kOMUNIKÁCIE
COMMUNICATIONS

OPERATION AND ECONOMICS

MECHANICAL ENGINEERING

ELECTRICAL ENGINEERING

CIVIL ENGINEERING

MANAGEMENT SCIENCE AND INFORMATICS

SAFETY AND SECURITY ENGINEERING

Jozef Gnap - Jana Kupculjakova - Stefania Semanova*

DETERMINATION OF TIME SAVINGS FOR PASSENGERS BY APPLYING THE PUBLIC PASSENGER TRANSPORT PREFERENCE IN CITIES

The paper deals with the issue of delays of public transport vehicles at the signal controlled junctions. Based on the road traffic survey results, the average values of vehicle delays were determined. By using those values, it is possible to define time savings for the vehicles and mainly for passengers that are transported. The time savings were determined based on the relationships and coefficients defined in this paper.

Keywords: *public passenger transport, delay time, time savings, preference of public passenger transport*

1. Introduction

The current trend of growing individual transport can be observed in central areas of Slovak cities where congestion, causing delays of passenger cars, as well as public transport vehicles arises during the peak hours. In the case of foreign cities, the problem of increasing intensity of individual transport is solved by introducing preferential measures for public passenger transport. Those measures contribute to the smooth movement of public transport vehicles on the urban infrastructure, higher speeds and travel time decreases. While maintaining the same interval at a line, the measures may also contribute to a decrease of a number of vehicles dispatched and thus to lower operating costs. This can be reflected in the level of fares, which are one of the important factors influencing the choice of transport means when travelling [1].

By introducing the preferential measures, it is possible to increase punctuality and reliability of the vehicles on lines. This is particularly essential in the case of integrated transport systems where the great emphasis is put on the coordination of individual transport modes [2]. Based on the above mentioned facts, it can be concluded that public passenger transport preference increases the efficiency of public transport and it significantly contributes to the attractiveness and competitiveness of public transport in comparison to individual transport.

Preferential measures can be divided into direct and indirect measures. The direct measures are directly linked to vehicles and roads. In general, they are further divided into physical/spatial measures, *preference at signal controlled junctions* and integrated measures combining the previous two [3], [4].

2. Determination of the delay extent of public transport vehicles at signal controlled junctions

As a great emphasis is nowadays put on the speed of transferring, public passenger transport appears as a slower way of transport and thus less attractive from a passengers' point of view. This is affected by the nature of movement of public transport vehicles, which must stop at the stops; therefore, their average speed is lower compared to passenger cars. Another negative factor affecting the speed of public transport vehicles represents the time losses, a significant part of which is caused by the delays at signal controlled junctions. The extent of such delays increases with increasing intensity of passenger cars and decreasing capacity of junctions [5], [6], [7]. In the case of the road sections with high traffic intensity, the junctions are integrated into the coordinated traffic management with regard to the main directions. However, this does not allow the vehicles, which stop at the stops located at the sections between junctions, to smoothly pass through the road network. As a result, time delays arise [6], [8], [9].

The delay time at a signal controlled junction consists of various forms of delays (Figure 1) [10], [11]:

- Delay due to stopping at a junction (D_1),
- Delay due to the time needed to approach and exit a junction; it also includes time of stopping at the junction (D_2),
- Travel time delay, which can be considered as the total value of the incurred delay (D_3).

Figure 1 depicts the differences between the delay times relating to stopping at a junction, approach and exit the junction as well as the overall travel time delay for one vehicle crossing the junction. The desired vehicle path (position in time) and the actual vehicle path are shown in this figure. The desired path represents the situation when a vehicle smoothly passes through a junction at the desired speed without any deceleration i.e. without any delay. The actual path represents the case in which a vehicle crossing a junction must reduce its speed or stop completely and then it must again achieve the required (previous) speed [10], [11]. Thus, the overall vehicle delay (CZ_v) can be expressed as the sum of the above mentioned partial delays:

$$CZ_v = D_1 + (D_2 - D_1) + (D_3 - D_2) \quad (1)$$

* **Jozef Gnap, Jana Kupculjakova, Stefania Semanova**
Department of Road and Urban Transport, Faculty of Operation and Economics of Transport and Communications, University of Zilina, Slovakia
E-mail: jozef.gnap@fpedas.uniza.sk

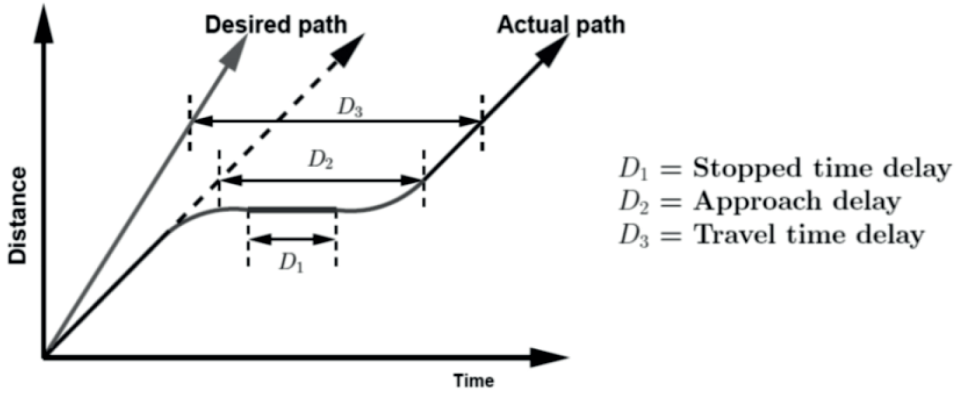


Figure 1 *Diagram of the delay time measurement principle, [10], [11]*

The extent of individual partial delays is influenced by the driving technique, driver's experience, and vehicle technical condition (this relates mainly to delay time labelled as D_2). On the other hand, traffic intensity and junction capacity also have an impact on the delay time. As a part of the research work, the extent of vehicle delays was determined by using the road traffic surveys in cities Zilina, Povazska Bystrica, Puchov and Prievidza [12], [13], [14]. The measurements were focused only on determination of the delay time at junctions, labelled as D_1.

Taking into account all the above mentioned facts, it is possible to pose the following research questions: "What delay time of public transport vehicles does occur at junctions? Is the delay extent significantly influenced by the traffic intensity on the road network?" Further, we assume that the delay time during the peak hours is at least 10 % higher compared to delays during the off-peak hours.

Time values of delays were determined based on the predetermined methodological procedure during the road traffic surveys. The survey distinguished between the vehicles crossing the junction without delay and the vehicles crossing the junction with a certain delay. Each vehicle that had to stop at the junction was considered as a delayed vehicle. Further, the time intervals between stopping the vehicle at the junction and its moving again was measured and recorded. The difference between those time values represented duration of the vehicle delay. In the case that a vehicle could not cross the junction during a single green phase and it had to stop several times, the delay times were measured unless the vehicle passed through the junction. Therefore, the delay D_1 was determined according to the methodological procedure as follows:

$$D_1 = \sum_{i=1}^{n} (CR_i - CZ_i) \tag{2}$$

where:

D_1 – delay due to stopping at a junction,
CR – time recorded when the vehicle starts to move,
CZ – time recorded at the moment when the vehicle stops at the junction,
n – number of vehicle stops at the junction.

A vehicle that slowed down due to the red signalling while approaching the junction but it did not stop was considered as a vehicle crossing the junction without delay.

3. Evaluation of delays and a definition of the relationships to determine the time savings for passengers

By carrying out the road traffic surveys, *964 public transport vehicles passing through* the *signal controlled junctions* were evaluated. The data were acquired by using people (counters) placed directly in public transport vehicles, as well as camera records, which were subsequently evaluated. Measured delay times are shown in Figure 2.

The individual points in the graph (Figure 2) depict the delay time measured for each vehicle under consideration. The horizontal solid line represents the average value of delay which was calculated as the arithmetic mean of all measured values. Based on the results, it can be concluded that *each vehicle crossing the signal controlled junction must stop for 28 seconds on average*. However, it should be noted that the extreme values of delays were also included into the mean value calculation. For this reason, the median was also used to determine the average value. This resulted in the average delay of 20 s (shown by the dashed line in Figure 2).

The above mentioned average values represent the average vehicle delay when crossing the junction. However, the values do not take into account the change of traffic intensity during a day. Assuming that a higher traffic intensity during the peak hours causes higher vehicle delays, it was necessary to determine the average value separately for each considered time interval. For this reason, the delays were divided according to the periods in which they occurred (peak or off-peak hours). The results are shown in Table 1.

During the peak hours, the average value of delay calculated as the arithmetic mean was at the level of 30 s and the median represented 23 s. This average value was calculated based on 622 recorded values of delays. During the off-peak hours, the average value of delays was determined at the lower level i.e. 24 s (arithmetic mean) or 14 s (median). The calculated results are shown graphically in Figure 3.

Based on the results obtained, it can be stated that the average time of delay increases with the increasing traffic intensity (from 24 s or 14 s up to 30 s or 23 s, respectively). This represents a 25 % increase of delay time. Taking into account median, the increase of the delay time during peak hours represents 64 % in comparison with off-peak hours.

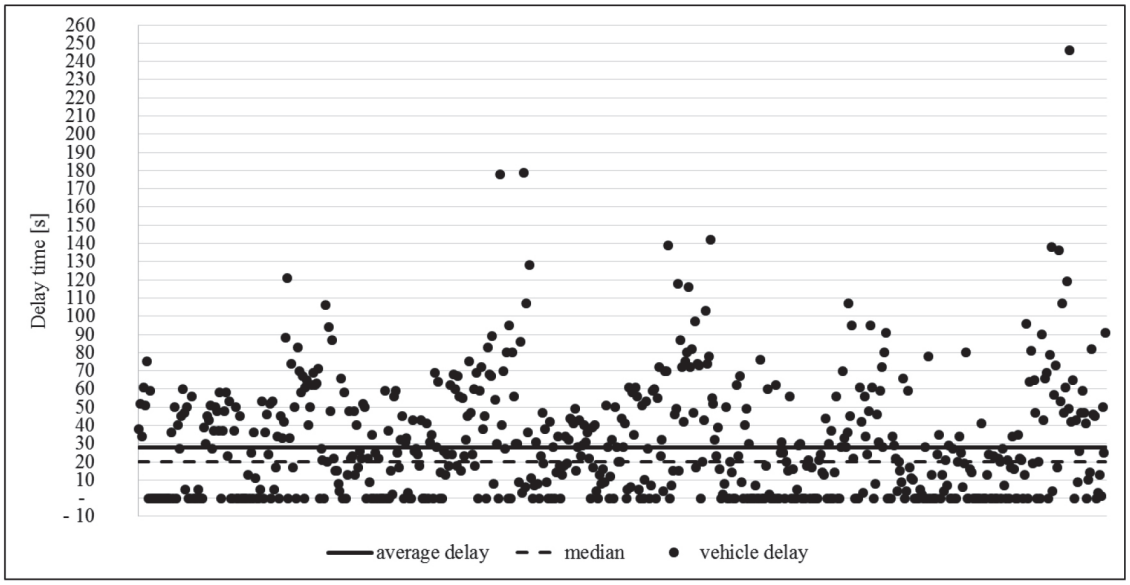


Figure 2 *Delay times measured during road traffic surveys*

Table 1 *Delays of public transport vehicles during daily periods*

Measurement	Number of considered values	Arithmetic mean [s]	Median [s]	Mode [s]
All values	964	28	20	0
Peak hours	622	30	23	0
Off-peak hours	342	24	14	0

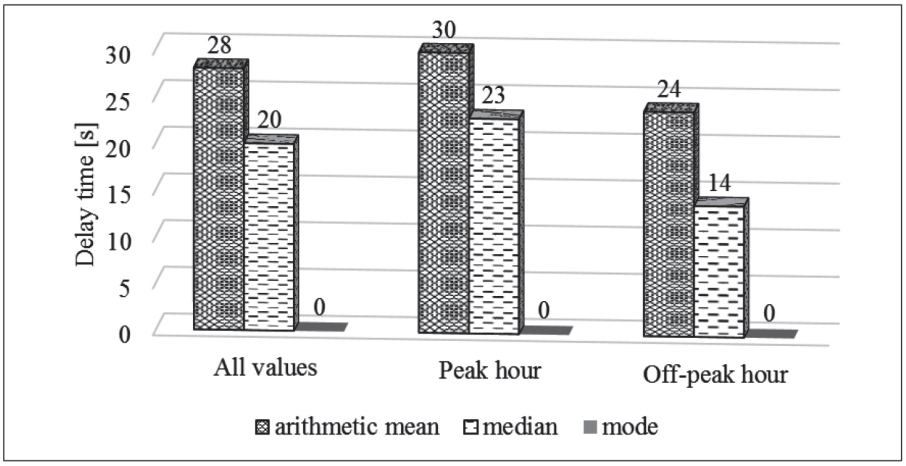


Figure 3 *Comparison of delays during the peak and off-peak hours*

3.1 Definition of relationships to determine time savings for passengers

Based on the obtained results, it can be concluded that public transport vehicles reach relatively large time losses at signal controlled junctions, in particular during the peak hours. The 25 % increase of vehicle delay time was recorded during peak hours compared to off-peak hours. It is possible to eliminate the mentioned time losses by introducing the public transport preference, which ensures that a public transport vehicle approaching the junction triggers the green signal and then it passes through the junction without delay or only with minimal delay. Let us assume that the delay time D_1 can be considered as assumed time savings of one vehicle crossing the junction:

$$D_1 = U_1 \tag{3}$$

$$D_{1ds} = U_{1ds} \tag{4}$$

$$D_{1dse} = U_{1dse} \tag{5}$$

If we further assume delay elimination by applying the public transport preference, the assumed time savings of vehicles

Table 2 *Number of vehicles crossing the junction with or without delay and the average delay time*

Measurement	Delay	Number of considered values	Parameter
All values	without (0 s)	291	Share – 30 %
	with	673	Average delay 40 s
			Median 34 s
Peak hours	without (0 s)	175	Share – 28 %
	with	447	Average delay 41 s
			Median 36 s
Off-peak hours	without (0 s)	116	Share – 34 %
	with	226	Average delay 36 s
			Median 29 s

crossing the junction during the working days can be determined according to the following formula:

$$PUV_R = ((U_{1ds} \cdot PS_{ds}) + (U_{1dse} \cdot PS_{dse})) \cdot PPD_R \quad (6)$$

where:

D_1 – delay time of one vehicle when crossing the junction,

U_1 – time savings of one vehicle crossing the junction regardless of the period during a day,

PUV_R – assumed time savings of public transport vehicles per annum,

U_{1ds} – time savings of one vehicle crossing the junction during the peak hours,

PS_{ds} – number of public transport vehicles during the peak hours per working day,

U_{1dse} – time savings of one vehicle crossing the junction during off-peak hours,

PS_{dse} – number of public transport vehicles during the off-peak hours per working day,

PPD_R – number of working days per annum.

In the case that the values of assumed time savings, divided according to the period during a day (peak and off-peak hours) are not available and the average value of delay/time savings is calculated only based on all measured values, it is possible to determine time savings by using the coefficients, which take into account the daily periods. Coefficients were determined based on the results obtained from the road traffic surveys and thus the assumed time savings value can be calculated according to the following formula:

$$PUV_R = (U_1 \cdot (PS_{ds} \cdot k_{1ds} + PS_{dse} \cdot k_{1dse})) \cdot PPD_R \quad (7)$$

where:

k_{1ds} – coefficient of time savings during the peak hours is 1.07 determined by using the arithmetic mean and 1.15 calculated as a median,

k_{1dse} – coefficient of time savings during the off-peak hours is 0.86 when applying the arithmetic mean and 0.7 calculated as a median.

Another possibility of determining the assumed time savings is that the vehicles that are not delayed when crossing the junction are not taken into account into calculations (the delay time D_1 represents 0 s). This means that every vehicle crossing the junction is considered as delayed when calculating the time savings. However, this does not have to correspond to the real situation, especially during the off-peak hours. The mentioned statement is confirmed by the road traffic surveys, the results of which showed that a part of the vehicles passed through the junction without any delay (Figure 2).

The average values of the delay time, calculated based on the previously mentioned procedure are at the higher level (Table 2), which results in *the unrealistic value of time savings*. Among the traffic survey results, we selected the vehicles that passed through the junction during the off-peak hours when the traffic intensity is lower. The reason for this was the assumption that the lower delays occur during that period or the vehicles passed through the junction without any delay during a single green phase (cycle).

To eliminate the incurred error, it is possible to use coefficients, the value of which represents the percentage share of vehicles passing through the junction with delay during the peak and off-peak hours. The survey results showed that almost one third of the vehicles crosses the junction without stopping. By using the ascertained average delays and taking into account the vehicles with no delays, it is possible to calculate the assumed time savings according to the following formula:

$$PUV_R = ((U_{1ds} \cdot PS_{ds} \cdot p_{ds}) + (U_{1dse} \cdot PS_{dse} \cdot p_{dse})) \cdot PPD_R \quad (8)$$

where:

p_{ds} – coefficient expressing the proportion of vehicles crossing the junction with delay during the peak hours – 0.72,

p_{dse} – coefficient expressing the proportion of vehicles crossing the junction with delay during the off-peak hours – 0.66.

Based on Equation (7) for determination of the assumed time savings by using the value U_1 and coefficients of the time savings, it is possible to calculate PUV_R according to the following formula:

$$PUV_R = (U_1 \cdot (PS_{ds} \cdot k_{1ds} \cdot p_{ds} + PS_{dse} \cdot k_{1dse} \cdot p_{dse})) \cdot PPD_R \quad (9)$$

where the coefficients are as follows:

k_{1ds} – coefficient of the time savings during the peak hours – 1.03 when applying the arithmetic mean and 1.06 calculated as median,

Table 3 *The assumed time savings of vehicles and passengers per a year*

Indicators	Public transport vehicles		Passengers	
	Arithmetic mean [hours/year]	Median [hours/year]	Arithmetic mean [hours/year]	Median [hours/year]
Assumed time savings (6) (10)	316.7	230.6	13 833.3	10 361.1
Assumed time savings (7)	316.6	230.6	13 821.1	10 361.1
Assumed time savings (8)	312.0	269.2	13 620.0	11 863.3
Assumed time savings (9)	313.2	269.2	13 680.0	11 871.7

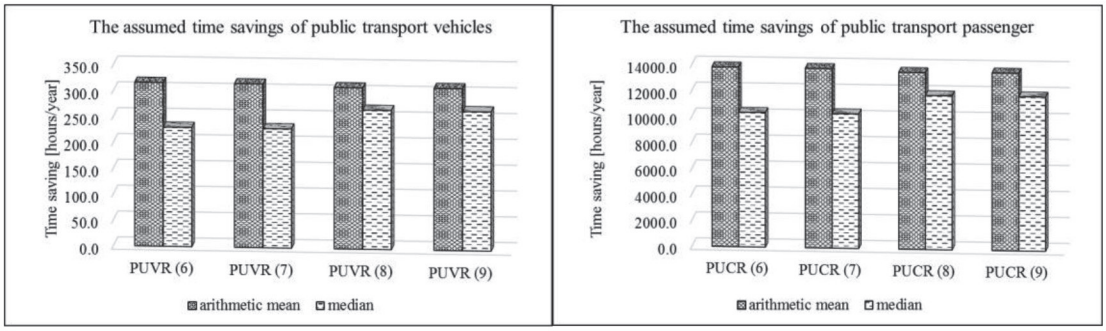


Figure 4 *The assumed time savings of vehicles and passengers*

k_{1dse} – coefficient of the time savings during the off-peak hours – 0.9 when applying the arithmetic mean and 0.8 calculated as median.

At present practice, the determination of the benefit from introducing the preferential measures in public passenger transport usually includes only the calculation of time savings for the vehicles passing through junctions. However, the public transport preference is primarily introduced for current, as well as potential passengers, in order to shorten travel time spent by commuting, travelling to schools etc. The travel time plays an important role in a decision-making process, and therefore, we recommend calculating the time savings with regard to passengers, as well, as follows:

$$PUC_R = ((U_{1ds} \cdot PS_{ds} \cdot PC_{ds}) + (U_{1dse} \cdot PS_{dse} \cdot PC_{dse})) \cdot PPD_R \quad (10)$$

where:

PUC_R – assumed time savings for passengers per annum,

PC_{ds} – average number of passengers in a public transport vehicle during the peak hours per working day,

PC_{dse} – average number of passengers in a public transport vehicle during the off-peak hours per working day.

The assumed time savings for passengers can be determined also by using the previous Equations (7), (8) and (9) into which it is necessary to add the average number of passengers.

3.2 Application of the defined relationships

The assumed time savings can be calculated by using the mentioned relationships and under consideration of the following assumptions. During the working day, 120 vehicles with the average occupancy of 50 passengers pass through the junction during the peak hours and 40 vehicles with the average occupancy of 20 passengers pass through the junction during the off-peak hours. One year is considered to include 250 working days. The results of time savings are shown in Table 3.

Values of the assumed annual time savings were determined separately for vehicles and passengers. To calculate these values, the relationships defined in the previous section were used. The values are also graphically shown in Figure 4.

4. Conclusion

The issue of public transport vehicle delays cannot be considered only as a delay of some vehicle at the junction but it is necessary to perceive the delay time mainly as the time losses of passengers.

Based on the survey results, it can be concluded that each public transport vehicle is delayed for 28 s on average when passing through one signal controlled junction (or 20 s if applying a median). However, this delay time does not take into account the traffic intensity changing during the day. For this reason, the average values of delay time during the peak hours (30 s or 23 s) and the off-peak hours (24 s or 14 s) were also calculated. Furthermore, the assumption that the delay time increases by more than 10 % during the peak hours was confirmed. The delay time during the peak hours increases by 25 % compared to the off-peak hours. When comparing the values calculated by using a median, this increase represents 64 %.

If we consider the average delay time as the assumed time savings and we use the relationships defined in this paper for the calculation, the assumed annual time savings for passengers will exceed 10 000 hours. Taking into account the average hourly salary, this can result in enormous financial losses.

By applying preferential measures, it is possible to eliminate delays and thus to provide significant time savings for vehicles, as well as for passengers. However, the preferential measures must be applied scrupulously in order to maximize the effects of public transport preference. The most appropriate method is a combination of spatial preference and preference at signal controlled junctions as well as introduction of the priority in driving for public transport vehicles.

References

[1] DYDKOWSKI, G., URBANEK, A.: Prices as an Element of Comparative Studies on Peoples Travels in Cities. Communications - Scientific Letters of the University of Zilina, 19(2), 43- 49, 2017.

[2] POLIAK, M., MRNIKOVA, M., JASKIEWICZ, M., JURECKU, R., KACIAKOVA, B.: Public Transport Integration. Communications - Scientific Letters of the University of Zilina, 19(2), 127- 132, 2017.

[3] HAVLENA, O., NOVOTNY, V.: Catalogue of Preference Measures for Public Transport [online]. CVUT, Praha, 2016. Available: http://preferencevhd.info/wp-content/uploads/2016/01/PREFOS-Katalog-preferen%C4%8Dn%C3%ADchopat%C5%99en%C3%AD.pdf.

[4] KUPCULJAKOVA, J.: Possibilities of Ensuring Urban Public Transport Priority. Proceedings of 12th International Conference Transport Systems Telematics (TST´12), Poland, p. 51, 2012.

[5] CERNICKY, L., KALASOVA, A., MIKULSKI, J.: Simulation Software as a Calculation Tool for Traffic Capacity Assessment. Communications - Scientific Letters of the University of Zilina, 18(2), 99-103, 2016.

[6] KALASOVA, A., CERNICKY, L., KUPCULJAKOVA, J.: The Impact of Public Transport Priority on the Traffic in the Chosen Part of the City of Zilina. Transport Problems: International Scientific Journal, 9(2), 19-26, 2014.

[7] KALASOVA, A., KUPCULJAKOVA, J.: Microscopic Simulation of Public Transport Priority at Signal Controlled Junction in Zilina City. Journal of Slovak Company for Systematic Integration, 1, p. 8, 2013.

[8] KALASOVA, A., CERNICKY, L., KUBIKOVA, S.: Microscopic Simulation of Coordinated Route in the City of Zilina. Communications - Scientific Letters of the University of Zilina, 16(2), 46-50, 2014.

[9] KUBIKOVA, S., KALASOVA, A., CERNICKY, L.: Microscopic Simulation of Optimal Use of Communication Network. Communications in Computer and Information Science, 471, 414-423, 2014.

[10] CERNICKY, L.: Time Delay at the Intersection – Comparison of Analytical and Simulation Method. Proceedings of Conference Transport Engineering 2016, Slovakia, 28-39, 2016.

[11] CERNICKY, L., KALASOVA, A., KAPUSTA, J.: Signal Controlled Junctions Calculations in Traffic-Capacity Assessment - Aimsun, OmniTrans, Webster and TP 10/2010 Results Comparison. Transport Problems, 11(1), 121-130, 2016. https://doi.org/10.20858/tp.2016.11.1.12

[12] Completion of Transport Service Plan by Public Passengers Transport in Accordance with the Act of the National Council of the Slovak Republic No. 56/2012 Collection of Laws on Road Transport for City of Puchov – Public Transport Priority in the City of Puchov. Department of Road and Urban Transport, University of Zilina, Zilina, 2016.

[13] Public Transport Priority in the City of Povazska Bystrica. Department of Road and Urban Transport, University of Zilina, Zilina, 2016.

[14] Public Transport Priority of Prievidza City - Optimal Variant. Department of Road and Urban Transport, University of Zilina, Zilina, 2017.

Oldrich Hyks - Kristyna Neubergova - Pavel Pribyl*

INFLUENCE OF DRIVING FLUENCY ON ECONOMIC AND ECOLOGICAL ASPECTS OF TRANSPORT

The paper is focused on the influence of driving behaviour on fuel consumption and the subsequent environmental impacts. The main issue is to estimate the economic and environmental costs connected with the forced deceleration or stopping and the subsequent acceleration of a motor vehicle in urban traffic. Authors introduce a physical model of a vehicle with variable parameters that allows the calculation of economic and ecological losses in congestions on the basis of data generated by an arbitrary floating car. The losses include the lost kinetic energy of a vehicle during the forced braking, accelerated degradation of vehicle components, the loss of time of drivers and passengers and the increased ecological footprint. Energetic costs are estimated from the vehicle engine efficiency with which the fuel energy is transformed during an acceleration to kinetic energy of a vehicle lost during braking. Further, these costs are estimated from fuel calorific value and fuel price. Costs resulting from the degradation of vehicle components are estimated from their average stated lifetime.

Keywords: *vehicle dynamics, fuel consumption, forced vehicle deceleration*

1. Introduction

The aim of this paper is to provide an estimation of the economic and environmental costs due to the forced deceleration or stopping and the subsequent acceleration of a motor vehicle in urban traffic. This estimation is based on a dynamic vehicle model with variable parameters and thus it can be obtained from the data generated by an arbitrary floating car. The problem discussed in this treatise is particularly topical in rapidly developing urbanized areas around cities - see [1].

2. Vehicle dynamics

2.1 Forces acting on a vehicle

Consider a vehicle moving with an acceleration α. In the travel direction, a driving force F is acting. The movement is opposed by a vehicle resistance that includes:[1]

tire rolling resistance: $R_f = fmg$ (1)

aerodynamic drag: $R_v = \frac{1}{2}\rho c_d A v^2$ (2)

acceleration resistance of rotating parts: $R_R = ma\xi$ (3)

breaking force: $R_B = \mu mgx$ (4)

Here f denotes a rolling resistance coefficient, m vehicle mass, g gravitational acceleration, ρ air density, c_d aerodynamic drag coefficient, A projected front area, v vehicle velocity, a vehicle acceleration, ξ coefficient including the inertia of rotating parts, μ tire adhesion coefficient, x breaking intensity expressed by a value from the interval [0,1].

As a model vehicle for calculations, Skoda Octavia 1.4 MPI/55 kW with gasoline ignition engine has been chosen, since it belongs to the best-selling cars in the Czech Republic. Therefore, the following values are considered (see [2]): f = 0.015, m = 1560 kg, c_d = 0.30, A = 2.069 m^2, ξ = 0.04,[2] g = 9.81 ms^{-2}, ρ = 1.25 kgm^{-3}, μ = 0.8 (for asphalt).

If the vehicle velocity v is constant, the first Newton's motion law implies that the sum of acting forces is equal to zero:

$$0 = F - R_v - R_f = F - \frac{1}{2}\rho c_d A v^2 - fmg \quad (5)$$

The driving force F can therefore be expressed in the form:

$$F = \frac{1}{2}\rho c_d A v^2 + fmg \quad (6)$$

For the mentioned model vehicle, the driving force necessary to maintain a constant velocity of e.g. v = 50 kmh^{-1} ≐ 13.89 ms^{-1} is

[1] A movement in a horizontal plane is considered. For an uphill journey, a climbing resistance $R_C = mg\sin\alpha$ should be added, where α denotes an angle between a road and a horizontal direction. Similarly, all mentioned forces containing a term mg would be multiplied by $\cos\alpha$. Nevertheless, we suppose that a driver getting over an elevation h, which requires an additional work $W = \int_A^B mg\sin\alpha\,\mathrm{d}s = mgh$, later returns to the initial altitude and earns the same energy back. In the case of the most frequent grades, the term $\cos\alpha$ would influence the price of one stopping and a subsequent acceleration at the third decimal position or further, the fuel consumption estimation would be affected at the fourth decimal position. Therefore, we put $\alpha = 0$ and thus $\alpha = 1$ in this paper.

[2] We consider the least value from the interval between 0.04 and 0.7, convenient for personal vehicles.

* [1]Oldrich Hyks, [2]Kristyna Neubergova, [2]Pavel Pribyl
[1]Department of Applied Mathematics, Faculty of Transportation Sciences, Czech Technical University in Prague, Czech Republic
[2]Department of Transportation systems, Faculty of Transportation Sciences, Czech Technical University in Prague, Czech Republic
E-mail: hyks@fd.cvut.cz

approximately equal to F = 304.39 N. The corresponding power (and thus also a consumption) is very small:

$$P = F \cdot v = 4227.64 \text{ W} \tag{7}$$

If the vehicle moves with an acceleration a (during the engine engagement, i.e., without braking), the second Newton's motion law implies:

$$ma = F - R_R - R_v - R_f \tag{8}$$

The driving force F therefore satisfies the equation:

$$F = ma + ma\xi + \frac{1}{2}\rho c_d A v^2 + fm \tag{9}$$

Denoting $\vartheta = 1 + \xi$, it can be expressed in the form:

$$F = ma\vartheta + \frac{1}{2}\rho c_d A v^2 + fmg \tag{10}$$

To reach the maximal acceleration (from 0 to 100 km in 15.5 s), the engine force acting on a vehicle has to be F = 3208.48 N. For an instantaneous velocity of 50 kmh^{-1} and the stated acceleration, the engine power would increase more than 10 times compared to the constant-velocity movement:

$$P = F \cdot v = 44\,562.27 \text{ W} \tag{11}$$

It is almost equal to the maximal engine power, the fuel consumption would also be almost maximal in this case. Nevertheless, the consumed energy is not yet lost. It is stored in the form of kinetic energy and consumed later during the vehicle deceleration or stopping. The ratio of a part utilized for a vehicle movement (when a car is put in neutral and moves by inertia) and a part dissipated to heat (during braking) is influenced by a driving style.

2.2 Motion equations

In the authors' paper [3], the investigation of the vehicle dynamics related to acceleration and braking was based on the corresponding first-order differential equations. In the former case, it is sufficient to express the acceleration a, which is a derivative of velocity, i.e., $a(t) = \dot{v}(t)$ Then one obtains the equation of motion as:

$$\dot{v}(t) = \frac{1}{\vartheta}\left(\frac{F}{m} - \frac{1}{2m}\rho c_d A v^2(t) - fg\right) \tag{12}$$

Further, the braking of various intensities with a non-zero initial velocity and putting a car to neutral were investigated. In this case, the following equation holds:

$$0 = ma\vartheta + \frac{1}{2}\rho c_d A v^2 + fmg + \mu mgx \tag{13}$$

From this relation, an acceleration can be expressed (a negative value would indicate a deceleration). Putting $a(t) = \dot{v}(t)$ leads to the following equation of motion:

$$\dot{v}(t) = \frac{1}{\vartheta}\left(-\frac{1}{2m}\rho c_d A v(t)^2 - fg - \mu gx\right) \tag{14}$$

By solving Equation (12) and Equation (14), the time developments of acting forces, acceleration, velocity and travelled distance were obtained. Based on those results, an economic analysis of consumed fuel and lost time for various braking intensities was carried out. It was shown that the highest economic losses were connected to vehicle deceleration and stopping, and that they were caused not only by an increased consumption, but also by time losses and losses given by a vehicle fading. The obtained relations demonstrated a decisive influence of congestions on traffic economy.

3. Energetic analysis

3.1 Consumption

In the subsequent research, the mentioned relations were used for the analysis of data provided by a floating vehicle that was moving in a real traffic and recorded instantaneous velocity and travelled distance. An expression of an amount of energy utilized for vehicle movement for the given velocity behaviour was considered as follows.

The expended energy is equal to the work done by the driving force of an engine through tires acting on a road surface in a given segment of journey:

$$W = \int_{s_0}^{s_1} F ds \tag{15}$$

Due to Equation (10), a work element can be expressed in the following form:

$$dW = F \cdot ds = \left(ma\vartheta + \frac{1}{2}\rho c_d A v^2 + fmg\right) \cdot ds \tag{16}$$

The amount of fuel consumed to earn this energy (for the sake of clarity, it is expressed directly in volume units) is given by the relation:

$$dV_W = \frac{dW}{\eta H_v} \tag{17}$$

where η denotes an engine efficiency and H_v stands for the caloric capacity of the used fuel, transformed to volume units. For the model vehicle considered in this paper, values η = 0.32 and H_v = 31.61 MJ/l were used. In reality, the total amount of the consumed fuel would be increased by a certain basal consumption that is necessary to keep an engine running and that is constant in time. For the considered model vehicle, this basal consumption is approximately equal to V_0 = 1 litre per hour. The total amount of the fuel consumed to carry out a work dW in time dt is therefore equal to the sum:

$$dV = \frac{dW}{\eta H v} + V_0 \cdot dt \tag{18}$$

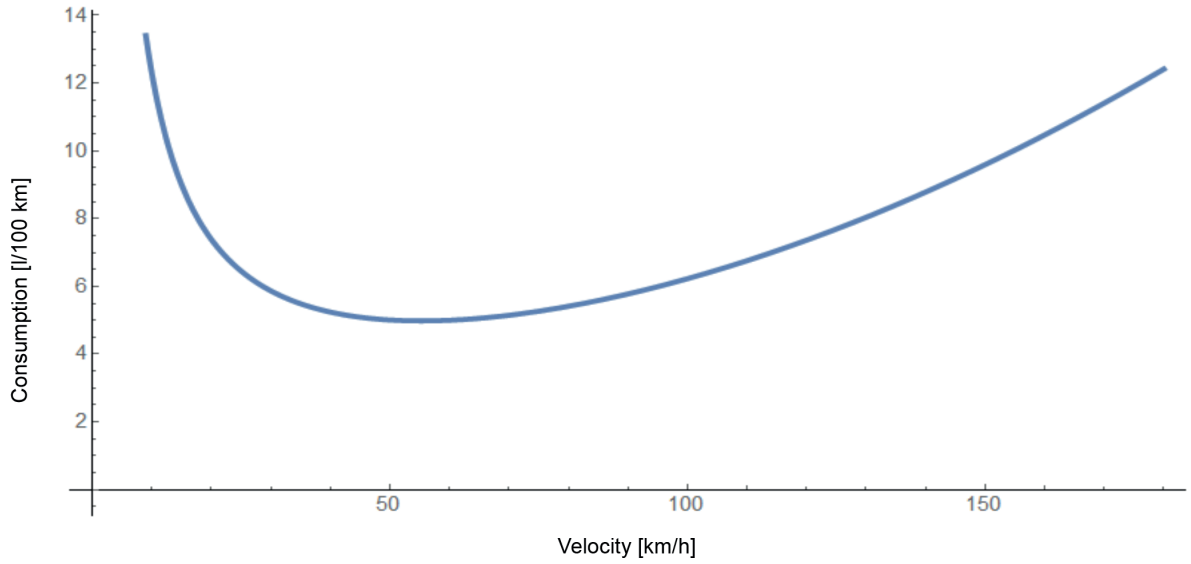


Figure 1 *Dependence of the fuel consumption on velocity*

For a known velocity $v = ds/dt$, the Equation (16) and Equation (18) imply:

$$dV = \frac{F \cdot ds}{\eta H_V} + V_0 \cdot \frac{ds}{v} = \\ = \left[\frac{1}{\eta H_V} \cdot \left(ma\vartheta + \frac{1}{2}\rho c_d A v^2 + fmg\right) + \frac{V_0}{v}\right] \cdot ds \tag{19}$$

An instantaneous fuel consumption can thus be expressed as:

$$\frac{dV}{ds} = \frac{1}{\eta H_V} \cdot \left(ma\vartheta + \frac{1}{2}\rho c_d A v^2 + fmg\right) + \frac{V_0}{v} \tag{20}$$

From Equation (20) it is obvious that the increase of fuel consumption in traffic congestions is caused not only by the kinetic energy losses. Another reason is that during the low velocities, the proportion of basal fuel consumption increases and so does the total fuel consumption, as well. The graph in Figure 1 demonstrates the dependence of the fuel consumption expressed in litres per 100 km on the velocity expressed in km/h (for each velocity value, a stationary movement with $a = 0$ is considered). Notice that the lowest consumption is reached for the velocity close to 50 km/h, i.e., close to the typical urban velocity.

3.2 Fuel consumption for a model vehicle

Using Equation (19), the values of the fuel consumption of the floating car driving on a part of the Prague circuit (Jizni spojka - South Connection) between the street 5. Kvetna and the Mrazovka tunnel, in various traffic densities, were estimated. Authors could work with a survey of instantaneous velocities v_i and the corresponding distances s_i (in 10 up to 30 metres). The volume of the consumed fuel was determined numerically as a sum of partial volumes consumed on individual segments.

For the segment $\langle s_i, s_{i+1} \rangle$, the motion was considered as uniformly accelerated with a constant acceleration

$$a = \frac{\Delta v}{\Delta t} = \frac{v_{i+1} - v_i}{s_{i+1} - s_i} \cdot \frac{v_{i+1} + v_i}{2} = \frac{v_{i+1}^2 - v_i^2}{2(s_{i+1} - s_i)} \tag{21}$$

where the time Δt was expressed as a function of an average velocity in the form $\Delta t = \Delta s / v_a$ with $v_a = (v_{i+1} + v_i)/2$.

An absolute fuel consumption on a given segment was then obtained on the basis of the Equation (19):

$$\Delta V = \left[\frac{1}{\eta H_V} \cdot \left(m\vartheta \cdot \frac{v_{i+1}^2}{2(s_{i+1} - s_i)} + \frac{1}{2}\rho c_d A \cdot \right.\right. \\ \left.\left. \cdot \left(\frac{v_{i+1} + v_i}{2}\right)^2 + fmg\right) + \frac{V_0 \cdot 2}{v_{i+1} + v_i}\right] \cdot (s_{i+1} - s_i) \tag{22}$$

If a vehicle deceleration is higher than the deceleration caused by natural resistances (i.e., in the case of braking), a computed driving force is negative, which means that the amount of fuel should be subtracted. Since the considered model car does not possess a recuperation system and a kinetic energy would be dissipated to heat, the driving force was set to be equal to zero.

At the end, the absolute fuel consumption during the travel on the segment of the length 7.61 km was replaced by a usually considered fuel consumption per 100 km.

The two examples are presented in the following paragraphs. The first one is based on data recorded by a floating car in an afternoon peak hour, when the car was forced to break and even stop very often. It took 27.10 min to finish the whole journey, and the average velocity of the vehicle was 16.85 km/h. The graph illustrating the behaviour of instantaneous velocities is depicted in Figure 2. The value of an absolute fuel consumption, calculated on the basis of a model described in the present paper, is equal to 1.06 l, which corresponds to the consumption of 13.91 l/100 km.

The second example is based on the record from a morning peak when the vehicle movement was also nonstationary, but it was not necessary to slow down up to the zero velocity. This time, the same journey was finished in 7.98 min and the average velocity of the floating car was 57.19 km/h. The behaviour of instantaneous velocities is depicted in Figure 3. The calculated absolute fuel

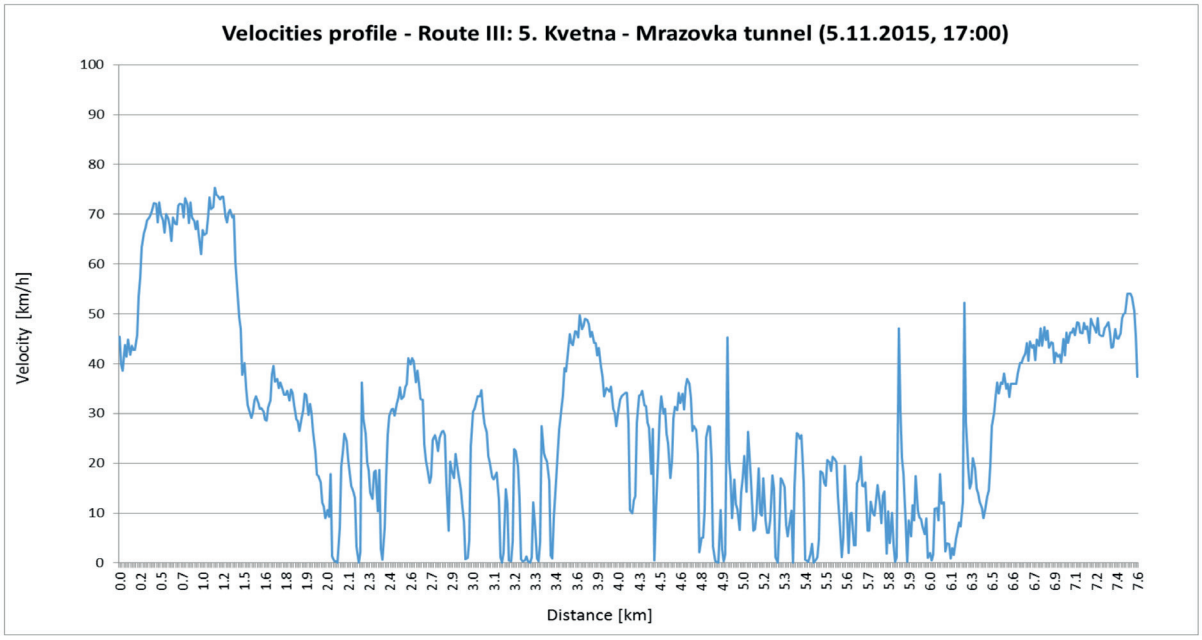


Figure 2 Behaviour of instantaneous velocities of a floating car in an afternoon peak

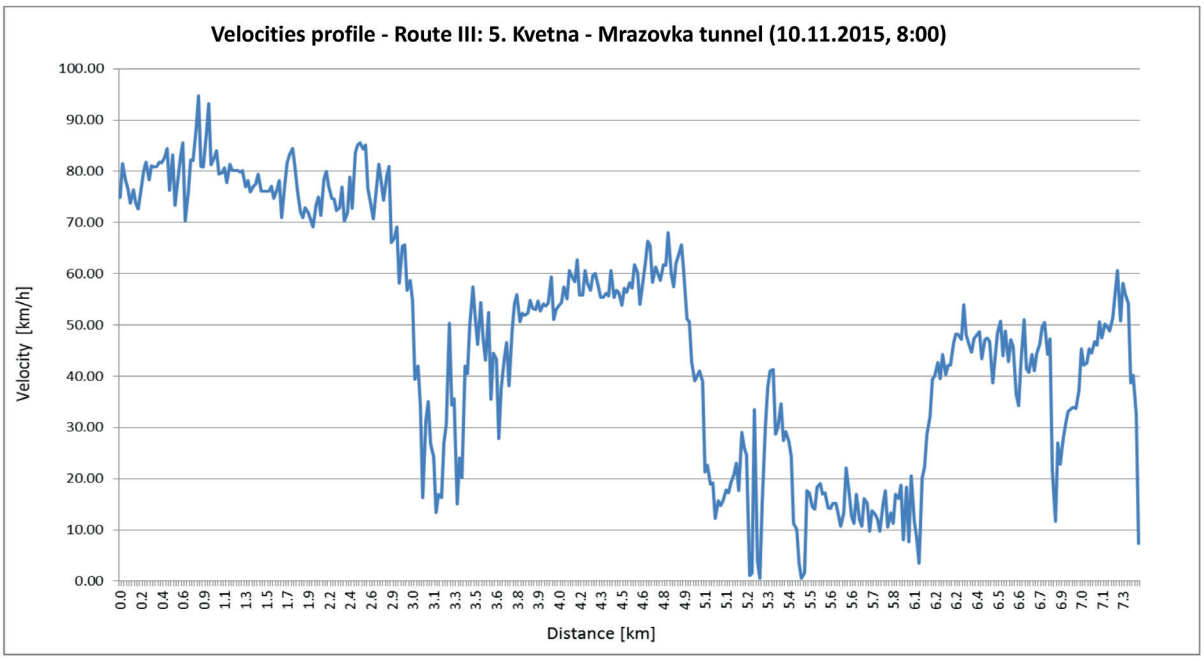


Figure 3 Behaviour of instantaneous velocities of a floating car in a morning peak

consumption is now equal to 0.64l, which corresponds to the consumption of 8.45 l/100 km.

For a comparison, one can easily calculate the fuel consumption for a constant velocity of 80 km/h, namely for the maximal allowed velocity on the considered road. The absolute consumption would be equal to 0.41l, which corresponds to the consumption of 5.41 l/100 km. The total time would be 5.71 min in this case.

The described model demonstrates that the fuel consumption indeed noticeably increases in congestions, which leads to a substantial increase of economic costs, too. On the short journey considered in the previous example, the absolute consumption increases approximately by 0.65l of gasoline in a peak hour compared to a stationary movement, which corresponds to the increase of economic costs by 15 CZK (approximately 0.59 EUR).

4. Emissions of carbon dioxide CO_2

During an ideal burning of fossil fuels, which are pure hydrocarbons, only CO_2 a H_2O would be emitted. Nevertheless, a great amount of other products are often formed in actual combustion processes, and only the following ones are being monitored: Carbon monoxide CO, hydrocarbons HCs, nitrogen oxides NO_x, sulphur dioxide SO_2 and soot. Their ratio to the amount of CO_2 and thus to the amount of consumed fuel is given

by the quality of combustion and additional technologies, and it is defined by a Euro-norm I-V. Therefore, if one determines the amount of fuel consumed additionally due to stopping and subsequent acceleration, one can easily determine the corresponding amount of CO_2 and according to the mentioned norm, the amounts of other selected combustion products, too. Moreover, CO_2 is a greenhouse gas and its production represents an important parameter. Its production can only be decreased by decreasing the absolute fuel consumption.

For the estimation of actually emitted CO_2, a methodology published by the United States Environmental Protection Agency (EPA) was used - see [4]. The calculation is based on chemical equations expressing amount of carbon in fuel during gasoline and diesel combustion. It also respects the recommendation of the Intergovernmental Panel on Climate Change (IPCC) to take into account the fact that only 99 % of fuel volume is burned perfectly. The EPA methodology is based on the estimate that one gallon (3.785 l) of gasoline contains 2.421 kg of carbon C, which corresponds to 0.644 kg of carbon in 1 liter of gasoline. The amount of emitted CO_2 can be approximated relatively precisely by the multiplication of this value by the factor 0.99 expressing the degree of imperfection of fuel combustion, and further by the ratio of molar masses of carbon dioxide CO_2 (44 g/mol) and carbon C (12 g/mol), which states how many times the mass of a molecule of carbon dioxide, created by bonding an aerial oxygen to an atom of carbon during the combustion, is greater than the mass of an atom of carbon alone:

$$m = \left(0.644 \cdot 0.99 \cdot \frac{44}{12}\right)\text{kg} = 2.338\ \text{kg} \tag{23}$$

For the considered model vehicle and the journey (less than 8 km long) discussed in part 3.3, one obtains the conclusion that in an afternoon peak, this individual vehicle produces 1.55 kg of CO_2 more than during the transit by a constant velocity of 80 km/h.

5. Evaluation of the time costs

There exist several possible approaches to evaluating the cost of time. Many of them are based on meta-analytic studies, e.g., the study by Wardman et al. [5]. For business trips, one of the most frequently used methods is the cost savings principle [6], when the time value is supposed to correspond to gross salary costs (i.e., gross salary plus an overhead), which are published in statistical overviews. In the Czech Republic, these costs are contained in the overviews of the Czech Statistical Office (CSU, see [7]). Travel compensations are also stated in the regulation published by the Ministry of Labour and Social Affairs (MPSV, see [8]). The case of private travels is certainly more difficult. Estimates based on willingness to pay is sometimes used, based either on manifested or on expressed preferences. In the Czech Republic, only several studies, aimed at this problem, have been performed (see e.g. [9], [10] and [11]), and a necessary representative investigation of travel behaviour of citizens has not yet been conducted. The values are therefore usually borrowed from foreign sources, e.g., from studies [5] and [6]. Nevertheless, those estimates are usually higher than the true domestic values. Based on the detailed search and analysis of existing methodologies, the value 300 CZK/h has been selected for the sake of this paper.

In the example considered in part 3.2, the travel time for the given journey was 21.39 minutes longer in an afternoon peak than in the case of a constant velocity of 80 km/h, which corresponds to a cost of 106.95 CZK (4.20 EUR).

6. Costs of vehicle component replacement

To estimate the costs connected to the need to maintain vehicle components, only those components, the lifetime of which is rather a function of kilometres driven and driving style rather than the vehicle age as such, were considered. The cycle of their replacement is usually stated in terms of kilometres driven. As it was shown above, the fuel consumption is influenced by the driving style, the amount of stoppings and accelerations and braking intensity. Therefore, if one transforms the component price to a price per kilometre driven and consequently to the price per one litre of consumed fuel, one can evaluate the forced stoppings in traffic congestions not only with respect to the price of an additional amount of fuel, but with respect to increased component replacement costs, as well:

$$C_f = \sum_{i=1}^{n} \frac{\text{component price}\,[\text{CZK}]}{\text{lifetime}\,[\text{km}]} \cdot \frac{100\,[\text{km}]}{\text{fuel consumption}\,[\text{l}]} \tag{24}$$

To get a rough estimate, one can consider only the components that have to be replaced most frequently, their replacement rates depend substantially on a driving style and their prices are significant, namely gears (ca. 18 000 CZK (708 EUR) per 120 000 km, brakes (ca. 18 000 CZK (708 EUR) per 120 000 km) and tires (ca. 4 000 CZK (157 EUR) per 40 000 km). For the model vehicle, one obtains the following estimation of costs given by component replacement per litre of a fuel:

$$C_f = \left(\frac{18\,000}{120\,000} + \frac{18\,000}{120\,000} + \frac{4\,000}{40\,000}\right) \cdot \frac{100}{6.9} \doteq \doteq 5.80 \frac{\text{CZK}}{1} \quad \left(0.23 \frac{\text{EUR}}{1}\right) \tag{25}$$

In the example discussed in part 3.2, the fuel consumption in an afternoon peak was 0.65 l of gasoline higher that in the case of a constant velocity of 80 km/h. It corresponds to the loss caused by additional component replacement in congestions in the range of 3.77 CZK (0.15 EUR).

7. Conclusion

A question may be raised, what is the value of calculating fuel consumption based on the physical model and data (velocity and distance) recorded by a floating vehicle, when consumption can be measured directly by the testing car. There exist several reasons for that. First of all, the measurement of an actual consumption is technically difficult. Data from a dashboard computer are only approximate, the consumption is estimated on the basis of the engine revolutions and openings of a fuel flap. The more exact

measurement of the actual amount of consumed fuel would require the replacement of a standard fuel tank by a special one enabling reading fuel decrease. This solution is unusual and demanding with respect to realization, as well as operation and maintenance.

The main reason is the fact that a direct consumption measure would only allow to measure the fuel consumption for a unique type of a vehicle used as a floating one. The described physical model allows the modification of the input parameters completely arbitrarily, and thus allows e.g. the estimation of a consumption of a truck based on data from a floating personal vehicle, even in segments where trucks are not allowed to drive.

The model can also become a part of a virtual transit simulation on some journey after its planned modification or even on a journey that does not exist at all.

In the forthcoming research, we would like to acquire data from an accelerometer of a floating car to find the behaviour of acceleration as a function of time, with the location control based on the GPS, and to use those data for the further improvement of the model.

References

[1] NEUBERGOVA K.: Urban Sprawl Syndrome and Transport. Proceedings of IV. Czech-Slovak Scientific Conference "Transport, Health and Environment", Czech Republic, 203-208, 2010.

[2] Technical Specification of Skoda Octavia Cars [online] (in Czech). Available: http://www.auto.cz/kompletni-technicka-data-nove-octavie-17169.

[3] HYKS O., NEUBERGOVA K.: Influence of Congestions on Economy and Ecological Footprint of Individual Vehicle Transportation (in Czech). Proceedings of X. Czech-Slovak Scientific Conference "Transport, Health and Environment", Czech Republic, 137-145, 2016.

[4] EPA - ENVIRONMENTAL PROTECTION AGENCY: Greenhouse Gas Emissions from a Typical Passenger Vehicle. EPA, Ann Arbor, 2014.

[5] WARDMAN M., CHINTAKAYALA P., DE JONG G., FERRER D.: European Wide Meta-Analysis of Values of Travel Time, Final Report to the European Investment Bank. University of Leeds, Leeds, 2012.

[6] HEATCO: Developing Harmonised European Approaches for Transport Costing and Project Assessment [online]. IER, 2004. Available: http://heatco.ier.uni-stuttgart.de/.

[7] CZECH STATISTICAL OFFICE: Consumer Price Indexes. Average Fuel Prices by Month from 2015 to 2017 [online] (in Czech). Available: https://www.czso.cz/csu/czso/indexy-spotrebitelskych-cen-zivotnich-nakladu-zakladni-cleneni-prosinec-2017.

[8] MINISTRY OF LABOUR AND SOCIAL AFFAIRS: Regulation of the Ministry of Labour and Social Affairs Nr. 385/2015 Sb. [online] (in Czech). Prague, 2016. Available: http://www.mpsv.cz/files/clanky/23440/vyhlaska_X2016.pdf.

[9] BRUHA J, BRUHOVA - FOLTYNOVA H.: Estimation of the Value of Time in the Czech Republic [online]. KIT, Kolin. Available: https://www.czp.cuni.cz/czp/images/stories/ Vystupy/TranExt/4-Bruha-Econometric_Estimation_of_Value_of%20Time-Pilsen_study.pdf, 2011.

[10] MACA V., MELICHAR J., et al: Methodology of Traffic Externalities Quantification (in Czech). UK and ATEM, Prague, 2013.

[11] MACA V., et al: TranExt, Quantification of External Traffic Costs in the Czech Republic [online] (in Czech). Available: https://www.czp.cuni.cz/czp/index.php/cz/ukoncene/18-tranext-vav-md-20072011.

Anna Goral - Wojciech Zorawski - Otakar Bokuvka*

INFLUENCE OF THE TITANIA CONTENT ON THE MICROSTRUCTURE AND PROPERTIES OF PLASMA SPRAYED ALUMINA-TITANIA COATINGS

The composite alumina-titania coatings have been widely used in industry fields to protect metallic components from wear, corrosion and thereby prolonging their service time. This paper presents an analysis of the microstructure and mechanical properties of Al_2O_3-TiO_2 composite coatings, plasma sprayed on steel substrates from feedstocks containing various concentrations of the TiO_2 powder. The coatings revealed lamellar microstructure, formed by the rapid solidification of molten droplets of the powder on previously deposited splats. The main phase identified in the coatings was γ-Al_2O_3 phase, the others were TiO_2 and α-Al_2O_3, as well as amorphous phase. The results indicated significant improvement of hardness, Young's modulus and fracture resistance of coatings with an increase in TiO_2 from 3 wt. % to 13 wt. %. The friction coefficient was found to be the same for both coatings.

Keywords: *Al_2O_3-TiO_2 coating, plasma spraying, hardness, fracture resistance*

1. Introduction

The very good combination of mechanical and thermal properties of oxide-based coatings makes them a desirable coating working in harsh conditions that involve mechanical and corrosive factors. Nowadays, the most common methods for a ceramic coating deposition are an atmospheric plasma spraying (APS) [1] and the high-velocity oxygen fuel (HVOF) thermal spraying [2]. The APS sprayed ceramic coatings are widely applied to structural materials and various machine parts in order to improve wear, oxidation and corrosion resistance [3]. The effect of spraying on fatigue properties of structural materials in the ultra-high-cycle regime was also examined; only a weak influence was observed [4]. Among the alumina-based coatings, particularly those containing TiO_2 are of interest since they possess enhanced toughness and wear resistance compared to Al_2O_3 coatings [5]. They are conventionally used hard coatings for their resistance to chemical and abrasive wear characteristics. Plasma sprayed Al_2O_3-TiO_2 coatings have been widely used as wear, corrosion and thermal resistant coatings in machinery, textile, printing, and aerospace industries [6], [7]. Addition of various amounts of TiO_2 to Al_2O_3 feedstock can change their mechanical properties. Titania in alumina is intended to lower the melting temperature of the mixture due to the lower melting temperature of titania compared to the melting temperature of alumina, which in turn facilitates the melting process of alumina-titania powders [8]. In addition, the fracture toughness is an important phenomenon affecting tribological properties and this is allowed by the TiO_2 additive to improve the wear resistance [9]. The wear of ceramic coatings is sensitive to their mechanical properties and therefore they should be characterized in details.

This study aims to identify the effects of TiO_2 addition to Al_2O_3 powder on the mechanical behavior of plasma-sprayed Al_2O_3-TiO_2 coatings.

2. Experimental

The coatings were plasma sprayed on steel substrates using the Plancer PN-120 system equipped with the Thermal Miller 1264 powder feeder. The feedstock powders were commercially available powder mixtures Al_2O_3:TiO_2 wt. % ratio = 87:13 and 97:3 (Amdry 6228, FST C-336.31, respectively). The microstructures of obtained alumina-titania coatings were characterized using the scanning electron microscope FEI E-SEM XL30 and X-ray diffraction diffractometer (D8 Discover, Bruker with CoKα radiation). Indentation tests were performed at the 100 mN maximum load, the 200 mN/min loading rate and unloading rate with the use of the CSEM-MCT equipment. On each coating, at least 7 indentations were made, and for further analysis, the average values were taken. The scratch tests were carried out on the MCT equipment using a Rockwell C diamond with a radius of 200 μm. The scratch distance was 5 mm. The maximum load value P_{max} was 30 N. Changes in the load value were linear, and the range over the entire scratch distance was from 0.01 N to 30 N. The rate of the indenter movement was 5 mm/min. The test and analysis were carried out in accordance with the PN-EN 1071-3 standard.

3. Results and discussion

The Al_2O_3-3TiO_2 and Al_2O_3-13TiO_2 coating feedstocks consisted of particles of α-Al_2O_3 (rhombohedral, a = 4.76 Å,

* [1]Anna Goral, [2]Wojciech Zorawski, [3]Otakar Bokuvka
[1]Institute of Metallurgy and Materials Science PAS, Krakow, Poland
[2]Faculty of Mechatronics and Mechanical Engineering, Kielce University of Technology, Poland
[3]Faculty of Mechanical Engineering, University of Zilina, Slovakia
E-mail: a.goral@imim.pl

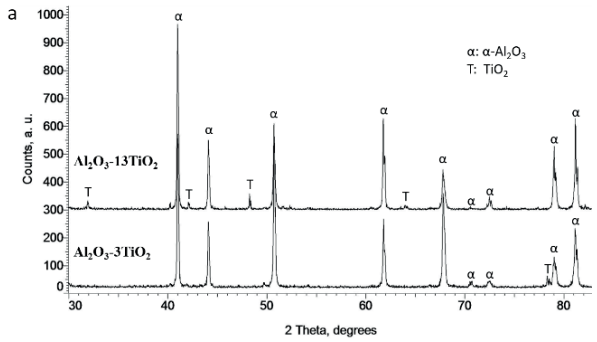


Figure 1 *X-ray diffraction patterns of the Al_2O_3-3TiO_2 and Al_2O_3-13TiO_2 powders*

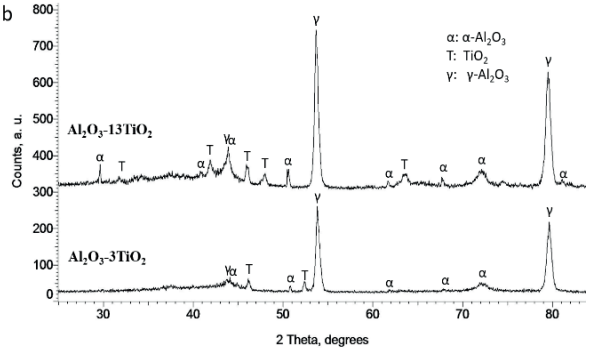


Figure 2 *X-ray diffraction patterns of plasma sprayed alumina - titania coatings*

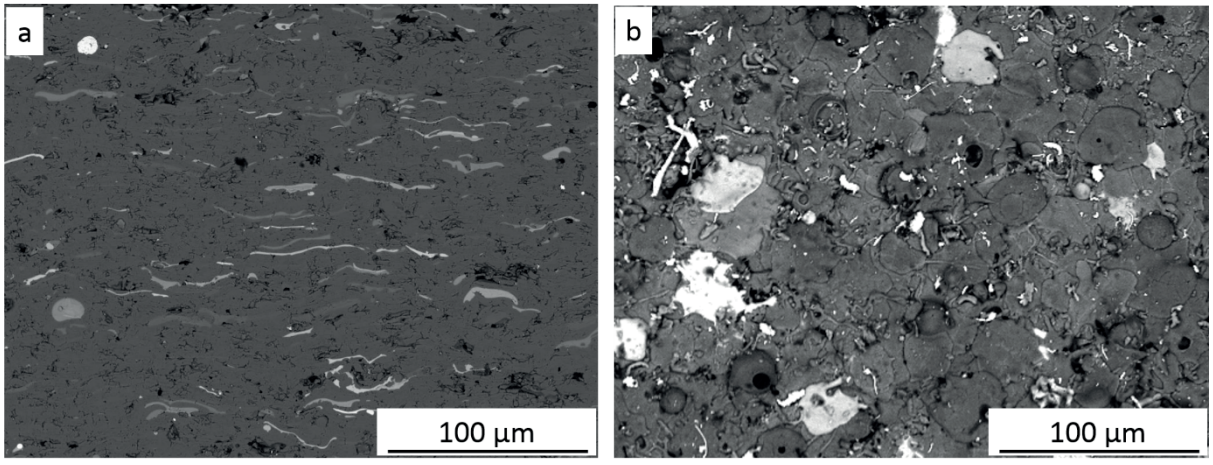


Figure 3 *Microstructure of Al_2O_3-3TiO_2 coating a) cross section, b) surface*

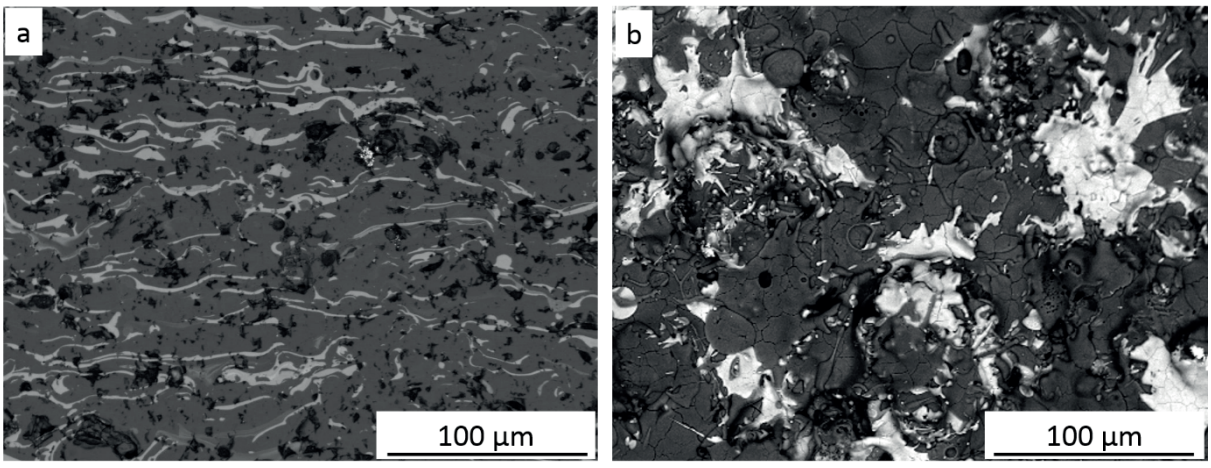


Figure 4 *Microstructure of the Al_2O_3-13TiO_2 coating a) cross section, b) surface*

c = 12.99 Å) and TiO_2 (tetragonal, a = 4.59 Å, c = 2.96 Å) phases identified using the X-ray diffraction method, Figure 1. The powders had irregular shapes and sizes in the range of 10-100 µm. The plasma sprayed coatings revealed the above mention phases but they were constituted mainly of γ-Al_2O_3 (cubic, a = 7.98 Å), Figure 2. The γ-Al_2O_3 phase is produced during the spray process from the α-Al_2O_3 present in the feedstock powder due to the high flattening cooling rate (between 10^{-6} and 10^{-8} K/s). Because of a higher cooling rate γ-Al_2O_3 nucleates in preference to α-Al_2O_3 phase [1], [10]. Additionally, the slightly raised background on the coating diffractograms in Figure 2 indicated a presence of an amorphous phase in their microstructures. That was confirmed by the TEM analysis presented in our previous paper [11].

During the plasma spraying the coatings were constituted as a result of an impaction and rapid solidification of the high-temperature and high-speed feedstock on the substrate or on the previously deposited layers. Such particles flattened and solidified rapidly at the surface of the part to be covered, forming a stacking

Table 1 *Hardness (HV and H) and Young's modulus measured for $Al2O_3$-3TiO_2 and Al_2O_3-13TiO_2 coatings under load of 100 mN*

Coating	HV	H [GPa]	E [GPa]	H/E
Al_2O_3-3TiO_2	683 ± 99	7.2 ± 1.0	102 ± 15	0.07
Al_2O_3-13TiO_2	980 ± 186	10.3 ± 1.9	124 ± 14	0.08

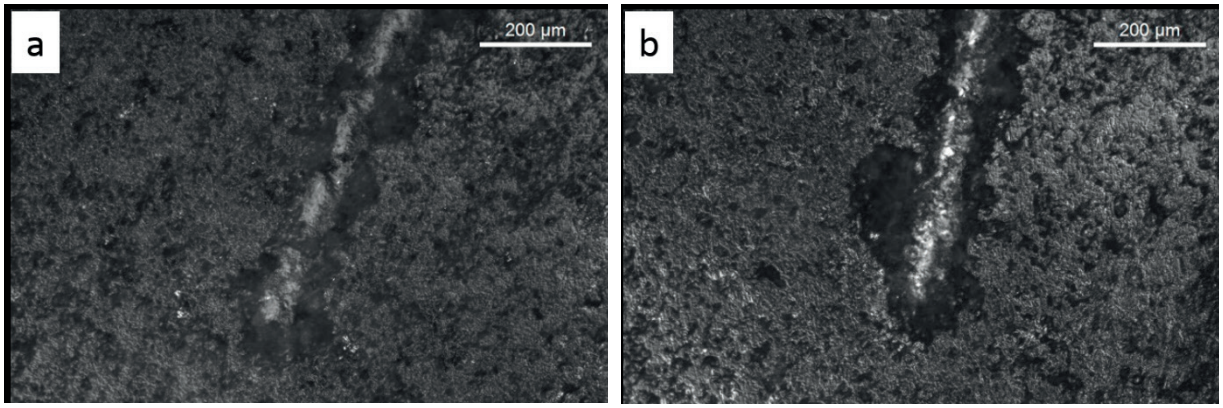


Figure 5 *Microstructure of wear tracks after the scratch test of coatings a) Al_2O_3-3TiO_2, b) Al_2O_3-13TiO_2*

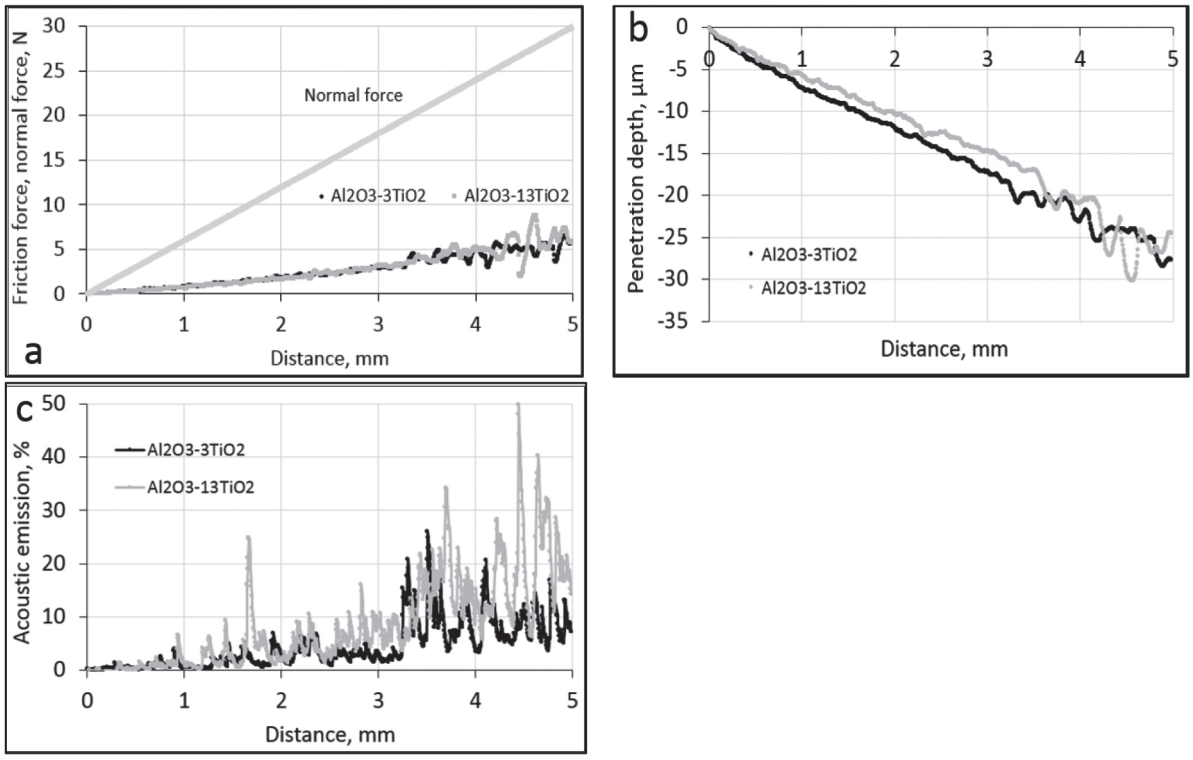


Figure 6 *Properties measured for Al_2O_3-3TiO_2 and Al_2O_3-13TiO_2 coatings during the scratch tests a) friction force and normal force, b) penetration depth, c) acoustic emission*

of lamellas. The layered microstructure of the Al_2O_3-3TiO_2 and Al_2O_3-13TiO_2 coatings constituted by the fully melted in the plasma flame initial powders, was composed of the Al_2O_3 matrix with areas composed of TiO_2. The strips with white contrast, visible on the coating cross sections in Figure 3 and Figure 4, correspond to TiO_2 phase. The phase appearing in grey contrast was mainly γ-Al_2O_3, although the α-Al_2O_3 phase was also present as was indicated in Figure 2. The α-Al_2O_3 phase was formed in the areas of coatings where the particles cooled more slowly. The detailed analysis of chemical and phase composition changes of the coatings was reported by Goral et al [11] and Zorawski et al [12]. The coatings showed a porosity content due to the voids left by the staking process and it was slightly lower for Al_2O_3-3TiO_2 coatings. Pores appear black in Figure 3 and Figure 4.

The indentation tests showed that the Al_2O_3-13TiO_2 coatings revealed for 43% higher microhardness than coatings containing less amount of TiO_2 in the feedstock. Those coatings were also characterized by 21% higher Young's modulus. Moreover, the ratio between the microhardness (H) and the elastic modulus (E) was calculated, because, according to Leyland et al. [13], the

most durable coatings can be obtained if H/E has a high value, i.e. sufficiently high hardness, but with a low elastic modulus. Results presented in Table 1 show that the slightly better wear resistance should be exhibited by the Al_2O_3-13TiO_2 coating.

To obtain the critical loads that are related to adhesion properties of coatings, the scratch tests were performed. The failure events were examined by an optical microscope. Only the cohesive cracks, resulting from over-strength of the coating, were observed as a result of the high tensile stresses behind the indenter, Figure 5.

The load at which such cracks occurred (L_{c1} - the first critical load) was 12 N and 16 N for Al_2O_3-3TiO_2 and Al_2O_3-13TiO_2 coatings, respectively. A fracture resistance was determined as the value of normal force when the first fracture occurs. This event was followed by a fluctuation in both force and penetration depth visible in Figure 6a-b. The corresponding acoustic emission (AE) was also measured during the test and presented in Figure 6c. An increase of the load to 23 N results in a significant peeling of the coating. Since no adhesion fracture was observed up to the maximum force of 30 N, it was not possible to determine the second critical load (L_{C2}) parameter. This was also caused by the high thickness of the coatings. The friction coefficient determined in this test was the same for both coatings and it was equal 0.21.

4. Conclusions

The plasma sprayed Al_2O_3-3TiO_2 and Al_2O_3-13TiO_2 coatings revealed layered structure with the titania strips uniformly distributed in the alumina matrix. The Al_2O_3-13TiO_2 coatings were characterized by enhanced mechanical properties, such as hardness, Young's modulus and fracture resistance with respect to the Al_2O_3-3TiO_2 coatings. The friction coefficient determined in the scratch test was the same for both coatings.

References

[1] VARGAS, F., AGEORGES, H., FOURNIER, P., FAUCHAIS, P., LOPEZ, M. E.: Mechanical and Tribological Performance of Al2O3-TiO2 Coatings Elaborated by Flame and Plasma Spraying. Surface and Coatings Technology, 205, 1132-1136, 2010. http://dx.doi.org/10.1016/j.surfcoat.2010.07.061

[2] SINGH, H., GREWA, M. S., SEKHON, H. S., RAO, R. G.: Sliding Wear Performance of High Velocity Oxy-Fuel Spray Al2O3/TiO2 and Cr2O3 Coatings. Journal of Engineering Tribology, 222, 601-610, 2008. https://doi.org/10.1243/13506501JET362

[3] UTU, I. D., MARGINEAN, G., HULKA, I., SERBAN, V. A., CRISTEA, D.: Properties of the Thermally Sprayed Al2O3-TiO2 Coatings Deposited on Titanium Substrate. International Journal of Refractory Metals and Hard Materials, 51, 118-123, 2015. https://doi.org/10.1016/j.ijrmhm.2015.03.009

[4] NOVY, F., MINTACH, R., BOKUVKA, O.: Effect of Galvanic Metallization, PVD Processes and Thermal Spraying on Fatigue Properties of Plain Structural Steels in the Ultra-High-Cycle Regime. Communications-Scientific Letters of the University of Zilina, 12(4), 49-54, 2010.

[5] SHAW, L. L., GOBERMAN, D., REN, R., GELL, M., JIANG, G., WANG, Y., XIAO, T. D., STRUTT, P. R.: The dependency of microstructure and properties of nanostructured coatings on plasma spray conditions. Surface and Coatings Technology, 130, 1-8, 2000. https://doi.org/10.1016/S0257-8972(00)00673-3

[6] WANG, M. D., SHAW, L. L.: Effects of the Powder Manufacturing Method on Microstructure and Wear Performance of Plasma Sprayed Alumina-Titania Coatings. Surface and Coatings Technology, 202(1), 34-44, 2007. https://doi.org/10.1016/j.surfcoat.2007.04.057

[7] HABIB, K. A., SAURA, J. J., FERRER, C., DAMRA, M. S., GIMENEZ, E., CABEDO, L.: Comparison of Flame Sprayed Al2O3/TiO2 Coatings: Their Microstructure, Mechanical Properties and Tribology Behavior. Surface and Coatings Technology, 201, 1436-1443, 2006. https://doi.org/10.1016/j.surfcoat.2006.02.011

[8] GHAZALI, M. J., FORGHANI, S. M., HASSANUDDIN, N., MUCHTAR, A., DAUD, A. R.: Comparative Wear Study of Plasma Sprayed TiO2 and Al2O3-TiO2 on Mild Steels. Tribology International, 93, 681-686, 2016. https://doi.org/10.1016/j.triboint.2015.05.001

[9] KUSOGLU, I. M., CELIK, E., CETINEL, H., OZDEMIR, I., DEMIRKURT, O., ONEL, K.: Wear Behavior of Flame-Sprayed Al2O3-TiO2 Coatings on Plain Carbon Steel Substrates. Surface and Coatings Technology, 200, 1173-1177, 2005. https://doi.org/10.1016/j.surfcoat.2005.02.219

[10] WANG, D., TIAN, Z., SHEN, L., LIU, Z., HUANG, Y.: Microstructural Characteristics and Formation Mechanism of Al2O3-13 wt. % TiO2 Coatings Plasma-Sprayed with Nanostructured Agglomerated Powders. Surface and Coatings Technology, 203, 1298-1303, 2009. https://doi.org/10.1016/j.surfcoat.2008.10.027

[11] GORAL, A., ZORAWSKI, W., LITYNSKA-DOBRZYNSKA, L.: Study of the Microstructure of Plasma Sprayed Coatings Obtained from Al2O3-13TiO2 Nanostructured and Conventional Powders. Materials Characterization, 96, 234-240, 2014. https://doi.org/10.1016/j.matchar.2014.08.016

[12] ZORAWSKI, W., GORAL, A., BOKUVKA, O., LITYNSKA-DOBRZYNSKA, L., BERENT, K.: Microstructure and Tribological Properties of Nanostructured and Conventional Plasma Sprayed Alumina-Titania Coatings. Surface and Coatings Technology, 268, 190-197, 2015. https://doi.org/1016/j.surfcoat.2014.09.014

[13] LEYLAND, A., MATTHEWS, A.: On the Significance of the H/E Ratio in Wear Control: A Nanocomposite Coating Approach to Optimized Tribological Behavior. Wear, 246, 1-11, 2000. https://doi.org/10.1016/S0043-1648(00)00488-9

Petr Jonsta - Irena Vlckova - Zdenek Jonsta - Vladimir Tomasek - Tatana Fenclova*

RESISTANCE OF HIGH-STRENGTH STEELS TO STRESS CORROSION CRACKING DEPENDING ON EXTERNAL ENVIRONMENT pH FACTOR

The paper deals with the study of stress corrosion cracking of high-strength steels in an aqueous environment with a varying pH factor ranging from 5.5 to 12.0. Steels were studied after quenching and tempering, one of the steels was prone to temper embrittlement.

Single-edge notched pre-cracked specimens were used for the experiments. Changes in the pH factor at the crack tip were measured using an antimony electrode. The pH factor values at the crack tip dropped to 2.0. Steel prone to temper embrittlement showed significantly shorter incubation period and more accelerated development of corrosion process compared to the optimized heat treatment of the second steel. Proneness to intergranular fracture was observed close to the fatigue crack tip. The obtained results expand the existing knowledge about localized corrosion processes leading to the refinement of the stress corrosion cracking model when changing the pH factor on the crack tip.

Keywords: *high-strength steels, heat treatment, stress corrosion cracking, pH factor, intergranular fracture*

1. Introduction

Steel structures are often multi-axially stressed and may be damaged by a different degradation mechanisms which cause to a loss of stability, excessive corrosion, corrosion fatigue etc. High-strength steels are materials with a higher yield and tensile strength which allows to use a structures with a lower weight. Due to that fact, the high-strength steels are used in many transportations applications. High stability of mechanical properties at high strain rate is used in automotive industry (crash-testing of vehicle). Mining and military industry offers the possibility of further applications, e.g. loaders, excavators, armored transporters etc.

Stress corrosion cracking (SCC) is one of the predominant failures of high-strength steels. The SCC is characterized by initiation and stable crack growth rate due to the simultaneous effect of tensile stress and long-term exploitation in aqueous environment. Initiation and stable crack growth rate, due to the SCC, occur when environmental characteristics, stress state and material properties simultaneously reach certain critical levels [1], [2]. The stable crack growth rate is characterized by the presence of the stress intensity factor threshold value ($K_{ISCC)}$. Shipilov [3] argued that the SCC is a special case of the corrosion fatigue.

The SCC is affected by a number of factors that can be divided into three groups. The stress mode, especially the stress state at critical points of the steel construction (static or cyclic tensile stress, residual stress), strain rate and frequency of cycling are included into the first group.

The second group involves properties of aqueous environment and temperature regime and the third group involves metallurgical factors (chemical composition, microstructure), the yield stress and heat treatment.

The SCC is related to electrochemical reactions, especially to changes in the *pH* factor at the tip of the growing crack, as Hirth states [4]. Two types of the SCC, which are often differentiated as the high *pH* or intergranular SCC (IGSCC) and the near-neutral *pH* or transgranular SCC (TGSCC) have been identified in works [5], [6], [7], [8].

The mechanism of the IGSCC is attributed to anodic dissolution resulting from selective dissolution at the grain boundaries and repeated rupture of passivate films that form over the crack tip [9].

The increase of deformation at the crack tip needed to rupture the oxidation layer may be caused by the monotonically increasing load, respectively by the creep at the constant load. When the oxidation layer is ruptured, the crack tip is exposed. Subsequently, the crack is affected by an aqueous environment with aggression depending on concentration of anions at the crack tip.

Aggression of environment at the crack tip is increasing by dissolution of the MnS inclusions and/or by gradient of potential between mouth and tip of the crack. Gradient of potential depends on oxygen concentration, which is dissoluted in water. A crack tip grows as a result of anodic dissolution. Simultaneously, the crack tip is progressively covered by the increasing passivation layer and the crack growth slows down.

In case of the TGSCC, great efforts have been made to understand the fundamental role of hydrogen and its synergism with stress in a few last years, as presented in works [10], [11]. The stable crack growth rate is controlled by successive partial processes, which included transport of corrosive environment

* [1]Petr Jonsta, [2]Irena Vlckova, [1]Zdenek Jonsta, [1]Vladimir Tomasek, [1]Tatana Fenclova
[1]Faculty of Metallurgy and Materials Engineering, VSB-TU Ostrava, Czech Republic
[2]RMTSC, Material & Metallurgical Research Ltd., Remote Site Ostrava, VUHZ a.s., Dobra, Czech Republic
E-mail: petr.jonsta@vsb.cz

Table 1 *The chemical composition of the steels (wt. %)*

Steel	C	Mn	Si	Cr	Ni	Mo	P	S
Mn-Si-Mo	0.30	1.42	1.38	0.22	0.36	0.23	0.013	0.009
30Cr2Ni2Mo	0.32	0.55	0.32	1.95	1.82	0.32	0.012	0.010

Table 2 *Mechanical properties of the steels*

Steel	Yield Strength 0.2 % offset (MPa)	Tensile Strength (MPa)	Elongation in 50 mm (%)	Reduction of Area (%)	Notch Toughness ($J \cdot cm^{-2}$)
Mn-Si-Mo	1490	1720	12	56	46
30Cr2Ni2Mo	1600	1762	12	52	27

into the crack tip, electrochemical reactions of the corrosive environment with newly created surfaces causing release of hydrogen, absorption of hydrogen at the crack tip, transport of hydrogen to areas where it is causing the embrittlement, the hydrogen-metal interaction that leads to embrittlement (reduced cohesive strength of grain boundaries, affecting mobility of dislocations, etc.). The resulting crack growth rate is then controlled by the slowest of the above processes.

The evaluation of structural metallurgical characteristics of the SCC has shown that in high-strength steels with a structure of martensite tempered at low temperatures there is a preferential occurrence of intergranular fracture when exposed to an aqueous environment [1], [4]. It is debatable whether this occurrence can be influenced by the heat treatment in a mentioned fracture mechanism, in other words, how the microstructure of low-tempered martensite contributes to the fracture development during the SCC in an aqueous solution.

The presented paper is focused on the SCC of two high-strength steels used in nuclear and classical power engineering. Steels were studied after quenching and tempering, one of the steels was prone to temper embrittlement. The aim of the work was to study the influence corrosion environment near the crack tip, in order to strengthen the knowledge about the mechanisms of the SCC processes of high-strength steels. In this context, attention was primarily focused on the study of changes in the *pH* factor of the solution near the crack tip during the, corrosion tests with a different *pH* baseline.

2. Materials and experimental procedure

Thermo-mechanically treated 14 mm thick sheets made of Mn-Si-Mo and 30Cr2Ni2Mo high-strength steels were used. Final rolled sheets were quenched and tempered by regime:

- 850 °C/water + 260 °C/4 h/air (Mn-Si-Mo steel)
- 880 °C/water + 250 °C/4 h/air (30Cr2Ni2Mo steel)

The chemical composition of the steels is given in Table 1.

Structural analysis was performed by optical metallography and scanning electron microscopy (SEM). The tensile properties were determined using an MTS 100 kN machine on flat specimens, which were taken from the mid-thickness of the materials in the longitudinal direction. The notch toughness was measured on the standard Charpy impact specimens at ambient temperature. The mechanical properties of the steels are summarized in Table 2.

The SCC susceptibility measurements were determined using the creep machines on single-edge notched pre-cracked specimens which were a 10 mm thick and 25 mm width. An angle of notch was α = 45°. A fatigue crack of approx. 2 mm in length was initiated at the root of the notch. Crack opening was determined by an inductive sensor and automatically registered. The specimen was placed in a specific plastic and organic glass cell.

The SCC tests were performed in distilled water and in an aqueous solution of 3.5 % NaCl solution. Tests were carried out at *pH* = 6 baseline value and after addition of 1 M of NaOH solution at *pH* = 12. Antimony electrodes of the authors' own design, with a wide range of applications, were used to measure the *pH* factor in the crack. For each measurement, 5 pairs of electrodes were used. The function of the electrodes has been verified by long-term measurements together with commercially saturated calomel electrodes. Both electrodes were placed in the holder together due to ensuring a constant distance during measurement a calibration.

In the SCC tests, the ratio of Fe^{2+} / Fe^{3+} concentrations, at different crack and notch points, was also measured by measuring the oxidation-reduction potential. In this case, the antimony electrodes were replaced by platinum.

3. Results and discussion

The microstructure of the Mn-Si-Mo steel after the heat treatment was formed by a mixture of lath and partially plate martensite, containing approx. 3 % retained austenite. The microstructure is presented in Figure 1.

The 30Cr2Ni2Mo steel, after the heat treatment, was also characterized by a microstructure of lath and plate martensite with a retained austenite content of 5-6 %. Because of the low Si content, tempering of the martensitic structure caused the low-temperature temper embrittlement. The embrittlement is characterized by the preferential etching of the original austenitic grains after etching in aqueous picric acid solution and surfactant, as shown in Figure 2.

The overall results of the *pH* factor changes in the Mn-Si-Mo steel, at their different baselines, are presented in Table 3. For further details on ongoing changes in the *pH* factor baseline, this table also provides data on *pH* factor values in the solution volume.

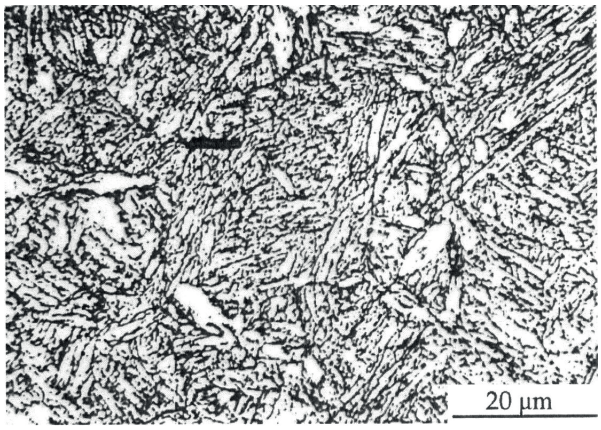


Figure 1 *Microstructure of the Mn-Si-Mo steel (mixture of lath and partially plate martensite, containing approx. 3% retained austenite)*

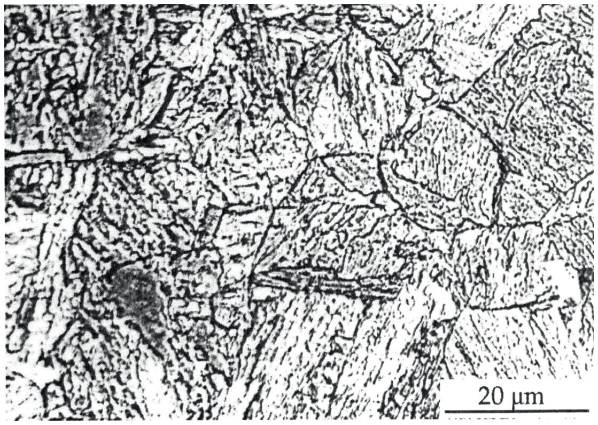


Figure 2 *Microstructure of the 30Cr2Ni2Mo steel (lath and plate martensite with a retained austenite content of 5-6%)*

Table 3 *pH values of stress corrosion cracking test*

Corrosion solution	Volume		Crack tip		Notch mouth	
	pH_{start}	pH_{finish}	pH_{min}	t (min)	pH_{max}	t (min)
3.5% NaCl	6.54	6.32	2.7	20	12.0	400
neutral	6.07	6.37	2.9	20	9.5	220-350
	6.40	7.43	1.5	-	9.5	-
3.5% NaCl	12.07	11.17	3.6	1500-1800	13.5	320
alkaline	12.16	11.20	1.9	350-650	13.1	200
	11.40	-	1.5	0-50	-	-
distilled water	5.50	6.05	2.8	1600-2200	7.5	200
alkaline water	12.15	13.13	1.2	3300	12.7	-

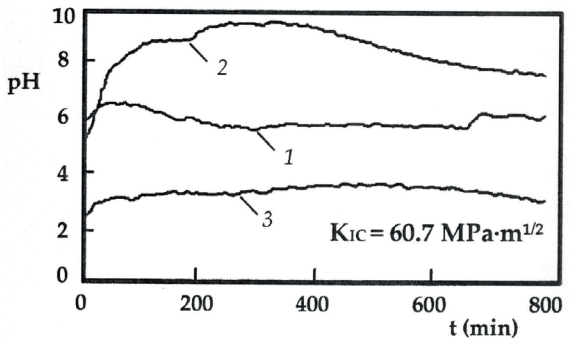


Figure 3 *Time flow of measured values of the pH factor for baseline pH = 6*

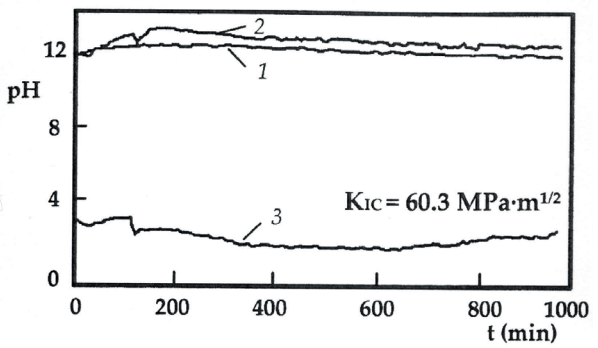


Figure 4 *Time flow of measured values of the pH factor for baseline pH = 12*

The timelines of measured *pH* values during the corrosion cracking tests in 3.5% aqueous, neutral and alkalized NaCl solutions, are presented in Figure 3 and Figure 4 for both baselines *pH* = 6 and *pH* = 12.

Curve 1 represents the measured values of the *pH* factor changes determined at the top of the notch, curve 2 at the notch mouth and the curve 3 in close proximity to the crack tip. Differences in minimum measured *pH* values are due to the different positions of the specific antimony electrode in relation to the crack tip. This fact is based on the experience that the lowest values were found in cases where the crack tip progressed through the location of the scanning position of the antimony electrode, as found by Sojka [12] and Sung et al. [13].

Based on presented results is evident that there is a significant acidification of the solution at the crack tip and the development of atomic hydrogen. This leads in the Mn-Si-Mo steel, optimally heat-treated in terms of the occurrence of irreversible temper embrittlement, to potential embrittlement of the matrix [14].

When the crack is filled with corrosive solution, the iron starts dissolving at its tip. The Fe^{2+} ions are, in the presence of oxygen, oxidized to Fe^{3+}. In a small crack volume, an increase in the concentration of metal ions occurs, as well as a decrease in the solution *pH* factor during their hydrolysis. An even more

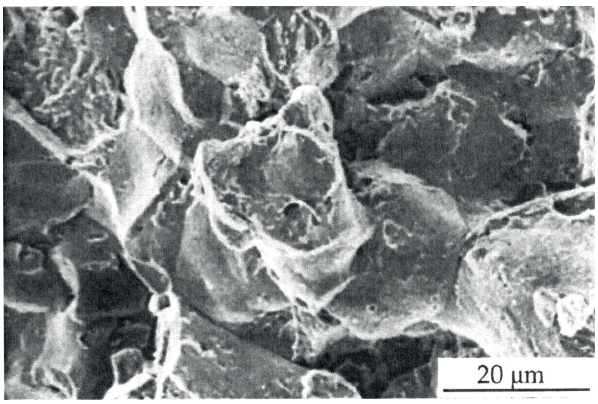


Figure 5 *Fracture area of the Mn-Si-Mo steel near the crack tip (intergranular fracture)*

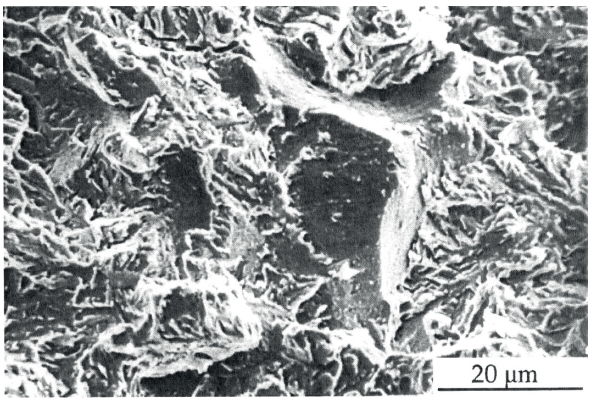


Figure 6 *Fracture area of the Mn-Si-Mo steel 2 mm in front of the crack tip (intergranular fracture with quasi-cleavage areas)*

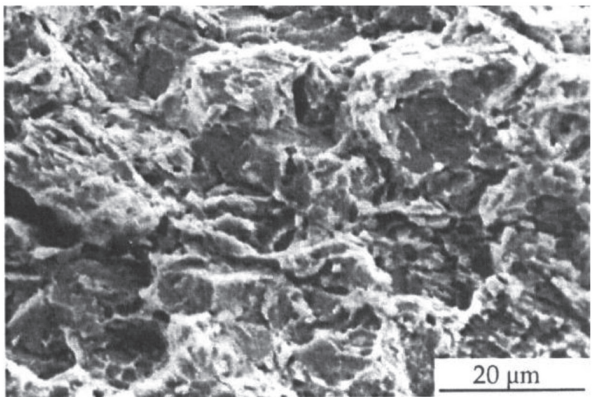


Figure 7 *Final fracture area of the Mn-Si-Mo steel (quasi-cleavage character)*

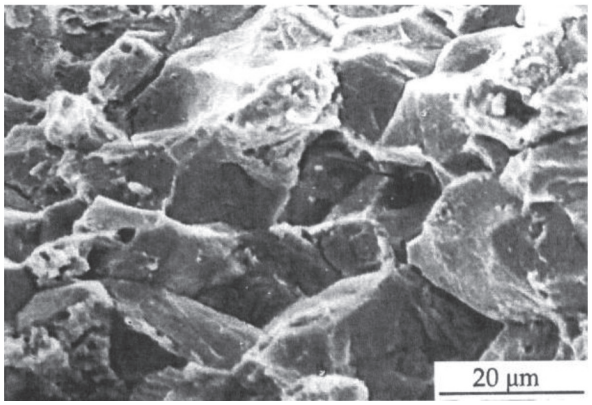


Figure 8 *Fracture area of the 30Cr2Ni2Mo steel near the crack tip (intergranular fracture)*

significant decrease in the *pH* factor can be observed probably during the adsorption of OH-ions on freshly exposed iron surfaces.

The electrons released by iron oxidation reduce H^+ ions to atomic hydrogen, which is adsorbed on the surface of the steel and preferably diffuses into the matrix along the grain boundaries of the initial austenite. When the critical concentration of hydrogen in the matrix is reached in front of the crack tip, intergranular decohesion occurs as evident from Figure 5. The intergranular fracture in corrosion cracking tests in aqueous solutions was located in a relatively narrow area just near the front of the fatigue crack.

With the crack development there is an increase in the stress intensity factor at its tip and the acceleration of its propagation. Consequently, in addition to intercrystalline areas, it is possible to observe also quasi-cleavage fracture areas at the fracture surface, see Figure 6, because the material in front of the crack tip cannot be sufficiently saturated by hydrogen [12].

The further development of the crack leads to an acceleration of its growth, which subsequently leads to a decrease in the effect of hydrogen and, in the event of a final fracture, the occurrence of an intergranular fracture is eliminated. Therefore, the fracture areas are predominantly of a quasi-cleavage character, as shown in Figure 7.

Based on the findings, it can be stated that for the 30Cr2Ni2Mo steel, the minor content of adsorbed hydrogen,

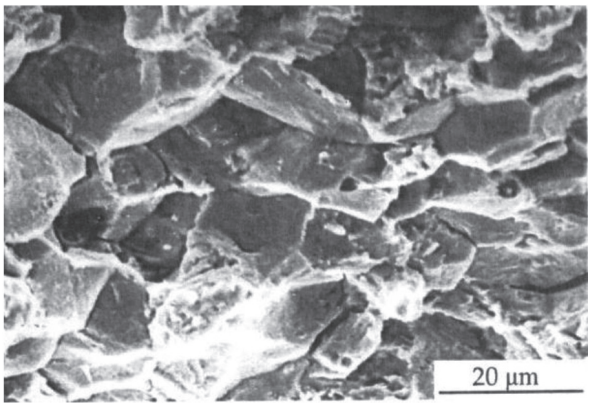


Figure 9 *Fracture area of the 30Cr2Ni2Mo steel near the crack tip (intergranular fracture)*

diffusing into the matrix to trigger an intergranular cleavage fracture, is sufficient due to the development of the low-temperature temper embrittlement and the associated "a priori" intergranular decohesion (see Figure 8 and Figure 9), and namely into a significant distance from the fatigue crack tip. It can be stated that the development of temper embrittlement is a significant degradation characteristics of the high-strength steels with the tempered martensite structure.

4. Conclusions

The presented research was focused on the analysis of conditions for development of the stress corrosion cracking of high-strength steels in an aqueous environment with different initial *pH* factor values. The Mn-Si-Mo and 30Cr2Ni2Mo steels were tested. Data on modification of electrochemical characteristics on the crack tip and in its volume, were obtained following two variants of the heat treatment. The obtained results expand the existing knowledge about localized corrosion processes leading to the refinement of the stress corrosion cracking model when changing the *pH* factor at the crack tip.

Comparison of susceptibility of the studied steels towards the development of the stress corrosion cracking showed that the susceptibility to this fracture mechanism depends on the heat treatment parameters. In the development of the low-temperature temper embrittlement, the incubation period was significantly shortened and development of the fracture was accelerated, compared to the state of optimized heat treatment where the susceptibility to intergranular fracture was located in close proximity of the fatigue crack tip.

Acknowledgement

This paper was created at the Faculty of Metallurgy and Materials Engineering, VSB - Technical university of Ostrava as part of the Project No. LO1203 "Regional Materials and Technology Centre-Feasibility Programme" funded by the Ministry Of Education, Youth and Sports of the Czech Republic and the Project SP 2017/58 "Specific Research in the Metallurgical, Materials and Process Engineering."

References

[1] LYNCH, S. P.: RAJA, V., SHOJI, T. (Eds.): Stress Corrosion Cracking: Theory and Practice. Woodhead Pub., Philadelphia, USA, 90-129, 2011.

[2] DJUKIC, M. B., SIJACKI ZERAVCIC, V., BAKIC, G., SEDMAK, A., RAJICIC, B.: Hydrogen Embrittlement of Low Carbon Structural Steel. Procedia Materials Science, 3, 1167-1172, 2014. https://doi.org/10.1016/j.mspro.2014.06.190

[3] SHIPILOV, S. A.: Mechanisms for corrosion fatigue crack propagation. Fatigue and Fracture of Engineering Materials and Structures, 25(3), 243-259, 2002. https://doi.org/10.1046/j.1460-2695.2002.00447.x

[4] HIRTH, J. P.: Effects of Hydrogen on the Properties of Iron and Steel. Metallurgical Transactions A, 11(6), 861-890, 1980. https://doi.org/10.1007/BF02654700

[5] ARAFIN, M. A., SZPUNAR, J. A.: A New Understanding of Intergranular Stress Corrosion Cracking Resistance of Pipeline Steel through Grain Boundary Character and Crystallographic Texture Studies. Corrosion Science, 51(1), 119-128, 2009. https://doi.org/10.1016/j.corsci.2008.10.006

[6] ZHANG, G. A., CHENG, Y. F.: Micro-Electrochemical Characterization of Corrosion of Pre-Cracked X70 Pipeline Steel in a Concentrated Carbonate/Bicarbonate Solution. Corrosion Science, 52(3), 960-968, 2010. https://doi.org/10.1016/j.corsci.2009.11.019

[7] PARKINS, R. N., BEAVERS, J. A.: Some Effects of Strain Rate on the Transgranular Stress Corrosion Cracking of Ferritic Steels in Dilute Near-Neutral-pH Solutions. Corrosions 59(3), 258-273, 2003. https://doi org/10.5006/1.3277559

[8] WANG, J. Q., ATTRENS, A.: Microstructure and Grain Boundary Microanalysis of X70 Pipeline Steel. Journal of Materials Science, 38(2), 323-330, 2003. https://doi.org/10.1023/A:1021169700779

[9] BUENO, A. H. S., MOREIRA, E. D., GOMES, J. A. C. P.: Evaluation of Stress Corrosion Cracking and Hydrogen Embrittlement in an API Grade Steel. Engineering Failure Analysis, 36, 423-431, 2014. https://doi.org/10.1016/j.engfailanal.2013.11.012

[10] LI, M. C., CHENG, Y. F.: Mechanistic Investigation of Hydrogen-Enhanced Anodic Dissolution of X-70 Pipe Steel and its Implication on Near-Neutral pH SCC of Pipelines. Electrochimica Acta, 52(28), 8111-8117, 2007. https://doi.org/10.1016/j.electacta.2007.07.015

[11] KANGA, Y., CHEN, W., KANIA, R., VAN BOVEN, G., WORTHINGHAM, R.: Simulation of Crack Growth during Hydrostatic Testing of Pipeline Steel in Near-Neutral pH Environment. Corrosion Science, 53(3), 968-975, 2011. https://doi.org/10.1016/j.corsci.2010.11.029

[12] SOJKA, J.: Resistance of Steels to Hydrogen Embrittlement. VSB-TU Ostrava, Faculty of Metallurgy and Materials Engineering, Ostrava, Czech Republic, 108, 2007.

[13] SUNG, J. K., HWAN, G. J., KYOO, Y. K.: Effect of Post-Weld Heat Treatment on Hydrogen-Assisted Cracking Behavior of High-Strength Process Pipe Steel in a Sour Environment. Scripta Materialia, 67, 895-898, 2012. https://doi.org/10.1016/j.scriptamat.2012.08.021

[14] LYNCH, S. P.: MOODY, N. R. (Ed.): Hydrogen Effects on Material Behavior and Corrosion Deformation Interactions. TMS, Warendale, PA, 449-466, 2003.

Daniel Kajanek - Branislav Hadzima - Filip Pastorek*

ELECTROCHEMICAL CHARACTERIZATION OF AZ31 MAGNESIUM ALLOY TREATED BY ULTRASONIC IMPACT PEENING (UIP)

Electrochemical characteristics of AZ31 magnesium alloy after impacting by ultrasonic impact peening (UIP) were measured in order to evaluate its effect on corrosion resistance of mentioned alloy. For this purpose, potentio-dynamic polarization (PD) tests and electrochemical impedance spectroscopy (EIS) of ground and impacted samples were performed in 0.1 M NaCl solution at the temperature of 22±2 °C after 5 minutes of potential stabilization. The obtained PD curves were analysed using the Tafel extrapolation method and data from EIS measurements in form of Nyquist plots were quantified by the equivalent circuit method. The results showed clearly that application of the UIP significantly deteriorated corrosion resistance of AZ31 alloy in the chosen testing solution.

Keywords: *magnesium alloy, corrosion, ultrasonic impact peening*

1. Introduction

Despite of all the well-known convenient properties of magnesium and its alloys, such as low weight, highest specific strength of all metal materials, good damping capacity etc., their applications are still not widely spread due to the insufficient corrosion performance [1], [2], [3], [4], [5]. Generally it is proposed that this derogation is caused by synergic effect of negative standard potential value (-2.36 V vs. SHE), impurities (Fe, Ni, Cu,...), which leads to a micro galvanic effect and non-protective quasi-passive film based on magnesium oxide or hydroxide depending on the nature of environment [6], [7], [8]. This surface film is stable only in highly alkaline solutions [9], [10]. Magnesium and its alloys exhibit strange behaviour during the anodic polarisation called the negative difference effect (or according to the latest researches it is better to call it ''anomalous hydrogen evolution''), which is represented by the hydrogen evolution (HE) during the anodic reaction [11], [12], [13]. Normally, hydrogen evolution is an accompanying feature of cathodic reaction during the corrosion of metals so there is an intense effort to explain this strange behaviour [14], [15], [16]. Nowadays, researches are focused on increasing of corrosion resistance of Mg alloys by preparation of resistant barrier e.g. coatings, plating or by modification of surface characteristics [17], [18], [19]. Recent works, aimed at application of ultrasonic impact peeing (UIP), propose that this technique has a positive effect on mechanical properties of various materials (steels, Co and Ti alloys and welded materials, etc.), increase the fatigue characteristics and also enhance their corrosion resistance. Due to those pros, the UIP has become a common technology in automotive and civil engineering. Its main principle is represented by deformation of surface by impacts with ultrasonic frequency, which induce favourable compressive residual stresses and decrease harmful tensile stresses in a surface layer. The more detailed description of this method could be found elsewhere [20], [21], [22], [23], [24].

To our best knowledge, the extensive research on influence of the UIP application on corrosion properties (and also on mechanical properties) of magnesium alloys has not been done yet. Therefore, the aim of this contribution is to compare electrochemical characteristics of AZ31 magnesium alloy before and after the UIP by means of potentio-dynamic polarization tests and electrochemical impedance spectroscopy in sodium chloride solution and thus evaluate its effect on corrosion resistance of this alloy.

2. Experimental material and methods

The AZ31 magnesium alloy has been prepared by continual casting and subsequently heat treated by homogenizing at temperature of 420 °C for 16 hours. Chemical composition of AZ31 Mg alloy (Table 1) has been obtained by the EDXRF analysis on ARL QUANT´X EDXRF Spectrometer. In order to evaluate the microstructure of the material, samples have been ground by an emery paper p1000, then polished by polishing discs MD Dac and MD Chem using proper polishing emulsions on Struers Tegramin-30 machine. Samples have been etched for 10 seconds by etchant composed of picric acid, acetic acid, ethanol and water [18], [25]. The microstructure has been observed using the ZEISS Axio Imager.Z1M light optical microscope.

The ultrasonic impact peening (UIP) has been performed using the CNC mill with mounted ultrasonic impact head (see Figure 1), equipped with a cooling system, which kept the constant temperature during impacting. The samples have been ground by an emery paper p1000 to obtain the uniform surface

* [1,2]Daniel Kajanek, [2]Branislav Hadzima, [2]Filip Pastorek
[1]Department of Materials Engineering, Faculty of Mechanical Engineering, University of Zilina, Slovakia
[2]Research Centre, University of Zilina, Slovakia
E-mail: daniel.kajanek@fstroj.uniza.sk

Table 1 *Chemical composition of AZ31 magnesium alloy*

Component	Al	Zn	Mn	Si	Cu	Fe	Mg
wt. %	3.31	0.65	0.23	0.114	0.004	0.001	balance

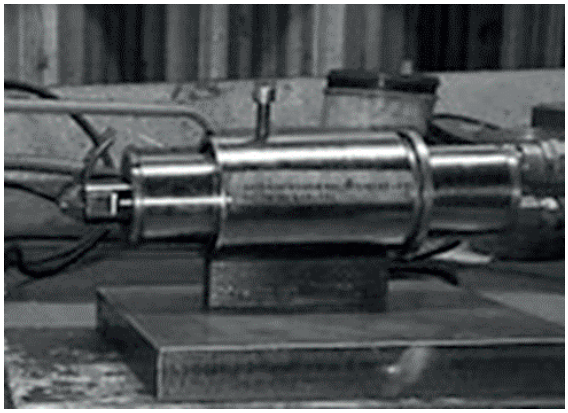


Figure 1 *Ultrasonic impact head*

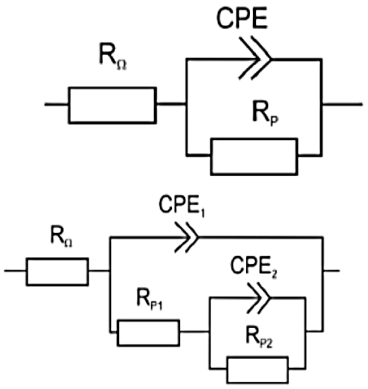


Figure 2 *Equivalent circuits for diagrams with one capacitance loop (up) and two capacitance loops (down)*

roughness before application of the UIP. The parameters of the UIP process have been set as follows: pressure force of 85 N, ultrasonic frequency of 20 kHz and the treatment time has been 5 minutes.

To observe influence of the UIP on surface of AZ31, photo documentation of ground and the UIP treated samples have been created by the stereomicroscope OLYMPUS SZX16 and supported by documentation of the cross-section view made by the ZEISS Axio Imager.Z1M light microscope with corresponding software.

With the aim to evaluate the corrosion resistance, electrochemical characteristics of ground samples and samples treated by the UIP have been measured by electrochemical impedance spectroscopy (EIS) and potentio-dynamic polarization (PD) tests. All of the experiments have been carried out in 0.1 M NaCl [26] at the temperature of 22 ± 2 °C on potentiostat VSP Biologic SAS, using the three electrode cell system with following configuration: AZ31 - working electrode, Pt electrode - counter electrode, saturated calomel electrode - reference electrode (+0.2446 V vs. SHE). The stabilization of potential between the working electrode and electrolyte has been set on 5 minutes for both EIS measurements and PD tests. The software EC Lab V10.34 has been used for analysis of data obtained by all the mentioned electrochemical techniques.

During the EIS measurements, amplitude of applied voltage has been set to 15 mV and the frequency ranged from 100 kHz to 10 mHz with change of 10 points per decade. The average value of voltage has been set on open circuit potential value [27]. The measured curves have been plotted in the shape of Nyquist diagrams. Those diagrams have been quantitatively described by the equivalent circuit method. The equivalent circuits used for the analysis are shown in Figure 2. Circuit represented in Figure 2 (left) has been used for the Nyquist plots with one capacitance loop and describes situation with homogenous conditions on the surface. Figure 2 (right) shows the equivalent circuit used for analysis of Nyquist diagram with two capacitance loops and interprets occurrence of locations with different electrochemical behaviour. For the correct interpretation of the Nyquist diagrams it is important to know the meaning of the circuit components. Solution resistance is represented by the R_Ω component and the CPE is a constant phase element, which replaces capacitor in the circuit and describes the heterogeneity of electrode´s surface. The component R_p is called the polarization resistance of sample/electrolyte interface. If the Nyquist plot is composed of two capacitance loops, final polarization resistance will be given by the sum of partial polarization resistances R_{p1} and R_{p2} [28]. Their interpretation depends on the particular system and will be explained in the discussion part.

Potential, ranged from -200 mV to +500 mV, has been applied on the AZ31 samples with the rate of 1 $mV.s^{-1}$ during the PD tests. The range of potentials have been set with respect to open circuit potential [18]. The measured data of PD tests are represented by potentiodynamic polarization curves which have been analysed by the Tafel extrapolation method in order to obtain electrochemical characteristics e.g. corrosion potential E_{corr}, corrosion current density i_{corr} and calculated corrosion rate r_{corr}.

3. Results and discussion

The AZ31 alloy´s microstructure (Figure 3) consists of polyhedral grains of solid solution of aluminium, zinc and other alloying elements in magnesium. Intermetallic phase $Mg_{17}Al_{12}$ is localised along the grain boundaries. There are also areas of twins caused by the deformation during the cutting and preparation of samples.

The difference between the ground surface and surface after the impact treatment can be observed from the documentation shown in Figure 4. Application of ultrasonic impacts has caused deformation of the surface and subsequently increased reactivity, which has led to the accelerated formation of film on the sample´s surface. According to nature of the UIP process and surrounding environment, it can be assumed that the main constituent of the film is MgO with small amount of $Mg(OH)_2$ as the environment has certain amount of moisture [9], [10]. It has to be kept in mind that constituents from the cooling system could be present,

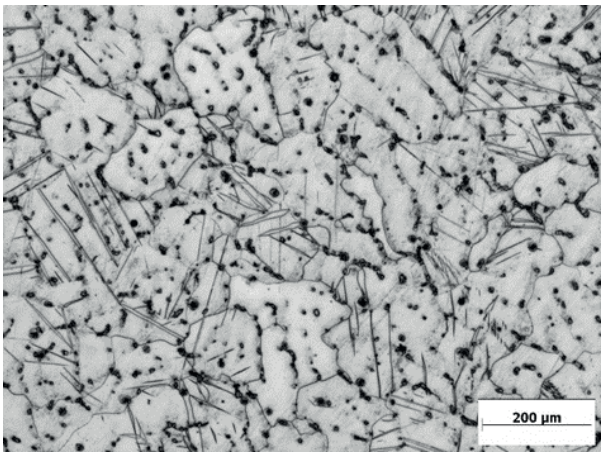


Figure 3 *Microstructure of AZ31 Mg alloy*

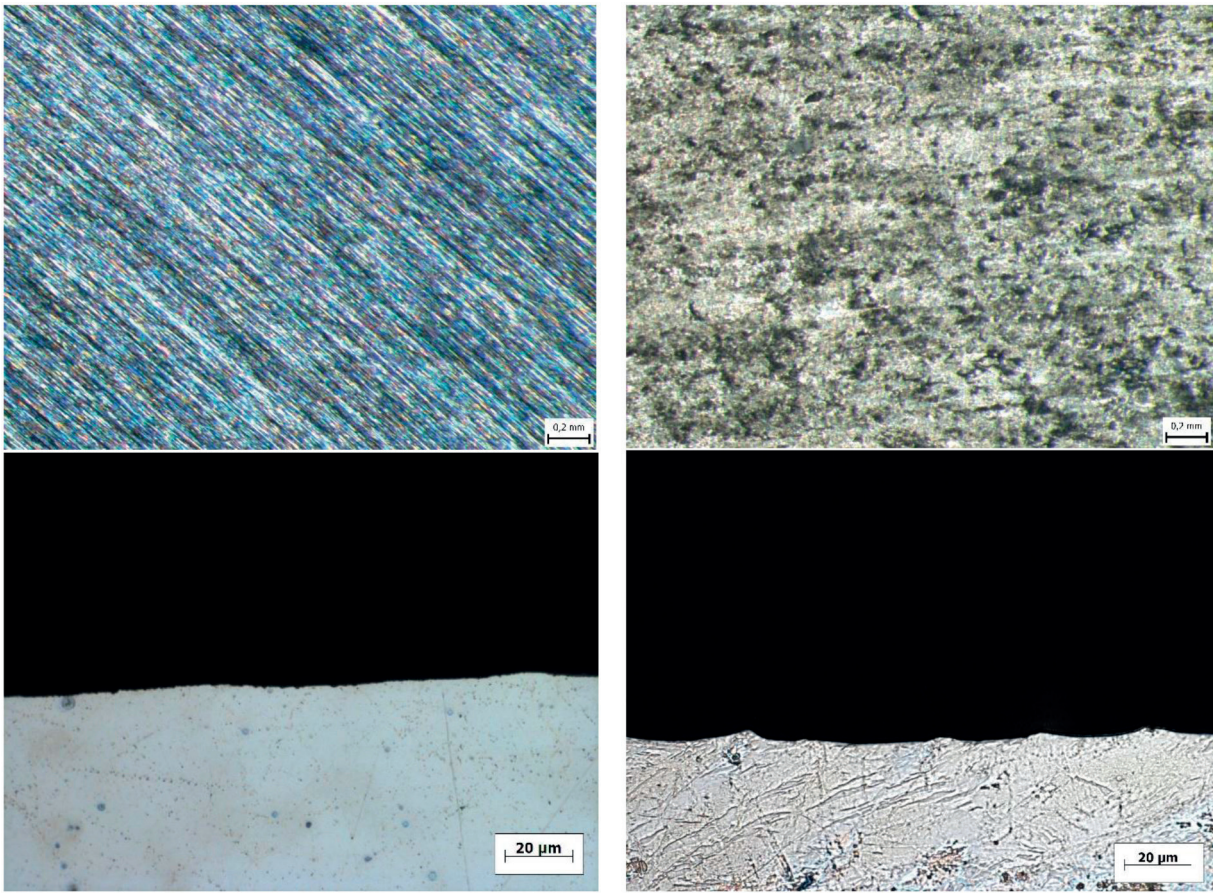


Figure 4 *Surfaces of ground (upper left) and UIP influenced samples (upper right) with cross sectional view of ground (left bottom) and UIP surface (right bottom)*

as well. According to literature, the thickness of such a film is in nanoscale range and thus it is difficult to observe it by the light microscopy [29]. The exact chemical composition of the film is going to be examined by XPS analysis in further research studies. As can be seen from the UIP influenced surface and from the cross- sectional views, several defects and holes are introduced to the surface. This kind of defects increases electrochemically active area, which is in contact with environment and could negatively affect corrosion performance of Mg parts [30].

The Nyquist diagrams as results of the EIS measurements in 0.1 M NaCl, for ground samples and the UIP influenced samples, are displayed in Figure 5. Values of electrochemical characteristics obtained by equivalent circuit analysis are listed in Table 2. As can be seen, diagram of ground samples consists of two capacitance loops and diagram of the UIP samples is created by one capacitance loop, which pointed to different electrochemical behaviour. In the case of ground samples, high capacitance loop represented by value of R_{p1} describes resistance of the corrosion product barrier ($Mg(OH)_2$ and $MgCl_2$ considering testing electrolyte) created on the surface while exposed to the solution. The low capacitance loop represented by value R_{p2}, is analysed as charge transfer resistance of the surface/electrolyte

Table 2 *Electrochemical characteristics obtained by EIS measurements in 0.1M NaCl*

	R_Ω ($\Omega.cm^2$)	R_{p1} ($\Omega.cm^2$)	R_{p2} ($\Omega.cm^2$)	R_p ($\Omega.cm^2$)
Ground AZ31	124±6	1270±23	919±14	2189±37
UIP AZ31	135±9	68±7	-	68±7

Table 3 *Electrochemical characteristics obtained by PD tests of in 0.1M NaCl*

	E_{corr} (mV)	i_{corr} ($\mu A/cm^2$)	β_a (mV/dec.)	β_c (mV/dec.)	r_{corr} (mm/year)
Ground AZ31	-1508±20	18±1.1	76±6	131±9	0.8±0.03
UIP AZ31	-1433±17	98±6.4	193±11	205±21	4.5±0.7

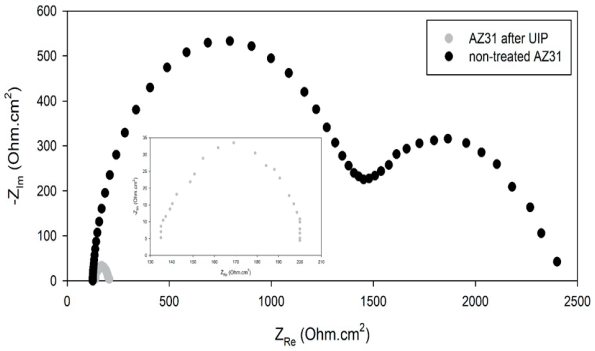


Figure 5 *Nyquist plots of ground and UIP influenced samples measured in 0.1M NaCl*

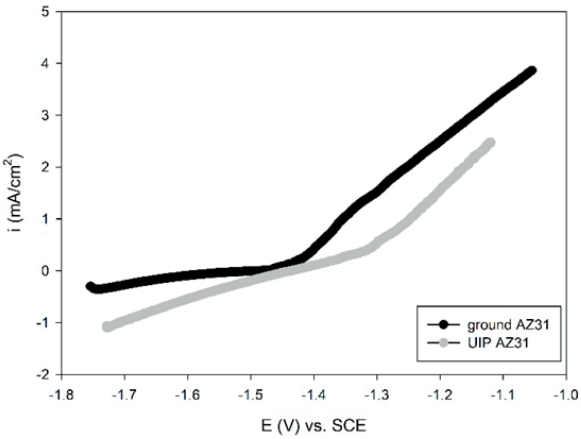


Figure 6 *Potentiodynamic polarization curves of ground and UIP influenced samples measured in 0.1M NaCl*

interface. Diagram of UIP samples consist of one capacitance loop. It indicates homogenous conditions on the surface during the exposition. The polarization resistance of these samples is represented by charge transfer resistance of sample/electrolyte interface. According to shape of diagrams for both states it can mentioned that degradation mechanism is under charge transfer control [31]. Final values of polarization resistance R_p shown in Table 2 point to significant difference between the ground and UIP samples. The ground samples have reached the polarization resistance value of 2189 $\Omega.cm^2$ and UIP samples have reached R_p of 68 $\Omega.cm^2$ which is difference of more than 32-times. It means that application of UIP on AZ31 has led to significant reduction of corrosion resistance as the polarization resistance is in direct relation with corrosion performance.

The PD tests have been performed in 0.1M NaCl to obtain information about corrosion thermodynamics (corrosion potential E_{corr}) and kinetics (corrosion current density i_{corr}) of samples. Potentiodynamic polarization curves of ground and UIP samples are shown in Figure 6 and corresponding values of electrochemical characteristics are stated in Table 3. The UIP samples have reached the E_{corr} value of -1433 mV. This value is more positive compared to value of ground samples (-1508 mV). Despite small potential difference, results point to slightly higher thermodynamic nobility of UIP samples compared to ground samples. In terms of corrosion kinetics, the UIP samples have reached more than 5-times higher value of corrosion current density (98 **$\mu A/cm^2$**) in comparison to value reached by ground samples (18 **$\mu A/cm^2$**). Current density is in direct proportionality with corrosion rate and it means that also corrosion rate r_{corr} is higher in case of UIP samples as it mentioned in Table 3. On the one hand UIP samples show better performance from thermodynamic point of view but on the other hand UIP samples are worse regarding corrosion kinetics. It has to be noted that corrosion of magnesium and its alloys is mainly driven by its kinetics hence the UIP has negatively influenced corrosion resistance of chosen alloy [10], [12]. These results are in good agreement with values obtained by EIS measurements.

It has to be bear in mind that there is possibly number of reasons which could influence corrosion process and their combination have led to worse corrosion resistance of UIP samples. According to the several authors, the film created on the surface (with composition mentioned above) provide a corrosion protection only in highly alkaline environment and thus is not able to protect Mg alloy for sufficient time in acidic and neutral solutions [9], [10], [29]. Furthermore, if Cl^- ions are presented in solution, this film is broke down after short period and corrosion of underneath substrate takes place [12], [29]. The presence of quasi passive film can be also proved by shift of corrosion potential to more positive values. Another factor is that application of UIP leads to the creation of defects (enlarged active surface) and to inducing dislocation and vacations. The surface of samples became highly electrochemically active and thus formation of corrosion products barrier, which could slow down the degradation, becomes more complicated compared to ground samples which have reached better corrosion resistance [30], [31].

4. Conclusion

Based on the performed experiments, measurements and analysis of results, these conclusions can be stated:

- Application of ultrasonic impact peening led to the creation of surface film and defects on the surface of AZ31 magnesium alloy.
- Samples after UIP reached more than 32-times lower value of polarization resistance in 0.1M NaCl (68 $\Omega.cm^2$) compared to value of R_p reached by ground samples (2189 $\Omega.cm^2$).
- UIP improved thermodynamic stability of AZ31 as the corrosion potential values of treated samples reached more positive E_{corr} (-1433 mV) compared to ground ones (-1508 mV), although improvement is not significant. However corrosion kinetics became worse as the UIP samples reached more than 5-times higher value of corrosion current density i_{corr} (98 **µA/cm**^{2} vs. 18 **µA/cm**^{2}) and higher corrosion rate r_{corr} compared to the non-treated samples.

Acknowledgement

The research was supported by project ITMS 26220220048 and by Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the contract VEGA no. 1/0045/17.

References

[1] MARUYAMA, K., SUZUKI, M., SATO, H.: Creep Strength of Magnesium-Based Alloys. Metallurgical and Materials Transactions A, 33, 875-882, 2002. https://doi.org/10.1007/s11661-002-0157-7

[2] MENG, X., WU, R., ZHANG, M., WU, L., CUI, CH.: Microstructures and Properties of Superlight Mg-Li-Al-Zn Wrought Alloys. Journal of Alloys and Compounds, 486, 722-725, 2009. https://doi.org/10.1016/j.jallcom.2009.07.047

[3] DARGUSCH, M. S., PETTERSEN, K., NOGITA, K., NAVE, M. D., DUNLOP, G. L.: The Effect of Aluminium Content on the Mechanical Properties and Microstructure of Die Cast Binary Magnesium-Aluminium Alloys. Materials Transactions, 47(4), 977-982, 2006. https://doi.org/10.2320/matertrans.47.977

[4] JUN, C., QING, Z., QUAN-AN, L.I.: Effect of Y and Ca Addition on the Creep Behaviors of AZ61 Magnesium Alloys. Journal of Alloys and Compounds, 686, 375-383, 2016. https://doi.org/10.1016/j.jallcom.2016.06.015

[5] BUKOVINOVA, L., BUKOVINA, M., PASTOREK, F.: Influence of Microstructural Features on the Corrosion Behaviour of AZ91 Alloy in Chloride Media. Materials Engineering - Materialove inzinierstvo, 21, 153-158, 2014.

[6] ATRENS, A., SONG, G.-L., LIU, M., SHI, Z., CAO, F., DARGUSCH, M. S.: Review of Recent Developments in the Field of Magnesium Corrosion. Advanced Engineering Materials, 17(4), 400-453, 2015. https://doi.org/10.1002/adem.201400434

[7] FRIEDRICH, H. F., MORDIKE, B. L.: Magnesium Technology, 1st edition. Springer, New York, 2006.

[8] ABBOT, T. B.: Magnesium: Industrial and Research Developments over the Last 15 years. Corrosion, 71(2), 120-127, 2015. https://doi.org/10.5006/1474

[9] GUSIEVA, K., DAVIES, C. H. J., SCULLY, J. R., BIRBILIS, N.: Corrosion of Magnesium Alloys: The Role of Alloying. International Materials Reviews, 60, 169-194, 2014. https://doi.org/10.1179/1743280414Y.0000000046

[10] SONG, G. L.: Corrosion of Magnesium Alloys, 1st edition. Woodhead Publishing, Philadelphia, 2011.

[11] ATRENS, A., DIETZEL, W.: The Negative Difference Effect and Unipositive Mg. Advanced Engineering Materials, 9(4), 292-297, 2007. https://doi.org/10.1002/adem.200600275

[12] ESMAILY, M., SVENSSON, J. E., FAJARDO, S., BIRBILIS, N., FRANKEL, G.S., VIRTANEN, S., ARRABAL, R., THOMAS, S., JOHANSSON, L. G.: Fundamentals and Advances in Magnesium Alloy Corrosion. Progress in Materials Science, 89, 92-193, 2017. https://doi.org/10.1016/j.pmatsci.2017.04.011

[13] CZERWINSKI, F.: Magnesium Alloys-Corrosion and Surface Treatments, 1st edititon. InTech, Rijeka, 2011.

[14] BIRBILIS, N., KING, A. D., SINU, T., FRANKEL, G., SCULLY, J. R.: Evidence for Enhanced Catalytic Activity of Magnesium Arising from Anodic Dissolution. Electrochimica Acta, 132, 277-283, 2014. https://doi.org/10.1016/j.electacta.2014.03.133

[15] LAMAKA, S. V., HOCHE, D., PETRAUSKAS, R. P., BLAWERT, C., ZHELUDKEVICH, M. L.: A New Concept for Corrosion Inhibition of Magnesium: Suppression of Iron Re-Deposition. Electrochemistry Communications, 62, 5-8, 2016. https://doi.org/10.1016/j.elecom.2015.10.023

[16] IZQUIERDO, J., FERNANDEZ-PEREZ, B. M., FILOTAS, D., ORI, Z., KISS, A., MARTIN-GOMEZ, R. T., NAGY, L., NAGY, G., SOUTO, R. M.: Imaging of Concentration Distributions and Hydrogen Evolution on Corroding Magnesium Exposed to Aqueous Environments Using Scanning Electrochemical Microscopy. Electroanalysis, 28, 1-14, 2016. https://doi.org/10.1002/elan.201600265

[17] APELFELD, A., KRIT, B., LUDIN, V., MOROZOVA, N., VLADIMIROV, B., WU, R. Z.: The Characterization of Plasma Electrolytic Oxidation Coatings on AZ41 Magnesium Alloy. Surface and Coatings Technology, 322, 127-133, 2017. https://doi.org/10.1016/j.surfcoat.2017.05.048

[18] KAJANEK, D., PASTOREK, F., FINTOVA, S., BACA, A.: Study of Corrosion Behavior of Dicalcium Phosphate-Dihydrate (DCPD) Coating Prepared by Large Amplitude Sinusoidal Voltammetry (LASV) Technique on ZW3 Magnesium Alloy. Proceedia Engineering, 192, 399-403, 2017. https://doi.org/10.1016/j.proeng.2017.06.069

[19] OMASTA, M., HADZIMA, B., PASTOREK, F.: Study of Calcium Phosphate (OCP) Electrodeposition Process on Elektron 21 Magnesium Alloy Surface. Materials Science Forum, 818, 115-120, 2015. https://doi.org/10.4028/www.scientific.net/MSF.818.115

[20] MALAKI, M., DING, H.: A Review of Ultrasonic Peening Treatment. Materials and Design, 87, 1072-1086, 2015. https://doi.org/10.1016/j.matdes.2015.08.102

[21] DONG, Z., LIU, Z., LI, M., LUO, J. L., CHEN, W., ZHENG, W., GUZONAS, D.: Effect of Ultrasonic Impact Peening on the Corrosion of Ferritic-Martensitic Steels in Supercritical Water. Journal of Nuclear Materials, 457, 266-272, 2015.

[22] MORDYUK, B. N., PROKOPENKO, G. I.: Ultrasonic Impact Peening for the Surface Properties' Management. Journal of Sound and Vibration, 308, 855-866, 2007. https://doi.org/10.1016/j.jsv.2007.03.054

[23] RODOPOULOS, C. A., PANTELAKIS, S. G., PAPADOPULOS, M. P.: The Effect of Ultrasonic Impact Treatment on the Fatigue Resistance of Friction Stir Welded Panels. Journal of Materials Engineering and Performance, 18(9), 1248-1257, 2009. https://doi.org/10.1007/s11665-009-9385-y

[24] FENG, Y., HU, S., WANG, D., CUI, L.: Formation of Short Crack and its Effect on Fatigue Properties of Ultrasonic Peening Treatment S355 Steel. Materials and Design, 89, 507-515, 2016. https://doi.org/10.1016/j.matdes.2015.10.009

[25] VOORT, G. F. V.: ASM Handbook Vol. 9: Metallography and Microstructures. ASM International, New York, 2004.

[26] RAI, X. J., NISHIMURA, T.: Studies on Galvanic Corrosion of Iron-Magnesium Couple by Scanning Electrochemical Microscopy in 0.1 M NaCl Solution. Journal of Industrial and Engineering Chemistry, 41, 141-150, 2016. https://doi.org/10.1016/j.jiec.2016.07.020

[27] KAJANEK, D., HADZIMA, B., PASTOREK, F., NESLUSAN JACKOVA, M.: Electrochemical Impedance Spectroscopy Characterization of ZW3 Magnesium Alloy Coated by DCPD Using LASV Deposition Technique. Acta Metallurgica Slovaca, 23(2), 147-154, 2017. https://doi.org/10.12776/ams.v23i2.900

[28] AMIRUDIN A., THIERRY D.: Application of Electrochemical Impedance Spectroscopy to Study the Degradation of Polymer-Coated Metals. Progress in Organic Coatings, 26(1), 1-28, 1996.

[29] GHALI, E.: Corrosion Resistance of Aluminium and Magnesium Alloys, 1st edition. John Wiley & Sons Inc., New Jersey, 2010. https://doi.org/10.1002/9780470531778

[30] MHAEDE, M., PASTOREK, F., HADZIMA, B.: Influence of Shot Peening on Corrosion Properties of Biocompatible Magnesium Alloy AZ31 Coated by Dicalcium Phosphate Dihydrate (DCPD). Materials Science and Engineering C, 39, 330-335, 2014. https://doi.org/10.1016/j.msec.2014.03.023

[31] ZHOU, J., YANG, Y., FRANK, M.A., DETSCH, R., BOCCACCINI, A.R., VIRTANEN, S.: Accelerated Degradation Behavior and Cytocompatibility of Pure Iron Treated with Sandblasting. Applied Materials and Interfaces, 8, 26482-26492, 2016. https://doi.org/10.1021/acsami.6b07068

https://doi.org/10.26552/com.C.2018.3.30-35

Slavica Miladinovic - Sandra Velickovic - Blaza Stojanovic - Ruzica Nikolic*

OPTIMIZATION OF PARAMETERS OF A PLANETARY GEARBOX USING THE TAGUCHI-GREY RELATIONAL ANALYSIS

Optimization of parameters of central gears of a planetary gearbox was performed in this study. Since the optimization of multi-performance characteristics is more complex compared to optimization of a single-performance characteristics, the Taguchi-Grey analysis was used. That approach combines the Taguchi orthogonal array L27 with the Grey relational analysis to determine the influence of the material, module and gear width on the safety coefficient of the planetary gearbox. A Grey relational grade obtained from the Grey relational analysis is used as the performance characteristics in the Taguchi method and Analysis of variance (ANOVA). The optimal combination of parameters for the safety coefficient values was obtained, when the value of the Grey relational grade is 1. Based on the ANOVA analysis, it was established that the most influential parameter is the module, then the gear width and the least influential one is the material.

Keywords: *Taguchi method, Grey relational analysis, ANOVA, optimization, planetary gearbox*

1. Introduction

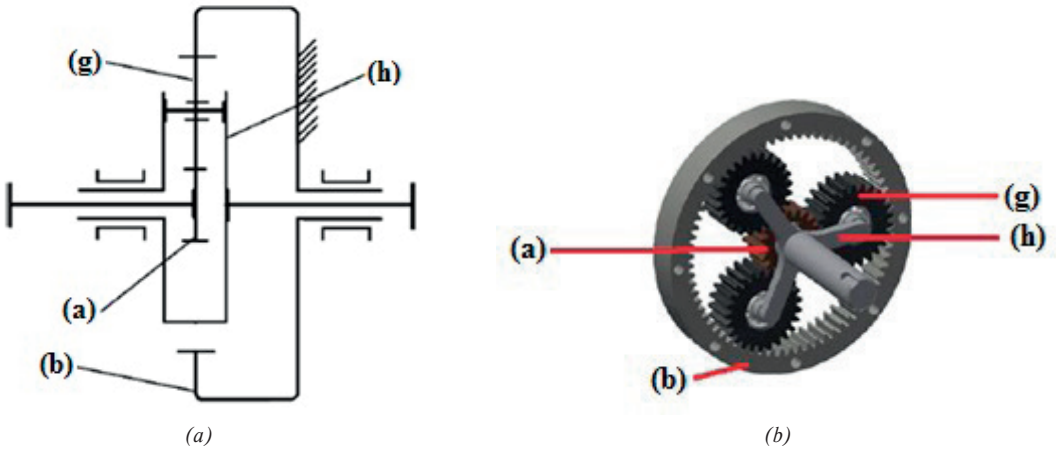
The planetary gear transmissions are widely used in industrial applications for motion and power transmission such as automotive, aerospace, robotics, machinery etc., because of their significant advantages including the compact structure, self-centering capability, high power density, high transmission ratios, high degree of efficiency, durability in operation, low level of noise and vibrations.

Design of planetary gear transmissions depends on a large number of parameters, some of those are: material, module, addendum, profile shift, tooth thickness, dedendum, cutter tip radius, pressure angle, gear width etc. The analysis and optimization of one simple planetary gearbox marked as A^{b}_{ha} was performed in this paper. The letters in the mark A^{b}_{ha} mean: A-the single-sided satellite, b-the fixed gear (internal toothed gear), a-the central gear (driving gear), h-the satellite carrier (connected to the output shaft) [1], [2]. Planetary gearbox of this type is shown in Figure 1, and next to already mentioned gears and satellite carrier, it consists of three satellites (g), as well.

The optimization of a simple planetary gear train was performed by authors of [3], they were observing the results obtained by changing the number of gear teeth, module, number of satellites and gear width. For development of a mathematical model for multi objective optimization they used the Puerto and ELECTRE methods and noticed that the correlations between used methods are obvious. This kind of approach shows that the other planetary gear transmission types can be subjects of the multi objective optimization. The multi objective approach in optimization is much faster with the use of the Taguchi-Grey method. For optimization of the plastic gear production authors in [4] applied the hybrid integration of Taguchi parametric design, the Grey relational analysis and the principal component analysis. For demonstration of efficiency and validity of the proposed hybrid optimization methods, in controlling all the influential injection molding processing parameters during the plastic gear manufacturing, a plastic gear was used. The optimal combination of different process parameters is determined in order to minimize the shrinkage behavior in tooth thickness, addendum circle and dedendum circle of the molded gear. They concluded that the proposed optimization method can produce a plastic-molded gear with minimum shrinkage behavior of 1.8% in tooth thickness, 1.53% in addendum circle and 2.42% in dedendum circle and it should be noted that these values are less than the values in the main experiment. Therefore, defects related to shrinkage that lead to severe failure in plastic gears can be effectively minimized while satisfying the demand of the global plastic gear industry [5].

The Taguchi method is a well-known optimization method, which is used in a variety of applications [5], [6]. Application of this method ensures savings in labor and time costs, with simultaneously examining several parameters in several experimental conditions, providing the quantitative information, [7], [8]. Limitation of the Taguchi method is that it can only be applied to solve individual objective problems and cannot be used for problems with multiple objectives optimization, especially when there are different performance characteristics needed. However, this problem has been overcome and the optimization of multiple objectives can be performed by combining the Taguchi method and the Grey relational analysis to optimize the multiple characteristics [9]. Combination of the Taguchi method and the Grey relational analysis was used for optimization of the EDM process (Electrical Discharge Machining) in [10]. Parameters of machining (workpiece polarity, pulse on time, duty factor, open discharge voltage, discharge current and dielectric fluid) were optimized considering the multiple performance characteristics including material removal rate, surface roughness and electrode wear ratio. The Taguchi-Grey method was really

* [1]Slavica Miladinovic, [1]Sandra Velickovic, [1]Blaza Stojanovic, [2]Ruzica Nikolic
[1]Faculty of Engineering, University of Kragujevac, Serbia
[2]Research Center, University of Zilina, Slovakia
E-mail: ruzicarnikolic@yahoo.com



(a) *(b)*

Figure 1 *Planetary gearbox of type* A^{b}_{ha} *: (a) Schematic presentation and (b) the 3D appearance of the planetary gearbox*

Table 1 *Control factors with their levels*

Control factors	Units	Level I	Level II	Level III
Material (A)	/	16MnCr5	28Cr4	C15E
Module (B)	[mm]	2.50	2.75	3.00
Gear width (C)	[mm]	30	33	36

needed for optimization of the EDM process because material removal rate is the higher-the-better performance characteristics and the surface roughness and electrode wear ratio are the lower-the-better performance characteristics - improvement of one performance characteristics may require a degradation of the other performance characteristics. This is the main reason that optimization of the multiple performance characteristics is much more complicated than optimization of a single performance characteristics. The Taguchi-Grey method has proven in this case that the machining performance in the EDM process can be effectively improved. Another successful application of the Taguchi-Grey method was in the color filter manufacturing [11] for improving the yield of Chrome (Cr) thin-film sputtering process of the black matrix (BM) and to find the robust parameters of the process with multiple quality characteristics, as well. For the experiment, the Taguchi L9 orthogonal array was chosen and the two quality characteristics and four control factors with three levels were considered. The correlation between the quality characteristics was considered and, for elimination of the multiple colinearity, the principal component analysis was applied. The entropy measurement method was applied for obtaining the weights of the quality characteristics. After confirmation of the experiment, it was concluded that manufacturing with optimal parameters, obtained by the Taguchi-Grey method, can reduce the opportunities of the repair and rework of the Cr thin-film and raise the yield.

The Taguchi-Grey method can be applied in various aspects of industry, from optimization of processes to optimization of a single machine or part of a machine.

In this paper, the Taguchi-Grey method was applied for optimization of the safety coefficient for the surface durability of the entire planetary gearbox. The safety coefficients for the surface durability of the first gear pair a-g (S_{H1-2}) and the second gear pair g-b (S_{H2-3}) were considered. Influence of parameters, such as the material, module and the gear width, was analyzed to obtain the largest value of the safety coefficient. The planetary gearbox considered in this paper was chosen since it is the simplest one; the Taguchi-Grey method can be successfully applied for other types of planetary gearboxes that are more complex than the considered gearbox, with taking into account some restrictions.

2. Experiments' design

The Taguchi method is one of the most useful techniques for a single-response optimization studies used for a number of engineering problems. When there is a multi-response optimization of processes, the use of Taguchi methodis is not reliable and it should be combined with some other method, e.g. the Grey relational analysis.

The experiments were designed using the Taguchi design of experimental methods, which are based on an orthogonal array of experiments, resulting in optimal setting of the process control parameters. The orthogonal array provides a set of balanced experiments with less number of experimental runs, in order to evaluate the optimal parameters. For computing of the Taguchi data and data analysis the MiniTab 16 was used.

In this paper, the orthogonal array L27 was chosen, which provides the least combination of parameters in a matrix. Those parameters were different and in this way the direct effect on the observed output was achieved. The chosen input parameters of the planetary gearbox are material, module and the gear width and they are all parameters of the third level. Levels of chosen parameters and their marks are listed in Table 1.

The Taguchi method employs the signal to noise (S/N) ratio, to measure the performance of the process response. The (S/N) ratio, being the ratio of the mean to standard deviation, can effectively consider the variation encountered in a set of

Table 2 *Experimental design using the L27 orthogonal array*

Exp. No	A	B	C	Experimental results		The Grey grade	S/N for the Grey grade	Rank
				S_{H1-2}	S_{H2-3}			
1	16MnCr5	2.50	30	1.31	2.36	0.36716	-8.70297	25
2	16MnCr5	2.50	33	1.37	2.47	0.40345	-7.88425	21
3	16MnCr5	2.50	36	1.42	2.57	0.44152	-7.10098	19
4	16MnCr5	2.75	30	1.44	2.59	0.45408	-6.85741	17
5	16MnCr5	2.75	33	1.51	2.71	0.51924	-5.69263	12
6	16MnCr5	2.75	36	1.57	2.82	0.59492	-4.51088	8
7	16MnCr5	3.00	30	1.58	2.82	0.60218	-4.40554	7
8	16MnCr5	3.00	33	1.64	2.95	0.71780	-2.87995	4
9	16MnCr5	3.00	36	1.71	3.07	0.89541	-0.95958	2
10	28Cr4	2.50	30	1.33	2.39	0.37743	-8.46331	24
11	28Cr4	2.50	33	1.39	2.5	0.41588	-7.62063	20
12	28Cr4	2.50	36	1.45	2.61	0.46306	-6.68725	16
13	28Cr4	2.75	30	1.46	2.63	0.47241	-6.51365	15
14	28Cr4	2.75	33	1.53	2.75	0.54334	-5.29850	11
15	28Cr4	2.75	36	1.59	2.87	0.63134	-3.99473	6
16	28Cr4	3.00	30	1.51	2.86	0.57470	-4.81114	10
17	28Cr4	3.00	33	1.67	2.99	0.77659	-2.19617	3
18	28Cr4	3.00	36	1.74	3.12	1.00000	0.00000	1
19	C15E	2.50	30	1.24	2.24	0.33333	-9.54243	27
20	C15E	2.50	33	1.3	2.34	0.36149	-8.83814	26
21	C15E	2.50	36	1.35	2.44	0.39174	-8.14002	23
22	C15E	2.75	30	1.37	2.46	0.40161	-7.92385	22
23	C15E	2.75	33	1.43	2.57	0.44544	-7.02428	18
24	C15E	2.75	36	1.49	2.68	0.50000	-6.02060	13
25	C15E	3.00	30	1.49	2.68	0.50000	-6.02060	13
26	C15E	3.00	33	1.56	2.8	0.58017	-4.72887	9
27	C15E	3.00	36	1.62	2.92	0.68159	-3.32956	5

experimental runs. The (S/N) ratios, which are log functions of a desired output, serve as the objective functions for optimization, help in the data analysis and prediction of the optimum results. There are three types of the (S/N) ratio analyses that are generally applicable: the-higher-the-better, the-nominal-the-better and the-lower-the-better [7].

In this study, the higher-the-better quality characteristics has been used for calculating the (S/N) ratio for the Grey grade of the responses. The (S/N) ratio can be calculated using the following equation:

$$S/N = -10\log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i^2}\right) \quad (1)$$

where y_i is the performance value of the i-th experiment and n is the number of experiments.

Results of all 27 experiments and the corresponding (S/N) ratios for the Grey grade are shown in Table 2.

3. Optimization using the Grey relational analysis

When there are multiple outputs, a combination of the Taguchi method and the Grey relational analysis is usually used for optimization. In this case, the Grey relational analysis is used to generate a single response for the two performance characteristics. Like in the Taguchi design, in the Grey relational analysis, there are also three quality characteristics: the-higher-the-better, the-nominal-the-better and the-lower-the-better.

In Table 2 are given experimental results, the Grey grade, the signal-to-noise ratio (S/N) for the Grey grade and results are ranked. Basically, the larger the Grey relational grade is, the closer the product quality will be to the ideal value.

In this paper, for the desired higher values of the safety coefficients of surface durability of a planetary gearbox at the output (the-higher-the-better quality characteristics), normalizing in the range of 0 - 1 was done according to the following equation [12]:

$$x_i(k) = \frac{y_i(k) - \min y_i(k)}{\max y_i(k) - \min y_i(k)} \quad (2)$$

Table 3 *Analysis of variance for the (S/N) ratios for the Grey grade*

Source	DF	Seq SS	Adj SS	Adj MS	F	P	Pr %
A	2	15.748	15.748	7.8738	118.89	0.000	10.09
B	2	106.377	106.377	53.1883	803.11	0.000	68.17
C	2	28.120	28.120	14.0602	212.30	0.000	18.02
A*B	4	1.140	1.140	0.2849	4.30	0.038	0.73
A*C	4	0.820	0.820	0.2049	3.09	0.081	0.53
B*C	4	3.304	3.304	0.8260	12.47	0.002	2.12
Residual Error	8	0.530	0.530	0.0662			0.34
Total	26	156.038					100.00

DF - degree of freedom, Seq SS - Sequential sum of squares, Adj SS - Adjusted sum of squares, Adj MS - Adjusted mean square, F-value, Pr-percentage of contribution

Table 4 *Response for the Grey grade*

Level	A	B	C
1	-5.444	-8.109	-7.027
2	-5.065	-5.982	-5.796
3	-6.841	-3.259	-4.527
Delta	1.776	4.850	2.500
Rank	3	1	2

where: k is the number of factors, in range 1-3, i is the experiment number, in range 1-27, $\max y_i(k)$ is the highest experimental value for the k-th output and $\min y_i(k)$ is the lowest experimental value for the k-th output [13].

For the-smaller-the-better quality characteristics and the-nominal-the-better quality characteristics Equation (3) and Equation (4) are used, respectively [13]:

$$x_i(k) = \frac{y_i(k) - \min y_i(k)}{\max y_i(k) - \min y_i(k)} \tag{3}$$

$$x_i(k) = \frac{|y_i(k) - OB|}{\max\{\max y_i(k) - OB, OB - \min y_i(k)\}} \tag{4}$$

The absolute difference value is calculated according to expression [13]:

$$\Delta_{0i} = \| x_0(k) - x_i(k) \| \tag{5}$$

where: $x_0(k)$ is the referential normalized value $(x_0(x) = 1)$.

The Grey relational coefficient is calculated according to [9], [12]:

$$\xi_i(k) = \frac{\Delta_{\min} + \zeta\Delta_{\max}}{\Delta_{0i}(k) + \zeta\Delta_{\max}} \tag{6}$$

where: $\Delta_{\min}$ and $\Delta_{\max}$ are the lowest and the highest values of absolute difference, respectively, ζ is the distinguishing coefficient in the interval $0 \le \zeta \le 1$; if the weighting of output is equal, then $\zeta = 0.5$ [5], [12].

Then, the Grey relational grade (γ_i) can be calculated according to:

$$\gamma_i = \frac{1}{n}\sum_{k=1}^{n} \xi_i(k) \tag{7}$$

where n is the number of outputs; γ_i has to be as high as possible.

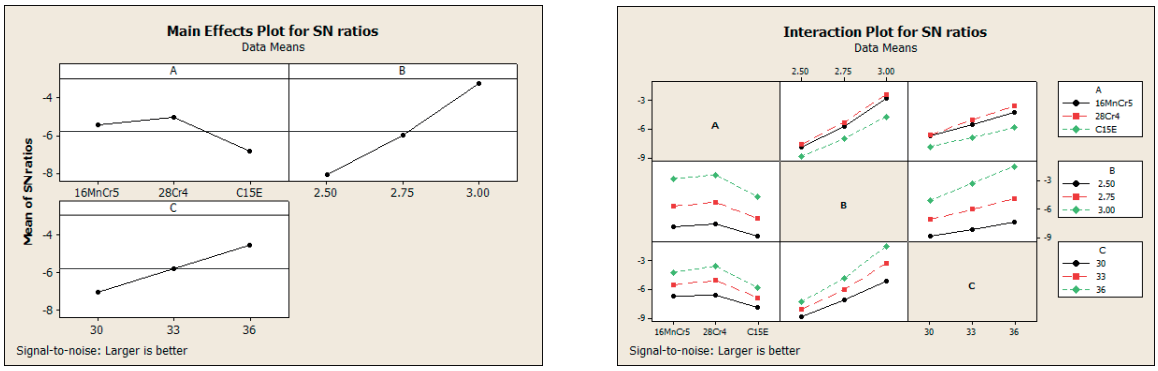
The experimental data are analyzed by using the Taguchi signal-to-noise ratio to measure the performance of the process response. The (S/N) ratio is the objective function for optimization and with the use of a logarithmic function it helps in data analysis and in prediction of the optimal results. In this paper, the-higher-the-better quality characteristics was used for calculating the (S/N) ratios of the responses [5], [6], [7]. In this way, the multiple response problems are converted into a single response problem.

4. Analysis of variance

Analysis of variance (ANOVA) is a statistical model used for finding the most influential parameter that contributes towards the outputs and how the variations in inputs affect the outputs. This method is easy to understand and use when the number of parameters is small.

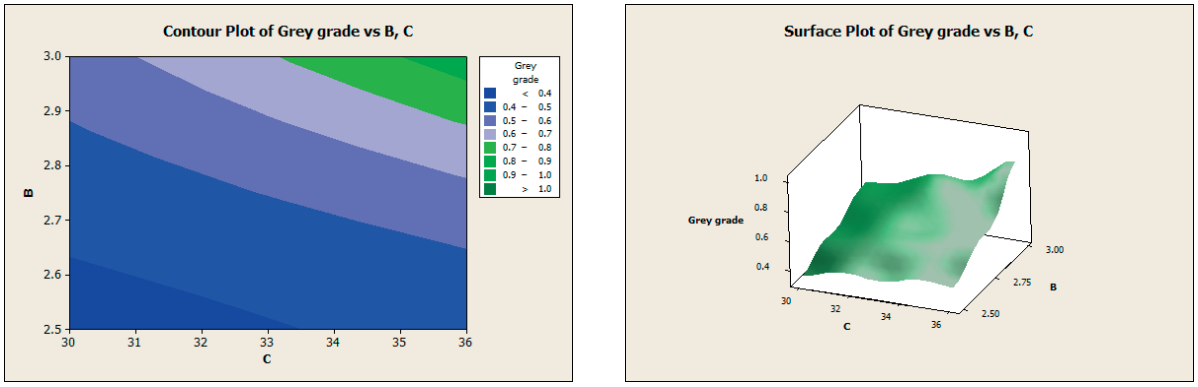
The effect of material, module and gear width on the safety coefficients for the surface durability of a planetary gearbox was analyzed with the 95% confidence level and 5% significance level [13]. The probability values show the level of significance of each factor, the parameter is highly statistically significant if the corresponding P value is less than 0.05. Table 3 shows the results of the ANOVA for multi-performance characteristics.

According to the ANOVA, the most significant control parameter is the module with contribution of about 68.17%, followed by the gear width with 18.02% and material with 10.09%. Interactions of observed parameters are insignificant because their effects on the safety coefficients were negligible. Analysis of parameters' influences was obtained from the response tables of the mean (S/N) ratio and the results are listed in Table 4. The larger values of the (S/N) ratio correspond to the better quality, thus the optimal combination of design parameters was obtained as A2B3C3, which is also shown in Figure2.



a) *b)*

Figure 2 *Plots of the (S/N) ratio for the safety coefficients for the durability of a planetary gearbox: a) main effects plot; b) interaction plot*



a) *b)*

Figure 3 *Response plots of combined effect of module and gear width on the surface durability of a planetary gearbox: a) contour plot; b) surface plot*

In Figure 3 are presented the contour plot (Figure 3a) and the surface plot (Figure 3b) of combined effect of module and the gear width on the Grey grade for the safety coefficients of surface durability of a planetary gearbox.

Based on the carried out analysis and on the plots shown in Figure 3, it can be concluded that the highest surface durability of a planetary gearbox is achieved for the highest values of the module and the gear width.

5. Conclusions

In this study, the Taguchi's orthogonal array with the Grey relational analysis was used to optimize the multiple performance characteristics and the following conclusions were drawn:

- According to results of the ANOVA analysis, the main contributing factors, affecting the safety coefficients for the surface durability of a planetary gearbox, were material, module and the gear width, with contributions of 10.09%, 68.17% and 18.02%, respectively. Interaction of parameters B*C (module*gear width) has influence of 2.12%, interactions of other parameters had negligible influences;
- The optimal combination of parameters, according to the Taguchi-Grey analysis is A2B3C3, which means that the highest values of the safety coefficients for the surface durability of a planetary gearbox are for material 28Cr4, module of 3 mm and the gear width of 36 mm. The Grey grade for this optimal combination of parameters is 1 and the (S/N) ratio is 0;
- Based on results from the applied methods, one can conclude that the Taguchi-Grey method can be applied for optimization of the safety coefficients for the surface durability of a planetary gearbox.

Results of this research suggest that the complex optimization of the multi-response problems with the Taguchi method can be easily simplified by adding the Grey relational analysis. Joined strength of these two methods has given more precise safety coefficients for the surface durability of a planetary gearbox.

Note: *Shorter version of this research was presented at "SEMDOK 2018" Conference, reference [14].*

Acknowledgement

This research was financially supported by European regional development fund and Slovak state budget by the project "Research Centre of the University of Zilina" and by the Ministry of Education, Science and Technological Development of Republic of Serbia through grant TR 35021.

References

[1] MILADINOVIC, S., VELICKOVIC, S., NOVAKOVIC, M.: Application of Taguchi method for the Selection of Optimal Parameters of Planetary Driving Gear. Applied Engineering Letters, 1(4), 98-104, 2016.

[2] TANASIJEVIC, S., VULIC, A.: Mechanical Transmitters - Planetary Transmitters and Variators (in Serbian). Faculty of Mechanical Engineering, Kragujevac, 2006.

[3] STEFANOVIC-MARINOVIC, J., PETKOVIC, M., STANIMIROVIC, I.: Application of the ELECTRE Method to Planetary Gear Train Optimization. Journal of Mechanical Science and Technology, 29(2), 647-654, 2015. https://doi.org/10.1007/s12206-015-0124-z

[4] MEHAT, N. M., KAMARUDDIN, S., OTHMAN, A. R.: Hybrid Integration of Taguchi Parametric Design Grey Relational Analysis and Principal Component Analysis Optimization for Plastic Gear Production.Chinese Journal of Engineering, 1-11, 2014. http://dx.doi.org/10.1155/2014/351206

[5] MILADINOVIC, S., RADOSAVLJEVIC, S., VELICKOVIC, S., ATYAT, R., SKULIC, A., SLJIVIC, V.: Optimization of Efficiency of Worm Gear Reducer by Using Taguchi-Grey Method. Applied Engineering Letters, 2(2), 69-75, 2017.

[6] RAMDUMAR, R., KEN, M.: Minimization of Shrinkage in Injection Molding Process of Acetal Polymer Gear Using Taguchi DOE Optimization and ANOVA Method. International Journal of Mechanical and Industrial Technology, 4(2),72-79, 2017.

[7] STOJANOVIC, B., BLAGOJEVIC, J., BABIC, M., VELICKOVIC, S., MILADINOVIC, S.: Optimization of Hybrid Aluminum Composites Wear Using Taguchi Method and Artificial Neural Network. Industrial Lubrication and Tribology, 69(6),1005-1015, 2017. https://doi.org/10.1108/ILT-02-2017-0043

[8] MILADINOVIC, S., VELICKOVIC, S.: Optimization and Prediction of Safety Coefficient for Surface Durability of Planetary Gearbox using Taguchi Design and Artificial Neural Network. Proceedings of the 3rd International Scientific Conference COMETa, Bosnia and Herzegovina, 139-146, 2016.

[9] YUNUS, M., ALSOUFI, M. S., MUNSHI, S. M.: Taguchi-Grey Relation Analysis for Assessing the Optimal Set of Control Factors of Thermal Barrier Coatings for High-Temperature Applications. Mechanics of Advanced Materials and Modern Processes, 2(4), 1-8, 2016. https://doi.org/10.1186/s40759-016-0011-z

[10] LIN, J. L.,LIN, C. L.: The Use of the Orthogonal Array with Grey Relational Analysis to Optimize the Electrical Discharge Machining Process with Multiple Performance Characteristics. International Journal of Machine Tools and Manufacture,42, 237–244, 2002.

[11] CHIANG, Y.M. , Hsieh, H. H.: The Use of the Taguchi Method with Grey Relational Analysis to Optimize the Thin-Film Sputtering Process with Multiple Quality Characteristic in Color Filter Manufacturing, Computers and Industrial Engineering, 56(2), 648-661, 2008. https://doi.org/10.1016/j.cie.2007.12.020

[12] SUDEEPANA, J., KUMARB, K., BARMANC, T. K., SAHOOC, P.: Study of Tribological Behavior of ABS/$CaCO_3$ Composite Using Grey Relational Analysis. Procedia Materials Science, 6, 682-691, 2014. https://doi.org/10.1016/j.mspro.2014.07.084

[13] STOJANOVIC B., BABIC M., IVANOVIC L.: Taguchi Optimization of Tribological Properties of Al/SiC/Graphite Composite. Journal of the Balkan Tribological Association, 22(3), 2592-2605, 2016.

[14] MILADINOVIC, S., VELICKOVIC, S., STOJANOVIC, B., NIKOLIC, R.: The Taguchi-Grey Relational Analysis for Optimization of the Safety Coefficient for the Surface Durability of a Planetary Gearbox. Proceedings of the 23rd International Seminar of Ph.D. Students SEMDOK 2018, Slovakia, 118-123, 2018.

Stefano Monti*

REDUCTION OF THE LOCAL STRESS FIELD AROUND HOLES THROUGH POROUS SHAPED STRUCTURES

Geometrical discontinuities in mechanical components are detrimental for the mechanical properties of the product itself. Actually, in proximity of such features, the stress increases due to the stress concentration factor, that in the case of a circular hole is equal to 3. Several solutions have been proposed to reduce the stress concentration value. In the present article, the application of a particular porous pattern that can be obtained by laser cutting with the appropriate finishing requirements is introduced in order to modify the local stress field and reduce the stress concentration value near the hole boundary.

Keywords: void pattern, stress concentration factor, stress field

1. Introduction

Holes are crucial features in mechanical components due to the stress concentration factor arising in the proximity of hole boundary. However, details with holes are widely used in engineering to guarantee some function, such as bolt or rivet joints. Anyway, the local stress concentration introduced by holes strongly limits the fatigue strength and is still a subject able to attract the interest of researchers and engineers. Through the years, many studies were proposed aimed at estimating the stress concentration factor both for isotropic [1], [2], [3] and orthotropic plates [4], [5], [6], [7]. Holed components are interesting even today, due to the increased manufacturing and machining possibilities offered by the new technologies that allow to define and put in practice original solutions for the local stress reduction. Meguid [8] introduced the so-called defense hole system in a uniaxially loaded plate with a reduction of 7.5-11 % of the stress concentration factor. In addition, other authors employed the defense hole systems for reducing the stress concentration factor [9], [10]. Jindal [11] investigated the influence of auxiliary holes on the static stress concentration around a central hole numerically by the finite element method (FEM) and experimentally by a two-dimensional photo-elastic method. Providakis et al. [12] presented a parametric study for the reduction of the dynamic stress concentration around a circular hole in an elastic finite plate subjected to a suddenly applied loading acting in its plane by using the BEM.

Alternatively, shaped porous structures have been recently studied and applied to metal sheet in order to achieve a specific material behavior, e.g. auxetic behavior, and manufactured by using additive manufacturing techniques. Grima et al. [13] found that conventional materials containing diamond or star shaped holes, arranged in a periodic structure, exhibits various Poisson's ratio values that can reach even negative values (auxetic behavior). Such evidence was found by Taylor et al. [14] by applying a pattern of elliptical voids, investigating different aspect ratios. As a consequence of the cited studies, Javid et al. [15] moved their interest to the stress behavior rather than on the material deformation and proposed a method for enhancing the fatigue life of turbine blades by replacing the commonly used circular holes with a novel orthogonal S-shaped void pattern, realized by means of the laser drilling. Such a novel pores design induces a soft mode of deformation based on rotation of the material around said voids, redistributing the stress field and modifying even the crack propagation phenomenon within the structure. The literature lacks in studies about employing the void patter for enhancing component's mechanical performances. As a consequence, there is a wide range of fields in which the void patterns can be employed for mechanical performance improvement.

The present study considers the case of a circular hole in an infinite plate under remote uniaxial load and proposes a C-shaped void pattern to redistribute the local stress field around the hole; in such a way the stress concentration factor can be consequently reduced. Moreover, the shape hereafter introduced is designed to be realized by means of the laser cutting technology. Furthermore, the introduced pattern is defined passing through the following steps:

- Introduction of the analytical description of the case study in order to comprehend which is the critical region;
- In defining the said pattern, it is necessary to define a suitable feature shape that is able to reduce the induced stress concentration factor as much as possible;
- Then, the distance between the two contiguous pattern rows is defined as the one able to better redistribute the stress field in the critical region;
- The same as before is done for defining the most suitable distribution of the features, considering a three-row pattern.

* **Stefano Monti**
Department of Mechanical Engineering, Politecnico di Milano, Italy
E-mail: stefano.monti@polimi.it

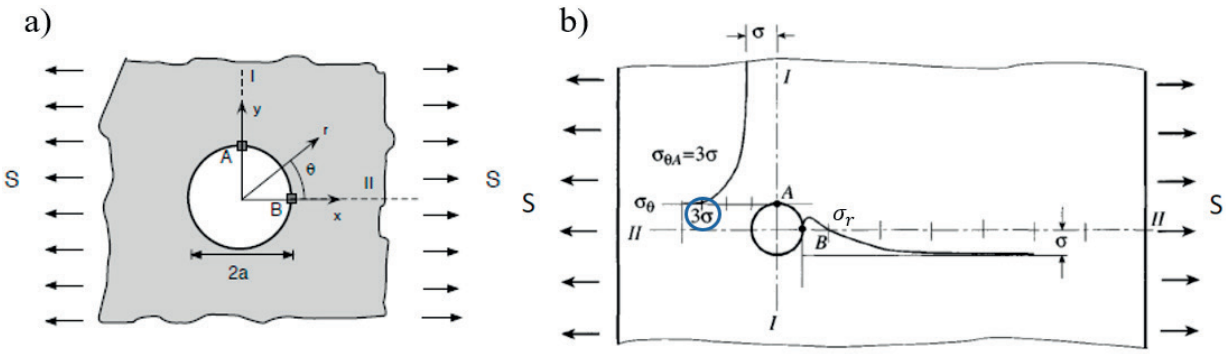


Figure 1 *Analytical model (a) and stress distribution (b) for a uniaxially loaded infinite holed plate*

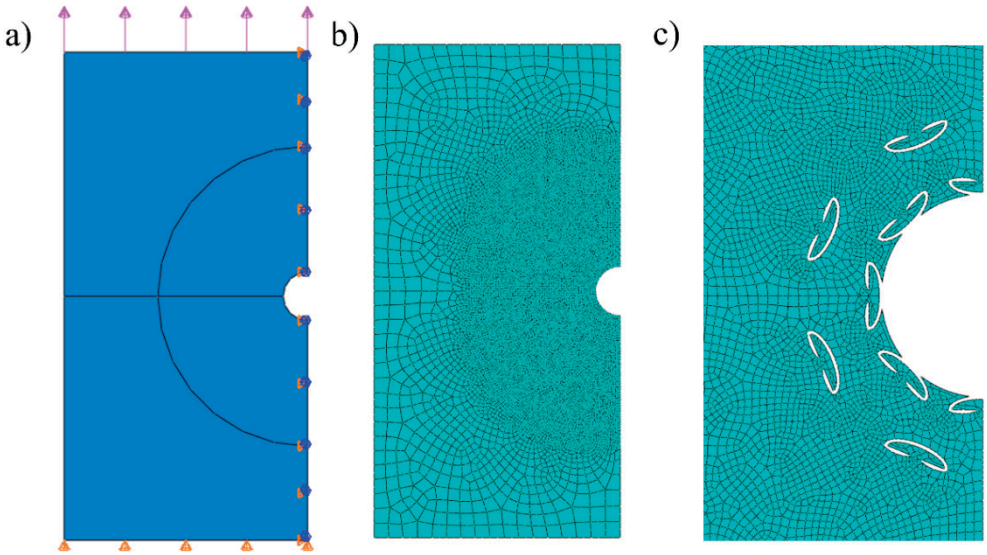


Figure 2 *Numerical model*

2. Stress concentration around a circular hole

The case of an infinite plate with a circular hole of radius R and subjected to a remote uniaxial tensile stress S, Figure 1a, is considered. The analytical formulation of the stress distribution is given by Airy's stress functions. For axisymmetric problems, such formulas can be conveniently represented in cylindrical coordinates, because of their dependence on the radial coordinate only, and presented in Equation 1.

$$\begin{cases} \sigma_r = \frac{1}{2}S\left(1-\frac{a^2}{r^2}\right)+\frac{1}{2}S\left(1-\frac{4a^2}{r^2}+\frac{3a^4}{r^4}\right)\cos(2\theta) \\ \sigma_\theta = \frac{1}{2}S\left(1-\frac{a^2}{r^2}\right)-\frac{1}{2}S\left(1+\frac{3a^4}{r^4}\right)\cos(2\theta) \\ \tau_{r\theta} = -\frac{1}{2}S\left(1+\frac{2a^2}{r^2}-\frac{3a^4}{r^4}\right)\sin(2\theta) \end{cases} \quad (1)$$

In Figure 1b the stress field described by Airy's stress functions is presented. The most critical point is placed perpendicular to the loading direction, for which at the hole boundary the stress value increases by a factor of 3. Clearly this is the crucial issue related to fatigue strength and justifies the efforts devoted to reducing the notch effect. A method based on the introduction of appropriate void patterns is presented and the results are critically discussed in view of a possible more general application to mechanical components for mitigating the present stress concentration. This kind of void patterns could be easily realized employing the laser cutting technology.

3. Feature design

As remarked in previous section, the case of an infinite holed plate is well known in literature and an analytical formulation is given by the Airy's stress function. However, the introduction of additional geometrical discontinuities makes difficult to obtain a simple and valid analytical formulation as well.

For this reason, a numerical model has been built to estimate the stress field for a holed infinite plate subjected to uniaxial load, Figure 2. The ABAQUS CAE software is employed for all numerical analysis. The plane stress 2D elements CPS4 type were used for discretizing the model (Figure 2b), whereas in Figure 2c a magnified view of the meshed region in proximity of the hole boundary is presented.

In particular, the mesh dimension goes from 3 mm (in the region far from the hole) to 0.5 mm (in the region interested by the hole and pattern). It is important to highlight that the mesh underwent a convergence analysis by changing the

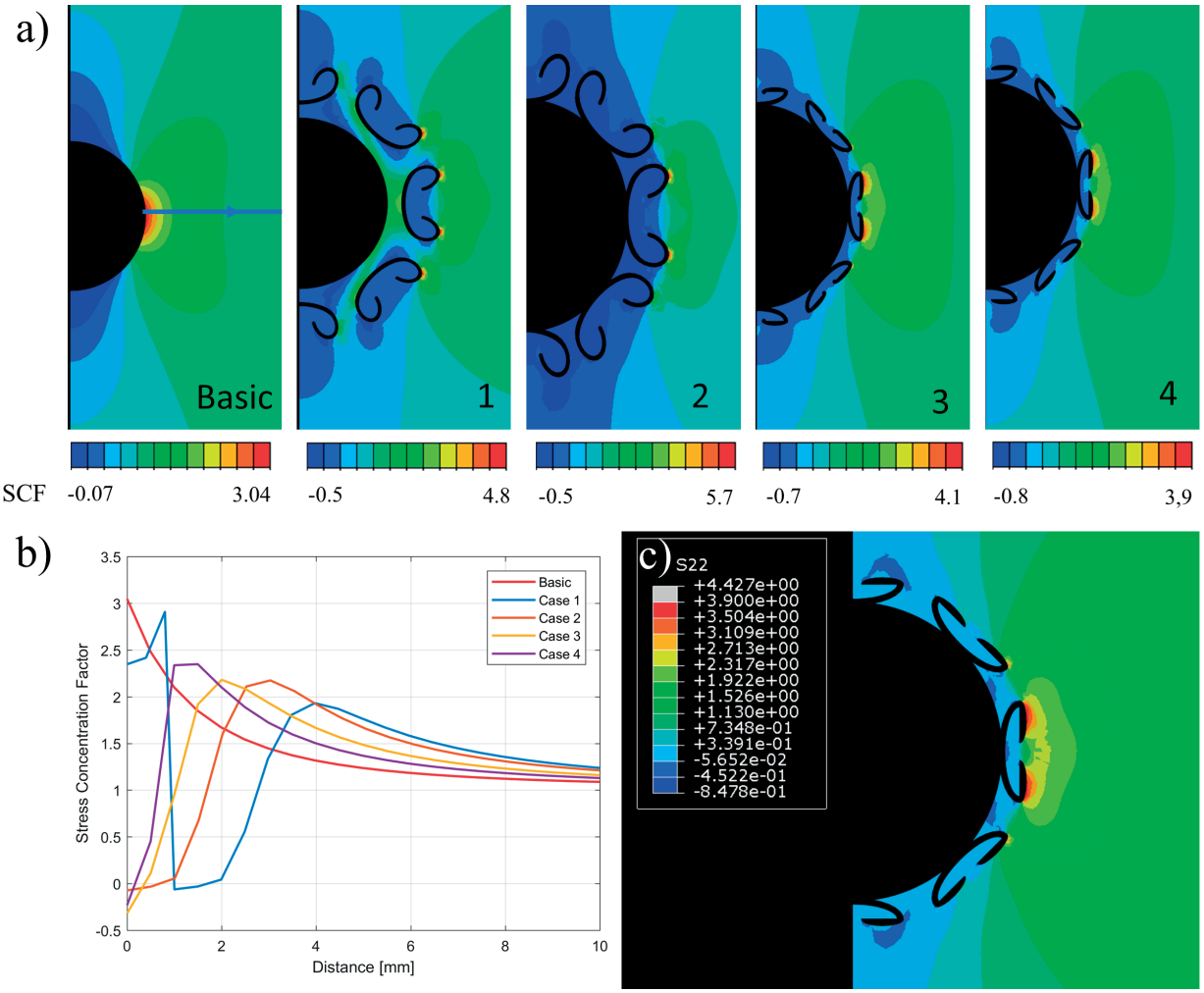


Figure 3 *(a) Feature shape evolution; (b) Stress distribution along the "critical direction"; (c) particular view of the stress field near the chosen feature shape. (SCF means "Stress Concentration Factor")*

finest mesh dimension from 2 to 0.1 mm. The 0.5 mm value is chosen as a compromise between the computational effort and being in a surely converged condition. Moreover, the material considered is a generic stainless steel having Young Modulus E = 193 GPa and Poisson's ratio ν = 0.33, as used in [15]. To reduce the computational effort, a symmetric boundary condition has been applied. The model is subjected to a remote stress with magnitude of S = 1 MPa; in such a way the results can be considered directly as Stress Concentration Factor values.

As previously anticipated, a C-shaped void pattern is introduced with the aim of redistributing the stress field around the hole in an infinite plate. Grima et al. [13] claim that the S-shaped void induces a soft mode of deformation in the surrounding material based on rotation of the material around. However, it is considered valuable to exploit the polar distribution, instead of the Cartesian. Clearly, a C-shaped element has a symmetry axis that makes it more valuable for the polar distribution than the S-shaped one (anti-symmetric). The definition of an optimized pattern passes through a series of design steps in which the feature shape, as well as its position relatively to the hole is modified with the aim of reducing the stress concentration factor along the *critical direction*, that is the one perpendicular to the load direction. Clearly, once a suitable feature-to-hole relative position has been defined, the feature design improvements objective consists of a compromise between the minimization of the stress along the critical direction and the induced stress concentration at the feature extremes. In Figure 3a, the feature shape evolution series is shown, as well as the stress distribution along the direction of interests in Figure 3b. In Figure 3c, a magnified view of the stress field around the final feature design is presented.

4. Ligament length

It is well known that introducing geometrical discontinuities in mechanical components generates stress concentration effects that are detrimental for the mechanical performances. The adoption of the C-Shape void pattern showed, in the previous section, the possibility of managing the stress distribution, meaning that a compromise must be reached. The modification of stress field clearly moves the potential crack initiation point from the hole boundary to the feature tip. Furthermore, considering that numerical analysis shows

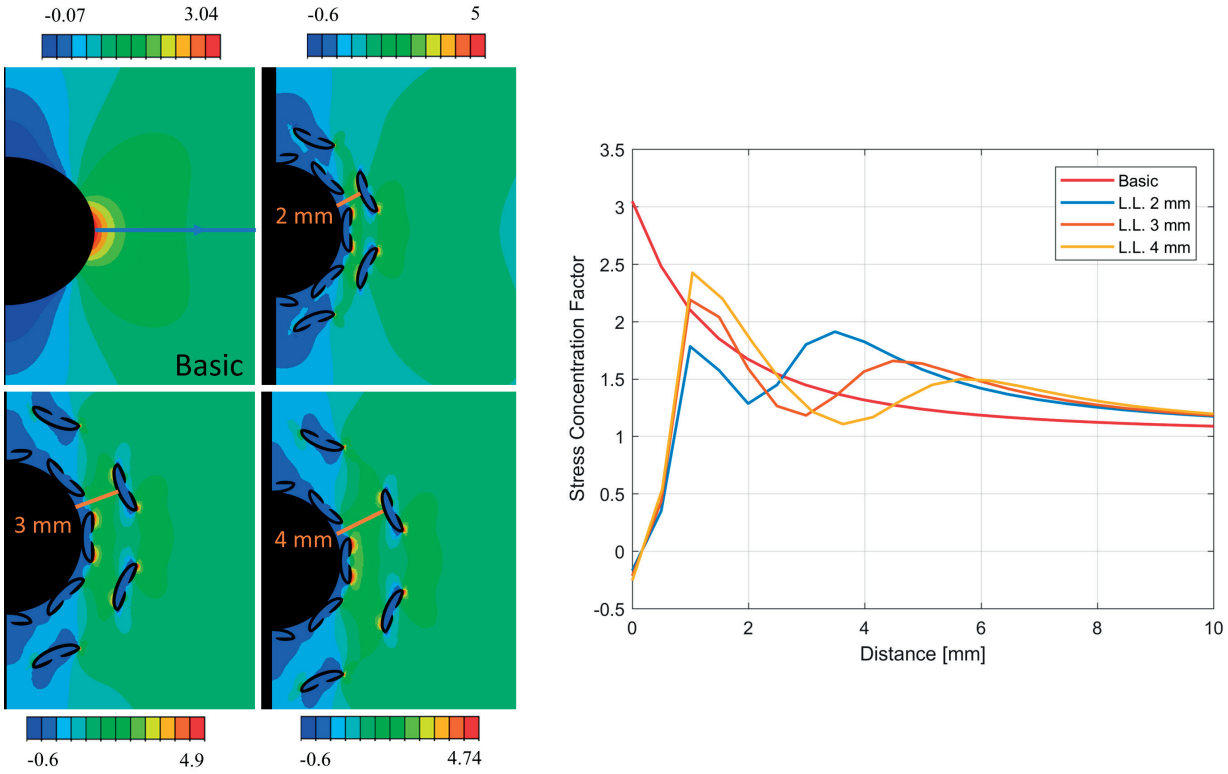


Figure 4 *Stress distribution along the "critical direction" for two-row- pattern case (considering the S22 as shown in Figure 3c)*

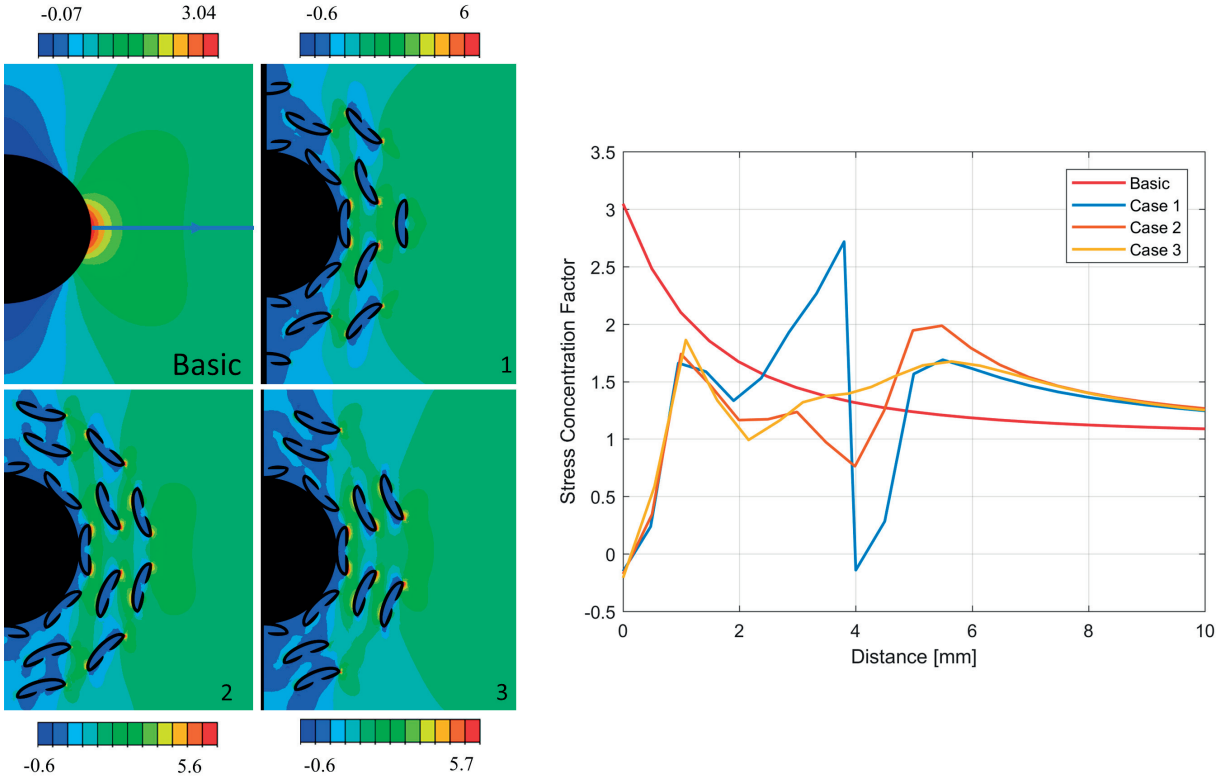


Figure 5 *Stress distribution along the "critical direction" for three-row pattern case (considering the S22 as shown in Figure 3c)*

the stress concentration values slightly higher than for the case of a simple holed plate, the presented solution can be more susceptible to crack fatigue damage. Clearly, to manage the crack propagation phenomena it is not possible to adopt a single row pattern. In [15] the so-called Ligament Length is used for representing the minimum distance between two contiguous pattern features. Even in the present case, it is interesting to determine the Ligament Length, defined as a distance between the two contiguous rows, whose value helps to reduce the stress concentration factor along the

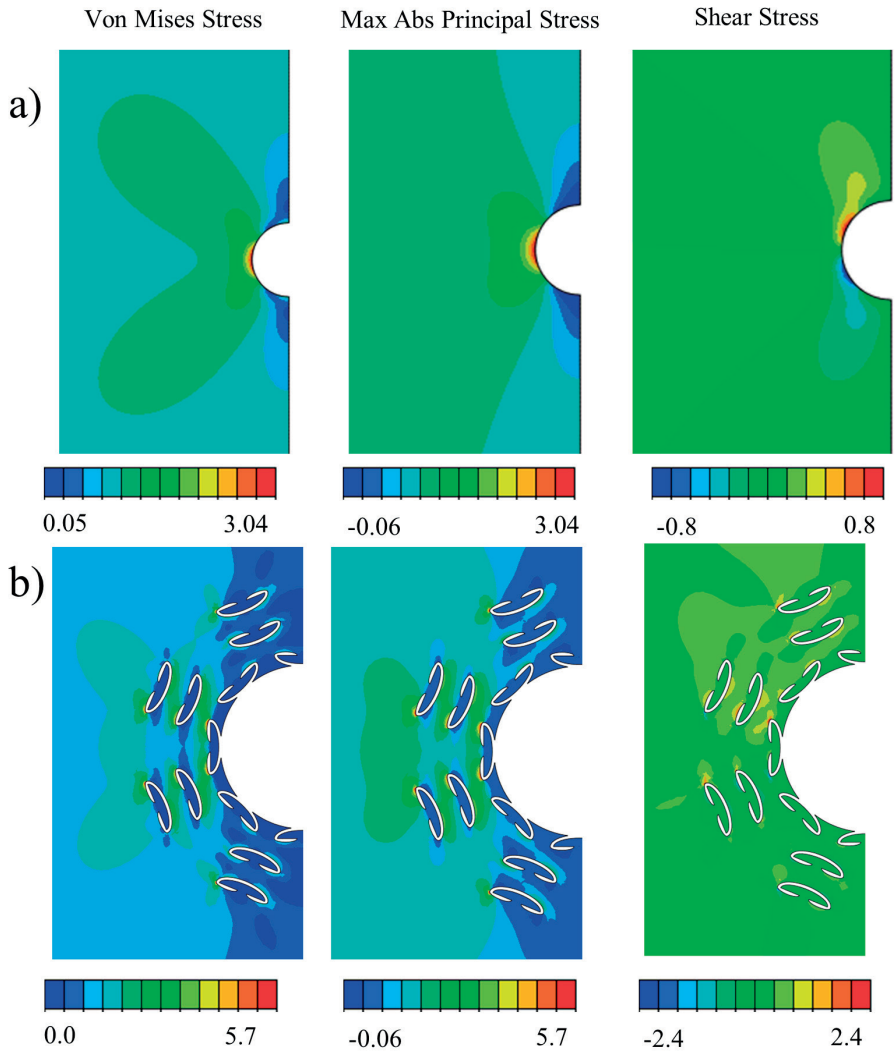


Figure 6 *Results in terms of Von Mises Stress, Max. Abs. Principal Stress and Shear Stress for: (a) the basic case; (b) the chosen pattern*

critical direction. Three cases have been considered adopting a Ligament Length of 2 mm, 3 mm and 4 mm, as shown in Figure 4. The resulting stress profiles shows that the most convenient Ligament Length is 2 mm since this solution provides the more effective stress concentration reduction.

5. Pattern improvement

Considering the stress profile obtained in the case of a single row pattern, comparison to one obtained with two-row pattern, it appears that the introduction of additional pattern rows contributes the redistribution of the stress field and reduce the maximum stress value reached. For this reason, a third row is added even adopting a Ligament Length of 2 mm. Looking at Figure 5, immediately appears that the resulting stress profile is affected by the features position, as already assessed in Section 3. For this reason, the three different cases, shown in Figure 5, are studied. The most effective stress reduction is reached adopting the case number 3, with a final Stress Concentration Factor reduction of nearly 38 % with respect to the initial value.

6. Evaluation of the whole model for the chosen solution

Despite the results obtained relatively to the critical direction, it is clear that the introduction of geometrical discontinuities introduces regions of high stress concentration requiring an evaluation with respect to the basic case of a single hole. In order to make this comparison, the Von Mises Stress, the Maximum Principal Stress, as well as the Shear Stress are evaluated for the chosen solution (presented in Figure 6a for the basic case, Figure 6b for the chosen pattern). A great difference in the stress distribution can be observed, as well as different areas of the stress concentration. Such stress concentration area is very small at the extremes of the single pattern feature, leading to a strong stress gradient. Despite the value itself of the stress concentration introduced by the void pattern, for the scope of this work the effort was focused to understanding the influence of the void pattern relatively to the stress along the critical direction. Consequently, the introduced C-shaped feature must be optimized in order to reduce the induced stress concentration at its extremes as much as possible.

7. Conclusion

Despite the several studies found in literature about introducing the auxiliary holes to provide a global stress concentration factor reduction by redistributing the load as an equivalent ellipse, the present study successfully introduced the C-shaped void patter to manage the stress field and reaching a reduction of the stress concentration factor of about 38 %. Clearly, this result is concerned with the "critical direction" that is well known in literature as the one perpendicular to the loading direction. However, the introduction of the strange shaped discontinuities generates stress concentration effects at the tips of these features, leading to a stress concentration increment of about 1.8. Simultaneously, unlike in the basic case, in the chosen pattern the stress gradient is very high, meaning that the stress increment is limited to a small area. The potential of introducing a structured pattern instead of a single feature stands in the possibility of "controlling" the crack propagation phenomenon, or at least to provide a method to detect the propagating crack before that reaches the critical value.

The present study introduces several research possibilities, in particular it is interesting to investigate whether the fatigue resistance is improved by introduction of the porous structure, as well as the application of such a solution to a defined case study. In addition, both feature design, as well as the pattern definition must be optimized for reaching the two main objectives: reducing the induced stress concentration factor at the feature tip and making the performance improvement independent of the load case.

Acknowledgements

The author wants to thank prof. Mario Guagliano for his support during the development of this work.

References

[1] DHIR, S. K.: Optimization in a Class of Hole Shapes in Plate Structures. Journal of Applied Mechanics, 48(4), 905-908, 1981.

[2] WU, Z.: Optimal Hole Shape for Minimum Stress Concentration Using Parameterized Geometry Models. Structural and Multidisciplinary Optimization, 37(6), 625-634, 2009. https://doi.org/10.1007/s00158-008-0253-4

[3] YOUNIS, N.T.: Assembly Stress for the Reduction of Stress Concentration. Mechanics Research Communications, 33(3), 837-845, 2006.

[4] WU, H.-C., MU, B.: On Stress Concentrations for Isotropic/Orthotropic Plates and Cylinders with a Circular Hole. Composites Part B: Engineering, 34(2), 127-134, 2003.

[5] KO, W. L.: Stress Concentration around a Small Circular Hole in the HiMAT Composite Plate. NASA Technical Memorandum 86038, 1-15, 1985.

[6] TOUBAL, L., KARAMA, M., LORRAIN, B.: Stress Concentration in a Circular Hole in Composite Plate. Composite Structures, 68(1), 31-36, 2005. https://doi.org/10.1016/j.compstruct.2004.02.016

[7] DAN-JUMBO, E., KELLER, R., CHAN, W. S., SELVARAJ, S.: Strength of Composite Laminate with Multiple Holes. Proceedings of 17th International conference on composite materials (CCM-17), United Kingdom, 27-31, 2009.

[8] MEGUID, S. A.: Finite Element Analysis of Defence Hole Systems for the Reduction of Stress Concentration in a Uniaxially-Loaded Plate with Two Coaxial Holes. Engineering Fracture Mechanics, 25(4), 403-413, 1986. https://doi.org/10.1016/0013-7944(86)90254-7

[9] AKOUR, S.N., AL-HUSBAN, M., NAYFEH, J.F.: Design and Optimization of Defense Hole System for Hybrid Loaded Laminates. Technology Engineering and Management in Aviation: Advancements and Discoveries, IGI Global, 151-160, 2012.

[10] JAIN, D. N.: The Reduction of Stress Concentration in a Uni-Axially Loaded Infinite Width Rectangular Isotropic/Orthotropic Plate with Central Circular Hole by Coaxial Auxiliary Holes. IIUM Engineering Journal, 12(6), 2012.

[11] JINDAL, U. C.: Reduction of Stress Concentration around a Hole in a Uniaxially Loaded Plate. The Journal of Strain Analysis for Engineering Design, 18(2), 135-141, 1983.

[12] PROVIDAKIS, C. P., SOTIROPOULOS, D. A., BESKOS, D. E.: BEM Analysis of Reduced Dynamic Stress Concentration by Multiple Holes. International Journal for Numerical Methods in Biomedical Engineering, 9(11), 917-924, 1993. https://doi.org/10.1002/cnm.1640091108

[13] GRIMA, J. N., GATT, R.: Perforated Sheets Exhibiting Negative Poisson's Ratios. Advanced Engineering Materials, 460-464, 2010. https://doi.org/10.1002/adem.201000005

[14] TAYLOR, M., FRANCESCONI, L., GERENDAS, M., SHANIAN, A., CARSON, C., BERTOLDI, K.: Low Porosity Metallic Periodic Structures with Negative Poisson's Ratio. Advanced Materials, 26(15), 2365-2370, 2014. https://doi.org/10.1002/adma.201304464

[15] JAVID, F., LIU, J., RAFSANJANI, A., SCHAENZER, M., PHAM, M. Q., BACKMAN, D., YANDT, S., INNES, M. C., BOOTH-MORRISON, CH., GERENDAS, M., SCARINCI, T., SHANIAN, A., BERTOLDI, K.: On the Design of Porous Structures with Enhanced Fatigue Life. Extreme Mechanics Letters, 13-17, 2017. https://doi.org/10.1016/j.eml.2017.08.002

Katarzyna Szwedziak - Tomasz Lusiak - Zaneta Grzywacz - Kacper Drozd*

NUMERICAL CFD ANALYSIS OF AN AERODYNAMIC HEAD COVER OF A ROTORCRAFT MOTOR

Autogyros can become an alternative for the use of rotorcrafts in various fields of life, including agroforestry. They have better economic performance than helicopters, owing to, among other things, the presence of a bearing rotor. Most autogyros also have other advantages in terms of no need for the compliance with stringent regulatory regulations – with respect to new constructions, lower combustion, noise and emissions of toxic elements. The cover of the bearing rotor head is an important element of rotorcrafts, which demonstrates that aerodynamics plays an important role in aerodynamic designs. Therefore, in this article, air flow model testing is carried out for two types of the bearing rotor blades of an autogyro with and without a cover using the ANSYS Fluent program. An aerodynamic drag analysis was also performed.

Keywords: *CFD numerical analysis, autogyro, aerodynamic cover, rotor head*

1. Introduction

Aviation is one of the most advanced technical disciplines used in many fields of life, including agriculture. A large number of experts and resources allocated to the technological development of this field of knowledge enable the establishment and development of the new and unconventional ideas and the popularization of their applications.

The intensive development of helicopters has been slowing down due to the limitations of the operation of rotors at the highest revolutions [1]. The development of the so-called "gyrocopters", which combine the construction of autogyros, helicopters and planes, is an alternative direction for rotorcrafts. They have very desirable options for vertical takeoff, landing and drift.

Autogyros, economically better than helicopters, can also be a good choice. The advantage is that their bearing rotors are driven by the incoming air, causing the effect of autorotation, which means that, unlike helicopters, they are not propelled by the engine on the fly. Every autogyro has a motor, of course, but it serves to propel the bearing blade exactly as in the case of motorcycles [2], [3], [4].

Autogyros most commonly make use of piston engines and it is the use of piston engines that makes them cheaper to operate. The strength of the machines is also the possibility to execute a short take-off and landing, which eliminates the need for a hardened and long runway. Autogyros contrast with most of the mentioned rotorcrafts, also with reference to the legalization of new constructions. Unlike the previously mentioned machines, most autogyros are not subject to rigorous regulations.

The four-person PAV (Personal Air Vehicle) by Carter Copter is an example of such means of travel [5], [6]. The highly innovative PAV consumes three times less fuel than the standard helicopters. What is more, the model is characterized by low emission of toxic elements and low noise during operation. It seems that the presented strengths provide a favorable future for this type of a flying craft.

The cover of the bearing rotor is an important element affecting the performance of autogyros - its aerodynamics plays an important role in the design of aircrafts. The shape of the mast significantly influences traveling at high speeds. Therefore, they are rarely used in the low-speed rotorcrafts. The structures of the cover of bearing rotor heads, depending on the design, are full- or half-covers. Therefore, it seems that the analysis of this element can affect the direction of production and thus improve the performance which, in turn, can also contribute to the increase of the popularity of autogyros in Poland and in the world [1], [7], [8].

2. Study aim

The aim of the study was to perform the air flow model testing for two types of autogyro bearing rotor heads using the ANSYS Fluent program and to perform an aerodynamic drag analysis.

3. Assumptions and scope of simulation tests

The scope of the model study included four airflow simulations for two types of the bearing rotor heads of a rotorcraft - with and without a cover. The simulations were made for two air velocities of 140 and 200 km/h. For each velocity, an individual angle of the bearing rotor adjustment was selected. Different parameters of the rake angle of the rotor are due to the fact that for different velocities different bearing rotor rake angles are used. The rake angles of the bearing rotor heads for each velocity are determined in Figure 1 [9]. It was decided to refer to a graph showing the parameters of the classic autogyro rotary engine. The rotor rake

* [1]Katarzyna Szwedziak, [2]Tomasz Lusiak, [1]Zaneta Grzywacz, [2]Kacper Drozd
[1]Department of Biosystems Engineering, Faculty of Production Engineering and Logistics, Opole University of Technology, Poland
[2]Department of Thermodynamics of Fluid Mechanics and Air Drives, Mechanical Engineering Faculty, Lublin University of Technology, Poland
E-mail: z.grzywacz@po.opole.pl

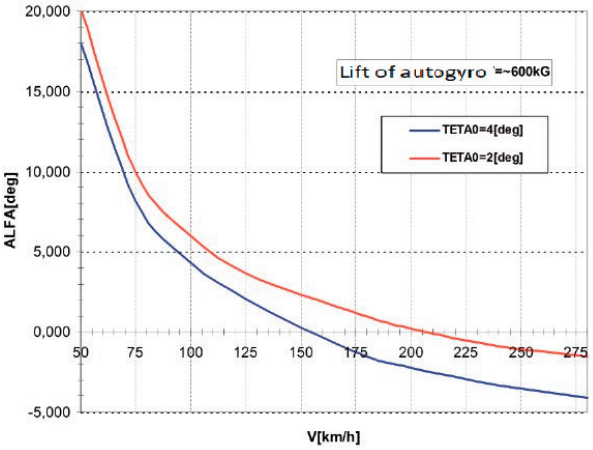
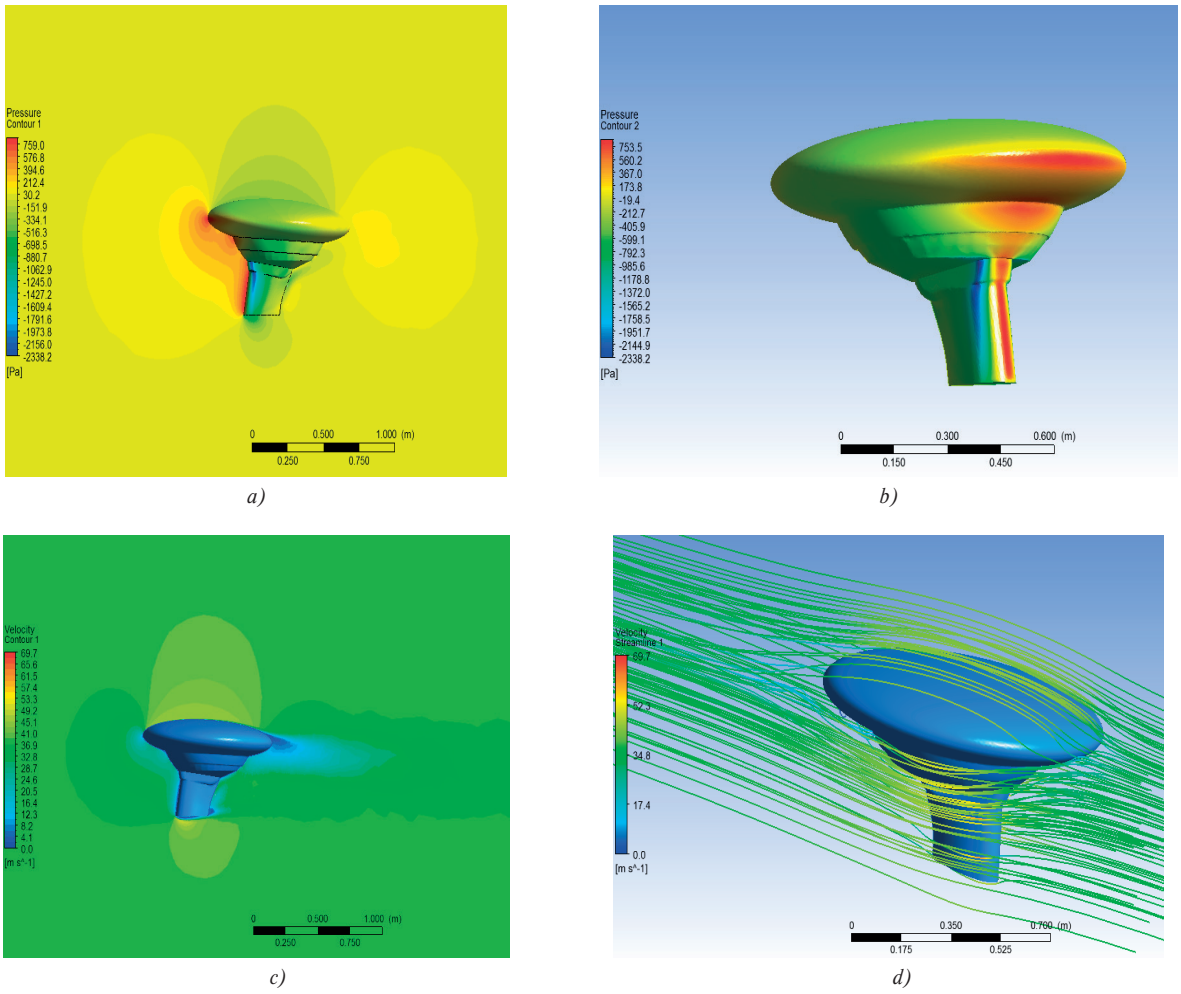


Figure 1 *Graph of the rake angles of the bearing rotor head as a function of the autogyro velocity for two blade position angles [9]*

Table 1 *Bearing rotor rake angles for different flight speeds*

Speed	Rake angle
140	3°
200	0°



a) *b)* *c)* *d)*

Figure 2 *Calculation results for model I. a) pressure gradient for the head rake angle 3°; b) Pressure gradient for the head rake angle 3°; c) velocity gradient for the head rake angle 3°; d) current lines for the head rake angle 3°*

angles were adjusted to the blade position at an angle of 2 degrees. The selected rake angles are shown in Table 1.

For model testing, the following assumptions were made:

- flow through the geometry is incompressible; this is due to the Mach number not exceeding 0.3,
- the phenomenon under consideration is stationary,
- turbulent airflow was assumed,
- aluminum was assumed as the material for the autogyro from the Ansys Fluent material base,
- it was assumed that the calculation solver that is used to solve the fluid mechanics equations is pressure-based,
- the two-tone turbulence model k-ω SST was assumed,
- the second-order interpolation for energy equations, momentum, kinetic energy of turbulence, turbulence dissipation energy was assumed.

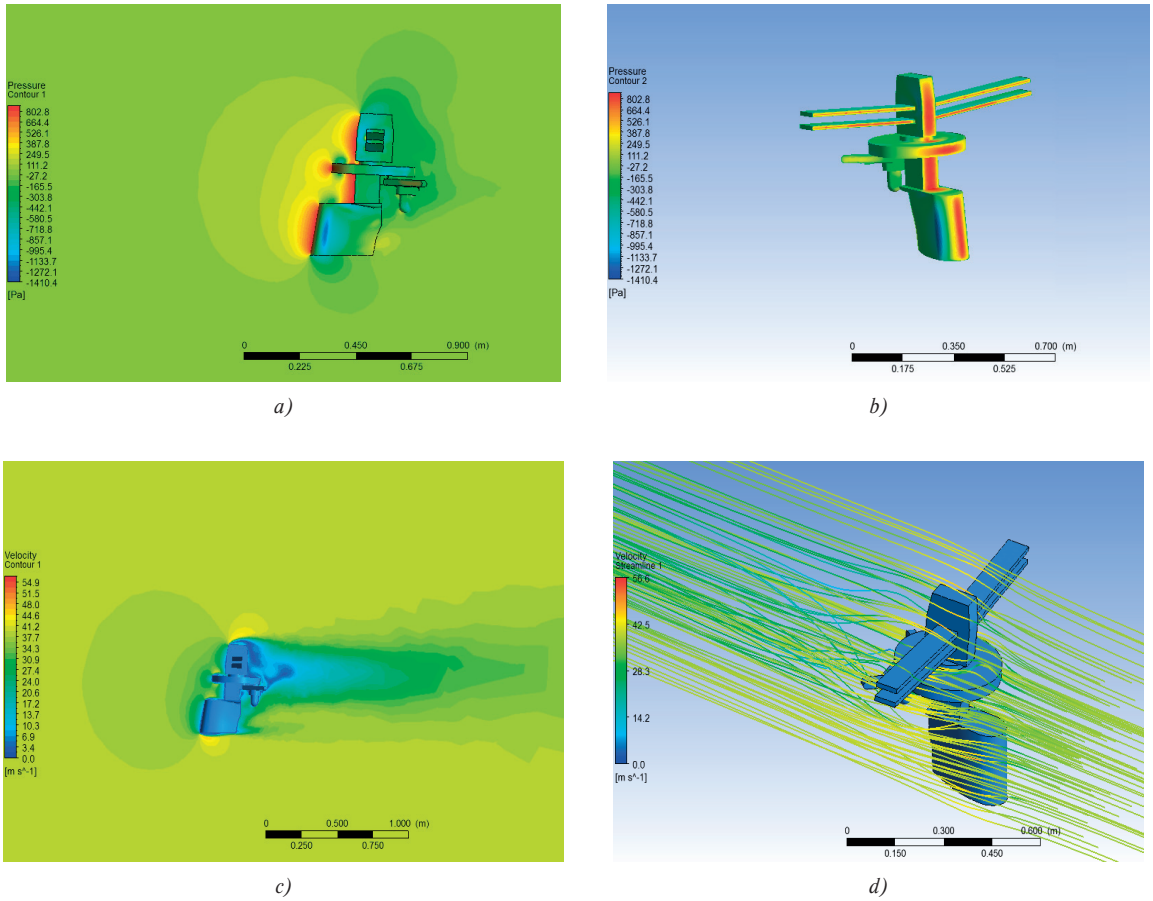


Figure 3 *Calculation results for model II: a) pressure gradient for the head rake angle 3°; b) pressure gradient for the head rake angle 3°; c) velocity gradient for the head rake angle 3°; d) current lines for the head rake angle 3°*

In development of the assumptions and the scope of testing, literature [9] and [10] have been applied, as well as the general state of knowledge on airflow simulation and its effects on the body.

4. Analysis of the test results

The analysis of the results is based on the illustration calculations concerning the pressure gradients, velocities and currents running through a computational domain. The current lines represent the flow behavior. The speed of the stream is shown in colors. Pressure gradients are shown in individual areas - in the symmetry plane of the heads and on their surface. The values of the characteristic aerodynamic parameters were also analyzed.

The results of the calculations are presented in order from the minimum velocity tested (140 km/h) to the maximum velocity tested (200 km/h), starting with the head with an aerodynamic cover.

4.1 Model I

Figure 2 shows the calculation results for the head with an aerodynamic cover. The results show a model with an air velocity of 140 km/h and a rake angle of the bearing rotor head of 3°.

After the preliminary analysis of results obtained for Model I, it can be seen that the face surface of the head is characterized by the largest areas of high pressure, while the side walls of the mast exhibit areas of vacuum.

4.2 Model II

Figure 3 shows the results of the calculation for the standard head. The results show a model with an air velocity of 140 km / h and a rake angle of the bearing rotor head of 3°.

Based on the results of model II, it can be seen that there are high pressure areas on the flat face areas. The values are higher than in the model with an aerodynamic cover.

4.3 Model III

Figure 4 shows the calculation results for the head with an aerodynamic cover. The results show a model with an air velocity of 200 km/h and a rake angle of the bearing rotor head 0°.

When analyzing the results of calculations for the Model III, it can be seen that the highest pressure is present on the face surfaces of the head aerodynamic cover. The value of the pressure there is twice as great as in the case of the first model. The values of the pressures on the side walls of the mast showed the same characteristics.

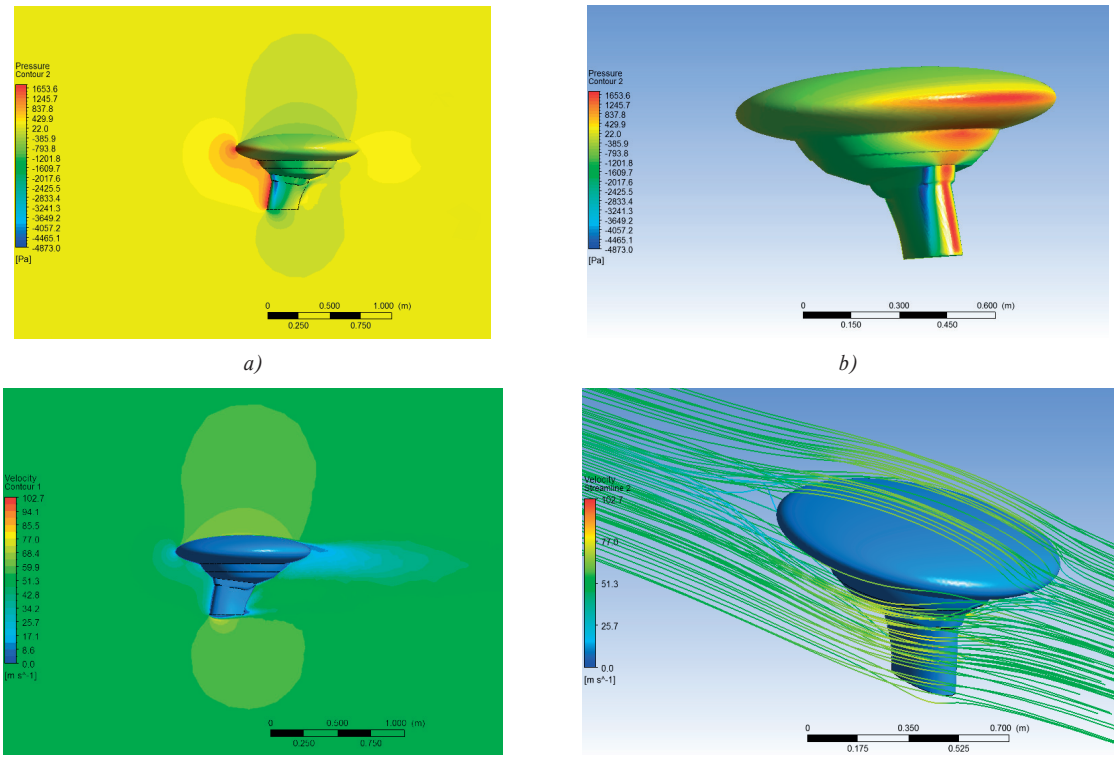


Figure 4 *Calculation results for model III: a) pressure gradient for the head rake angle 0°; b) pressure gradient for the head rake angle 0°; c) velocity gradient for the head rake angle 0°; d) current lines for the head rake angle 0°*

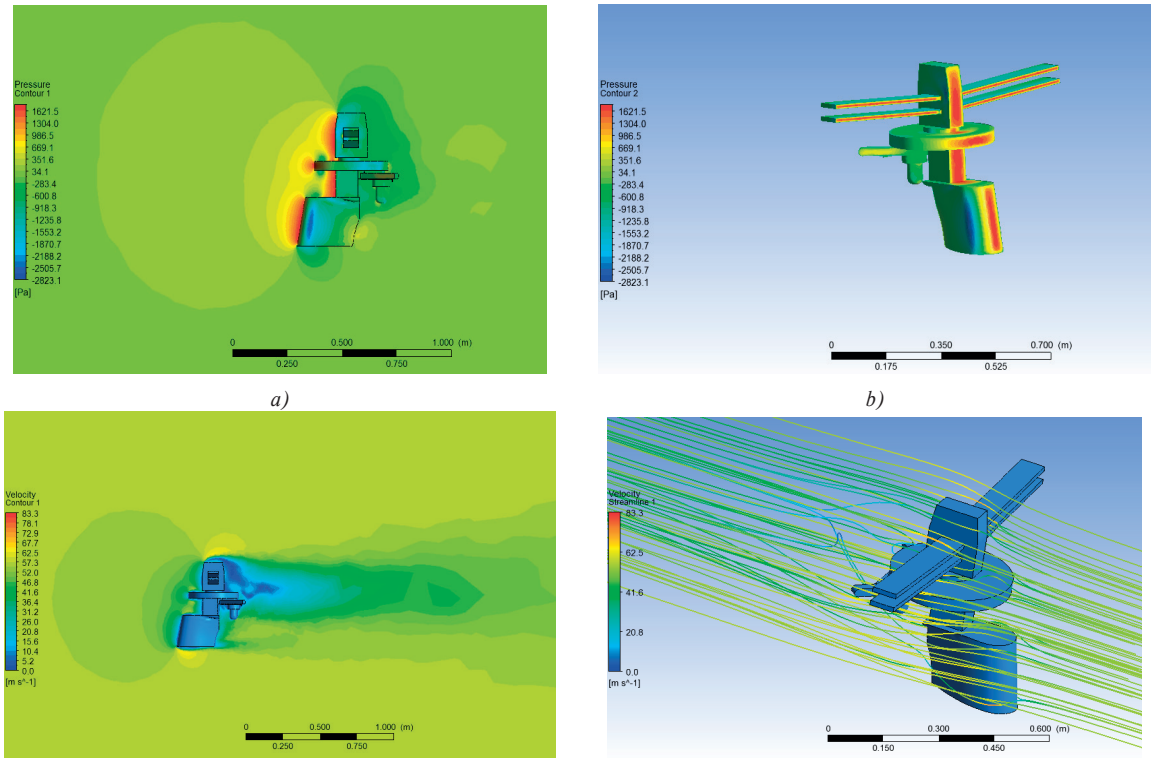


Figure 5 *Calculation results for model IV: a) pressure gradient for the head rake angle 0° b) pressure gradient for the head rake angle 0° c) velocity gradient for the head rake angle 0° d) current lines for the head rake angle 0°*

Table 2 *Values of the aerodynamic parameters of the analyzed heads*

rake angle of the head	F_x	F_y	F_z	C_x	C_y	C_z
without a cover / 3°	71.435	-1.786	-3.584	0.743	-0.019	-0.037
with a cover / 3°	36.468	-1.577	49.578	0.142	-0.006	0.193
without a cover / 0°	149.746	-0.807	-11.253	0.844	-0.005	-0.063
with a cover / 0°	75.280	4.790	45.479	0.145	0.009	0.087

F_x - aerodynamic resistance,
F_z - lifting force,
F_y - aerodynamic lateral force
C_x - Spring of aerodynamic resistance,
C_z - Youth supporting,
C_y - aerodynamic lateral force.

4.4 Model IV

Figure 5 shows the results of the calculation for the standard head. The results show a model with an airflow velocity of 200 km/h, while the rake angle of the bearing rotor is 0°.

The analyzed model IV has recorded the high pressure areas that occur on the front surfaces. The overpressure value of the last model is twice that of the second model, but has a value similar to the model analyzed for the same flight speed. After analyzing the results obtained for the two types of the bearing rotor heads and the two speeds of the autogyro, the highest pressure values were found on the front areas. It was also observed that the pressure gradient for flat surfaces is much larger than for the spherical areas of the aerodynamic cover. The above illustrations also show areas of vacuum on the mast side surfaces. Low pressure is present on the mast walls and is a result of the Bernoulli's law influence. Reducing the flow surface of the model results in an increase in air velocity that results in a vacuum in the areas. This phenomenon can also be seen in the area above the rotor head cover. Since the geometry of the aerodynamic cover is similar to the airplane profile, the distribution of pressure fields in the areas is similar. For the models analyzed, the value of aerodynamic forces was measured. The readings were used for calculations of the most important aerodynamic coefficients which are presented in Table 2.

The results show the beneficial effect of the aerodynamic cover. The highest coefficient of aerodynamic drag was recorded for the standard head at the rake angle of the bearing rotor head 0°. The bearing rotor head with an aerodynamic cover has more than four times the lower value of the aerodynamic drag coefficient. An additional advantage of the aerodynamic cover design is the positive value of the lift force coefficient. The value of the lift force coefficient for the bearing rotor head rake angle changed by 3° increased more than twice.

5. Summary

One of the most important goals in aircraft design is to achieve satisfactory aerodynamic performance. The strive for improving and increasing the quality can be achieved through changing the geometry of aircrafts.

The results obtained after four airflow simulations for two types of bearing rotor heads of a rotorcraft (with and without a cover) show that the heads with aerodynamic covers have more favorable aerodynamic properties. The aforementioned heads have a four times lower aerodynamic drag coefficient, which significantly affects the aircraft's flying performance.

The presented studies provide a good basis for conducting further simulations. In the next stages, it will be possible to perform a simulation using the whole autogyro model or a head with rotating blades. The simulation studies under discussion can be extended by both a vertical and diagonal flight of an aircraft.

References

[1] ABLAMOWICZ, A., NOWAKOWSKI, W.: Fundamentals of Aerodynamics and Mechanics (in Polish). WKiL, Warszawa, 1980.
[2] CZYZ, Z., LUSIAK, T., MAGRYTA, P.: Numerical Studies on the Impact of Wind Turbines on the Aerodynamic Characteristics of the Windmill (in Polish). PIL, Warszawa, 2013.
[3] CIESLAK, S.: Study of the Impact of the General Stroke Kata of the Rotor on the Properties of Volatile Gyroplane (in Polish). Warszawa, 2015.
[4] SZABELSKI K., JANCELEWICZ B., LUCJANEK W.: Introduction to the Construction of Helicopters (in Polish). WKiL, Warszawa, 1995.
[5] JURIEW, B. N.: Aerodynamics of Snails and Helicopters (in Polish). Publisher of the Ministry of National Defense, 1959.
[6] SZCZEPANIK, T., DABROWSKA, J.: Windmills, as the Anticipated Direction of Development of Centrifuges in the 21st Century (in Polish). PIL, Warszawa, 2009.
[7] BENO, L., BUGAJ, M., NOVAK, A.: Application of RCM Principles in the Air Operations. Communications – Scientific Letters of the University of Zilina, 7(2), 20-24, 2005.
[8] DELEGA, M.: IL-28 Windmill Head for Vertical Start (in Polish) . PIL, Warszawa, 2009.

[9] KAZIMIERSKI, Z.: Basics of Mechanics of Liquids and Computer Methods of Flow Simulation (in Polish). WIST, Lodz, 2004.
[10] STALEWSKI, W.: Calculation Analysis of Aerodynamic Properties of the Windmill Rotor in the Established Flight Condition (autorotation) (in Polish). PIL, Warszawa, 2011.

Tomasz Linek - Tomasz Tanski - Wojciech Borek*

INFLUENCE OF SURFACE ROUGHNESS ON THE CAVITATION WEAR OF P265GH AND X2CrNi18-9 STEEL CAVITATION GENERATORS

The aim of this paper was to determine the effect of surface roughness of cavitation generators made from two different materials: P265GH steel, with a ferritic-pearlitic structure, and X2CrNi18-9 (304L) steel with an austenitic structure on the mass loss and cavitation wear. Cavitation generators were tested in the conditions of cavitation wear environment continuously for 500 PMHs in a specially designed and constructed author's stream and flow device. Based on the carried out experiments was confirmed that the highest mass loss - 0.1752 g is seen for a sample of P265GH steel wet sanded with paper with the grain size of 1000. The smallest mass loss was recorded for the cavitation generator made of X2CrNi18-9 (304L) steel, sanded with sandpaper with the grain size of 2500. Certainly, the smallest number of cavitation wear effects was found for a cavitation generator made of austenitic steel X2CrNi18-9 (304L). Few places were identified based on macroscopic photographs, especially near the edges of the straight-through openings, but their number was much smaller than for the ferritic-pearlitic steel, which is associated most of all with the properties of austenitic chromium - nickel steel.

Keywords: *cavitation, cavitation wear, cavitation generators, surface roughness, mass loss, P265GH steel, X2CrNi18-9 (304L) steel*

1. Introduction

The cavitation destruction phenomenon occurs in industry wherever a quickly flowing working medium, e.g. fluid or saturated steam, is moving due to the influence of differential pressure in a technological process. The basic wear of constructional materials working in a cavitation environment consists of an abrupt change of pressure of the fluid rinsing a given surface of a constructional element of machines, devices and industrial apparats. The key methods of preventing a harmful effect of the cavitation destruction phenomenon of constructional elements are such constructional solutions as: optimum parametrisation of machine parts' geometry and the designing of a streamlined surface shape, as well as additional technological solutions such as: selection of materials and application of the adequate surface engineering treatment [1], [2], [3], [4], [5].

The purpose of this work was to examine the effect of surface roughness of the following investigated steels, i.e. P265GH and X2CrNi18-9 (304L) steel, on the mass loss and cavitation wear of cavitation generators subject to operation in a stream and flow device. Constructional elements, which were designed and modelled in advance in Computational Fluid Dynamics (CFD) numerical software, were made within the investigations [1], [2], [3]. The constructions meeting the majority of geometrical criteria were selected for further investigations in real conditions in a stream and cavitation device with the flow character of work. The selected constructional elements were made of two structural steels, the first one of P265GH steel with a ferritic-pearlitic structure, and the second of X2CrNi18-9 (304L) steel with an austenitic structure. The author of this work will utilise the results of the investigations and the conclusions found after the operation of steel elements for 500 operating hours [PMH - Productive Machine Hour] to implement appropriate protection of surface against cavitation wear, choosing from the available surface engineering technologies, e.g. Physical Vapour Deposition (PVD) or others [6], [7], [8], [9], [10], [11].

2. Research material and methodology

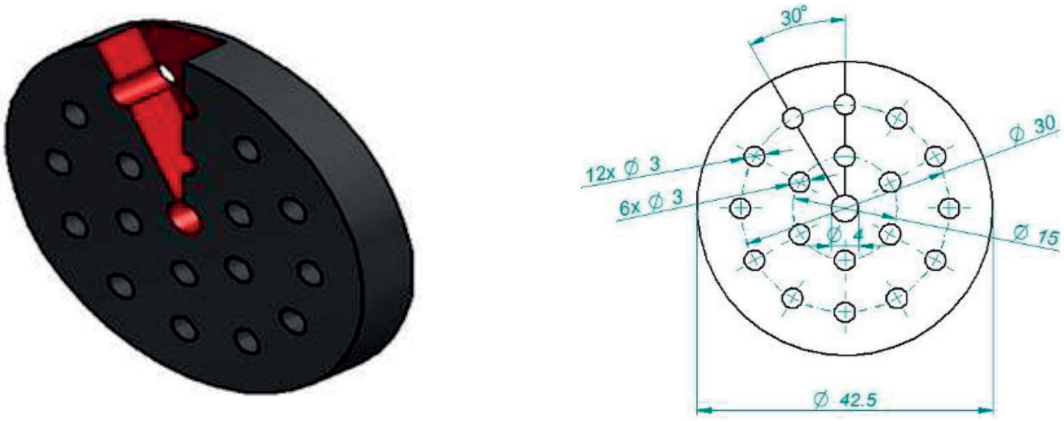
Following the pre-optimisation of cavitation generator dimensions and shape [1], two grades of steel were used for the purpose of investigations aimed at determination of potential places of intensive cavitation wear. The first grade is steel commonly used for pressure devices working at elevated temperatures, P265GH, with a ferritic-pearlitic structure, and the second is from a group of stainless steels - chromium - nickel X2CrNi18-9 (304L) steel with an austenitic structure. The P265GH steel - due to its unlimited availability and attractive, low market prices - is used for constructing heat distribution devices and heating devices, and for less important constructional parts. The X2CrNi18-9 (304L) steel, which exceeds the price of P265GH steel five times, is used for production of devices, apparats and fittings in the chemical, food, power and petrochemical industry and for constructional elements in the aviation and shipbuilding sector. The chemical composition of the structural steels, tested in the conditions of cavitation wear, is presented in Table 1.

Constructional elements made of the two steels, with their shape and dimensions selected based on the analysis of results of numerical simulations in ANSYS FLUENT software, described in

* [1,2]Tomasz Linek, [2]Tomasz Tanski, [2]Wojciech Borek
[1]Odlewnie Polskie S.A, Starachowice, Poland
[2]Institute of Engineering Materials and Biomaterials, Silesian University of Technology, Gliwice, Poland
E-mail: tomasz.linek@odlewniepolskie.pl

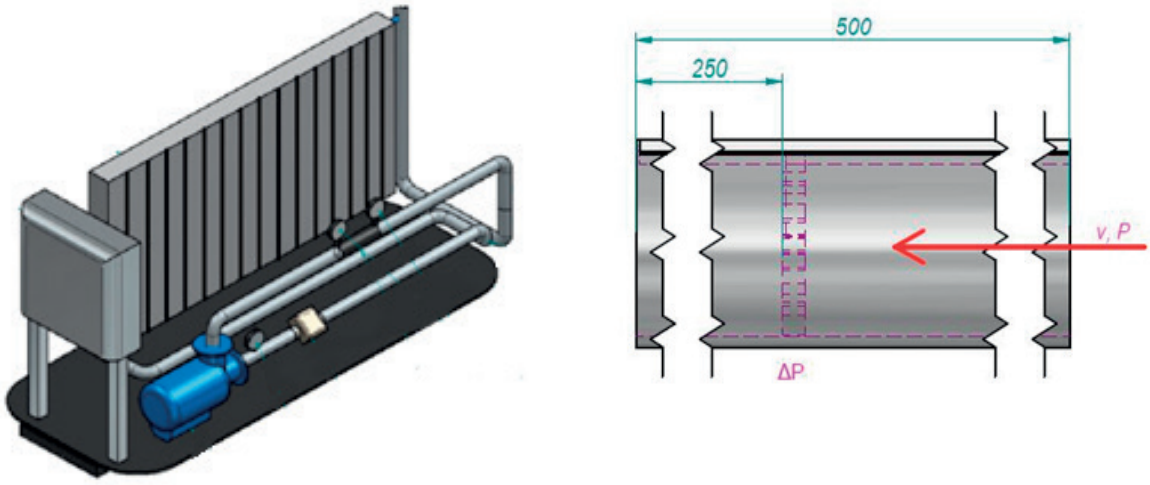
Table 1 *Chemical composition of the structural steels tested in the conditions of cavitation wear: P265GH by PN-EN 10028:2010; X2CrNi18-9 by PN-EN 10088 [mass fraction, %]*

Chemical composition		C	Mn	Si	Al	Cr	Ni	Cu	Ti	N	S	P
Steel		[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
P265GH	max	-	-	0.4	-	0.3	0.3	0.3	0.03	0.012	-	-
		0.16	0.99	0.23	0.047	0.027	0.013	0.026	0.001	0.003	0.008	0.019
X2CrNi18-9	min.					17.50	8.00					
(304L)	max	< 0.03	< 2.0	< 1.0	-	19.50	10.50	-	-	< 0.11	< 0.045	< 0.015



a) *b)*

Figure 1 *Cavitation generator model's dimensions and shape selected based on the analysis of results of numerical simulations in ANSYS FLUENT software; cavitation generator thickness of 5 [mm], relative clearance of Pp=11.1 [%]*



a) *b)*

Figure 2 *Model of a stream and flow device generating a cavitation environment; a) isometric diagram of the device, testing and measuring system; b) simplified computer model of the cavitation generator location with the medium (water) flow direction marked*

[1], shown in Figure 1, were tested in the conditions of cavitation wear continuously for 500 PMHs in a specially designed and constructed author's stream and flow device (Figure 2) generating a cavitation environment. The detailed process parameters are presented and described in [1].

The purpose of this work was to examine the effect of surface roughness of the investigated steels on the mass loss and cavitation wear of cavitation generators in a stream and flow device. Due to a large number of samples and, most of all, a long exposure time in a stream and flow device of 500 PMHs, 3 cavitation generators made of P265GH steel, wet sanded with Struers sandpapers with the grain size of, respectively 200, 1000 and 2500, and one cavitation generator made of X2CrNi18-9 (304L) steel, sanded with sandpaper with the grain size of 2500, were selected for the investigations. The samples, prior to installation in a stream and flow device generating a cavitation environment (Figure 2), as well as after 500 PMHs of continuous work in such a device, were cleaned in an ultrasound scrubber,

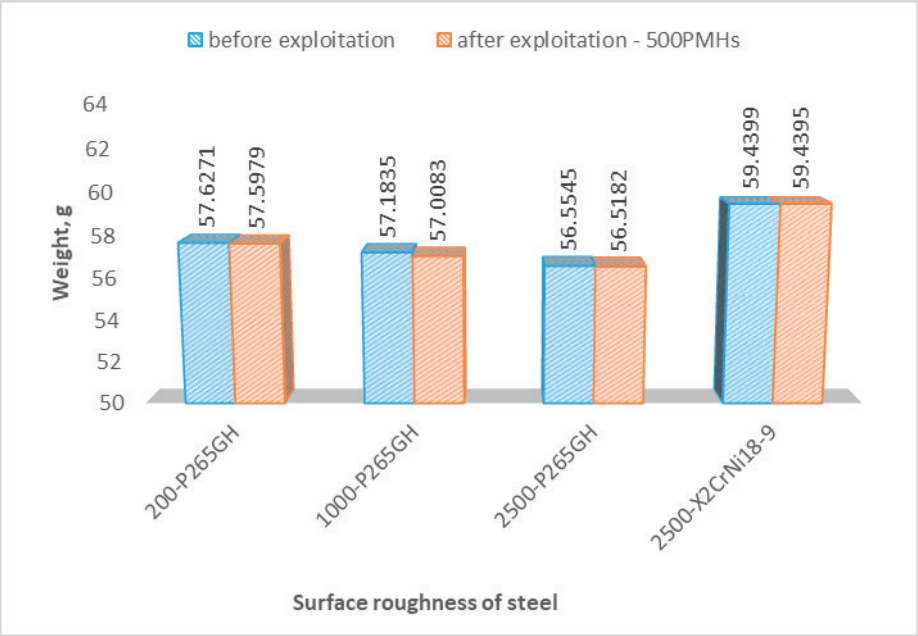


Figure 3 *Loss of mass, i.e. the mass of cavitation generator made of P265GH and X2CrNi18-9 (304L) steels after use in cavitation wear conditions in a stream and flow device*

and then weighed on an analytical scale, AS/X, by RADWAG. Surface roughness was determined with a Surtronic 25 contact profilometer by Taylor Hobson. At least four measurements, along the length of 16 mm in different areas of the cavitation generator, were made to determine the surface roughness of each sample. A motorised stereoscope microscope, Stereo Discovery V12, by Zeiss, with a magnification of 8 to 100 x, featuring an AxioVison image analysis system, was used for preliminary elimination of cavitation effect results. Detailed macroscopic examinations of the samples surface after cavitation wear were carried out with a scanning electron microscope, SUPRA 35, at the accelerating voltage of 5 to 20 kV using secondary electrons (SE) detection, with the magnification of 150-750 x.

3. Results

The cavitation generators, both before and after use in the conditions of cavitation wear in a stream and flow device continuously for 500 PMHs, were weighed and subjected to surface roughness measurements. The results of weight and surface roughness measurement examinations are shown, respectively, in Figure 3 and Figure 4.

The results of macroscopic examinations of the applied P265GH and X2CrNi18-9 (304L) steels in cavitation wear conditions made with a stereoscopic microscope, Stereo Discovery V12, by Zeiss, with the magnification of 8 to 100 x, are shown in Figure 5, Figure 6, Figure 7 and Figure 8. Meanwhile, the detailed results of examinations of the samples surface after the cavitation wear, carried out with a scanning electron microscope, SUPRA 35, using the secondary electrons (SE) detection, are shown in Figure 9. A cavitation generator made of ferritic- pearlitic steel designated as 200-P265GH, wet sanded with sandpaper with the grain size of 200, weighed 57.6271g before use and featured a surface roughness coefficient R_a of 0.627, thus falling to the 8th surface roughness class according to PN-EN ISO 1302:2004. As a result of operating the cavitation generator in a stream and flow device for 500 PMHs, numerous corrosion centres have appeared on the generator front surface according to the flow direction of the medium, i.e. water with an inlet pressure of 244 kPa, especially along the cavitation generator sanding direction. Cavitation wear was found near the straight-through openings (Figure 5). A negligible loss of mass of approx. 0.03 g was also found as a result of operation and a roughness coefficient R_a fell from 0.627 to 0.41, which is further classified as the 8th surface roughness class. For a cavitation generator made of ferritic-pearlitic steel, designated as 1000-P265GH, i.e. wet sanded with sandpaper with the grain size of 1000, with the weight of 57.1835 g and the surface roughness class R_a of 0.15, which classifies it into the 10th roughness class, due to work in a cavitation environment for 500 PMHs, numerous corrosion centres have developed, especially in the area of the generator's straight-through openings (Figure 6). The highest mass loss, by as much as 0.1752g in relation to all the operated generators, and the growth of the roughness coefficient R_a from 0.15 to about 0.5, was found after operating a generator marked as 1000-P265GH, which decreases the roughness class from 10 to 8. The last sample made of ferritic-pearlitic steel, designated as 2500-P265GH, operated for 500 PMHs, features a surface with numerous cavitation corrosion centres on the surface (Figure 7a) and damages, breaks and flaking of the cavitation generator openings' edges were found (Figure 7b and Figure 9). A small mass loss of the sample was seen for this sample from 56.5545g to 56.5182g and the roughness coefficient increased from 0.09 to 0.285, i.e. declined from the 10th to 9th roughness class.

Much better results were achieved by a cavitation generator made of austenitic steel X2CrNi18-9 (304L) wet sanded with sandpaper with the grain size of 2500, with the weight of 59.4399 g and with the surface roughness coefficient R_a of 0.1, which is ascribed to the 10th roughness class. The weight dropped

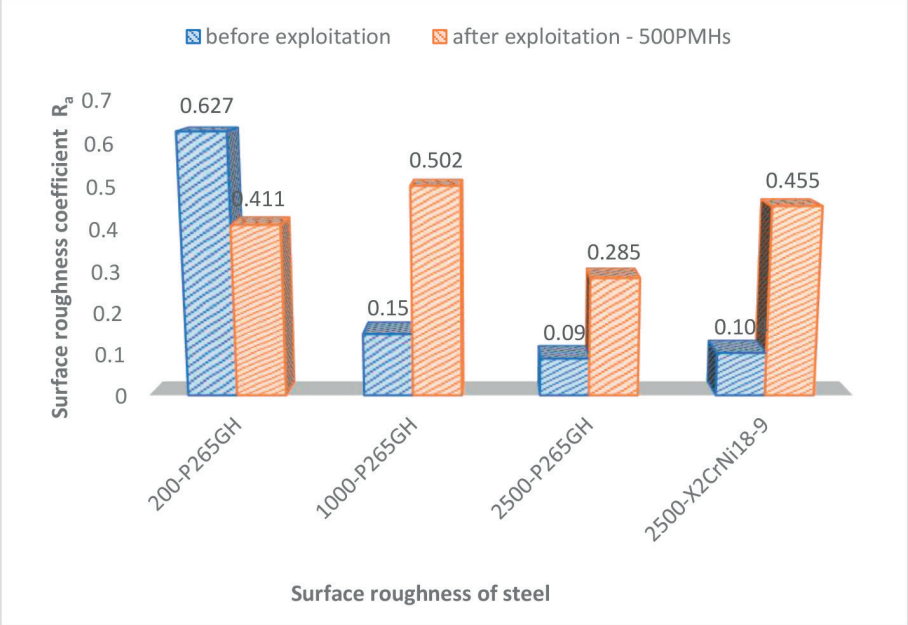
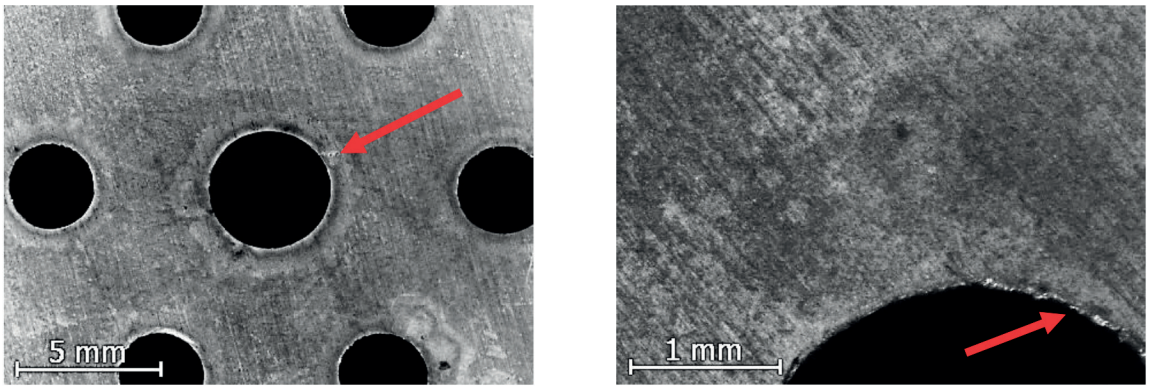
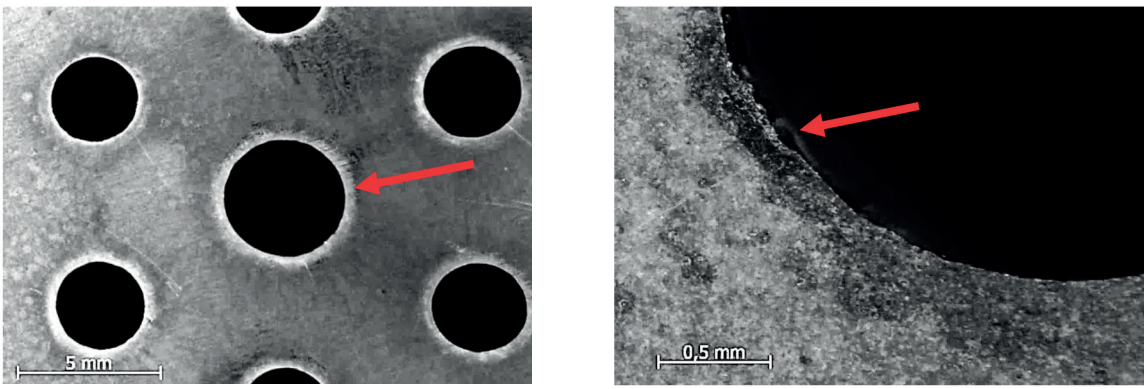


Figure 4 *Variation of the surface roughness coefficient R_a of cavitation generators made of P265GH and X2CrNi18-9 (304L) steels after use in cavitation wear conditions in a stream and flow device*



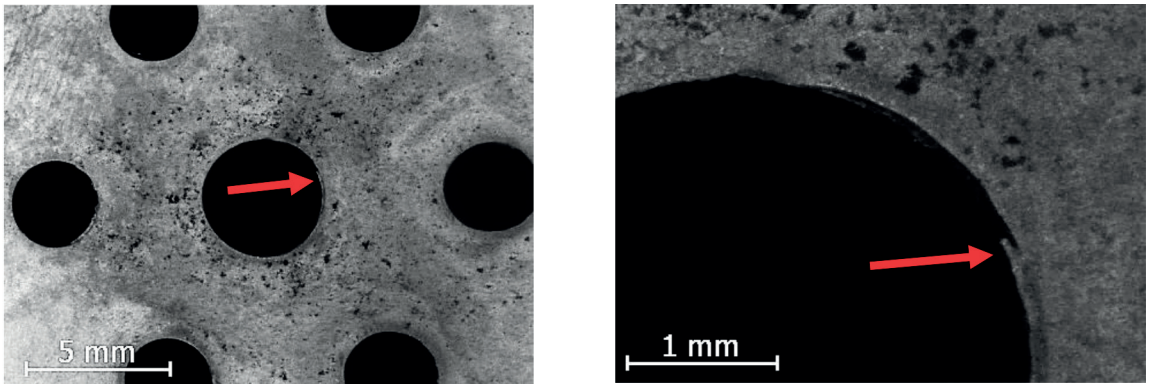
a) *b)*

Figure 5 *Result of cavitation wear of the surface of a constructional element made of P265GH steel wet sanded with sandpaper with the grain size of 200 and after operation in a stream and flow device for 500 PMHs a) magnification of 8 x; b) magnification of 40 x*



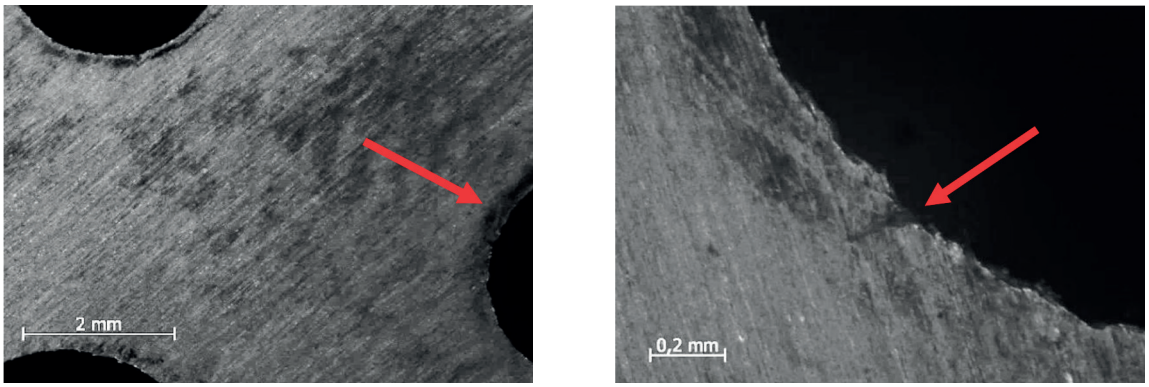
a) *b)*

Figure 6 *Result of cavitation wear of the surface of a constructional element made of P265GH steel wet sanded with sandpaper with the grain size of 1000 and after operation in a stream and flow device for 500 PMHs a) magnification of 8 x; b) magnification of 60 x*



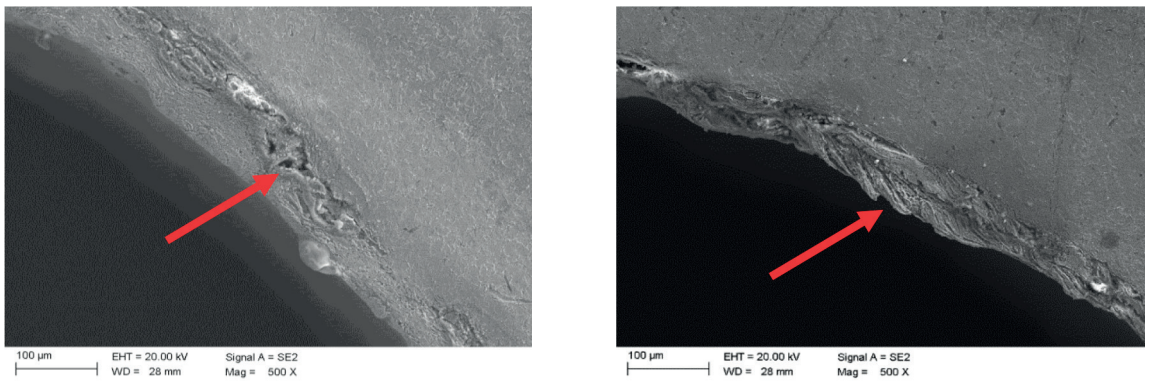
a) b)

Figure 7 *Result of cavitation wear of the surface of a constructional element made of P265GH steel wet sanded with sandpaper with the grain size of 2500 and after operation in a stream and flow device for 500 PMHs a) magnification of 8 x; b) magnification of 40 x*



a) b)

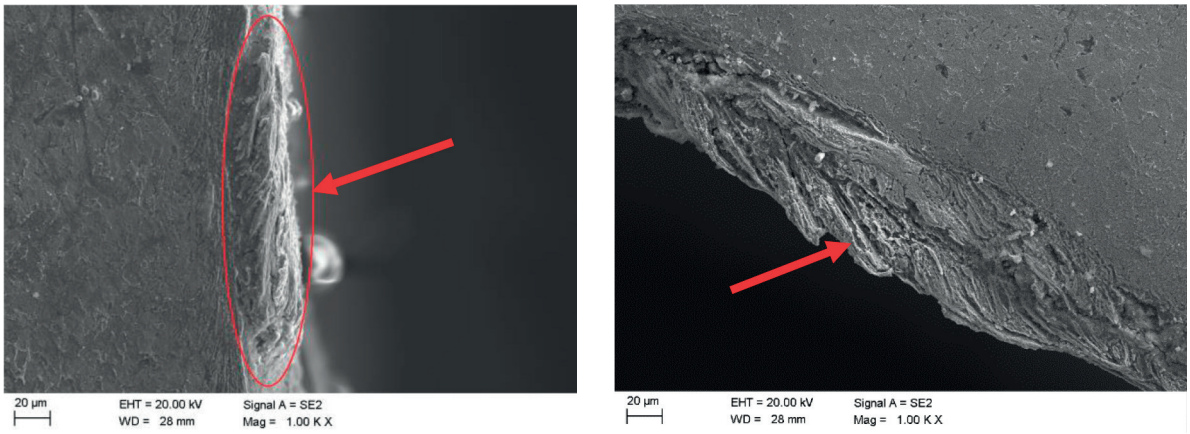
Figure 8 *Result of cavitation wear of the surface of a constructional element made of X2CrNi18-9 (304L) steel wet sanded with sandpaper with the grain size of 2500 and after operation in a stream and flow device for 500 PMHs a) magnification of 20 x; b) magnification of 100 x*



a) b)

Figure 9 *Result of cavitation wear of the surface of a constructional element made of P265GH steel wet sanded with sandpaper with the grain size of 2500 and after operation in a stream and flow device for 500 PMHs*

to just 0.0002 g after operation, i.e. was within the measurement error range, and the roughness coefficient rose to 0.455, which ranks it in the 8th roughness class. Deterioration of the surface roughness with practically negligible mass loss indicates that the tested cavitation generator reached surface distortion, but no significant material losses were found, especially near the edges of the straight-through openings, but only the micro cracks. It means that the material has deformed but without losing weight, a chrome-nickel steel generator that is resistant to corrosion is characterized by better stability and resistance in conditions of the flowing medium. A few places in X2CrNi18-9 (304L) cavitation generator were identified based on macroscopic photographs, especially near the edges of the straight-through openings, where cavitation corrosion was initiated. Axial brittle cracks going deep inside the material, were also identified (pointed by an arrow,



a) *b)*

Figure 10 *Result of cavitation wear of the surface of a constructional element made of P265GH steel wet sanded with sandpaper with the grain size of 2500 and after operation in a stream and flow device for 500 PMHs*

Figure 8b), initiated with an eroded edge of the straight-through opening.

In addition, based on the macroscopic examinations undertaken using a scanning electron microscope (SEM) for the cavitation generator made of P265GH steel, it was confirmed that on the edges of the straight-through openings, especially on the edge with the biggest opening area, exist numerous places, which were being spun by the flowing water. Cavitation craters and pits were formed in the first stage, then such craters were piling up and more corrosion centres were being formed, leading to either complete damage and breaking of the material part from the edge or to material flow inside the opening according to the medium flow direction (Figure 9 and Figure 10).

4. Summary

Based on the experiments carried out, consisting of the operation of specially designed and produced cavitation generators in the author's stream and flow device, generating a cavitation environment continuously for 500 PMHs, made of P265GH and X2CrNi18-9 (304L) steel, it was confirmed that the highest mass loss - 0.1752 g is seen for a sample of P265GH steel wet sanded with paper with the grain size of 1000. The smallest mass loss, with its value at the level of the measurement error, was recorded for the cavitation generator made of X2CrNi18-9 (304L) steel, sanded with sandpaper with the grain size of 2500. For the cavitation generators, with surface roughness in the input state of the 8th roughness class (R_a = 0.627 for 200-P265GH generator), it was found that operation in a stream and flow device for 500 PMHs causes small improvements in surface quality - R_a = 0.41. For the generators, sanded with sandpapers with the grain size of 1000 to 2500 for the both tested steels, it was found that the surface roughness has deteriorated from the coefficient value of about R_a = 0.1 to R_a = 0.3-0.5, depending on the material and sanding method of generators; hence the roughness class has decreased from 10 to 9 or even 8. Numerous corrosion centres, existing on the surface of cavitation generators, especially in the P265GH steel with a ferritic-pearlitic structure, were found based on the macroscopic examinations with a stereoscopic microscope and with a scanning electron microscope. Damages, breaks and flaking of the cavitation generator straight-through opening were found in the steel.

Certainly, the smallest number of cavitation wear effects was found for a cavitation generator made of austenitic steel X2CrNi18-9 (304L). Few places were identified based on macroscopic photographs, especially near the edges of the straight-through openings, where the cavitation corrosion was initiated and also axial brittle cracks were found going deep inside the material, initiated with the eroded edge of the straight-through opening, whereas their number was much smaller than for the ferritic-pearlitic steel, which is associated most of all with the properties of austenitic chromium - nickel steel, the basic characteristic of which is corrosion resistance. The medium pressure (244 kPa) in a stream and flow device and the operating time have initiated the process of cavitation wear, whereas the results are not as significant as for the ferritic-pearlitic steel. The X2CrNi18-9 (304L) steel has definitely the best service life in such operating conditions, unless both engineering materials are coated with special coatings, e.g. PVD, which is planned by the authors of this publication in further research experiments, and this may largely extend the service life of cavitation generators, thus reducing the operating costs.

References

[1] LINEK, T., TANSKI, T., BOREK, W.: Numerical Analysis of the Cavitation Effect Occurring on the Surface of Steel Constructional Elements. Archives of Materials Science and Engineering, 85(1), 24-34, 2017. https://doi.org/10.5604/01.3001.0010.1555

[2] KRELLA, A., STELLER, J.: Experimental Research on the Erosion of Selected Materials Depending on the Cavitation Load (in Polish). Instytut Maszyn Przeplywowych PAN, Gdansk, 2006.

[3] AZAR, G. T. P., YELKARASI, C., URGEN, M.: The Role of Droplets on the Cavitation Erosion Damage of TiN Coatings Produced with Cathodic Arc Physical Vapor Deposition. Surface and Coatings Technology, 322, 211-217, 2017. https://doi.org/10.1016/j.surfcoat.2017.05.050
[4] PN-EN 10028:2010 Flat Products Made of Steels for Pressure Purposes, Part 2: Non-Alloy and Alloy Steels with Specified Elevated Temperature Properties (in Polish).
[5] SZKODO, M.: Cavitation Erosion of Metal Construction Materials (in Polish). Wydawnictwo PG, Politechnika Gdanska 2008.
[6] BRENNEN, C. E.: Cavitation and Buble Dynamics. Cambridge University Press, New York, 2014.
[7] GLOWACKA, M., HUCINSKA, J.: The State of Research on Cavitation Destruction of Metal Alloys and their Protection against this Process (in Polish). Inzynieria Materialowa, 2, 79-84, 2001.
[8] SANTOS, J. F, GARZON, C. M., TSCHIPTSCHIN, A. P.: Improvement of the Cavitation Erosion Resistance of an AISI 304L Austenitic Stainless Steel by High Temperature Gas Nitriding, Materials Science and Engineering A, 382, 378-386, 2004. https://doi.org/10.1016/j.msea.2004.05.00310.1016/j.msea.2004.05.003
[9] LI, Z., HAN, J. S., LU, J. J., ZHOU, J. S., CHEN, J. M.: Vibratory Cavitation Erosion Behavior of AISI 304 Stainless Steel in Water at Elevated Temperatures. Wear, 321, 33-37, 2014. https://doi.org/10.1016/j.wear.2014.09.012
[10] LAZAROVA, R., PETROV, R. H., GAYDAROVA, V., DAVIDKOV, A., ALEXEEV, A., MANCHEV, M., MANOLOV, V.: Microstructure and Mechanical Properties of P265GH Cast Steel after Modification with TiCN Particles. Materials and Design, 32(5), 2734-2741, 2011. https://doi.org/10.1016/j.matdes.2011.01.024
[11] PN-EN ISO 1302:2004 Geometrical Product Specifications (GPS) – Indication of Surface Texture in Technical Product Documentation (in Polish).

Peter Drgona - Rastislav Stefun - Lubomir Stefke*

DESIGN OF CNC DEVICE DESIGNED FOR AUTOMATION OF THE APPLICATION OF VISCOUS MATERIALS

This article deals with the design and practical realization of a CNC device designed for the application of viscous materials, such as a thermal conducting paste. Nowadays, growing trend of usage of electronic circuits in transportation systems is observed. Switched mode power supplies are widely used in all types of means of transport and their infrastructure. For correct operation of power converters, produced heat must be dissipated. Thermal conducting paste is used for better heat dissipating but its application on surface can be an issue. Designed CNC device serves for easier and more accurate application in production process in transport industry.

Keywords: *automation in transportation systems, CNC design, CNC control, ball screw, stepper motor*

1. Introduction

At present, we are monitoring the growing trend of CNC devices deployment in the production process. Transport industry is not an exception. Automation generally increases productivity. Human factor is limited, which means that products are better and more accurate, and the number of manufacturing errors is rapidly declining. Another very important advantage of automation machines is to relieve people from physically strenuous, monotonous and psychologically dysfunctional work, what is a genuine issue in automotive industry. For this reason, requirement for the CNC dispending machine, which is the subject of this article, arises. From semiconductor converters, the heat generated by the dissipation power of semiconductor components must be drained, in most cases by passive heatsinks. Very often during the production a thermal conducting paste, which is often stored in tubes, is applied to the radiator by a piston pistol. The problem arises if the operator is not physically enough disposed to be able to apply the paste to the surface during the entire working shift. The role of the designed CNC device is therefore to automate the application process using standard tubes.

The basic requirements include:

- *X*-axis positioning: min. 300 mm
- *Y*-axis positioning: min. 400 mm
- *Z*-axis positioning: min. 80 mm with manual offset adjustment
- Positioning accuracy ± 1 mm in all axes
- Self-calibration at zero point
- Ability to position 5 kg tube
- Extraction from standardised 1x400 ml and from 2x400 ml tube, too
- Tolerance in dispensing two - part paste <0.1%
- Usage worm gear for extrusion
- Ability to perform any trajectory
- Configuration via GUI for standard PCs running MS Windows 7 or higher
- Configurable parameters: positions in *X*, *Y*, *Z*, positioning speed, adjustable amount of paste applied

2. State of the art

Nowadays, there are currently many CNC dispensing solutions available on the market. For example, we can present professional solutions of dispense devices from Bdtronic. Beckhoff offers excellent software solutions for them. Fisnar also offers dispensing machines using worm gear instead of air pressure but only for small cartridges. Another option is modular solution. For example, Mitsubishi offers interesting PLC-based CNC solutions M70 or C70 series [1]. However, no company on the market offered a solution that would fully satisfy our requirements. The main problem was that the devices were not able to work directly with standard tubes, and the offered solutions required additional development, which increased the price dramatically. In addition, the SW configuration of the device and the programming of the correct behavior of the device in various operating situations was also required, e.g. during the tube replacement. Very popular among hobby enthusiasts is Mach3. However, SW is not compatible with Windows 7 and it is not an open source, too. Frequently the drivers used for stepper motors include breakout boards for the MPS MP6500, Allegro's A4988 or TI's DRV8825 [2]. However, their power is poor for our application. That's why we decided to try to develop our own device, as cheap as possible. Block scheme of our proposed solution is shown in Figure 1. In next chapters are described additional information of separate blocks.

3. Mechanical construction

We have decided for the Compass H-1000 GS KIT from CauCau company. It is shown in Figure 2. This is a construction

* [1]Peter Drgona, [1]Rastislav Stefun, [2]Lubomir Stefke
[1]Department of mechatronics and electronics, Faculty of electrical engineering, University of Zilina, Slovakia
[2]Bel Power Solutions, Dubnica nad Vahom, Slovakia
E-mail: rastislav.stefun@fel.uniza.sk

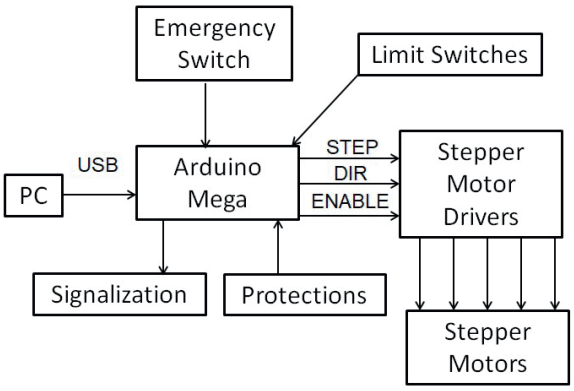


Figure 1 *Block scheme*

Figure 2 *Selected mechanical construction [3]*

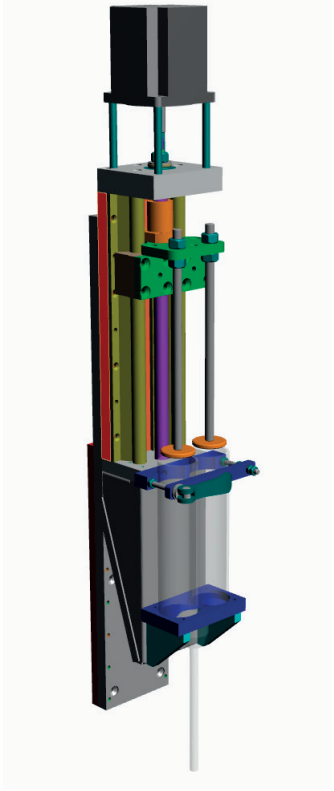


Figure 3 *3D model of paste extrusion module*

with a movable portal frame that performs a working movement. On the portal, there should be a spindle, in our case a paste extrusion module (Figure 3).

When designing the paste extrusion module, we decided to use the SBR16S. It is a system of a support rod on which there are carts with linear ball bearings with open cases. The support rod is turbid and its support is made of aluminium. The lead pair is mounted on a supporting aluminium plate. We have decided to use the ball screws to transform the rotational motion of the motor on the linear motion, because it has a higher load ability, durability, efficiency and accuracy [4], which is essential for our application.

4. Actuator selection and power electronics design

In our application, we decided to use two-phase hybrid stepper motors with bipolar winding. All motors work in open position loop [5]. For positioning needs MotionKing motor 23HS2430 was selected, with step angle 1.8°. 4 motors were needed, namely 2 on the X axis and one on the Y and Z axes.

When choosing the extruding motor, we firstly needed to know the amount of force needed to reliably push the paste out of the tube. We calculated the parameters from a pneumatic pistol used for this purpose. Working pressure is 600 kPa at the actuator area of $5.67 \cdot 10^{-3}$ m^2. It follows that in case of omission of friction, pistol is able to develop the force according to the equation:

$$F_a = p \cdot S = 6 \cdot 10^5 \cdot 5.67 \cdot 10^{-3} = 3402\ N \tag{1}$$

The total required torque M_c, which the motor must overcome, consists of several components [6].

$$M_c = M_{acc} + M_d + M_p + M_f \tag{2}$$

where M_{acc} – acceleration torque [Nm],
M_d – the torque necessary for load movement [Nm],
M_p – torque caused by pretension of the nut [Nm],
M_f – torque caused by friction in bearings and carts [Nm].

Torque M_{acc} directly depends on the momentum of the inertia of the system (extrusion pistons, flange, ball screw, rotor) and acceleration [4]. Since both components are low in our application, we can neglect the acceleration torque to simplify the calculation. The nut pretension is zero. We consider friction torque and it is reflected in the resulting formula as efficiency. For the final calculation of the total torque, M_c is therefore valid [7]:

$$M_c = \frac{F_a \cdot p \cdot o}{2000\pi \cdot \eta} = \frac{3.4 \cdot 10^3 \cdot 5 \cdot 0.5}{2000\pi \cdot 0.9} = 4.5\ Nm \tag{3}$$

where F_a – axial force [N], we calculated it in Equation (1),
p – screw thread pitch [mm],
η – efficiency of screw drive [-], although [4] states 95%, we chose 90%, because we consider not always ideal conditions of power transfer. Practically maintenance is often neglected which means that the lubrication of the bolt and nut contact will not always be ideal,
o – oversizing of the entire drive by 50% recommended by [8].

Motor SX4570 from Microcon company fulfills torque requirements. Like 23HS2430 its step angle is 1.8°. As we can see in the Figure 4 [8], it is possible to achieve 2.5 rot/s with enough torque at divided step.

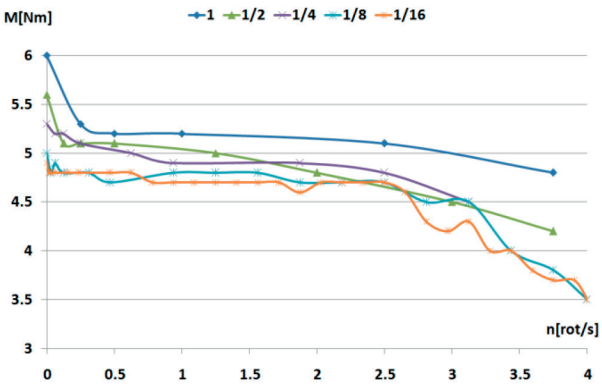


Figure 4 *Torque characteristics of motor SX4570 for current amplitude 3.8 A*

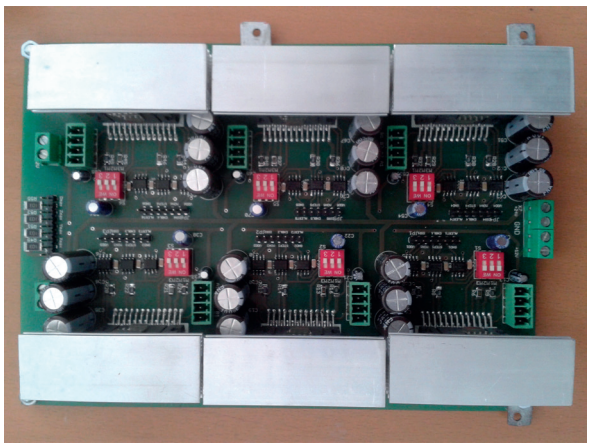


Figure 5 *Mounted PCB, top view*

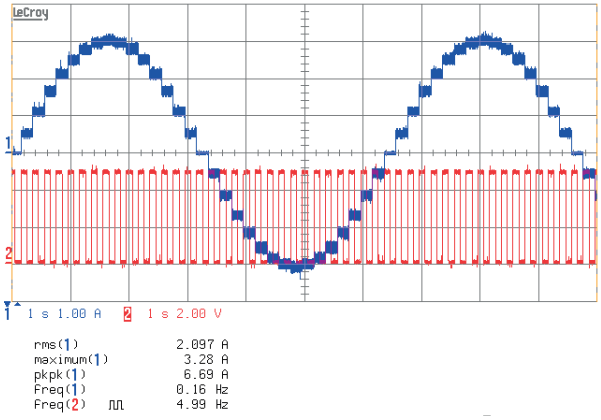


Figure 6 *Phase current of motor 23HS2430 for 1/8 step (blue) and signal STEP (red)*

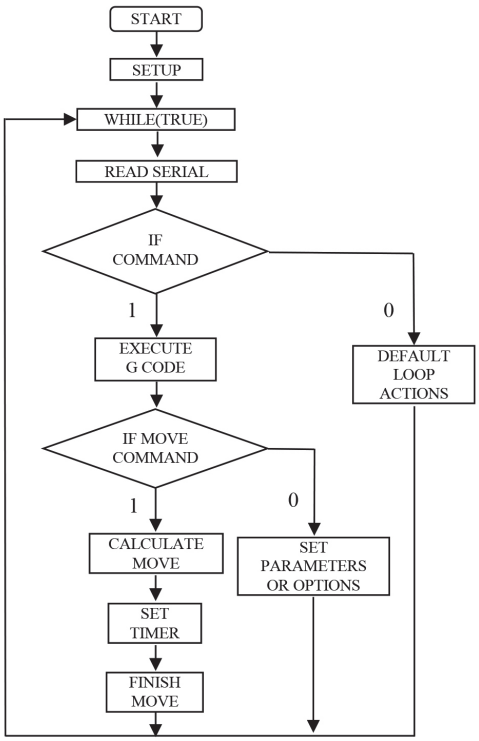


Figure 7 *Simplified Repetier flowchart*

There are many topologies of stepper motor driving circuits. According to [9], for maximum efficiency, we have chosen circuits with PWM current control. We decided to use integrated circuits and therefore IC TOSHIBA TB6600 was chosen. It provides integration of logical circuits and two power H-bridges capable of working with a peak current up to 4.5A. The logical part provides sequential changes in the current reference, depending on the chosen step size division. Around this reference, the output current is switched by a hysteresis controller (Figure 6) [10]. IC allows the stepper motor to be controlled by three signals only, namely STEP, DIR and ENABLE [11]. Rising edge of STEP signal causes motor to move one step. Logic level of signal DIR determines direction of motor rotation. Low level is for CW direction and high level is for CCW direction. The designed PCB allows to connect 6 stepper motors (Figure 5). The size of the step angle is possible to divide up to 1/16 using DIP switch. Digital isolators ISO722x isolate apart power and logic parts of circuit. As a power supply was used FNP300-1024, which provides 24V for power part and 12V for logic part. Using voltage regulator LP2950CDT−5.0 we lowered 12V to TTL level, which is suitable for TB6600 [11].

5. Selection of control electronics and firmware

As can be seen in Figure 1 choice of control electronics fell on Arduino Mega based on ATMega 2560, which is 8-bit microprocessor from Atmel. Because we are unable to debug program using boot loader, we used AVR Dragon board. This requires the OCD feature that the ATMega2560 meets. As a base for the control program, we chose open source firmware Repetier, made for 3D printers, as we needed to configure the device behavior [12].

In Figure 7, there is shown a simplified Repetier flow chart. It operates in an infinite loop. After start of programm all registers and GPIOs are set. Then communication with PC can start. Commands are in ISO code language. It is a standard for CNC machines. If no command is sent to the processor, then the flag interrupts, temperature of hot end or the state of communication is controlled. But if the command is present in the communication buffer, processor decides whether it is moving instruction or instruction for setting some parameters or options. At moving instruction processor calculates amount of steps to reach the desired point in space according to the desired velocity

and acceleration. The logical level of DIR signal and other attributes are also set. The timer makes appropriate PWM signal for the motor control. After the finish of the move the processor reads the serial communication link again.

Because of industrial application we had to implement security features. We connected emergency switch to the Reset pin. Pressing it will bring logical 0, which will cause processor to reset and to shutdown the device immediately.

Overheating of the circuit is solved by the interrupt. When the chip temperature reaches 160°C, IC switches the ALERT signal from log.1 to log.0. The falling edge causes interrupt. In service routine all ENABLE signals first switch to log.0. Subsequently, the green LED on the signaling device goes off and the red signal starts lighting to indicate the fault condition. Before the program remains stuck in an infinite loop, the GUI application will show message that overheating has occurred. Since interrupts are used to communicate with PC, they firstly must be enabled. ATmel processors automatically disable all interrupts when the program is in the service routine of one of them. After GUI message all interrupts are again disallowed with the cli() command. The only way to get a device from an endless loop is to reset the processor after cooling of the drivers.

The role of the infrared gate is to tell the processor that a foreign body, most often a part of the human body is in the working area of the machine. The processor must respond with an immediate break of the motion activity. Engines must be stopped throughout the presence of the body in the detected area. We tried to use an interrupt but then there was a loss of steps, so we had to change the method where Bresenham interpolation algorithm is calculated.Return value of this method is the value of the comparator register for the timer. The output compares register changes for each step, depending on which axis the movement to reach the desired point in space is performed. It also depends on the required speed and acceleration. The timer then generates the appropriate STEP that signals for the individual motors to be rotated with respect to the required acceleration ramps by changing the pulse width. The current calculated ISO code line is pointed by pointer named as cur. Its private attributes include initial speed, desired speed, acceleration, accelerator steps, and so on. The Bresenham algorithm calculates some of its attributes, mainly the difference and deviation, but also the direction. Some of these attributes changes with each step, so this method is called by the processor when one of the motors moves and recalculates the attributes again and thereforwe performed the implementation here. So before processing the next step, the processor checks the logical level on the input from the gate. If the log. 1 level is there, the red LED goes on, the green goes off, and the processor executes blank "nop" instructions until the logic level changes again to log.0. All motors are stopped at that moment. After the change of logical level, the green LED lights up again and the red goes off. Then the attributes to which the pointer points are reset by the command cur = NULL and then recalculated, when the program is resumed. This has resulted in a ramp increase of the rotation speed, which prevents the steps being lost.

When the device operates, it may happen that the tube is emptied. Therefore, we also implemented algorithm for this case. In the method with Bresenham algorithm, is monitored whether or not a switch is activated. The switch is activated when the push pistons reach the lower position. If so, the attributes pointed by the 'cur' pointer are reseted and the red LED lights up, and the green goes off. In the GUI is then printed "Change_Tub" message. Current values of the STEP, DIR and ENABLE signals of the extruder are stored in auxiliary variables. Subsequently, the ENABLE signal is set to log.1 as well as the DIR signal is. The rising edge of the STEP signal in the cycle is generated until the pistons reach the top position. Next switch gives the processor information that the pistons are in the upper position. Consequently, it is necessary to change the level of the logical signal DIR back to log.0 to be able to move the pistons downwards. Then the device waits for the human operator to replace the tube. The operator sets the pistons to position where the paste starts to flow from the mixer by pushing the button. If everything is ready to continue the application process, operator presses the OK button. Next, for cycle, which generates rising edge of STEP signal, stops and the LED lights switch to the original combination. STEP, DIR, and ENABLE signals are rewritten to the original logical values. By doing so, the tube replacement algorithm is terminated and the device can continue its work.

6. Practical realization and evaluation of achieved results

We attached the paste extrusion module onto the *Z*-axis aluminium plate of positioning mechanism. For the connection of the selected motors to the structure, we used spacers in combination with flexible couplings. To control the actuators, we used the proposed PCB. It works reliably. The positioning motors are operated on a 1/8 step at a current amplitude 3.3 A. For a more fluent tube extrusion, we chose for the SX4570 1/16 step at a current amplitude 3.8 A. We placed the wires in the energy chains to prevent them from being twisted (Figure 8). The demonstration of the application over the circular trajectory is shown in Figure 9. As far as the specification in Part I is concerned, we have successfully fulfilled all the requirements. Reached dimensions of the workspace and maximal values of axes speeds are stated in Table 1.

The ability of manual *Z* axis offset adjustment is possible thanks to the appropriate design. The change range is 364.5 mm. The overall positioning accuracy is given by the precision of the stepper motor, the ball screw and the nut. After counting all deviations, we obtain maximum deviation that approximately is ± 0.0014 mm, which does not exceed the required accuracy of ± 0.1 mm.

Calibration at zero point was achieved by using limit switches on each axis. During the tests, we were able to reliably position 5 kg of weight. When designing a paste extrusion module, we used a worm gear with accentuate on making possible to rework the module as easily as possible for the use of both one and two-part tubes. During the application of two-part paste, both pistons are set in one plane and driven by the same motor, so there is no deviation in the application. By selecting the appropriate firmware, we have determined that the device is able to perform any trajectory whether it is a straight line or a circle. Thanks to the Repetier Host GUI application, compatible with Windows 7 and

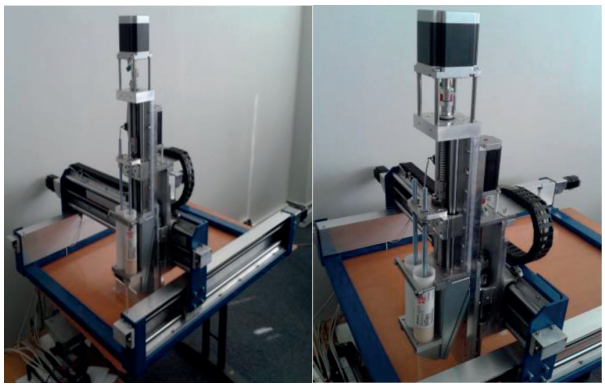


Figure 8 *Practical realization of the CNC device*

Figure 9 *Result of paste application along the circular path.*

Table 1 *Reached dimensions of the workspace and maximal values of axes speeds at the selected acceleration 500mm/s²*

Axis	Required dimension [mm]	Reached dimension [mm]	Maximal reached speed [mm/s]
X	300	665	185
Y	400	570	242.5
Z	80	80	51

higher, we are able to easily configure all the required parameters. These are the positions in the *X*, *Y*, *Z* axes as well as the position of the motor for extrusion purposes, the speed and acceleration of the drives. It is also possible to show the path in the application. The amount of applied paste can be adjusted to a certain level by changing the speed of the extrusion motor.

7. Conclusion

We have built a functional device capable of applying viscous materials according to the specified specification. We have selected mechanical construction on which we have attached the designed module for the paste extrusion. The selected motors are controlled by an open source firmware which was improved according to our needs. The control algorithm for the tube exchange and some safety features were added. The device will serve to automate the application process in the production process. In the future, the device will be placed in the cage and the only accessible area will be covered with an infrared gate. It is also necessary to program the device so that there is not too much paste dispensed during the application of the material.

Acknowledgment

Authors would like to thanks to APVV-15-0571 Research of the optimum energy flow control in the electric vehicle system, KEGA 073ZU-4/2017 Implementation of modern tools for teaching automotive electronics and electromobility, Vega 0479/17 Research of optimal methods of energy transfer control in systems with accumulating elements and Bel Power Solutions, Dubnica nad Vahom.

References

[1] STEFANEC, P., PASKALA, M., DOBRUCKY, B., SUSLIK, B.: Control System for Preheating of Molds with using PLC. Proceedings of International Conference on Industrial Technology and Management Science (ITMS 2015), China, book series: ACSR-Advances in Comptuer Science Research, 34, 1647-1652, 2015.

[2] DUCHON, F., HUBINSKY, P., HANZEL, J., BABINEC, A., TOLGYESSY, M.: Intelligent Vehicles as the Robotic Applications. Proceedings of 5th International Conference on Modelling of Mechanical and Mechatronics Systems (MMaMS 2012), Slovakia, book series: Procedia Engineering, 48, 105-114, 2012.

[3] Online. Available: http://www.cnc1.eu/sk/KompasH1000GSKIT.htm [accessed 2017-03-11].

[4] Online. Avialable: http://www.thomsonlinear.com/downloads/screws/Leadscrews_Ballscrews_Splines_cten.pdf [accessed 2017-04-14].

[5] URICA, T., SIMONOVA, A., KASCAK, S.: Control Design and Simulation of Continuous Systems. Proceedings of IEEE 18th International Scientific Conference Electric Power Engineering (EPE 2017), Czech Republic, USB key, 125-129, 2017.

[6] FEDAK, V., DUROVSKY, F., UVEGES, R., KYSLAN, K., LACKO, M.: HIL Simulator of Drives of an Industrial Robot with 6 DOF. Elektronika ir Elektrotechnika, 21(2), 14-19, 2015.

[7] Online. Available: http://www.linearmotiontips.com/calculate-motor-drive-torque-ball-screws/ [accessed 2017-05-10].

[8] Online. Available: http://microcon.cz/ [accessed 2017-05-10].

[9] TEREN, A., SPANIK, P., FENO, I.: DC/DC Converters with Soft (ZVS) Switching. Proceedings of International Conference ELEKTRO 2001, Slovakia, 82-87, 2001.

[10] SIMONOVA, A., HARGAS, L., KONIAR, D., LONCOVA, Z. HRIANKA, M.: Mathematical Analysis of Complete Operation Cycle of a System with Two-Position Controller. Proceedings of IEEE 11th International Conference ELEKTRO 2016, Slovakia, CD-ROM, 213-215, 2016.

[11] Toshiba, Datasheet TB6600 [online]. Available: www.toshiba.com.

[12] HOCK, O., DRGONA, P., PASKALA, M.: Simulation Model of Adjustable Arm Using Denavit-Hartenberg Parameters. Proceedings of 10th International Conference ELEKTRO 2014, Slovakia, 176-179, 2014.

Denisa Cihlarova - Petr Mondschein - Silvia Capayova*

COMPACTION DENSITY DETERMINATION OF THE ROAD ASPHALT LAYERS

This paper analyses the destructive and non-destructive methods for checking the degree of compaction of asphalt layers and their multi-criteria evaluation. The data obtained during laying of ACO 11+50 mm thickness was selected for the analysis. The bulk density was determined by using probes Troxler 3440, Troxler 2701 and from the drill cores in these given locations. According to the analysis accuracy, the probe Troxler 2701 used to determine the final degree of compaction of the construction asphalt layer and for the routine control of compaction is not suitable, even though when applying AHP (Analytic Hierarch Process) it was assessed as the second most suitable one.

Keywords: *asphalt mixture, compaction, quality control, destructive method, non-destructive method, nuclear density gauge, electrical density gauges, bulk density, drill core, multi-criteria evaluation*

1. Introduction

The asphalt is a significant building material used both in roads and for other building purposes. Up to 97 percent of the roads have a bituminous surfacing in the Czech Republic and 99 percent in Slovakia (situation to 1st July 2017). Road asphalt layers make the upper part of the flexible pavements, which is exposed to direct vertical and tangential wearing effects of the vehicles, which are then transmitted to further layers of the construction. The wearing course layer is immediately submitted to the effects of atmospheric and climatic influences.

Thus, the cover of the road should be water-resistant, flat, and should have proper anti-skid qualities, so that safe, fast and comfortable drive is ensured. To meet these requirements, it is necessary not only to ensure adequate building material [1], [2], [3] but also to keep the technology of the construction; otherwise there could be various defects such as permanent deformation [4] and adhesion failure between individual asphalt layers, [5], and when talking about wearing course layers one means inadequate surface properties. The extent of compaction influence all the above mentioned qualities and so its control plays a significant role in the process of building road communications.

2. Compaction quality control of asphalt mixtures

A proper compaction of the asphalt mixtures is crucial for reaching the required properties and ensuring the loading capacity, performance and length of life. By the compaction process, the permeability of the mixture decreases and by this the carrying capacity, weariness and rutting resistance increase. The extent of compaction depends to a certain level on the base, type of mixture, thickness of layers, used compaction technology and local conditions during laying. A well-designed and managed compaction process is crucial in order to reach good quality and long lifespan of the asphalt roads and vice versa; inadequate compaction leaves a high percentage of air pores in the road complex of layers, which becomes sensitive to moisture infiltration, oxidation and making of cracks [6]. On the other hand, over-compaction will result in very small amount of air pores in the construction layer, which could lead to asphalt bleeding on the surface of the road in the construction layer during the summer season, or it could result in a crushing of the aggregate, thus change of the mechanical qualities of the road, or rather the construction layer. It follows that when the road is inadequately compacted, one can see a non-standard degradation as a consequence of this and thus a decrease in its length of life.

The technology of compaction process is mainly determined empirically. However, there are studies that try to predict the performance of a mixture during a compaction process in the field conditions by various models, see [7]. Finding of the adequate model would enable designing the mixture so that after the compaction in the given local conditions, the mixture would have the required qualities.

The compaction control is carried out in two ways, the destructive method (drill core) and non-destructive methods NDG (Nuclear Density Gauge), or EDG method (Electrical Density Gauges). Globally, there is a rise in research, particularly in the area of non-destructive tests for determination of the compaction degree, e. g. by using sensors FBG (Fibre Bragg Grating) [8], or by Intelligent Asphalt Compaction Analyzer (IACA), which uses an artificial neuron net (ANN). By this ANN, the estimated value of the road bulk density under a roller during the compaction process is obtained [9]. Those two methods of finding the bulk density could be characterised as a continuous control of the compaction process; for now, they cannot be used as a tool for proving the concordance with the final layer, though, the advantages of the continuous method cannot be denied. The

* [1]Denisa Cihlarova, [2]Petr Mondschein, [3]Silvia Capayova
[1]Department of Transport Constructions, Faculty of Civil Engineering, VSB - Technical University of Ostrava, Czech Republic
[2]Department of Road Structures, Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic
[3]Department of Transportation Engineering, Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Slovakia
E-mail: silvia.capayova@stuba.sk

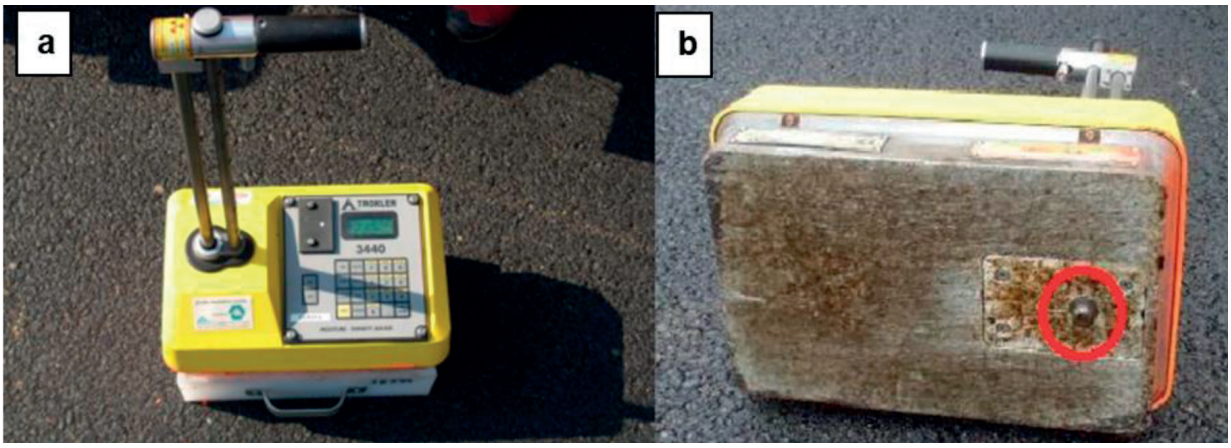


Figure 1 *Troxler 3440, a - a view from the top; b - a view from the bottom [11]*

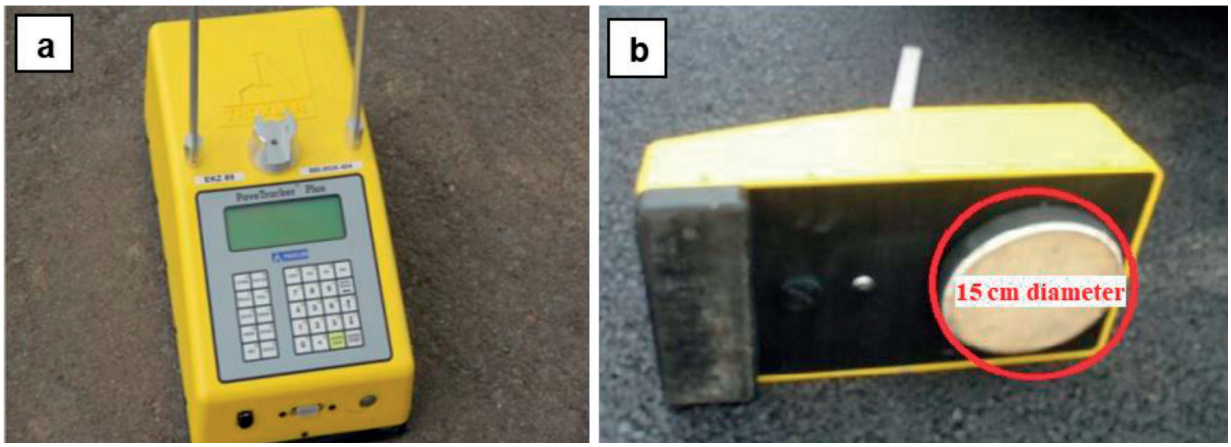


Figure 2 *Troxler 2701-B, PaveTrackerTM, a - a view from the top; b - a view from the bottom [11]*

study [10] focuses on examination of the achievable compaction quality in the whole length of the roadway. It was discovered that the bulk density of the layers of the asphalt road accidentally changes during the compaction process and that is caused by the re-orienting the aggregate to a random complex structures. Thus, there is 1.9 % possible difference in the density in the areas only 20 cm apart.

The control principle of the finished layers is described in the standard of Czech Republic *CSN 73 6121(Road building - Asphalt Pavement Courses - Construction and conformity assessment)*. The degree of the compaction (CD) is possible to control destructively on the drill cores, or after an agreement between the ordering party and the contractor also by the non-destructive method. Depending on the type of the asphalt mixture, the 96 or 98 % compaction degree is required. The asphalt mixtures marked S must reach 98 % of the compaction degree on the average and its compaction degree cannot decrease below 96 %. The compaction degree of the asphalt mixtures with the label + or "without any label" cannot decrease below 96 %.

3. Experiment

Individual methods for determination of the compacted bulk density of the asphalt differs from the point of view of time demands, technical equipment, calibration, principle of determination of the bulk density and its exactness. The aim of the experiment was to conduct a comparison of the methods: destructive method, NDG and EDG.

3.1 Destructive method

For obtaining core drill holes, the road core drilling equipment 60-0100 was used. Samples of the 100 mm diameter were acquired by using this equipment. After separating the layers, the bulk density was determined on the drill cores by using the procedure for a saturated dry surface.

3.2 Nuclear Density Gauge (NDG)

Measuring by this method was carried out by the radiometric set Troxler 3440, Figure 1. It is a transportable gauge for the fast determination of moisture, density and extent of compaction, especially soils, concrete and asphalt surfaces, without disturbing the construction of the measured material. The measuring probe contains shielded Cs 137 source of the 0.3 GBq activity and 241Am/Be of the 1.48 GBq activity. The weight of the NDG is 13 kilograms. Prior to each measurement, it is necessary to determine the daily calibration response, i.e. radiation intensity. Every 2 years the gauge needs to be inspected for the long-life stability in the certified laboratory; moreover, it is necessary to record once a month dose neutron equivalents of

Table 1 *Measured sections overview*

Section No.	Communication	Structural layer	Layer thickness	Asphalt mixture
1	III/4673 Stitina - Dehylov	wearing course	50 mm	ACO 11 +
2	II/467 Stitina - Kravare - Stepankovice	wearing course	50 mm	ACO 11 +
3	II/467 Stitina - Kravare I part	wearing course	50 mm	ACO 11 +
4	II/467 Kravare - Stepankovice II part	wearing course	50 mm	ACO 11 +
5	II/467 Stepankovice - Koberice III part	wearing course	50 mm	ACO 11 +

Table 2 *Statistical assessment of the bulk densities and compaction degree*

Section No.	No. of measurement		Bulk density in kg.m^{-3}			Compaction degree in %		
			Troxler 3440	Troxler 2701	Drill core	Troxler 3440	Troxler 2701	Drill core
1	8	min.	2 261	2 089	2 294	96.6	89.3	98.0
		mean	2 301	2 187	2 319	98.3	93.5	99.1
		max.	2 356	2 302	2 357	100.7	98.4	100.7
		min.-max.	95	213	63	4.1	9.1	2.7
2	7	min.	2 258	2 078	2 274	96.3	88.7	97.5
		mean	2 287	2 192	2 301	97.6	93.5	98.2
		max.	2 328	2 305	2 324	99.8	98.8	99.2
		min.-max.	70	227	50	3.5	10.1	1.8
3	10	min.	2 284	2 259	2 274	97.9	96.8	97.5
		mean	2 328	2 286	2 292	99.8	98.0	98.3
		max.	2 365	2 331	2 310	101.4	99.9	99.1
		min.-max.	81	72	36	3.5	3.1	1.6
4	14	min.	2 250	2 214	2 284	96.1	94.5	97.5
		mean	2 335	2 260	2 296	99.7	96.5	98.0
		max.	2 393	2 306	2 323	102.2	98.4	99.2
		min.-max.	143	92	39	6.1	3.9	1.7
5	16	min.	2 284	2 128	2 309	96.9	90.2	97.9
		mean	2 313	2 170	2 317	98.1	92.0	98.2
		max.	2 345	2 206	2 324	99.4	93.6	98.6
		min.-max.	61	78	15	2.5	3.4	1.7

the operators, which are continuously assessed in the national personal dosimetry service. The operators must be trained for work with a radioactive material.

3.3 Electrical Density Gauges (EDG)

The non-nuclear sensing device Troxler 2701-B, Pave Tracker TM, Figure 2, was used for the measurement. This device is using electromagnetic density indication. The PaveTracker Plus does not contain any radioactive material and so there is no need for any licence or any special training for operating this gauge. Its weight is approximately 6 kg. A regular calibration every 2 years must be executed in a certified laboratory, as well.

3.4 Course of experiment

The selection of the assessing parameters was chosen in order to make the mutual comparison of individual methods possible. Parameters (time for measurement preparation, time of measuring, bulk density assessment time and other costs) were found out when reconstructing roads in real traffic of the construction and from personal experience. For evaluation of the exactness, at one place the measuring NDG and EDG in both positions was carried out, and at the same place the drilling was executed. Data obtained by this measurement were subsequently divided into two evaluating files, which had the same input attributes of the constant thickness and composition of the laying mixture; the measured bulk density obtained from the drill was taken as a reference value.

3.5 Evaluation exactness

Measuring the level of compaction by all three methods was carried out on five individual sections on the wearing course layers, see Table 1. All measurements on five experimental sections were performed on one type of asphalt mix - ACO 11+ asphalt concrete with 50/70 paving grade. All the fractions, except aggregate filler, which were used for the production of bituminous mixture, were from Bucovice quarry. From a petrographic point of

Table 3 Compaction degree with the standard requirements and mutual comparison of them

Section No.	No. of measurement	No. of measurements with CD ≥ 96 %			CD Drill core > CD Troxler 3440	CD Drill core > CD Troxler 2701
		T 3440	T 2701	Drill core		
1	8	8	2	8	7	7
2	7	7	1	7	5	6
3	10	10	10	10	1	6
4	14	14	10	14	1	12
5	16	16	0	16	10	16

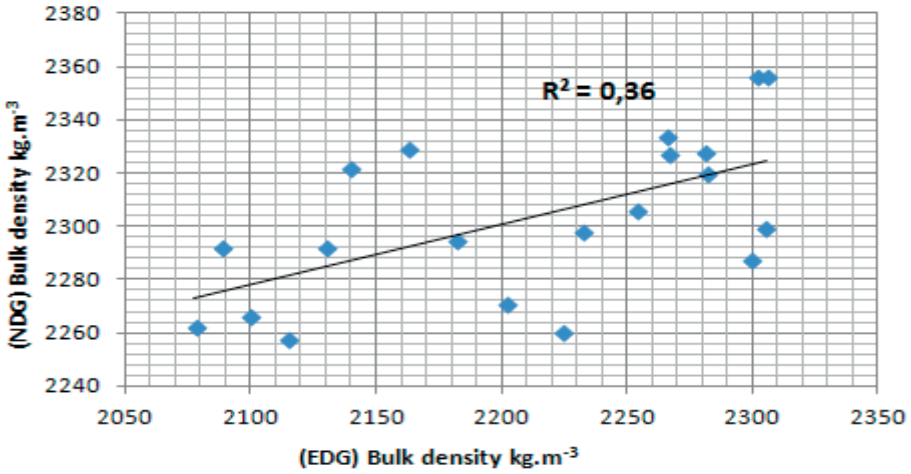


Figure 3 The relationship between the method NDG and EDG

view, it is a greywacke of gray to dark gray color with a brownish or greenish shade. The compaction degree is determined as a ratio of the compacted bulk density determined on the structural layer and reference bulk density determined on the Marshall Test specimen, i.e. cylinder test specimen. In the case of the destructive method, the bulk density of the drill core and Marshall specimen are determined by the same method, by using the hydrostatic scales, which means that this method can be considered to be a standard method that has the highest informative value. Both of the above described non-destructive methods were compared to the destructive method.

The basic statistical parameters of the compacted bulk density and the compaction degree of the compared sections are given in Table 2. One can conclude that the bulk density variance is the smallest in all the five assessed cases when determining the drill cores. The difference is by one third smaller when comparing to the non-destructive method by Troxler 3440. At the section number 5, the difference between maximal and minimal measured bulk density at the drill cores is four times smaller than by the Troxler 3440 measurement. The Troxler 2701 probe measurement displayed much worse results. The same results can be concluded when comparing compaction intensities determined by three methods. The difference between the compaction density determined on the drill cores and the Troxler 3440 probe is at individual sections equal to: No.1+0.8 %, No. 2 +0.6 %, No. 3 - 1.5%, No. 4 - 1.7%, No.5 - 0.8%. The difference between the compaction density determined on the drill cores and the Troxler 2701 probe is substantially higher, it reaches even cca. 6 %.

The bulk density assessment with the demand of the standard CSN 73 6121, which requires minimal compaction intensity of 96 % for the mixture ACO 11, is displayed in Table 3. During the checking tests carried out on the drill cores, it was possible to state that all five sections were compacted enough in their whole lengths. When evaluating the compaction density by the non-destructive probe Troxler 3440, exactly the same result can be concluded. When assessing the compaction density by the non-destructive probe T 2701, the number of satisfactory measurement was 23 out of 55.

Within finding out whether there is a certain dependency between the destructive method and both of the non-destructive methods, the conclusion is that none has been found. The coefficient of determination was near zero. When comparing both non-destructive methods, the coefficient of determination was between (0.3-0.4), Figure 3. However, even in this case it would be daring to claim that the dependency exists between these two methods.

3.6 Multi-criteria evaluation

In order to determine the relative weights, the quantitative method of pair comparison (the Saaty's method) was used. The essence is to construct a matrix expressing the relative values of a set of attributes. Five experts are asked to choose whether cost is very much more important, rather more important, as important, and so on down to very much less important, than operability. Each of these judgements is assigned a number on a scale. One common scale (adapted from Saaty), can be seen in Table 4.

A basic, but very reasonable, assumption is that if attribute A is absolutely more important than attribute B and is rated at 9, then B must be absolutely less important than A and is valued at 1/9. The degree of importance of individual criteria

Table 4 *The Saaty Rating Scale*

Intensity of importance	Definition	Explanation
1	Equal importance	Two factors contribute equally to the objective.
3	Somewhat more important	Experience and judgement slightly favour one over the other.
5	Much more important	Experience and judgement strongly favour one over the other.
7	Very much more important	Experience and judgement very strongly favour one over the other. Its importance is demonstrated in practice.
9	Absolutely more important	The evidence favouring one over the other is of the highest possible validity.
2,4,6,8	Intermediate values	When compromise is needed

Table 5 *Pair preference of the criteria*

	Measurement accuracy	Price 1 measurement	Equipment price	Price adjustment	Time measurement	Time calibration	Activation time	Time finding density	Maintenance	Energy	Additional equipment	Weight
Measurement accuracy	1	9	9	9	9	9	9	9	9	9	9	9
Price 1 measurement	1/9	**1**	5	7	5	9	9	7	9	9	9	9
Equipment price	1/9	1/5	**1**	7	5	9	5	5	5	7	5	9
Price adjustment	1/9	1/7	1/7	**1**	1/3	5	1/3	1/5	1	1	1/5	1/5
Time measurement	1/9	1/5	1/5	3	**1**	9	7	5	5	7	5	9
Time calibration	1/9	1/9	1/9	1/5	1/9	**1**	1/7	1/9	1/7	1/5	1/7	1/9
Activation time	1/9	1/9	1/5	3	1/7	7	**1**	1/5	1/7	1/5	3	3
Time finding density	1/9	1/7	1/5	5	1/5	9	5	**1**	5	7	5	9
Mainentance	1/9	1/9	1/5	1	1/5	7	7	1/5	**1**	1	1/5	3
Energy	1/9	1/9	1/7	1	1/7	5	5	1/7	1	**1**	1/5	1
Additional equipment	1/9	1/9	1/5	5	1/5	7	1/3	1/5	5	5	**1**	5
Weight	1/9	1/9	1/9	5	1/9	9	1/3	1/9	1/3	1	1/5	**1**

Table 6 *Multi-criteria method evaluation of the bulk density determination on construction*

	Measurement accuracy	Price 1 measurement	Equipment price	Price adjustment	Time measurement	Time calibration	Activation time	Time finding bulk density	Mainentance	Energy	Additional equipment	Weight	*u* (Ai)	Range number
Troxler 3440	0.8	0.9	0.6	0	0.9	0	0.8	0.8	0.6	1	1	0.8	0.80	1
Troxler 2701	0	1	1	1	1	1	1	0.9	1	1	1	0.9	0.71	2
Drill Core	1	0	0.3	1	0	1	0.1	0.1	0.8	0.8	0.8	0	0.44	3
Scales	0.28	0.18	0.12	0.02	0.09	0.01	0.02	0.18	0.03	0.02	0.04	0.02		

was determined based on experience of the author, see Table 5. To determine the order of advantages of individual methods, from the point of view of the chosen criteria, the method of the Weighted sum product - WSA was used, see Table 6.

The order assessed by the WSA method is as follows that the most suitable method of the bulk density determination according to the chosen criteria was the Troxler 3440 equipment. Even though the exactness was unsatisfactory, The Troxler equipment 2701 is the second most suitable method. The third place is taken by the method of bulk density determination by using drill core. The exactness is the most important in this case. Despite using the AHP assessment, it is clear that destructive method is necessary especially when handing the construction for determination of the layer thickness and their connections. According to the standard, the drill cores are carried out every 1,500 m^2. It is without doubt that by this method one obtains the most reliable values of the bulk density; still this method is both equipment and time-demanding.

4. Conclusion

From the analysis of the measured results of the compacted bulk densities, or more precisely the compaction degree, it can be stated that the biggest informative value has the determination of the compaction intensity by the destructive method on the drill cores. In the case of resolving tests, it is possible to use only this method. The non-destructive method, when the compaction degree is determined by the Troxler 3440 equipment, is feasible in the process of compacting, and even when controlling the work performed. The results of the compaction intensity can differ from the real degree by cca 1 %. When measuring the degree of compaction by the radio-metric probe, it would be advisable that the obtained compaction degree was at the level of 99 %, or rather 97 % depending on the type of the asphalt mixture. From the presented results, it can be stated that the second equipment, i. e. Troxler 2701 probe, does not reach the sufficient exactness, not even the repeating measurement. It is neither possible to recommend it for the determination of the final compaction degree of the construction asphalt layer, nor for the continuous inspection of the compaction. Based on the obtained results it can be concluded that the AHP as a tool for the multi-criteria evaluation is not suitable for this particular case.

Acknowledgement

The paper was elaborated with the support of the Competence Centre programme of the Technology Agency of the Czech Republic (TACR) within the Centre for Effective and Sustainable Transport Infrastructure (CESTI), the project number TE01020168, project VEGA 1/0501/17 and ITMS: 26240220084 - University Science park Slovak University of Technology in Bratislava.

References

[1] STEHLIK, D., DASEK, O., HYZL, P., COUFALIK, P., KRCMOVA, I., VARAUS, M.: Pavement Construction Using Road Waste Building Material - From a Model to Reality. Road Materials and Pavement Design, 16, 314-329, 2015. http://dx.doi.org/10.1080/14680629.2015.1029680

[2] KOMACKA, J., REMISOVA, E., LIU, G., LEEGWATER, G., NIELSEN, E.:. Influence of Reclaimed Asphalt with Polymer Modified Bitumen on Properties of Different Asphalts for a Wearing Course. Proceedings of the 3rd Conference on Transportation Infrastructure, Italy, 179 -185, 2014.

[3] WU, S. P., ZHONG, J. J., ZHU, J. Q., WANG, D. M.: Influence of Demolition Waste Used as Recycled Aggregate on Performance of Asphalt Mixture. Road Materials and Pavement Design, 14, 679-688, 2013. http://dx.doi.org/10.1080/14680629.2013.779304

[4] REMISOVA, E.: Effect of Film Thickness on Resistance to Permanent Deformation in Asphalt Mixtures. Journal of Road and Bridge Engineering, 10(4), 333-339, 2015. http://dx.doi.org/10.3846/bjrbe.2015.42

[5] HAMZAH, M. O., KAKAR, M. R., QUADRI, S. A., VALENTIN, J.: Quantification of Moisture Sensitivity of Warm Mix Asphalt Using Image Analysis Technique. Journal of Cleaner Production, 68, 200-208, 2014. http://dx.doi.org/10.1016/j.jclepro.2013.12.072

[6] ARAMBULA, E., MASAD, E., MARTIN, A.: Influence of Air Void Distribution on the Moisture Susceptibility of Asphalt Mixes. Journal of Materials in Civil Engineering, 19(8), 655-664, 2007. http://dx.doi.org/10.1061/(ASCE)0899-1561

[7] KONERU, S., MASAD, E., RAJAGOPAL, K. R.: A Thermomechanical Framework for Modeling the Compaction of Asphalt Mixes. Mechanics of Materials, 40(10), 846-864, 2008. http://dx.doi.org/ 10.1016/j.mechmat.2008.03.008

[8] TAN, Y. Q., WANG, H. P., MA, S. J., XU, H. U.: Quality Control of Asphalt Pavement Compaction using Fibre Bragg Grating Sensing Technology. Construction and Building Materials, 54, 53-59, 2014. http://dx.doi.org/10.1016/j.conbuildmat.2013.12.032

[9] BEAINY, F., COMMURI, S., ZAMAN, M.: Quality Assurance of Hot Mix Asphalt Pavements Using the Intelligent Asphalt Compaction Analyzer. Journal of Construction Engineering and management, 138(2), 178-187, 2012. http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000420

[10] BEAINY, F., SINGH, D., COMMURI, S., ZAMAN, M.: Laboratory and Field Study on Compaction Quality of an Asphalt Pavement. International Journal of Pavement Research and Technology, 7(5), 317-332, 2014. http://dx.doi.org/ 10.6135/ijprt.org.tw/2014.7(5).317

[11] SPILAK, L.: Determine the Degree of Compaction of Asphalt Layers Destructive and Non-Destructive Methods, and Comparing Them (in Slovak). Dissertation, VSB - Technical University of Ostrava, Czech Republic, 2015.

Libor Izvolt - Jozef Harusinec - Michal Smalo*

OPTIMISATION OF TRANSITION AREAS BETWEEN BALLASTLESS TRACK AND BALLASTED TRACK IN THE AREA OF THE TUNNEL TURECKY VRCH

The paper follows previously published papers in Communications – Scientific Letters of the University of Zilina No. 4/2014, No. 4/2015 and No. 4/2016, where were presented the results of the diagnostics of the track geometry of the experimental sections surrounding the portals of a new building of the Tunnel Turecky vrch. Based on the data obtained by comprehensive diagnostics, there can be stated that the problematic section in the slab track are the transition areas allowing the transfer of the vehicle between the line structures with a different bending stiffness. The comprehensive diagnostics was realised by the continuous method (measuring trolley KRAB), the results of which were verified by the results of the measuring car of ZSR and by the geodetic monitoring of the layout of the track, as well. Mathematical modelling of the transition areas was realized in the environment of the software ANSYS with the aim to optimise the structural solution of the existing transition areas, as well as of in the future built ones, within the modernization of the railway infrastructure of the Slovak Republic. The model of the existing transition area was structurally optimised –models were realised with application of the stiffening rails and with usage of the under sleeper pads. The model of the current state and the models of the optimised structural solutions of the transition areas were subsequently compared by the values of the vertical displacement, vertical speed and vertical acceleration.

Keywords: *railway track, railway superstructure, ballastless track, ballasted track, numerical modelling of transition area*

1. Introduction

Until now, the track layout and the track geometry parameters have been monitored nine times with the time interval of 6 months in the area of the portals of the Tunnel Turecky vrch. The monitoring was carried out within the comprehensive diagnostics with a continuous method by a manual measuring trolley *KRAB™-Light*. There was an initial measurement prior to the construction commissioning (MSO) and eight operational measurements (PO1, PO2, PO3, PO4, PO5, PO6, PO7 and PO8). The quality mark with the worst development tendency seems to be the quality mark of the vertical profile of the lines VP, VL QM_{VK} based on measurements implemented thus far. A problematic section is, in the long term, the transition area from the slab track (PJD) of the type RHEDA 2000® to the ballasted track (the track skeleton is placed into the ballasted trackbed) located on the southern portal of the Tunnel Turecky vrch [1]. The trend lines were created by means of the values VL and VP measured within PO5, PO6, PO7 and PO8 in the section of the transition from the slab track to the ballasted track. Based on these trend lines, it can be assumed that the maintenance limit (±6 mm) of the parameter VL was expected to be reached approximately in 9/2017 and of the parameter VP will be reached in 5/2018 in the section of the track Nr. 1 in the southern portal of the tunnel. In the track Nr. 2 in the southern portal of the tunnel the values of the maintenance limit of the VL and VP parameters were expected to be reached approximately in 9/2017, or cca. in 6/2018, respectively, [2]. These overrun levels of the maintenance limits were detected within the third measurement campaign of the measuring trolley of the ZSR (in the autumn of 2017). The trend lines of the other quality marks do not indicate any necessity of the corrective action in the coming years and the parameter values seem to be more or less stabilised. The local defects were recorded within the safety limits before the installations of the sections. Specifically, there were the sections with the slab track (9 local defects), with the ballasted track (30 local defects) and in the transition area of the track Nr. 2 in the southern portal (1 local defect). Those defects were removed prior to installation of the sections. This was also verified by the results of the measurement PO1 within which no local defects were recorded. The local defects were recorded in the transition area of the track Nr. 1 in the southern portal (2 local defects in the measurement PO2 and 4 local defects in the measurements PO3 and PO4) within the other operational measurements. After the executions of the corrective actions, no other local defects were recorded [3].

The railway infrastructure manager has provided so far the results from eleven measurement campaigns of the measuring trolley. The campaigns were executed in the years 2013 (in summer and autumn), 2014 (in spring, summer and autumn), 2015 (in spring, summer and autumn) and 2016 (in spring, summer and autumn). On the measurement results in the sections with the slab track one can the recurrence of the defect of the longitudinal elevation deviation of the left (VL) and the right line (VP) in the section 1.1 (track Nr. 1 of the southern portal of the tunnel) and the defects of the longitudinal elevation deviation of the left line VL in the section 2.2 (track Nr. 2 of the northern portal of the tunnel). The defects of the longitudinal elevation deviation of the left (right) line VL, VP were recorded in the track Nr. 2 of the

* [1]Libor Izvolt, [2]Jozef Harusinec, [1]Michal Smalo
[1]Department of Railway Engineering and Track Management, Faculty of Civil Engineering, University of Zilina, Slovakia
[2]Department of Transport and Handling Machines, Faculty of Mechanical Engineering, University of Zilina, Slovakia
E-mail: libor.izvolt@svf.uniza.sk

transition area of the southern portal of the tunnel (section 2.1). The direction fault of the left line SL in the transition area of the adjacent track Nr. 1 (section 1.1) has been detected once so far. In the area of the northern portal of the tunnel in the transition area of the track Nr. 1 defects of the longitudinal elevation deviation VL, VP were recorded in both lines, [4].

The measurements of the track layout (geodetic surveying) have been realised 6 times so far - the initial measurement in the autumn of 2012, the second measurement in the spring of 2013, the third measurement in the autumn of 2013, the fourth measurement in the spring of 2014, the fifth measurement in the spring of 2015 and the sixth measurement in the spring of 2016. The outputs of the geodetic monitoring of the track proved the results of the comprehensive diagnostics by the continuous method and the results acquired from the measuring trolley of ZSR, as well. Based on the implemented analysis of the measurements, it can be stated that no significant changes in the horizontal or vertical profile of both lines in the section with the slab track were recorded. The same applies also for the slab track where rest the particular rails. The recorded lateral or longitudinal displacements of the track are within the specified measurement accuracy. The lateral displacements of both tracks, within the range from 0 to 12 mm in the area of the southern portal and from 15 to 20 mm in the area of the northern portal of the tunnel, were diagnosed in the same section with the ballasted track [5].

The diagnostics by a visual method of the sections with the slab track was realised in the summer of 2015 and the size of the cracks in the reinforced concrete plate of the slab track and in the reinforced concrete blocks was documented. The results of this diagnostics were compared to the documentation of the railway infrastructure manager processed based on the diagnostics realised in the autumn of 2012 and in the spring of 2013. The highest frequency of cracks was detected in the section of the track Nr. 2 of the northern portal of the tunnel in the area between the railway bridge and the transition area. Based on the comparison to the results of 2013, it is obvious that the extent and amount of the cracks do not show any changes. In the section with the highest frequency of cracks no significant deviations by the measuring trolley *KRAB™-Light*, measuring car and neither by the geodetic monitoring were measured.

Based on evaluation of measurements of the track geometry and the layout of the track, it can be stated that the monitored sections with the slab track are stabilised and that they do not show any relevant changes of the particular parameters between the individual measurements. The sections with the ballasted track do show changes; however, those changes result from the character of this construction. The ballasted track has the higher stiffness values in comparison to the slab track. Therefore, the ballasted track needs the adjustment of the track geometry, cleaning and supplement with the ballast material and other maintenance and repair works in certain intervals (depending on the operational load) with the aim to maintain the quality parameters throughout the life span of the track. However, the measurements of the track geometry pointed to the irregularities in the transition areas between the ballasted track and the slab track, especially in both of the tracks in the area of the southern portal of the Tunnel Turecky vrch. These results were subsequently proved by the geodetic survey and also by the outputs of the measuring trolley. The decreases recorded in these transition areas, from the beginning of the measurements, resulted in the limit overruns of the vertical profile quality marks of the rails VP, VL QM_{VK}, as well as in the appearance of the local defects. Based on these facts, the railway infrastructure manager executed the corrective action. The corrective action slightly slowed down the tendencies of the defect development; however, the track geometry quality of both lines is still being reduced. As it was expected, the next limit overrun occurred during the second half of the year 2017 and resulted in the necessity to execute another corrective action.

The transition areas, between the ballasted track and the slab track in the railway infrastructure of the Slovak Republic, have been for the first time exposed to such difficult conditions (the operational load in combination with the passenger train speeds up to 160 kmph) especially in the area of the portals of the Tunnel Turecky vrch. The other operated transition areas, e.g. in the area of the portals of Bratislava tunnels Nr. 1 and Nr. 2, are also exposed to a high operational load, however, in combination with significantly lower train speeds (and by that also significantly lower dynamic forces affecting the transition area construction). There is the expectation that the other sections with the slab track will be built within the modernization of the railway network of the Slovak Republic. Therefore, it would be convenient to improve the further adjusted transition areas structurally and, in this way, to eliminate the defects formation in the track geometry. This will exclude or at least eliminate the corrective actions on the exposed track sections (bridge structures, tunnels). The engineers, who currently work on the projects of modernisation of the railway infrastructure of ZSR in Slovakia, normally design those constructions at the exposed track sections.

2. The numerical modelling of the transition areas

The numerical modelling was realised with the aim to structurally optimise the transition areas. The modelling was based on the faults and defects occurrence in the transition areas from the slab track to the ballasted track recorded within the diagnostics of the monitored sections in the areas of the portals of the Tunnel Turecky vrch (especially in the area of the southern portal). The optimisation of the construction solution of the transition areas was realised to enhance the continuity of the vehicle transfer between the two construction systems of the railway with different stiffness. In this way, the assumptions will also be created for elimination or minimisation of the extent of fault and defect formation in the transition areas and their immediate surrounding.

In the case of the ballasted track the wheel forces are transferred from the rail through the sleeper to the track ballast and the subgrade. Zimmermann [6] introduced a model of calculation for the specification of the rail lowering, which is based on the elastic embedded longitudinal beam. This beam reacts to the track ballast and the sub-ballast on all the sleepers as to springs of the same size and independent from each other, which are characterized by means of the modulus of subsoil reaction *C*. The bending coefficient of the sleeper support *C* depends on the rail lowering and surface pressure between the sleeper and the track ballast. It is expressed by the formula [7]:

$$C = \frac{p}{z} \ (\text{N.mm}^{-3}) \tag{1}$$

where:
C – bending coefficient of the sleeper support, (N.mm^{-3}),
p – surface pressure between the sleeper and the ballasted trackbed, (N.mm^{-2}) and
z – rail lowering, (mm).

Modulus of the subsoil reaction is determined from the measurements of the rail lowering under the static load by a wheel (wheel power) in accordance with the formula:

$$C = \frac{Q}{4 \cdot b \cdot z} \cdot \sqrt[3]{\frac{Q}{E \cdot I \cdot z}} \ (\text{N.mm}^{-1}) \tag{2}$$

where:
Q – the wheel force, (N),
E – modulus of elasticity, (N.mm^2),
I – moment of inertia of the rail, (mm^4),
b – fictional width of the elastic embedded longitudinal beam, (mm), whereby

$$b = \frac{F}{a} \ (\text{N.mm}^{-1}) \tag{3}$$

where:
F – half efficient bottom surface of the sleeper, (mm2) and
a – axial distance of the sleepers, (mm).

For the railway track with the long-term operation the modulus of subsoil reaction is ranging from 0.05 up to 0.15 N.mm^{-3}. For the new buildings, the value of the modulus of subsoil reaction is considered to range from 0.3 up to 0.4 N.mm^{-3} [7].

In the case of the slab track, the consolidated (bonded) and unconsolidated bearing layers contribute to the elasticity with only a small portion. Therefore, in the model-based calculation, as the elastic components are considered only the resilient pads (a part of the rail fixing placed under the rail). The resilient pads enter the calculation with the so-called stiffness of the fixing node c, which is being calculated from the formula:

$$c = \frac{S}{z} \ (\text{N.mm}^{-1}) \tag{4}$$

where:
S – force acting on the fixing node, (N).

Unlike for the ballasted track, with a layer of the ballasted trackbed, for the slab track is instead of the modulus of subsoil reaction C used the stiffness of the fixing node c. For the relationship between the modulus of the subsoil reaction C and the stiffness of the fixing node c applies the following formula:

$$C = \frac{S}{F \cdot z} = \frac{c}{F} \ (\text{N.mm}^{-3}) \tag{5}$$

where:
F – half efficient bottom surface of the sleeper, (mm^2).

The static track stiffness c_G is unlike for the stiffness of the fixing node c not defined by means of the force acting on the fixing node, but by means of the static wheel force Q. The track stiffness is calculated in accordance with the formula:

$$c_G = \frac{Q}{z} \ (\text{N.mm}^{-1}) \tag{6}$$

where:
c_G – the static track stiffness, (N.mm^{-1}).

After substituting Equation (4) and Equation (5) into the Equation (2), one obtains the formula for the relation between the stiffness of the fixing node c and the track stiffness c_G for the slab track:

$$c_G = \sqrt[4]{\left(\frac{4 \cdot c}{a}\right) \cdot E \cdot I} \ (\text{N.mm}^{-1}) \tag{7}$$

The dimensioning of the slab track is being carried out with the aim to ensure the sufficient elasticity of the track form. The tests have shown that the rail lowering z with approximately 1.5 mm and the static track stiffness c_G of the value 64 ± 5 kN.mm^{-1} correspond most closely to this objective.

The formula for the modulus of subsoil reaction C and the track stiffness c_G for the ballasted track is obtained by inserting the Equation (6) into the Equation (2):

$$c_G = \sqrt[4]{64 \cdot E \cdot I \cdot b^3 \cdot C^3} \ (\text{N.mm}^{-3}) \tag{8}$$

If the Equation (6) for the stiffness of the fixing node c is adjusted, one obtains:

$$c = \sqrt[3]{\frac{c_G^4 \cdot a^3}{64 \cdot E \cdot I}} \ (\text{N.mm}^{-1}) \tag{9}$$

Both construction systems – the conventional track (the so-called ballasted track), as well as the slab track (the so-called ballast-free track), are the structures proved by operation during decades. The transition areas represent the potentially problematic places in the line structure where, in general, every change in construction can result in change of its dynamic load. The increase of the dynamic load directly influences the track position and most often results into the rail lowering which increases over time.

The initial (comparing) model for the realization of the numerical modelling is the transition area without any components – so the transit from the slab track directly to the ballasted track (model A). Another stage of the numerical modelling was creation of the transition area model in the surrounding of the southern portal of the Tunnel Turecky vrch. The transition area is created as a reinforced concrete tub of the concrete C 30/37, 20 m long, longitudinally graded and covered with an elastic mat (model B). The thickness of the ballast bed in this type of transition area, is decreasing under the sleepers in the direction to the slab track, whereby the stiffness of the trackbed is gradually increasing. The last stage of the process of the numerical modelling was the creation of the models optimising the transition area applied in the surrounding of the southern portal of the Tunnel Turecky vrch by adding other structural elements. Those were:

- Two stiffening rails (profile 49E1) 20 m long. The rails are placed in the length of 5 m in the slab track section and in the length of 15 m in the ballasted track section – model C.
- Under sleeper pads *Sylomer SLB 2210 G* (producer *Getzner*) 10 mm thick. They are placed in the transition tub directly behind the slab track, in the section with the length that the fastest train passes in 0.5 s. This, at speed of 160 kmph, represents the section 22.2 m long. The pads, based on recommendations from the supplier, were used in the length

Table 1 *Three-dimensional models for monitoring of the dynamic behaviour of different types of transition areas*

Model definition	Transition area solution
Model A (initial model)	• without any transition area (the direct transfer of the vehicle from the slab track to the ballasted track)
Model B (current state)	• 20 m long reinforced concrete tub with the longitudinal gradient (the track ballast of variable thickness) covered by the elastic mat
Model C (current state + two stiffening rails)	• 20 m long reinforced concrete tub with the longitudinal gradient (the track ballast of variable thickness) covered by the elastic mat and • two stiffening rails (profile 49E1), 20 m long, placed in the length of 5m in the slab track section and in the length of 15 m in the ballasted track section
Model D (current state + under sleeper pads in the transition tub)	• 20 m long reinforced concrete tub with the longitudinal gradient (the track ballast of variable thickness) covered by the elastic mat and • application of the under sleeper pads *Sylomer SLB 2210 G*, 10 mm thick, placed in the whole length of the transition area (20 m)
Model E (current state + under sleeper pads partially in the slab track section and partially in the transition area constructed in the reinforced concrete tub)	• 20m long reinforced concrete tub with the longitudinal gradient (the track ballast of variable thickness) covered by the elastic mat and • application of the under sleeper pads *Sylomer SLB 2210 G*, 10 mm thick, placed in the section 22.2 m long, whereby the half of the section length is on the slab track (11.1 m – 18 sleepers) and the half of the section length is in the transition section in the shape of the reinforced concrete tub (11.1 m – 18 sleepers)

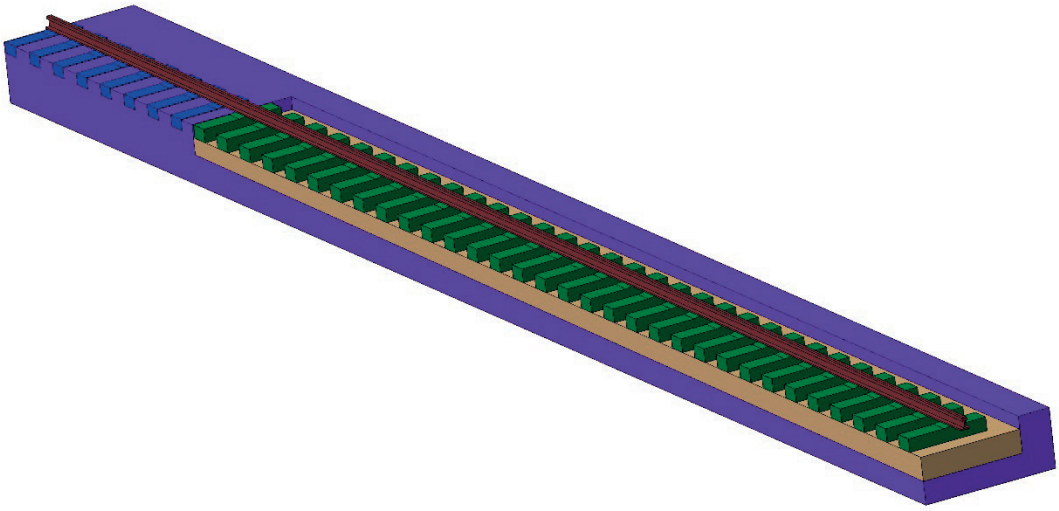


Figure 1 *Model of the transition area with the reinforced concrete tub in the program CATIA*

of 20 m (i.e. in the whole length of the transition area) – model D.

- Under sleeper pads *Sylomer SLB 2210 G* (producer *Getzner*) 10 mm thick. They are placed in the slab track section and in the transition area in the shape of the reinforced concrete tub directly behind the slab track in the section with the length that the fastest train passes in 0.5 s. This, at speed of 160 kmph, represents the section 22.2 m long. The section was divided into two parts of the same length. The first part is situated in the section with the slab track (11.1 m – 18 sleepers) and the second one in the transition area (in the reinforced concrete tub) immediately behind the slab track (11.1 m – 18 sleepers). The technical and technological possibilities of the installation of the under sleeper pads into the slab track are to be consulted with the supplier of the slab track system - model E.

During the application process of the under sleeper pads, the supplier of the pads of the producer *Getzner* - the company *Hydrobeton,* Ltd was consulted.

Initially, the aim of the numerical modelling was the modelling of the transition area with the stabilization of the railway ballast by chemical binders, as well. However, because of the large variation of the input data values, the numerical modelling was finally not realised since the relevance of the comparisons to other numerical models was questionable.

The overview of the numerical models for monitoring of the dynamic behaviour of different types of transition areas is introduced in Table 1.

The numerical modelling of the transition areas was carried out by means of the three-dimensional (3-D) models A, B, C, D and E created in the program *CATIA V5-6R2016* (Figure 1) based on the two-dimensional drawings from the program *Auto CAD 2016*. The individual models were subsequently transferred into the working environment of the program *ANSYS R17.2*. by means of the exchange files STEP.

The models were initially implemented as the model of the 20 m long section with the slab track, the 20 m long transition area (with the exception of the model without the transition

Table 2 *Material characteristics of the line structure elements*

Line structure element	Volume weight ($kg.m^{-3}$)	Young Modulus (GPa)	Poisson's ratio (-)
Rail	6186	210	0.3
Rail pad	950	0.21	0.3
Under sleeper pad	4200	1.3	0.42
Sleeper (in the section with the track ballast)	2054	30	0.2
Double block sleeper (in the slab track section)	2400	70	0.2
Track ballast	1800	0.2	0.1
Material of the sub-ballast layer	2200	0.3	0.2
Embankment material (in the section with the ballasted track)	2200	0.4	0.2
Concrete bearing layer (CBL)	2400	34	0.2
Hydraulically bound bearing layer (HBL)	2400	5	0.2
Frost protection layer (FPL)	2400	0.12	0.2

Table 3 *Results of the static load tests in the area of the southern portal of the tunnel*

Km position	Structural part	Characteristic of the permanent way construction	Track number	Value of the static deformation modulus $E_{s,i}$ (MPa)	Degree of compaction E_{def2}/E_{def1}
102.365	Subgrade	Ballasted track	2	225.0	4.33
	Subgrade	Ballasted track		250.0	3.50
102.375	Top surface of the sub-ballast	Ballasted track	1	75.0	3,60
	Ballast bed	Ballasted track		93.8	3.30
102.425	Subgrade	Ballasted track	1	173.1	3.00
102.465	Ballast bed	Transition area	1	22.0 [1)]	1.67

Note:

[1)] The value was influenced by the elastic mat, which was deformed under the applied load.

area) and the 20 m long ballasted track section. At the start of tackling the issue was presumed that one of the problems can be the insufficient length of the transition area whereas the foreign regulations, e.g. [8], recommend the length of the transition area in accordance with the formula:

$$L_{\min} = 0.5 \cdot \frac{V_{\max}}{3.6} \text{ (m)} \qquad (10)$$

where:

L_{min} – minimum recommended length of the transition area, (m) and

V_{max} – maximum speed in the section, ($km.h^{-1}$).

The transition areas in the surrounding of the portals of the Tunnel Turecky vrch should have, according to Equation (10), the length minimum 22.2 m for the line speed of 160 kmph. This means that the transition areas are 2.2 m shorter than the ones recommended by the foreign regulations. However, in the course of tackling the issue, it was proved that the increased length of the transition area, e.g. to the length of 25 m, would bring no or only a negligible improvement. Therefore, the numerical modelling of the transition areas with an optimised length was not realised. The numerical modelling concentrated on the optimisation of the critical place of the transition areas – the place of the end of the slab track and the beginning of the transition area.

Due to the challenge and the long time needed to launch each of the models and also in view of the fact that there occurred no significant deviations of the values in the section with the slab track and the part of the transition area in the shape of the reinforced concrete tub (circa 15 m) adjacent to the ballasted track, the model length was reduced. The model was ca. 5 m long section with the slab track and the transition area in the whole length. That was especially the place of the end of the slab track and the beginning of the transition area that proved to be crucial. For that reason, the shortening of the model length did not have any influence on the result of the numerical modelling of the transition areas concerned.

Each model was implemented only for one rail on which the wheel of the railway vehicle moves. Since this is a symmetrical problem and the railway line is in a direct section and without any super elevation, there was no point in realization of the model on the track as a whole. The so called "mirroring"" of the model would cause the increase in the number of nodes, elements and couplings, and logically also the extension of mathematical calculations, in connection with the slowing-down the numerical modelling process.

Each model was simultaneously made by application of the concerned material characteristics, introduced in Table 2, [9], [10], and the values of the static deformation modulus $E_{s,I}$, which were identified by the static load tests on the particular levels of subsoil of the monitored structure. Those values were acquired

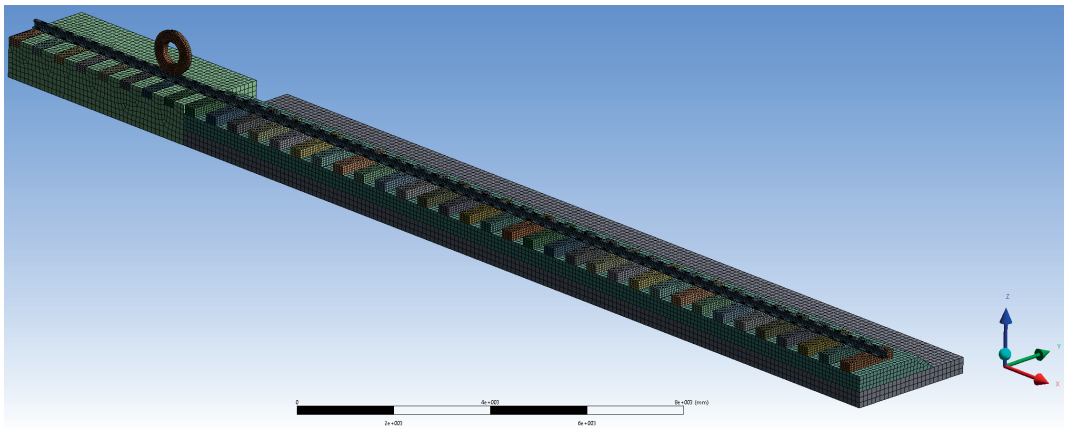


Figure 2 *Model A - (initial model)*

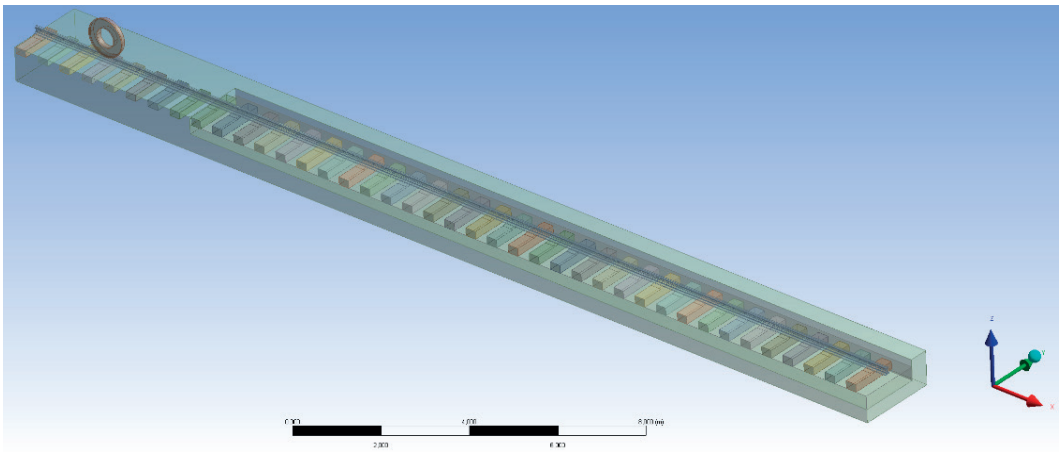


Figure 3 *Model B - (current state)*

within the quality verification of the realised works during the installation of the modernized line structure, in terms of the methodology of the Regulation by ZSR S4 Subgrade, Appendix No. 20. That diagnostics of the quality of the realised geotechnical works was carried out by the employees of the Department of the Railway Engineering and Track Management during the building of the transition sections in the area of the portals of the Tunnel Turecky vrch. The values of the static deformation modulus $E_{s,i}$ which were identified by the static load tests, are introduced in Table 3. The values of the static deformation modulus $E_{s,i}$ are linearly interpolated between the places of the measuring for the needs of the numerical modelling.

The model of the transition area in the surrounding of the southern portal of the Tunnel Turecky vrch (model B) comprises of 101 369 node points 74 598 elements and 129 contact couplings. Each calculation was divided into 120 integration steps. Since this is a non-linear problem, it was necessary to carry out from 337 up to 456 integration steps for each calculation.

For the value of the time interval t is valid (for the vehicle speed 160 kmph) the formula:

$$t = \frac{s}{v} = \frac{0.065}{44.44} = 0.0014625\ s \tag{11}$$

where:
s – length of the iterative step, (m),
v – vehicle speed, (m.s^{-1}),

The value 0.065 m was chosen as the length of the iterative step. It is considered, that the change of the vertical position parameter on this interval causes the change of the vertical speed. For the value of the vertical speed v_{zv} applies the formula:

$$v_{zv} = \frac{z_n - z_{n-1}}{t}\ (\text{m.s}^{-1}) \tag{12}$$

where:
z_n – vertical displacement in point n, (m),
z_{n-1} – vertical displacement in point n -1, (m),
t – time (s).

For the value of the vertical acceleration a_{zv} applies the formula:

$$a_{zv} = \frac{v_{zv,n} - v_{zv,n-1}}{t}\ (\text{m.s}^{-2}) \tag{13}$$

where:
$v_{zv,n}$ – vertical speed in the point n, (m),
$v_{zv,n-1}$ – vertical speed in the point n-1, (m).

Creation of the individual models started by the creation of basic components of the railway line (the rails, fastenings, sleepers, trackbed layers, models of the check rails and under sleeper pads). A separate task was the creation and testing of the effects of the body movement on the rail embedded on the subgrades of different material characteristics and the creation of contact pairs between these materials. The basic objective was to reflect, as closely as possible, the reality of the load transfer.

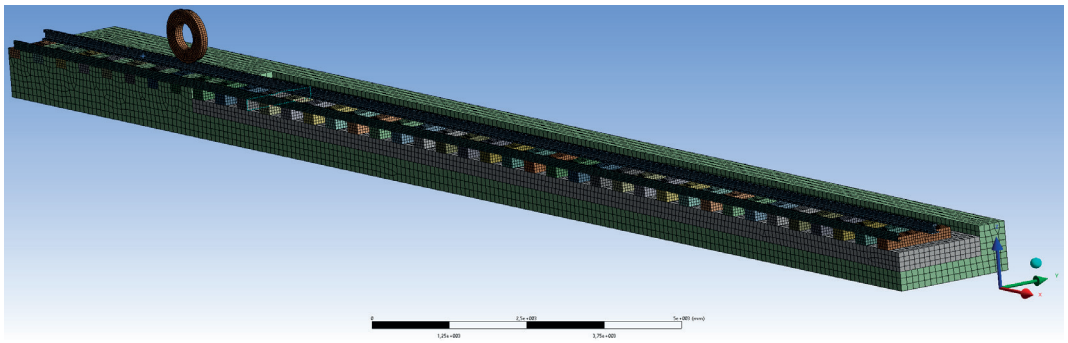


Figure 4 *Model C – (current state + two stiffening rails)*

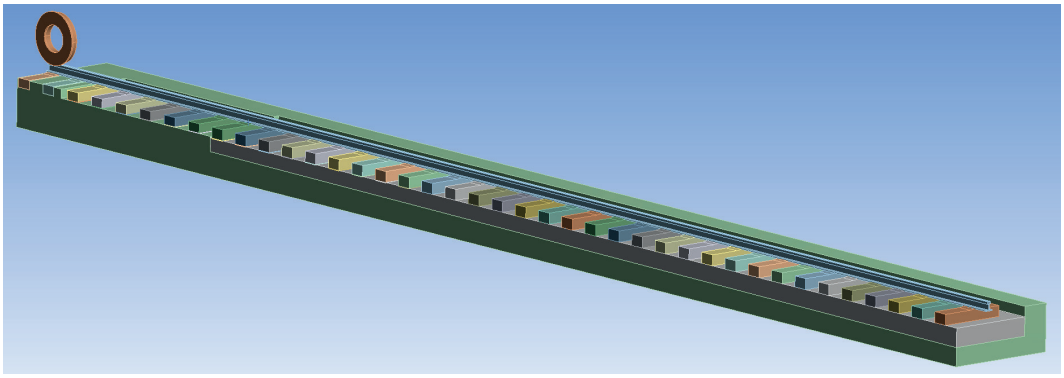


Figure 5 *Model D – (current state + under sleeper pads in the transition tub)*

Subsequently, the moving wheel was given the vertical power adequate to the axle load (half of the value 225 kN on the axle, so 112.5 kN), which replaced the original unit force.

A model of the current state – labelled as model B (Figure 3) was gradually created, into which were added other structural elements. T the model C was created with the two stiffening rails (Figure 4) and models D and E with application of the under sleeper pads *Sylomer SLB 2210 G* (Figure 5).

Finally, the transition area from model B was omitted, creating the initial (comparing) model A. This model should serve as a comparison of impairment of the individual parameters, if there is no transition area applied on the line (Figure 2).

3. Evaluation the numerical modelling results

Comparison of the results of the numerical modelling was realised by comparison of values of the vertical displacement, vertical speed and vertical acceleration. The graphic presentation of the introduced parameters is depicted in Figure 6, Figure 7 and Figure 8. For better visualisation, the graphs illustrate only a segment of the modelled section where the biggest parameter changes are. They depict the last sleeper of the slab track and the first three sleepers (in case of the vertical displacement the first six sleepers) of the transition area.

The numerical modelling of the transition areas was implemented by the load (wheel) transfer between areas with different stiffness values. The graph area is depicted in such a way that the value 0 of the horizontal axis represents the axis of the last sleeper in the slab track section. For the better visualisation the colour coding was used for the area with the slab track (blue background) and the transition area (white background, with the exception of model A, which is marked out with a chain line and in which behind the slab track follows immediately the ballasted track). For the better orientation, the rail supports (double block sleepers in the section with the slab track as well as the reinforced concrete sleepers in the section with the track ballast) are marked out by a vertical dashed line in each graph.

According to the initial assumption, the worst results are recorded for model A (without the transition area) where the line structure is the most stressed. Significantly better values are depicted on the model of the current state (model B), where the application of the transition tub lowered the values of the vertical displacement and the vertical speed, as well. Unlike the initial assumption, the better results are obtained from the model C (with a couple of stiffening rails). In that case, an improvement of the values compared to model B was recorded, but only a minimal one. As an appropriate solution, based on the carried out numerical modelling, seems to be the application of the under sleeper pads *Sylomer SLB 2210 G.* These under sleeper pads reduced the values of the virtual displacement and also the virtual speed in both numerical models (models D and E) in comparison to the current state (model B), as well as compared to the model of the current state improved by the application of the stiffening rails (model C). Model E seems to be optimal, as the transition between the slab track and the transition area is the most fluent for all the three monitored parameters.

4. Conclusion

The three-dimensional models, created for monitoring of the dynamic behaviour of different types of the transition areas from the slab track to the ballasted track, are by means of the

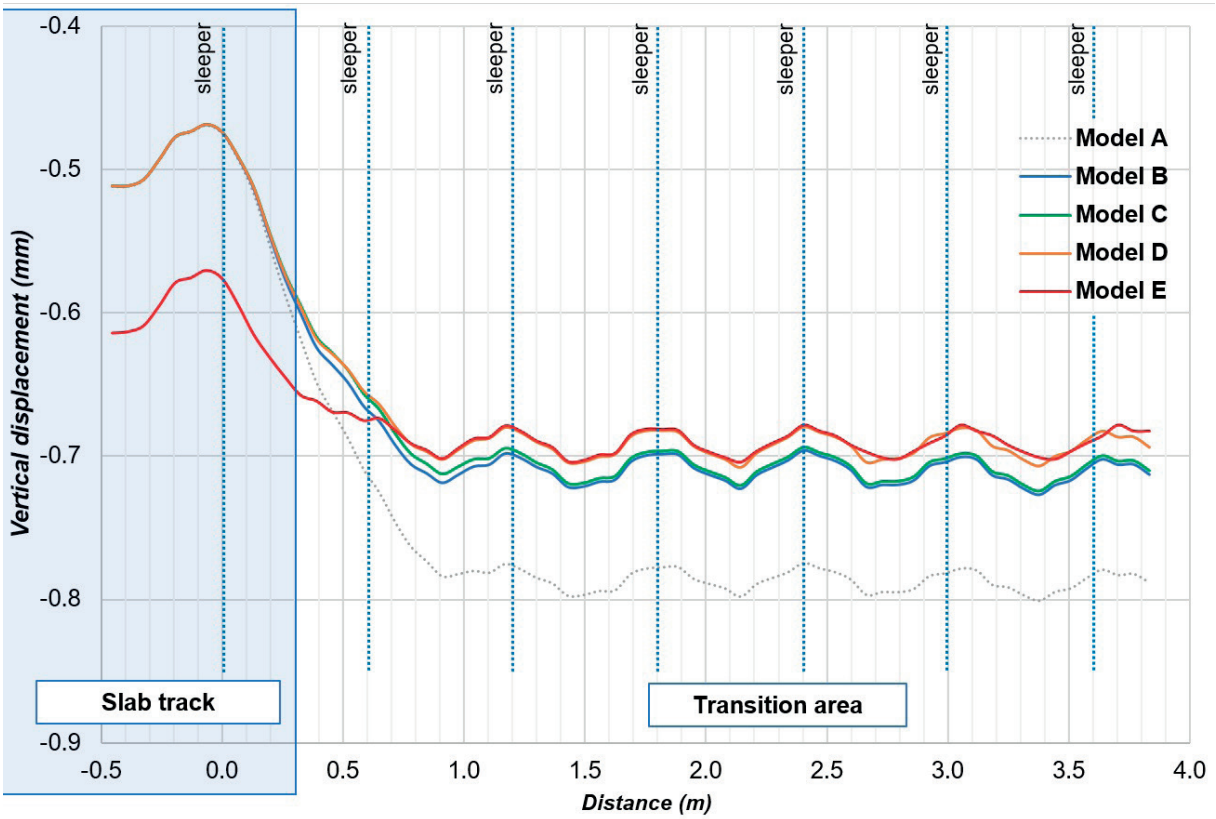


Figure 6 Vertical displacement - model comparison

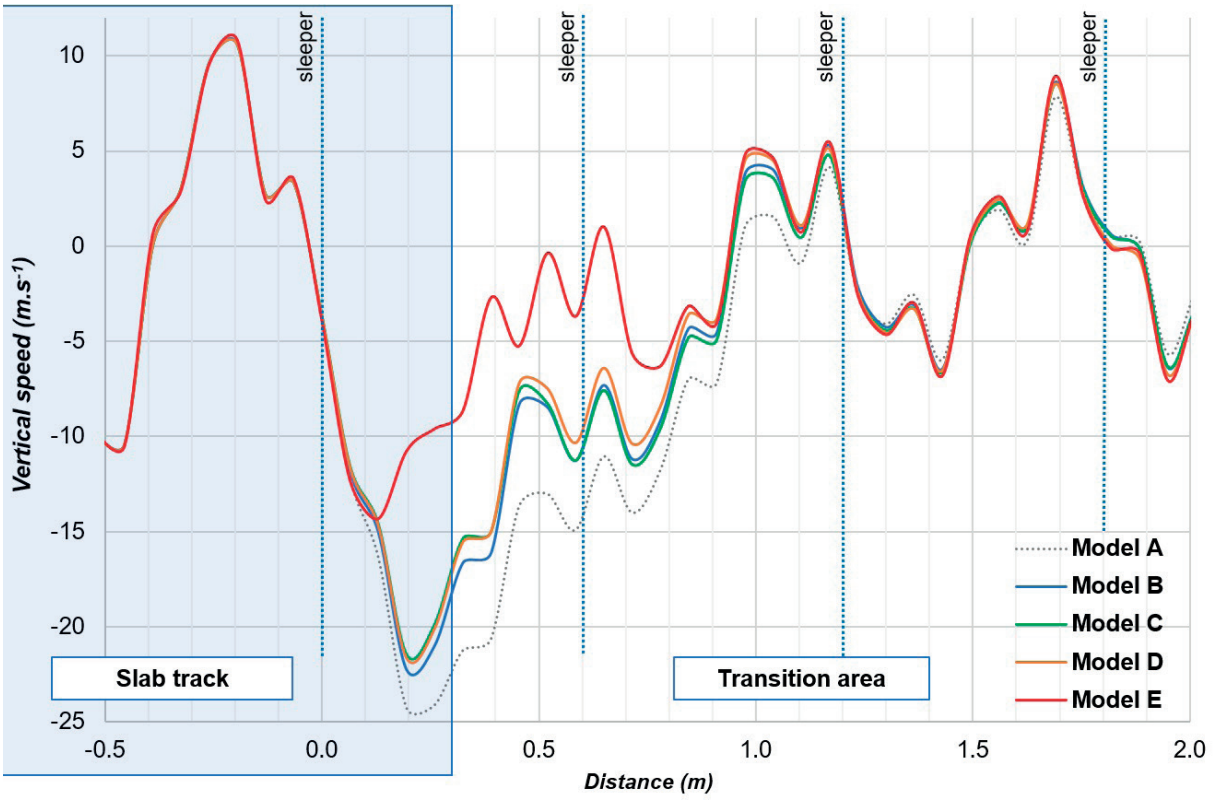


Figure 7 Vertical speed - model comparison

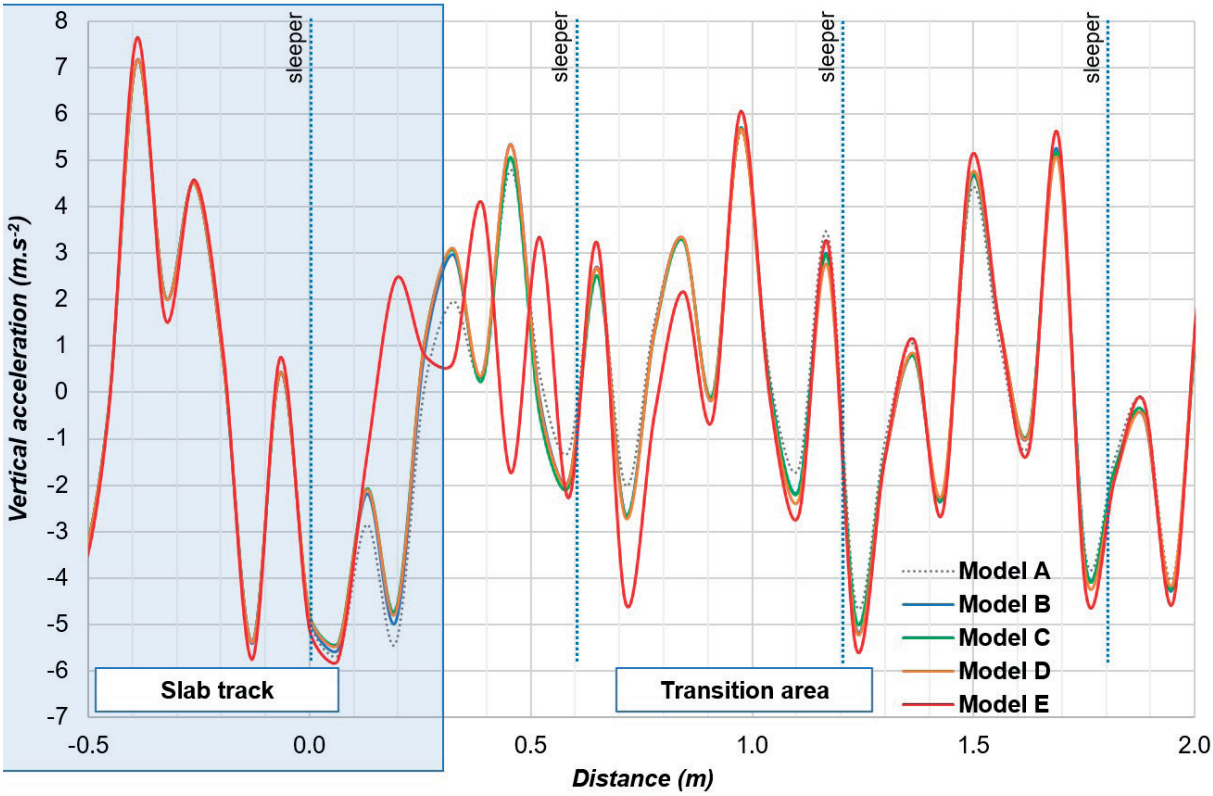


Figure 8 *Vertical acceleration - model comparison*

numerical modelling confronted to each other in the environment of the program *ANSYS*. The basis for their confrontation are the values of the vertical displacement, vertical speed and vertical acceleration.

Model B is created according to the transition area of the southern portal of the Tunnel Turecky vrch using the real values of the static deformation modulus of the trackbed and the material characteristics of the particular components of the line structure. The model shows that the currently built transition area is relevant. It is for that reason that the transition area significantly absorbs the dynamic effects of the passing trains, which is also proved by the comparison with the initial model A. This model was created to be compared to the case when there is no transition area applied on the line.

Taking into account the defects occurrence in the track geometry of the transition areas (they were recorded within the diagnostics by the manual measuring trolley *KRAB*^{TM}*-Light,* measuring car of ZSR and stated by the geodetic survey of the layout of the track), other models were created in order to optimise the existing transition area. The models were implemented with the aim to create the transfer of vehicles between the two areas of different stiffness as fluent as possible. The optimisation of the transition area will decrease the dynamic effects on the line structure. In this way, the defects occurrence in the track geometry, which results in the necessity to realise the cost-intensive corrective actions related to the lockouts, diminishing the attractiveness and competitiveness of the rail transport, will be eliminated or minimized.

The improvement of the transfer between the two areas of different stiffness should be originally numerically modelled by means of three models, which would improve the existing transition area of another structural element. These are: a model with a couple of stiffening rails, a model applying the under sleeper pads and a model with the stabilization of the railway ballast of the transition area in the shape of reinforced concrete tub by chemical binders (the so called bonding of the railway ballast). Finally, only the first two cases were modelled, due to the lack of the input data and the large variance of values of the material characteristics of the chemically stabilized railway ballast.

Despite the initial expectations, the application of the stiffening rails was not considered as a significant improvement. The placement of the stiffening rails into the track brought the lowering of the values of the vertical displacement and also the vertical speed, but only minimally. Application of the under sleeper pads has proved to be an interesting way of optimisation of the transition area. The placement of the pads, 10 mm thick, in the whole length of the transition area (model D) brought the relevant improvements. However, there were obvious restraints in the most problematic place – between the last sleeper in the slab track and the first sleeper in the transition area in the shape of reinforced concrete tub. Exactly at those places – in the first metres of this transition area in the contact with the slab track –the defects were detected by the diagnostics resulting in the necessity to carry out the corrective actions. This knowledge led to the optimisation of model D. In this model, the section with the under sleeper pads was partially moved forward into the slab track, a part of the pads was then situated in the area of the slab track and a part in the transition tub. It was exactly this model, which created the most fluent transfer between the areas concerned and it seems to be the most convenient one. However, in the case of the application directly on the line, it would be necessary to resolve technically

and technologically the embedding of the under sleeper pads under the double block sleepers of a monolithic construction directly with the producer (or supplier) of the slab track system.

Based on the realised numerical modelling, it can be stated that the best way to optimise the existing transition area in the area of the southern portal of the Tunnel Turecky vrch is to apply the under sleeper pads in the section of the transition area in the shape of reinforced concrete tub. The application of the under sleeper pads in the transition area seems to be beneficial in case of the future applications of the slab track systems within the railway infrastructure of the Slovak Republic. However, the most efficient would probably be the displacement of the section with the under sleeper pads also into a part of the slab track section.

The under sleeper pads can be globally evaluated as a relevant structural element for the optimisation of the transition areas between the slab track and the ballasted track. However, before their direct implementation into the line it would be necessary to verify this structure by realization of set of numerical modelling with different types of pads and a different section length with pads. It would be desirable to verify the section combination with different types of under sleeper pads, as well, which would be coherent in the transition area or which would partially reach into the slab track section. Subsequently, it would be desirable to verify the results of the numerical modelling on a test section with a real railway operation.

Acknowledgement

The paper contains results of the grant VEGA 1/0275/17 “Application of numerical methods to define the changes of geometrical track position”.

References

[1] SESTAKOVA, J.: Quality of Slab Track Construction - Track Alignment Design and Track Geometry. Civil and Environmental Engineering, 11(1), 2-9, 2015. https://doi.org/10.1515/cee-2015-0001

[2] IZVOLT, L., IZVOLTOVA, J., HODAS, S., MECAR, M., PISCA, P., SESTAKOVA, J., SMALO, M.: Monitoring of Sections of a Non-Conventional Constructions of the Railway Superstructure and the Transition Areas – 9th and 10th Stage. ZSR Modernization of Railway Track Nove Mesto nad Vahom - Puchov, km 100.500 to 159.100, part 24-32-01 Nove Mesto - Trencianske Bohuslavice (in Slovak). Department of Railway Engineering and Track Management, Faculty of Civil Engineering, University of Zilina, Zilina, 2016.

[3] SESTAKOVA, J., MECAR, M.: Evaluation of Track Design and Track Geometry of the Track with Unconventional Structure of Railway Superstructure. XXIV. R-S-P seminar Theoretical Foundation of Civil Engineering (24RSP) (TFoCE 2015), Procedia Engineering, 111, 709-716, 2015. https://doi.org/10.1016/j.proeng.2015.07.136

[4] SMALO, M., IZVOLT, L.: Assessment of Track Quality in Trial Test Sections by Spot and Continuous Method. Communications - Scientific Letters of the University of Zilina, 17(4), 28-33, 2015.

[5] IZVOLTOVA, J., VILLIM, A., PISCA, P.: Analysis of Height Changes of Ballastless Track. Geodesy, Cartography and Geographic Information System. Proceedings of VIII. International conference Tatranske Matliare, Slovakia, 2014.

[6] ZIMMERMANN, H.: Calculation of the Railway Superstructure. Verlag V. Ernst & Sohn, Berlin, 1941.

[7] Normative DB Netz: Requirements on Ballastless Track Construction – 4th edition (in German), 2002.

[8] Normative SZDC S9: Ballastless Track (in Czech), 2012.

[9] SHAHRAKI, M., WARNAKULASOORIYA, CH., WITT, K. J.: Numerical Study of Transition Zone between Ballasted and Ballastless Railway Track. Transportation Geotechnics, 2(3), 58-67, 2015. https://doi.org/10.1016/j.trgeo.2015.05.001

[10] SHAN, Y., ALBERS, B., SAVIDIS, S. A.: Influence of Different Transition Zones on the Dynamic Response of Track-Subgrade Systems. Computers and Geotechnics, 28(48), 21-28, 2013. https://doi.org/10.1016/j.compgeo.2012.09.006

Svitlana Kokhan - Antonina Moskalenko - Oleg Drozdivskyi*

QUANTITATIVE LAND SUITABILITY MAPPING FOR CROP CULTIVATION

Developing the structure of geodatabase and knowledge base to provide quantitative mapping of land suitability for cultivation the main crops has been shown in the research. The general model and catalogs of knowledge base were designed that are structuring and formalizing information for creating thematic and complex maps and geomodels.

The real world spatial problems give rise to multi-criteria decision-making based on geographical information systems (GIS). The GIS environment provides both aggregation and spatial analysis of various georeferenced data. In this paper, the two approaches of multi-criteria evaluation (MCE) are represented - overlaying with the AND operation and the lowest score assignment on one criterion. An example of multi-attribute decision analysis is done using criteria for the crop suitability assessment. A comparison of those two approaches has been made, based on results of the land-use suitability mapping for the study crops - winter wheat, sunflower and corn. The results demonstrated that at certain locations, a range of criteria values, according to the AND operations had the lower applicability and less flexibility than according to the second approach, based on the lowest score assignment on one criterion. Designed maps characterize the threshold status of the soil quality to provide yields of studied crops on certain areas.

A set of land suitability maps, designed as quantitative models of integration of environmental, soil and climatic conditions, would be very effective to manage the complex decisions under the crop cultivation.

Keywords: *knowledge base, geodatabase, land suitability, geoinformation mapping*

1. Introduction

Ukraine, being rich in unique land resources, has favorable climatic and soil conditions for cultivation of sustainable crop yields. However, both anthropogenic activity and intensification of agricultural production have a tendency to grow, what leads to increase of a soil degradation.

In order to enhance the efficiency of utilizing agricultural resources and improve land use planning, different approaches are applied. The climatic and soil conditions of territories are taken into consideration in majority of research.

Land suitability is the fitness of a given type of land for a specified kind of land use [1]. The aim of land evaluation is to select the optimum land-use according to the purpose and scale of land units. Assessment and grouping of certain areas, in terms of their suitability for proper use, is based on mapping a suitability classes or indexes. Physical resources (e.g. climate, vegetation, water, hydrology, landform and soil), human and capital resources are the major kinds of the land-use resources [2].

A set of the land-use suitability maps, designed as quantitative models of integration environmental, soil and climatic conditions, would be very effective to manage complex decisions in land-use planning.

Research in land suitability assessment combines different approaches, techniques and methods, represented in publications [3], [4], [5], [6]. The multi-criteria decision-making (MCDM) has also become one of the most useful methods for land-use and environmental planning [3], [7]. It is primarily concerned with how to combine the information from several criteria to form a single index of evaluation [8].

A number of multi-criteria decision rules implemented in the geographic information system (GIS) environment, including the weighted linear combination (WLC) or weighted summation/ Boolean overlay methods, have been used [9].

There has been an approach for land use planning in sustainable rural systems with a conceptual model, a process model and a future model [10]. The processes of developing soil erosion and soil salinity, as well, the increase in plowing arable lands in the modern Ukrainian agriculture for the latest decades, have been an issue of soil rational use and protection.

In order to ensure the effective quantitative land suitability evaluation and design a set of land suitability maps for crop cultivation, an information system is required, which would ensure solutions of the problem of land rational use and land protection.

2. Research methodology

The approach of geoinformation mapping, providing the development of formal geographic and cartographic knowledge, is based on application of geodatabases. The knowledge is obtained with heterogeneous geodata that is being basic in developing an appropriate information system [11].

The process of designing thematic maps requires both an improvement of geodatabase sectoral models and development of knowledge base models. Research on agricultural land monitoring

* **Svitlana Kokhan, Antonina Moskalenko, Oleg Drozdivskyi**
Department of Geoinformatics and Aerospace Research of the Earth, National University of Life and Environmental Sciences of Ukraine, Ukraine
E-mail: kokhan_s@nubip.edu.ua

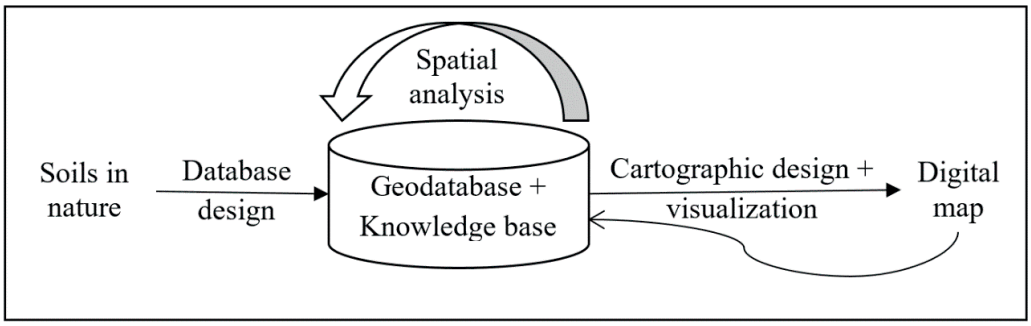


Figure 1 *Scheme of geoinformation mapping system for quantitative land suitability assessment*

and approaches to the development of GIS and geodatabases were considered in papers [12], [13], [14], [15]. The development of knowledge base structure with UML is represented in articles [6], [12], [16], [17], [18], [19].

The aim of the current research is to develop the structure of geodatabase and knowledge base to provide quantitative mapping of land suitability for main crops cultivation in the Northern Forest-Steppe of Ukraine. Basic structural components of geodatabase and the knowledge base should be explained to map quantitative land suitability sub-classes.

The research comprises of the following stages:

- designing conceptual, logical and physical models of geodatabase as basic components of geoinformation mapping system;
- determining the general structure of information blocks; development of a general model and catalogs of knowledge base;
- physical realization of the trial version of geoinformation mapping system using attribute and spatial data for land suitability evaluation in the Northern Forest-Steppe of Ukraine.

Land suitability evaluation for the crop production involves the interpretation of data related to climate, soils, slopes of the territory, type of crops, etc., in order to provide crop requirements to optimal growth conditions.

The Food and Agricultural Organisation [1] proposed land evaluation in terms of two broad classes, "suitable" (S) and "not suitable" (N), based on the suitability of land characteristics to various conditions. Those two are further sub-classified as follows:

Class S1 - Highly suitable: land having no significant limitations for sustained applications to a given use. Minor limitations are applied that would not significantly reduce the productivity.

Class S2 - Moderately suitable: land having limitations that in the aggregate are moderately severe for sustained application to a given use and may reduce the productivity marginally. Those lands have slight limitations and/or no more than three moderate limitations.

Class S3 - Marginally suitable: land with limitations that in the aggregate are severe for sustained application to a given use and as such reduce productivity significantly, but is still marginally economical. Those lands have more than three moderate limitations and/or more than one severe limitation that, however, does not preclude their use for the specified purposes.

An approach, described in [20], based on five sub-classes of land suitability was applied in this research. According to this classification, the arable lands in Ukraine are grouped into sub-classes in terms of the suitability of land characteristics to provide the crop requirements to soil and climatic conditions. Following to this approach, the input data comprises of digital map of soil agricultural production groups and topographic factors. Combining these initial data with the soil quality characteristics (e.g. pH and available nutrients), quantitative land suitability maps for the study area have been developed. To map a suitability sub-classes for study crops, the two approaches were used - overlaying with the AND operation and the lowest score assignment on one criterion. The GIS software was used for data analysis and mapping (ArcGIS 10.3), as well as for supporting spatial decisions.

For the Boolean intersection and overlay, all the criteria are assumed to be constraints and the result is the AND or Boolean minimum operation. This combination technique provides the lowest possible risk, since the only areas considered suitable in the result are those considered suitable in all the criteria.

In weighted linear combination method, criteria may include both weighted factors and constraints. Factor weights are very important in the WLC because they determine how individual factors will tradeoff relative to each other. Each criterion can be assigned a specific weight that reflects it importance relative to other criteria under consideration [5]. Derivation of weights is a central step in defining the decision maker's preferences. This procedure is sometimes too complicated. An approach that is more flexible is used in this study. It is based on the lowest score assignment on one criterion for each study crop.

3. Results and discussion

Geoinformation mapping is the process of informational mapping and modeling of geosystems [14]. The maps designing is connected with the use of the standard GIS and development of specialized geoinformation systems and new mapping methods, based on them.

The scheme of geoinformation mapping system for quantitative land suitability assessment, composed of several steps, is illustrated in Figure 1.

Physical realization of the trial version of geoinformation mapping system for quantitative land suitability assessment, was carried out in the form of Personal geodatabase for the territory of Agronomic research station of National University of Life

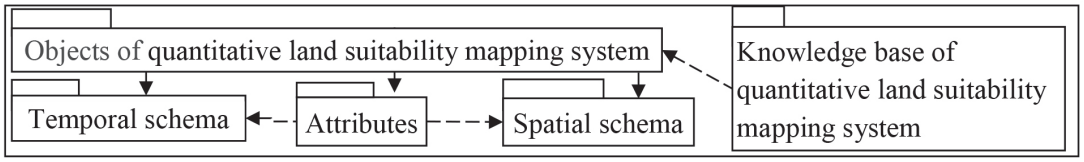


Figure 2 *The UML-model of geodatabase packages*

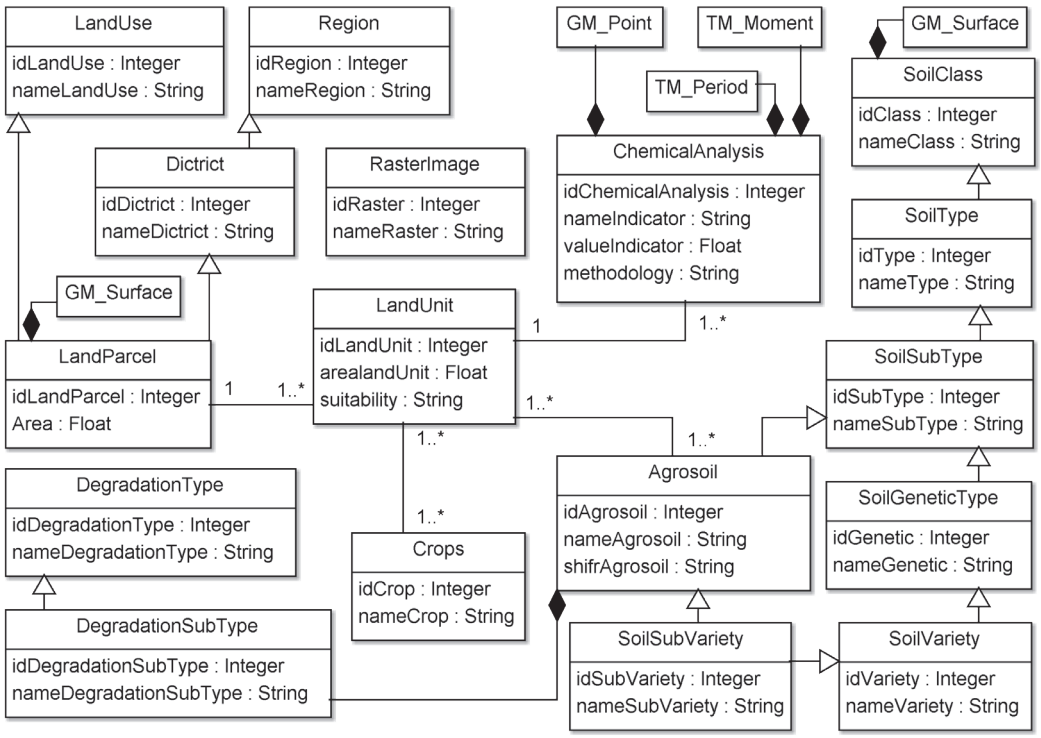


Figure 3 *The UML-model of geodatabase classes*

and Environmental Sciences of Ukraine, which is located in the Northern Forest-Steppe zone. A set of geoimages, obtained in the research, is based on results of the geoinformation analysis and modeling.

3.1 Geodatabase

In general, the geoinformation mapping system for quantitative land suitability assessment could be described as a set of input, intermediate and output data, as well as the processes of data processing, interaction and representation. The input data of the system comprises of non-spatial (attributive) and spatial data. The attributive data included statistical reporting and results of field agrochemical survey. The input data included digital map of soil agricultural production groups (scale 1: 10 000) and digital elevation model (the Shuttle Radar Topography Mission, SRTM).

Large volumes of heterogeneous data require the construction of geodatabase to accumulate, store, analyze spatial information and to create a knowledge base containing descriptions of rules, methods of determination and rules for presentation cartographic materials.

A key feature of geodatabases is the ability to display various data including spatial data. In order to design the conceptual model of geodatabase, the following models have been developed (Figure 2): the scheme of non-spatial characteristics (attributes) of objects (according to the ISO 19110 Geographic information - Methodology for feature cataloguing); the schema of spatial representation of objects, expressed by geometric primitives (according to ISO 19107 Geographic information - Spatial scheme), and the timing representation of spatial information (according to ISO 19108 Geographic information - Temporal scheme).

The land suitability for crop cultivation strongly depends on the soil quality. Such soil characteristics as pH, hydrolizable nitrogen, available phosphorus and exchangeable potassium, were mapped for each soil unit within the study area in order to obtain a set of intermediate cartographic models.

The model of conceptual geodatabase design of geoinformation mapping system is illustrated in Figure 3.

3.2 The knowledge base of geoinformation mapping system

Geoinformation mapping system for quantitative land suitability assessment requires creating a set of complex thematic maps for visualization of the real soil status. In this concern, each mapping object has rules of description, methods of determination, rules for representing the cartographic material

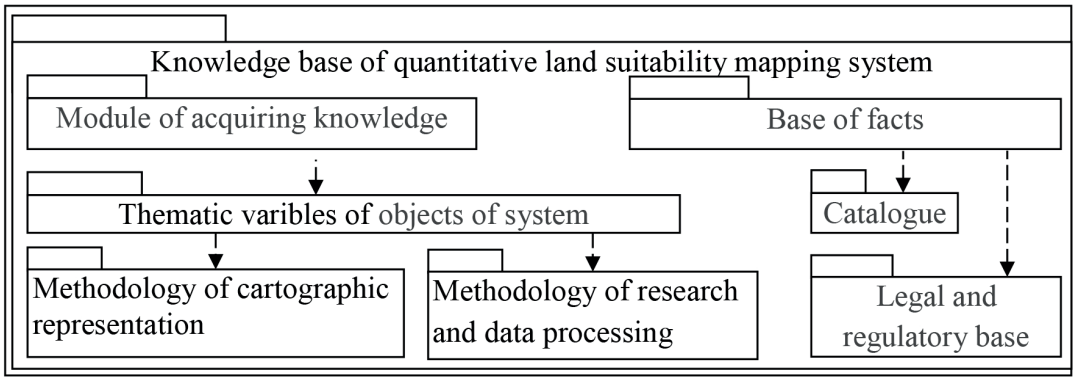


Figure 4 *The UML-model of Knowledge base packages*

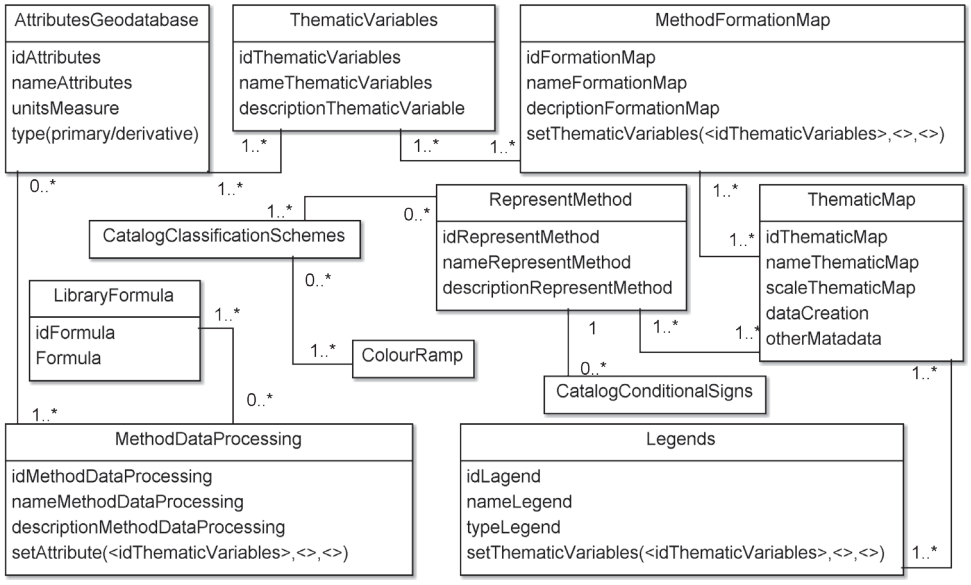


Figure 5 *The UML-model of knowledge base classes of quantitative land suitability mapping system*

and a set of rules for spatial analysis [12], [15], [21]. These elements are structured with knowledge base [17]. The knowledge base provides development of unified and formalized sets of complex digital thematic maps in the GIS environment.

Structural elements of the process of knowledge base development is shown in Figure 4 by the UML-model of Knowledge base packages.

For each attribute, the appropriate classifiers and codifiers were organized. A set of catalogs and rules of knowledge base and their structure has been defined. It unifies the information, methods of data processing and data representation, determines the set and composition of metadata, as well as rules of cartographic representation of objects (Figure 5).

The following catalogs of the main components of this model are represented by catalog of attributes, methods of data processing, mapping variables, legends, methods of graphical representation (classification schemes, colour ramps).

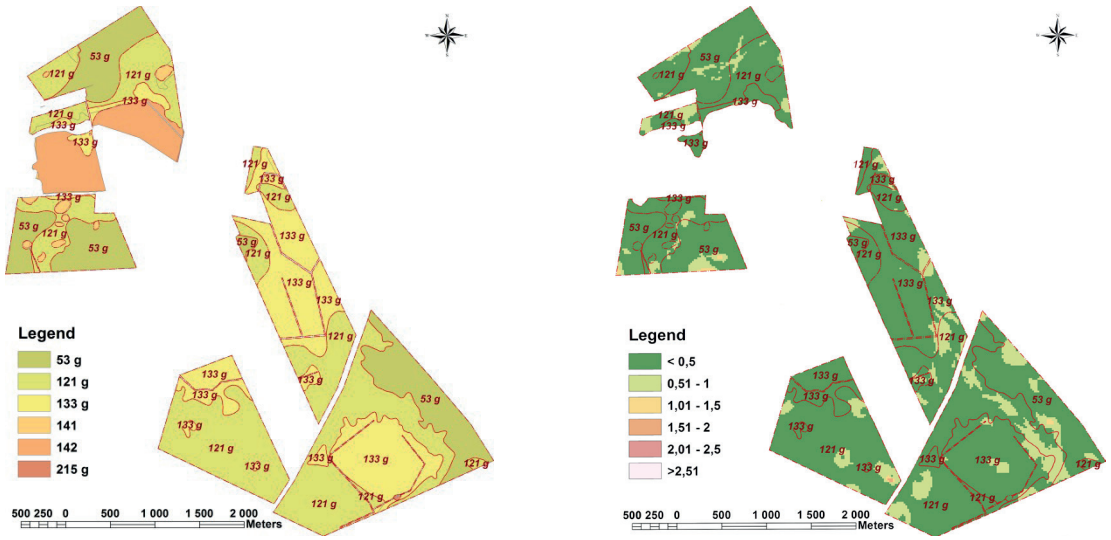
3.3 Land suitability mapping

The input data of geoinformation mapping system for the quantitative land suitability evaluation comprised of statistical reporting and results of field agrochemical survey, map of soil agricultural production groups, digital elevation model (DEM) and slope map, derived from the DEM. Basic maps of the research area are shown in Figure 6. Slopes are represented to be less than 1 degree. Soils belong to the following soil agricultural production groups: 53g - Chernozem typical and Chernozem strongly regraded; 121g - Chernozem-like soil; 133g - Meadow soil; 141/142 Meadow-boggy soils, undrained/drained; 215g - Degraded soils. The group of degraded soils is not involved in crop cultivation being under pastures. Criteria for the crop suitability assessment are represented in Table 1. Values were scaled according to Chirikov method for both P_2O_5 and K_2O, and N according to Kornfield [22].

The classification is known to be one of the most general methods for visualization of analyzed results of spatial and non-spatial data. In order to determine the land suitability for crop cultivation, the classification procedure has been combined with overlay analysis (the AND operation) according to the following sub-classes: S1 - the most suitable lands; S2 - lands with moderate suitability; S3 - limited suitability; S4 - low suitability; N - unsuitable lands. Since the criteria are measured on different scales, the factors were standardized with linear scaling before combination, so that all the factor maps are correlated with suitability.

Table 1 *Criteria for crop suitability assessment (ppm for N, P_2O_5 and K_2O)*

	Criteria	S1 high	S2 moderately	S3 marginally	S4 low	N not suitable
Wheat	N	>175	151-175	76-150	35-75	<35
	P_2O_5	>100	76-100	61-76	23-60	<22
	K_2O	>100	81-100	41-80	31-40	<30
	pH	6.3-7.3	6.0-6.3/ 7.3	5.8-6.0 / 7.4	4.5-5.8 / 7.5	<4.5/ > 7.5
Sunflower	N	>190	166-190	101-165	40-100	<40
	P_2O_5	>100	91-100	51-90	23-50	<22
	K_2O	>200	186-200	101-185	50-100	<50
	pH	6.0-6.8	5.8-6.0/ 6.9-7.0	5.5-5.8/ 7.1-7.3	5.0-5.5/ 7.4-7.5	< 5.0/ >7.5
Corn	N	>190	166-190	101-165	40-100	<40
	P_2O_5	>100	91-100	51-90	23-50	<22
	K_2O	>200	186-200	101-185	50-100	<50
	pH	6.0-6.8	5.8-6.0/ 6.9-7.0	5.5-5.8/ 7.1-7.3	5.0-5.5/ 7.4-7.5	< 5.0/ >7.5



a *b*

Figure 6 *Basic maps: a - soil agricultural production groups, b - slopes (degree)*

According to the land suitability classification [20] the study area belongs to S1 and S3 sub-classes for winter wheat cultivation, however the four sub-classes (S1 to S4) are determined by classification based on the soil quality characteristics being mainly represented by the sub-classes with limited and low suitability lands. The same results are achieved to specify the land suitability areas for corn and sunflower.

The lowest score on one criterion has been assigned (according to Table 1) for each study crops. Two criteria - both pH and K_2O were the most sensitive criteria, which causes suitability class modification from S3 to S4. S1 and S2 were relatively stable classes despite a certain degree of variations in the values of N, P_2O_5, K_2O and pH (Figure 7). Designed maps characterize the threshold status of the soil quality to provide yields of studied crops on certain areas.

4. Conclusions

The process of developing the structure of geoinformation mapping system has been defined by a set of attributes of the soil unit objects. Temporal and spatial characteristics of objects have been established being the basis of the conceptual model of geodatabase. In addition, the possibility of recording data into the geodatabase, related to soil properties and those of being based on agrochemical survey and characterized soil in terms of agrochemical indicators, has been taken into account.

The knowledge base of geoinformation mapping system includes the following components as thematic variables of geoinformation mapping objects: set of research methods, methods of data processing, methods of cartographic representation and spatial analysis. The structure and composition of the knowledge base library and rules of geoinformation mapping system has have been determined in the study.

Application of the geoinformation mapping approaches, based on the constructed knowledge base, provides creating a set of land suitability maps. The developed structure of the knowledge base could be the basis for creating a set of actual land suitability maps.

A number of multi-criteria evaluation methods has been introduced in the GIS environment. Two approaches - overlaying with the AND operation and the lowest score assignment on

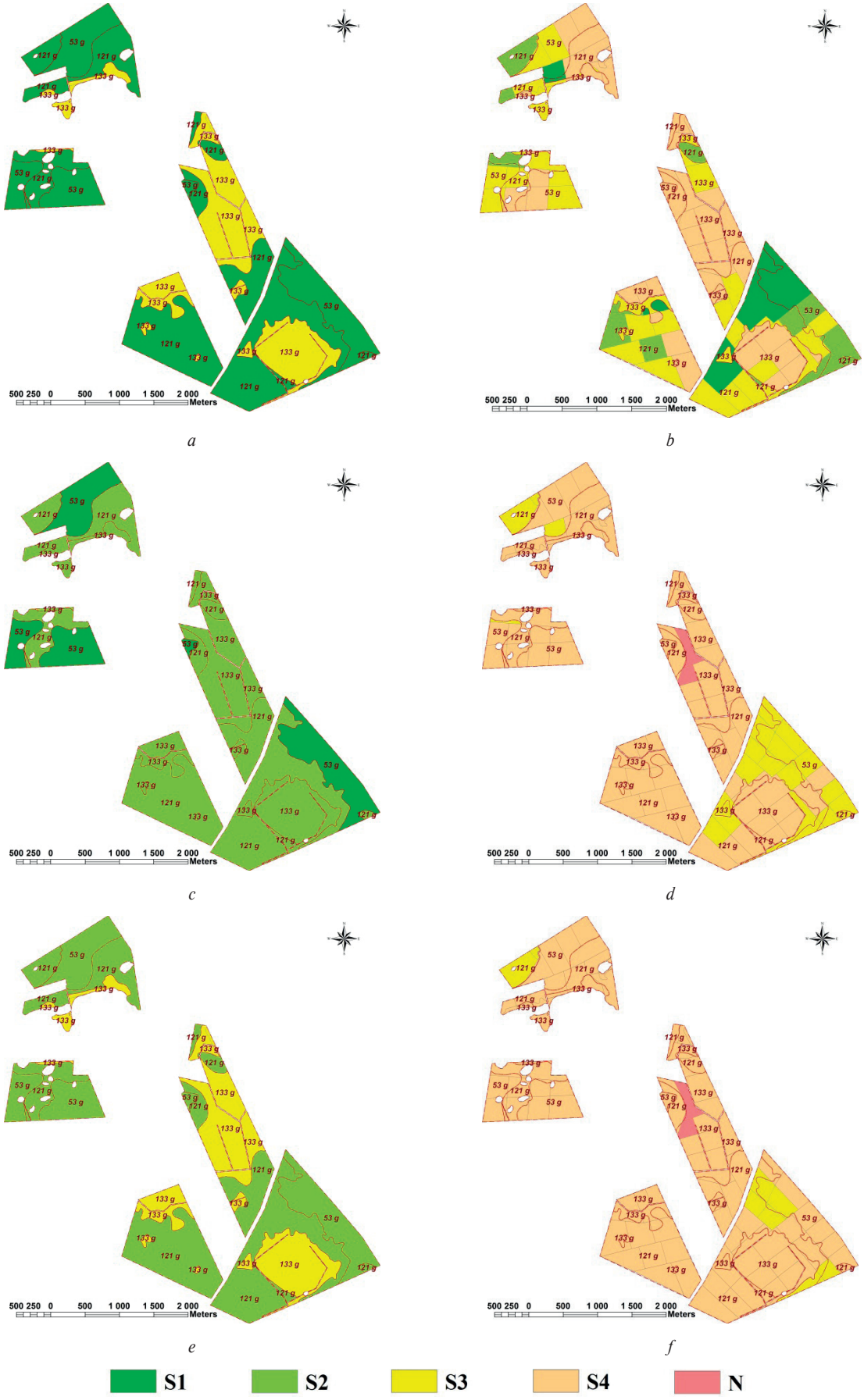


Figure 7 *Land suitability maps: a - winter wheat (according to soil groups); b - winter wheat (according to soil quality); c - corn (according to soil groups); d - corn (according to soil quality); e - sunflower (according to soil groups); f - sunflower (according to soil quality)*

one criterion were used in the study. Results demonstrated that at certain locations, a range of criteria values, according to the AND operations had the lower applicability and less flexibility than according to the second approach, based on the lowest score assignment on one criterion. The results presented in this paper have demonstrated the application of both approaches within a GIS for determining the land suitability areas for the studied crops.

There are still some topics, related to the multi-criteria decision analysis, that should be explained, improved and developed. Those are methods of attribute aggregation, weighing techniques, standardization of factors, error estimation and fuzzy analysis.

References

[1] FAO. A Frame Work for Land Evaluation. Soils Bulletin, 32 [online]. Rome, 1976. Available: http://www.fao.org/docrep/x5310e/x5310e00.HTM_1976.

[2] VAN RANST, E., DEBAVEYE, J.: Land Evaluation Part I, Principles in Land Evaluation and Crop Production Calculation. Agricultural Publication, No. 7. General Administration for development cooperation, Brussels, Belgium, 1991.

[3] AHAMED, T. R. N., RAO, K. G., MURTHY, J. S. R.: GIS-Based Fuzzy Membership Model for Crop-Land Suitability Analysis. Agricultural Systems, 63, 75-95, 2000.

[4] CROSETTO, M., TARANTOLA, S: Uncertainty and Sensitivity Analysis: Tools for GIS-Based Model Implementation. International Journal of Geographical Information Science, 15(5), 415-437, 2010. https://doi.org/10.1080/13658810110053125

[5] MALCZEWSKI, J.: GIS-Based Land-Use Suitability Analysis: A Critical Overview. Progress in Planning, 62(1), 3-65, 2004. https://doi.org/10.1016/j.progress.2003.09.002

[6] SENES, G., TOCCOLINI, A.: Sustainable Land-Use Planning in Protected Rural Areas in Italy. Landscape and Urban Planning, 41, 107-117, 1998. https://doi.org/10.1016/S0169-2046(97)00064-9

[7] WU, F.: SimLand: A Prototype to Simulate Land Conversion through the Integrated GIS and CA with AHP-Derived Transition Rules. Geographical Information Science, 12(1), 63-82, 1998.

[8] CHEN, Y., YUA, J, KHAN, S.: Spatial Sensitivity Analysis of Multi-Criteria Weights in GIS-Based Land Suitability Evaluation. Environmental Modelling and Software, 25(12), 1582-1591, 2010. https://doi.org/10.1016/j.envsoft.2010.06.001

[9] CARVER, S. J.: Integrated Multi-Criteria Evaluation with Geographical Information Systems. International Journal of Geographical Information Systems, 5(3), 321-339, 2007. https://doi.org/10.1080/02693799108927858

[10] VAN LIER, H.: The Role of Land Use Planning in Sustainable Rural Systems. Landscape and Urban Planning, 41(2), 83-91, 1998.

[11] KARPINSKI, Y., LYASCHENKO, A.: Content and Means of Modern Geo-Mapping. Proceedings of International Scientific-Practical Conference "Integration of geospatial data in studies of natural resources", Ukraine, 72-76, 2014.

[12] BULYGIN, S., ACHASOV A.: Using GIS Technology for Soil Mapping, Bulletin of Agricultural Science, 10, 52-56, 2012.

[13] KARPINSKI, Y., LYASCHENKO, A., RUNETS, R.: Reference Model of Topographic Database. Bulletin of Geodesy and Mapping, 2, 28-36, 2010.

[14] LYASCHENKO, A., CHERIN, A.: Architecture Modern GIS Based on Spatial Databases. Bulletin of Geodesy and Mapping, 5, 45–50, 2010.

[15] ZATSERKOVNYY, V., KRIVOBERETS, S.: Analysis of GIS Technology Possibilities to Monitoring of Arable Land. Chernihiv Scientific Journal, Series 2, Engineering and Nature, 2(2), 88-94, 2011.

[16] KOKHAN, S., MOSKALENKO, A.: Development of Knowledge Base Structure of Geoinformation System for Evaluation of Quality Status of Agricultural Lands. Eastern-European Journal of Enterprise Technologies, 5(2), 32-37, 2015.

[17] MOHD, S. A., EVANS, A., BENEST, I., KIMBLE, C.: Modelling Knowledge-Based Systems Using UML Profile [online]. Available: http://www.chriskimble.com/Publications/Documents/Abdullah_2004b.pdf.

[18] SIGOV, A., NECHAEV, V., KOSHKAREV, M.: Architecture Subject-Oriented Knowledge Base of Intellectual. International Journal of Open Information Technologies, 2(12), 1-6, 2014.

[19] ZAGORULKO, Y.: On the Concept of the Integrated Model of Knowledge Representation. Bulletin of the Tomsk Polytechnic University, 5(322), 98-103, 2013.

[20] DOBRYAK, D., KANASH, O., BABMINDRA, D., ROZUMNYI, I.: Classification of Agricultural Lands as Scientific Prerequisite of their Ecological Secure Use. Urozhay, Kyiv, 248-447, 2009.

[21] KOKHAN, S., MOSKALENKO, A., SHILO, L.: Geoinformation Service of Soil Quality Evaluation. Eastern-European Journal of Enterprise Technologies, 6(3), 18-25, 2013.

[22] GORODNIY, M.: Agrochemistry: Manual, the 4th edition. Aristey, Kyiv, 138-145, 2008.

Peter Czimmermann - Michal Kohani*

CHARACTERISTICS OF CHANGES OF TRANSPORTATION PERFORMANCE FOR PAIRS OF CRITICAL EDGES

When the robustness of a public service system design is tested, we can often use scenarios where possible random failures can occur and they can influence the time the service is accessible which is provided for system users. The construction of a suitable scenario is based on the choice of links of the transportation network which influence the system performance in a substantial way. In such scenarios one or multiple arcs can be affected by this failure. In our contribution we present characteristics of pairs of critical arcs that can be used to develop an algorithm for the creation of critical scenarios.

Keywords: *critical scenario, system robustness, transportation performance, public service system*

1. Introduction

The characteristics of public service systems in transportation networks are given by the location of its service centres. Examples of such systems include emergency, health care and supply systems. The service for users is realized on the shortest paths between the users and service centres. A design of such a system can be solved as a weighted *p*-median problem in a network. The algorithms and methods for solving these types of problems are known and can be found in [1], [2], [3].

However, the service systems are usually suggested for ideal conditions, when the traverse time is constant for every edge. In real transportation networks, various random events may occur. These events can elongate the traverse time of the affected edges (arcs) [4]. The collapse of an individual arc was studied in [5], [6], [7]. Two characteristic functions were suggested for measuring the system robustness regarding possible disruptive events on network arcs. The network robustness index was studied in [6]. The network trip robustness was suggested in [7]. Characteristic function of individual network arcs was also studied in [8].

In this paper, we focus on the characterization of pairs of affected edges. We suggest an approach that allows us to compute the function of the elongation of transportation performance when two edges are affected by random events at the same time.

2. Definition of transportation performance

In this section we introduce basic definitions. We work with the network $G = (V,E,w,t)$, where V is the set of vertices, E is the set of edges, $w(u)$ is the weight of the vertex u, and $t(e)$ is the driving time through the edge e. The driving time from vertex u to vertex v is denoted by $d(u,v)$ and it is the length of the shortest path from u to v when we consider the times on the edges from E. Similarly, we can define the value

$$d(u,X) = \min\{d(u,v); v \in X\} \tag{1}$$

for every non-empty subset of vertices $X \subset V$.

Let the set of customers be denoted by U and the set of facilities be denoted by S, where $U \subset V$, $S \subset V$. The transportation performance [8] is

$$W = \Sigma_{u \in U} w(u) \cdot d(u,S) \tag{2}$$

This expression is similar to the total weighted distance, however we use travel times on edges, instead of distances.

3. Changes of the transportation performance

It is not hard to imagine situations when the travel time on the edge is extended. In this section we describe the impact of such extensions on transportation performance. We focus on the case with two critical edges.

Let the set

$$Y = \{e_1, e_2, \ldots, e_k\} \subseteq E \tag{3}$$

be given. We suppose that $\delta_i \geq 0 (i = 1, 2, \ldots, k)$ is an extension of the driving on the edge e_i. We denote the transportation performance with extended times

$$t(e_i) + \delta_1, t(e_2) + \delta_2, \ldots, t(e_k) + \delta_k \tag{4}$$

on edges of Y by

$$W_Y(\delta_1, \delta_2, \ldots, \delta_k) \tag{5}$$

We can see that W_Y is a function of k variables $\delta_1, \delta_2, \ldots, \delta_k$ with domain R_+^k.

The case $k = 1$ is studied in [8]. We are concerned with the case $k = 2$ and consider some possibilities for $k > 2$.

* **Peter Czimmermann, Michal Kohani**
Department of Mathematical Methods and Operations Research, Faculty of Management Science and Informatics, University of Zilina, Slovakia
E-mail: peter.czimmermann@fri.uniza.sk

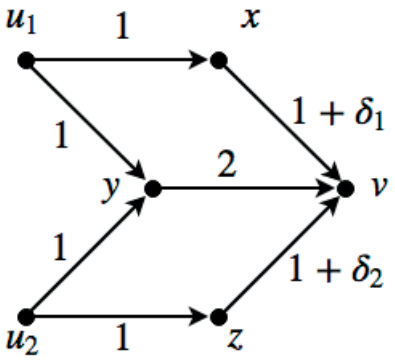


Figure 1 *Network from Example 1*

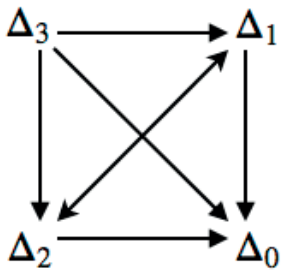


Figure 2 *Comparison of sets* Δ_3, Δ_1, Δ_2, Δ_0

Example 1. We consider the network G in Figure 1. Let $U = \{u_1,u_2\}$ and $S = \{v\}$. Transportation performance $W_Y(\delta_1,\delta_2)$ is given by formulas:

1. $4+\delta_1+\delta_2$, if $0 \le \delta_1,\delta_2 \le 1$
2. $5+\delta_1$, if $0 \le \delta_1$ and $\delta_2 > 1$ (6)
3. $5+\delta_2$, if $0 \le \delta_2$ and $\delta_1 > 1$
4. 6 otherwise

We can see that the following facts hold for every network and pair of its edges.

Theorem 1. For the given vertex u, the transportation performance is given by, at most, four formulas.

Proof. There are four cases to consider:

Case c_3 - the edges e_1, e_2 belong to the shortest path from u to S for given values δ_1, δ_2.

Case c_1 - the edge e_1 belongs (e_2 does not) to the shortest path from u to S for given values δ_1, δ_2.

Case c_2 - the edge e_2 belongs (e_1 does not) to the shortest path from u to S for given values δ_1, δ_2.

Case c_0 - the edges e_1, e_2 do not belong to the shortest path from u to S for given values δ_1, δ_2.

The number of formulas is equal to the number of cases c_3, c_1, c_2, c_0 that can occur. We denote the sets of points (δ_1, δ_2), for which cases c_3, c_1, c_2, c_0 hold, by Δ_3, Δ_1, Δ_2, Δ_0. We can compare these sets by the following relation:

$\Delta_i \to \Delta_j \Leftrightarrow$ there exists point $(a, b) \in \Delta_i$ such that for all $(c, d) \in \Delta_j$ we have $a \le c$, $b \le d$ and at least one inequality is sharp.

It is possible to show by case by case analysis that the relationships between the sets Δ_3, Δ_1, Δ_2, Δ_0 are in Figure 2. It is easy to consider that every case is represented by one formula (we have one formula for every set), as we can see below:

- case c_3: $w(u) \cdot (d_3(u,S) + \delta_1 + \delta_2)$, where $d_3(u,S)$ denotes the length of the shortest path from u to S, when $(\delta_1, \delta_2) \in \Delta_3$, (this path is the shortest among all paths from u to S, which contain the edges e_1 and e_2)
- case c_1: $w(u) \cdot (d_1(u,S) + \delta_1)$, where $d_1(u,S)$ denotes the length of the shortest path from u to S, when $(\delta_1, \delta_2) \in \Delta_1$,
- case c_2: $w(u) \cdot (d_2(u,S) + \delta_2)$, where $d_2(u,S)$ denotes the length of the shortest path from u to S, when $(\delta_1, \delta_2) \in \Delta_2$,
- case c_0: $w(u) \cdot d_0(u,S)$, where $d_0(u,S)$ denotes the length of the shortest path from u to S, when $(\delta_1, \delta_2) \in \Delta_0$.

Theorem 2. The function $W_Y(\delta_1,\delta_2)$ is continuous and piecewise linear function in two variables δ_1, δ_2.

Proof. Let the vertex $u \in U$ be given. Let the edges e_1, e_2 belong to the shortest path from u to S. If the travel times on edges e_1, e_2 are increased by values δ_1, δ_2, then the travel time $d(u,S)$ from u to S is changed to $d(u,S) + \delta_1 + \delta_2$. If the value of δ_2 achieves the level which involves omitting the edge e_2, then we obtain a new formula for the travel time from u to S: $d_1(u,S) + \delta_1$. Moreover, the transition between these formulas is given by equality:

$$d(u,S)+\delta_1+\delta_2 = d_1(u,S)+\delta_1 \tag{7}$$

It follows from the equality that this transition is continuous. The same can be checked for all cases. The general formula for the transportation performance is

$$W_Y(\delta_1,\delta_2) = \sum_{u\in U} w(u)\cdot(d_i(u,S)+a_i\cdot\delta_1+b_i\cdot\delta_2) \tag{8}$$

where $i \in \{0,1,2,3\}$; its value depends on $(\delta_1, \delta_2) \in \Delta_i$ and $a_0 = b_0 = 0$, $a_1 = 1$, $b_1 = 0$, $a_2 = 0$, $b_2 = 1$, $a_3 = b_3 = 1$. It is known that the sum of continuous functions is a continuous function and from the general formula it follows that $W_Y(\delta_1,\delta_2)$ is piecewise linear function.

3.1 Computation of transportation performance for a given vertex *u*

It was mentioned in the proof of Theorem 2 that the general formula for $W_Y(\delta_1,\delta_2)$ is the sum of formulas derived for each vertex from the set U. Hence, we show the approach for the computation of formulas for every vertex from this set. We also show the computation of boundaries of sets with the same formula. Composition of these sets provides the partition of the domain of $W_Y(\delta_1,\delta_2)$. We show the approach which gives the formula for transportation performance for a given vertex $u \in U$.

1. We find the shortest path from u to S in G (we use values $\delta_1 = \delta_2 = 0$). The length of this path is $d_i(u,S)$. Now, when we carefully increase the values of δ_1, δ_2 (the word carefully means that the shortest path is not changed), we obtain these possible starting formulas:

$$\begin{aligned} c_3 &\to w(u)\cdot(d_3(u,S)+\delta_1+\delta_2)\\ c_1 &\to w(u)\cdot(d_1(u,S)+\delta_1)\\ c_2 &\to w(u)\cdot(d_2(u,S)+\delta_2)\\ c_0 &\to w(u)\cdot d_0(u,S) \end{aligned} \tag{9}$$

The starting formula is given by the occurrence of edges e_1 and e_2 in shortest path for $\delta_1 = \delta_2 = 0$.

2. We find out which cases (from c_1, c_2, c_0) can also occur, by omitting the edges e_1, e_2. Formulas for these cases are similar to the formulas described above.
3. If the case c_3 occurs for $\delta_1 = \delta_2 = 0$, then the bounds of Δ_3 are

$$\begin{aligned} w(u)\cdot(d_3(u,S)+\delta_1+\delta_2) &= w(u)\cdot(d_1(u,S)+\delta_1)\\ w(u)\cdot(d_3(u,S)+\delta_1+\delta_2) &= w(u)\cdot(d_2(u,S)+\delta_2)\\ w(u)\cdot(d_3(u,S)+\delta_1+\delta_2) &= w(u)\cdot d_0(u,S) \end{aligned} \tag{10}$$

Hence, the set Δ_3 is given by bounds

$$\begin{aligned} \delta_1,\delta_2 &\geq 0\\ d_3(u,S)+\delta_2 &\leq d_1(u,S)\\ d_3(u,S)+\delta_1 &\leq d_2(u,S)\\ d_3(u,S)+\delta_1+\delta_2 &\leq d_0(u,S) \end{aligned} \tag{11}$$

The set Δ_1 (when the case c_1 occurs) is given by inequalities

$$\begin{aligned} \delta_3(u,S)+\delta_2 &\geq d_1(u,S)\\ d_1(u,S)+\delta_1 &\leq d_2(u,S)+\delta_2\\ d_1(u,S)+\delta_1 &\leq d_0(u,S) \end{aligned} \tag{12}$$

Similarly, Δ_2 (when the case c_2 occurs) is given by

$$\begin{aligned} \delta_1,\delta_2 &\geq 0\\ d_3(u,S)+\delta_1 &\geq d_2(u,S)\\ d_1(u,S)+\delta_1 &\geq d_2(u,S)+\delta_2\\ d_2(u,S)+\delta_2 &\geq d_0(u,S) \end{aligned} \tag{13}$$

The set Δ_0 (when the case c_0 occurs) is given by

$$\begin{aligned} \delta_1,\delta_2 &\geq 0\\ d_3(u,S)+\delta_1+\delta_2 &\geq d_0(u,S)\\ d_1(u,S)+\delta_1 &\geq d_0(u,S)\\ d_2(u,S)+\delta_2 &\geq d_0(u,S) \end{aligned} \tag{14}$$

4. If the case c_1 is the starting case, then the bounds of Δ_1 are

$$\begin{aligned} \delta_1,\delta_2 &\geq 0\\ d_1(u,S)+\delta_1 &\leq d_2(u,S)+\delta_2\\ d_1(u,S)+\delta_1 &\leq d_0(u,S) \end{aligned} \tag{15}$$

Similarly, Δ_2 (when the case c_2 occurs) is given by

$$\begin{aligned} \delta_1,\delta_2 &\geq 0\\ d_1(u,S)+\delta_1 &\leq d_2(u,S)+\delta_2\\ d_2(u,S)+\delta_2 &\leq d_0(u,S) \end{aligned} \tag{16}$$

The bounds of Δ_0 (when the case c_0 occurs) are

$$\begin{aligned} \delta_1,\delta_2 &\geq 0\\ d_1(u,S)+\delta_1 &\leq d_0(u,S)\\ d_2(u,S)+\delta_2 &\leq d_0(u,S) \end{aligned} \tag{17}$$

5. The starting case c_2 is very similar to the previous starting case.
6. If c_0 is the starting case, then we have only one set Δ_0 given by bounds $\delta_1, \delta_2 \geq 0$.

Bounds for formulas of transportation performance can be obtained from bounds computed for every vertex $u \in U$.

4. Algorithms for computation of pairs of critical edges

In this section we introduce a fast algorithm, which finds the set of critical edges in the case where the extensions of travel times do not force us to avoid the critical edges.

We start with finding the pair of critical edges, when the values δ_1 and δ_2 are small - we do not avoid any edge in the shortest path.

1. Find the shortest paths from u_i to S for all $u_i \in U = \{u_1, u_2, \ldots, u_m\}$. The set of all such paths is denoted by $P = \{p_1, p_2, \ldots, p_m\}$. (The starting vertices of these paths are from U.)
2. For every edge e we have

$$s(e) = \sum_{i=1}^{m} w(u_i)s_i(e), \text{ where } s_i(e) = \begin{cases} 1, if\ e \in p_i \\ 0, if\ e \in p_i \end{cases} \tag{18}$$

The sum is taken over all vertices $u \in U$ such that the edge e is the edge of the shortest path from the set P with the starting point u.

3. We order the edges of G by $s(e)$ decreasingly. The first two edges form a critical pair of edges. This is a polynomial algorithm which can be generalised for every $k > 2$. Let us suppose that the extensions of times are too large, we need to find the new shortest paths without the critical edges. It is, in general, an NP-hard problem. However, for $k = 2$, we can solve the problem by brute force.

5. Conclusions and further research

We studied a possible characterization of critical pairs of edges in transportation networks. We suggested the approach which allows us to determine the characteristic function of the time elongation of transportation performance for pairs of affected edges.

Future research in this field will focus on the tests on real transportation networks and study the elongation of transportation performance as a function of $k > 2$ variables – it means that we will also consider scenarios with $k > 2$ critical edges.

Acknowledgment

This work was supported by the research grants VEGA 1/0342/18 "Optimal dimensioning of service systems", VEGA 1/0463/16 "Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities" and APVV-15-0179 "Reliability of emergency systems on infrastructure with uncertain functionality of critical elements".

References

[1] GARCIA, S., LABBE, M., MARIN, A.: Solving Large p-median Problems with a Radius Formulation. INFORMS Journal on Computing, 23(4), 546-556, 2011. https://doi.org/10.1287/ijoc.1100.0418

[2] JANACEK, J., KVET, M.: Relevant Network Distances for Approximate Approach to Large p-median Problems. Operations Research Proceedings 2012, selected papers of the International Annual Conference of the German Operations Research Society (GOR), Germany, 123-128, 2012. https://doi.org/10.1007/978-3-319-00795-3_18

[3] KVET, M.: Computational Study of Radial Approach to Public Service System Design with Generalized Utility. Digital Technologies, Proceedings of the 10th International IEEE Conference, Slovakia, 187-197, 2014. https://doi.org/10.1109/DT.2014.6868713

[4] KVET, M., MATIASKO, K.: Concept of Dynamic Index Management in Temporal Approach Using Intelligent Transport Systems. ROCHA, A., CORREIA, A. M., ADELI, H., REIS, L. P., COSTANZO, S. (Eds.): Recent Advances in Information Systems and Technologies, Vol. 1, Springer, p. 549-560, 2017.

[5] JENELIUS, E.: Network Structure and Travel Patterns: Explaining the Geographical Disparities of Road Network Vulnerability. Journal of Transport Geography, 17(3), 234-244, 2009. https://doi.org/10.1016/j.jtrangeo.2008.06.002

[6] SCOTT, D. M., NOVAK, D. C., AULTMAN-HALL, L., GUO, F.: Network Robustness Index: A New Method for Identifying Critical Links and Evaluating the Performance of Transportation Networks. Journal of Transport Geography, 14(3), 215-227, 2006. https://doi.org/10.1016/j.jtrangeo.2005.10.003

[7] SULLIVAN, J. L., NOVAK, D. C., AULTMAN-HALL, L., SCOTT, D. M.: Identifying Critical Road Segments and Measuring System-Wide Robustness in Transportation Networks with Isolating Links: A Link Based Capacity Reduction Problem. Transportation Research Part A., 44, 323-336, 2010. https://doi.org/10.1016/j.tra.2010.02.003

[8] JANACEK, J., KVET, M.: Characteristic of a Critical Network Arc in a Service System. Transport Problems, 12, 141-146, 2017. https://doi.org/10.20858/tp.2017.12.se.12

Petr Kozel - Lucie Orlikova - Sarka Michalcova*

THE MODIFIED RURAL POSTMAN PROBLEM IN VEHICLE ROUTE OPTIMIZATION

The submitted paper deals with designing routes of the vehicles, which provide the transport network services. We limit our focus to such tasks, where the priority is the edge service in the transport network and the initial problem is finding an Eulerian path. Regarding to real-life problems, this contribution presents such procedure of solving, which takes into account both the existence of a mixed transport network containing one-way roads and the existence of a wider transport network. In this network, there are only selected edges with possibility of the effective passages. This problem can be solved by the modified Rural Postman Problem assuming the strongly connected network. Linear programming is a suitable tool for designing optimal routes of service vehicles.

Keywords: *Linear programming, Eulerian path, vehicle routing tasks, the Rural Postman Problem, municipal waste collection*

1. Introduction

Vehicle route planning is an inherent part of the decision making process of the all subjects providing a transport network service (such as cleaning and maintenance of roads, municipal waste collection, separated municipal waste collection etc.). Generally, the underlying problem is to design service vehicles routes while taking into account the chosen optimization criterion and respecting all of the substantial constraints resulting from the real traffic, too. Particularly, the most commonly used optimization criterion for designing service vehicle routes is minimizing the total distance travelled. In general, tasks relating to the route planning and optimization can be divided into two basic groups. The first one includes such problems that solve a point in a transport network - *a vertex* (customer, depot etc.) [1], [2] and the aim of these tasks is to find the minimum Hamiltonian cycle. The next group represents such tasks, in which the subject of service is a road - *an edge* (snow cleaning, street cleaning etc.) and the basic objective is to determine an Eulerian path. In this paper, we limit our focus to problems of the second set of tasks.

2. Basic problem formulation

Let the real transport network be represented by graph N (V, E, l) with three attributes: V = 1, ..., m as set of vertices, E = 1,..., m denotes a set of edges and l_{ij} (where $i, j \in V$) represents an edge evaluation and denote the road length (in kilometres). Generally, the aim is to find a route that passes through each edge at least once with the minimal length and also begins and ends at the depot. This problem was first formulated by Mei-Ko-Kuan and in the professional literature is known as the Chinese Postman Problem [3]. Assuming that all conditions for existence of Eulerian path are met the Edmonds' algorithm [4] can be used to find a solution of this problem. However, if some of these conditions fails to be met, the application of an alternative approach becomes necessary to eliminate this shortage. Principally, it is vital to identify those edges, which are used repeatedly on the route. In practice, these are unproductive passages through some sections. For this, the so called minimum matching principle [5] is commonly applied to identify the repeatedly included edges, in detail in [6]. Nevertheless, the disadvantage of this approach is the necessity to design some substitute transport network, in which the duplicate edges can be identified by using the minimum matching principle. In this respect, we consider more appropriate procedure for identifying duplicate edges by determining the number of passages through each edge of the transport network [5]. Above all, it is not necessary to create a substitute network in the framework of this approach. Information in the matrix of distances l_{ij} corresponding with the default network N is included in the mathematical model. In addition, this matrix contains only values corresponding to the vertices that are connected by the edge. This fact is expressed as "exists" l_{ij}.

2.1 Basic problem formulation - mathematical model

Mathematical model of the above mentioned issues can be formulated as follows:

$$\text{Min} \sum_{i=1}^{m} \sum_{\substack{j=1 \\ \text{exists } l_{ij}}}^{m} l_{ij} \cdot z_{ij} \tag{1}$$

$$z_{ij} + z_{ji} \geq 1 \qquad \text{for } i,j = 1,\ldots,m, \text{, where exists } l_{ij} \tag{2}$$

$$\sum_{\substack{j=1 \\ \text{exists } l_{ij}}}^{m} z_{ij} = \sum_{\substack{j=1 \\ \text{exists } l_{ij}}}^{m} z_{ij} \quad \text{for } i = 1,\ldots,m \tag{3}$$

$$z_{ij} \in Z^{+} \qquad \text{for } i = 1,\ldots,m \text{, where exists } l_{ij} \tag{4}$$

* **Petr Kozel, Lucie Orlikova, Sarka Michalcova**
Department of Mathematical Methods in Economics, Faculty of Economics, Technical University of Ostrava, Czech Republic
E-mail: petr.kozel@vsb.cz

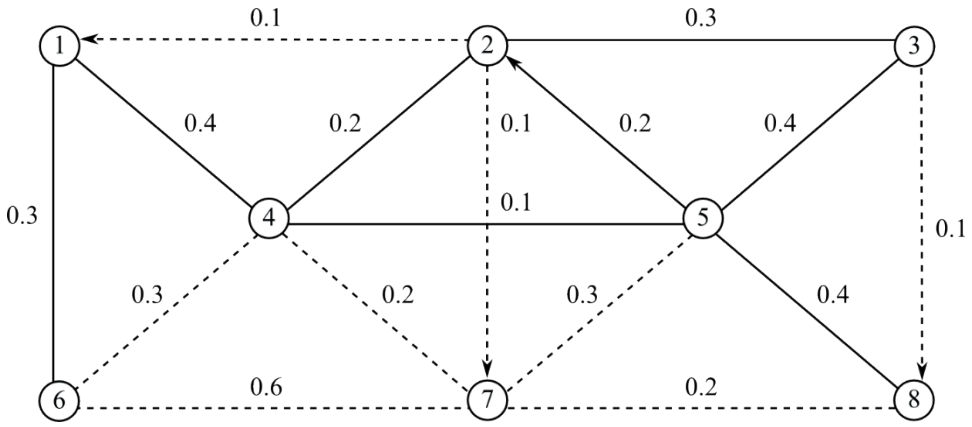


Figure 1 *The network N with marked subnets N_1 and N_2*

The objective function (1) gives e.g. the total distance travelled during the waste municipal collection. The variable z_{ij} is an integer representing the number of passages through the edge *ij*. The conditions (2) ensure that every existing edge (exists l_{ij}) is included at least once. The condition (3) ensures that each visited vertex will be subsequently abandoned. The obligatory condition (4) indicates that the variable z_{ij} is a nonnegative integer. Numerical experiments with this model were published in [6].

3. Mathematical model for real networks - modified approach

The mathematical model (1) - (4) can be used only for undirected transport network without further specific requirements. However, with regard to the practice, it is necessary to take into account several other requirements:

- existence of a mixed transport network, which contains directed and undirected edges (e.g. one-way roads in practice);
- existence of a wider transport network, in which the service of some edges is required (e.g. this corresponds to the situation when the certain set of roads is served and other roads can be used for efficient passages).

Taking into account the above-formulated requirements of the real traffic, both the approach and the mathematical model are presented in the following text.

3.1 Formulation of modified task formulation

The specified network N (V, E, l) defines a *mixed graph*. This network has three attributes and can be divided to several subnets. In our case there are two subnets. The first one N_1 (V_1, E_1, d) represents a set of edges, which have to be served at least once. The second subnet N_2 (V_2, E_2, e) determines a set of edges, which can be used for efficient passages. The constant l_{ij} denotes edge evaluation *ij*, $d_{ij} \subset l_{ij}$ and $e_{ij} \subset l_{ij}$. Hence, the mutual relation can be expressed mathematically as follows: $N = N_1 \cup N_2$. The aim of this problem is to find such route, which passes through all the edges of the subgraph N_1 at least once with possibility of using the edges from the subgraph N_2 so that the total length is minimal. The request route starts and ends at the initial vertex.

The following requirement must be met at the same time: firstly, the directed edges, located both in subnet N_1 and N_2, can be passed only in the edge direction. This requirement corresponds to the one-way roads in practice.

The model case is presented in Figure 1, where the network N is created by 8 vertices and 16 edges. Subnet N_1 (indicated by solid lines) represents the edges, which have to be served at least once. The subnet N_2 (indicated by dashed lines) shows the edges, which may be used for efficient passages.

3.2 Sufficient preconditions to admissibility of the solution

Unlike the classical version of the Rural Postman Problem [7], the strongly connected subnet N_1 is a basic assumption of the approach presented in this paper. If this subnet is not strongly connected, the resulting solution obtained by the proposed method may include subnets in the form of isolated closed trails. However, assuming the strongly connected subnet N_1, this situation cannot occur. This claim is described in detail and confirmed by the proof in the following text. There are introduced two auxiliary concepts to demonstrate this proof:

- an obligatory edge - an edge belonging to the subset E_1,
- an obligatory vertex - a vertex that incidents with an obligatory edge at least once.

Assumptions:

1. The edge evaluation is the set of positive real numbers.
2. All just required obligatory vertices are located in the subgraph N_1.
3. The subgraph N_1 is strongly connected, thus for each two vertices *i*, *j* it is possible to find an oriented path from *i* to *j* and *j* to *i*, too.

The statement: The resulting route obtained by the proposed method is a single closed trail including all required obligatory edges, therefore it does not include two or more isolated closed trails that do not have a common vertex.

Proof by contradiction:

1. Let the resulting route contains an isolated closed trail that does not have a common obligatory vertex with a rest of the route.
2. Both isolated closed trail and the rest of the route must contain a obligatory edge with at least one positive evaluation.
3. However, all obligatory vertices belong to the subgraph N_1 that is strongly connected and therefore there is a route, which

Table 1 *The model example*

	Input data															
ij	14	16	23	24	25	32	35	41	42	45	52	53	54	58	61	85
d_{ij}	0.4	0.3	0.3	0.2	T	0.3	0.4	0.4	0.4	0.1	0.2	0.4	0.1	0.4	0.3	0.4
ij	12	21	27	38	46	47	57	64	67	72	74	75	76	78	83	
e_{ij}	T	0.1	0.1	0.1	0.3	0.2	0.3	0.3	0.6	T	0.2	0.3	0.6	0.2	T	

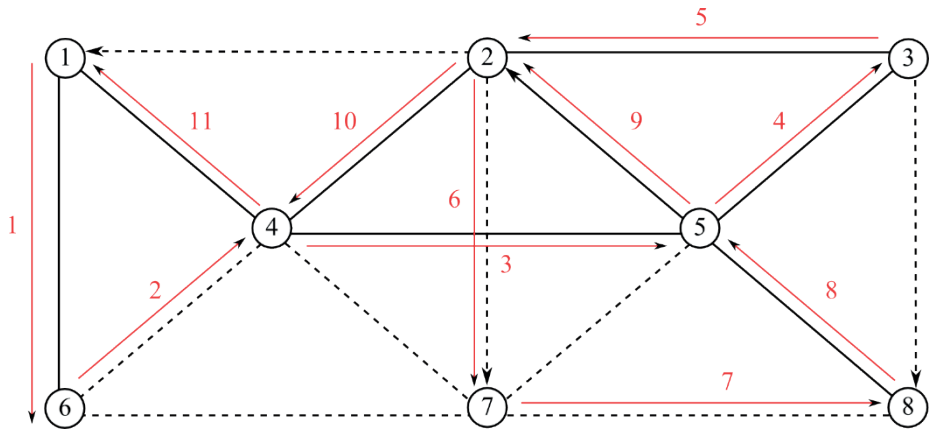


Figure 2 *The use of edges of the subnets N_1 and N_2 in the optimal solution*

is composed only of the obligatory edges and connects any selected obligatory vertex from the isolated closed trail with one of the final (obligatory) vertices of the obligatory edge in the rest of the route. On this route, there must necessarily be an obligatory vertex, which is the final vertex of the obligatory edge of the isolated closed trail and at the same time, it is the final vertex of the obligatory edge that belongs to the rest of the route. It is a dispute with that assumption.

4. Mathematical model of the modified task

To find the optimal route, it is possible to use a mathematical model that determines, how many times the edge *ij* is used in the route. This model (5) - (9) is based on the previously mentioned model (1) - (4). Data corresponding to the subnets N_1 and N_2 enter into the mathematical model. The evaluation of existing edges in subnet N_1 is represented by a constant d_{ij} and the evaluation of existing edges in subnet N_2 is represented by a constant e_{ij}. This method of constant determination ensures that the existing edges are used and what is more, it corresponds to the input matrices, which are sparse. The dynamic declaration is used in the computational environment.

The possibility of passages through the edge against its direction is penalized by a prohibitive constant *T*, the corresponding relationship is defined in [6]. An example of the sparse matrix with the prohibitive constant *T* corresponding to Figure 1 is given in Table 1.

The number of the edge passages *ij* - mathematical model [8]:

$$\text{Min} \sum_{i=1}^{m} \sum_{\substack{j=1 \\ \text{exists } d_{ij}}}^{m} d_{ij} \cdot z_{ij} + \sum_{i=1}^{m} \sum_{\substack{j=1 \\ \text{exists } e_{ij}}}^{m} e_{ij} \cdot y_{ij} \tag{5}$$

$$z_{ij} + z_{ji} \geq 1 \qquad \text{for } i,j = 1,\ldots,m, \text{, where exists } d_{ij} \tag{6}$$

$$\sum_{\substack{j=1 \\ \text{exists } d_{ij}}}^{m} z_{ij} + \sum_{\substack{j=1 \\ \text{exists } e_{ij}}}^{m} y_{ij} = \sum_{\substack{j=1 \\ \text{exists } d_{ij}}}^{m} z_{ij} + \sum_{\substack{j=1 \\ \text{exists } e_{ij}}}^{m} y_{ij} \text{ for } i = 1,\ldots,m \tag{7}$$

$$z_{ij} \in Z_0^+ \text{ for } i = 1,\ldots,m \text{, where exists } d_{ij} \tag{8}$$

$$y_{ij} \in Z_0^+ \text{ for } i = 1,\ldots,m \text{, where exists } e_{ij} \tag{9}$$

The objective function (5) gives e.g. the total distance travelled during the service route. The first part of the expression is the distance traveled during the edges service of the subnet N_1, and the second part represents the distance traveled by using edges from the subnet N_2. The variable z_{ij} is an integer variable that expresses the number of edge runs in the subnet N_1. The variable y_{ij} is an integer variable which determines the number of passages through edges of the subnet N_2. Condition (6) ensure that each existing edge in the subnet N_1 is included in the route at least once. The condition (7) means that each visited vertex will be subsequently abandoned. The obligatory conditions (8) and condition (9) ensure that variables z_{ij} and y_{ij} may be only the nonnegative integer.

5. Computational experiments

This chapter focuses on experiments, which were realized with mathematical model (5)-(9) in computational environment *Xpress-IVE [9], [10]*. Prior to presenting final results, the solution procedure is given by an illustration example. Let us suppose the network as depicted in Figure 1. Firstly, the solution may be obtained assuming that it is not possible to use the other edges than the edges of the subnet N_1. The optimal route (1-4-5-8-5-2-3-5-2-4-1-6-1) length corresponding to this solution is 3.6 km.

Secondly, the optimal solution can be also achieved by using edges of the subnet N_2. This optimal route (1-6-4-5-3-2-7-8-5-2-4-1) with length 2.9 km is shown in Figure 2. In accordance with this

Table 2 *Results of numerical experiments 1 - 6*

Exp. 1: (133 vert. / 168 edg. / 8 one-way)				Exp. 4: (114 vert. / 147 edg. / 15 one-way)			
V1	104 / 3 108.3 [km]	V2	64 / 5 82.1 [km], 24.2 %	V1	99 / 8 70.3 [km]	V2	48 / 7 67.1 [km], 4.6 %
Exp. 2: (128 vert. / 166 edg. / 18 one-way)				Exp. 5: (102 vert. / 126 edg. / 7 one-way)			
V1	130 / 10 112.3 [km]	V2	36 / 8 95.8 [km], 14.7 %	V1	84 / 4 62.1 [km]	V2	42 / 3 57.6 [km], 7.2 %
Exp. 3: (72 vert. / 98 edg. / 12 one-way)				Exp. 6: (104 vert. / 131 edg. / 6 one-way)			
V1	70 / 5 58.6 [km]	V2	28 / 7 50 [km], 14.7 %	V1	86 / 3 69.6 [km]	V2	45 / 3 66.9 [km], 3.9 %

Table 3 *Results of numerical experiments 7 - 10*

Exp. 7: (185 vert. / 239 edg. / 9 one-way)				Exp. 9: (157 vert. / 213 edg. / 10 one-way)			
V1	139 / 7 112.3 [km]	V2	100 / 2 95.9 [km], 14.6 %	V1	126 / 7 90.5 [km]	V2	87 / 3 77.8 [km], 14 %
Exp. 8: (147 vert. / 186 edg. / 8 one-way)				Exp. 10: (139 vert. / 186 edg. / 7 one-way)			
V1	123 / 5 96.9 [km]	V2	63 / 3 82.9 [km], 14.4 %	V1	122 / 5 90.4 [km]	V2	64 / 2 74.4 [km], 17.7 %

Table 4 *Results of numerical experiment - 11 (real problem)*

Exp. 11: (3706 vert. / 6446 edg. / 1012 one-way)			
V1	2063 / 249 280 [km]	V2	4383 / 763 268 [km], 4.3 %

solution, 3 edges of the subnet N_2 are used for the effective vehicle transits. When comparing to the original solution, the achieved saving is 0.7 km, which is 12.1 %.

In the following we will present such results, which were realized using the mathematical model (5)-(9) on a set of 11 nontrivial tasks [11], [12]. For the sake of usage, the size and conditions of these tasks correspond to real problems. For the first ten tasks, all input data were generated by *Wolfram Mathematica [13], [14]*. Input data for last task were obtained from the real traffic of municipal waste collection, which are provided by Technical services in Olomouc. All calculations have been made for two variants (V1 and V2). To design the optimal route, the first variant enables using only edges of the subnet N_1 and the second one allows to use edges from the subnet N_2. Results with experiments 1-10 are summarized in Table 2 and Table 3. Results with the experiment 11 are summarized in Table 4.

Each of the experiments 1-11 shown in Table 2, Table 3 and Table 4 are characterised briefly by the following aspects: the number of vertices / the number of edges / the number of one-way streets and it is divided into two variants of solution V1, V2. The basic network N_1 is used in the variant V1 (the number of edges / the number of one-way streets in the network N_1). The second variant V2 works with the networks N_1 and N_2 (the number of edges / the number of one-way streets in N_2). The total route length (in km) is stated in each of the variant. Furthermore, the percentage improvement of V2 compared to V1 is shown for each experiment. The computation time of all experiments was about 2 seconds.

6. Conclusion

This present paper deals with the planning of the optimal service vehicle route using the integer programming. We have introduced the mathematical model with constraints corresponding to the requirements of a real traffic. In particular, the constrains are the existence of a mixed transport network that contains one-way roads and the existence of a wider transport network, in which the only selected edges are used with possibility of effective passages. Data for the presented numerical experiments were both based on the real traffic values and artificially generated by the software *Wolfram Mathematica*. On the basis of these experiments performed in the computational environment *Xpress-IVE* was found, that this procedure can be appropriately applied in practice. As shown by the results, average savings computed by the mathematical model (5)-(9) are 13 % in the case of variant V2 comparing to the variant V1. While using the real data, achieved savings of the total distance traveled are 4.3 %. Thanks to the effective passages, the objective function has been decreased from 280 km to 268 km.

Acknowledgements

This paper was supported within Operational Programme Education for Competitiveness - Project No. CZ.1.07/2.3.00/20.0296 and programme for the support of applied research and experimental development, Technology Agency of the Czech Republic (EPSILON): Effective approaches to economical and adaptable systems of maintenance and operation of transport networks TH02010930.

References

[1] BORCINOVA, Z., PESKO, S.: New Exact Iterative Method for the Capacitated Vehicle Routing Problem. Communications - Scientific Letters of the University of Zilina, 18(3), 19-21, 2016.
[2] SIMCHI-LEVI, D., CHEN, X., BRAMEL, J.: The Logic of Logistics: Theory, Algorithms and Applications for Logistics Management. Springer, New York, 2014.
[3] SKIENA, S., PEMMARAJU, S.: Computational Discrete Mathematics. Cambridge University Press, New York, 2003.
[4] BONDY, A., MURTY, U.: Graph Theory. Springer, San Francisco, 2008.
[5] PELIKAN, J.: The Discrete Models in Operations Research. Professional Publishing, 2001.
[6] KOZEL, P., MICHALCOVA, S.: The Using of Linear Programming for Solving the Municipal Waste Collection Problem. Proceedings of 32nd International Conference Mathematical Methods in Economics, Czech Republic, 483-486, 2014.
[7] EISELT, H. A., GENDREAU, M., LAPORTE, G.: Arc Routing Problems, Part II: The Rural Postman Problem. Operational Research, 43(3), 399-414, 1995. https://doi.org/10.1287/opre.43.3.399
[8] KOZEL, P., MICHALCOVA, S.: The Use of Linear Programming to Solve Routing Tasks in Practice. Proceedings of 33rd International Conference Mathematical Methods in Economics, Czech Republic, 395-400, 2015.
[9] GUERET, CH., PRINS, CH., SEVAUX, M., HEIPCKE, S.: Applications of Optimization with Xpress-MP. Blisworth: Dash Associates, United Kingdom, 2002.
[10] XPRESS-Mosel "User guide". Blisworth: Dash Associates, United Kingdom, 2005.
[11] SOBEK, M.: Municipal Waste Collection Routes Optimization. Bachelor thesis (in Czech). VSB - TU Ostrava, Ostrava, 2018.
[12] SPACEK, F.: Mathematical Programming Application in Routing Problems Solving (in Czech). Master thesis. VSB - TU Ostrava, Ostrava, 2018.
[13] FRIEDRICH, V.: Mathematica for non-mathematician (in Czech). VSB - TU Ostrava, Ostrava, 2013.
[14] WOLFRAM, S.: The Mathematica Book. Cambridge University Press, New York, 1999.

COMMUNICATIONS – Scientific Letters of the University of Zilina
Author guidelines

1. Submitted papers must be unpublished and must not be currently under review for any other publication.
2. Submitted manuscripts should not exceed 8 pages including figures and graphs (in Microsoft WORD – format A4, Times Roman size 12, page margins 2.5 cm).
3. Manuscripts written in good English must include abstract and keywords also written in English. The abstract should not exceed 10 lines.
4. Submission should be sent by e-mail – as an attachment – to the following address: komunikacie@uniza.sk.
5. Uncommon abbreviations must be defined the first time they are used in the text.
6. Figures, graphs and diagrams, if not processed in Microsoft WORD, must be sent in electronic form (as JPG, GIF, TIF, TTF or BMP files) or drawn in high contrast on white paper. Photographs for publication must be either contrastive or on a slide.
7. The numbered reference citation within text should be enclosed in square brackets - in numerical order. The reference list should appear at the end of the article (in compliance with ISO 690).
8. The numbered figures and tables must be also included in the text.
9. The author's exact mailing address, full names, E-mail address, telephone or fax number, the name and address of the organization and workplace (also written in English) must be enclosed.
10. The editorial board will assess the submitted paper in its following session. If the manuscript is accepted for publication, it will be sent to peer review and language correction. After reviewing and incorporating the editor's comments, the final draft (before printing) will be sent to authors for final review and minor adjustments.

komunikacie
COMMUNICATIONS

VEDECKÉ LISTY ŽILINSKEJ UNIVERZITY
SCIENTIFIC LETTERS OF THE UNIVERSITY OF ZILINA
VOLUME 20 Issue 3

https://doi.org/10.26552/com.J.2018.3

Address of the editorial office:
University of Zilina
EDIS – Publishing House
Univerzitna 8215/1
010 26 Zilina
Slovakia

E-mail: komunikacie@uniza.sk

Individual issues of the journal can be found on:
http://www.uniza.sk/komunikacie

Each paper was reviewed by two reviewers.

Journal is excerpted in **SCOPUS, EBSCO** and **COMPENDEX**.

Published quarterly by University of Zilina in EDIS – Publishing House of University of Zilina

Registered No: EV 3672/09

ISSN (print version) 1335-4205
ISSN (online version) 2585-7878

ICO 00397 563

September 2018