. Gáspár, J. Bokor, A. Soumelidis, An inverted pendulum tool for teaching linear optimal and model based control, Periodica Polytechnica Transportation

Engineering 25 (1-2) (1997) pp. 9–19.
https://pp.bme.hu/tr/article/view/6592.

[7] V. A. Arya, E. G. Ashni, Stabilisation of cart inverted pendulum using the combination of PD and PID control, International Journal of Innovative Research in Science, Engineering and Technology 7 (4) (2018) pp. 3559–3565.

[8] B. Messner, D. Tilbury, Control tutorials for Matlab Simulink [cited 2019-02-23].
http://ctms.engin.umich.edu/CTMS/

[9] Scilab [cited 2019-02-23].
https://www.scilab.org/

[10] Mistems Ltd., MaxWhere [cited 2019-02-23].
https://www.maxwhere.com/

[11] T. Budai, M. Kuczmann, Development of a VR capable virtual laboratory framework, Pollack Periodica 13 (3) (2018) pp. 83–93.
doi:10.1556/606.2018.13.3.9.

[12] B. Lantos, M. Lőrinc, Nonlinear control of vehicles and robots, Springer, 2011.

[13] A. D. Drexler, Nonlinear and robust control, Budapest University of Technology and Economics, 2015, typescript in Hungarian.

[14] K. György, L. Dávid, A. Kelemen, Theoretical study of the nonlinear control algorithms with continuous and discrete-time state dependent riccati equation, Procedia Technology 22 (2016) pp. 582–591.
doi:10.1016/j.protcy.2016.01.123.

[15] W. Liuping, Model predictive control system design and implementation using Matlab, Springer, 2009.

Acta
Technica
Jaurinensis

Vol. 12, No. 2, pp. 148-160, 2019
DOI: 10.14513/actatechjaur.v12.n2.500
Available online at acta.sze.hu

Analysation of the Effect of the Contact Properties on Soil's Penetration Resistance in Discrete Element Simulations

K. Kotrocz[1], Gy. Kerényi[1]

[1]Budapest University of Technology and Economics, Department of Machine and Product Design
Műegyetem rkp. 3-9., 1111 Budapest, Hungary
e-mail: kotrocz.krisztian@gt3.bme.hu

Abstract: Cone penetration in-situ tests are commonly used to measure the cone index of the soils. This measurement process can be modelled very well with Discrete Element Method (DEM) if the parameters of the correct contact model are defined properly. In this paper the Hertz-Mindlin with bonding contact model are used and the effect of the properties of this contact model on soil's penetration resistance is investigated. Our aim was to determine those contact parameters which play important role in the penetration process, thus are necessary to take into account while calibrating the discrete element soil model to the results of real penetration tests.

Keywords: DEM; soil; penetration resistance; cone penetrometer

1. Introduction

Soil compaction is one of the most known problem in agriculture, and it has negative effect on crop growth and yield. In the last few decades, the size of the agricultural machines has been increased which resulted in increasing the mass of the machines, thus increasing the vertical load applied into the soil during tillage operations as well. This normal stress is generated in the soil by the driven and non-driven wheels, and is responsible for soil compaction.

One of the most common method to measure soil strength is by cone penetrometers [1]. During the measurement, the penetrometer cone is pressed into the soil up to given depth, while the vertical force acting on the tip of the cone is

measured. By dividing this force with the projected area of the cone, the soil's resistance to penetration can be determined [2].

Thanks to the development of the information technology in the last few decades, researchers can simulate mechanical processes using numerical methods. One of these is the Finite Element Method (FEM) which can be used to model continuous materials by approximating the geometry with the finite element mesh [3], and calculating the displacement of the nodes. This method is very useful in multi mass simulations (e. g. in modelling the vibration of a one-cylinder engine crankshaft) [4] or performing static analysis on continuous materials as well [5]. It is also used for simulation of the penetration process [6], but because of soil consists of individual particles, these calculations can be used only to determine the location of the hard layers in the soils [7] and are not able for proper simulation of the soil's deformation. More suitable method seems to be the Discrete Element Method (DEM) which is established by Cundall and Strack [8] and is used to model bulk materials such as soils. Tanaka et al created a two-dimensional (2D) discrete element model for modelling of the penetration in-situ test [9], however Butlanska et al investigated this phenomenon in three dimension (3D) [10]. They concluded that the results highly depend on whether full, half or quarter circle soil geometries are used, and the use of spherical elements with non-cohesive or non-bonded contact models results in large error [11].

In our work, we used spherical elements with the Hertz-Mindlin with bonding contact model to simulate the penetration process in cohesive soils. Our aim was to investigate the effect of the contact properties on simulation results, namely on the penetration resistance-penetration depth curve. The results of the paper can be used further when calibrating the contact properties to the results of in-situ penetration tests.

2. Materials and Methods

2.1 Discrete element contact models

DEM is based on dividing the process into small timesteps of Δt, and calculating the displacement vectors of the individual elements according to Newton's 2nd law in each timestep. In most cases, the particles are non-deformable during the simulations, thus the importance of the contact models between the elements are very high. In our calculations, we used the Hertz-Mindlin with bonding model available in EDEM 2.7 software and shown in Fig. 1, which consists two separate model, the Hertz-Mindlin model (Fig. 1/a) and the Parallel Bond contact model (Fig. 1/b).

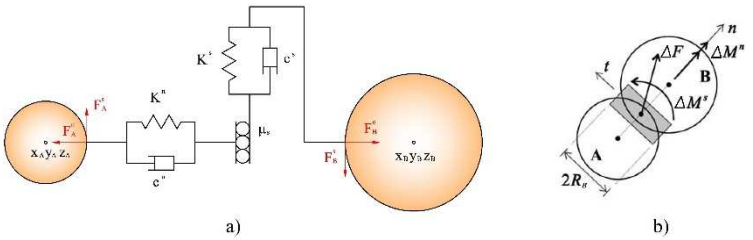


Figure 1. The Hertz-Mindlin (a) and the Parallel Bond contact models (b) [15].

With the first, the friction between the elements can be simulated, and with Parallel Bonds, the cohesion between soil particles can be taken into account. In the Hertz-Mindlin contact model [12, 13, 14], the contact force transferred from one particle to another, is divided into normal (F^n) and shear force (F^s). The normal component can be determined using Eq. 1:

$$F^n = \frac{4}{3} \cdot E^* \cdot \sqrt{R^*} \cdot (U^n)^{\frac{3}{2}}, \tag{1}$$

The E^* represents the equivalent elastic modulus, R^* is the equivalent radius and U^n is the normal overlap of the contacting elements. These can be determined from the parameters, E, R and the location of the particles. The contact shear force (F^s) can be calculated from the equivalent shear modulus (G^*) and tangential displacement (U^s) according to Eq. 2:

$$F^s = -\left(8 \cdot G^* \cdot \sqrt{R^* \cdot U^n}\right) \cdot U^s, \tag{2}$$

In addition, it has a limit according to Coulomb's law of friction, namely it cannot be higher than the value from Eq. 3:

$$F^s \leq F^n \cdot \mu_s. \tag{3}$$

In Eq. 3, the μ_s denotes to the friction coefficient between the soil particles. In addition, there are damping forces to model the energy dissipation of the particle's collisions, they can be divided into normal ($F^n{}_d$) and shear components ($F^s{}_d$) as well and can be calculated as:

$$F^n{}_d = -2 \cdot \sqrt{\frac{5}{6}} \cdot \beta \cdot \sqrt{K^n \cdot m^*} \cdot v^n{}_{rel}, \tag{4}$$

$$F^s{}_d = -2 \cdot \sqrt{\frac{5}{6}} \cdot \beta \cdot \sqrt{K^s \cdot m^*} \cdot v^s{}_{rel}. \tag{5}$$

In Eq. 4 and Eq. 5, K^n, K^s, m^*, v^n_{rel} and v^s_{rel} are the normal and shear contact stiffness (Fig. 1/a), the equivalent mass and the normal and shear component of the relative velocity (v_{rel}) of the contacting elements, respectively. In addition, β can be calculated with Eq. 6 using the coefficient of restitution (e).

$$\beta = \frac{\ln e}{\sqrt{\ln^2 e + \pi^2}} \,. \tag{6}$$

Parallel Bonds can be added to the model at the time of t_{Bond} to represents the cementinous effect between the elements. This contact model can be envisioned as a set of elastic springs around the contact point with bond radius of R_B [15]. The bonds act parallel with the Hertz-Mindlin contact model, therefore additional normal (ΔF^n) and shear forces (ΔF^s) are summed to the corresponding components:

$$\Delta F^n = -k^n{}_B \cdot A_B \cdot \Delta U^n \,, \tag{7}$$

$$\Delta F^s = -k^s{}_B \cdot A_B \cdot \Delta U^s \,. \tag{8}$$

In Eq. 7 and Eq. 8, $k^n{}_B$, $k^s{}_B$, A_B and ΔU^n and ΔU^s are the bond's normal and tangential stiffness, the area of the bond and the normal and shear component of the relative displacement of the contacting elements, respectively. These displacements are incremental, thus are calculated from the time of t_{Bond}, when both of them are set to zero. Because Parallel Bond behaves as beam contact between the particles, it can transmit moments through the elements in both normal and tangential direction as well. These moments can be calculated similar to the bond forces, but the area of the bond should be changed to the polar moment of inertia (J_B) and the relative displacements to the relative rotations of the elements ($\Delta\theta^n$ and $\Delta\theta^s$), as it is shown in Eq. 9 and Eq. 10.

$$\Delta M^n = -k^n{}_B \cdot J_B \cdot \Delta\theta^n \,, \tag{9}$$

$$\Delta M^s = -k^s{}_B \cdot \frac{J_B}{2} \cdot \Delta\theta^s \,. \tag{10}$$

Similar to the relative displacements of ΔU^n and ΔU^s, the relative rotations are set to zero at the bond's formation time of t_{Bond} as well. In addition, there are limit stresses (i. e. the normal and tangential Parallel Bond strengths). If the stress in the bond obtains the strength value in normal or tangential direction, the bond will break in the next timestep, and the elements will move only according to the Hertz-Mindlin contact model. The maximum normal (σ_{max}) and tangential bond stresses (τ_{max}) can be determined using Eq. 11 and Eq. 12:

$$\sigma_{max} = \frac{-\Delta F^n}{A} + \frac{2 \cdot \Delta M^s}{J_B} \cdot R_B \,, \tag{11}$$

$$\tau_{max} = \frac{-\Delta F^s}{A} + \frac{\Delta M^n}{J_B} \cdot R_B \,. \tag{12}$$

2.2 Discrete element simulations of cone penetration test.

To analyse the effect of the contact properties on simulation results, a lot of numerical cone penetration simulations were performed. First, the initial geometry of the soil was created by filling up a cylinder with diameter of Ø135,4 mm and height of 130 mm with spherical elements (Fig. 2/a). The size of the model was chosen according to our earlier research [16] where it is proved, that the boundary of the model is far enough from the cone penetrometer, thus it has negligible effect on simulation results. The particles were created using the Simple Sequential Inhibition (SSI) technique [17], which places elements with random diameter to random locations. The contact properties of the Hertz-Mindlin model for soil particles are summarized in Table 1. After the creation of the elements, the particles fell down to the bottom of the cylinder because of Earth gravity. Finally, when the whole system obtained the equilibrium state (the element's maximum velocity got smaller than 1e-2 mm/s), the Parallel Bonds were formed between the particles, then the geometry of the penetrometer cone (shown in Fig. 2/b) was imported into the program using an .stl file. The properties of the soil's Parallel Bond contact model and of the wall elements (as penetrometer cone) can be also seen in Table 1. Note, that some contact parameters of the soil were determined according to the process which is published in our earlier research [18], the remaining properties were chosen for sensitivity tests. These are presented in Table 1 as grey background, and the ranges where the effect of these contact properties were investigated, were also shown in the table. The properties of the penetrometer cone were chosen as general steel material. Additionally, the initial geometry of the system can be seen in Fig. 2/a.

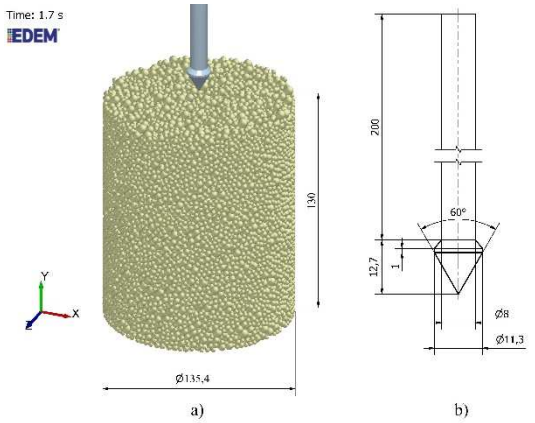


Figure 2. The initial geometry of the soil (a) and the dimension of the used penetration bar (b).

In the next step, the geometry of the penetrometer cone was pressed into the soil model with vertical speed of 20 mm/s, which is typical in cone penetration in-situ tests. During the simulations, the force acting on the penetrometer cone (as soil resistance) was calculated and was saved in each $0{,}05^{th}$ calculation time. This process was repeated in all discrete element simulations where the timestep was set to the value of 5e-6 s.

Table 1. The settings of the discrete element cone penetration tests

Parameter	***Value***
Geometrical properties	
Particle radius distribution (mm)	1,33…3
Contact radius (mm)	1,6…3,6
Initial porosity (before the penetration bar pressed into the soil) (-)	0,425
Properties of the Hertz-Mindlin with bonding contact model between the soil elements	
Density (kg/m^3)	1,6e+03…2,0 e+03
Shear modulus (Pa)	1,44e+06…1,44e+08
Poisson ratio (-)	0,2…0,3
Coefficient of restitution between the soil particles (-)	0,5
Coefficient of restitution between the soil particles and walls (-)	0,5
Friction coefficient between ball and ball (-)	0,4
Friction coefficient between ball and walls (-)	0,6
Bond radius (mm)	0,5…5,0
Bond normal stiffness (Pa/m)	9,5e+06…4,75e+07
Bond shear stiffness (Pa/m)	
Bond normal strength (Pa)	3,131e+4
Bond shear strength (Pa)	4,428e+4
Properties of the Hertz-Mindlin contact model between the penetrometer (wall) elements	
Density (kg/m^3)	7,8e+03
Shear modulus (Pa)	7,692e+10
Poisson ratio (-)	0,3
Coefficient of restitution between walls (-)	0,5
Friction coefficient between wall and wall (-)	0,1

3. Results and Discussion

In this section, the results of the sensitivity test of numerical discrete element penetration simulations are shown. First all simulations were evaluated qualitatively which means that we analysed the broken Parallel Bonds, the compression force and the velocity of the particles before evaluating the vertical forces that acts on the penetration cone. Fig. 3 shows typical results of the simulations, the broken bonds are presented as blue lines in Fig. 3/a. It can be asserted, that they break near the penetration cone which can be acceptable. In the part *b)* and *c)* of the figure, the compression force of the particles and their velocities are shown, respectively. It can be clearly seen that the highest forces arise in front of the cone. Similar to that can be said in case of the element's velocity, however high velocities can be experienced near the penetration bar as well. This is because of the friction between the penetrometer and the soil particles, and is similar to the results in Tanaka et al [9]. According to these, we concluded that the simulation results are in accordance with our expectations and the results of other researches, thus the calculations provided good results qualitatively.

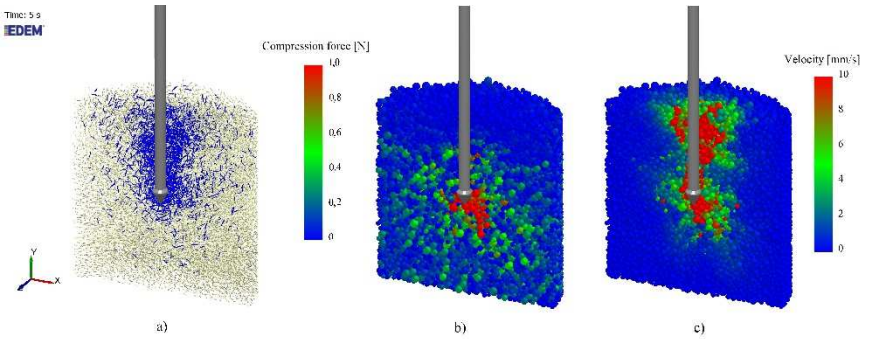


Figure 3. The broken bonds (a), the compression force between the elements (b) and the particle's velocity (c) at penetration depth of 64 mm.

In the next step, we investigated the results quantitatively as well, the penetration resistance-penetration depth curves were compared in case of different contact properties values. First, the effect of the particle's density and Poisson-ratio on soil resistance to penetration were analysed. Fig. 4 shows the results in case of different density values. It can be asserted, that this contact property does not have significant effect on penetration resistance, the curves are almost the same. Similar to that can be concluded in case of different Poisson-ratio (Fig. 5), because there are small differences in the penetration resistance-depth curves.

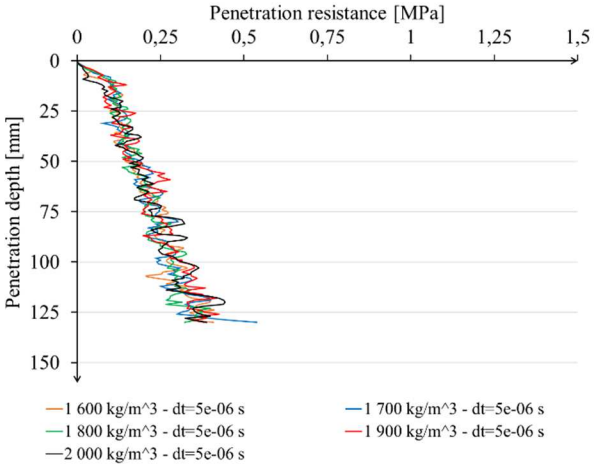


Figure 4. The effect of the particle's density on penetration resistance.

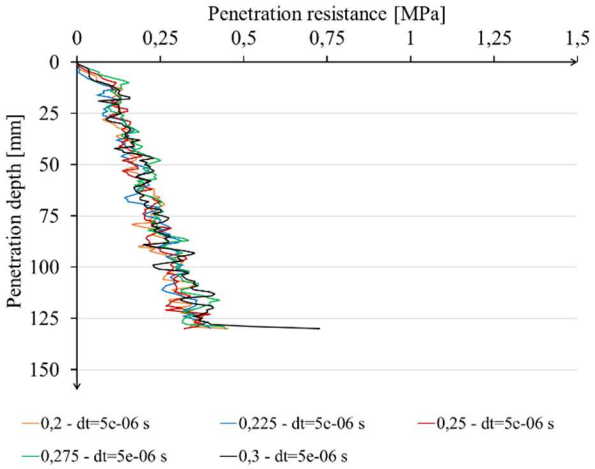


Figure 5. The effect of the particle's Poisson-ratio on penetration resistance.

However, this is not the case in the simulations with different particle's shear modulus. Fig. 6 shows that the penetration resistance is increasing when this contact property is increased as well, but this tendency is stopped in case of shear modulus of 4,32e+07 Pa. If higher than this value is set up in the simulations, this parameter has negligible effect on soil's resistance to penetration.

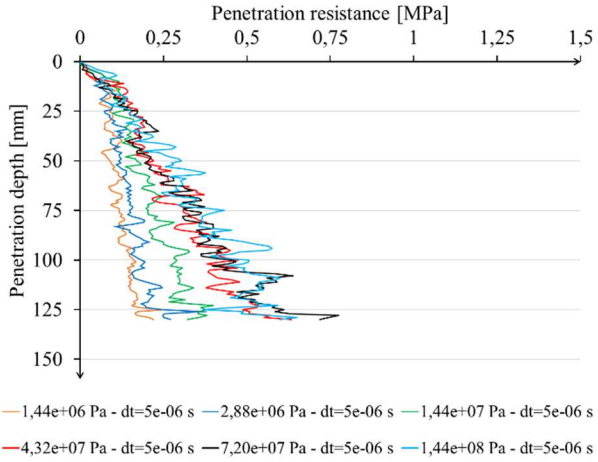


Figure 6. The effect of the particle's shear modulus on penetration resistance.

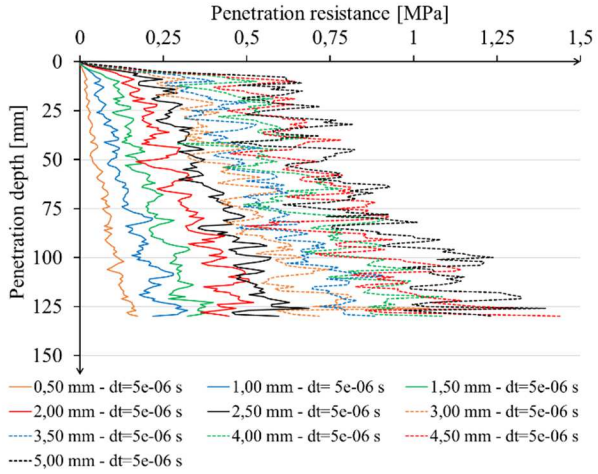


Figure 7. The effect of the Parallel Bond radius on penetration resistance.

Parallel Bond radius also has large effect on simulation results. According to Fig. 7, the higher the value of this parameter is, the force acting on the penetration cone, thus the penetration resistance of the soil model will be higher as well. This conclusion stands in the whole range (namely the value from 0,5 mm to 5,0 mm), where this contact property was investigated.

In Fig. 8, the results can be seen in case of different Parallel Bond stiffness. It can be asserted, that the effect of this parameter is very similar to the experienced one in case of particle's shear modulus. Namely, if the bond stiffness is increasing from 9,5e+06 Pa/m to 2,85e+07 Pa/m, the model's resistance to penetration will increase as well. But in case of the highest Parallel Bond stiffness values, the penetration resistance-depth curves are similar to each other, thus this parameter has no significant effect on simulation results in this range of value.

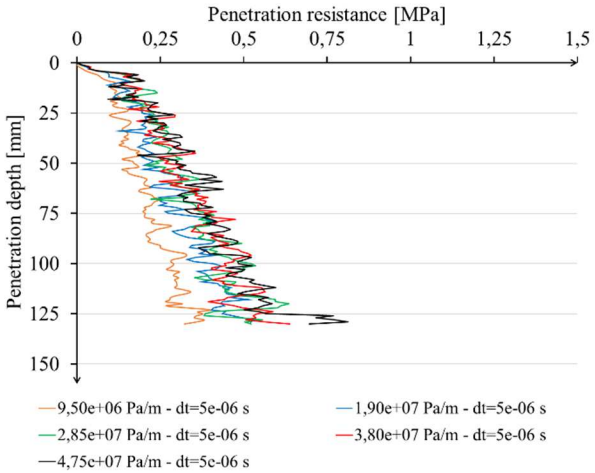


Figure 8. The effect of the Parallel Bond stiffness on penetration resistance.

4. Conclusion

In this paper, the discrete element method was adopted to model the cone penetration in-situ measurements. We used the Hertz-Mindlin with bonding contact model to simulate cohesive soil and our aim was to investigate the effect of the contact properties on simulation results, namely on penetration resistance-penetration depth curve. We concluded that the Parallel Bond radius, stiffness and the elements' shear modulus affect the results, thus they play important role in calibration of the contact properties to the result of in-situ penetration test. On the other hand, the particle's density and Poisson-ratio have negligible effect on the penetration resistance-depth curve in the investigated range, therefore it is not necessary to take these into account in the calibration process.

Acknowledgement

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).

References

[1] L. Laib (Ed.), Terepen mozgó járművek (Moving off-road vehicles). Szaktudás Kiadó Ház, Budapest, 2002, in Hungarian.

[2] E. McKyes, Soil Cutting and Tillage, The University of Michigan, Michigan, 1985.

[3] M. Kuczmann, Overview of the Finite Element Method. Acta Technica Jaurinensis 8 (2015) pp. 347-383.
doi: https://doi.org/10.14513/actatechjaur.v8.n4.393

[4] P. Horváth, J. Égert, Dynamic Analysis of a One-cylinder Engine Crankshaft. Acta Technica Jaurinensis 8 (2015) pp. 280-295.
doi: https://doi.org/10.14513/actatechjaur.v8.n4.379

[5] P. T. Zwierczyk, K. Váradi, Frictional contact FE analysis in a railway wheel-rail contact. Periodica Polytechnica Mechanical Engineering 58 (2014), pp. 93-99.l
doi: https://doi.org/10.3311/PPme.7229

[6] W. A. Foster Jr., C. E. Johnson, R. C. Chiroux, T. R. Way, Finite element simulation of cone penetration. Applied Mathematics and Computation 162 (2005), pp. 735-749.
doi: https://doi.org/10.1016/j.amc.2004.01.012

[7] M. Z. Tekeste, R. L. Raper, E. W. Tollner, T. R. Way, Finite element analysis of cone penetration in soil for prediction of hardpan location. Transactions of the ASABE 50 (2007), pp. 23-31.

[8] P. A. Cundall, O. D. L. Strack, A discrete numerical model for granular assemblies. Géotechnique 29 (1979), pp. 47-65.
doi: https://doi.org/10.1680/geot.1979.29.1.47

[9] H. Tanaka, M. Momozu, A. Oida, M. Yamazaki, Simulation of soil deformation and resistance at bar penetration by the Distinct Element Method. Journal of Terramechanics 37 (2000), pp. 41-56.
doi: https://doi.org/10.1016/S0022-4898(99)00013-0

[10] J. Butlanska, M. Arroyo, A. Gens, Homogeneity and Symmetry in DEM Models of Cone Penetration. AIP Conference Proceedings 1145 (2009), pp. 425-428.
doi: https://doi.org/10.1063/1.3179952

[11] J. Butlanska, M. Arroyo, A. Gens, C. O'Sullivan, Multi-scale analysis of cone penetration test (CPT) in a virtual calibration chamber. Canadian Geotechnical Journal 51 (2013), pp. 51-66.
doi: https://doi.org/10.1139/cgj-2012-0476

[12] H. Hertz, On the contact of elastic solids. J. reine und angewandte Mathematik 92 (1882), pp. 156-171.

[13] R. D. Mindlin, Compliance of elastic bodies in contact. Journal of Applied Mechanics 16 (1949), pp. 259-268.

[14] R. D. Mindlin, H. Deresiewicz, Elastic spheres in contact under varying oblique forces. ASME (1953) pp. 327-344.

[15] D. O. Potyondy, P. A. Cundall, A bonded-particle model for rock. International Journal of Rock Mechanics and Mining Sciences, Rock Mechanics Results from the Underground Research Laboratory, Canada 41 (2004), pp. 1329-1364.
doi: https://doi.org/10.1016/j.ijrmms.2004.09.011

[16] K. Kotrocz, A. M. Mouazen, Gy. Kerényi,. Numerical simulation of soil cone penetrometer interaction using discrete element method. Computers and Electronics in Agriculture 125 (2016):(C) pp. 63-73.
doi: https://doi.org/10.1016/j.compag.2016.04.023

[17] J. W. Evans, Random and cooperative sequential adsorption. Reviews of Modern Physics 65 (1993), pp. 1281-1329.
doi: https://doi.org/10.1103/RevModPhys.65.1281

[18] K. Kotrocz, Gy. Kerényi, Numerical Discrete Element Simulation Of Soil Direct Shear Test, 31st Conference on Modelling and Simulation, ECMS 2017, Budapest, Hungary, pp. 510-515. Edited by: Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics.
doi: https://doi.org/10.7148/2017-0510

Acta
Technica
Jaurinensis

Vol. 12, No. 2, pp. 161-177, 2019
DOI: 10.14513/actatechjaur.v12.n2.501
Available online at acta.sze.hu

Comparative Analysis of Axlebox Accelerations in Correlation with Track Geometry Irregularities

Cs. Ágh[1]

[1]**Széchenyi István University, Department of Transport Infrastructure and Water Resources Engineering**
Egyetem tér 1, 9026 Győr, Hungary
e-mail: agh.csaba@sze.hu

Abstract: Stochastic track irregularities influence additional dynamic forces developed in the vehicle-track interaction that lead to faster deterioration of track. For economical track maintenance it is important to understand the relationship between the irregularities recorded by conventional track geometry measuring car and the resulting dynamic vehicle responses. This paper focuses on the correlation between lateral and vertical axlebox accelerations and differently processed track geometry parameters based on a real measurement run on straight track. Decolouring of chord-offset measurement results was performed and derivatives of track geometry parameters were calculated for comparison. Those track geometry parameters have been selected which provide the most accurate information about the recorded wheelset accelerations caused by track geometry irregularities, eg. second order derivative of cross level.

Keywords: track geometry; axlebox acceleration; vehicle response; decolouring

1. Introduction

The deterioration of the railway track is primarily caused by static and dynamic forces transferred from the vehicles. The railway track alignment always deviates from nominal geometry; it contains vertical and lateral geometric irregularities. The railway wheelset passing through the geometric defects is forced to 'follow' these irregularities, because of that the railway vehicle's wheelsets, bogies and car body have three dimensional trajectory movement. Therefore, vertical and lateral accelerations of wheelsets, bogies and car body can be measured at any position. If

wheelset, bogies and car body are considered as rigid bodies coupled with springs and dampers, the forces between the track and the vehicle can be determined from the accelerations on rigid bodies [1]. These dynamic forces accelerate the deterioration of track geometry and the presence of rolling contact fatigue type rail defects , e.g. track twists, squats, rail cracks, etc. [2].

The forces required for irregular movements of certain rigid bodies (wheelsets, bogies, car body) are directly proportional to body mass and acceleration according to Newton's laws. Although the mass of the car body is large, its accelerations are usually small. The wheelsets and bogies suffer significant acceleration, the forces required for their displacement ultimately dominate the dynamic forces between the railway vehicle and the track. [3] In this article the relationship between the acceleration of the wheelset and the track irregularities recorded by track geometry measurement is analysed. Vertical irregularities are often caused by an inhomogeneous longitudinal subgrade stiffness/damping distribution. [4]

This paper attempts to determine the relationship between the simultaneously recorded conventional track geometry parameters (longitudinal level, alignment, cross level) and vertical/lateral accelerations of the railway wheelset on straight track segments, i.e., how they are related to forces exerted by irregular vehicle movements that cause further geometric deterioration of the track (Figure 1).

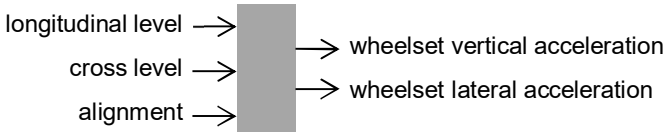


Figure 1. System view of track geometry parameters and wheelset dynamic responses

Nowadays, more and more in-service passenger trains are equipped with inertial sensors providing an useful estimation of the current track quality on daily basis. Therefore, it is important to understand the relationship between 'vehicle response' and track geometry deviation from the nominal alignment. [5] [6] [7]. In the past period, track geometry assessment systems of infrastructure managers adopted different theoretical and empirical approaches, and a demand emerged to create an intervention limit system based on dynamical vehicle responses. [8] It is also important to understand the multiplicity of vehicle responses by testing and simulation for the acceptance of running characteristics of railway vehicles. [9]

Based on track geometric irregularities, complex multi-body software is used to estimate acceleration on the vehicle, and the literature deals with the development of so-called performance-based track geometry [10]. The vehicle-track system contains many nonlinearities, and neural networks are particularly suitable for analyzing them. [11]. The recent Dynotrain project also paid particular attention to vehicle reactions to track geometric defects. [12]

2. Methodology

The measurement data required for the experiments now presented were recorded by track geometry measuring system and the vehicle dynamic measurement system operating simultaneously on the measuring car FMK-007, which is originally an intercity wagon and it was adapted for track diagnostic purposes by MÁV Central Rail and Track Inspection Ltd.

The longitudinal level and alignment results used in the experiments now presented are derived from a chord offset measurement system. Therefore, from these measurement results for the correct calculation, the distortion of the chord system has to be compensated by so-called decolouring method. Although European standards [13] require the use of a band-pass filter for a given wavelength range for the assessment of track maintenance measurements, in reality the dynamic response of the vehicle depends on all wavelength components in the track, in this way tests are also performed on decoloured but unfiltered data as well.

2.1. Track geometry measurement data

Track geometry measuring system of the FMK-007 consists of 3 laser units per rail, complemented by an inertial unit between the central laser units. Measurement of longitudinal level and alignment bases on all laser units, as detailed below. Cross level is calculated according to inclination measurement of inertial unit mounted on car body and neighbouring laser sensors compensate for the motion of the car body relative to the rails.

The measuring system works with $\Delta x = 0.25$ m step equidistant sampling. To avoid numerical errors, the lengths of the chord parts for the longitudinal level and alignment resulting from the positioning of the laser units were rounded to 0.25 m. The complex transfer function of the chord system $H(\lambda)$ can be calculated with Eq. (1) known from the literature [14]:

$$H(\lambda) = 1 - \frac{a}{L} e^{\frac{2\pi b}{\lambda} i} - \frac{b}{L} e^{-\frac{2\pi a}{\lambda} i}, \qquad (1)$$

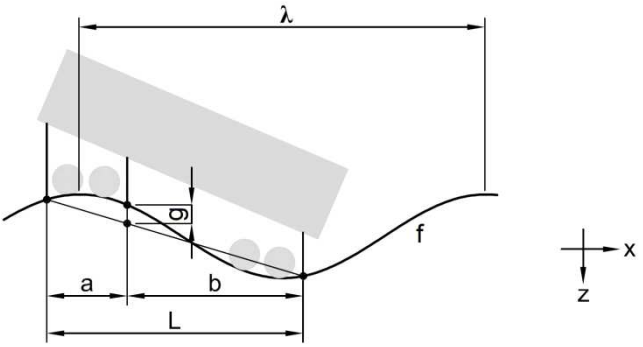


Figure 2. Asymmetric chord offset measurement system of track recording car FMK-007 (drawn by the author)

where the value of *a, b* and *L* (Figure 2) depending on the asymmetric chord are 4.00 m, 19.00 m, 23.00 m, respectively. The parameter λ is the wavelength and *i* is the imaginary unit. In the opposite direction of measurement, *a* and *b* are reversed. The amplitude characteristics (2) and phase characteristics of the system (3) is the same as the magnitude (Figure 3) and phase (Figure 4) of complex function *H*, respectively:

$$|H(\lambda)| = \sqrt{\left[1 - \frac{a}{L}\cos\left(\frac{2\pi}{\lambda}b\right) - \frac{b}{L}\cos\left(\frac{2\pi}{\lambda}a\right)\right]^2 + \left[\frac{b}{L}\sin\left(\frac{2\pi}{\lambda}a\right) - \frac{a}{L}\sin\left(\frac{2\pi}{\lambda}b\right)\right]^2}, \tag{2}$$

$$\angle H(\lambda) = \arctan\frac{\frac{b}{L}\sin\left(\frac{2\pi}{\lambda}a\right) - \frac{a}{L}\sin\left(\frac{2\pi}{\lambda}b\right)}{1 - \frac{a}{L}\cos\left(\frac{2\pi}{\lambda}b\right) - \frac{b}{L}\cos\left(\frac{2\pi}{\lambda}a\right)}, \tag{3}$$

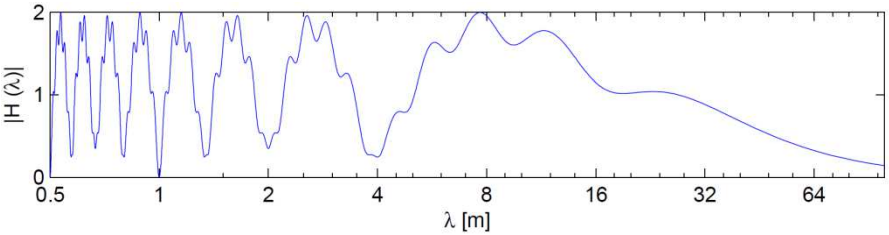


Figure 3. Magnitude of transfer function of the asymmetric chord measurement system (4+19 m)

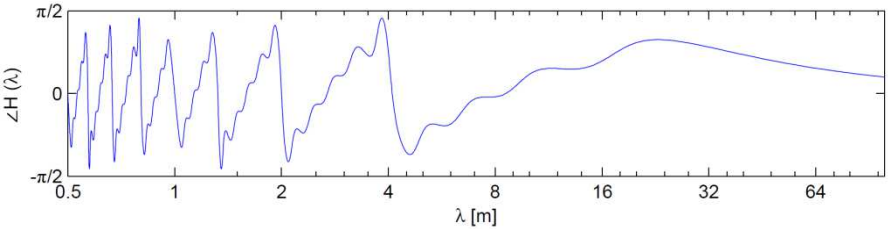


Figure 4. Phase of the transfer function of the asymmetric chord measurement system (4+19m)

The calculations were carried out with data series of straight track section of 1 km, which contained a total of N = 4001 values per measurement parameter. Raw track geometry measurement parameters used:

- cross level (*CL*),
- longitudinal level of left rail ($LL_{l,chord}$), longitudinal level of right rail ($LL_{r,chord}$),
- alignment of left rail ($AL_{l,chord}$), alignment of right rail ($AL_{r,chord}$).

The data of cross level parameter could be used directly, but for the longitudinal level and alignment parameters had to be decoloured in order to remove the distortion of the chord offset measurement. To do this, the Fourier transform of the sequence must be multiplied by the inverse of the transfer function:

$$F(\lambda) = G(\lambda)\, H^{-1}(\lambda)\,, \tag{4}$$

where F is the Fourier transform of the real track shape labelled with f, and G is the Fourier transform of the measured track geometry data g (which may be $LL_{l,chord}$, $LL_{r,chord}$, $AL_{l,chord}$, $AL_{r,chord}$):

$$G_k = \sum_{n=0}^{N-1} g_n \; e^{-2\pi k \frac{n}{N} i}\,, \tag{5}$$

where $k = 0,1,\ldots N–1$ and

$$\lambda = \frac{N\,\Delta x}{k}\,, \tag{6}$$

where λ represents considered wavelength. (If $k = 0$, then the general shift of g is in question.)

Reciprocating by definition

$$H^{-1} = \frac{\mathrm{Re}(H) - i\,\mathrm{Im}(H)}{|H|^2}, \tag{7}$$

at low values of $|H|$ makes the calculation numerically unstable and it cannot be applicable if zero value is used. Managing this problem requires great care, as H needs to be modified and this also affects the result. According to *Insa* [15] complex numbers less than a given value should be replaced by 1. However, in this study, in case of $|H| < c$ the Eq. (7) was modified as follows:

$$H_{|H|<c}^{-1} = \frac{\mathrm{Re}(H) - i\,\mathrm{Im}(H)}{c^2}, \tag{8}$$

where c is 0.2, determined after multiple trials. The decoloured parameter f (which may be LL_r, LL_l, AL_r, AL_l), was calculated using (4) and after that an inverse-Fourier transform was performed on function $F(\lambda)$:

$$f_n = \frac{1}{N}\sum_{k=0}^{N-1} F_k \; e^{2\pi k \frac{n}{N} i}, \tag{9}$$

Of the converted 1 km track section, only the middle 800 m could be used well because of numerical issues.

On function f, the band-pass filtering for D1 wavelength range (3 m < λ < 25 m) according to EN 13848 [13] resulted in sequences $LL_{r,D1}$, $LL_{l,D1}$, $AL_{r,D1}$, $AL_{l,D1}$.

2.2. Acceleration measurement data

In this paper only some accelerometer sensors of the vehicle dynamic measurement system of the car FMK-007 were used. The accelerometer sensors located at the axlebox of leading wheelset considering the measurement direction was taken into account. Positioning of sensors are (Figure 5):

- there are one-axle vertical accelerometer sensors above the axleboxes on both sides of the leading wheelset, the labelling of which are $\ddot{z}_l$ and $\ddot{z}_r$,
- there is a one-axle lateral accelerometer sensor above the axlebox on left side of the wheelset, the labelling of which is $\ddot{y}$.

The accelerometer sensors are located at a height of h = 0.65 m above the rail top level on a vertical console attached to the end of the axlebox (Figure 6), the distance between the vertical accelerometers is q = 2.5 m, the average distance of the wheel-rail contact patches is t = 1.5 m. During the study, it was focused on the relevant wavelengths of track geometry, which starts, according to Salvador [16] at around 2 metres. There was no goal to investigate rail surface defects (e.g. squat). Therefore, the acceleration signals were passed through a 16 Hz second-order Butterworth low-pass filter and recorded at 300 Hz sampling rate (gain: ± 100g).

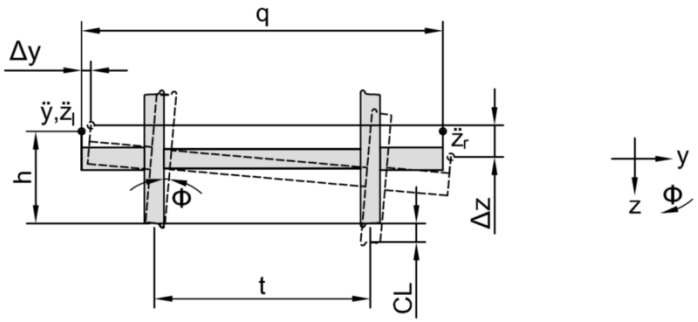


Figure 5. Accelerometer sensors $\ddot{y}$, $\ddot{z}_l$, $\ddot{z}_r$ mounted above the axleboxes and visualization of rotated wheelset due to cross level defect (drawn by the author)

Figure 6. Accelerometer sensor box mounted above the axlebox (photo taken by the author)

2.3. Connection between lateral and vertical axlebox accelerations

In the case of level irregularity of one of the rails, due to the change in the cross level *CL* the wheelset is rotated around one rail as marked by Φ in Figure 5. Because of small angles:

$$CL \approx \mathrm{t}\,\phi\,, \tag{10}$$

height difference between vertical accelerometers located above the end of the axleboxes:

$$\Delta z \approx \mathrm{q}\,\phi\,, \tag{11}$$

lateral displacement of the accelerometer mounted on the axlebox at height *h*:

$$\Delta y \approx \mathrm{h}\,\phi\,. \tag{12}$$

Therefore, the following relationship can be assumed between the vertical axlebox acceleration and the lateral axlebox acceleration when passing through a track twist:

$$\ddot{y}_{\Delta\ddot{z}} = \frac{\mathrm{h}}{\mathrm{q}}(\ddot{z}_l - \ddot{z}_r)\,. \tag{13}$$

2.4. Axlebox acceleration calculation based on track geometry data

If the wheels 'follow' the rail irregularities perfectly, the trajectory movement of the wheels will be same as the track geometry which is recorded. If a constant measuring speed is assumed, the accelerations of the wheelset can be calculated based on track geometry results by double deriving. [17] Taking into account the actual speed, an estimation was made for the axlebox acceleration. Compared to the acceleration signals, it was found that the Δx = 0.75 m step is the most favourable for numerical derivation. Derivatives are:

$$f'(x) = \frac{f(x+\Delta x)-f(x)}{\Delta x}, \tag{14}$$

$$f''(x) = \frac{f(x+\Delta x)-2f(x)+f(x-\Delta x)}{\Delta x^2}, \tag{15}$$

$$f'''(x) = \frac{f(x+2\Delta x)-3f(x+\Delta x)+3f(x)-f(x-\Delta x)}{\Delta x^3}, \tag{16}$$

where $f^{(n)}(x)$ is the *n*-order derivative of a track geometry parameter at section *x* and $f(x-\Delta x)$, $f(x+\Delta x)$ are the values of *LL, AL, CL,* 0.75 m before and after section *x*, respectively.

The resulting acceleration is proportional to the square of the speed. The expected track acceleration based on track geometry is therefore calculated by following equations. Considering the permanent travel speed *v*, the estimate of left vertical wheel acceleration based on the left longitudinal level is:

$$\ddot{z}_{l,LL,l} = LL_l''(x)\,v^2. \tag{17}$$

Taking into account the permanent travel speed, the estimate of lateral wheelset acceleration based on the left alignment is (assuming that wheelset directly 'follows' lateral track irregularities):

$$\ddot{y}_{AL,l} = AL_l''(x)\,v^2, \tag{18}$$

Taking into account the permanent travel speed, accelerometer positioning and wheelset measures, the estimate of lateral wheelset acceleration based on cross level is:

$$\ddot{y}_{CL} = \frac{\mathrm{h}}{\mathrm{t}} CL''(x)\, v^2. \tag{19}$$

Because of noise, the value of $\ddot{y}_{cl}$ was smoothed by a 0.75 m moving average.

2.5. Calculation of correlation coefficient

When acceleration signals were compared to track geometry measurements, a resampling of 0.25 m step was performed on the acceleration data. Synchronisation of signals was carried out manually. The method of calculating the correlation coefficient was the same as that used by *Karis et al.* [18]:

$$r_{u,v} = \frac{\mathrm{cov}(u,v)}{\sigma_u \sigma_v} = \frac{\Sigma_{j=1}^{N}(u_j-\bar{u})(v_j-\bar{v})}{\sqrt{\Sigma_{j=1}^{N}(u_j-\bar{u})^2\ \Sigma_{j=1}^{N}(v_j-\bar{v})^2}}. \tag{20}$$

3. Results

3.1. Correlation between vertical and lateral axlebox accelerations

Correlations between vertical and lateral axlebox accelerations were investigated on a Hungarian railway line (Budapest–Kelebia), with variable measurement speeds (0–80 km/h), omitting the curved sections, over 45 km length, as well as over 1 km long straight section with constant speed of 79 km/h (Table 1).

The standard deviation of the decoloured and D1 filtered parameters on the examined 1 km long section is given as follows. Standard deviations (in millimetres) of left longitudinal level, right longitudinal level, left alignment, right alignment are 2.71, 2.33, 1.22 and 1.28, respectively.

The correlation coefficient between $\ddot{z}_l - \ddot{z}_r$ and $\ddot{y}$, over the tested 45 km inhomogeneous and variable speed section is 0.75 and over the 1 km homogeneous section 0.79. There is also connection between $\ddot{y}$ and separately recorded vertical axlebox accelerations. The lateral accelerometer is located on the left side and the numbers indicate that the connection with the left vertical accelerometer shows higher correlation.

Table 1. Correlation coefficients between vertical and lateral wheelset accelerations (45 km various track and various measuring speeds)

	45 km track, various speed	*1 km track, constant speed*
	$\ddot{y}$ [m/s^2]	
$\ddot{z}_r$ [m/s^2]	–0.39	–0.57
$\ddot{z}_l$ [m/s^2]	0.46	0.68
$\ddot{z}_l - \ddot{z}_r$ [m/s^2]	**0.75**	**0.79**

Figure 7 shows the measured and the estimated lateral axlebox acceleration which was calculated by Eq. (13) based on $\ddot{z}_l - \ddot{z}_r$.

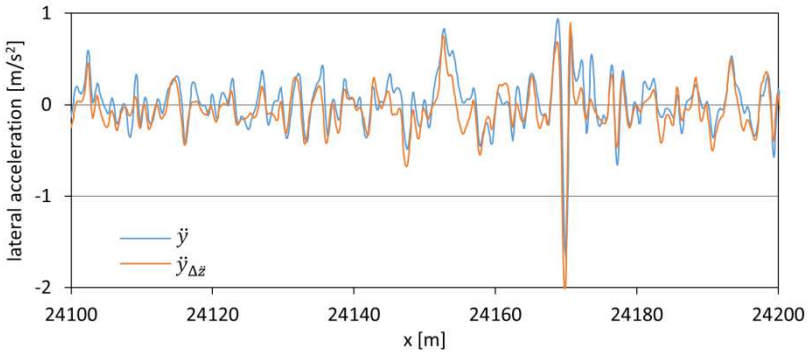


Figure 7. Measured lateral axlebox acceleration and estimated wheelset lateral acceleration based on vertical axlebox acceleration difference

3.2. Correlation between vertical axlebox accelerations and longitudinal level

Correlation coefficients were calculated for the 1 km (800 m) long section detailed above, comparing the left axlebox acceleration and the left longitudinal level (Table 2).

The acceleration measured on the left axlebox correlates with the second order derivative of the de-coloured left longitudinal level (LL_l'') significantly (0.63). The correlation with the second order derivative of the original chord measurement data

is close to this (0.56). The statistical relationship to the D1 filtered data is in low level.

Table 2. Correlation coefficients between vertical axlebox accelerations and longitudinal level derivatives (0.8 km track, 79 km/h speed)

	$\ddot{z}_l$ [m/s^2]
$LL_{l,chord}$ [mm]	0.40
$LL''_{l,chord}$ [mm]	0.56
$LL_{l,D1}$ [mm]	0.21
$LL''_{l,D1}$ [mm]	0.14
LL_l [mm]	0.38
LL'_l [mm]	0.42
LL''_l [mm]	**0.63**
LL'''_l [mm]	0.43

A Figure 8 shows the measured left vertical axlebox acceleration and the estimated left lateral wheel acceleration which was calculated by Eq. (17) based on LL''_l. It should be taken into consideration that the vertical acceleration and the longitudinal level were measured in two separate planes: at the axleboxes and at the rail, respectively.

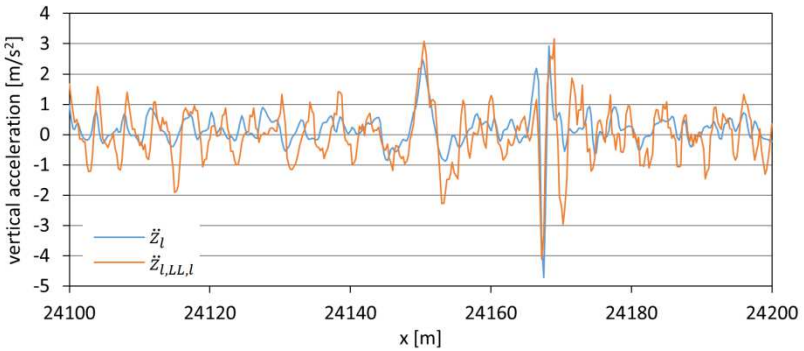


Figure 8. Measured vertical axlebox acceleration and estimated vertical wheel acceleration based on longitundinal level (left)

3.3. Correlation between lateral axlebox acceleration and alignment and cross level

Correlation coefficients were calculated for the 1 km (800 m) long section detailed above, for the cross level and for left alignment (Table 3). In the alignment parameter, neither the original chord nor the decoloured data showed a noticeable statistical relationship with the lateral axlebox acceleration, not even the second derivative. However, the second order derivative of the cross level is closely related to the lateral axlebox acceleration, where the correlation coefficient is 0.77. First order derivative of cross level (CL') was also calculated, which corresponds to traditional 'track twist' on 0.75 m base, but its correlation is weaker.

Table 3. Correlation between vertical axlebox acceleration and alignment and cross level (0.8 km track, 79 km/h speed)

	$\ddot{y}$ [m/s^2]
$AL_{l,chord}$ [mm]	0.07
$AL_{r,chord}$ [mm]	0.12
$AL''_{l,chord}$ [mm]	–0.02
$AL''_{r,chord}$ [mm]	0.05
$AL_{l,D1}$ [mm]	0.02
$AL_{r,D1}$ [mm]	0.05
$AL''_{l,D1}$ [mm]	–0.07
$AL''_{r,D1}$ [mm]	0.02
AL_l [mm]	0.07
AL_r [mm]	0.09
AL''_l [mm]	0.02
AL''_r [mm]	0.03
CL [mm]	0.35
CL' [mm]	0.53
CL'' [mm]	**0.77**
CL''' [mm]	0.60

Figure 9 shows the measured lateral axlebox acceleration and the estimated lateral axlebox acceleration which was calculated by Eq. (18) based on AL''_l and the estimated lateral axlebox acceleration which was calculated by Eq. (19) based on CL''.

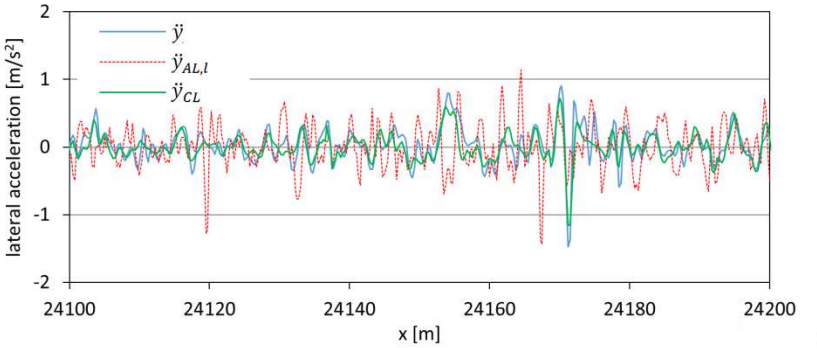


Figure 9. Measured lateral axlebox acceleration and estimated lateral acceleration based on alignment (left) and cross level

4. Conclusions

For economical track maintenance it is important to understand the relationship between the irregularities recorded by conventional track geometry measuring car and the resulting dynamic vehicle responses. The conclusions reached in this article are based on experiment carried out on only straight track in average condition which contains only stochastic track geometry irregularities but no curves or transition curves.

The difference of the signals of the vertical accelerometers (mounted above the left and right axleboxes) correlated with the signal of the lateral axlebox accelerometer significantly. Therefore, it can be concluded that the lateral accelerations of the axlebox are decisively influenced by the roll movements of the wheelset resulting from rate of the change of cross level. The differences of the vertical axlebox accelerations and the lateral axlebox acceleration are practicably proportional, and the proportionality constant results come from the geometric position of the accelerometers.

Based on the correlation analysis of the longitudinal level and the measured vertical accelerations it can be stated that the highest correlation was found in case the chord measurement was decoloured using the inverse of the transfer function and the second derivative of the resulting data series was produced. This means that the amplitude-based local fault evaluation used in the current track maintenance practice is not entirely adequate to limit the force exerted in the vehicle-track system, but rather the second order derivative of the longitudinal level. This coincides with the often stated statement in the international and Hungarian literature: not only the

amplitude of the local defect, but also its 'wavelength' is essential. However, based on the present study, it can be said that second order derivative is determinant. The optimum step of numerical derivation from the point of view of correlation is believed to depend on the speed of the vehicle and the low pass filtering rate applied on the acceleration signals: for 80 km/h and 16 Hz low-pass filter frequency a step of 0.75 m was favourable. Vertical acceleration and the longitudinal level were measured in two separate planes, in this way left axlebox acceleration could be influenced by both left and right longitudinal level.

Filtering to the D1 wavelength range greatly reduces the correlation between longitudinal level and measured vertical accelerations. This is explained by the fact that larger accelerations occur in short-wave track geometry defects that are reduced or eliminated by D1 filtering.

On the straight line examined, the lateral axlebox acceleration and any derivatives of the alignment parameters appeared to be statistically independent. This is due to the fact that the wheelset does not directly 'follow' the stochastic lateral rail irregularities with small amplitude, so there is no linear statistical relationship between the two sets of data. Small alignment irregularities cause only wheel-rail contact patch displacement, not a lateral wheelset displacement.

On the basis of the calculated correlation between the cross level derivatives and the lateral axlebox acceleration, it can be concluded that the lateral accelerations of wheelsets (and bogies) were influenced by the cross level changes primarily. The first derivative of the cross level (which is similar to the 'track twist') produced a lower level correlation, but based on the second derivative of the cross level, the lateral axlebox acceleration can be estimated well. Therefore, it is also advisable to consider the second order derivatives of the cross level for the analysis of the effects from vehicle on the track. Cross level defects can therefore cause not only vertical extra forces but significant lateral forces as well. From the point of view of safety against derailment, the lateral force component between the wheel and the rail is of key importance and attention should be paid.

However, the author dealt only straight tracks in this article, in the future dynamic parameters are able to be analysed not only on straight track, but also in curves and transition curves. The European railway track design standard [19] contains the calculation possibility (method) of different dynamic parameters (eg. 'angular acceleration around roll axis', 'angular jerk around roll axis'). Some researchers investigated the transition curves from geometrical and design aspects [20] which results can be used in the future analyses.

References

[1] L. Császár, Cs. Pálfi: Determination of the wheel/rail contact forces by different measurement methods, in: Proceedings of the 9th International Conference on Railway Bogies and Running Gears, Budapest, Hungary, 2013, pp. 153–166.

[2] M. Molodova, Z. Li, R. Dollevoet: Axlebox acceleration: Measurement and simulation for detection of short track defects. Wear 271 (1-2) (2011) pp. 349–356.
doi: http://dx.doi.org/10.1016/j.wear.2010.10.003

[3] A. Haigermoser, B. Luber, J. Rauh, G. Gräfe: Road and track irregularities: measurement, assessment and simulation. Vehicle System Dynamics 53 (7) (2015) pp. 878–957.
doi: http://dx.doi.org/10.1080/00423114.2015.1037312

[4] V. Zoller, I. Zobory: On dynamics of the track/vehicle system in presence of inhomogeneous rail supporting parameters. Periodica Polytechnica Transportation Engineering 39 (2) (2011) pp. 83–85.
doi: http://dx.doi.org/10.3311/pp.tr.2011-2.06

[5] P. Weston, C. Roberts, G. Yeo, E. Stewart: Perspectives on railway track geometry condition monitoring from in-service railway vehicles. Vehicle System Dynamics, 53 (7) (2015) pp. 1063–1091.
doi: http://dx.doi.org/10.1080/00423114.2015.1034730

[6] X. Wei, F. Liu, L. Jia: Urban rail track condition monitoring based on in-service vehicle acceleration measurements. Measurement 80 (2016) pp. 217–228.
doi: http://dx.doi.org/10.1016/j.measurement.2015.11.033

[7] Á. Vinkó, P. Bocz, Z. Posgay: A practical approach to tramway track condition monitoring: vertical track defects detection and identification using timefrequency processing techniques. SSP – Journal of Civil Engineering, 13 (s1) (2018) pp. 135–146.
doi: http://dx.doi.org/10.1515/sspjce-2018-0013

[8] Sz. Fischer, F. Horvát: Speed-dependence of railway superstructure's geometric tolerances, in: G. Köllő (Ed.) XIII. Nemzetközi Építéstudományi Konferencia: ÉPKO 2009, Cluj-Napoca, Romania, 2009, pp. 137–143, in Hungarian.

[9] European Standard EN 14363. Testing and Simulation for the acceptance of running characteristics of railway vehicles. Running Behaviour and stationary tests.

[10] Y. Liu, E. Magel: Performance-based track geometry and the track geometry interaction map. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 223 (2) (2009) pp. 111–119. doi: http://dx.doi.org/10.1243/09544097JRRT225

[11] T. Karis: Correlation between Track Irregularities and Vehicle Dynamic Response Based on Measurements and Simulations. Doctoral dissertation, KTH Royal Institute of Technology (2018).

[12] K. U. Wolter, M. Zacher, B. Slovak: Correlation between track geometry quality and vehicle reactions in the virtual rolling stock homologation process, in: 9th World Congress on Railway Research, 2011, pp. 22–26.

[13] European Standard EN 13848. Railway applications. Track. Track geometry quality.

[14] B. Lichtberger: Track compendium. EurailPress, Hamburg, 2005, pp. 400–405.

[15] R. Insa, J. Inarejos, P. Salvador, L. Baeza: On the filtering effects of the chord offset method for monitoring track geometry. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 226 (6) (2012) pp. 650–654. doi: http://dx.doi.org/10.1177/0954409712447481

[16] P. Salvador, V. Naranjo, R. Insa, P. Teixeira:. Axlebox accelerations: Their acquisition and time–frequency characterisation for railway track monitoring purposes. Measurement 82 (2016) pp. 301–312. doi: http://dx.doi.org/10.1016/j.measurement.2016.01.012

[17] M. Li, I. Persson, J. Spännar, M. Berg: On the use of second-order derivatives of track irregularity for assessing vertical track geometry quality. Vehicle System Dynamics 50 (sup1) (2012) pp. 389–401.
doi: http://dx.doi.org/10.1080/00423114.2012.671947

[18] T. Karis, M. Berg, S. Stichel, M. Li, D. Thomas, B. Dirks: Correlation of track irregularities and vehicle responses based on measured data. Vehicle System Dynamics 56 (6) (2018) pp. 967–981.
doi: http://dx.doi.org/10.1080/00423114.2017.1403634

[19] European Standard EN 13803:2017. Railway applications. Track. Track alignment design parameters. Track gauges 1435 mm and wider.

[20] Sz. Fischer: Comparison of railway track transition curves. Pollack Periodica 4 (3) (2009) pp. 99–110.
doi: https://doi.org/10.1556/Pollack.4.2009.3.9