FORECASTING USING SIMPLE EXPONENTIAL SMOOTHING METHOD

Eva OSTERTAGOVÁ[*], Oskar OSTERTAG[**]

[*]Department of Mathematics and Theoretical Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Němcovej 32, 042 00 Košice, Slovak Republic, tel.: +421 55 602 2447, e-mail: eva.ostertagova@tuke.sk

[**]Department of Applied Mechanics and Mechatronics, Faculty of Mechanical Engineering, Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic, tel.: +421 55 602 2460, e-mail: oskar.ostertag@tuke.sk

ABSTRACT

In the paper a relatively simple yet powerful and versatile technique for forecasting time series data – simple exponential smoothing is described. The simple exponential smoothing (SES) is a short-range forecasting method that assumes a reasonably stable mean in the data with no trend (consistent growth or decline). It is one of the most popular forecasting methods that uses weighted moving average of past data as the basis for a forecast. The procedure gives heaviest weight to more recent observations and smaller weight to observations in the more distant past. The accuracy of the SES method strongly depends on the optimal value of the smoothing constant α. To determine the optimal α value in the paper was used a traditional optimalization method based on the lowest mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE).

Keywords: *time series, simple exponential smoothing model, forecast, smoothing constant, root mean square error*

1. INTRODUCTION

A time series is a sequence of observations indexed by time, usually ordered in equally spaced intervals and correlated. In our days it is well known the importance of time series studies. These studies provide indicators about a country economy, the unemployment rate, the export and import product rates, etc. The most interesting and ambitious task in time series analysis is to forecast future values. Models are commonly fitted in order to predict future values of a time series [4].

Exponential smoothing methods are the most widely used forecasting methods. The formulation of exponential smoothing forecasting methods arose in the 1950s from the original work of Brown (1959, [2]) and Holt (1957, [6]) who were working on creating forecasting models for inventory control systems.

Exponential smoothing is an intuitive forecasting method that weights the observed time series unequally. Recent observations are weighted more heavily than remote observations. The unequal weighting is accomplished by using one or more smoothing parameters, which determine how much weight is given to each observation [9].

The simplest technique of this type, simple exponential smoothing (SES), is appropriate for a series that moves randomly above and below a constant mean (stationary series). It has no trend and no seasonal patterns [16].

The Holt-Winters method, also referred to as double exponential smoothing, is an extension of exponential smoothing designed for trended and seasonal time series. Holt-Winters smoothing is a widely used tool for forecasting business data that contain seasonality, changing trends and seasonal correlation [5].

Exponential smoothing model is a widely used method in time series analysis. This popularity can be attributed to its simplicity, its computational efficiency, the ease of adjusting its responsiveness to changes in the process being forecast, and its reasonable accuracy [11].

Generally, exponential smoothing is regarded as an inexpensive technique that gives good forecast in a wide variety of applications. In addition, data storage and computing requirements are minimal, which makes exponential smoothing suitable for real-time application.

2. SIMPLE EXPONENTIAL SMOOTHING MODEL

The simple exponential smoothing (SES) model is usually based on the premise that the level of time series should fluctuate about a constant level or change slowly over the time [9].

2.1. Mathematical Formulation

The SES model is given by the model equation

$$y(t) = \beta(t) + \varepsilon(t), \tag{1}$$

where $\beta(t)$ takes a constant at the time t and may change slowly over the time; $\varepsilon(t)$ is a random variable and is used to describe the effect of stochastic fluctuation.

Let an observed time series be y_1, y_2, …, y_n. In any case, in this simple model, to predict y_t is merely to predict (estimate) β. To estimate, it makes sense to use all the past observations, but due to declining correlation as you go back into the past, to down-weight older observations.

Formally, the simple exponential smoothing equation takes the form of

$$F_{t+1} = \alpha y_t + (1 - \alpha) F_t, \tag{2}$$

where y_t is the actual, known series value at the time t; F_t is the forecast value of the variable Y at the time t; F_{t+1} is the forecast value at the time $t + 1$; α is the smoothing constant [3].

The forecast F_{t+1} is based on weighting the most recent observation y_t with a weight α and weighting the most recent forecast F_t with a weight of $1 - \alpha$.

To get started the algorithm, we need an initial forecast, an actual value and a smoothing constant.

Since F_1 is not known, we can:

- Set the first estimate equal to the first observation. Further we will use $F_1 = y_1$.

- Use the average value of the first few observations of the data series for the initial smoothed value.

Smoothing constant α is a selected number between zero and one, $0 < \alpha < 1$.

Rewriting the model (2) we can see one of the neat things about the SES model

$$F_{t+1} - F_t = \alpha(y_t - F_t), \qquad (3)$$

change in forecasting value is proportional to the forecast error. That is

$$F_{t+1} = F_t + \alpha e_t, \qquad (4)$$

where residual

$$e_t = y_t - F_t \qquad (5)$$

is the forecast error at the time t.

So, the exponential smoothing forecast is the old forecast plus an adjustment for the error that occurred in the last forecast [1, 12].

By iterating formula (2) we get:

$$F_1 = y_1; \qquad F_2 = \alpha \cdot y_1 + (1-\alpha) \cdot F_1 = y_1;$$

$$F_3 = \alpha \cdot y_2 + (1-\alpha) \cdot F_2 = \alpha \cdot y_2 + (1-\alpha) \cdot y_1 =$$

$$\alpha \cdot y_2 + \alpha(1-\alpha) \cdot y_1 + (1-\alpha)^2 \cdot y_1;$$

$$F_4 = \alpha \cdot y_3 + (1-\alpha) \cdot F_3 =$$

$$\alpha y_3 + (1-\alpha)(\alpha \cdot y_2 + \alpha(1-\alpha) \cdot y_1 + (1-\alpha)^2 \cdot y_1) =$$

$$\alpha\big(y_3 + (1-\alpha)y_2 + (1-\alpha)^2 y_1\big) + (1-\alpha)^3 \cdot y_1; \quad \dots$$

The forecast equation in general form is

$$F_{t+1} = \alpha \sum_{k=0}^{t-1} (1-\alpha)^k y_{t-k} + (1-\alpha)^t y_1, \quad t \in \mathbb{N}, \qquad (6)$$

where F_{t+1} is the forecast value of the variable Y at the time $t+1$ from knowledge of the actual series values y_t, y_{t-1}, y_{t-2} and so on back in time to the first known value of the time series, y_1 [3, 13].

Therefore, F_{t+1} is the weighted moving average of all past observations.

The series of weights used in producing the forecast F_{t+1} is

$$\alpha, \ \alpha(1-\alpha), \ \alpha(1-\alpha)^2, \ \dots \qquad (7)$$

It is obviously from (7) that the weights are exponential; hence the name exponentially weighted moving average [1]. The exponential decline of the weights toward zero is evident. This is shown in Figure 1.

The decay is slower for small values α, so we can control the rate of decay by choosing α appropriately.
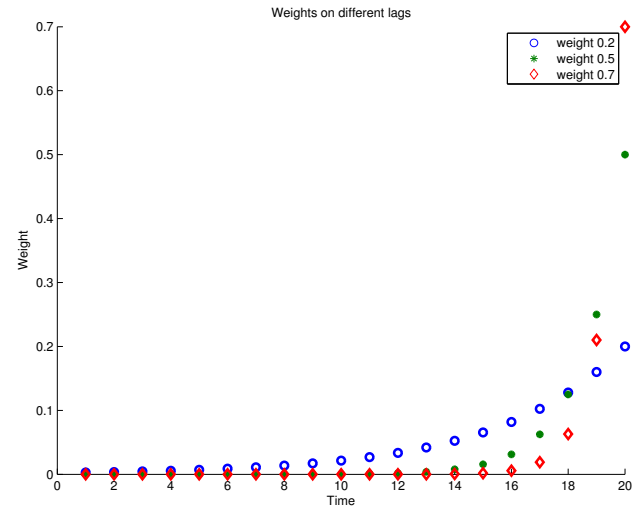

Fig. 1 Exponentially declining weights

2.2. Measuring Forecast Error

After the model specified, its performance characteristics should be verified or validated by comparison of its forecast with historical data for the process it was designed to forecast.

This is no consensus among researchers as to which measure is best for determining the most appropriate forecasting method. Accuracy is the criterion that determines the best forecasting method; thus, accuracy is the most important concern in evaluating the quality of a forecast. The goal of the forecast is to minimize error [14].

Some of the common indicators used to evaluate accuracy are *MAE* (Mean absolute error), *MSE* (Mean squared error), *RMSE* (Root mean squared error) or *MAPE* (Mean absolute percentage error):

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |e_t|, \qquad (8)$$

$$MSE = \frac{1}{n} \sum_{t=1}^{n} e_t^2, \qquad (9)$$

$$RMSE = \sqrt{MSE} \qquad (10)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \frac{|e_t|}{y_t} \cdot 100\%, \qquad (11)$$

where y_t is the actual value at the time t; e_t is residual at the time t; n is the total number of the time periods.

MAE is a measure of overall accuracy that gives an indication of the degree of spread, where all errors are assigned equal weights. If a method fits the past time series data very good, *MAE* is near zero, whereas if a method fits the past time series data poorly, *MAE* is large. Thus, when two or more forecasting methods are compared, the one with the minimum *MAE* can be selected as most accurate [14].

MSE is also a measure of overall accuracy that gives an indication of the degree of spread, but here large errors

are given additional weight. It is a generally accepted technique for evaluating exponential smoothing and other methods [7].

Often the square root of *MSE*, *RMSE*, is considered, since the seriousness of the forecast error is then denoted in the same dimensions as the actual and forecast values themselves.

MAPE is a relative measure that corresponds to *MAE*. It is the most useful measure to compare the accuracy of forecasts between different items or products since it measures relative performance. It is one measure of accuracy commonly used in quantitative methods of forecasting [10]. If *MAPE* calculated value is less than 10 %, it is interpreted as excellent accurate forecasting, between 10–20 % good forecasting, between 20–50 % acceptable forecasting and over 50 % inaccurate forecasting [8].

Selection of an error measure has an important effect on the conclusions about which of a set of forecasting methods is most accurate.

2.3.   Choosing the Best Value for Smoothing Constant

The accuracy of forecasting of SES technique depends on smoothing constant. Choosing an appropriate value of exponential smoothing constant is very crucial to minimize the error in forecasting.

Selecting a smoothing constant is basically a matter of judgment or trial and error, using forecast errors to guide the decision. The goal is to select a smoothing constant that balances the benefits of smoothing random variations with the benefits of responding to real changes if and when they occur.

The smoothing constant serves as the weighting factor. When α is close to 1, the new forecast will include a substantial adjustment for any error that occurred in the preceding forecast. When α is close to 0, the new forecast is very similar to the old forecast.

The smoothing constant α is not an arbitrary choice but generally falls between 0.1 and 0.5. Low values of α are used when the underlying average tends to be stable; higher values are used when the underlying average is susceptible to change.

In practice, the smoothing constant is often chosen by a grid search of the parameter space; that is, different solutions for α are tried starting, for example, with $\alpha = 0.1$ to $\alpha = 0.9$, with increments of 0.1 [1,12]. The value of α with the smallest *MAE*, *MSE*, *RMSE* or *MAPE* is chosen for use in producing the future forecasts.

3.   EXAMPLE FROM TECHNICAL PRACTICE

The simple exponential smoothing model can be illustrated by using data about primary production of electricity (y_t) in terajoules (TJ) in Slovakia over the years 2001–2009 that are in the Table 1 [17].

We want to forecast primary production of electricity in Slovakia for the year 2010.

We can see from the plot that there is roughly constant level. Thus, we can make forecasts using SES method. It is illustrated in Figure 2.

Table 1 Observed values of primary production of electricity

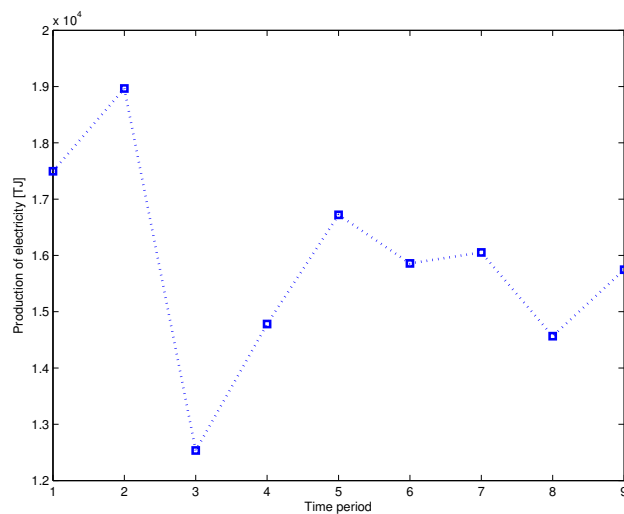| Time t	Observed values y_t
1	17496
2	18965
3	12535
4	14781
5	16722
6	15858
7	16053
8	14565
9	15747

Fig. 2   Observed time series data

In this study were adopted to assess the accuracy of forecasting methods three accuracy models – *MAE*, *MAPE* and *RMSE*.

These error measures were calculated using equations (2), (5) and (8)–(11) for different values of exponential smoothing constant using MATLAB. Table 2 shows the values of *MAE*, *MAPE* and *RMSE* for different α.

Table 2 *MAE*, *MAPE*, *RMSE* for different values of α

α	*MAE*	*MAPE*	*RMSE*
0.25	1405.09	9.97	2139.29
0.26	1402.04	9.95	2138.86
0.27	1399.27	9.93	2138.99
0.28	1396.76	9.91	2139.63
0.29	1394.53	9.89	2140.74
0.30	1396.23	9.90	2142.30

Values of *MAE*, *MAPE*, and *RMSE* depend on the choice of the smoothing constant.

MAE and *MAPE* are decreased with increasing α up to 0.29 and after that *MAE* and *MAPE* are increased. Variation of *MAE* with α is shown in Figure 3 and variation of *MAPE* with α is shown in Figure 4.
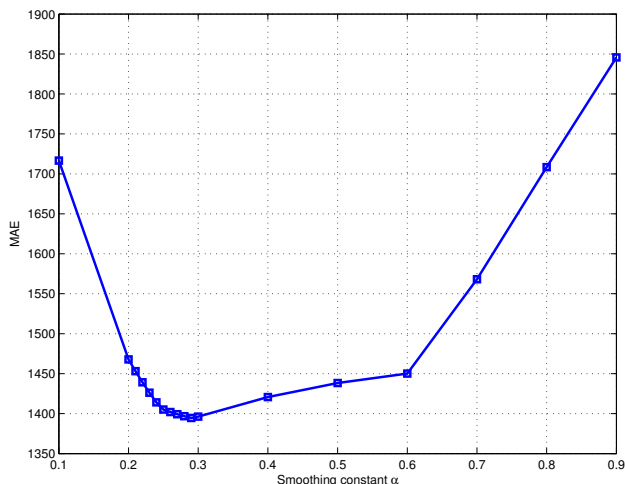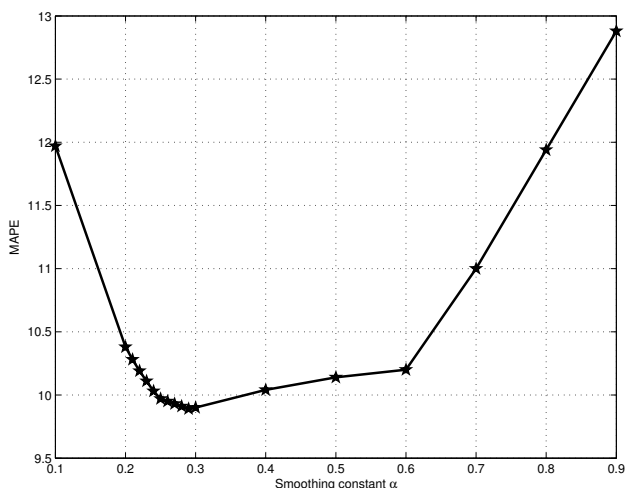
Fig. 3 Variation of *MAE* for different values of α

Fig. 4 Variation of *MAPE* for different values of α
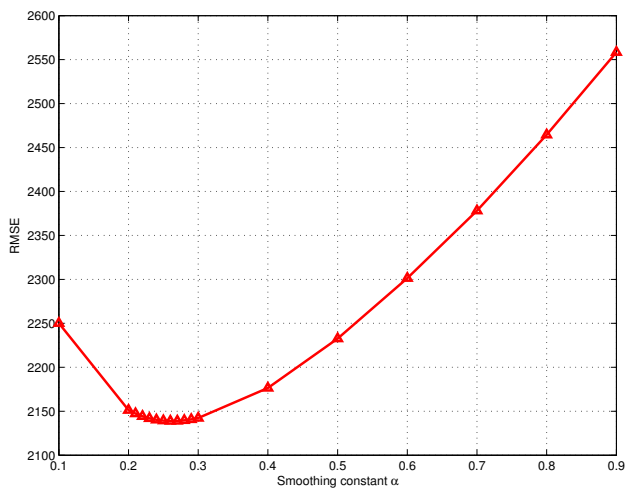

Fig. 5 Variation of *RMSE* for different values of α

RMSE is decreased with increasing α up to 0.26 and after that *RMSE* is increased. Variation of *RMSE* with α is shown in Figure 5.

To find the optimal value of smoothing constant, minimum values of *MAE*, *RMSE* and *MAPE* are selected and corresponding value of smoothing constant is the optimal value for this problem [13].

By calculating the forecast values using equation (1) for $\alpha = 0.29$ and $\alpha = 0.26$, we obtain the values presented in the columns 3–4 in the Table 3.

In the last line in the Table 3 are forecast results for the time period 10, therefore for year 2010.

So the forecast of primary production of electricity in Slovakia for year 2010 by using $\alpha = 0.29$ is about 15660 TJ and by using $\alpha = 0.26$ it is about 15717 TJ.

For both values of alpha was obtained result value of *MAPE* less than 10 % and it means the excellent accurate forecasting.

Table 3 Forecast values for minimum *MAE*, *MAPE* and *RMSE*

Period t	Observed values y_t	Forecast F_t ($\alpha = 0.29$)	Forecast F_t ($\alpha = 0.26$)
1	17496	17496.00	17496.00
2	18965	17496.00	17496.00
3	12535	17922.01	17877.94
4	14781	16359.78	16488.78
5	16722	15901.93	16044.75
6	15858	16139.75	16220.84
7	16053	16058.04	16126.50
8	14565	16056.58	16107.39
9	15747	15624.02	15706.37
10	–	15659.69	15716.93

This concept is illustrated in Figure 6 which shows a time series observed for periods 1 to 9 and corresponding forecast values for period 10 by using $\alpha = 0.29$ and $\alpha = 0.26$.
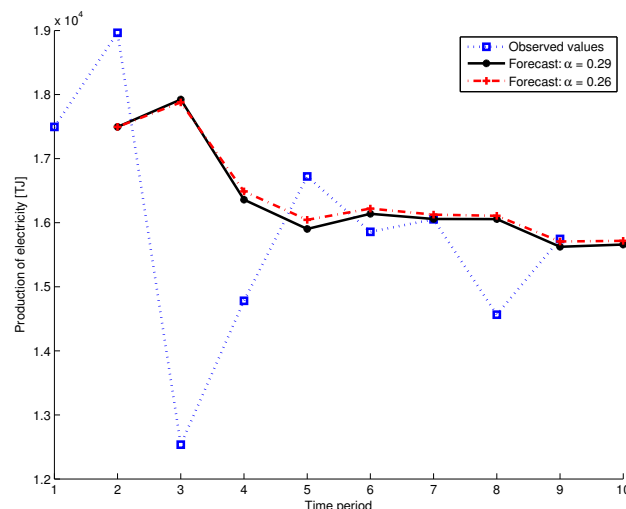

Fig. 6 Comparison of observed values and forecast values using $\alpha = 0.29$ and $\alpha = 0.26$

4. RESULTS

The purpose of this paper was to evaluate forecast accuracy by using data of primary production of electricity in Slovakia. It is therefore aimed to analyse SES forecasting method. To estimate optimal value of smoothing constant, forecasts are computed with $\alpha = 0.1$ to $\alpha = 0.9$, with increments of 0.01. Three forecasting accuracy techniques, such as *MAE*, *MAPE*, and *RMSE* are used to select the most accurate forecast for one year ahead forecast.

5. DISCUSSION/CONCLUSIONS

The exponential smoothing provides an idea that the most recent observations usually give the best guide to the future, therefore we want a weighting scheme with decreasing weights for older observations. The choice of the smoothing constant is important in determining the operating characteristics of exponential smoothing. The smaller the value of α, the slower the response. Larger values of α cause the smoothed value to react quickly – not only to real changes but also random fluctuations [11]. Simple exponential smoothing model is only good for non-seasonal patterns with approximately zero trend and for short-term forecasting because if we extend past the next period, the forecasted value for that period has to be used as a surrogate for the actual demand for any forecast past the next period. Consequently, there is no ability to add corrective information (the actual demand) and any error grows exponentially.

REFERENCES

[1] ACZEL, A. D.: *Complete Business Statistics*, Irwin, 1989, ISBN 0-256-05710-8.

[2] BROWN, R. G.: *Statistical Forecasting for Inventory Control*, McGraw-Hill: New York, 1959.

[3] BROWN, R. G. – MEYER, R. F.: *The Fundamental Theory of Exponential Smoothing*, Operations Research, 9, 673–685, 1961.

[4] CORDEIRO, C. – NEVES, M. M.: *Bootstrap and Exponential Smoothing Working together in Forecasting Time Series*, Proceedings in Computational Statistics, Physica-Verlag, 891–899, 2008.

[5] GELPER, S. – FRIED, R. – CROUX, CH.: *Robust Forecasting with Exponential and Holt Winters Smoothing*, Journal of forecasting, Vol. 29, No. 3, 285–300, 2010.

[6] HOLT, C. C.: *Forecasting trends and seasonals by exponentially weighted averages*, O.N.R. Memorandum 52/1957, Carnegie Institute of Technology, 1957.

[7] JARETT, J.: *Business Forecasting Methods*, Cambridge, MA: Basil Blackwell, 1991.

[8] LEWIS, C. D.: *Industrial and Business Forecasting Methods*, London, Butterworths, 1982.

[9] LI, Z. P. – YU, H. – LIU, Y. C. – LIU, F. Q.: *An Improved Adaptive Exponential Smoothing Model for Short Term Travel Time Forecasting of Urban Arterial Street*, Acta automatica sinica, Vol. 34, No. 11, 1404–1409, 2008.

[10] MAKRIDAKIS, S. – WHEELWRIGHT, S. C. – HYNDMAN, R. J.: *Forecasting Methods and Applications*, New York, Wiley, 1998.

[11] MONTGOMERY, D. C. – JOHNSON, L. A. – GARDINER, J. S.: *Forecasting and Time Series Analysis*, McGraw-Hill, Inc., 1990, ISBN 0-07-042858-1.

[12] OSTERTAGOVÁ, E.: *Applied Statistics*, [in Slovak], Elfa, Košice, 2011, ISBN 978-80-8086-171-1.

[13] OSTERTAGOVÁ, E. – OSTERTAG, O.: *The Simple Exponential Smoothing Model*, Proceedings of the 4th International Conference on Modelling of Mechanical and Mechatronic Systems, Technical University of Košice, Slovak Republic, 380–384, 2011.

[14] RYU, K. – SANCHEZ, A.: *The Evaluation of Forecasting Methods at an Institutional Foodservice Dining Facility*, The Journal of Hospitality Financial Management, Vol. 11, No. 1, 2003.

[15] SANJOY, K. P.: *Determination of Exponential Smoothing Constant to Minimize Mean Square Error and Mean Absolute Deviation*, Global journal of research in engineering, Vol. 11, 2011.

[16] YORUCU, V.: *The Analysis of Forecasting Performance by Using Time Series Data for Two Mediterranean Island*, Review of Social, Economic & Business Studies, Vol. 2, 175–196, 2003.

[17] http://www.statistics.sk.

BIOGRAPHIES

Eva Ostertagová (PhDr, PhD) graduated (MSc) at the Faculty of Natural Sciences at the P. J. Šafárik University in Košice. She defended her PhD at the Faculty of Mechanical Engineering at Technical University in Košice. She is working as an assistant professor the applied mathematics at the Faculty of Electrical Engineering and Informatics, Technical University in Košice. Her scientific research is application of the mathematical modelling in automation and control in industry.

Oskar Ostertag (Doc, Ing, PhD) graduated (MSc) at the Faculty of Mechanical Engineering at Technical University in Košice. He defended his PhD at the Faculty of Mechanical Engineering at Technical University in Košice. He is working as an associated professor of the applied mechanics at the Faculty of Mechanical Engineering at Technical University in Košice. His scientific research is application of the analytical and numerical methods in construction parts designing.