

## Research Article

# Using Empirical Modal Decomposition to Improve the Daily Milk Yield Prediction of Cows

Zhiyong Cao,<sup>1,2</sup> Zhijuan Cao,<sup>3</sup> Hongwei Zhao,<sup>4</sup> Jiajun Xu,<sup>1</sup> Guangyong Zhang,<sup>1</sup> Yi Li,<sup>1</sup> Yufei Su,<sup>1</sup> Ling Lou,<sup>1</sup> Xiujuan Yang <sup>2,5</sup> and Zhaobing Gu <sup>2,5</sup>

<sup>1</sup>College of Big Data, Yunnan Agricultural University, Kunming, China

<sup>2</sup>Yunnan Provincial Key Laboratory of Animal Nutrition and Feed Science, Kunming, China

<sup>3</sup>School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, China

<sup>4</sup>College of Science, Yunnan Agricultural University, Kunming, China

<sup>5</sup>College of Animal Science and Technology, Yunnan Agricultural University, Kunming, China

Correspondence should be addressed to Xiujuan Yang; 2006040@ynau.edu.cn and Zhaobing Gu; 2012044@ynau.edu.cn

Received 25 March 2022; Accepted 17 June 2022; Published 11 July 2022

Academic Editor: Kuruva Lakshmana

Copyright © 2022 Zhiyong Cao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, the daily lactation data of Holstein dairy cows in one lactation period (305 days) were used as lactation time series data. Empirical mode decomposition (EMD) was used to decompose milk yield series. The nonstationary milk yield series with multiple oscillation modes was decomposed into several components. After eliminating the interference components, the interference components were superimposed. Remaining component reconstruction was used to get the denoising milk yield series. The denoising milk yield series retained the basic characteristics of the original milk yield series and corrected the errors of the original data. The back propagation neural network (BPNN) was used to predict and compare the original milk yield series and the denoising milk yield series. The results showed that it was feasible to use EMD to smooth the original daily milk production data. The noise reduction milk production series was beneficial to the learning of prediction model and could improve the accuracy of prediction of daily milk production of dairy cows.

## 1. Introduction

With regard to the predictive study on the milk yield of dairy cows, Brody et al. [1] had used a mathematical model to describe the functional relationship between milk yield and lactation time in 1923. Thereafter, a large amount of mathematical model research is carried out to predict the milk yield [2–4], and the basic thinking of selecting or constructing model is generally consistent, which is to fit the mean milk yield of cows in a certain period of time and then to predict the milk yield using model. It is useful to predict the population milk yield of dairy cows using the population mean statistical data for investigating the nutritional requirements of population dairy cows [5–7]; however, such treatment suggests the difference in production performance between individual cows. Clearly, applying such prediction model in predicting the milk yield of individual dairy cows will increase the prediction error [8–10]. With the increas-

ingly demanding standards of fine feeding for dairy cows, the milk yield prediction accuracy of individual dairy cows should be improved, the nutritional requirements of individual cows should be obtained, and the diet nutrition concentration for individual cows is clustered, thus obtaining the optimal grouping scheme and group feeding formula of dairy cows.

The biological difference of individual dairy cows, together with the metabolic difference in different feed nutrients, has resulted in changes in the daily milk yield of dairy cows with time. Therefore, within a lactation period, the daily milk yield of individual dairy cows shows a nonlinear time series. Measuring error during production or the unpredictable external effects will affect partial daily milk yield and become the noise term. Consequently, the partial daily milk yield represents the unpredictable part in the milk yield series from the perspective of the timing characteristics of the entire lactation period, which will disturb the learning of major data characteristics by the

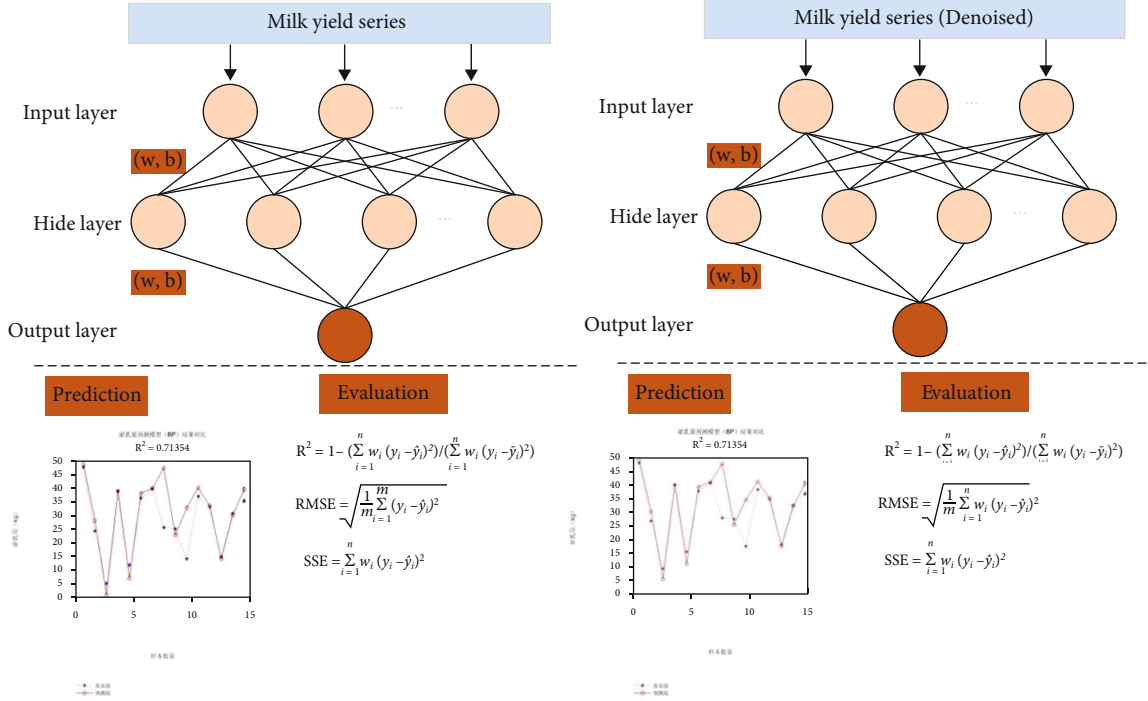


FIGURE 1: Simulation architecture. The left architecture is a BPNN model fitted with the original daily milk yield series, and  $w$  and  $b$  represent the connection *weights* and *bias* between neurons in each BPNN model. The current daily milk yield enters the input layer of the BPNN model as an input value, and the output layer is the model predictor of the next day's milk yield. The accuracy of the model prediction value is used as the criterion for evaluating the model. The right architecture is a BPNN model fitted with the denoised daily milk yield series.

prediction model and reduce the model prediction performance [11–13]. Traditional forecasting methods and models have not been flexible enough to predict lactation dynamically and reliably due to the complex physiological mechanisms of dairy cows, and in addition, single-forecasting models usually do not provide accurate prediction results. Combined forecasting models combine the advantages of different techniques or methods to predict data trends more effectively. This study employed empirical model decomposition (EMD) [14] to decompose the original milk yield data series according to the time scale characteristics, then to produce the denoised milk yield series after reconstruction, and to carry out simulation forecast using the BPNN model. The results suggested that the denoised milk yield series could well preserve the series properties of the original data, which contributed to improving the model prediction accuracy. The proposed combined model enhances the learning capability of the prediction model.

## 2. Materials and Methods

**2.1. Data.** In this study, daily milk yield data was converted to 4% FCM [15]. A lactation period of dairy cow was defined as a milk yield series, while the lactation period started from cow calving to the 305<sup>th</sup> lactation days after delivery, and the last milk yield data was not collected for prediction. Data in a milk yield series were randomly divided into training set and verification set datasets, among which the training set samples ( $n = 202$ ) were used to train the model, while the

verification set samples ( $n = 102$ ) were utilized for model verification.

**2.2. Relevant Theories and Techniques.** EMD-based milk yield data decomposition: the original milk yield series was carried out stationary processing using EMD and decomposed into a residual sum and a series of finite intrinsic mode functions (IMFs) with low quantity [14]. The original milk yield series was  $x(t)$ , which represented the cumulative sum of several timing components:

$$x(t) = \sum_{i=1}^n C_i(t) + R_n(t). \quad (1)$$

$C_i(t)$  stands for the IMFs of the original milk yield series arranged in the order of high frequency to low frequency;  $i$  represents the number of IMFs, and  $R_n(t)$  indicates the trend of original milk yield series  $x(t)$ . The original milk yield series was decomposed by EMD, as shown below:

- (1) All local maximum values and local minimum values on the original milk yield series  $x(t)$  were identified, and the upper envelop [16] curve  $u(t)$  comprised the local maximum values, together with the lower envelop curve  $l(t)$  constituted by local minimum values were obtained by the cubic spline function [17]

The means of upper envelop curve  $u(t)$  and lower envelop curve was  $m_1(t) = (u(t) + l(t))/2$ , and the difference between

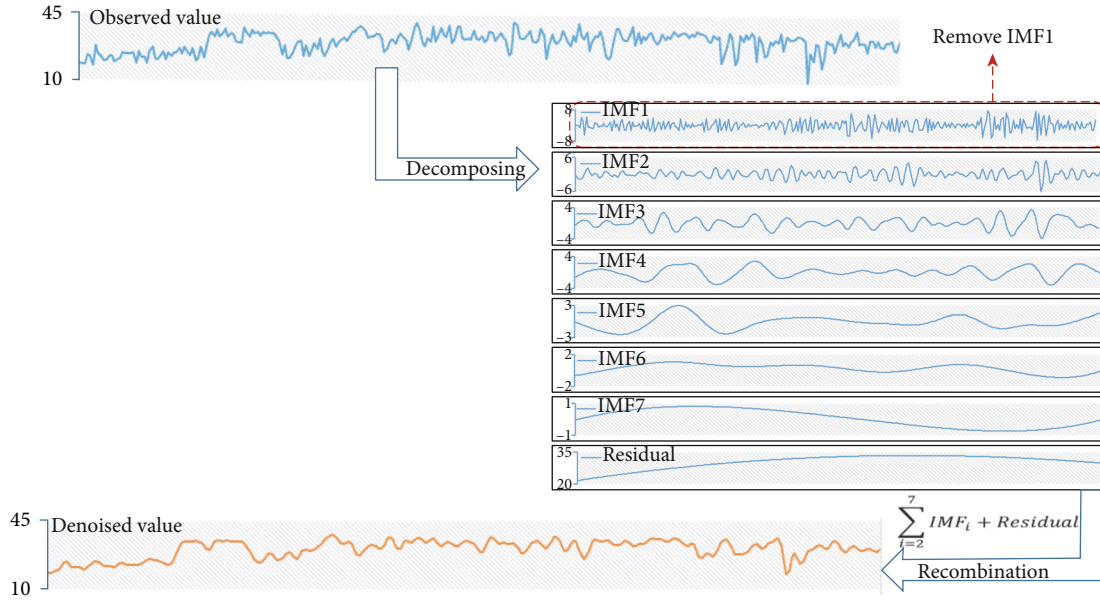


FIGURE 2: EMD denoised. The observed value is the original milk yield value, and the value varies between 10 and 45 kg per day. The original milk yield series was decomposed through EMD, to obtain IMFs and the residual series. The resultant IMF after decomposition showed obvious physical significance, which represented the hidden fluctuation components in the original milk yield series at various scale. Residual series reflect a hidden prolactin change, i.e., most milk yield value changes between 20 and 35. The high-frequency IMF1 was removed, then other IMFs were cumulatively added, which were then added with residual to obtain the denoised milk yield series.

the original milk yield series  $x(t)$  and the mean  $m_1(t)$  was  $h_1(t) = x(t) - m_1(t)$

- (2) If  $h_1(t)$  was an IMF, then  $h_1(t)$  was  $IMF_1$ , and  $C_1(t) = h_1(t)$ . Otherwise, the original milk yield series  $x(t)$  was replaced by  $h_1(t)$ , and the calculation was returned to step (1), until  $IMF_1$  was extracted
- (3) The residual was calculated as follows  $R_1(t) = x(t) - C_1(t)$ , if  $R_1(t)$  was a monotonic function, then  $x(t)$  was replaced by  $R_1(t)$ ; and steps (1)-(4) were executed. Otherwise, multiple different decomposition results would be obtained, the high frequency  $IMF_1$  was removed [18], then other IMFs were cumulatively added, which were then added with residual to obtain the denoised milk yield series

**2.3. BPNN Model.** BPNN is a kind of multilayer feedforward neural network with signal forward transfer and error back propagation, which is frequently used in nonlinear prediction, and it is also one of the most extensive neural network models applied at present. A three-layer BPNN model is constructed for simulation forecast, the numbers of input layer and output layer neurons in the model topology are related to the prediction task, and the Hecht-Nielsen method is utilized in obtaining the neuron number in the hidden layer [19]. If the neuron number in the input layer is  $m$ , then the neuron number in the hidden layer should be  $2m + 1$ . BPNN is greatly dependent on the initial weight and bias, and a group of superior weight and biasing was of crucial importance to the network prediction performance. Under the premise of guaranteeing the minimal

training error, the improved  $z$  distribution iteration is adopted to obtain these parameters, to replace the method of standard BPNN to randomly produce parameters.

**2.4. Simulation Forecast.** Simulation forecast was constituted by the original milk yield series/denoised milk yield series,  $z$  distribution iteration adaptive operator finds, and BPNN sets (Figure 1), as described below:

- (1) The normalized formula  $(x' = 2 \times ((x - x_{min}) / (x_{max} - x_{min})) - 1)$  was used to normalize the original milk yield series/denoised milk yield series to interval  $[-1,1]$
- (2)  $z$  distribution iteration adaptive operators were found, BPNN was initiated, and the self-adapting operator find algorithm was used to optimize the BPNN initial weight and bias
- (3) BPNN was trained using the optimal weight and bias, the training set milk yield was used as the input variable, while the morrow milk yield of training set was used as the output variable, and finally the verification set was used for test
- (4)  $R^2$ , RMSE, and SSE were employed to evaluate the prediction results

**2.5. Evaluation Indexes of Prediction Performance.**  $R$ -square (coefficient of determination) was used to characterize the fitting effect of data, and the value closer to 1 indicates better

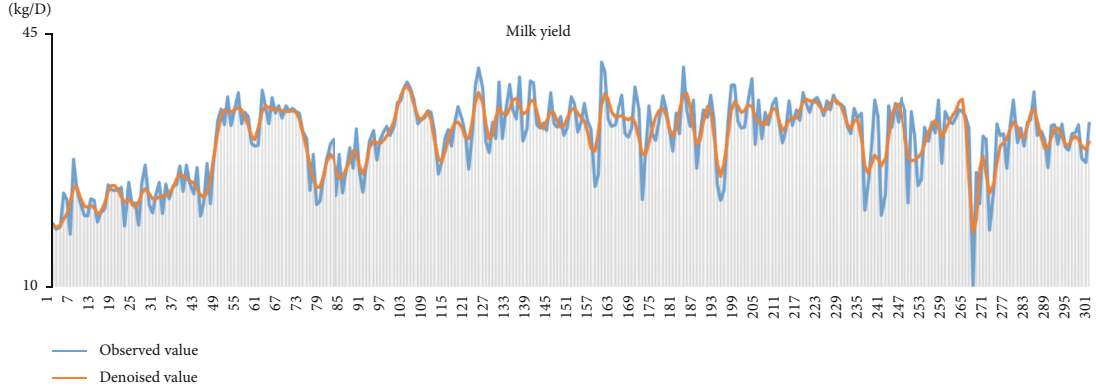


FIGURE 3: Original value and denoised value of the milk yield series. The observed value is the original milk yield value. The horizontal axis represents the time of day (305 days), and the vertical axis represents the milk yield value.

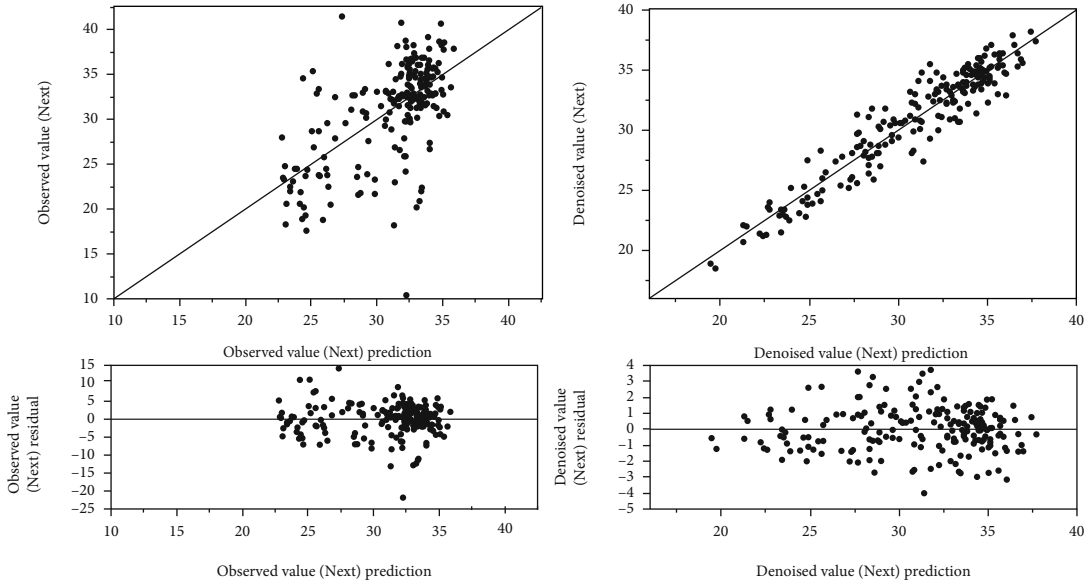


FIGURE 4: Training set comparison. The correlation between the predicted values obtained by the training set fitting the prediction model and the original values is represented in the scatter plot on the left. The correlation between the predicted values obtained by the training set fitting the prediction model and the denoised values is represented in the scatter plot on the right.

fitting degree of data by the model.

$$R^2 = 1 - \frac{\left(\sum_{i=1}^n w_i (y_i - \hat{y}_i)^2\right)}{\sum_{i=1}^n w_i (y_i - \bar{y}_i)^2}, \quad (2)$$

$$R^2 = 1 - \frac{\left(\sum_{i=1}^n w_i (y_i - \hat{y}_i)^2\right)}{\left(\sum_{i=1}^n w_i (y_i - \bar{y}_i)^2\right)}.$$

Root mean squared error (RMSE) was the fitting standard deviation of the regression system.

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}. \quad (3)$$

The sum of squares due to error (SSE) closer to 0 indicated better model selection and fitting and more successful

data prediction.

$$\text{SSE} = \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2. \quad (4)$$

### 3. Results and Discussion

**3.1. Denoised of the Milk Yield Series.** In Figure 2, the observed daily milk yield data of every cow represented one timing value of the milk yield series within a lactation period; then, this series was characterized by the nonstability, nonlinearity, complexity, and noise. As IMF included various scale features, EMD was utilized to decompose the original milk yield series, and all finite IMFs and residual series were extracted. The remaining IMFs and residuals after removing the first IMF were recombined to obtain the denoised milk yield series.

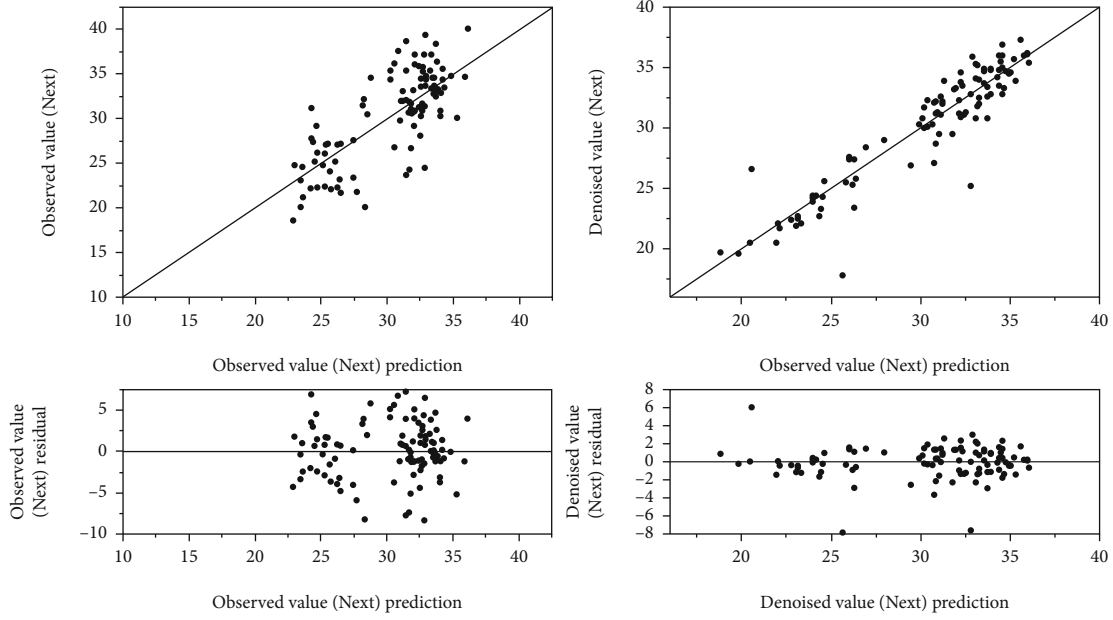


FIGURE 5: Verification set comparison. The correlation between the predicted values obtained by the validation set using the prediction model and the original values is represented in the scatter plot on the left. The correlation between the predicted values obtained by the training using the prediction model and the denoised values is represented in the scatter plot on the right.

TABLE 1: Comparison of the prediction model evaluation between original milk yield series and denoised milk yield series.

| Evaluation items | Training dataset |          | Validate dataset |          |
|------------------|------------------|----------|------------------|----------|
|                  | Observed         | Denoised | Observed         | Denoised |
| $R^2$            | 0.37             | 0.90     | 0.56             | 0.86     |
| RMSE             | 4.35             | 1.37     | 3.35             | 1.79     |
| SSE              | 3827.68          | 379.90   | 1146.28          | 326.81   |
| Frequency        | 202              | 202      | 102              | 102      |

The milk yield series was decomposed through EMD, to obtain IMF and the residual series, and the resultant IMF included the local characteristic signals of the original milk yield series at various time scales, which was the “roguing” [20, 21] reflection of the entire original milk yield series. The resultant IMF after decomposition showed obvious physical significance, which represented the hidden fluctuation components in the original milk yield series at various scale, while the residual indicated the basic trend of the milk yield trend, and the residual trend was consistent with the group statistical data trend features [22–24].

The front milk yield of the original milk yield series was gradually increased (early lactation period), which peaked in the second or third lactation month and maintained for a period of time (peak lactation period), and thereafter decreased slowly (middle-late lactation period). The trend of milk yield is consistent with the known reports [25, 26]. The denoised milk yield series also well reflected the lactation peak features at the period of 49-79 days, corrected the measuring error of the 271 lactation days, and preserved the milk yield characteristics of individuals in the original milk yield series on the whole (Figure 3).

**3.2. BPNN Prediction.** The training datasets of the original milk yield series and denoised milk yield series were used to fit the prediction equation using BPNN. The prediction equation scatter diagram of the training set of the original milk yield series is shown in the left of Figure 4, while that of the training set of the denoised milk yield series is presented in the right of Figure 4. The prediction equation scatter diagram of the verification set of the original milk yield series is displayed in the left of Figure 5, while that of the verification set of the denoised milk yield series is shown in the right of Figure 5.

The prediction scatter diagram of the training set of the original milk yield series was clearly scattered, while obvious correlation could be seen in the training set of the denoised milk yield series. The minimum of 10.4kg was seen on the 271<sup>st</sup> lactation day in the original milk yield series, which belonged to the measuring error upon verification; the predicted value was 32.3kg, which had resulted in the maximum residual of 21.9. Such abnormal value was repaired to be 17.8kg after denoised processing, and the residual was reduced. At the same time, the milk yield values in the training set of the original milk yield series mostly concentrated on 30 kg/d to 35 kg/d, and such feature was reflected in the scatter diagram of the training set of the denoised milk yield series. The greatest scatter density was seen at 30-35 in the right of Figure 4.

The prediction scatter diagram of the verification set of the original milk yield series was not clearly scattered, which also demonstrated that the errors and noise data randomly divided to the verification set of the original milk yield series were less than those to the test set. The obvious correlation could still be observed in the verification set of the denoised milk yield series. With regard to the residual, the verification

set of denoised milk yield series also performed better than that of the original milk yield series.

In this study, the prediction effect of the denoised milk yield series was better than that of the original milk yield series, and the indexes of correlation and residual in the denoised milk yield series were better than those of the original milk yield series, which partially overcame the problem of prediction distortion of the measuring errors and noise data.

**3.3. BPNN Evaluation.** The training datasets of the original milk yield series and denoised milk yield series were fit using BPNN (Table 1), respectively. The determination coefficient  $R^2$  of the original milk yield series fitting model was 0.37, RMSE was 4.35, and SSE was 3827.68. The determination coefficient  $R^2$  of the denoised milk yield series fitting model was 0.90, RMSE was 1.37, and SSE was 379.90. The original milk yield series fitting model was applied to the original test set samples, and the prediction result  $R^2$  was 0.56, RMSE was 3.35, and SSE was 1146.28. The denoised milk yield series fitting model was applied to the denoised test set samples, and the prediction result  $R^2$  was 0.86, RMSE was 1.79, and SSE was 326.81. Clearly, the fitting degree obtained upon directly using the original milk yield series was substantially decreased, and the error was greatly increased, suggesting apparently poor effect of the prediction model constructed based on the original milk yield series. By contrast, the prediction model constructed based on the denoised milk yield series had better selection and fitting, as well as more successful data prediction [27].

In Table 1, the observed value is the original milk yield value. Data in a milk yield series were randomly divided into training set and verification set datasets, among which the training set samples ( $n = 202$ ) were used to train the model, while the verification set samples ( $n = 102$ ) were utilized for model verification. The samples are frequency.

The training set and verification set were compared in the same prediction model. The  $R^2$  of the verification set of the original milk yield series was increased by 19% compared with that of the training set, RMSE was decreased by 1 unit, and SSE was reduced by 2681.4 units. The  $R^2$  of the verification set of the denoised milk yield series was decreased by 4% compared with that of the training set, RMSE was increased by 0.42 unit, and SSE was reduced by 53.09 units. The differences in  $R^2$ , RMSE, and SSE between the training set and verification set of the original milk yield series were great, which suggested poor stability of the prediction model fitted using the original milk yield series; at the same time, the model was associated with the problem of excessive fitting.

## 4. Conclusion

The original milk yield data series of dairy cow within a lactation period displays dynamic complexity. Prior to construction of the prediction model, EMD is adopted to reduce the noise of the original milk yield series, which can remove the abnormal points in the series, obtain the basic trend of milk yield within the lactation period and daily milk

yield fluctuations of various scales, and restore the clear lactation features of individual cows. The original milk yield data series of individual cows is subjected to EMD denoised processing. Besides, fitting a nonlinear model (such as BPNN) for prediction is a feasible strategy, which can not only improve the daily milk yield prediction performance but also enhance the stability of the prediction model.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no competing interest.

## Authors' Contributions

Zhiyong Cao and Zhijuan Cao contributed equally to this work as co-first author.

## Acknowledgments

This research was supported financially by the project "Major Science and Technology in Yunnan Province/2018ZF012" and the project "Soil Pollution Prevention and Control in Yunnan Province/201839# A3008784."

## References

- [1] S. Brody, A. C. Ragsdale, and C. W. Turner, "The rate of decline of milk secretion with the advance of the period of lactation," *The Journal of General Physiology*, vol. 5, no. 4, pp. 441–444, 1923.
- [2] T. E. Ali and L. R. Schaeffer, "Accounting for covariances among test day milk yields in dairy cows," *Canadian Journal of Animal Science*, vol. 67, no. 3, pp. 637–644, 1987.
- [3] I. R. Johnson, J. France, and B. R. Cullen, "A model of milk production in lactating dairy cows in relation to energy and nitrogen dynamics," *Journal of Dairy Science*, vol. 99, no. 2, pp. 1605–1618, 2016.
- [4] J. H. M. Thornley and J. France, *Mathematical Models in Agriculture*, CAB International, Wallingford, UK, 2007.
- [5] L. F. Dong, T. Yan, C. P. Ferris, and D. A. McDowell, "Comparison of maintenance energy requirement and energetic efficiency between lactating Holstein-Friesian and other groups of dairy cows," *Journal of Dairy Science*, vol. 98, no. 2, pp. 1136–1144, 2015.
- [6] P. Kliš, D. Pivczyński, A. Sawa, and B. Sitkowska, "Prediction of lactational milk yield of cows based on data recorded by AMS during the periparturient period," *Animals*, vol. 11, no. 2, p. 383, 2021.
- [7] V. M. Thorup, D. Edwards, and N. C. Friggens, "On-farm estimation of energy balance in dairy cows using only frequent body weight measurements and body condition score," *Journal of Dairy Science*, vol. 95, no. 4, pp. 1784–1793, 2012.
- [8] A. Liseune, M. Salamone, D. Van den Poel, B. van Ranst, and M. Hostens, "Predicting the milk yield curve of dairy cows in the subsequent lactation period using deep learning,"

- Computers and Electronics in Agriculture*, vol. 180, article 105904, 2021.
- [9] N. C. Friggens, C. Ridder, and P. Løvendahl, "On the use of milk composition measures to predict the energy balance of dairy cows," *Journal of Dairy Science*, vol. 90, no. 12, pp. 5453–5467, 2007.
- [10] W. Grzesiak, D. Zaborski, I. Szatkowska, and K. Królaczyk, "Lactation milk yield prediction in primiparous cows on a farm using the seasonal auto-regressive integrated moving average model, nonlinear autoregressive exogenous artificial neural networks and Wood's model," *Animal bioscience*, vol. 34, no. 4, pp. 770–782, 2021.
- [11] R. Thakur, A. Chandel, and R. K. Gupta, "Prediction of cow milk yield in Himachal Pradesh and Northern Himalayan Province of India," *Biological Forum*, vol. 14, no. 1, pp. 1078–1082, 2021.
- [12] C. Tantithamthavorn, S. McIntosh, A. E. Hassan, A. Ihara, and K. Matsumoto, "The impact of mislabelling on the performance and interpretation of defect prediction models," in *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, vol. 1, pp. 812–823, Florence, Italy, May 2015.
- [13] F. Zhang, J. Upton, L. Shalloo, and M. D. Murphy, "Effect of parity weighting on milk production forecast models," *Computers and Electronics in Agriculture*, vol. 157, pp. 589–603, 2019.
- [14] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [15] W. L. Gaines, *The Energy Basis of Measuring Milk Yield in Dairy Cows. Bulletin (University of Illinois (Urbana-Champaign campus). Agricultural Experiment Station)*, no. 308, 1928.
- [16] H. H. Zhu and N. Huang, *Random Signal Analysis*, Beijing Institute of Technology Press, 2011.
- [17] Y. E. Hou and H. B. Liu, "The computer solution to cubic spline," *Journal of Zhongzhou University*, vol. 3, pp. 111–112, 2004.
- [18] C. Li, Z. Li, J. Wu, L. Zhu, and J. Yue, "A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features," *Information Processing in Agriculture*, vol. 5, no. 1, pp. 11–20, 2018.
- [19] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural networks for perception*, pp. 65–93, Academic Press, 1992.
- [20] S. Mou, Y. Ji, and C. Tian, "Retail time series prediction based on EMD and deep learning," in *2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, pp. 425–430, Guiyang, China, August 2018.
- [21] J. J. Ruiz-Aguilar, I. Turias, J. González-Enrique, D. Urda, and D. Elizondo, "A permutation entropy-based EMD-ANN forecasting ensemble approach for wind speed prediction," *Neural Computing and Applications*, vol. 33, no. 7, pp. 2369–2391, 2021.
- [22] I. Harder, E. Stamer, W. Junge, and G. Thaller, "Lactation curves and model evaluation for feed intake and energy balance in dairy cows," *Journal of Dairy Science*, vol. 102, no. 8, pp. 7204–7216, 2019.
- [23] V. E. Olori, S. Brotherstone, W. G. Hill, and B. J. McGuirk, "Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd," *Livestock Production Science*, vol. 58, no. 1, pp. 55–63, 1999.
- [24] P. D. P. Wood, "Algebraic model of the lactation curve in cattle," *Nature*, vol. 216, no. 5111, pp. 164–165, 1967.
- [25] Q.-Y. Luo, B. H. Xiong, Y. Ma, Z. H. Pang, and W. Deng, "Study on lactation curve models of Chinese Holstein for the second parity," *Scientia Agricultura Sinica*, vol. 43, no. 23, pp. 4910–4916, 2010.
- [26] B. H. Xiong, Y. Ma, Q. Y. Li, Z. H. Pang, and W. Deng, "Study on lactation curve models of Chinese Holstein for the third parity," *Scientia Agricultura Sinica*, vol. 44, no. 2, pp. 402–408, 2011.
- [27] J. Dong, W. Dai, L. Tang, and L. Yu, "Why do EMD-based methods improve prediction? A multiscale complexity perspective," *Journal of Forecasting*, vol. 38, no. 7, pp. 714–731, 2019.