Short-term time series forecasting based on internal smoothing of Padé interpolants

Minvydas Ragulskis*, Kristina Lukoseviciute, Tadas Telksnys, and Zenonas Navickas

Research Group for Mathematical and Numerical Analysis of Dynamical Systems,
Kaunas University of Technology
Studentu 50-147, Kaunas LT-51368, Lithuania
minvydas.ragulskis@ktu.lt,kristina.lukoseviciute@ktu.lt,
tadas.telksnys@ktu.lt,zenonas.navickas@ktu.lt
http://www.minvydasragulskis.com/

Abstract. Short-term time series forecasting technique based on Padé interpolants and adaptive internal smoothing is presented in this paper. Adaptive corrections of time series data in the window of observation allows to construct near-optimal Padé extrapolant. Computational experiments with real world time series are used to demonstrate the efficiency of the proposed approach.

Keywords: time series, forecasting, Padé interpolant, internal smoothing

1 Introduction

Time series forecasting is and important technique used in a large variety of applications in different areas of science, engineering, finance and economics in general. The basic idea of any time series prediction algorithm is to identify a mathematical model generating the analyzed series and project this model into the future. Many different time series forecasting models and techniques have been developed during the recent decades. Conditionally, there methods can be classified into long-term and short term time series forecasting algorithms [1].

Time prediction horizon correlates with this classification – usually only short predictions suffice for short-term time series. It is true that predictors with even one time step forward horizons are important in a variety of applications [1]. Such techniques are widely used in finance [2–4]; electricity demand and the associated price forecasting problem [5–7]; wind power; passenger demand [8] and many others.

One time step forward prediction algorithms are usually based on the extrapolation of the available data. It is well known that Padé interpolants can be used for generating mathematical models of complex nonlinear processes [9].

* Corresponding author.

Padé functions, defined as ratios of univariate polynomials of, in general, different orders, have classically been used to approximate smooth functions with known Taylor series [10]. Padé functions have also been applied in split-step approximations of the solutions to differential equations [11–13], approximation of elliptic-type functions [14], generalized Euler transforms [15] and cosmographic analysis [16].

Significant attention has been recently devoted to Padé interpolation schemes. Padé-type rational and barycentric interpolation is considered in [9]. A rational interpolation scheme with a superpolynomial rate of convergence that reduces the Runge effect and can be used on discontinuous functions has been developed in [17]. A robust and efficient implementation of the rational interpolation scheme can be found in [18].

Padé approximants have also been applied to time series analysis and forecasting, primarily as a tool for the construction of ARMA models. In [19], the Padé approximation is used to accomplish the LS identification of an unstable ARMA equations. A method to identify the order of an ARMA time series model and to compute its coefficient based on the Padé approximant is presented in [20]. These methods have been applied to real-life time series in the field of economics [21, 22].

The main objective of this article is to present a new application of Padé-type methods in short-term time series forecasting. This paper is organized as follows: internal smoothing of algebraic is discussed in Section 2; an overview of Padé interpolants in presented in Section 3; the pre-processing algorithm of time series data is given in Section 4; the fitness function construction is discussed in Section 5; computational experiments are discussed in Sections 6 and 7; concluding remarks are given in the last section.

2 Internal smoothing of algebraic interpolants – preliminaries

Internal smoothing of an algebraic interpolant has been introduced in [23]. The main idea of this smoothing procedure is based on a projection of the reconstructed algebraic model into the future. However, instead of trying to make a straightforward projection of the model, a conciliation between the variability of the algebraic interpolant and the smoothness of moving average time series estimates is considered. We will use the standard industrial moving average algorithm to smooth the time series:

$$\mathrm{M}A_t = \frac{1}{s}\sum_{j=0}^{s-1} x_{t-j-1}, \tag{1}$$

where $\mathrm{M}A_t$ is a smooth value at time moment t; s is the averaging window. In general, the width of the averaging window should be preselected for each time series is not related to the length of the observation window used for the algebraic interpolant.

Now, time series elements in the observation window are individually perturbed by corrections $\varepsilon_1, \ldots, \varepsilon_M$. It is clear that additional constraints for the corrections are required in order to make this extrapolation problem well-posed. A fitness functions for the set of corrections $\varepsilon_1, \ldots, \varepsilon_M$ can be maximized in order to reconstruct a near-optimal algebraic skeleton representing the underlying dynamics in the observation window [23]:

$$F(\varepsilon_1, \ldots, \varepsilon_M) = \frac{1}{\alpha \sum_{j=1}^{M} |\varepsilon_j| + |\widetilde{x}_t - \mathrm{MA}_t|}, \tag{2}$$

where $\widetilde{x}_t$ is an exact algebraic extrapolant constructed over the perturbed elements of the time series in the observation window; the parameter $\alpha > 0$ determines the penalty proportion between the sum of corrections and the difference of forecast produced by algebraic extrapolant and moving average. It is clear that the target function would be unbounded at all zero corrections if algebraic extrapolant constructed over non-perturbed elements of the time series would coincide to moving average prediction.

The objective of this paper is to employ Padé interpolants for the algebraic prediction of the time series evolution and to enhance the intelligent perturbation of the analyzed time series.

3 Discrete Padé approximation scheme

Traditionally, Padé approximations are used to approximate a smooth function with a Taylor series expression by means of a rational function. In this paper, we apply the Padé approximant to time series data.

Suppose the time series $(t_1, x_1), \ldots, (t_M, x_M)$ is given; t_k denotes the time variable and x_k denotes the measurement taken at time t_k. Note that for time series where the length of the time interval is unknown, it can be taken that $t_k = k, k = 1, \ldots, M$ with no impact on the approximation.

The order (m, n) Padé function reads:

$$[m/n]_x(t) := \frac{\sum_{j=0}^{m} a_j t^j}{1 + \sum_{j=1}^{n} b_j t^j}; \quad a_j, b_j \in \mathbb{R}. \tag{3}$$

It is recommended to select $m \geq n$ [10]. The function (3) applied to the time series data yields the following system of linear equations with respect to the parameters $a_0, \ldots, a_m, b_1, \ldots, b_n$:

$$[m/n]_x(t_k) = x_k; \quad k = 1, \ldots, M. \tag{4}$$

Inserting (3) into (4) and simplifying yields:

$$\sum_{j=0}^{m} a_j t_k^j - x_k \sum_{l=1}^{n} b_l t_k^l = x_k; \quad k = 1, \ldots, M. \tag{5}$$

The system (5) can be rewritten in matrix form:

$$\mathbf{W}\mathbf{p} = \mathbf{x}, \tag{6}$$

where

$$\mathbf{W} = \begin{bmatrix} 1 & t_1 & \ldots & t_1^m & -x_1 t_1 & \ldots & -x_1 t_1^n \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_M & \ldots & t_M^m & -x_M t_M & \ldots & -x_M t_M^n \end{bmatrix}; \tag{7}$$

and

$$\mathbf{p} = \begin{bmatrix} a_0 \ldots a_m \; b_1 \ldots b_n \end{bmatrix}^{\mathrm{T}}; \quad \mathbf{x} = \begin{bmatrix} x_1 \ldots x_M \end{bmatrix}^{\mathrm{T}}. \tag{8}$$

Let us denote $N = \dim \mathbf{p} = m + n + 1$ the total number of parameters in (3). Noting that $\mathbf{W}$ is nonsingular if $t_k \neq t_l, x_k \neq x_l$; $k \neq l$ (which is always satisfied for time series, since $t_k < t_l$ for $k < l$) yields the unique least-squares solution to (6) for $N \leq M$:

$$\mathbf{p} = \left(\mathbf{W}^{\mathrm{T}}\mathbf{W}\right)^{-1} \mathbf{W}^{\mathrm{T}}\mathbf{x}. \tag{9}$$

Note that for $N = M$, the interpolant is obtained. However, this is impractical for time series analysis, because of the large number of parameters required and the negative impact of the Runge effect [24].

4 Pre-processing algorithm

The pre-processing algorithm of the time series data is given below. This algorithm normalizes the time series and selects optimal parameters for the Padé extrapolant using full sort.

Algorithm 1: Time series pre-processing

Input : $x_1, \ldots, x_M$ – time series;
$\overline{M}$ – maximum number parameters in Padé function (3);
L – number of time-forward steps for RMSE evaluation.

Output: $\widetilde{x}_1, \ldots, \widetilde{x}_M$ – time series normalized to the range of $[-1, 1]$;
(m_*, n_*) – parameters of optimal Padé function for given time series.

Normalize time series:
for $k = 1, \ldots, M$ **do**
$$\widetilde{x}_k = \frac{2x_k - \max_{1 \le l \le M} x_l - \min_{1 \le l \le M} x_l}{\max_{1 \le l \le M} x_l - \min_{1 \le l \le M} x_l};$$
end

Minimize $RMSE\,(m,n)$ **using full sort:**
for $j = 2, \ldots, \overline{M}$ **do**
  for $N = 2, \ldots, M$ **do**
    for $m = \lfloor \frac{N}{2} \rfloor, \ldots, N-1$ **do**
      $n = N - m - 1$;
      form $\mathbf{W}, \mathbf{x}$ with n, m and $x_{\overline{M}-j+1}, \ldots, x_{\overline{M}}$;
      obtain parameters: $\mathbf{p} = \left(\mathbf{W}^{\mathrm{T}}\mathbf{W}\right)^{-1} \mathbf{W}^{\mathrm{T}}\mathbf{x}$;
      compute Padé forecast for L forwards steps: $\widehat{x}_{\overline{M}+1}, \ldots, \widehat{x}_{\overline{M}+L}$;
      compute error: $RMSE\,(m,n) = \dfrac{1}{L}\sqrt{\sum_{k=1}^{L} \left(x_{\overline{M}+k} - \widehat{x}_{\overline{M}+k}\right)^2}$;
    end
  end
end
Choose parameters with smallest error: $(m_*, n_*) = \arg\min_{m,n} RMSE\,(m,n)$.

5 The construction of the fitness function

An evolutionary strategy is used in [25] to identify the algebraic skeleton sequence in the observation window of the predicted time series by removing the unknown additive noise. The idea is based on the assumption that the time series comprises some sort of deterministic skeleton describing the dynamics of the time series which is contaminated by the additive noise.

Let us denote $\tilde{x}_k = x_k - \varepsilon_k; k = 1,2,3,\ldots$ as the corrected values of the sequence (ε_k are unknown corrections). The F-measure as in [26] becomes the fitness measure of a genetic algorithm that identifies predictive patterns in the sequence of events [27].

The F-measure consists of two parts that embody different objectives: PRECISION, the model precision, requires from the model that it faithfully recon-

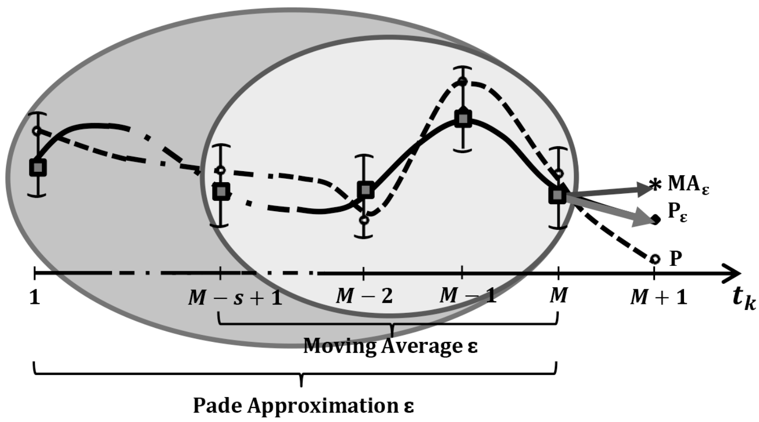

Fig. 1. Schematic diagram illustrating the proposed Padé method, where dots denote the original time series; squares denote the corrected time series; P^{ε} the Padé forecasting for corrected time series; P forecasting for original time series.

struct the last known time series values and RECALL requires that the prediction repeats past dynamical behaviour:

$$F\left(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M\right) = \frac{\left(\gamma^2+1\right) \cdot PRECISION \cdot RECALL}{\gamma^2 \cdot PRECISION + RECALL}. \tag{10}$$

In equation (10) the value γ controls the relative importance of precision to recall. If $\gamma = 0$ then the fitness function evaluates only the PRECISION part. If $\gamma = \infty$ then the fitness function evaluates the RECALL values only. In our case we build PRECISION and RECALL functions in such a way that the minimal value of the fitness function is reached when the corrections are small and the improved Bernstein extrapolation (through points $\tilde{x}_k$) is close to the moving average prediction (also based on $\tilde{x}_k$):

$$PRECISION = \frac{1}{M-1} \sum_{i=1}^{M} \left| x_i - \hat{x}_i \right|, \tag{11}$$

$$RECALL = \alpha \sum_{i=1}^{M} \left| \varepsilon_i \right| + \beta \left| \mathrm{MA}_{M+1} - P^{\varepsilon}_{M+1} \right|; \alpha > 0, \beta > 0, \tag{12}$$

where an array ε_0,ε_1,...,ε_M represents near-optimal corrections of the original time series; $\mathrm{M}A_{M+1}$ stands for the moving average through the last s time series values; P^{ε}_{M+1} stands for the Padé extrapolation through last M values of $\tilde{x}_k$; parameter α determines the penalty proportion between the sum of weighted corrections and the difference of forecasts based on MA_{n+1} and P^{ε}_{M+1}. Fig. 1 illustrates this technique.

To calibrate the fitness function in order to obtain optimal forecasts, an analysis on the impact of parameters α, β, γ to the fitness function must be performed. As shown in Fig. 5, the prediction results are closest to the original time series elements for $\alpha = 1, \beta = 0.5, \gamma = 2$ ($RMSE = 0.0480$). The parameter values obtained in this computational experiment are fixed for the subsequent computations.

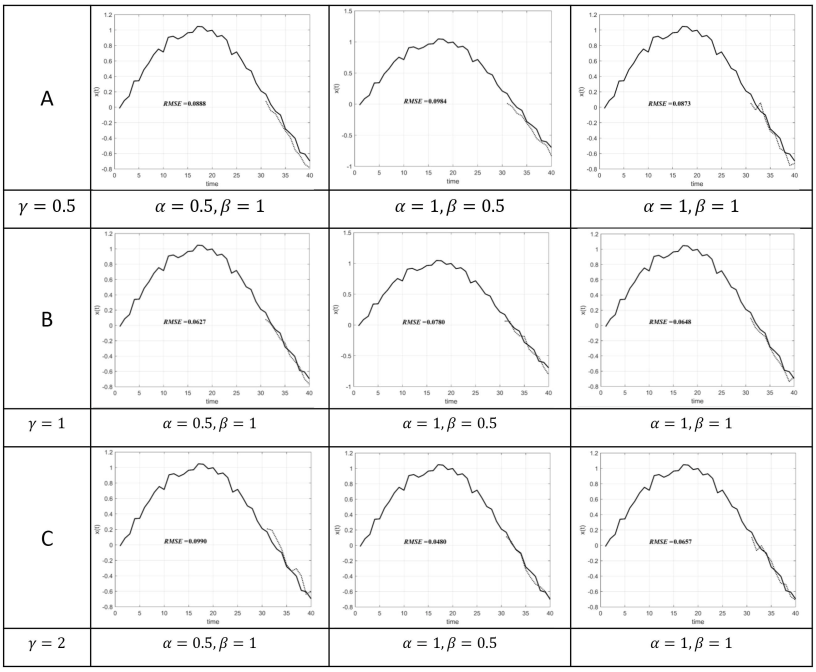

Fig. 2. Plots of the prediction accuracy for different values of the fitness function parameters α, β, γ.

6 Forecasting strategy for the Dow Jones time series

In previous section the optimal parameters of fitness function were selected. We apply this forecasting model with preselected fitness function parameters to real world time series: Dow Jones Industrial Average (DJIA) time series (data range provides 1896-05-26 to 2013-08-27 monthly index observations made up of 11 US stocks) [29]; DJIA time series is normed into the range $[-1; 1]$. The pre-processing is executed for $\{x_0, \ldots, x_{30}\}$. The Padé model is built using the initial 21 observations; RMSE is computed for the last 10 observations. Padé polynomial parameters read: $M = 20$, $m = n = 4$. Results of the prediction for these parameter values are displayed in Fig. 3. It can be noted that this model can be improved by taking into account dynamical nature of the Dow Jones time series.

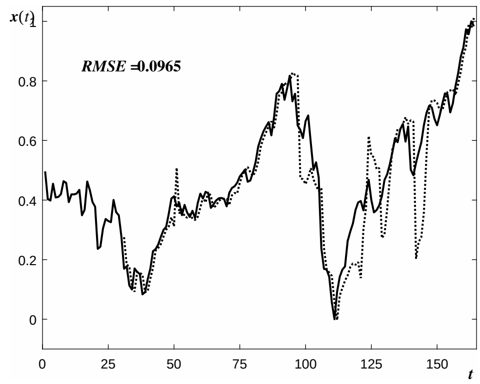

Fig. 3. Dow-Jones time series forecasting.

7 Adaptive forecasting of the Dow Jones time series

In this section, a strategy of adaptive selection of parameters M, m, n is proposed and used to predict different parts of time series. The proposed analysis is based on the idea of algebraic segmentation of short nonstationary time series [28]. The error level $\delta = 0.9$ is the key parameter that is preselected before the prediction is done. The prediction of the Dow Jones time series is shown in Fig. 4 (A); Fig. 4 (B) is the error plot for the prediction displayed in Fig. 4 (B).

Note that in the time interval $[31; 49]$, the time series has been predicted using Padé extrapolation with parameters $M = 20, n = m = 4$. In the interval $[50; 53]$, parameters $M = 20, n = m = 9$ have been used. In the next time series segment $[54, 96]$ optimal predictions are obtained for parameter values $M = 20, m = 5, n = 6$; for segments $[97; 99], [100; 106], [107; 117]$ the parameters values read $M = 11, m = 9, n = 1$; $M = 15, m = 8, n = 4$; $M = 20, n = m = 4$ respectively. For the remaining segment of the series, parameter values $M = 7, n = m = 1$ have been used. This parameter selection scheme has enabled the reduction of forecasting error down to $RMSE = 0.0707$.

8 Concluding remarks

Time series forecasting technique based on adaptive one step forward extrapolation of Padé extrapolants with internal smoothing is presented in this paper. Special optimization problem is developed for the identification of a near-optimal

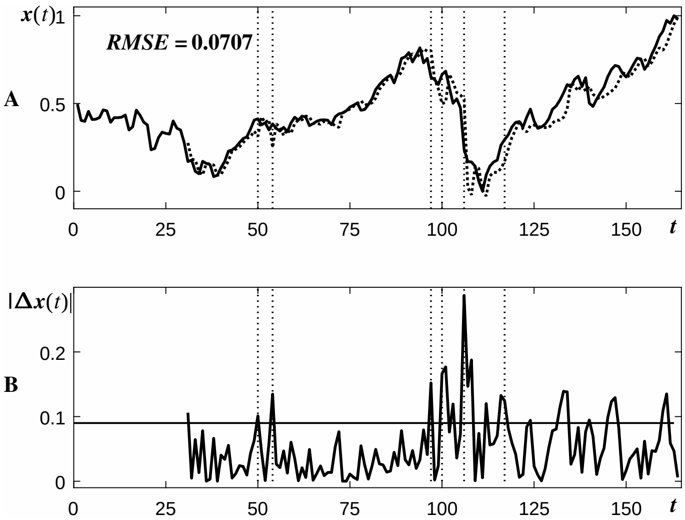

Fig. 4. Adaptive forecasting Dow Jones adaptive time series: part (A) is the time series forecast; part (B) is the error between the forecast and original time series values.

set of corrections used for the identification of the algebraic model in the observation window of the analyzed time series. Such an approach enables to unleash the power of Padé interpolants by ensuring the optimal smoothness of the extrapolant. Computational experiments with Dow Jones time series demonstrate the efficiency and the applicability of the proposed techniques for the prediction of real-world time series.

Acknowledgment

This research was funded by a grant (No. MIP078/2015) from the Research Council of Lithuania.

References

1. Parras-Gutierrez, E., Rivas, V. M., Garcia-Arenas, M., del Jesus, M. J.: Short, medium and long term forecasting of time series using the L-Co-R algorithm. Neurocomputing 128, 433–446 (2014)
2. Cao, Q., Leggio, K. B., Schniederjans M. J.: A comparison between Fama and Frenchś model and artificial neural networks in predicting the Chinese stock market. Computers & Operations Research, Vol. 32(10), 2499–2512 (2005)
3. Lam K., Lam, K. C.: Forecasting for the generation of trading signals in financial markets. Journal of Forecasting, Vol. 19(1), 39–52 (2000)
4. Podsiadlo, M., Rybinski, H.: Financial time series forecasting using rough sets with time-weighted rule voting. Expert Systems With Applications 66, 219–233 (2016)

5. Taylor, J. W., de Menezes, L. M., McSharry,P.E.: A comparison of univariate methods for forecasting electricity demand up to a day ahead. International Journal of Forecasting, Vol. 22(1), 1–16 (2006)
6. Darbellay, G. A., Slama, M.: Forecasting the short-term demand for electricity – Do neural networks stand a better chance?. International Journal of Forecasting, Vol.16(1), 71–83 (2000)
7. Nogales, F. J., Contreras, J., Conejo, A. J., Espnola, R.: Forecasting Next-Day Electricity Prices by Time Series Models. IEEE TRANSACTIONS ON POWER SYSTEMS, Vol. 17 (2), 342–348 (2002)
8. Kim, S., Shin,D. H.: Forecasting short-term air passenger demand using big data fromsearch engine queries. Automation in Construction 70, 98-108 (2016)
9. Brezinski, C., Redivo-Zaglia, M.: Padétype rational and barycentric interpolation. Numerische Mathematik , Vol. 125 (1), 89–113 (2013)
10. Baker, G. A. Jr., Graves-Morris P. R.: Padé Approximants Second edition, Cambridge University Press (1996)
11. Michael, D., Collins, A.: Split-step Padé Solution for the Parabolic Equation Method. J. Acoust. Soc. Am. 93, 1736 (1993)
12. Kamakura, T., Nomura, H., Clement, G. T.: Application of the split-step Padé approach to nonlinear field predictions Ultrasonics 53, 432–438 (2013)
13. Nagao, H.: The Padé Interpolation Method Applied to q-Painlevé Equations. Letters in Mathematical Physics, Vol. 105 (4), 503–521 (2015)
14. Baratchart, L., Yattselev, M. L.: Padé approximants to certain elliptic-type functions. Journal d'Analyse Mathématique , Vol. 121(1), 31–86 (2013)
15. Sablonnire, P.: Padé-type approximants for generalized Euler transforms. Numerical Algorithms, Vol. 66(2), 339–347 (2014)
16. Gruber, C., Luongo, O.: Cosmographic analysis of the equation of state of the universe through Pad approximations. Physical Review D 89, 103506 (2014)
17. Wang, Q., Moin, P., Iaccarino, G.: A Rational Interpolation Scheme with Superpolynomial Rate of Convergence. SIAM Journal on Numerical Analysis, Vol. 47(6), 4073–4097 (2010)
18. Gonnet, P., Pachon, R., Trefethen, L. N.: Robust rational interpolation and leastsquares. Electronic Transactions on Numerical Analysis 38, 146–167 (2011)
19. Gel,Y. R., Fomin, V. N.: Identification of an unstable ARMA equation. Mathematical problems in engerineering 7, 97–112 (2001)
20. Kumar, K.: Padé approximation and its application in time series analysis. Applied Mathematics and Computation, Vol. 48(2-3), 139–151 (1992)
21. Gil-Farina, M. C., Gonzalez-Concepcion, C., Pestano-Gabino, C.: Padé Approximation Modelling of an Advertising-Sales Relationship. J. Service Science & Management 3, 91–97 (2010)
22. Gonzalez-Concepcion, C., Gil-Farina, M. C.: Padé Approximation in Economics. Numerical Algorithms 33, 277–292 (2003)
23. Palivonaite, R., Ragulskis, M.: Short-term time series algebraic forecasting with internal smoothing. Neurocomputing, 127, 161–171 (2014).
24. Heath, M.: Scientific Computing. McGraw-Hill, New York (2000).
25. Palivonaite, R., Lukoseviciute, K.,Ragulskis, R.: Short-term time series algebraic forecasting with mixed smoothing. Neurocomputing 171, 854–865 (2016)
26. van Rijsbergen, C. J.: Information Retireval. Butterworths, London (1979)
27. Weiss,G. M.: Timeweaver: a Genetic Algorithm for Identifying Predictive Patterns in Sequences of Events. In Proceedings of the Genetic and Evolutionary Computation Conference 718–725 (1999)

28. Palivonaite, R.,Lukoseviciute, K., Ragulskis, M.: Algebraic segmentation of short nonstationary time series based on evolutionary prediction algorithms. Neurocomputing 121, 354–364 (2013)
29. Federal reserve bank of st. Louis, `http://research.stlouisfed.org/fred2/series/STLFSI/downloaddata`.

Proceedings

ITISE 2017

International work-conference on Time Series

Granada

September, 18-20 2017

Volumen 2

PROCEEDINGS ITISE 2017

Advanced in Time Series Forecast

Applications in Time Series

Forecasting Complex/Big data

Future of Mathematical and Logical Structures behind Time Series Analysis and History