ELEKTRYKA 2011
Zeszyt 3 (219) Rok LVII

Łukasz MALIŃSKI
Zakład Pomiarów i Systemów Sterowania, Politechnika Śląska w Gliwicach

AN IDENTIFICATION OF INDIVERTIBLE ELEMENTARY BILINEAR TIME-SERIES MODELS

Summary. One of the issues of identification of coefficient of elementary bilinear time-series model is indivertibility of the model due to displacement of global minimum of identification algorithm cost function. The paper shows that it is possible to change placement of this global minimum, so it is corresponding to real value of model coefficient.

Keywords: Identification, bilinear models, indivertible models, time-series

IDENTYFIKACJA NIEODWRACALNYCH ELEMENTARNYCH BILINIOWYCH MODELI CIĄGÓW CZASOWYCH

Streszczenie. Jednym z podstawowych problemów identyfikacji współczynnika elementarnego biliniowego modelu ciągu czasowego jest nieodwracalność modelu. Skutkuje ona przemieszczeniem minimum globalnego funkcji kosztu algorytmu identyfikacji, tak że odpowiada ono nieprawidłowej wartości współczynnika. Niniejszy artykuł przedstawia propozycję rozwiązania tego problemu oraz analizę jego powtarzalności.

Słowa kluczowe: identyfikacja, modele biliniowe, modele nieodwracalne, ciągi czasowe

1. INTRODUCTION

The purpose of this paper is to present the possible solution for identification of the coefficient of indivertible elementary bilinear time-series model (EB), which is simplest of all bilinear time-series models. Although, EB model has limited use in compare to more complex models (e.g. BARMA), the understanding of its properties and ability to correctly identifying it is the key to identification and understanding properties of more complex bilinear time-series models.

The EB time-series model was introduced by Granger and Andersen [1] in 1978 and further analysed by Tong [2] and Granger and Teräsvirta [3] in 1993. One year later Bielińska and Nabagło [11] proposed modification to the Recursive Least Square algorithm for EB models, which has improved parametric identification efficiency. Latter, in 1999, Martins [4], and Berlin Wu [5] have further extended knowledge of bilinear time-series model and in 2006 Bouzaachane, Harti and Benghbrit proposed use of minimum likelihood algorithm for chosen types of bilinear time-series models. Next year the extensive monograph [6] was written by Bielińska about properties of process obtained for EB models, different identification approaches and prediction algorithms with application examples. Moreover, there is enough evidence to be found in work of Brunner and Hess [8] form 1995 and more recently (2010) in work of Bielińska and Maliński [8,9] that identification of coefficients of EB models is a difficult task. Further research performed by Figwer and Maliński shown that random processes obtained from EB models are nonstationary [10], which implicates non-ergodicity of these processes and speaks against use of identification algorithms based on dependencies between empirically obtained statistical moments of the process and the analytically deduced formulas of statistical moments for the model iteself.

In following pages, the definition of EB model, discussion of identification difficulties due to indivertibility of the model and finally solution to this problem will be presented.

2. THEORETICAL BACKGROUND

The most general form of bilinear time-series model is called BARMA(dA,dC,dK,dL) and it is defined by the following equation:

$$y(n) = \sum_{i=1}^{dA} a_i y(n-i) + \sum_{j=0}^{dC} c_j e(n-j) + \sum_{k=1}^{dK} \sum_{l=1}^{dL} \beta_{kl} e(n-k) y(n-l) \tag{1}$$

where: $y(n)$ is an output sequence, n is a discrete time indicator, $e(n)$ is an innovation sequence considered to be Gaussian white noise of properties defined in (2), a_i and c_j are coefficients of the linear part of model, β_{kl} are bilinear part coefficients, dA, dC, dK and dL are ranks of structure parameters.

$$\begin{gathered} \mathrm{E}\{e(n)\} = 0; \quad \mathrm{E}\{e(n)^2\} = \lambda^2; \quad \mathrm{E}\{e(n)e(n-1)\} = 0; \\ \mathrm{E}\{e(n)^3\} = 0; \quad \mathrm{E}\{e(n)^4\} = 3\lambda^4 \end{gathered} \tag{2}$$

The properties of bilinear time-series model, even in its pure form (without linear part), are hard to determine and identification of its coefficients is also difficult to perform. Therefore, in order to better understand the behaviour of bilinear models some simplifications are necessary to be taken. The simplest type of bilinear time-series model is elementary bilinear time-series model EB(k,l), which is defined below:

$$y(n) = e(n) + \beta_{kl} e(n-k) y(n-l) \tag{3}$$

This model can be classified in three structure types depending on *k* and *l* relation:

For $k < l$ the EB time-series model is called *superdiagonal*

For $k = l$ the EB time-series model is called *diagonal*

For $k >$ l the EB time-series model is called *subdiagonal*

The stability condition of the EB model is the same for structure types and is defined by following equation:

$$\beta_{kl}^2 \lambda^2 < 1 \tag{4}$$

where: λ^2 is a variance of the white noise *e(n)*.

The general invertibility condition for EB model is defined bellow:

$$\beta_{kl}^2 m'^{(2)}_y < 1 \tag{5}$$

where: $m'^{(2)}_y$ is a second central moment (variance) of the output sequence *y(n)*.

Because, different EB model structure types, stimulated by the white noise *e(n)* with variance equal to λ^2 have different output variances [6], the general invertibility condition (5) can be specified for each of EB model structure type by formulas in Table 1.

Table 1

Invertibility conditions for different EB model structures.

	Superdiagonal	Diagonal	Subdiagonal
Invertibility condition	$\beta_{kl}^2 \lambda^2 < 0.5$	$\beta_{kl}^2 \lambda^2 < 0.36$	$\beta_{kl}^2 \lambda^2 < \sqrt[k-l+2]{0.5}$

3. IDIVERTIBILITY ISSUE

In order to estimate the value of EB model coefficient β_{kl} an identification algorithm has to be chosen. The most common approach to identification of discrete time-series models is minimisation of Mean Square Error between model output and time-series real values (MSE). Therefore, describing the value of error between model output $\hat{y}(n)$ and real time-series value *y(n)* (also called as prediction error) as follows:

$$\varepsilon(n) = y(n) - \hat{y}(n) = y(n) - \beta_{kl}\hat{e}(n-k)y(n-l) \tag{6}$$

the typical cost function (MSE) for identification algorithms can be defined as bellow:

$$J(\hat{\beta}_{kl}) = \frac{1}{N}\sum_{n=1}^{N}\varepsilon(n)^2 \tag{7}$$

The estimates of innovation signal $\hat{e}(n)$ in (6) are assumed to be equal to prediction error $\varepsilon(n)$. Therefore, the estimates of innovation signal $\hat{e}(n)$ are obtained from output sequence $y(n)$ which means use of inverted model. If the time-series has been obtained from indivertible model this $\hat{e}(n)$ sequence is the output of an unstable inverted model which means that it is not convergent to real innovation signal $e(n)$.

Because, the identification algorithm minimises MSE, the global minimum will be fund for wrong (explicitly lower) value of β_{kl}. Although, this is not considered as a bad solution, because this model can still recreate some statistical properties of process and can be used in prediction applications, from classification point of view this result can be misleading. Assuming that coefficient value corresponds to the magnitude of some phenomenon, it may not be possible to distinguish two clearly different in magnitude cases by the model coefficient value if one of them will corresponds to the value of coefficient of indivertible model.

This problem is visualised on following figures:

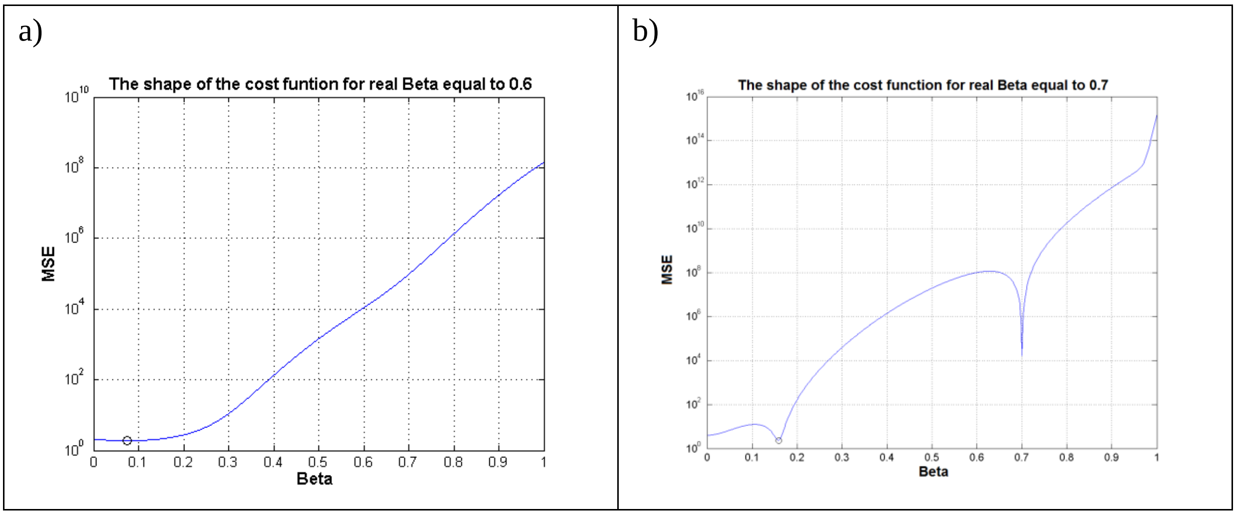


Fig. 1. Samples of the shape of the cost functions for indivertible models: a) $\beta_{kl} = 0.6$; b) $\beta_{kl} = 0.7$
Rys. 1. Przykłady kształtów funkcji kosztu dla przypadków nieodwracalnych: a) $\beta_{kl} = 0.6$; b) $\beta_{kl} = 0.7$

For simulated time-series obtained from indivertible diagonal EB model with coefficients $\beta_{kl} = 0.6$ (Figure 1a) and $\beta_{kl} = 0.7$ (Figure 1b) and with variance of $e(n)$ equal to 1, the shape of the cost function of identification algorithm (MSE) has been obtained for entire positive stability range with resolution of 10^{-4}. It is easy to observe that there is no global minimum in vicinity of the $\beta_{kl} = 0.6$ value (Fig. 1a) or there is only local minimum in place corresponding to real coefficient value $\beta_{kl} = 0.7$ (Fig. 1b). Moreover, global minimum is corresponding to β_{kl} values significantly lower than correct ones, which clearly provide with wrong identification result. In next chapter the possible solution to that problem is presented.

4. THE PROPOSED SOLUTION

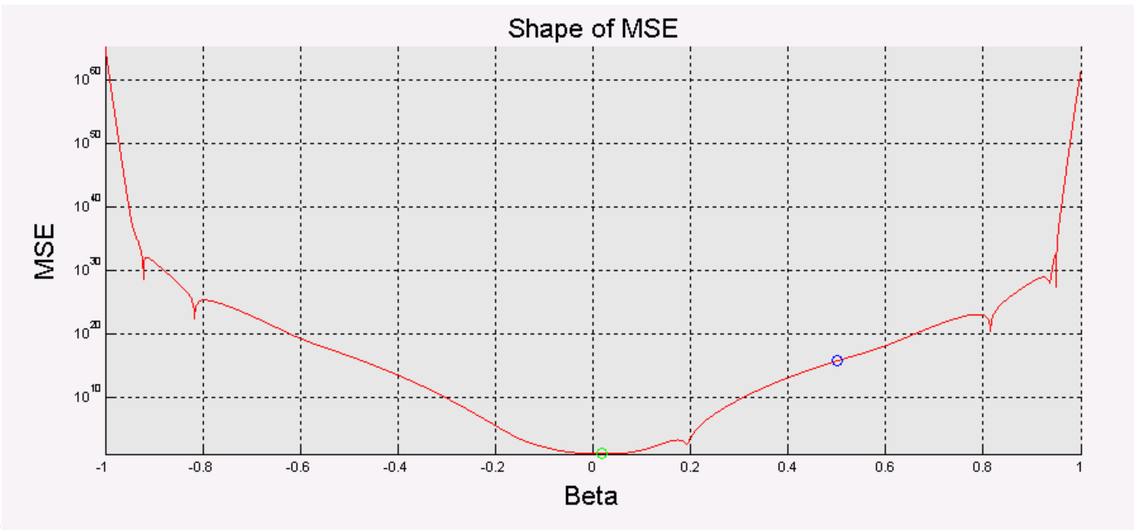
Let's consider equation (3) as a sum of two independent random variables. According to well known statistical theorem, the variance of the sum of two or more independent random variables is equal to sum of variances of these variables. It means that the variance of the output sequence *y(n)* is always larger then variance of the innovation signal *e(n)* used to generate the time-series.

Using this conclusion with assumption that prediction error *ε(n)* is an estimate of original innovation signal *e(n)* it is possible to state that the variance of the prediction error cannot be ever grater then the variance of the output sequence *y(n)*. Going further, assumption of Gaussian distribution of *e(n)* implicates that 99.7% of values of *e(n)* is within range of 3λ (where λ is standard deviation of *e(n)*). Although, standard deviation of *e(n)* is unknown, it is possible to estimate its value from above as S_y (standard deviation of output sequence). This way the range of possible $\hat{e}(n)$ values can be obtained (8):

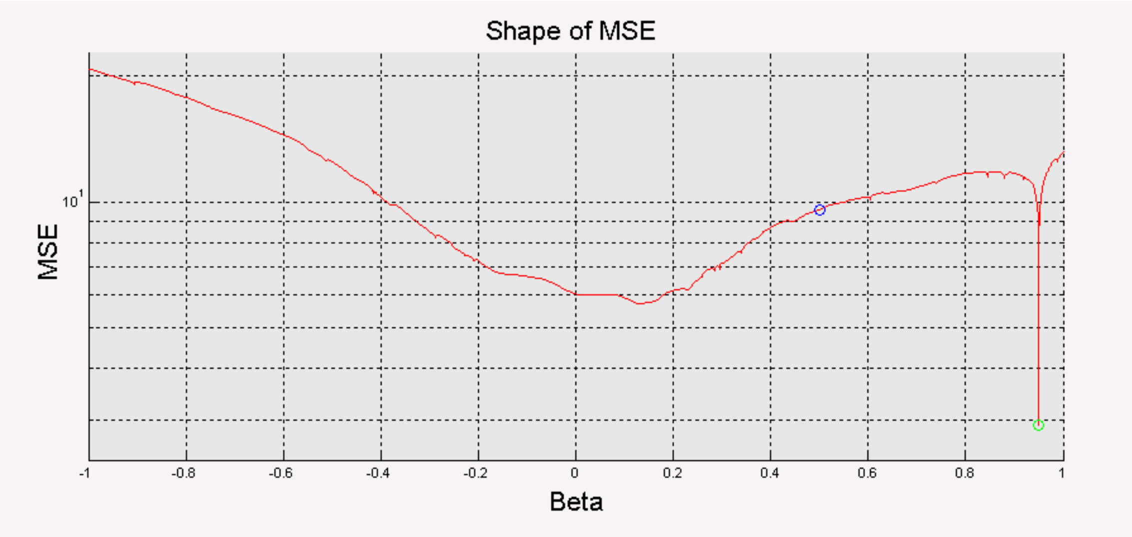
$$\hat{e}(n)=\begin{cases} w & \varepsilon(n)>w \\ \varepsilon(n) & -w\le\varepsilon(n)\le w; \quad for \quad w=qS_y \\ -w & \varepsilon(n)<-w \end{cases} \tag{8}$$

where: q is constant scaling factor (typically equal to 2 or 3), and S_y is the estimate of standard deviation of *y(n)*.

Proposed limit of $\hat{e}(n)$ which, was also used to stabilize RLS algorithm in [11], have a very interesting impact on the shape of the MSE function. On Figure 2, (a and b) the shapes of MSE function with and without limiting of $\hat{e}(n)$ are presented for clearly irreversible diagonal model of β_{kl} near stability threshold.



a)



b)

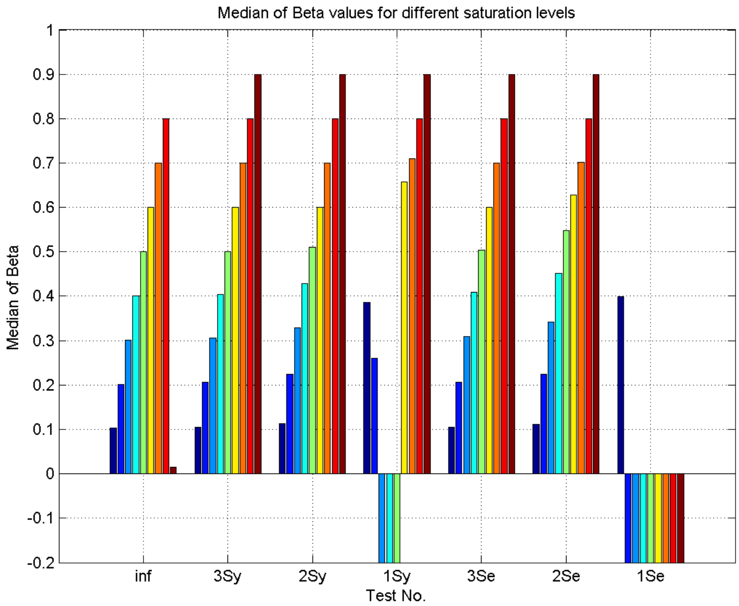
Rys. 2. Kształty funkcji kosztu dla tego samego ciągu czasowego: a) bez limitu $\hat{e}(n)$, b) z limitem $\hat{e}(n)$
Fig. 2. Shapes of the cost function for the same time-series: a) without $\hat{e}(n)$ limit, b)with $\hat{e}(n)$ limit

The time-series for which shapes of the MSE are presented on Figure 2 (a and b) has been obtained from EB model with β_{kl} = 0.95 and with variance of the white noise λ^2 = 1. The expected placement of the global minimum of the MSE should be in vicinity of argument equal to 0.95. The top chart (fig 2a) clearly denies that statement. In this case the global minimum of the MSE is near argument equal to 0, which means that identification result will be clearly wrong. Opposite situation can be observed on bottom chart (fig 2b) where limiting of q = 2 (8) has been applied. In this case explicit global minimum is observed exactly for argument equal to 0.95. Moreover, numerous empirical tests, presented in next chapter, shows that phenomenon of replacement of global minimum to right place after limiting of $\hat{e}(n)$ is repeatable.

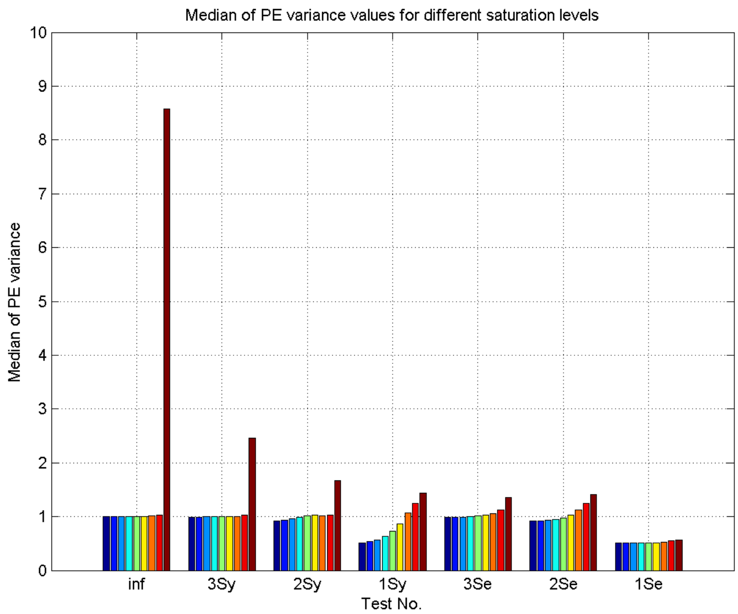
5. REPEATABILITY OF GLOBAL MINIMUM REPLACEMENT

In order to check repeatability of the global minimum replacement to correct position the large number of simulations have been performed. The set R_β of r = 200 realizations of random processes, obtained from superdiagonal EB(1,2) model, have been simulated for each coefficient value β_{kl} from set S:{0.1, 0.2, ... , 0.9} This way almost entire positive stability range has been covered, assuming $e(n)$ variance equal to λ^2 = 1. For each realization the global minimum has been found with grid-search algorithm for seven different values of $\hat{e}(n)$ limit w. First marked as “inf” level of limit w has been set equal to infinity (no limit) and is treated as reference case. Next three levels of limit w where based on standard deviation of output sequence $y(n)$ (which is known before identification in every experiment) as it is described in (8). The *q* values were equal to 1,2 and 3 and they have been marked accordingly as “1Sy”, “2Sy” and “3Sy”. Finally, last three levels of limit *w* have been based on standard deviation of *e(n)* sequence (which can be known only in simulations, but it is theoretically the best choice) also for *q* equal to 1,2 and 3 (marked as “1Se”, “2Se” and “3Se”).

The simulations of processes obtained from EB model have been performed for Gaussian distribution of *e(n)* sequence for 1200 samples, but only last 1000 samples have been taken into computation assuming, first 200 samples as the buffer to fade out initial conditions.



a)



b)

Fig. 3. Median values for each set T_β: a) for $\hat{\beta}_{kl}$ placement, b) for $J(\hat{\beta}_{kl})$ placement.
Rys. 3. Wartości median dla zbiorów realizacji T_β: a) dla położenia $\hat{\beta}_{kl}$, b) dla płożenia $J(\hat{\beta}_{kl})$

This way the each of R_β sets of time-series realizations contains 200 of simulated processes defined by the same EB model coefficient value. The corresponding sets T_β contain r = 200 estimated global minimum coordinates (represented by $\hat{\beta}_{kl}$ and $J(\hat{\beta}_{kl})$) obtained for chosen levels of w. Therefore, in order to obtain statistical properties of observed phenomenon, median values have been computed for $\hat{\beta}_{kl}$ and $\hat{e}(n)$ for each T_β set. The results obtained, are presented graphically on figure 3 (a and b).

On each chart seven sets of bars are shown. Each set of bars corresponds to different w level and each bar in particular set corresponds to different T_β set. For example first bar in each set corresponds to estimation results for processes obtained for EB(1,2) model witch coefficient value β_{kl} = 0.1.In an ideal situation for set $T_{0.1}$ all estimates $\hat{\beta}_{kl}$ should be equal to β_{kl} = 0.1 and all $J(\hat{\beta}_{kl})$ should be equal to $\lambda^2 = 1$. In reality it is expected, those estimates will be equal to its real values statistically, so if the median values from each T_β set are equal to original parameters the result is considered as satisfactory.

The analysis of results obtained and presented on figure 3 can be summarised as follows:

Estimates of β_{kl} values are satisfactory analysing the median values for w set to “2Sy” but clearly the best results have been obtained for w set to “3Se”.

Proper selection of w value also provides with good estimates of variance of $e(n)$ represented by $J(\hat{\beta}_{kl})$ (in all cases this variance should be equal to $\lambda^2 = 1$).

6. FINAL REMARKS

Presented results show that it may be possible to properly estimate the values of coefficient of indivertible EB model. At this point there are many further research to be done in order to check reliability of proposed solution. Because, invertibility of the model is unknown before identification this solution should be applied in all cases. Therefore, knowing that in case of invertible time-series with low values of original β_{kl} coefficient the value of variance of the output sequence $y(n)$ is very close to value of variance of the innovation signal $e(n)$, it is possible that application of proposed solution will damage the identification results due to too low value of w. Moreover, some improvement choosing the limit of $\hat{e}(n)$ level should be performed to optimise the proposed solution efficiency. At this point, it can be stated the proposed solution is very promising and it may be also possible to extend its application for different and more complex type of time series models.

BIBLIOGRAPHY

1. Andersen A., Granger C.: An Introduction to Bilinear Time Series Models. Vandenhoeck und Ruprecht 1978.
2. Tong H.: Non-linear time series, Clarendon Press. Oxford 1993.
3. Granger C., Teräsvirta T.: Modelling nonlinear Economic Relationships. Oxford University Press 1993.
4. Martins C.: A note on the third order moment structure of a bilinear model with non independent shocks. "Portugaliae Mathematica" 1999, Vol. 56.
5. Wu Berlin, Model-free forecasting for nonlinear time series: with example on ARCH and bilinear models. "Fuzzy Sets and Systems" 1999, Vol.108.
6. Bielińska E.: Bilinear time series models in signal analysis. Zeszyty Naukowe Politechniki Śląskiej, Gliwice 2007 (in Polish).
7. Brunner A., Hess G.: Potential problems in estimating bilinear time series. "Journal of Economics Dynamics & Control" Vol. 19, 1995.
8. Maliński Ł., Bielińska E.: Statistical Analysis of Minimum Prediction Error Variance in the Identification of a Simple Bilinear Time-Series Model. Advances in System Science, Academic Publishing House EXIT 2010, p. 183-188.
9. Maliński Ł.: An Analysis of Parameters Selection of the Recursive Least Squares Identification Method with Application to a Simple Bilinear Stochastic Model, Advances in System Science, Academic Publishing House EXIT 2010, p. 189-196.
10. Maliński Ł., Figwer J.: On Stationarity of bilinear time-series, The 16th International Conference on Methods and Models in Automation and Robotics, Międzyzdroje, Poland 2011.
11. Bielińska E., Nabagło I.: A modification of ELS algorithm for bilinear time-series model identification. Zeszyty Naukowe Politechniki Śląskiej "Automatyka" 1994 vol. 108.

Recenzent: Prof. dr hab. inż. Marian Pasko

Wpłynęło do Redakcji dnia 20 września 2011 r.

Mgr inż. Łukasz MALIŃSKI

Politechnika Śląska. Wydział Automatyki, Elektroniki i Informatyki
Zakład Pomiarów i Systemów Sterowania
ul. Akademicka 16; 44-100 GLIWICE

tel. (032) 237 19 04, e-mail: lukasz.malinski@polsl.pl