Metabolic grey early warning model for dam deformation based on wavelet denoising

WU Yunxing1,2,a, GU Yanchang1,2
1Nanjing Hydraulic Research Institute, Nanjing 210029, China
2Dam Safety Management Center of the Ministry of Water Resources, PRC, Nanjing 210029, China

Abstract. Influenced by environment and human factors, the observed data of dam deformation consist of real deformation value and observation error (noise). The conventional GM(1,1) model based on non-denoised observation data is not very effective. In order to improve the prediction effect of conventional GM(1,1) model, wavelet threshold denoising method is used to eliminate the noise in the original data and improve the smoothness of the data sequence. Then, based on the conventional GM(1,1) model, the metabolic GM(1,1) model is established by eliminating the oldest information and adding the newest information. The application results show that the wavelet threshold denoising can obviously remove the noise from the original data. The predicted vertical displacement of the metabolic GM(1,1) model based on the denoised data has little difference with the measured value, and the predicted precision is obviously higher than that of the conventional GM (1,1) model. Therefore, the metabolic GM(1,1) model based on wavelet denoising can be used for prediction and early warning of dam deformation.

1 Introduction

Surface deformation monitoring is an important part of dam safety monitoring, and it is of great significance for mastering dam operation behavior and safety condition. How to correctly analyze the dam deformation monitoring data and predict the future deformation trend is also a hot topic. At present, there have been many data analysis methods for dam safety monitoring, such as statistical model, support vector machine model, neural network model, grey system theory and so on. However, each method has its advantages and disadvantages. Among them, the grey system theory is suitable for analyzing the uncertain system of "small sample, poor information". Through excavating some known information, the valuable information can be extracted, and the correct description and effective monitoring of the system behavior and evolution law can be realized. Technical specification for earth-rockfill dam safety monitoring (SL 551-2012) stipulates that the frequency of surface deformation monitoring of earth-rock dams is 2-6 times per year. But due to the lack of professional technicians and financial constraints, surface deformation monitoring of many reservoirs is carried out only once a year or once a few years, which results in insufficient data for dam deformation monitoring. Therefore, it is difficult to fully and accurately reflect the deformation law of the dam, which brings about difficulties to dam deformation prediction and early warning. However, the grey system theory can well solve this problem. The grey GM (1,1) model, which has the advantages of simple operation and high modeling accuracy, can be established with a small amount of dam deformation monitoring data to accurately predict the future deformation trend of the dam. However, the predictive effect of GM (1,1) model is closely related to the smoothness of the data sequence. Because the dam deformation observation will be affected by weather and human factors, the observed data is not the real deformation value but the data containing noise (expressed as error). In addition, the grey GM (1,1) model only has a higher prediction accuracy for the 1-2 period data after the modeling data, and the later the time, the worse the accuracy.

Based on the above two reasons, this paper intends to adopt wavelet threshold denoising method to denoise the original dam deformation observation data and extract the deformation value close to the real dam so that the data sequence is smoother. Then, based on the conventional GM (1,1) model, the metabolic GM (1,1) model is established by using the denoised data to realize the prediction of dam deformation and improve the prediction accuracy.

2 Wavelet threshold denoising method

In dam deformation observation, due to the influence of surrounding environment and human factors, certain errors often exist in the deformation observation data, that is, deformation observation data consists of real signals and observation errors called "noise"[2]. Because noise interferes with the analysis of observed data, it is necessary to denoise the observed data before analysis. Wavelet threshold denoising method proposed by
Donoho and Johnstone has been widely used in signal denoising at present. The basic idea of wavelet threshold denoising method is that the wavelet coefficients of real signals become larger, and the wavelet coefficients of noises are smaller than those of real signals after wavelet transform. Then a suitable threshold is set. The wavelet coefficients which are larger than the threshold are considered to be generated by the real signal and are retained, while the wavelet coefficients which are smaller than the threshold are considered to be generated by noise and set to zero so as to achieve the purpose of denoising.

The steps of wavelet threshold denoising are as follows [3]:

1. Wavelet decomposition: appropriate wavelet function is selected and the decomposition level is confirmed. The decomposition coefficients of all levels are obtained via using the wavelet decomposition to analysis the original data.

2. Threshold processing: the decomposition coefficients of all levels are quantified after selecting appropriate thresholds and threshold function.

3. Wavelet reconstruction: optimal estimation of real signal is acquired after reconstructing the coefficients.

The effectiveness of threshold denoising is measure via RMSE, BIAS and SNR. The smaller RMSE and BIAS and the larger SNR, the better denoising result.

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (f(i) - \hat{f}(i))^2} \]  
\[ BIAS = \frac{1}{n} \sum_{i=1}^{n} (f(i) - \hat{f}(i)) \]  
\[ SNR = 10 \times \log\left(\frac{\text{power}_{\text{signal}}}{\text{power}_{\text{noise}}}\right) \]  
\[ \text{power}_{\text{signal}} = \frac{1}{n} \sum_{i=1}^{n} f^2(i) \]  
\[ \text{power}_{\text{noise}} = \frac{1}{n} \sum_{i=1}^{n} \left| f(i) - \hat{f}(i) \right|^2 \]  

where \( f(i) \) represents the original observed data; \( \hat{f}(i) \) represents the denoised data; \( \text{power}_{\text{signal}} \) is energy value of original date and \( \text{power}_{\text{noise}} \) is energy value of the noise.

### 3 Principle of the metabolic GM(1,1) model

#### 3.1 Conventional GM(1,1) model

As the basic model of grey prediction model, the model principle of GM(1,1) model is as follows [4]:

The non-negative straggling equidistant data sequence is

\[ x^{(0)} = \{ x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n) \} \]

After an accumulation of \( x^{(0)} \), the result is obtained:

\[ x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i); k = 1, 2, \ldots, n \]  

The new date sequence is

\[ x^{(1)} = \{ x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n) \} \]

The albinism differential equation of GM(1,1) is constructed according to the sequence \( x^{(1)} \).

\[ \frac{dx^{(1)}}{dt} + ax^{(1)} = u \]  

where \( a \) is development parameter, which reflects the developing trend of \( x \); \( u \) represents the grey action, revealing the changes in the relationship between data.

Winterization equation satisfies the initial condition \( t = t_0 \) and solution of \( x^{(1)} = x^{(1)}(t_0) \) is

\[ x^{(1)}(t) = \left[ x^{(1)}(t_0) - \frac{u}{a} \right] e^{-at} + \frac{u}{a} \]  

The discrete value \( x^{(1)}(t_0+1) \) is normalized as follows:

\[ x^{(1)}(k+1) = \left[ x^{(1)}(1) - \frac{u}{a} \right] e^{-ak} + \frac{u}{a} \]  

The coefficients \( a \) and \( u \) are estimated by least square method:

\[ \hat{U} = \left[ \hat{a} \right] = \left( B^T B \right)^{-1} B^T Y \]  

Inhere,

\[ B = \begin{bmatrix} -\frac{1}{2} (x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2} (x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2} (x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \]

The estimated values \( \hat{a} \) and \( \hat{u} \) are put into the equation (9) and its result is as follows:

\[ \hat{x}^{(1)}(k+1) = \left[ x^{(1)}(1) - \frac{\hat{u}}{\hat{a}} \right] e^{-ak} + \frac{\hat{u}}{\hat{a}} \]  

The result is obtained via restoring original data as follows:

\[ \hat{x}^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k) = \left( 1 - e^{-\hat{a}} \right) \left[ x^{(1)}(1) - \frac{\hat{u}}{\hat{a}} \right] e^{-ak} \]  

When \( k = 1, 2, \ldots, n \), the fitted value of original sequence \( x^{(0)} \) are obtained; when \( k = n \), the estimated value of original sequence \( x^{(0)} \) are obtained.
3.2 Metabolic GM(1,1) model

With time, some disturbance factors enter the grey system, which affect the development of the system. Although conventional GM(1,1) model has higher prediction accuracy of the first and second period date after original sequence, its prediction accuracy becomes worse since the effect of old data to describe systematic evolution is lessen with the development of the system. In order to reflect the future effect of grey system due to stochastic disturbance, the new grey metabolic GM(1,1) model is established, which continues to add the new information and delete the old information, thus assisting the model sequence to reflect the current features of the system better and improving the prediction accuracy.

The principle of metabolic GM(1,1) model is as follows: delete the oldest information \( \{x^{(0)}(1), \ldots, x^{(0)}(n)\} \) of the original sequence \( \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\} \); at the same time, add the newest information \( x^{(0)}(n+1) \); then GM(1,1) model is built by using the new data sequence \( \{x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n+1)\} \) and prediction value at \( n+2 \) is obtained; after that, divide \( x^{(0)}(2) \) by the obtained result, add \( x^{(0)}(n+2) \) and use \( \{x^{(0)}(3), x^{(0)}(4), \ldots, x^{(0)}(n+2)\} \) to construct the new GM(1,1) model and predict the value at \( n+3 \). Such repeated process is called the same dimension gray recurrence dynamic model, namely grey metabolic GM(1,1) model.

3.3 Accuracy test

The accuracy of grey prediction model is tested via residual qualified model and the testing equation is as follows:

\[
e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \tag{13}
\]

\[
\Delta_i = \left| \frac{e(k)}{x^{(0)}(k)} \right| \tag{14}
\]

\[
\overline{\Delta} = \frac{1}{n} \sum_{k=1}^{n} \Delta_i \tag{15}
\]

where \( e(k) \) is residual error; \( \Delta_i \) is relative error; \( \overline{\Delta} \) is average relative error.

When the conditions \( \overline{\Delta} < \alpha \) and \( \Delta_i < \alpha \) are satisfied, the residual error is qualified given \( \alpha \). The accurate levels of model are showed in Table 1.

Table 1. Model accuracy level comparison

| Prediction accuracy grade | \( \alpha \) |
|---------------------------|-------------|
| Frist level               | 0.01        |
| Second level              | 0.05        |
| Third level               | 0.10        |
| Fourth level              | 0.20        |

4 Application of the model

To verify the feasibility of the metabolic GM(1,1) model based on wavelet denoising proposed in this paper, ten periods vertical displacement data from deformation measuring point in Table 3 are used to build model and make the prediction.

4.1 Data denoising processing

The effect of wavelet denoising is closely related to wavelet basis function, decomposition level, the setting of threshold value and the selection of threshold function [5]. After reading relative literature, it is found that wavelet functions sym8 and db4 are widely used in practice and heursure and soft threshold function are usually applied in the setting of threshold value and the selection of threshold function respectively. The great effect results of decomposition layer-number are selected on the basis of sym8 and db4 in the paper shown in Table 2.

From Table 2, it is found that wavelet functions sym8 has achieved a better denoising effect than db4 and the effect of sym8 is decreased with the increase of the decomposition layer-number. When the decomposition layer-number is 2, the best denoising effect of sym8 is acquired. The denoising processing is finished via using heursure and soft threshold function, whose denoised vertical displacement data could be seen in Table 3.

4.2 Construction of prediction model

The conventional GM(1,1) model is constructed by using vertical displacement data before denoising in the first 7-period and denoised vertical displacement data respectively. Prediction error of conventional GM (1,1) model is shown in Table 4.
From Table 4, it is found that the prediction value of conventional GM(1,1) model after wavelet denoising is closer to the original value and the prediction average relative error is 1.53%. Additionally, the prediction accuracy is obviously higher than the conventional GM(1,1) model using the data before denoising. However, as the prediction period increases, the prediction accuracy of the conventional GM(1,1) model using denoised data obviously appears declining. Therefore, to predict the deformation more accurately, the metabolic GM(1,1) model is established.

The conventional GM(1,1) model is built based on the denoised data from 1 to 7 period; the metabolic GM(1,1) model I is constructed via using the denoised data from 2-8 period; the metabolic GM(1,1) II model is established by adopting the denoised data from 3-9 period.

(1) conventional GM(1,1) model
   
   The model sequence is: \( x^{(0)} = \{83.08, 83.94, 85.21, 86.68, 88.23, 89.81, 91.27\} \).  
   
   When the estimated parameter values \( \hat{\beta} = -0.016983 \) and \( \hat{\alpha} = 81.728498 \), the time response equation can be expressed as:
   
   \[
   x^{(0)}(k + 1) = 4895.4511 \left(1 - e^{-0.016983} \right)e^{0.016983k} \tag{16}
   \]

(2) Metabolic GM (1,1) Model I
   
   Excluding the first-period vertical displacement observation data, and adding the eighth-period vertical displacement observation data, the modeling sequence can be expressed as:
   
   \( x^{(8)} = \{83.94, 85.21, 86.68, 88.23, 89.81, 91.27, 92.49\} \)
   
   When the estimated parameter values \( \hat{\beta} = -0.016609 \) and \( \hat{\alpha} = 83.194650 \), the time response equation can be expressed as:
   
   \[
   x^{(8)}(k + 1) = 5092.9502 \left(1 - e^{-0.016609} \right)e^{0.016609k} \tag{17}
   \]

(3) Metabolic GM (1,1) Model II
   
   Subsequently, excluding the second-period vertical displacement observation data, and adding the ninth-period vertical displacement observation data, the modeling sequence can be expressed as:
   
   \( x^{(9)} = \{85.21, 86.68, 88.23, 89.81, 91.27, 92.49, 93.36\} \)
   
   When the estimated parameter values \( \hat{\beta} = -0.015040 \) and \( \hat{\alpha} = 85.011171 \), the time response equation can be expressed as:
   
   \[
   x^{(9)}(k + 1) = 5737.5485 \left(1 - e^{-0.015040} \right)e^{0.015040k} \tag{18}
   \]

### 4.3 Analysis of prediction results

According to the Conventional GM (1,1) model, the Metabolic GM (1,1) model I and the Metabolic GM (1,1) model II, the vertical displacements of 8th to 10th periods were predicted. The prediction results of each model are shown in Table 5. The residual accuracy model was used to evaluate the prediction accuracy level of each model. The evaluation results are shown in Table 6.

### Table 4. Prediction error of conventional GM (1,1) model

| Observation period | Original / mm | Predicted value / mm | Residual / mm | Relative error (%) | Mean relative error (%) | Denoised |
|-------------------|----------------|----------------------|--------------|--------------------|------------------------|----------|
|                    |                |                      |              |                    |                        |          |
| 8                 | 92.32          | 92.20                | -0.88        | 0.95               | 92.84 -0.52            | 0.56     |
| 9                 | 93.03          | 94.89                | -1.86        | 2.00               | 94.43 -1.40            | 1.50     |
| 10                | 93.68          | 96.61                | -2.93        | 3.13               | 96.05 -2.37            | 2.53     |

### Table 5. Prediction results of each model

| Observation period | Conventional GM (1,1) model | Metabolic GM (1,1) model I | Metabolic GM (1,1) model II |
|-------------------|-----------------------------|----------------------------|----------------------------|
|                    | Predicted value / mm | Residual / mm | Relative error (%) | Predicted value / mm | Residual / mm | Relative error (%) | Predicted value / mm | Residual / mm | Relative error (%) |
| 8                 | 92.32          | 92.84                | -0.52            | 0.56               | 92.68 -0.36 | 0.39             | 92.34 -0.02 | 0.02          |
| 9                 | 93.03          | 94.43                | -1.40            | 1.50               | 94.23 -1.20 | 1.29             | 93.71 -0.70 | 0.75          |
| 10                | 93.68          | 96.05                | -2.37            | 2.53               | 95.81 -2.13 | 2.27             | 95.15 -1.47 | 1.57          |

### Table 6. Prediction accuracy grade of each model

| Model                  | Conventional GM (1,1) model | Metabolic GM (1,1) model I | Metabolic GM (1,1) model II |
|------------------------|-----------------------------|-----------------------------|-----------------------------|
| Mean relative error/%  | 1.53                        | 1.32                        | 0.78                        |
| Prediction accuracy grade | Second level               | Second level               | First level                 |

According to Table 5 and Table 6, the predicted values of the Metabolic GM (1,1) Models are closer to the original measured value than those of the Conventional GM(1,1) Model. Besides, the prediction average relative error of the Metabolic GM (1,1) Models is also significantly reduced. In addition, with the continuous establishment of a new metabolic GM (1,1) model, the prediction accuracy level of the new model is also improved. To be specific, the accuracy level is improved from the second level of the Metabolic GM (1,1) Model I to the first level of the Metabolic GM (1,1) Model II, and the model accuracy level meets the requirements.

### 5 Conclusion

(1) The observation data is processed by the wavelet threshold denoising method, which effectively removes the noise in the observation data and improves the smoothness of the data sequence.

(2) The conventional GM(1,1) models were established by using the data without denoising and the denoised data, respectively. The prediction results show that the conventional GM(1,1) model established by using denoised data has better prediction effect than that established with noise data.

(3) Using the denoised data to establish a metabolic GM (1,1) model, the prediction accuracy is significantly improved compared with the conventional GM (1,1) model, indicating that the metabolic GM (1,1) model based on wavelet denoising is practical and feasible in the prediction and warning of dam deformation.
Acknowledgment

Fund project: Supported by national key research and development projects (2016YFC0401608); Key projects of special funds for basic research and business expenses of public welfare research institutes at the central level of Nanjing Hydraulic Research Institute (Y717007, Y716014); Construction project of scientific research innovation team of special funds for basic research and business expenses of public welfare research institutes at the central level of Nanjing Hydraulic Research Institute (Y717012).

References

1. SL 551-2012 Earth and Rock Dam Safety Monitoring Technical Specification [S]. Beijing: China Water Resources and Hydropower Press, 2012.

2. Liang Xueqing, Guo Xianglin, Pan Liangjun. Application of Grey Theory Based on Wavelet in Settlement Data[J]. Energy and Environment Protection, 2017, 39(9):140-144.

3. Cui Naidan, Xiang Wanli, Meng Xuelei, et al. Research on freight volume forecast based on wavelet gray GM(1,1) model[J]. Journal of Railway Science and Engineering, 2017, 14(11): 2480-2486.

4. Liu Sifeng. Grey System Theory and Its Application [M]. Beijing: Science Press, 2017.

5. Liang Yong-ping, Yan Liping. Study on Dynamic Grey Bridge Pier Settlement Prediction Model Based on Wavelet Denoising[J]. Geographical Surveying and Mapping, 2017, 42(6):65-68.