Economic Index Forecasting via Multi-Scale Recursive Dynamic Factor Analysis

Haoyi Yuan[1], Yang Yuan[2], H. C. Wu[3] and Yuexian Zou[4*]

[1] Jiaxing University Nanhu College, Jiaxing, China

[2] Northeastern University, Seattle, USA

[3] The University of Hong Kong, Pokfulam, Hong Kong, China

[4] ADSPLAB, Peking University Shenzhen Graduate School, Shenzhen, China
[*] Correspondence Author: zouyx@pkusz.edu.cn

Abstract. In this paper, we propose a new multi-scale recursive dynamic factor analysis (MS-RDFA) algorithm for economic index foresting (EIF). The proposed MS-RDFA algorithm first employ empirical mode decomposition (EMD), which is a powerful tool for multi-scale analysis and modeling on the non-linear and non-stationary signal such as economic index data. Moreover, an efficient RDFA algorithm using recursive subspace tracking is adopted to explore the correlated nature of the adjacent intervals of the economic index data. The one-step prediction of PC scores is modeled as an AR process and can be recursively tracked by Kalman filter (KF). The major advantage of the proposed MS-RDFA method is its low arithmetic complexity and simple real-time updating, which is different from other conventional algorithms. This makes it as an attractive alternative to other conventional approaches to EIF on mobile services. The experiments show that the proposed MS-RDFA algorithm has better forecasting results than other EIF methods.

Keywords: Economic Index Forecasting, Empirical Mode Decomposition (EMD), Recursive Dynamic Factor Analysis (RDFA).

1. Introduction

Economic index including the commodity index and stock price index not only has a strong influence in the commodity futures market and the securities market, but also provides an early warning signal for macroeconomic regulation. Therefore, economic index forecasting (EIF) becomes to be one of the most attractive tasks encountered by financial organizations and private investors. And there are a large number of prediction models proposed to effectively mitigate the risk and to gain high investment return. Thanks to making use of information technology, the research on EIF can change from low-frequency data to high-frequency data. For instance, the Hang Seng Index (HIS) will be updated every minute which offers many opportunities for more detailed analysis of market activity. However, such high-frequency time-series data is non-normality and highly nonlinear [1, 2]. It thus calls for more desired approaches to deal with this new challenging task.

Different methods for economic forecasting had been proposed due to the increasing need on the practical applications such as transaction decision and investment. A traditional approach is based on autoregressive (AR) model in time series theory [3]. The AR model can be specifically solved using the ordinary least squares (LS). Other variants such as the autoregressive integrated moving average (ARIMA) methods have also been used in socket research [4]. For ARIMA, the processing speed is fast and the algorithm is easy to implement. However, when it comes to the non-stationary time series analysis, the performance may be significantly degraded, which results in large prediction error. Recently, support vector machine regression (SVR) has also been more envisioned in nonlinear regression estimation [5]. On the other hand, nonlinear variability in the data has led to people's considerable interest in neural networks [6]. Besides, hybrid models such as BPNN with genetic algorithm [7] and ANNs with metaheuristics [8] are put forward to economic index analysis. Hybrid method has potentially higher accuracy than other single forecasting method, but its arithmetic complexity is very high.

According to recent research [3, 5, 9], several common behaviors of economic time-series data are: a) highly non-linear and non-stationary and b) strong correlation between adjacent values. One of the major challenges of EIF is its non-linearity and non-stationary characteristics of caused by numerous influence factors, such as economy, affecting government, enterprise and investors [10]. This makes the modeling and prediction of the economic time-series data be very difficult. Therefore, multi-scale analysis methods can be employed to decompose signal into several sub-frequency components. In each sub-frequency components, signal will be more stationary and easier to be modeled and predicted. On the other hand, factor analysis (FA) techniques such as functional principal component analysis (FPCA) [11] is useful for extracting the correlation between adjacent intervals of the data. Online batch processing is usually desirable, where the forecasting is performed by applying the forecasting algorithm to a data block making up of consecutive economic samples. Whenever a new data is available, the existing data block is appended with the new sample and the earliest sample is discarded. This procedure is repeated for each incoming sample or blocks of samples in each update. However, this may also lead to high arithmetic complexity.

In this paper, we propose a multi-scale recursive dynamic factor analysis (MS-RDFA) algorithm to cope with the prediction problem and high arithmetic complexity incurred by such online real-time estimation. It first employs empirical mode decomposition (EMD) [12] to decompose the economic time-series data into several intrinsic mode functions (IMFs) along with a residue which stands for the trend. EMD is an effective approach to obtain instantaneous frequency data from non-stationary and nonlinear data. Moreover, an efficient RDFA algorithm using efficient recursive subspace tracking, called the orthonormal projection approximation subspace tracking with rank-1-modification (OPASTr) [13], to track recursively the major subspace spanned by the PCs. Since only the most recent sample is used for the updating, the memory storage required is also reduced. The one-step prediction of PC scores is modeled as an AR process. By assuming that the innovation is Gaussian distributed, the Kalman filter (KF) algorithm [14] can be used to recursively tracked the PC scores. An outline and major contributions are summarized below: a) A new foresting algorithm composed of EMD algorithm and RDFA algorithm is proposed. This algorithm reduces memory requirement and arithmetic complexity, which is appropriate for practical applications on mobile devices. b) A novel application of the proposed multi-scale RDFA (MS-RDFA) to the economic data such as commodity index and stock price index. Experimental results show that the proposed approach can achieve better ahead forecast accuracy than other methods. A real-time updating model with higher accuracy on EIF is obtained.

The following contents are organized as follows: The proposed MS-RDFA algorithm is described in section 2. Afterwards, its application to EIF and comparison to existing methods are presented in section 3. Finally, conclusions are drawn in section 4.

2. Proposed MS-RDFA Algorithm

2.1. Empirical Mode Decomposition

EMD is an effective multi-scale analysis method [12] which decomposes a signal into a sum of oscillatory functions which is called intrinsic mode functions (IMFs). An IMF is a function that has only one extreme between zero crossings, along with a mean value of zero. The IMFs can be extracted from the time series data set through an iterative decomposition process, which is described as follows:

Step 1: For a time-series signal $x(t)$, find its upper envelop $v_{\max}(t)$ and lower envelope $v_{\min}(t)$ by a cubic-spline interpolation of their local maximas and minimas.

Step 2: Calculate the means of the upper and lower envelops $m(t) = (v_{\max}(t) + v_{\min}(t))/2$.

Step 3: Extract the difference between the data and the mean of the upper and lower envelops as $d(t) = x(t) - m(t)$.

Step 4: Check whether $d(t)$ satisfy one of the following stop criterion: (i) the numbers of zero-crossings and extrema of $d(t)$ differs at most by one, or (ii) the predefined maximum iteration is reached. If the stop criterion is satisfied, then calculate the residue $r(t) = x(t) - d(t)$ and replace $x(t)$ by the residue $r(t)$. Otherwise, replace $x(t)$ by $d(t)$ and repeat step 1-3.

Step 5: Repeat steps 1-4 until the residual becomes a constant, a monotonic function, or a function with only one maximum and one minimum from which no more IMF can be extracted.

As a result, the original time-series signal is decomposed as the sum of these IMFs and a residual as

$$x(t) = \sum_{k=1}^{M} F_k(t) + r(t) \quad (1)$$

where M is the number of IMFs.

2.2. Functional Principal Component Analysis

Since the economic index data for adjacent intervals may be correlated, it is advantageous to employ FPCA [11] to explore and capture the correlations. The economic index data $z(n)$ can be approximated by a linear combination of orthogonal basis functions and their associated coefficients, which are referred to as principal component (PC) and PC score respectively. More precisely, suppose we are given the economic index data of N intervals (day or hour), i.e. $z(n)$ $n = 1, 2, \boxed{?}, N$. The number of samples in each interval is *J*. The samples in the *n*-th interval can be grouped into a vector: $\mathbf{z}(n) = [x((n-1)J+1), x((n-1)J+2), \boxed{?}, x(nJ)]^T$. For instance, if the interval is one day, then $J = 24$ for hourly index data and each vector represents the hourly data for the *n*-th day. The value of J can be adjusted for other time scales, such as $J = 48$ for half-hourly data. $z(n)$ is usually "centered", i.e. with its mean removed, before the PC functions are computed. Hence, the mean of $z(n)$, $n = 1, 2, \boxed{?}, N$ is first computed and is subtracted from each of the measurement vector to form $\bar{z}(n)$. In PCA, we wish to express the centered economic index vector $\bar{z}(n)$:

$$\bar{\mathbf{z}}(n) = \sum_{m=1}^{B} t_m(n)\mathbf{p}_m(n) + \mathbf{e}(n) \quad (2)$$

where *B* is an appropriately chosen number of PCs to achieve a sufficiently small approximation error $\mathbf{e}(n)$, $\boldsymbol{p}_m$ is the *m*-th PC, and $t_m(n)$ is its associated score for $\bar{z}(n)$. The batch eigen-decomposition (ED) of the following covariance matrix,

$$\boldsymbol{C}_{zz}(n) = \boldsymbol{U}(n)\boldsymbol{\Lambda}(n)\boldsymbol{U}^T(n) \quad (3)$$

is adopted to update the PCs.

To perform EIF, AR-based time-series model can be built for each PC score. More precisely, one-step ahead forecasting is given as:

$$\hat{\mathbf{z}}(n+1) = \boldsymbol{\mu}(n) + \sum_{m=1}^{B} \hat{t}_m(n+1)\,\mathbf{u}_m(n), \tag{4}$$

where $\boldsymbol{\mu}(n)$ is the mean vector and $\hat{t}_m(n+1)$ is one-step ahead PC score which should be predicted. However, it will require high arithmetic complexity to update the PCs using online implementation of ED in equation (3). This is not appropriate for low-light devices especially for mobile service.

2.3. Subspace Tracking

Motivated by the PAST algorithm in [15], the signal subspace spanned by the major PCs $\mathbf{U}_B(n)$ is tracked recursively instead of computing the entire ED. In the OPASTr algorithm [13], the PCs are extracted from the signal subspace tracking. Given the signal subspace $\boldsymbol{W}(n)$, the covariance matrix $\mathbf{C}_{zz}(n) = \mathbf{U}(n)\mathbf{\Lambda}(n)\mathbf{U}^T(n)$ is projected onto the signal subspace $\boldsymbol{W}(n)$ to obtain

$$\begin{aligned} \boldsymbol{C}_{yy}(n) &= \boldsymbol{W}^T(n)\boldsymbol{U}(n)\boldsymbol{\Lambda}(n)\boldsymbol{U}^T(n)\boldsymbol{W}(n) \\ &= \boldsymbol{W}^T(n)\boldsymbol{U}_B(n)\boldsymbol{\Lambda}_B(n)\boldsymbol{U}_B^T(n)\boldsymbol{W}(n) = \boldsymbol{\Phi}(n)\boldsymbol{\Lambda}_B(n)\boldsymbol{\Phi}^T(n), \end{aligned} \tag{5}$$

where $\boldsymbol{\Phi}(n)$ is a $B \times B$ orthogonal transformation satisfying $\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n) = \boldsymbol{I}$ and

$$\mathbf{U}_B(n) = \mathbf{W}(n)\,\mathbf{\Phi}(n). \tag{6}$$

The covariance matrix $\boldsymbol{C}_{yy}(n) = E[\boldsymbol{y}(n)\boldsymbol{y}^T(n))]$ can be recursively updated as

$$\boldsymbol{C}_{yy}(n) = \beta\boldsymbol{C}_{yy}(n-1) + (1-\beta)\boldsymbol{y}(n)\boldsymbol{y}^T(n), \tag{7}$$

where $\mathbf{y}(n) = \mathbf{W}^T(n)\bar{\mathbf{z}}(n)$, which a projection of $\bar{\mathbf{z}}(n)$ on the subspace $\boldsymbol{W}(n)$. $\mathbf{\Phi}(n)$ can be recursively computed using the ED of $\mathbf{C}_{yy}(n)$. Firstly, let the ED of $\mathbf{C}_{yy}(n-1)$ be $\mathbf{\Phi}(n-1)\mathbf{\Lambda}_B(n-1)\mathbf{\Phi}^T(n-1)$. The expression in (7) can be rewritten as one rank-one modification [16] given by

$$\boldsymbol{C}_{yy}(n) = \boldsymbol{\Phi}(n-1)[\beta\boldsymbol{\Lambda}_B(n-1) + (1-\beta)\boldsymbol{s}(n)\boldsymbol{s}^T(n)]\boldsymbol{\Phi}(n-1)^T, \tag{8}$$

where $\mathbf{s}(n) = \quad ^{T}(n-1)\check{\ }(n)$.Let the corresponding ED be

$$\beta\boldsymbol{\Lambda}_B(n-1) + (1-\beta)\boldsymbol{s}(n)\boldsymbol{s}^T(n) = \boldsymbol{\Phi}(n)\boldsymbol{\Lambda}_B(n)\boldsymbol{\Phi}^T(n). \tag{9}$$

The ED of the rank-one update in (9) can be recursively computed using rank-one modification. Finally, the eigenvectors of $\mathbf{C}_{yy}(n)$ are given by

$$\boldsymbol{\Phi}(n) = \boldsymbol{\Phi}(n-1)\widetilde{\boldsymbol{\Phi}}(n), \tag{10}$$

Then, the PC scores $\mathbf{t}(n) = [t_1(n), t_2(n), \boxed{?} t_B(n)]^T$ can be computed recursively as

$$\mathbf{t}(n) = \mathbf{U}_B^T(t)\overline{\mathbf{z}}_s(n), \tag{11}$$

2.4. Kalman Filter for EIF

Once PCs and PC scores are recursively updated, time-series prediction model can be built for each PC score to perform EIF. In this paper, the KF [14] is employed to recursively track the PC score so that the EIF can be computed online in a real-time manner. More precisely, for each PC score $t_m(n)$, a L-th order AR model will be constructed. The solution of the AR model can be obtained by the following the least squares (LS) formulation [17],

$$\begin{aligned} \min \ \{ & \textstyle\sum_{i=1}^{n} \| R_m^{-1/2}(i)(t_m(i) - \Sigma_{j=1}^{L}\alpha_j(i)t_m(i-j)) \|_2^2 \\ & + \textstyle\sum_{i=1}^{n} \| \boldsymbol{Q}_m^{-1/2}(i)(\boldsymbol{\alpha}_m(i) - \boldsymbol{\alpha}_m(i-1)) \|_2^2 \}, \end{aligned} \tag{12}$$

where $\boldsymbol{\alpha}_m(i) = [\alpha_{m,1}(i), \alpha_{m,2}(i), \boxed{?}, \alpha_{m,L}(i)]^T$ are the AR coefficients and $R_m(i)$ is the covariances of the loss function $e_m(i) = t_m(i) - \Sigma_{j=1}^{L}\alpha_j(i)t_m(i-j)$. To reduce the variance of the estimator, we incorporate a regularization term $\|\boldsymbol{\alpha}_m(i) - \boldsymbol{\alpha}_m(i-1)\|_2^2$ into (12) where $\boldsymbol{Q}_m(i)$ are the covariances of the loss function $\boldsymbol{\varepsilon}_m(i) = \boldsymbol{\alpha}_m(i) - \boldsymbol{\alpha}_m(i-1)$. The inverses $R_m^{-1/2}(i)$ and $\boldsymbol{Q}_m^{-1/2}(i)$ are used to perform scaling on each variable (whitening) to achieve equal variance of the transformed variables. The regularization term requires the estimate to stay close to the previous estimate and hence the variance of the estimator will be reduced. It is shown in [18] that Eqn. (12) can be formulated as the following state space model (SSM):

$$\boldsymbol{\alpha}_m(n) = \boldsymbol{\alpha}_m(n-1) + \boldsymbol{\varepsilon}_m(n), \tag{13}$$

$$t_m(n) = \boldsymbol{h}_m(n)^T \boldsymbol{\alpha}_m(n) + e_m(n). \tag{14}$$

Eqn. (13) is the state equation and it describes the evolution of the AR coefficients over time, as a function of the previous AR coefficients $\boldsymbol{\alpha}_m(n)$ and $\boldsymbol{\varepsilon}_m(n)$ represents the modeling error. Eqn. (14) is the measurement equation which models the current PC score $t_m(n)$ with previous scores $\boldsymbol{h}_m(n) = [t_m(n-1), t_m(n-2), ..., t_m(n-L)]^T$, and the AR coefficients $\boldsymbol{\alpha}_m(n)$ represent the weighting of each previous score $t_m(n-i)$, $i = 1, 2, \boxed{?} L$, and $e_m(n)$ is the measurement noise. We can see that the state equation in (13) and the measurement equations in (14) are equivalent to the regularization term and the loss function in (12) respectively. The SSM in (13) and (14) can be recursively tracked using the Kalman filter (KF). Afterwards, the one-step ahead prediction $\hat{t}_m(n+1)$ is given by

$$\hat{t}_m(n+1) = \boldsymbol{h}_m^T(n+1)\boldsymbol{\alpha}_m(n+1 \mid n), \tag{15}$$

From (15), the one-step ahead prediction of the economic index data vector $\hat{\boldsymbol{z}}(n+1)$ can be determined as follows,

$$\hat{\boldsymbol{z}}(n+1) = \quad (n) + \Sigma_{m=1}^{B} \hat{t}_m(n+1) \quad {}_m(n), \tag{16}$$

To perform two-step ahead prediction, one can append the prediction $\hat{t}_m(n+1)$ into $\boldsymbol{h}_m(n+2) = [\hat{t}_m(n+1), t_m(n), t_m(n-1), \boxed{?}, t_m(n-L+2)]^T$, and then apply the KF again to predict $\hat{t}_m(n+2)$ using (15). This procedure can be repeated for h times to compute a h-step ahead prediction.

3. Application to Economic Index Forecasting

MS-RDFA based EIF consists of three steps: i) Use EMD to decompose the original economic time-series data into several IMFs and one residue; ii) Use RDFA algorithm in section 2 to extract the predicted results of IMFs and residue; and iii) Combine all the prediction results to reconstruct the final result. To verify the effectiveness of the proposed MS-RDFA algorithm, two most-widely used economic indices such as Goldman Sachs Commodity Index (GSCI) and Hang Seng Index (HSI) are employed for EIF experiments. The GSCI index is one of the most famous commodity indices for prediction of the economic trend when HSI index is a stock price index to track the stock quotes. The prediction error is determined by computing the difference between the predictor and the actual value. We employ the Absolute Percentage Error (MAPE) to evaluate the performance of the proposed EIF algorithm:

$$\text{MAPE} = \frac{1}{K}\sum_{k=1}^{K}\left|\frac{x^{(k)}(n) - \hat{x}^{(k)}(n)}{x^{(k)}(n)}\right|. \qquad (17)$$

Generally, lower MAPE indicates better accuracy of the algorithm. We choose two different time resolutions for them to verify the effectiveness and robustness of the proposed method. For GSCI index, the hourly data of from January 1, 2017 to December 31, 2017 is used for experiments. The step of ahead prediction is one day, which namely one-day ahead prediction. Each valid day has 6 hours for GSCI, hence $J=6$ and the order of AR model is chosen as $L=7$. The number of PCs is chosen as $B=4$, which is determined by the MDL method in [21]. For HSI Index, five-minute data from December 1 to December 31, 2017 is collected for experiments. Here the step of ahead prediction denotes one hour. Hence, $J=12$ and the order $L=5$ is chosen for the one-hour ahead prediction. The number of PCs is still chosen as $B=4$. An initial data block of length 10 days and 6 hours is used to initialize the MS-RDFA algorithm for the one-day ahead prediction and one-hour ahead prediction respectively. After each prediction step, the actual data has been obtained and the model should be updated using the most recent actual data. This procedure is repeated for onward prediction such that the latest actual data is incorporated into the model on each prediction step.

Four state-of-arts methods namely RDFA [17], Support Vector Machine Regression (SVR) [5], ANN [6] and hybrid model (HM) [7] are employed for comparison. Now, we compared the performance of RDFA, SVR, ANN, and HM algorithms by the quantitative measure of MAPE. The comparison of MAPE results are shown in Table 1 and Table 2. It can be observed from the Table 1 and Table 2 that the proposed MS-RDFA has the best results, whose error has been significantly reduced compared to ANN, SVR and HM. Unlike the other algorithms, which generally give larger prediction error when the forecasting period increases, we can see that the MS-RDFA and RDFA algorithm generally gives consistent forecasting error for 1-4 hours or 5-20 minutes ahead forecast. The better performance of the MS-RDFA and RDFA is partly contributed by the consideration of the temporal data structure. Moreover, thanks to the multi-scale decomposition of the proposed MS-RDFA algorithm, it can achieve lower errors than RDFA algorithm. The reason is that the signal in each IMF is more stationary to predict and analysis.

Table 1. Comparison of the forecasting errors, hourly data of GSCI index.

Index	1 hour	2 hour	3 hour	4 hour
MS-RDFA	0.1155	0.1185	0.1138	0.1127

RDFA	0.1231	0.1271	0.1282	0.1249
SVR	0.1801	0.2970	0.3421	0.3985
ANN	0.2269	0.3204	0.4036	0.4725
HM	0.1821	0.1976	0.2928	0.3491

Table 2. Comparison of the forecasting errors, five-minute data of HSI index.

Index	5 min	10 min	15 min	20 min
MS-RDFA	0.1856	0.1825	0.1833	0.1857
RDFA	0.1901	0.1970	0.1921	0.1915
SVR	0.2003	0.2572	0.2844	0.3312
ANN	0.2219	0.2599	0.2988	0.3196
HM	0.1987	0.2578	0.2889	0.3102

4. Conclusion

A new multi-scale RDFA (MS-RDFA) algorithm has been proposed based on the EMD algorithm and RDFA algorithm for economic index forecasting (EIF). Experimental results of economic indices such as Goldman Sachs Commodity Index (GSCI) and Hang Seng Index (HSI) show that the proposed MS-RFDA can achieve higher daily and hourly ahead forecast accuracy and stability than other conventional algorithms. Moreover, the low complexity of efficient recursive implementation of the proposed MS-RFDA makes it as an attractive alternative to other conventional approaches to EIF and other possible applications on mobile services.

Acknowledgement

This paper was partially supported by Shenzhen Key Laboratory for Intelligent Multimedia and Virtual Reality (ZDSYS201703031405467).

References

1. Jacquier E., Polson N. G., and Rossi P. E.: Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. J. Econ. 122(1), 185–212 (2004).

2. Jobson J. D., Korkie B.: Estimation for Markowitz efficient portfolios. J. Amer. Statist. Assoc. 75(371), 544–554 (1980).

3. Box G. E. P., Pierce D. A.: Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. J. Amer. Statist. Assoc. 65(332), 1509–1526 (1970).

4. Islam M.A., Hassan M.F., Imam M.F., and Sayem S.M.: Forecasting coarse rice prices in Bangladesh, Progress Agriculture 22, 193-201(2011).

5. Nayak S. C., Misra B. B., and Behera H. S.: An adaptive second order neural network with genetic-algorithm-based training (ASONN-GA) to forecast the closing prices of the stock market. Int. J. Appl. Metaheuristic Comput. 7(2), 39–57 (2016).

6. Göçken M., Özçalici M., Boru A., and Dosdogru A. T.: Integrating metaheuristics and artificial neural networks for improved stock price prediction. Expert Syst. Appl. 44, 320–331 (2016).

7. Zhang G., Patuwo B. E., and Hu M. Y.: Forecasting with artificial neural networks: The state of the art. Int. J. Forecasting 14(1), 35–62 (1998).

8. Lu C. J., Lee T. S., and Chiu C. C.: Financial time series forecasting using independent component analysis and support vector regression. Decision Support Syst. 47(2), 115–125, (2009).

9. Duan Q., Zhang L., Wei F., Xiao X., and Wang L.: Forecasting model and validation for aquatic product price based on time series GA-SVR. Trans. Chin. Soc. Agricult. Eng. 33(1), 308–314 (2017).

10. He, K., Zha, R., Wu, J., Lai, K.K.: Multivariate EMD-based modeling and forecasting of crude oil price. Sustainability 8(4), 387 (2016).

11. Shang H. L.: Nonparametric Modeling and Forecasting Electricity Demand: An Empirical Study, Department of Econometrics and Business Statistics, Monash University, Monash Econometrics and Business Statistics Working Papers, No. 19/10, 2010.

12. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Roy. Soc. London A. 454, 903–995 (1998).

13. Chan S. C., Wu H. C., and Tsui K. M.: Robust recursive eigen-decomposition and subspace-based algorithms with application to fault detection in wireless sensor networks. IEEE Trans. Instrum. Meas. 61(6), 1703-1718 (2012).

14. Liao B., Zhang Z. G., and Chan S. C.: A new robust kalman filter based subspace tracking algorithm in an impulsive noise environment. IEEE Trans. Circuits Syst. II, Exp. Briefs 57(9), 740–744 (2010).

15. Yang B.: Projection approximation subspace tracking. IEEE Trans. Signal Process. 43 (1), 95–107 (1995).

16. Golub G. H., Van Loan C. F.: Matrix Computations, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press (1996).

17. Wu H. C., Chan S. C., Tsui K. M. Hou Y.: A new recursive dynamic factor analysis for point and interval forecast of electricity price. IEEE Transactions on Power Systems 28(3), 2352-2365, 2013.

18. Fahrmeir L., Tutz G.: Multivariate Statistical Modelling Based on Generalized Linear Models, ser. Springer Series in Statistics, 2nd ed., New York: Springer (2001).