SELECTED ENGINEERING PROBLEMS
NUMBER 3
INSTITUTE OF ENGINEERING PROCESSES AUTOMATION
AND INTEGRATED MANUFACTURING SYSTEMS

Łukasz MALIŃSKI*

Institute of Automatic Control, Faculty of Automatic Control Electronics and Computer Science, Silesian University of Technology, Gliwice, Poland
*lukasz.malinski@polsl.pl

ON DIFFICULTIES IN IDENTIFICATION OF SIMPLE LINEAR-BILINEAR TIME-SERIES MODELS USING MEMETIC ALGORITHM

Abstract: The new approach to identification of linear-bilinear time-series models has been recently proposed. It is based on separated identification of linear and bilinear parts of the model and exploits the advantages of Memetic Algorithms Therefore, simple survey tests have been performed for different sets of time-series and some difficulties have been recognized. The results of the tests and possible explanations of problem are presented on following pages.

1. Introduction

The identification of time-series models and analysis of their properties is still dominated by linear models. They are very useful for prediction, analysis and classification and they are also easy to identify. Therefore, up to now the nonlinear modelling of time-series is significantly less explored and definitely more challenging. The first research in this area was performed by Granger and Andersen [1]. They have also the pioneers in theory of bilinear time-series models [2] which, are the subject of this paper and also are one of the simplest variants of nonlinear time-series models.

The bilinear time-series models were next concerned by Subba Rao [3] and Quinn. Leter, the Method of Lest Squares for estimation of diagonal variant of bilinear time-series model was proposed by Pham and Guegan. Moreover, Gooijger and Heuts [6] performed a statistical analysis of higher order moments of certain bilinear models which was a foundation of Method of Moments [11]. The stability condition for certain bilinear models has been proposed by Lee and Mathews [7]. Close to this time, Bielińska and Nabagło proposed a modification to a Least Squares (LS) method [8], which introduced the concept of limiting estimates of prediction error in identification of nonlinear time-series.

A general bilinear time-series model (BARMA) is very complex and proper identification of its coefficients is troublesome task. Therefore, many authors consider simplified variants, like the elementary linear-bilinear time series model (LEB). This model still requires many uncommon approaches in order to obtain unbiased estimates during identification of its coefficients.

Some analysis of identification difficulties has been presented by Brunner and Hess [9]. They surveyed the cost function of maximum likelihood algorithm and uncovered its complex and multimodal shape. The solution to this problem was proposed inter alia by Maliński, who used a evolutionary algorithm [12] to overcame the problem of multimodality of the Mean Square Error (MSE) cost function and proposed the procedure of estimating coefficients of all stable elementary bilinear time-series [13,15].

Basing on those accomplishments and concept of separated identification, proposed by Wang [10], the adaptive Memetic Algorithm has been designed [17]. In paper, survey tests, with use of multiple generated time-series, are presented. Their purpose is to check efficiency of the algorithm and point out possible problems.

2. Theoretical background

The BARMA(*dA*,*dC*,*dK*,*dL*) model [13] is defined below:

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

where: $y(t)$ is a discrete output signal, t is a discrete time indicator, coefficients a_i and c_j determine linear part of the model and β_{kl} are coefficients of the bilinear part. Parameters *dA*, *d*C, *dK* and *dL* describe the structure of the model and innovation signal $e(t)$ is assumed to be independent, identically distributed white noise.

The above model is to complex for analysis and there are also numerous problems to be found in attempts of identification of it. Therefore, simplified structures [4,8-12] are typically considered. Further, the elementary linear-bilinear LEB(*m*,*k*,*l*) model, defined in (2), will be taken under the consideration.

$$y(t) = e(t) + \alpha y(i-m) + \beta e(t-k) y(t-l) \tag{2}$$

Assuming following statistical properties of the innovation signal $e(t)$:

$$\begin{aligned} &E\{e(t)\} = 0; \quad E\{e(t)^2\} = \lambda^2; \\ &E\{e(t)e(t-1)\} = 0; \quad E\{e(t)^3\} = 0; \end{aligned} \tag{3}$$

The stability condition of bilinear part of the LEB(*m*,*k*,*l*) model can be defined [11]:

$$\beta^2 \lambda^2 < 1 \tag{4}$$

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

The stability of the linear part of the LEB(*m*,*k*,*l*) model is [14]:

$$|\alpha| < 1 \tag{5}$$

The linear part of the LEB(m,k,l) model can be identified with commonly used Recursive Least Squares (RLS) algorithm. However, according to [17] and [12] identification of the bilinear part requires more advanced approach capable of performing the optimization in multimodal solution space. This can be solved using adaptive Memetic Algorithm which was recently proposed in [17]. This solution already address the second major problem of bilinear part identification which is model indivertibility addressed in details in [13].

3. Identification algorithm

As it has been mentioned in previous section, the concerned identification algorithm is described in [17], thus only its general principles are adducted here and the main focus of the paper will is dedicated to performance survey, presented in next section.

The identification algorithm is implemented according to idea presented in Figure 1.

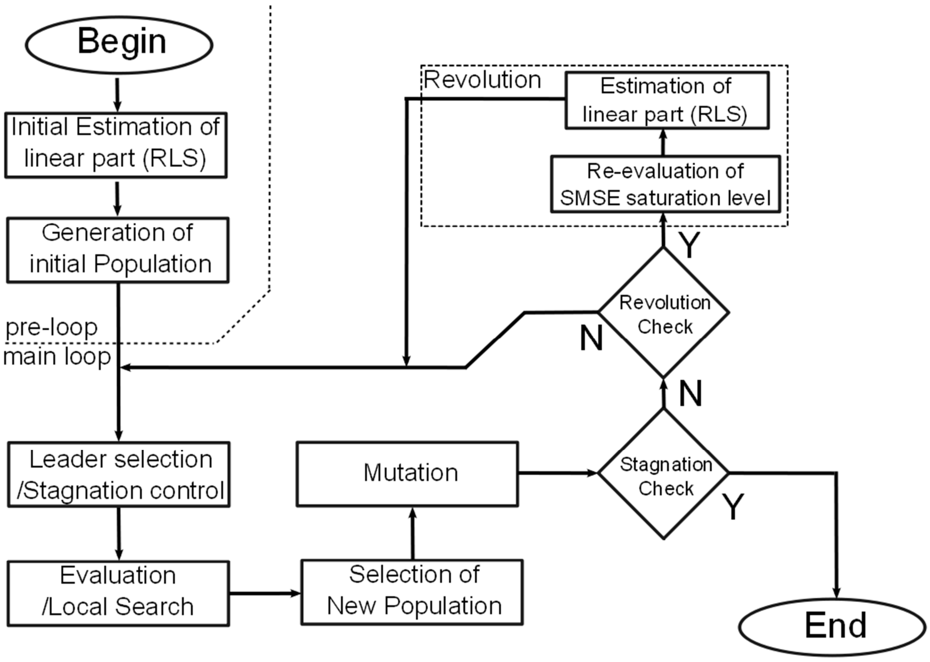

Fig.1. The Memetic Algorithm design

The cost function used in this optimization algorithm is Saturated Mean Square Error function [13] defined by equation (7) with support of (8) and (9):

$$J(w,y,\hat{\beta},\hat{\alpha},k,l)=\frac{1}{N}\sum_{t=1}^{N}\hat{e}(t)^2 \tag{7}$$

$$\varepsilon(t)=y(t)-\hat{\alpha}y(t-m)-\hat{\beta}\hat{e}(t-k)y(t-l) \tag{8}$$

$$\hat{e}(t)=\begin{cases} w & for \quad \varepsilon(t)\le w \\ \varepsilon(t) & for \quad -w<\varepsilon(t)<w \\ -w & for \quad \varepsilon(t)<-w \end{cases} \tag{9}$$

A similar saturation function (9) is also applied during computation of prediction error in commonly known RLS algorithm [14] used in estimation of linear part of LEB(m,k,l) model.

4. Survey results

The survey test have been prepared as follows:

- The set of testing values have been selected to be S = {0.2, 0.5, 0.8} for both model coefficients.
- R = 10 realizations of time-series generated using LEB(1,1,1) model have been obtained for each possible combination of coefficients values from set S. The variance of the white noise used as innovation signal was unary an each particular time-series consisted of N = 1000 samples. This way a set T of 90 testing time-series has been obtained.
- For each time-series from the set T the identification procedure using proposed algorithm [17] has been performed. Algorithm parameters were also set accordingly [17]. The Structure of the model has been assumed to be known.
- The identification results: estimates of model coefficients ($\hat{\alpha}, \hat{\beta}$), estimate of innovation signal ($\hat{\lambda}^2$), number of iterations (n) of the algorithm and final evaluation of saturation level (w), have been obtained and are presented in Tables 1-9.
- The mean value, standard deviation and median of results for each set of R time-series realizations with the same coefficient values have been computed and also presented in corresponding tables. Statistics for number of iterations (n) have been rounded up to maintain their natural meaning and their real values are presented in brackets.
- Some interesting results have been highlighted (bold font).

Tab. 1. Results for low values of both coefficients

Parameters	α	β	λ^2	n	w
Original values	0,2000	0,2000	1,0000	-	3,0000
1	0,2168	0,1842	1,0011	20	3,0017
2	0,1657	0,2050	0,9432	11	2,6843
3	0,1875	0,2118	0,9862	12	2,9462
4	0,2278	0,2271	1,0084	27	3,0126
5	0,1491	0,2308	0,9818	11	2,9398
6	0,1584	0,1975	0,9257	22	2,8863
7	0,2577	0,2522	0,9751	10	2,7814
8	0,1724	0,2152	0,9514	11	2,8705
9	0,2550	0,2114	0,9900	19	2,9742
10	0,2045	0,1884	0,8878	39	2,8267
Mean value	0,1995	0,2124	0,9651	18 (18,20)	2,8924
Std. Deviation	0,0392	0,0205	0,0377	9 (9,37)	0,1047
Median	0,1960	0,2116	0,9784	16 (15,50)	2,9131

Tab.2. Results for medium value of α and low value of β

Parameters	α	β	λ^2	n	w
Original values	0,5000	0,2000	1,0000	-	3,0000
1	0,5332	0,1989	0,9773	33	2,9657
2	0,5163	0,2124	0,9315	8	2,8311
3	0,4796	0,1724	0,9044	17	2,8415
4	0,5272	0,2136	0,9842	25	2,9760
5	0,4772	0,1886	0,9245	15	2,7856
6	0,4923	0,2207	0,9911	15	2,9749
7	0,5047	0,2377	0,9319	18	2,8968
8	0,4962	0,2001	0,9257	13	2,8866
9	0,5533	0,1829	0,9957	13	2,9936
10	0,4635	0,2047	0,9718	7	2,8090
Mean value	0,5044	0,2032	0,9538	16 (16,40)	2,8961
Std. Deviation	0,0281	0,0191	0,0334	8 (7,73)	0,0775
Median	0,5005	0,2024	0,9519	15 (15,00)	2,8917

Tab.3. Results for high value of α and low value of β

Parameters	α	β	λ^2	n	w
Original values	0,8000	0,2000	1,0000	nd	3,0000
1	0,7969	0,2004	0,9637	13	2,9745
2	0,8068	0,2149	0,9722	11	2,9569
3	0,7828	0,1893	0,9865	7	3,0064
4	0,7875	0,2173	0,9746	7	3,0003
5	0,8064	0,2371	0,9199	30	2,8770
6	0,7763	0,2115	0,9426	9	2,9520
7	0,7667	0,1772	0,9644	9	2,9646
8	0,7650	0,2036	1,0113	15	3,0169
9	0,8223	0,1843	0,9414	15	2,9114
10	0,8392	0,1942	0,9768	19	2,9649
Mean value	0,7950	0,2030	0,9653	14 (13,50)	2,9625
Std. Deviation	0,0241	0,0179	0,0258	7 (6,98)	0,0429
Median	0,7922	0,2020	0,9683	12 (12,00)	2,9648

Tab. 4. Results for low value of α and medium value of β

Parameters	α	β	λ^2	n	w
Original values	0,2000	0,5000	1,0000	-	3,0000
1	0,2159	0,5219	0,9812	13	2,9836
2	0,1806	0,4889	0,9821	21	2,9725
3	0,1958	0,5203	0,9685	11	2,9458
4	0,1917	0,4956	1,0246	11	3,0360
5	0,1915	0,5104	1,0126	16	2,9814
6	0,2657	0,4838	0,9540	7	2,9022
7	0,2322	0,5141	1,0182	20	3,0340
8	0,1905	0,5287	0,9574	17	2,9358
9	0,2439	0,5184	0,9559	14	2,9352
10	0,1039	0,5200	0,9765	20	2,9627
Mean value	0,2012	0,5102	0,9831	15 (15,00)	2,9689
Std. Deviation	0,0439	0,0154	0,0265	5 (4,62)	0,0427
Median	0,1937	0,5163	0,9789	15 (15,00)	2,9676

Tab. 5. Results for medium value of both coefficients

Parameters	*α*	*β*	λ^2	*n*	*w*
Original values	0,5000	0,5000	1,0000	-	3,0000
1	0,5221	0,5084	1,0030	13	3,0224
2	0,5617	0,4460	0,9912	21	2,9865
3	0,4112	0,5381	1,0108	24	3,0623
4	0,5224	0,4899	0,9916	20	2,9849
5	0,5474	0,4420	1,0181	9	3,0904
6	0,4920	0,5138	1,0103	12	3,0037
7	0,5478	0,4892	0,9670	17	2,9740
8	0,4990	0,5118	0,9532	18	2,9257
9	0,5547	0,4841	1,0137	18	3,0126
10	0,4738	0,5076	0,9505	16	2,9281
Mean value	0,5132	0,4931	0,9909	17 (16,80)	2,9991
Std. Deviation	0,0462	0,0302	0,0254	4 (4,49)	0,0522
Median	0,5223	0,4988	0,9973	18 (17,50)	2,9951

Tab. 6. Results for high value of α and medium value of β

Parameters	*α*	*β*	λ^2	*n*	*w*
Original values	0,8000	0,5000	1,0000	-	3,0000
1	**0,7815**	**0,4995**	**1,0993**	**28**	**3,1577**
2	0,7555	0,1386	1,6849	16	3,9233
3	**0,7029**	**0,4946**	**1,2022**	**16**	**3,1997**
4	**0,7045**	**0,5167**	**1,1113**	**24**	**3,0774**
5	0,9557	0,3523	1,1631	59	3,2448
6	0,8569	0,2840	1,4102	28	3,6373
7	0,8976	0,2735	1,3364	31	3,4697
8	0,8181	0,3361	1,2065	14	3,2920
9	**0,8202**	**0,4928**	**0,0627**	**20**	**3,1549**
10	**0,8102**	**0,4882**	**1,0105**	**52**	**3,0371**
Mean value	0,8103	0,3876	1,1287	29 (28,80)	3,3194
Std. Deviation	0,0800	0,1297	0,4211	15 (15,29)	0,2790
Median	0,8142	0,4203	1,1827	26 (26,00)	3,2223

Tab. 7. Results for low value of α and high value of β

Parameters	*α*	*β*	λ^2	*n*	*w*
Original values	0,2000	0,8000	1,0000	-	3,0000
1	0,1336	0,3384	1,7164	90	3,9168
2	0,2396	0,2967	1,7037	19	3,9340
3	**0,1799**	**0,7988**	**1,2212**	**46**	**3,1990**
4	0,0218	0,1916	2,2814	200	4,3456
5	0,3374	0,3181	1,8263	19	4,0836
6	0,3059	0,1403	2,0728	16	4,3375
7	0,3029	0,2707	1,9990	32	4,2586
8	-1,7686	-0,4529	4,3948	200	5,4018
9	-0,0670	0,0055	2,2536	15	4,5054
10	0,2289	0,3422	1,8807	7	4,2832
Mean value	-0,0086	0,2249	2,1350	64 (64,40)	4,2266
Std. Deviation	0,6317	0,3141	0,8512	75 (75,29)	0,5525
Median	0,2044	0,2837	1,9399	26 (25,50)	4,2709

Tab. 8. Results for medium value of α and high value of β

Parameters	α	β	λ^2	n	w
Original values	0,5000	0,8000	1,0000	-	3,0000
1	0,5193	0,1969	2,1879	172	4,4333
2	0,3837	0,2846	1,8038	47	4,0280
3	0,4789	0,2950	1,9823	35	4,2261
4	0,5763	0,0749	2,2105	29	4,4621
5	0,5657	0,2010	0,2118	27	4,3839
6	0,1337	0,1742	1,8786	200	4,1489
7	0,6255	0,2747	2,0107	20	4,2849
8	0,5033	0,0921	2,3422	11	4,6325
9	0,5640	0,1809	2,0715	20	4,3158
10	0,3907	0,3014	1,6956	33	3,9410
Mean value	0,4741	0,2076	1,8395	59 (59,40)	4,2857
Std. Deviation	0,1429	0,0815	0,6044	68 (67,75)	0,2085
Median	0,5113	0,1990	1,9965	31 (31,00)	4,3004

Tab. 9. Results for high values of both coefficients

Parameters	α	β	λ^2	n	w
Original values	0,8000	0,8000	1,0000	-	3,0000
1	0,9020	0,1390	2,4499	200	4,8448
2	0,9281	0,1162	2,8293	24	5,2102
3	0,8821	0,1148	2,2791	39	4,9582
4	0,7012	0,1983	2,9549	22	5,6182
5	1,6128	0,0900	4,5597	21	6,6345
6	0,7156	-0,0836	3,9696	47	5,9712
7	0,8651	0,0670	2,6967	22	5,0878
8	0,7644	0,0974	2,8456	13	5,3191
9	0,6093	-0,2150	2,8400	27	5,1305
10	0,7457	0,0235	4,5193	15	6,6941
Mean value	0,8726	0,0548	3,1944	43 (43,00)	5,5469
Std. Deviation	0,2795	0,1207	0,8365	56 (56,12)	0,6727
Median	0,8148	0,0937	2,8428	23 (23,00)	5,2647

The remarks can be summarised by the following:

- The estimates of coefficients obtained from identification of models for time-series with low original values of β are sufficiently accurate (see Tables 1-3). In these all cases the saturation level has been evaluated properly.
- The estimates of coefficients obtained for time-series with medium original values of β are mostly satisfactory (see Tables 4-6). The clearly biased results occurred in Table 6 only. They refer to time-series with high original value of α coefficient.
- Only one identification result for time-series obtained from LEB(1,1,1) model with high original values of β coefficient can be considered as satisfactory and it has been obtained for case with low original value of α coefficient (see Tables 7-9).
- The final remark is that all incorrect identification results has a common feature. In all this cases the saturation level for SMSE function has not been evaluated correctly. This way the proper placement of the global minimum of the solution space of identification task cannot be achieved.

5. **Possible sources of problem**

As it was shown in previous section some of the results obtained, especially for high coefficients values are biased. The direct cause of this problem seems to be an incorrectly evaluated saturation level during identification. Let's exam the one of the incorrectly identified models (first case from Table 7) in details:

- The history of changes of coefficients (α – top, β –bottom) values are presented in Figure 2.

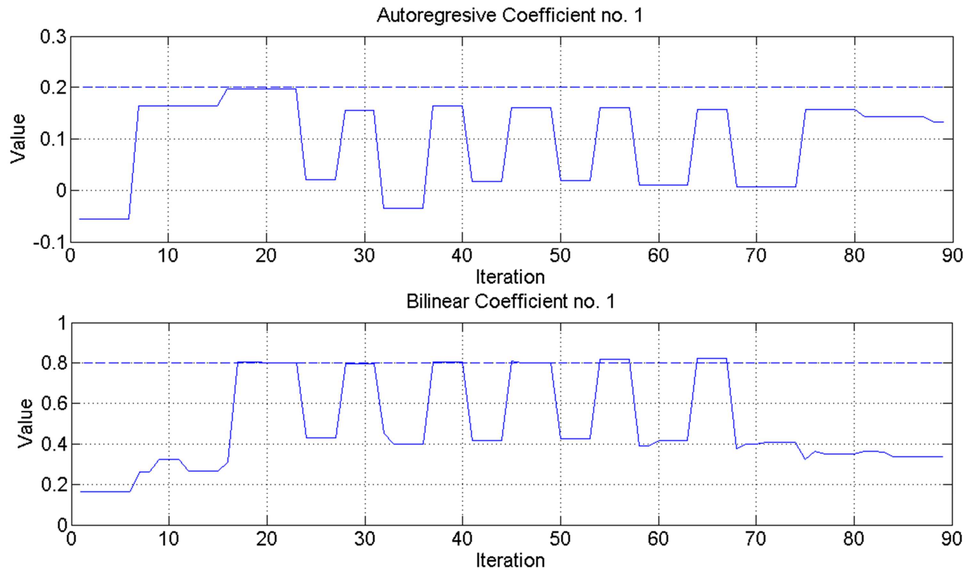

Fig.2. The History of changes of coefficient values

- Additionally the history of changes for estimate of variance of innovation signal (Min. of SMSE) and saturation level evaluation is presented in Figure 3.

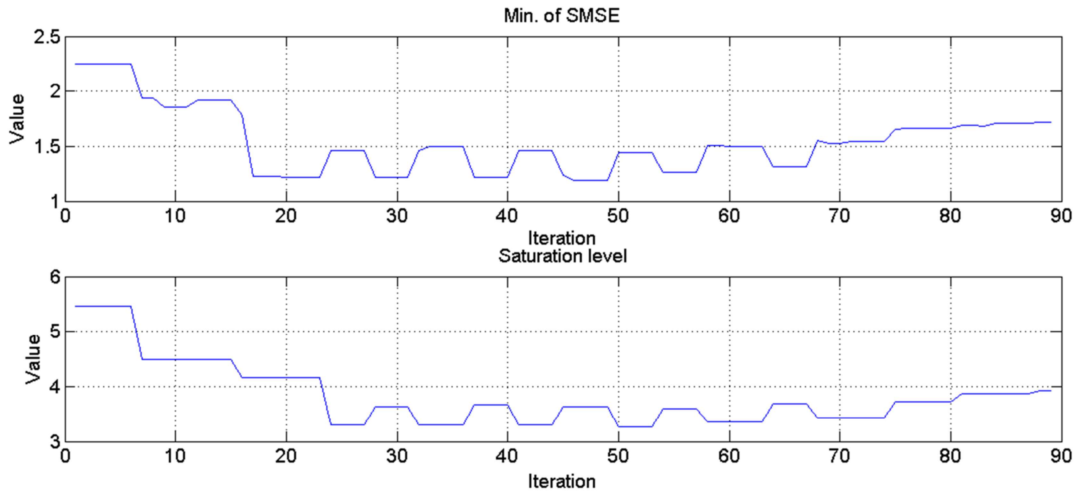

Fig.3. The History of changes of support identification results

- What can be concluded from figures above is that saturation level has a critical influence on identification results. However, it is not that simple that saturation level closer to desired value give better result. This particular case shows that sometimes larger value of saturation level can provide us with better coefficient estimates (iterations #15 - #25). Also change in estimates of coefficient values, even in a correct direction, not necessarily improve the saturation level evaluation.

- To support the thesis that incorrect saturation level is reason of the considered issue, another run of algorithm has been performed. This time saturation level was forced to be correct ($w = 3\lambda$ as proposed in [13,16,17]) The results are presented in Figure 4.

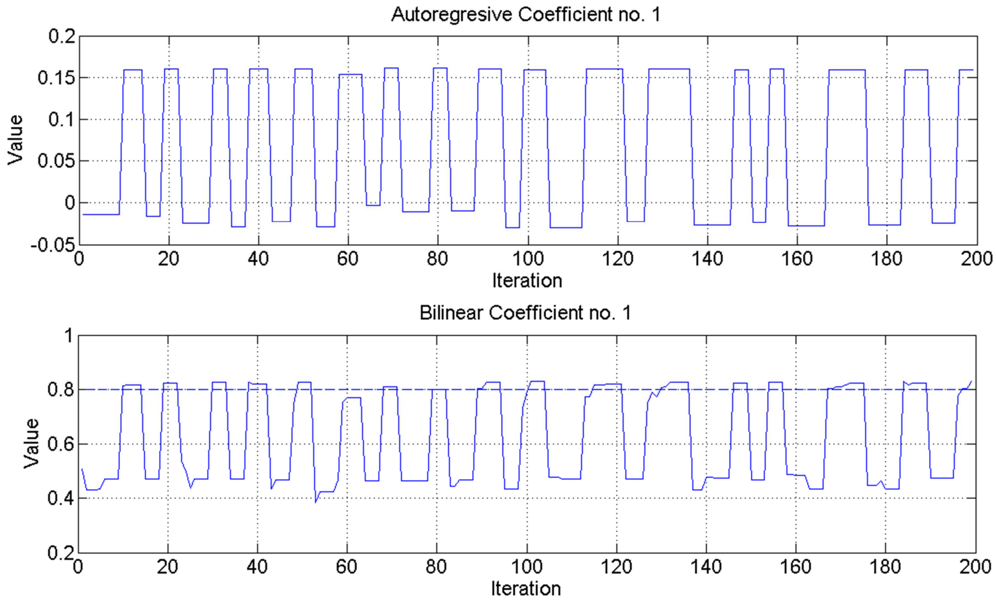

Fig.4. The History of changes of coefficients for constant, forced saturation level.

- These results (Fig. 4) clearly show that incorrect saturation level evaluation is not only reason of problems in identification.

6. Summary

Presented results shows that identification of the LEB model is not a simple task. Although, advanced and tested solutions have been applied, there is still hard to achieve a satisfactory effectiveness of algorithm especially for hard cases with high original values of model coefficients.

Up to now it is hard to point out all difficulties which have to be overtaken but clearly some light has been casted on the problem. Although, saturation level seems to have significant impact, on identification, the problem may be the separated identification itself because, a change in one coefficient estimate leads to the change in shape of solution space for other one and vice versa.

References

1. C., Granger, A. Andersen. Nonlinear time series modelling Applied Time series analysis, Academic Press, 1978.
2. C. Granger, A. Andersen. An introduction to bilinear time series models. Vandenhoeck and Ruprecht, 1978.
3. T. Subba Rao. On the Theory of Bilinear Time Series Models. Journal of the Royal Statistical Society, vol. B44, pp. 244-255, 1981.
4. B. Quinn. Stationarity and invertibility of simple bilinear models. Stochastic Processes and Their Applications, vol. 12, pp. 225-230, 1982.
5. D. Guegan, D. T. Pham. A Note on the Estimation of the Parameters of the Diagonal Bilinear Model by Method of Least Squares. Scandinavian Journal of Statistics, vol. 16, pp. 129-136, 1989.
6. J. Gooijger, R. Heuts. Higher order moments of bilinear time series processes with symmetrically distributed errors. Proceedings to Second International Tempere Conference in Statistics, pp. 467-478, 1987.
7. J. Lee, J. Mathews. A Stability Condition For Certain Bilinear Systems. Signal Processing, vol. 42, pp. 1871-1973, 1994.
8. E. Bielińska, I. Nabagło. A modification of ELS algorithm for bilinear time-series model identification. Zeszyty Naukowe Politechniki Śląskiej: Automatyka, vol. 108, pp. 7-24, 1994 (in Polish).
9. A. Brunner, G. Hess. Potential problems in estimating bilinear time-series models. Journal of Economic Dynamics and Control, vol. 19, pp. 663-681, 1995.
10. H. Wang. Parameter Estimation and Subset Detection For Separable Lower Triangular Bilinear Models. Journal of Time Series Analysis, vol. 26, pp. 743-757, 2004.
11. E Bielińska. Bilinear time series models in signal analysis. *Zeszyty Naukowe Politechniki Śląskiej*, 2007 (in Polish).
12. Ł. Maliński. An Identification Procedure For Elementary Bilinear Time-series Model Based on the Evolutionary Programming. Forum Innowacji Młodych Badaczy (FIMB) – II Ogólnopolskie Seminarium, Łódź, 2011.
13. Ł. Maliński. On identification of coefficient of indivertible elementary bilinear time-series model. Proceedings XIV Symposium: Fundamental Problem Of Power Electronics Electromechanics and Mechatronics PPEEm, 2011, Wisła, Poland.
14. E. Bielińska, "Prognozowanie ciągów czasowych", Wydawnictwo Politechniki Śląskiej, Gliwice 2007 (In Polish).
15. Ł. Maliński. 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, str. 189-196.
16. Ł. Maliński. The Evaluation of Saturation Level for SMSE Cost Function in Identification of Elementary Bilinear Time-Series Model. 17 International Conference on Methods and Models in Automation and Robotics, Międzyzdroje, Poland, 2012
17. Ł. Maliński. Memetic Algorithm for identification of linear-bilinear time-series model. XIV International PhD Workshop, Wisła 2012.