TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 5545 ~ 5552
ISSN: 2302-4046

Gross Error Denoising Method for Slope Monitoring Data at Hydropower Station

Wei Hu[1], Xingguo Yang[*1], Jiawen Zhou[1], Huige Xing[2], Jian Xiang[3]
[1]State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University
Chengdu, Sichuan 610065, China, Ph./Fax: +86-28-85465055
[2]College of Architecture & Environment, Sichuan University
Chengdu, Sichuan 610065, China, Ph./Fax: +86-28-85402897
[3]Sino Hydro Bureau 7 CO., LTD, Sino Hydro Group LTD.
Chengdu, Sichuan 610081, China, Ph./Fax: +86-28-83371597
Corresponding author, e-mail: 18982176030@163.com[1], hgxing@scu.edu.cn[2],
xiangj@7j.sinohydro.com[3]

Abstract

There are mainly two types of errors existed in monitoring displacement of a rock slope: gross errors and random errors. Monitoring data is very important for the safety construction and operation of the Hydropower Station. The use of slope monitoring data for safety evaluation is influenced by the gross errors during the monitoring process. This paper presents a gross error denosing method for a nonlinear time series based on the three-standard-deviation rule (3-σ rule), and then reconstructing the time series by a first-order Lagrange interpolation method. The present method is applied to the gross error analysis of the slope displacement monitoring data collected at the Jinping I Hydropower Station. Computed results show that the first-order difference values of the gross errors can be above or below the upper or lower three-standard-deviation boundary, and the gross errors can be removed effectively.

Keywords: *slope, nonlinear time series, gross error, data denosing, three-standard-deviation rule*

1. Introduction

With the rapid growth in energy demand, lots of Hydropower Stations are constructed or planned in future [1]. Technical progress has promoted the rapid development of geotechnical engineering, and it has also brought unprecedented high slope stability problems, especially for large hydropower stations [2]-[4]. Therefore, safety monitoring of high and steep slopes has become a key technological problem in geotechnical engineering [3]. Data analysis is a key element of safety monitoring; it includes data preprocessing, forecasting, and early warning. There are varieties of processing methods for noisy time series, such as information theoretic, adaptive filtering, dynamical systems and stochastic approaches, and wavelet transform [4]-[5]. The reliability of a slope stability evaluation is determined by the accuracy of the monitoring data. Errors in the monitoring data occur, and the data should be denoised to remove the errors [6]-[8]. High slope-safety monitoring data are in the form of time series, the study of which has many aspects, such as error analysis, nonlinear prediction, back analysis and comprehensive analysis.

During the monitoring process for a slope at Hydropower Station, several types of errors are existed in a nonlinear time series, gross error and random error are the two main errors impact on the monitoring data [9], all of these errors should be removed before the data is used for safety evaluation of slope, and the gross error is the most important one [10]. There are varieties of processing methods for noisy time series, such as information theoretic, adaptive filtering, dynamical systems and stochastic approaches, and wavelet transform [10]-[13]. Previous methods have mostly used unilateral reduction of gross errors or random errors, and the reconstructed signals still contain a great number of errors that affect the accuracy of subsequent analytical results [14]-[17]. In the present paper, the three-deviation rule for a difference sequence is applied for gross error identification and modification. It is based on statistical theory for data processing and error evaluation, and the gross error can be removed effectively.

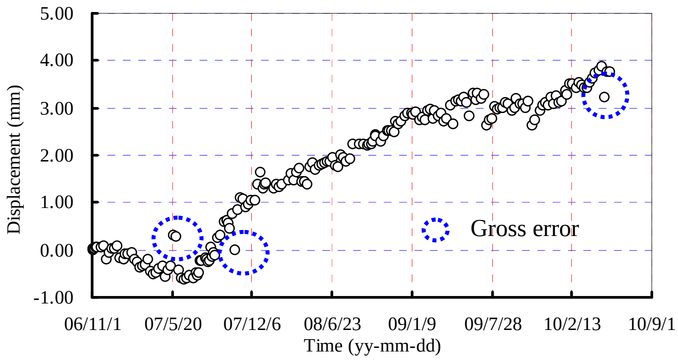
Received April 19, 2013; Revised June 21, 2013; Accepted July 9, 2013

2. Gross Error of Slope Monitoring Displacement

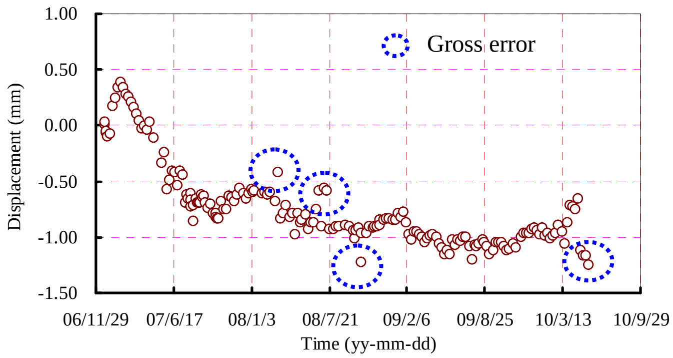
According to error theory for measurement data, the errors in monitoring data can be divided into three types: gross error, random error, and systematic error. Gross error is usually caused by an observational mistake, with a sudden, single outlier. When the monitoring time series is obtained under the same observation conditions, random error (also called accidental error) exists occasionally in the size; its impact can be reduced effectively during data processing as the measured number increases. Systematic error shows systematic characteristics in size and symbol; its impact can be reduced generally through timing calibration apparatus and a calibration datum mark [6].

In addition, some abnormal deformation values exist in the monitoring time series because slope deformation of soil or rock mass is influenced by the change in external loading and environmental variable mutation. An abnormal deformation value is different from the gross error [11]. It has potential and valuable information for the slope stability analysis and should be retained. Therefore, according to the nonlinear time series of the data, the gross error denoising process of displacement monitoring data during error analysis is: identification of the gross errors and abnormal deformation values and removal of gross errors.

Figure 1 shows a slope displacement time series from monitoring points M^4_4 and M^4_6 at the Jinping I Hydropower Station.



(a) monitoring point M^4_4



(b) monitoring point M^4_6

Figure 1. Gross errors exist in slope displacement time series

As shown in Figure 1, gross errors exist in the nonlinear time series of monitoring data of slope displacement. The time series are distorted by the gross errors during the monitoring process. The denoising effect is influenced by the gross errors, so they should be removed first, and then the smoothing of the nonlinear time series.

3. Gross Error Denoising Method

Tendency and non-stationary characteristics exist in the nonlinear time series of displacement monitoring data. Here a difference sequence method is applied to remove the variation trend of a time series and transform it into a stationary time series [17]. For the difference method, the first- or higher-order difference is computed for the non-stationary time series until it becomes a smooth difference sequence, and then the gross errors are detected from the smooth difference sequence.

According to the slope monitoring displacement data, the original nonlinear time series [11] obtained by monitoring displacement is assumed to be as follows:

$$[x_i] = \{x_1, x_2, \cdots, x_i, \cdots, x_n\}, \quad (i = 1,2,\cdots,n) \tag{1}$$

where *i* is the number of the monitoring displacement data values.

The first-order difference equation for one monitoring datum x_i of the nonlinear time series is as follows:

$$\nabla x_i = x_i - x_{i-1} \tag{2}$$

where ∇ is the first order difference, x_i is the monitoring displacement of point *i*, and x_{i-1} is the monitoring displacement of point *i*-1.

The k-1 order difference equation for one monitoring datum x_i of the nonlinear time series is as follows:

$$\nabla^k x_i = \nabla^{k-1} x_i - \nabla^{k-1} x_{i-1} \tag{3}$$

where ∇^k is the k order difference, and ∇^{k-1} is the *k*-1 order difference.

Then the first-order difference sequence and *k*-order difference sequence can be obtained:

$$[\nabla x_i] = \{\nabla x_1, \nabla x_2, \cdots, \nabla x_i, \cdots, \nabla x_n\}, \text{ (First order)} \tag{4}$$

$$[\nabla^k x_i] = \{\nabla^k x_1, \nabla^k x_2, \cdots, \nabla^k x_i, \cdots, \nabla^k x_n\}, \text{ (}k\text{ order)} \tag{5}$$

The gross errors in the nonlinear time series of displacement monitoring are detected from the difference sequence.

The gross errors in the nonlinear time series are detected by applying the three-standard-deviation rule (3-σ rule). The 3-σ rule is based on the hypothesis that a group of normal distribution monitoring data contains random error and gross error and that there is a standard deviation σ. Then an interval of $\pm 3\sigma$ is defined so that 99.74% of the data is expected to lie within this interval. If a data point lies outside this interval, then it is assumed as a gross error, should be smoothed.

For the first-order difference sequence of displacement monitoring data, when the 3-σ rule is applied to detect gross errors, the average value $\overline{\nabla x}$ of the difference values ∇x_i is substituted for the true value, and the standard deviation S calculated by the Bessel formula is substituted for σ. First the residual error is computed as follows:

$$\nabla v_i = \nabla x_i - \overline{\nabla x} \tag{6}$$

where ∇v_i is the residual error.

The residual error of suspicious monitoring data should satisfy the following equation:

$$|\nabla v_i| = |\nabla x_i - \overline{\nabla x}| > 3S, \quad (i = 1,2,\cdots,n) \tag{7}$$

where $\overline{\nabla x}$ is the mean value of the difference sequence and can be computed as follows:

$$\overline{\nabla x} = \frac{\nabla x_1 + \nabla x_2 + \cdots + \nabla x_n}{n} \tag{8}$$

where S is the standard deviation calculated by the Bessel formula:

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\nabla x_i - \overline{\nabla x}\right)} \tag{9}$$

As shown in Eq. (7), the residual error of suspicious monitoring data is computed from the first-order difference sequence, and the gross error may be detected by the 3-σ rule. The interval for removing the gross error of the difference sequence is as follows:

$$\left[E_{\min}, E_{\max}\right] = \left[\overline{\nabla x} - 3S, \overline{\nabla x} + 3S\right] \tag{10}$$

where $E_{\min}$ is the minimum vale of the error, and $E_{\max}$ is the minimum vale of the error.

The gross errors in the nonlinear time series from displacement monitoring can be detected by the 3-σ rule based on the first-order difference sequence, especially when the sample number is larger than 50, which occurs especially for long historical records from monitoring displacement of slope. If the gross errors can not be detected based on the first-order difference sequence, a k-order difference sequence should be used to apply the 3-σ rule, such as a second-order or third-order difference sequence [17].

After the data affected by gross errors have been removed from the nonlinear time series, the time series is no longer continuous, so repair becomes necessary. A polynomial interpolation method is applied to repair the monitoring data time series. The polynomial interpolation method includes two types: the first-order Lagrange interpolation method and the quadratic Lagrange interpolation method. In the present study, the first-order Lagrange interpolation method is used.

If the coordinates of two points near the interpolated point are (t_1, x_1) and (t_2, x_2), according to the first-order Lagrange interpolation method, then the coordinates of the interpolation point are obtained as follows:

$$x = \frac{t - t_2}{t_1 - t_2} x_1 + \frac{t - t_1}{t_2 - t_1} x_2 \tag{11}$$

where x is the displacement value of the interpolation point, and t is the monitoring time of the interpolation point.

In summary, the gross errors in the nonlinear time series of displacement monitoring data are detected by the 3-σ rule, and the time series is repaired by the first-order Lagrange interpolation method. The continuity of the time series is thereby restored by the interpolation points.

4. Validation

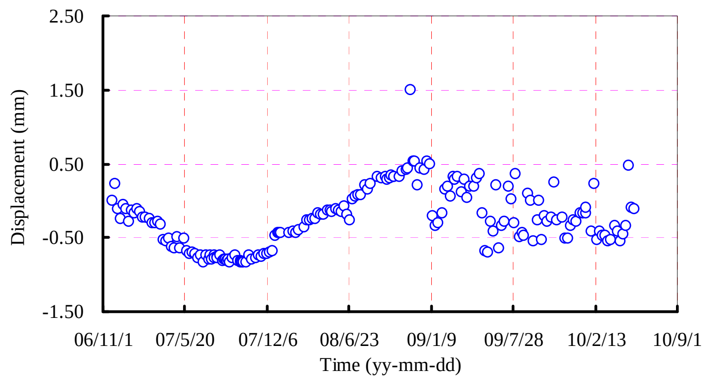
In this section, the nonlinear time series of displacement monitoring of the left bank slope at the Jinping I Hydropower Station are used to verify the present data denoising method. The method is implemented by Matlab 6.5 professional software. The gross errors in the nonlinear time series are detected by the 3-σ rule and repaired by the first-order Lagrange interpolation method.

Jinping I Hydropower Station is located at the big bend of the Yalong River, Muli County, Liangshan city, Sichuan province, southwest China. An arch dam 305 m high will be built and will be the highest dam in the world. A large amount of slope excavation is carried out for the construction of the dam, especially on the left bank slope. The maximum excavation height is about 540 m, and the excavation volume is about 5.5 million m^3. The slope stability is influenced by the poor geological conditions and high in-situ stress, so that substantial monitoring measures are planned for the safe control of the slope during excavation. Monitoring displacement is a real reflection of the stability status of slope, but its influenced by several errors.

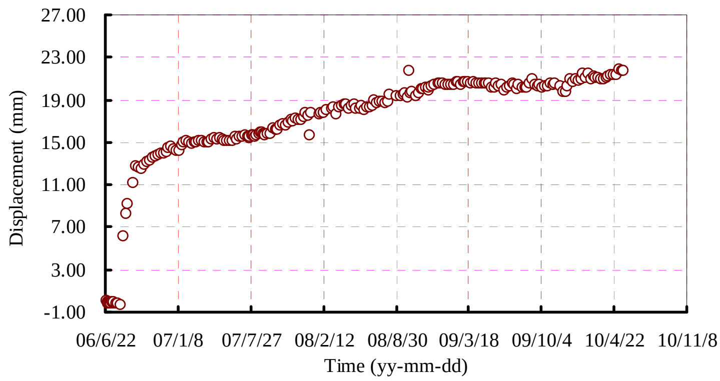
A large amount of slope excavation is carried out for the construction of the dam, especially on the left bank slope. The slope stability is influenced by the poor geological conditions and high in-situ stress, so that substantial monitoring measures are planned for the safe control of the slope during excavation. The slope deformation monitoring includes internal deformation monitoring, appearance deformation monitoring (surface deformation), and crack opening degree monitoring. Here we select two monitoring points, M^4_1 and M^4_3, which are near the horizontal monitoring section EL. 1990 m (EL. is the elevation).

During displacement monitoring, the displacement is influenced by external factors, such as construction disturbance, instrument precision, and external environment. Some errors exist in the monitoring data, and if the monitoring data were directly applied to slope stability evaluation, forecasting and early warning. Therefore, data preprocessing is the prerequisite for data analysis calculations.

Figure 2 shows the nonlinear time series of displacement monitoring at the points M^4_1 and M^4_3.



(a) monitoring point M^4_1.

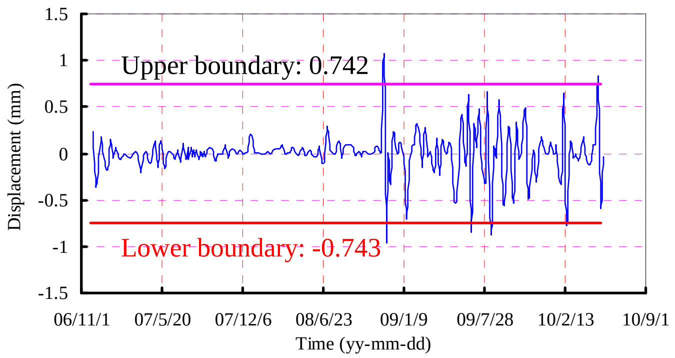


(b) monitoring point M^4_3.

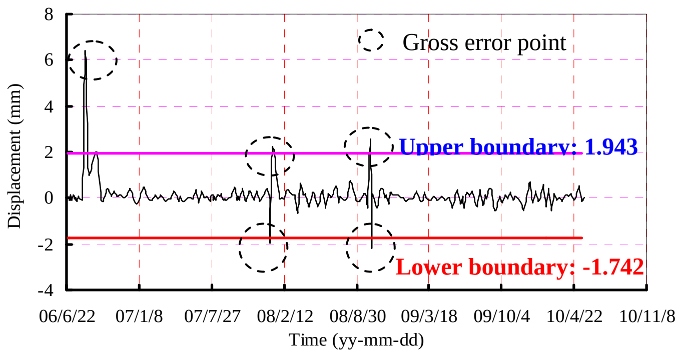
Figure 2. Nonlinear time series of monitoring displacement in slope

As shown in Figure 2(a), in the end of the nonlinear time serie, the data is distorted by serveral factors, the gross errors are hard to detected. As shown in Figure 2(b), the accumulative displacement at point M^4_3 slowly grows with the time, a characteristic of a typical stable time series with few abnormal data. Gross errors exist in these two time series, so they should be denoised first. Here the 3-σ rule is used to detect the gross errors, and then the time series is repaired by the first-order Lagrange interpolation method.

Firstly, the first-order difference of the original time series is computed. Figure 3 shows the first-order difference of the time series of displacement and the three-standard-deviation boundary of gross errors.



(a) monitoring point M^4_1



(b) monitoring point M^4_3

Figure 3. First-order difference of nonlinear time series

There is no obvious increasing tendency for the data distribution of the first-order difference series, so that the gross errors can be detected by the 3-*σ* rule. Table I shows the computed results of the first-order difference sequence by application of the 3-*σ* rule.

Table 1. Computed results of the first-order difference series by the 3-*σ* rule

Monitoring point	Mean value	Standard deviation of FODS sample	Interval of FODS sample
M^4_1	0.00	0.25	[-0.74,0.74]
M^4_3	0.10	0.61	[-1.74,1.94]

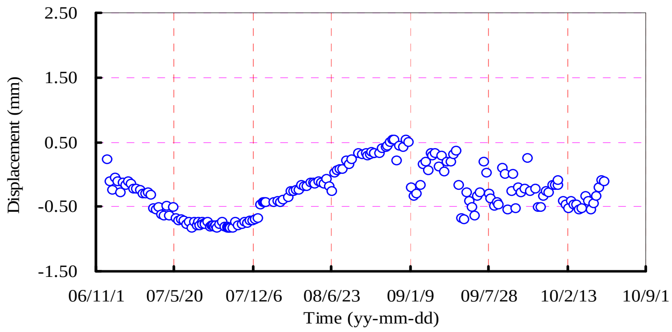
(FODS is the first order difference series; 3-*σ* is the three-standard deviation).

As shown in Table 1, the mean values of the first-order difference series for monitoring points M^4_1 and M^4_3 are 0.00 and 0.10, respectively; the standard deviations are 0.25 and 0.61, respectively. As shown in Fig. 3, the upper and lower three-standard-deviation boundaries of monitoring point M^4_1 are 0.742 and -0.743, respectively; the upper and lower three-standard-deviation boundaries of monitoring point M^4_3 are 1.943 and -1.742, respectively. There are 5 gross errors in the time series of monitoring point M^4_1. The observation dates are November 21, 2008, June 18, 2009, August 4, 2009, February 10, 2010 and May 6, 2010. There are also 5 gross errors in the time series of monitoring point M^4_3. The observation dates are August 12, 2006, January 6, 2008, January 10, 2008, October 6, 2008 and October 11, 2008. Table 2 shows the first-order difference values of the gross error points in the time series of monitoring point M^4_3.

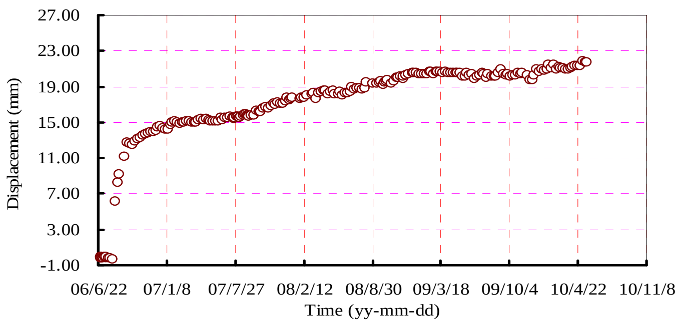
Table 2. First order difference values of the gross error points existing in the time series

Time (yy-mm-dd)	Difference value of the gross error point	Difference value of the former point	Difference value of the latter point
06-8-12	6.42	-0.12	1.90
08-1-6	-1.93	-0.24	2.14
08-1-10	2.14	-1.93	-0.04
08-10-6	2.55	-0.39	-2.12
08-10-11	-2.12	2.55	0.13

As shown in Table 2, the first-order difference values of the gross error points are above the upper or below the lower three-standard-deviation boundary. The gross error points exhibit a jump characteristic and should be removed from the time series. The gross errors must be removed and the time series reconstructed. In this paper, the first-order Lagrange interpolation method is applied to repair the time series. Figure 4 shows the reconstructed time series obtained after this processing.



(a) monitoring point M^4_1



(b) monitoring point M^4_3

Figure 4. Reconstructed time series by the first-order Lagrange interpolation method

As shown in Figure 4, the gross errors in the time series have been removed, and the time series is smoother, the gross errors existed in the nonlinear time series are removed effectively.

5. Conclusion

During safety monitoring of high rock slopes, abnormal information or errors exist in the time series of displacement monitoring data due to the influence of external objective conditions, such as construction conditions, climate factors, and monitoring instruments. There are several errors existed in the nonlinear time series of monitoring data. Gross error is usually caused by observation mistake and impact on the safety evaluation of slope very obviously. In this paper, a gross error denoising method is presented, the gross errors in the nonlinear time series are detected by the three-standard-deviation rule (3-σ rule), and then the time series is reconstructed by the first-order Lagrange interpolation method.

The nonlinear time series from monitoring of the displacement of the left bank slope at the Jinping I Hydropower Station is used to verify the present data denoising method. First the gross errors are detected by the 3-σ rule. The gross error points show a jump characteristic and should be removed from the time series. The computed results show that the first-order difference values of the gross errors are either above or below the upper or lower three-standard-deviation boundary. Then the nonlinear time series is repaired by the first-order Lagrange interpolation method, the gross errors are removed effectively.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 41102194, 51209156) and the China Postdoctoral Science Foundation (No. 2012T50785).

References

[1] Papadimitfiou S, Bezerianos A. Multiresolution analysis and denoising of computer performance evaluation data with the wavelet transform. *Journal of Systems Architecture*. 1996; 42(1): 55-65.
[2] Basu B, Nagarajaiah S, Chakraborty A. Online dentification of linear time-varying stiffness of structural systems by wavelet analysis. *Structural Health Monitoring*. 2008; 7(1): 21-36.
[3] Jones AJ, Evans D, Kemp SE. A note on the Gamma test analysis of noisy input/output data and noisy time series. *Physica D: Nonlinear Phenomena*. 2007; 229(1): 1-8.
[4] Ghanem R. Romeo F. A wavelet-based approach for the identification of linear time-varying dynamical systems. *Journal of Sound and Vibration*. 2000; 234(4): 555-576.
[5] Yen EC. Solubility and stability of recurrent neural networks with nonlinearity or time-varying delays. *Communications in Nonlinear Science and Numerical Simulation*. 2011; 16(1): 509-521.
[6] Maus A, Sprott JC. Neural network method for determining embedding dimension of a time series. *Communications in Nonlinear Science and Numerical Simulation*. 2011; 16(8): 3294-3302.
[7] Harikrishnan KP, Misra R, Ambika G. Combined use of correlation dimension and entropy as discriminating measures for time series analysis. *Communications in Nonlinear Science and Numerical Simulation*. 2009; 14(9-10): 3608-3614.
[8] Chang CC, Shi YF. Identification of time-varying hysteretic structures using wavelet multiresolution analysis. *International Journal of Non-Linear Mechanics*. 2010; 45(1): 21-34.
[9] Jayawardena AW, Gurung AB. Noise reduction and prediction of hydrometeorological time series: dynamical systems approach vs. stochastic approach. *Journal of Hydrology*. 2000; 228(3-4): 242-264.
[10] Kotnik B, KaČiČ Z. A noise robust feature extraction algorithm using joint wavelet packet subband decomposition and AR modeling of speech signals. *Signal Processing*. 2007; 87(6): 1202-1223.
[11] Zhong S, Oyadiji SO. Crack detection in simply supported beams without baseline modal parameters by stationary wavelet transform. *Mechanical Systems and Signal Processing*. 2007; 21(4): 1853-1884.
[12] Kumar K, Jain VK. Autoregressive integrated moving averages (ARIMA) modelling of a track noise time series. *Applied Acoustics*. 1999; 58(3): 283-294.
[13] Dongqing Zhang, Yubing Han, Xueyu Tang. Nonlinear/NonGaussian Time Series Prediction Based on RBFHMMGMM Model. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 2012, 10(6): 1214-1226.
[14] Coca D, Billings SA. Non-linear system identification using wavelet multi-resolution models. *International Journal of Control*. 2001; 74(18): 1718-1736.
[15] Tu CC, Juang CF. Recurrent type-2 fuzzy neural network using Haar wavelet energy and entropy features for speech detection in noisy environments. *Expert Systems with Application*s. 2012; 39(3): 2479-2488.
[16] Daubechies I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*. 1990; 36(5): 961-1005.
[17] Bin Fang, Xiaoqi Guo. Modified Allan Variance Analysis on Random Errors of MINS. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(3): 1227-1235.