Number 2, Volume VIII, July 2013

DATA DEPENDENT SYSTEMS AS A TOOL FOR MODELLING OF STOCHASTIC PROCESSES IN TRANSPORT AREA

Bohuš Leitner, Lucia Ďuranová[1]

Summary: The aim of the paper is short description of possibilities of data dependent systems (DDS) mathematical apparatus using. Basic form of DDS expression is form a stochastic time series of process which can be used to description and modelling of stochastic process, e.g. operational process. The paper contains a briefly characteristics of fundamental terms and equations of time series theory mathematical apparatus, algorithm of a statistically adequate discrete model of stochastically dynamic system and developed strategy application on real process.

Key words: data dependent system, stochastic operational process, time serie.

INTRODUCTION

One of possible way of complex systems analysis without loss of accuracy and without necessity of complicated mathematical apparatus utilisation is observation of system during its working and utilisation of produced data to its analyses. Such analysed system we call Data Dependent Systems (DDS). It means that we do not need to know anything about composition of system and all analysis and conclusions are made just based on observed system output. Data dependent systems are represented by sets of continuous output signals which are discretized with uniform sampling interval gets a series of data (values) which forms a base for description and analysis of investigated system. The main aim is to get possibility of system behaviour forecasting and eventually influencing of its behaviour not to get system in any unwanted state (2).

For description of DDS, one can use autoregressive models, advantageously Auto-Regressive Moving Average Models (next as ARMA models). Their advantages are mainly precisely formulated statistically criteria and relatively simple mathematical apparatus of statistically regressive analysis and testing of statistical hypothesis. Their disadvantages are complications of identification procedures and time consuming computer calculations despite relatively simple mathematical apparatus. Therefore, they are suitable just for off-line dynamic systems identification and modelling and their main application – forecasting of dynamic system behaviour is suitable just by stationary systems, which do not change their parameters with time.

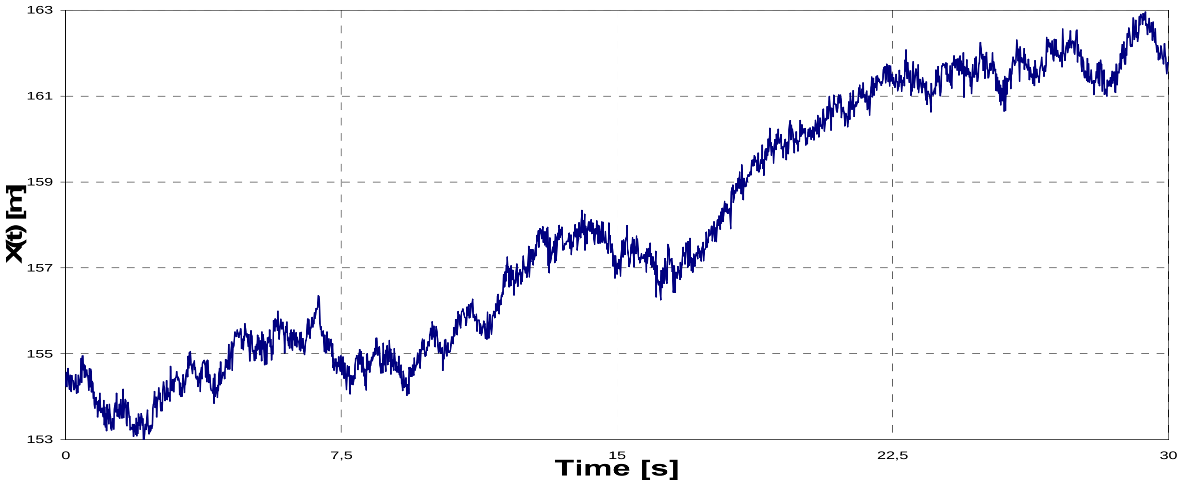
---

[1] doc. Ing. Bohuš Leitner, PhD., Ing. Lucia Ďuranová, Univerzity of Zilina, Faculty of Special Engineering, Department of Technical Sciences and Informatics, 1.mája 32, 010 26 Žilina, SR, Tel.: +421 41 513 6863, Fax: +421 41 513 6612, E-mail: Bohus.Leitner@fsi.uniza.sk, Lucia.Duranova@fsi.uniza.sk

1. DATA DEPENDENT SYSTEMS AND ITS THEORETICAL APPARATUS

First logical technique for identification is to derive the mathematical model from system behaviour and directly from nature of its physical characteristics or as the dependence on input factors influencing this behaviour. Mathematical representations of such relations are often stated in form of differential equations systems. Solutions of such systems whether numerical or exact is very time-consuming therefore not usable for the real time control. Another problem with this possible way of identification is: some of the influencing factors are not well known and some can not be quantified. Another possible technique for system identification is to derive the mathematical model from internal interdependencies of given process only.

Clear Autoregressive (AR) model is suitable mathematical model for discrete non-stationary signal (Fig.1) that is considered to be visible representation of system dynamics (2). Dependence of X (t) values on previous (in time) values X (t-1) is pictured in Fig. 2.



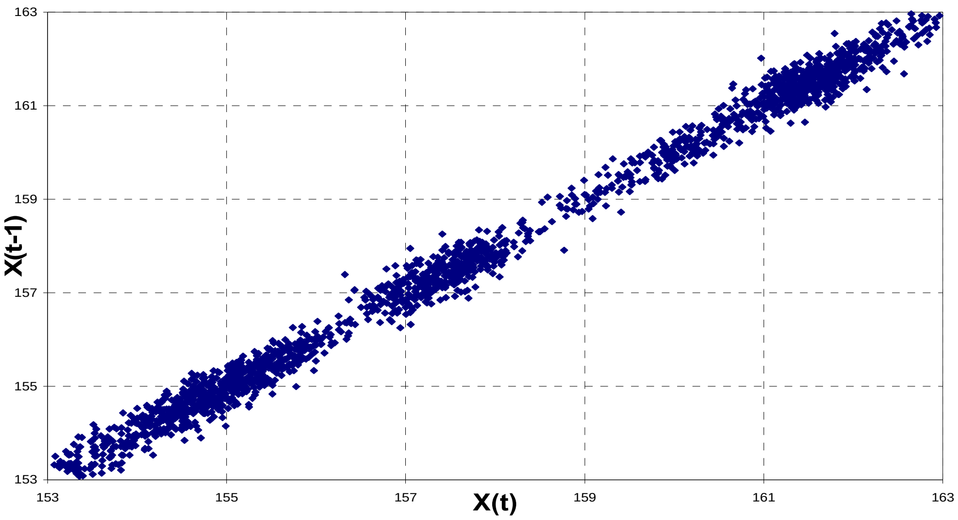
Source: (2)

Fig.1 - Example of real non-stationary process

Linear regression can be clearly seen. This relation can be described by simple 1st-order Autoregressive model (1) in form (2):

$$X(t) = a_1.X(t-1) + \varepsilon_t \tag{1}$$

where *t* is the time, *X(t)* is discrete signal value in time t, *X (t -1)* is discrete signal value in time (t-1), a_1 is coefficient of AR model and εt is error resulting from model imperfections.



Source: (2)

Fig.2 - Interdependence X (t)=f {X (t-1)}

This simplest model can be extended to universal n-th order of clear auto-regressive model AR (n)

$$X(t) = a_1.X(t-1) + a_2.X(t-2) + \ldots + a_n.X(t-n) + \varepsilon_t \tag{2}$$

where n is the model order, *X (t), X (t-1), X (t-2) ..., X (t-n)* are discrete signal values in time *t, t-1, t-2 ... t-n,* a_1, a_2, ..., a_n are coefficients of Autoregressive model and ε_t is error resulting from model imperfections [2,3]. If *X (t)* depends not only on preceding values *X (t-1), X (t-2)… X (t-n)* but also on the values of preceding errors ε_{t-1}, ε_{t-2} ... simple AR models are extended to Autoregressive moving average models (ARMA). The ARMA models can cover more complex character of process interdependencies and their coefficients relate closely to physical principles of observed process.

Universal ARMA model of n-th order in Autoregressive part and (n-1)-th order in moving average part is given by equation (3).

$$X(t) = a_1.X(t-1) + a_2.X(t-2) + \ldots + a_n.X(t-n) + \\ + b_1.\varepsilon_{t-1} + b_2.\varepsilon_{t-2} + \ldots + b_{n-1}.\varepsilon_{t-n+1} + \varepsilon_t \tag{3}$$

where *X(t), X(t-1), X(t-2) ... X(t-n)* are the discrete signal value in time t, t-1, t-2 ... t-n, values ε_t, ε_{t-1}, ε_{t-2} .. ε_{t-n+1} is error in the time t, t-1, t-2… t-n+1, values a_1, a_2 ... a_n are coefficients of Autoregressive part of model, b_1, b_2 ... b_n are coefficients of moving average part of model.

Main problem in identification, modelling and simulation by Autoregressive models is finding coefficients of AR and MA parts and determination of adequate order of the model.

Coefficients of AR model can be simply found using least square method (LSM). For universal ARMA models non-linear LSM must be used. Both methods use matrix calculations for finding needed coefficients, which are very time-consuming and therefore not usable for on-line process, control or identification and they also can not be used for modelling of non-stationary time-varying process (3).

2. ALGORITHM OF ADAPTIVE AR AND ARMA MODELS

Because of the above stated reasons, procedure based on theory of adaptive and self-learning systems is used for describing system behaviour in real time. Algorithm for adaptive modelling (3) is based on a gradient method (steepest descent method) and can also be used for non-stationary processes. Model is able to adapt itself to the changes in process character. It is supposed that n-th order Autoregresive model (2) is at any given time defined by the vector of its coefficients:

$$\mathbf{a}(k) = [a_1(k), a_2(k) \ldots a_n(k)]^T \tag{4}$$

Using the steepest descent method, point of least squares Σ ε_{t2} is searched for. Search begins with an initial guess as to where the minimum point of Σ ε_{t2} may be. Minimal sum of squares S is in the point where

$$\frac{\partial S}{\partial a_k} = \frac{\partial}{\partial a_k}\left(\sum \varepsilon_t^{\,2}\right) = 0 \tag{5}$$

The updated values of AR model coefficients are obtained from

$$\mathbf{a}(k+1) = \mathbf{a}(k) + \eta . \frac{\partial S}{\partial a} \tag{6}$$

where $\frac{\partial S}{\partial a} = -2\left[\varepsilon_{ft} . \mathrm{X}^T(k-1)\right]$ is the gradient direction and positive value of η in equation (6) scales the amount of readjustment of the model coefficients in one time step. Then, the iterative corrections of coefficients are

$$\mathbf{a}(k+1) = \mathbf{a}(k) + \mu . \left[e_t . \mathbf{X}^T(k-1)\right]. \tag{7}$$

In [3] adaptive AR models were extended to include also MA part to adaptive AutoRegresive models with moving average. To achieve this, vector of moving average coefficients must be considered

$$\mathbf{b}(k) = [b_1(k), b_2(k) \ldots b_n(k)]^T \tag{8}$$

Same procedure as for vector of AR coefficients was used to derive formula (10) for iterative corrections' calculations of MA part coefficients

$$\mathbf{b}(k+1) = \mathbf{b}(k) + \mu . \left[\varepsilon_t . \varepsilon^T(k-1)\right] \tag{9}$$

where ε t is error from the last iterative step and ε (k-1) is vector of preceding errors

$$\boldsymbol{\varepsilon}(k-1) = \left[\varepsilon_{k-1}, \varepsilon_{k-2}, \ldots \varepsilon_{k-n+1}\right]^T \tag{10}$$

Another problem arises when deciding value of convergence constant. It influences converging speed of algorithm and also its sensitivity to random or systematic changes in process environment character. Procedure for calculating μ constant, based on experimental work [2,3] was presented for use in area of adaptive control:

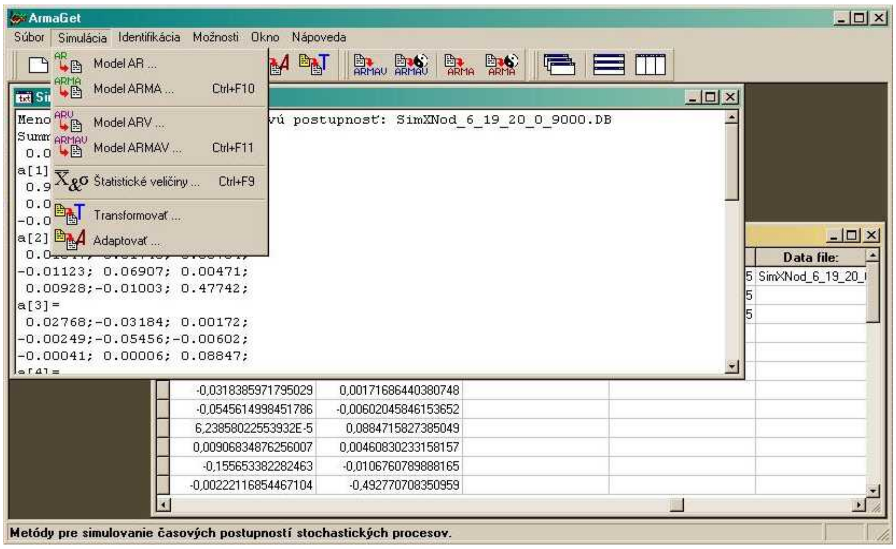
$$\mu_k = \varphi + \beta . C_k \quad (11) \qquad \text{and} \qquad C_k = \left(1 - \frac{1}{\alpha}\right) C_{k-1} + \frac{1}{\alpha} . \varepsilon_t^{\,2}, \quad (12)$$

where α is constant describing system memory, it influences model sensitivity to random process changes, β is constant characterising system dynamics and φ is constant for correction of numeric calculation errors. Actual values of these constants can be chosen so the model sensitivity to stochastic events and response speed to process character changes are as required.

3. SOFTWARE SUPPORT FOR STOCHASTIC PROCESSES MODELLING

The scalar models of a simple description of dynamic system can not express statistically adequate description of complex systems. For this reason there was developed an effective software system which enables to create the statistically models of dynamic stochastic system by using ARMA or vector autoregressive models VARMA. The final form of an identification software was created in such a way that it enables the use of an identification library and to realise the own identification of system parameters.

The result of a proposed application of this methodology is ArmaGet software, this was developed on authors department. Software support is fully compatible with Microsoft Windows systems. It contains users menu, which apart from basic functions with file, configurations, work with windows and help functions contains two submenus – submenu of “***Simulation***” and submenu of “***Identification”*** (Fig.3).



```
Meno                      vú postupnosť: SimXNod_6_19_20_0_9000.DB
Summ
 0.0
a[1]
 0.9
 0.0
-0.0
a[2]
 0.0
-0.01123; 0.06907; 0.00471;
 0.00928;-0.01003; 0.47742;
a[3]=
 0.02768;-0.03184; 0.00172;
-0.00249;-0.05456;-0.00602;
-0.00041; 0.00006; 0.08847;
a[4]=
```

-0,0318385971795029	0,00171686440380748
-0,0545614998451786	-0,00602045846153652
6,23858022553932E-5	0,0884715827385049
0,00906834876256007	0,00460830233158157
-0,155653382282463	-0,0106760789888165
-0,00222116854467104	-0,492770708350959

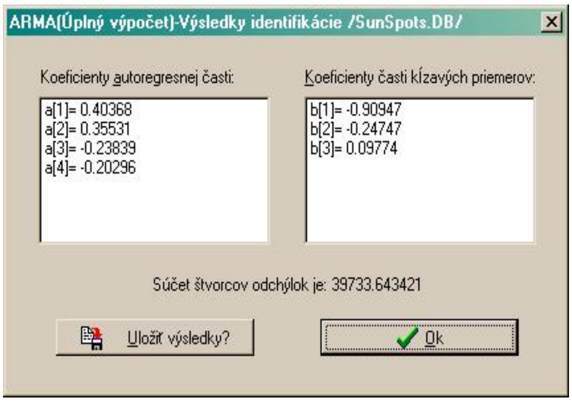
	Data file:
5	SimXNod_6_19_20_
5	
5	

Metódy pre simulovanie časových postupností stochastických procesov.

Fig. 3 - Main window of created application ArmaGet

The heart of the program is submenu “***Identification***”, by means, which is possible to make selection of the identification method and way of chosen time series, whereupon it is possible to use either adaptive algorithm of time series identification or make identification using non-linear least squares method. Identification by means of higher presented non-linear (respectively for AR models – linear) least square method is available in item Identification and its sub-menu NLINLS. Here are four options. First two - *Model AR* – after orders and *Model AR- complete calculation* give as results of identification AR model, described by (2).

Next item – *Model ARMA* - after orders gives coefficients of beforehand selected order of ARMA (n, n-1) models determination. It means, that we it is necessary beforehand to determine required order (known number of coefficients) of autoregressive part - a_k and moving average part - b_k which principally determine number of former values the calculated value depends on. The initial guess is of coefficients of ARMA (n, n-1) model. To coefficients of moving average part are assigned value of zero and coefficients of autoregressive part obtained by application of linear least square method. Simultaneously the sum of squares of deviations value expressing deviation of theoretical model from real model is calculated. Then the proper iterative calculation follows - outputs are the coefficients of model. The last important item from submenu "Identification" is - *Model ARMA - complete calculation* - which aim is to find an optimal ARMA (n, n-1) model. This model is the best describe of stochastic system, which output is a time serie. Because, in most cases we don't know optimal order of model, beforehand it is necessary to determine by an iterative procedure an optimal order of model for description of given system (Fig.4). An algorithm of optimum autoregressive model determination used in ARMAGET is identical as algorithm from part 2 of this paper.

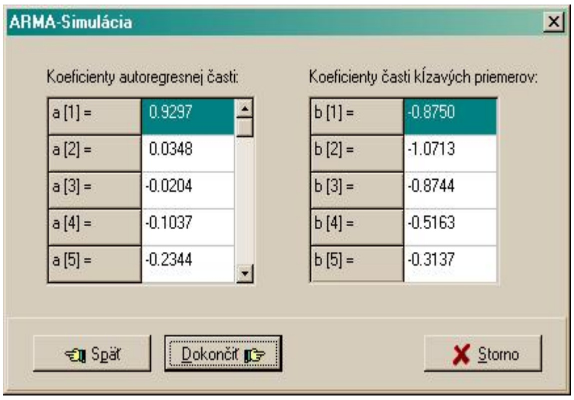


Koeficienty autoregresnej časti:
a[1]= 0.40368
a[2]= 0.35531
a[3]= -0.23839
a[4]= -0.20296

Koeficienty časti kĺzavých priemerov:
b[1]= -0.90947
b[2]= -0.24747
b[3]= 0.09774

Súčet štvorcov odchýlok je: 39733.643421

Fig.4 - Results of identification



Koeficienty autoregresnej časti:
a [1] = 0.9297
a [2] = 0.0348
a [3] = -0.0204
a [4] = -0.1037
a [5] = -0.2344

Koeficienty časti kĺzavých priemerov:
b [1] = -0.8750
b [2] = -1.0713
b [3] = -0.8744
b [4] = -0.5163
b [5] = -0.3137

Fig.5 - Results of simulation of time serie from parameters of ARMA model

Item "***Simulation***" enables adjustment and conversion of incompatible input files of time series to compatible ones and simulation (generation) of time series basing on given AR or ARMA models order and parameters with possibilities of mean and dispersion selection of simulated serie (Fig.5). The accuracy and the reliability of the developed methods and algorithms were verified by use of commercial software packages (MS Excel, MATLAB) and will be compared and verified by using of utility ARMASA Package [3] etc.).

4. MODELLING OF REAL WORKING CONDITIONS PROCESSES IN TRANSPORT SYSTEMS

Because the chosen mathematical apparatus allows getting of models based of physical principles of researched processes which are developed from recorded courses of examined values it was necessary to obtain experimental data, which characterise typical working conditions (1, 4). Therefore a tot of experimental measurements was done which describe typical working conditions of trucks (unevenness of road surface, segmentation of terrain, typical working speed during different types of working modes e.t.c). Longitudinal and transverse unevenness had a dominant position as a typical of transport machines working conditions generally.

The experimental measurement was realised which consisted of following steps:

- Finding of roads and their parts representing individual categories of roadways from point of view of quality of their coverage,
- Selection of actual parts of chosen length,
- Own measurement using suitable equipment,
- Filtering of trends and rough unevenness with suitable software package Statgraphics©
- Evaluation of chosen parts by means of selected dynamic methods [1,4] and verification of their classification for relevant categories of quality.

As a result of former steps there were obtained courses of longitudinal unevenness together from 6 sectors (5 qualitative categories of road and one sector of terrain). Because of limited scope of this paper just course of 5th category of road and the values of terrain are presented on Fig.6.

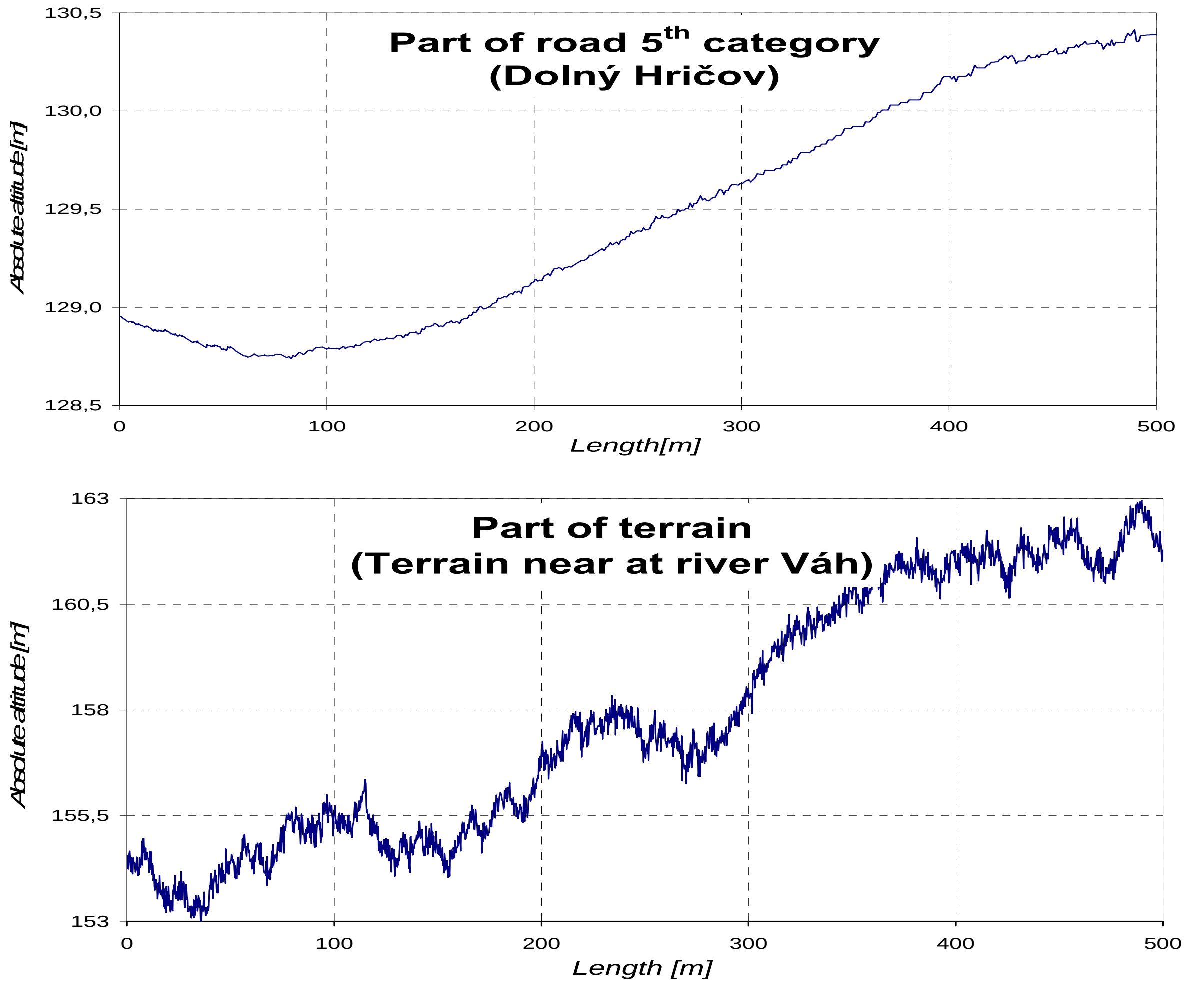


Fig.6 - Courses of absolute high unevenness of 5th category road and terrain

The values of evaluating parameters "C" and "IRI" are presented in Tab.1. If was determined that from obtained parameters of "C" and "IRI" chosen sections really are represented chosen qualitative categories and there for they can be judged as representative of qualitative different surfaces [1].

Tab.1 - Calculated values of „C“ and „IRI“ parameters for chosen sections of road

Location	"C"	"IRI"	Category
Žilina	*0,85*	*1,61*	***I***
Brodno	*2,31*	*2,62*	***II***
Čičmany	*3,92*	*4,28*	***III***
Varín	*13,85*	*8,54*	***IV***
Dolný Hričov	*22,81*	*14,37*	***V***
Povodie Váhu	*42,05*	*22,74*	***Terrain***

It was necessary from of point of view of further practical application of experimentally obtained models to separate the random part of unevenness from its trend. It was used the software pack Statgraphics© which allows directly selection of trends and seasonally from recorded values. To demonstrate this are on Fig.7 shown courses of stochastic parts of profile unevenness for series from Fig.6 .

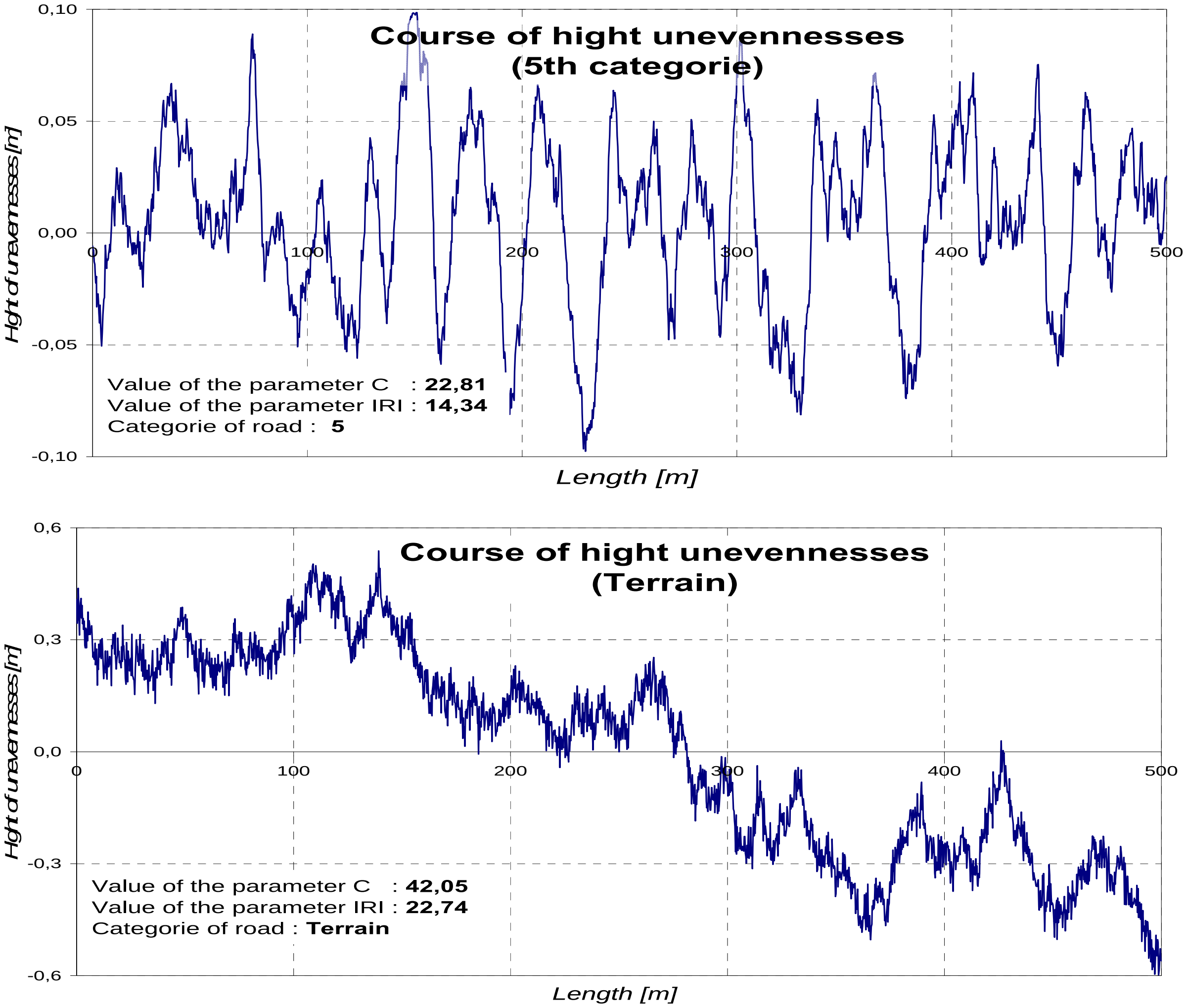


Fig.7 - Random components of altitude of unevenness for 5th category and terrain

It was used the ArmaGet software to get real and simulated data. Practically it means optimal autoregressive model and its parameter determination (order and coefficients of model) for all inputs of discrete values of altitude unevenness to selected sectors of roads.

Finding parameters of adequate models for sector of 5th category road and sector of terrain are presented in Tab. 2.

Tab.2 - Parameters of adequate ARMA models for sectors of 5th category and terrain

Cat.	**Coefficients of optimal ARMA model**				**Sum of squares**
5.	a(1) = 0,623808 a(2) = 0,095081 a(3) = 0,592155	a(4) = -0,050585 a(5) = 0,518931 a(6) = -0,797272	b(1) = -0,391528 b(2) = -0,335376 b(3) = 0,197765	b(4) = 0,196800 b(5) = 0,777805	SSC = 0,063145
Terr	a(1) = 0,867707 a(2) = 0,186025	a(3) = -0,057025 a(4) = 0,002244	b(1) = 0,542767 b(2) = 0,046748	b(3) = -0,047127	SSC = 2,545889

These models were used for generation of new, statistically adequate, series of analysing process values. For this purpose software ArmaGet contains the item "***Simulation***" which was successfully applied by generation of new courses altitude of unevenness of reference sectors. For presented parts of 5th category and terrain there are two new simulated courses and initial experimental course presents on Fig.8.

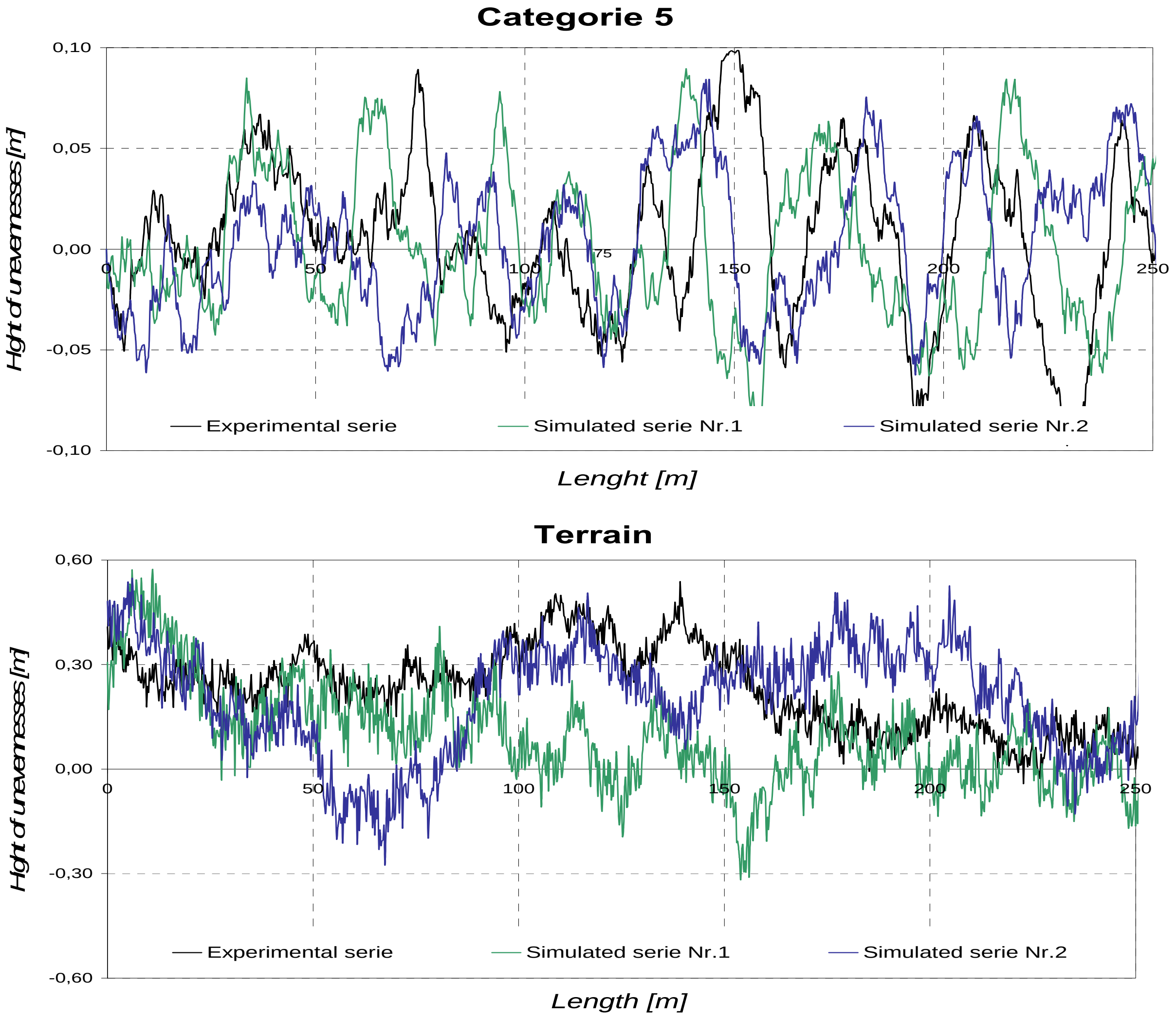


Fig.8 - The courses of simulated series of 5th category road and terrain sectors

For computer simulation of working condition characteristics available especially on computer stress analysis of critically parts of trucks, on influence analyses to selected working properties of transport machines and for automatic control of various load machines too was used the finding adequate models.

During verifying their physical adequacies were once more used dynamic methods of evaluation sectors of roads from point of view longitudinal unevenness. It was found and verified that determined mathematical models very good describe both analysed characteristics. Determined differences were in both of tested factors neglect able and statistically non-significant.

CONCLUSIONS

It is well known that working of majority of machines is significantly influenced by different kinds of stochastic loads. There is possible to respect the tendency a limitation of energetically and material consumption to oversize their dimensions. But it is necessary to look for some more ingenious methods to deal with these problems. Some of them are the ways to control (influence) the working of a mechanical system in respect to their proposed behaviour. But it needs to follow of the system behaviour in the real time and to make some necessary controlling interventions.

It was not possible from point of view of determined scope of paper to show all of obtained outputs from realised calculations and experiment (mainly made on computers). Therefore they are presented just some of them in this paper. Based on above-mentioned results it is possible to state that chosen theoretical apparatus and methodology of autoregressive modelling (namely adaptive one). Is a suitable tool for modelling and simulation of different working condition factors.

REFERENCES

(1) KOVÁČ, M., REMIŠOVÁ, DECKÝ, M., ĎURČANSKÁ, D., ČELKO, J.: *Diagnostika parametrov prevádzkovej spôsobilosti vozoviek.* 1. vyd., Žilina: Žilinská univerzita, 2012. 256 s., ISBN 978-80-554-0568-1

(2) LEITNER, B.: *Autoregressive models in modelling and simulation of transport means working conditions.* In: Transport means: the 14th international conference: Kaunas University of Technology, Kaunas, Lithuania. 2010, s. 21-24, ISSN 1822-296X.

(3) LEITNER, B.: *The software tool for mechanical structures dynamic systems identification..* In: 15th international conference Transport means 2011: October 20-21, 2011, Kaunas University of Technology, Lithuania, s. 38-41, ISSN 1822-296X.

(4) ČELKO, J., DECKÝ, M., KOVÁČ, M.: *An analysis of vehicle – road surface interaction for classification of IRI in frame of Slovak PMS.* In: Maintenance and reliability. Polish Maintenance society, Nr 1 (41) 2009, str.15-21, ISSN 1507-2711.