

ORIGINAL RESEARCH

Open Access



Time series modeling and filtering method of electric power load stochastic noise

Li Huang^{1*}, Yongbiao Yang¹, Honglei Zhao², Xudong Wang² and Hongjuan Zheng¹

Abstract

Stochastic noises have a great adverse effect on the prediction accuracy of electric power load. Modeling online and filtering real-time can effectively improve measurement accuracy. Firstly, pretreating and inspecting statistically the electric power load data is essential to characterize the stochastic noise of electric power load. Then, set order for the time series model by Akaike information criterion (AIC) rule and acquire model coefficients to establish ARMA (2,1) model. Next, test the applicability of the established model. Finally, Kalman filter is adopted to process the electric power load data. Simulation results of total variance demonstrate that stochastic noise is obviously decreased after Kalman filtering based on ARMA (2,1) model. Besides, variance is reduced by two orders, and every coefficient of stochastic noise is reduced by one order. The filter method based on time series model does reduce stochastic noise of electric power load, and increase measurement accuracy.

Keywords: Electric power load, Stochastic noise, ARMA model, Kalman filter

1 Introduction

Power load operation is complex. Accurate power load forecasting has great significance for designing power supply program and making a good power balance between supply and demand. The power load sequence contains relatively obvious white noise. With longer sampling time interval, the white noise becomes more intense [1, 2]. The prediction accuracy of power load is related to the length of historical observation data. With noise and chaos in the observed data, different time series have different upper limit of prediction accuracy [3, 4]. It is important to estimate the noise intensity directly from the observed data and to separate the noise from the observed data, which is very important to improve the accuracy of the power load forecasting result.

To improve the quality of power load data, stochastic noise present in the load data must be identified and filtered out [5, 6]. At present, there are mainly following methods in the power load forecasting field, such as regression analysis, combined forecasting, exponentially smoothing, neural network and wavelet methods, and so on. Moreover, in view of the uncertainties and randomness

of short-term load, innovative data processing strategies are proposed, such as frequency domain decomposition method and property matrix hierarchical analysis method [7–9]. However, the existing time series modeling methods may not meet the requirements of time series stationary. These methods neglect the pretreatment of load data and statistical checking [10]. Independent, steady, normal, zero-mean and trend-item processing of the required data is required, and non-stationary, non-random and non-normal characteristics of power load data are needed to be tested.

Time series method and Kalman filter algorithm are proposed to filter the power load stochastic noise by pretreating and statistically testing of power load data, then, the total variance method is used to evaluate the stochastic errors of the load data before and after filtering effectively.

2 Methods

2.1 Stochastic noise time series method in power load data

2.1.1 Timing sequence processing

Traditional load forecasting method adopts the regression analysis and the least square method. However, this method is difficult to reflect the new information of the load change during the operation of the power system to the model, and the prediction accuracy is low. According to the characteristics of power load data, the statistical

* Correspondence: 15850575576@163.com

¹NARI Technology Development CO., Ltd, Nanjing, Jiangsu Province 211106, China

Full list of author information is available at the end of the article

parameter model reflecting the running state of the system is established, and the time series of electric load is constructed. Then, the shortcomings of the existing methods can be effectively overcome [11–13].

The time series model is used to fit the stationary, normal sequence. An auto regressive moving average (ARMA) model ARMA(p, q) with appropriate order can be used to describe the stationary stochastic process of power load. ARMA(p, q) model of a stable normal time series $\{x_k\}(k = 1, \dots, n)$ can be obtained by

$$x_k = \phi_1 x_{k-1} + \phi_2 x_{k-2} + \dots + \phi_p x_{k-p} + a_k - \theta_1 a_{k-1} - \dots - \theta_q a_{k-q} \quad (1)$$

where $\{x_k\}$ is time series, x_k is the value of the time series $\{x_k\}$ at the k-th moment, and x_k can be estimated by the value of the timing in the past periods $x_{k-1}, x_{k-2}, \dots, x_{k-p}$, ϕ_p is autoregressive coefficient, θ_q is moving average coefficient, a_k is residual, p and q are orders of ARMA model.

The estimated error of x_k is obtained by

$$e = a_k - \theta_1 a_{k-1} - \dots - \theta_q a_{k-q} \quad (2)$$

The prerequisite for establishing the ARMA model is that the load data satisfy the requirements of stationarity and normality. Power load output data usually do not meet these requirements, then, it is necessary to make pre-processing operations and test of the corresponding characteristics for sampled data.

The first step is stationary test. The reverse order test is used to test the stationary state of the power load data sequence. If the stationarity requirement is not satisfied, the trend item extraction is carried out for the stochastic load sequence. The reverse order test method is carried out as following. $\{x_n\}$ is divided as subsequences $\{x_{j,n}\}$ with quantity of l. The mean value μ_i of each subsequence is obtained, and new subsequence is obtained with $\mu_1 \mu_2 \dots \mu_i$. With $i > j$, the reverse order A_j equals to the amount of $\mu_i > \mu_j$. The total reverse order number of sequences is obtained by

$$A = \sum_{j=1}^{l-1} A_j \quad (3)$$

where $j \in (0, 1, 2, \dots, l)$.

The theoretical mean and variance of the total number of reversal order are obtained as following [14]:

$$E[A] = \sum_{j=1}^{l-1} E[A_j] = \frac{l(l-1)}{4} \quad (4)$$

$$\sigma_A^2 = \frac{l(2l^2 + 3l - 5)}{72} \quad (5)$$

Statistics value u is obtained as following [14]:

$$u = \frac{(A + \frac{1}{2} - E[A])}{\sigma_A} \quad (6)$$

If $|u| \leq 1.96$, there is no significant difference between μ_i and $\{x_n\}$ can be determined to be a stationary sequence.

The second step is trend item extraction. The data sequence is processed by difference to get the new sequence. Data sequence subtracts the mean of the new sequence, then, obtains the mean value of the difference to complete the trend item extraction.

The third step is normality test. The power load data sequence was tested for normality [14], mainly including standard skewness coefficient ξ and standard kurtosis coefficient v.

Mean value is obtained by

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (7)$$

Variance value is obtained by

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (8)$$

Standard skewness coefficient is obtained by

$$\xi = \sqrt{\frac{1}{6n}} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{S} \right)^3 \quad (9)$$

Standard kurtosis coefficient is obtained by

$$v = \sqrt{\frac{n}{24}} \left[\sum_{i=1}^n \left(\frac{x_i - \bar{x}}{S} \right)^4 - 3 \right] \quad (10)$$

$\xi \approx 0$ and $v \approx 0$ indicates stochastic sequence satisfies the normality requirement.

2.1.2 Online timing modeling

After pretreatment and statistical tests of power load data, model order and parameters also need to be calculated. In addition, the applicability of the new model still need to be tested [14, 15]. Based on the new model, the system state equation and output equation can be established, and Kalman filter method can be used to deal with the power load data.

The common method Akaike Information Criterion (AIC) for judging the order of time series models is given by

$$AIC(p, q) = n \ln \sigma_a^2 + 2(p + q) \quad (11)$$

where p and q are orders of ARMA model, n is the number of data in the sequence, σ_a^2 is the Variance of noise $a(t)$.

The AIC criterion takes into account the interaction between model order and residuals, and the smallest AIC value is to be selected.

The applicability of the model is also a critical task for online modeling of power load data. The criterion is to

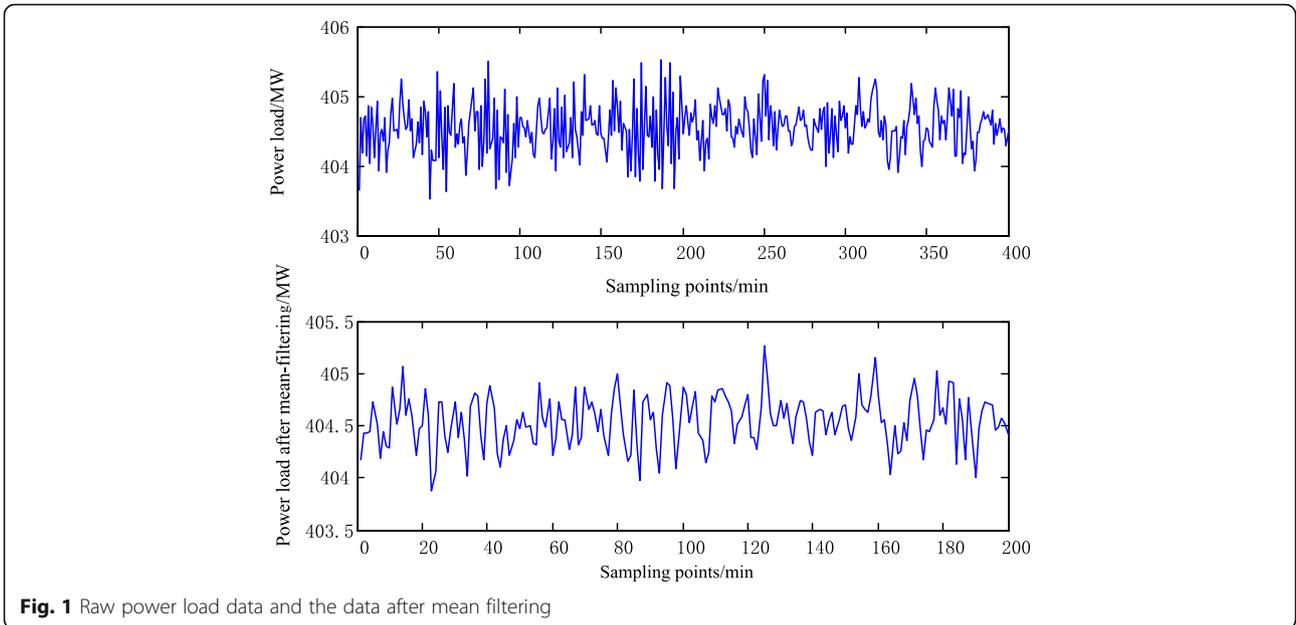


Fig. 1 Raw power load data and the data after mean filtering

check whether the model residuals are white noise. If the model residuals are white noise, the model is available; otherwise, it is not applicable.

2.2 Kalman filtering based on time series model

Kalman filtering method, a kind of effective recursive filtering method, estimates the system state according to a series of measurements including stochastic noise. Kalman filtering selects proper state space, builds state equation and measurement equation, based on the period and characteristic of load prediction. Parameter estimation

and load forecasting are implemented in the filtering, to be an organic whole.

According to the ARMA model, Kalman filtering method is adopted to suppress the stochastic noise of power load. System state equation is built by white noise of the stochastic noise of power load [16, 17].

State equation is as following:

$$X_k = AX_k + Bv_k \tag{12}$$

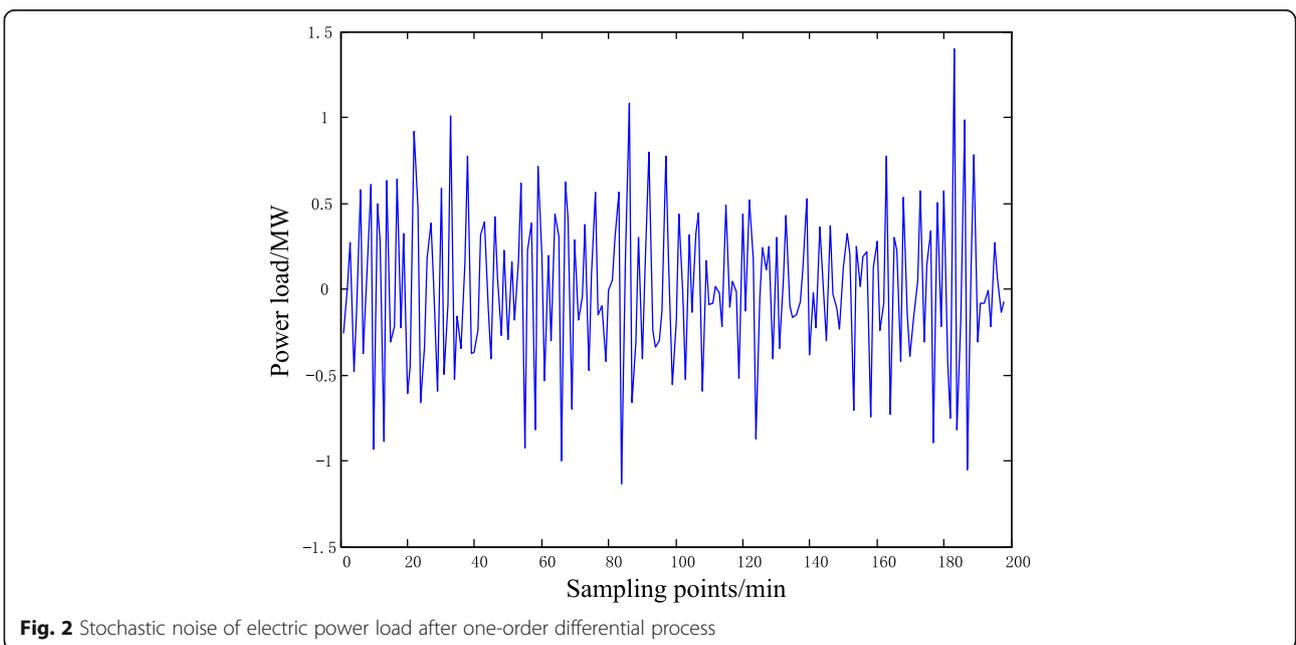


Fig. 2 Stochastic noise of electric power load after one-order differential process

Table 1 AIC values of ARMA model of power load

p	q	AIC value	p	q	AIC value
-	-	-	2	0	0.2145
0	1	-0.0268	2	1	-0.0295
0	2	-0.0265	2	2	-0.0212
0	3	-0.0144	2	3	-0.0273
1	0	0.3546	3	0	-0.0157
1	1	-0.0221	3	1	-0.0219
1	2	-0.0233	3	2	-0.0249
1	3	-0.0275	3	3	-0.0263

Assuming that W_k is estimation error of ARMA model, so there is an equation as following:

$$Y_k = X_k + W_k \tag{13}$$

System output is as following:

$$Z_k = Y_k \tag{14}$$

Output equation is as following:

$$Z_k = CX_k + W_k \tag{15}$$

The mean of both v_k and W_k is zero, white noise with constant autocorrelation function is independent of each other. The statistical properties satisfy the mean equals to zero, $E(W_k) = E(v_k) = 0$. Autocorrelation function $\phi_{vv} = R\delta_{ki}$, $\phi_{vv} = Q\delta_{kj}$, and cross-correlation function $\phi_{vw}(k, j) = 0$.

Kalman filtering equations of power load are built based on state equation and output equation, shown as following:

$$\begin{cases} \hat{X}_{k,k} = A\hat{X}_{k-1,k-1} \\ \hat{X}_{k,k} = \hat{X}_{k,k-1} + K_k[Y_k - C\hat{X}_{k,k-1}] \\ K_k = P_{k,k-1}C^T[CP_{k,k-1}C^T + R_k]^{-1} \\ P_{k,k-1} = AP_{k,k-1}A^T + BQ_{k-1}B^T \\ P_{k,k} = [I - K_kC]P_{k,k-1} \\ \hat{Y}_k = C\hat{X}_{k,k} \end{cases} \tag{16}$$

where, $\hat{X}_{k,k-1}$ is further estimation of filtering state, $\hat{X}_{k,k}$ is the filtering state at the time k, $Y_k - C\hat{X}_{k,k-1}$ is optimal estimate at the time k being the error between observation estimation and observation value, K_k is gain matrix of filter at the time k, R is error of system measurement noise, and Q is noise variance of system process, and \hat{Y}_k is the output of filter at the time k. Initial values need to be given in advance.

Kalman prediction process is the filtering process of state reconstruction. Known from Eq.(16), estimated information \hat{X} of state phasor X is updated constantly. Considering feedback unit, this part can avoid the effect of dynamic noise v_k . However, for estimation value \hat{Y} of output phasor Y, it can only be approximated owing to the influence of dynamic noise v_k .

3 Result

3.1 Application and analysis of time series model and Kalman filtering

To verify the validity of time series model and Kalman filtering method of power load stochastic noise, 100 power load data of some place in 2015 is analyzed as following. After mean-filtering, power load data is able to

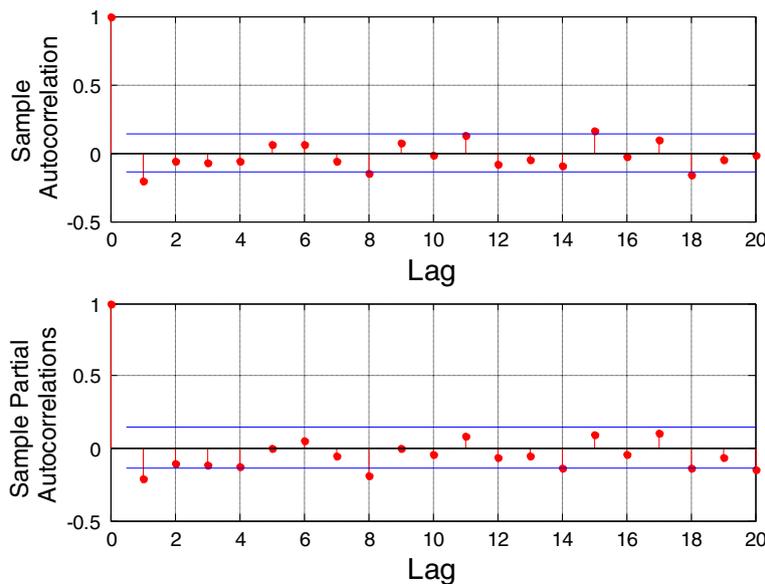
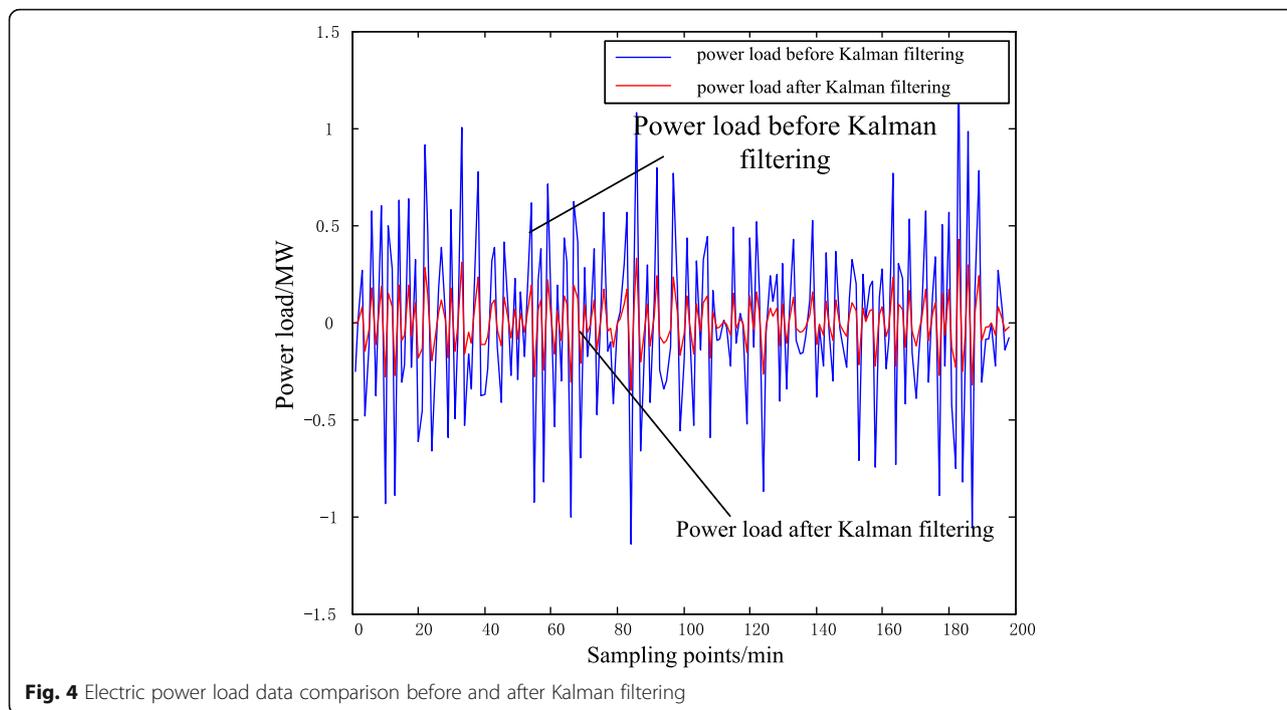


Fig. 3 ACF and PACF of model residual



effectively characterize raw power load data, both shown in Fig.1.

The first stationarity test results of power load data is as following: $|u| = 2.54 > 1.96$, which means it doesn't meet the stationarity requirement. After extracting trend, the result is this, namely $|u| = 0.5 < 1.96$, meeting the requirement. The results of normality test are as following: standard skewness coefficient $\xi = 0.0032 \approx 0$, standard kurtosis coefficient $\nu = 4.54 \times 10^{-4} \approx 0$. The new data sequence meets the normality requirements. After trend extraction, stochastic noise of power load is shown in Fig. 2. And new power load data is stationary, zero-mean and normal, satisfying the precondition of online modeling.

As for the power load suitable for modeling, AIC values are calculated. In addition, orders of ARMA model are relatively small, p and q is set to be less than 3. AIC values of chosen model are listed in Table 1.

Table 2 Stochastic noise coefficients before and after Kalman filtering

Noise coefficient	Before filtering	After filtering
Q	1.50e-3	1.91e-4
L	1.01e-5	1.26e-6
B	4.74e-4	5.66e-5
K	7.50e-3	8.89e-4
R	4.31e-2	5.10e-3

Table 1 demonstrates that ARMA(2,1) model shall be selected for power load stochastic noise model according to the minimum AIC value, built as following:

$$x_k = \phi_1 x_{k-1} + \phi_2 x_{k-2} + a_k - \theta_1 a_{k-1} \tag{17}$$

where, x_k is the model output, and a_k is the white noise, of which mean is 0 and variance is σ_a^2 . ϕ_1 , ϕ_2 and θ_1 is calculated by least square fitting (LSF)

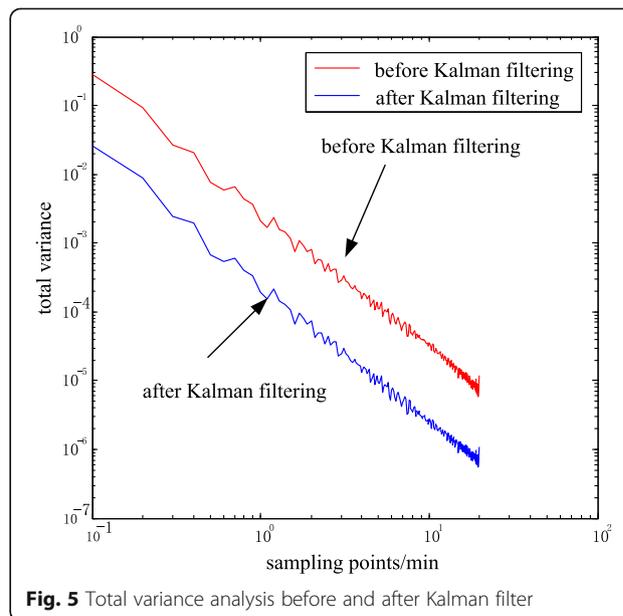


Fig. 5 Total variance analysis before and after Kalman filter

$$x_k = -0.7072x_{k-1} - 0.1325x_{k-2} + a_k - 0.1242a_{k-1} \quad (18)$$

ACF and PACF of model residual is shown as Fig.3, and both can be regarded as white noise input.

According to established model ARMA(2,1), corresponding system state equation is obtained as following:

$$x_k = -0.7072x_{k-1} - 0.1325x_{k-2} + a_k - 0.1242a_{k-1} \quad (19)$$

System output equation is as following:

$$X_k = AX_{k-1} + BV_k \quad (20)$$

where, $A = \begin{bmatrix} -0.7072 & -0.1325 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0.1242 \\ 0 & 0 \end{bmatrix}$, $C = [1 \ 0]$.

Initial value of co-variance matrix P is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, initial value of matrix X is $[0 \ 0]^T$, value of matrix R is variance of estimation error, and value of progress noise Q equals to $\begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_a^2 \end{bmatrix}$.

Kalman filtering method is used to denoise the stochastic noise of power load data, the curves before and after filtering shown in Fig. 4. Result demonstrates that noise amplitude in the stochastic noise data is significantly reduced by ARMA model and Kalman filter. Variance before filtering is 1.56×10^{-4} , after filtering it becomes 3.58×10^{-6} , reduced by two orders of magnitude. The filtered stochastic noise is obviously suppressed.

Stochastic noise of power load is presented with different correlation time and power spectral density, and total variance method is effective to evaluate five kinds of stochastic noise of power load data before and after filtering, including load random walk (L), bias instability (B), rate ramp walk (K), rate ramp (R) and quantization noise (Q). Table 2 is each stochastic noise coefficient before and after Kalman filtering, and Fig.5. is the total variance curve of power load before and after Kalman filtering.

Known from Table 2 and Fig.5, selected power load data mainly contains quantization noise, rate random walk and bias instability. However, each stochastic coefficient in the power load data is effectively reduced through time series modeling and Kalman filtering, each coefficient value is reduced by an order of magnitude. The proposed method can eliminate the stochastic noise of power load data and promote power load accuracy.

4 Conclusion

Suppressing power load stochastic noise is one of the important links in power load modeling and forecasting.

Based on the characteristics of power load data, time series analysis is used to model the data of power load on-line, realizing the pretreatment and inspection analysis of power load data. ARMA (2,1) model is established and Kalman filtering method is used to denoise load data. And total variance method is adopted to verify the effect of modeling and filtering, namely the stochastic error coefficients before and after filtering.

The results show that stochastic noise amplitude of power load data after time series modeling and Kalman filtering is significantly reduced, the variance value is decreased by two orders of magnitude, and each stochastic error coefficient of power load is reduced by an order of magnitude. The proposed time series modeling and filtering method can effectively suppress the stochastic noise of power load data and improve the prediction accuracy of power load.

Acknowledgments

This work was financially supported by Science and Technology Project of SGCC (SGTJDK00DWJS1600014).

Authors' contributions

LH, YY and HZ mainly wrote the paper together, LH is responsible for the most of paper, including abstract, Part 1 Stochastic noise time series method in power load data and Part 3 Application and analysis of time series model and Kalman filtering individually. In addition, LH and YY are responsible for the revised manuscript, finishing the reviewers' comments. YY is responsible for the Part 2 Kalman filtering based on time series model and participates in the abstract with LH. HZ writes Part 0. Introduction and participates in the Part 4 Conclusion with YY. And helps to submit the revised manuscript in the web. XW gives advice on the paper structure and helps to check the whole paper's grammar. HZ helps to check the whole paper's grammar and words spelling. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests. And the authors certify that none of the material in the paper has been published or is under consideration for publication elsewhere.

Author details

¹NARI Technology Development CO., Ltd, Nanjing, Jiangsu Province 211106, China. ²State Grid Tianjin Electric Power Company, Tianjin 300010, China.

Received: 11 January 2017 Accepted: 30 June 2017

Published online: 11 July 2017

References

1. Qian, C. H. E. N., Wenying, H. U. A. N. G., Cheng, L. I., et al. (2008). Component based on-line modeling for electric loads [J]. *Automation of Electric Power Systems*, 32(2), 7–10.
2. Yang, G. A. O. (2009). *The white noise separation in "mechanism model + identification model" strategy for short time micro-grid load forecasting [D]*. Tianjin University.
3. Yundong, G. U., Sujie, Z. H. A. N. G., & Junshu, F. E. N. G. (2015). Multi-model fuzzy synthesis forecasting of electric power loads for larger consumers [J]. *Transactions of China Electrotechnical Society*, 30(23), 110–115.
4. Shuqing, Z. H. A. N. G., Rongyan, S. H. I., & Liguu, Z. H. A. N. G. (2016). Improvement of chaotic application in power forecasting model and its daily load forecasting [J]. *Chinese Journal of Scientific Instrument*, 37(1), 208–214.
5. Chuanping, X. I. O. N. G., Junjie, C. A. O., Qian, C. H. E. N., et al. (2011). A wavelet-based method for de-noising data of electric load modeling [J]. *Journal of Hohai University (Natural Sciences)*, 39(4), 470–473.
6. Bai, X. I. A. O., Chao, Z. H. O. U., & Gang, M. U. (2013). Review and Prospect of the spatial load forecasting methods [J]. *Proceedings of the CSEE*, 33(25), 78–92.

7. Bai, X. I. A. O., Xiao, X. U., Kun, S. O. N. G., et al. (2013). Abnormal data identification and treatment in spatial electric load forecasting [J]. *Journal of Northeast China Institute of Electric Power Engineering*, 33(1), 45–50.
8. Shan, G. A. O., & Yuanda, S. H. A. N. (2001). A new method of load data error-correction and smoothing based on wavelet singularity detection [J]. *Proceedings of the CSEE*, 21(11), 105–108 113.
9. Xinyao, S. U. N., Xue, W. A. N. G., Jiangwei, W. U., et al. (2014). Feature weighting based hierarchical probabilistic load forecasting in distributed collaborative network[J]. *Chinese Journal of Scientific Instrument*, 35(2), 241–246.
10. Xiaoling, S. H. E. N. (2009). *Research on power system short-term load forecasting approaches [D]*. Tianjin University.
11. Yuna, Z. H. A. N. G., Quana, Z. H. O. U., Haijunb, R. E. N., et al. (2013). Application and research of improved artificial immune network to power short-term load forecasting [J]. *Journal of Chongqing University (Natural Science Edition)*, 36(4), 33–38.
12. Xinran, L. I., Xuejiao, J. I. A. N. G., Jun, Q. I. A. N., et al. (2010). A classifying and synthesizing method of power consumer industry based on the daily load profile [J]. *Automation Of Electric Power Systems*, 34(10), 56–61.
13. Qi, W. A. N. G., Wenchao, Z. H. A. N. G., Yong, T. A. N. G., et al. (2010). A new load survey method and its application in component based load modeling [J]. *Power System Technology*, 34(2), 104–108.
14. LI Xiaojing CHEN Jiabin, YONG Shangguan. A method to analyse and eliminate stochastic noises of fog based on ARMA and kalman filtering method[C]. *Intelligent Human Machine Systems and Cybernetics (IHMSC)*. 2014.
15. Charles, A., Greenhall, A. D., et al. (1999). Total variance, an estimator of long-term frequency stability [J]. *IEEE Transactions on Ultrasonic, Ferroelectrics and Frequency Control*, 46(5), 1183–1191.
16. Chicco, G., Ionel, O. M., & Prumb, R. (2013). Electrical load pattern grouping based on centroid model with ant colony clustering [J]. *IEEE Transactions on Power Systems*, 28(2), 1706–1715.
17. Liqi, W. A. N. G., Huifang, G. H., Guibin, L. I., et al. (2014). Characteristic wavelength variable optimization of near-infrared spectroscopy based on Kalman filtering [J]. *Guang pu xue yu guang pu fen xi = Guang pu*, 34(4), 958.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
