



Autoregressive Modeling and Prediction of the Activity of Antihypertensive Peptides

Xufen Xie¹, Chuanchuan Zhu¹, Di Wu^{2,3} and Ming Du^{2,3*}

¹School of Information Science and Engineering, Dalian Polytechnic University, Dalian, China, ²School of Food Science and Technology, Dalian Polytechnic University, Dalian, China, ³National Engineering Technology Research Center of Seafood, Dalian Polytechnic University, Dalian, China

Naturally derived bioactive peptides with antihypertensive activities serve as promising alternatives to pharmaceutical drugs. There are few relevant reports on the mapping relationship between the EC₅₀ value of antihypertensive peptide activity (AHTPA-EC₅₀) and its corresponding amino acid sequence (AAS) at present. In this paper, we have constructed two group series based on sorting natural logarithm of AHTPA-EC₅₀ or sorting its corresponding AAS encoding number. One group possesses two series, and we find that there must be a random number series in any group series. The random number series manifests fractal characteristics, and the constructed series of sorting natural logarithm of AHTPA-EC₅₀ shows good autocorrelation characteristics. Therefore, two non-linear autoregressive models with exogenous input (NARXs) were established to describe the two series. A prediction method is further designed for AHTPA-EC₅₀ prediction based on the proposed model. Two dynamic neural networks for NARXs (NARXNNs) are designed to verify the two series characteristics. Dipeptides and tripeptides are used to verify the proposed prediction method. The results show that the mean square error (MSE) of prediction is about 0.5589 for AHTPA-EC₅₀ prediction when the classification of AAS is correct. The proposed method provides a solution for AHTPA-EC₅₀ prediction.

OPEN ACCESS

Edited by:

Juan Wang,
Inner Mongolia University, China

Reviewed by:

Liang Cheng,
Harbin Medical University, China
Yanjuan Li,
Quzhou University, China

*Correspondence:

Ming Du
duming@dpu.edu.cn

Specialty section:

This article was submitted to
Statistical Genetics and Methodology,
a section of the journal
Frontiers in Genetics

Received: 25 October 2021

Accepted: 09 December 2021

Published: 11 January 2022

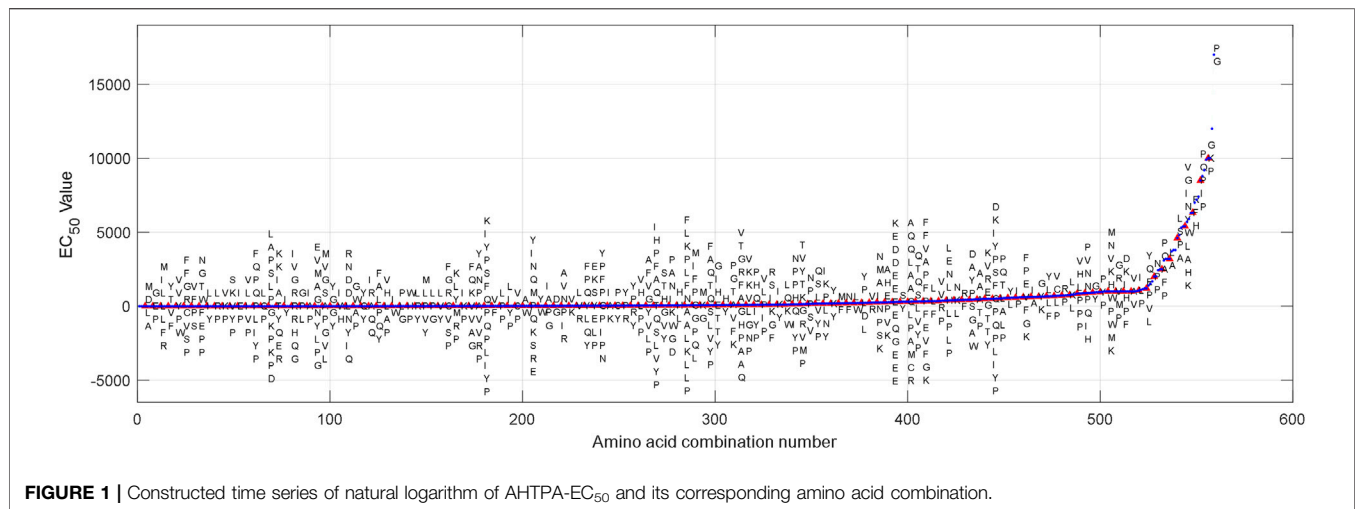
Citation:

Xie X, Zhu C, Wu D and Du M (2022)
Autoregressive Modeling and
Prediction of the Activity of
Antihypertensive Peptides.
Front. Genet. 12:801728.
doi: 10.3389/fgene.2021.801728

Keywords: antihypertensive peptides, NARXNN, fractal characteristics, EC₅₀ prediction, machine learning

1 INTRODUCTION

Hypertension is a clinical syndrome characterized by increased systemic arterial blood pressure, which can be accompanied by functional or organic damage of the heart, brain, kidney, and other organs. The renin-angiotensin system (RAS) controls blood pressure by regulating the volume of blood in blood vessels. The angiotensin-converting enzyme (ACE) is the core component of the RAS. The ACE can convert inactive angiotensin I into angiotensin II with vasoconstriction, which indirectly increases blood pressure (Zhang et al., 2000). Therefore, ACE inhibitors are widely used as drugs for the treatment of cardiovascular diseases (Stone, 2018). Antihypertensive active peptide is an effective ACE inhibitor (Tu et al, 2018a; Tu et al, 2018b; Wu et al, 2019), which has attracted great attention in the treatment and prevention of hypertension. The EC₅₀ value (sample concentration when the ACE inhibition rate is 50%) describes the activity of antihypertensive peptide, which is the most important index to select antihypertensive active peptide. Some research studies focus on feature representation (Tong,

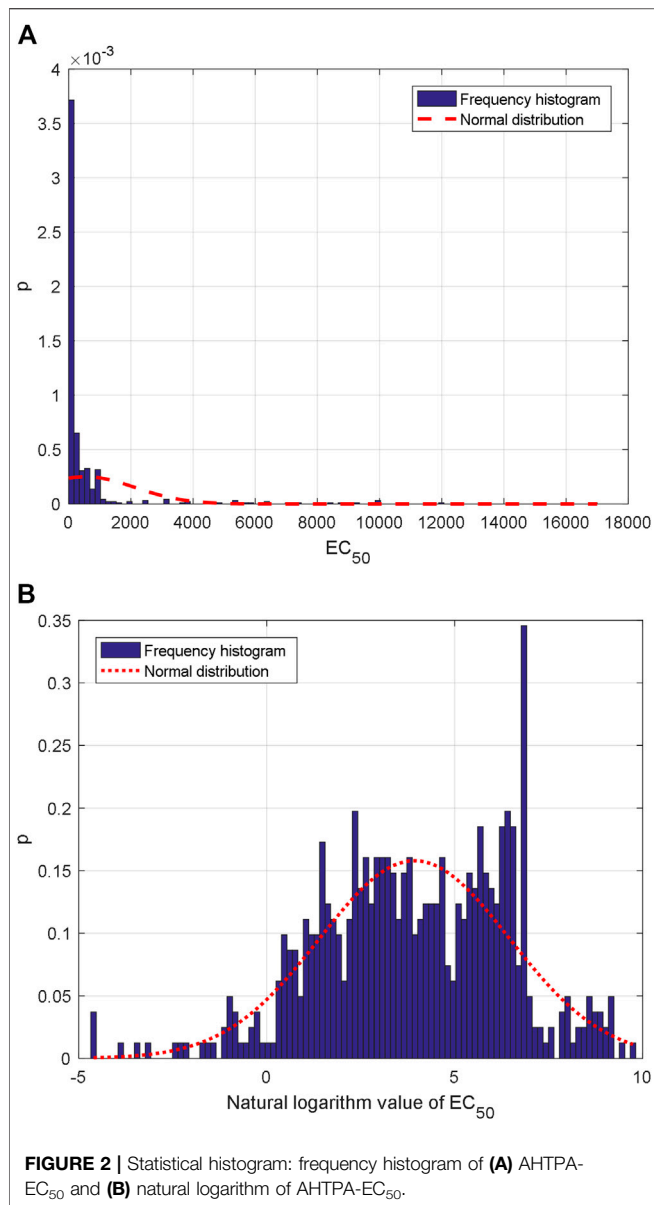


et al, 2008; Manavalan et al, 2019), and some research studies focus on identification (Majumder, and Wu, 2010). Machine learning (ML) approaches are becoming more and more popular in bioinformatics (Baldi et al., 2001; Libbrecht and Noble, 2015; Zou and Qiliu, 2019; Yang et al., 2020; Zhang et al., 2021). Some research studies are associated with classification, and some are associated with regression. In 2015, Kumar et al. developed four different model types for predicting AHTPs with varied lengths using ML approaches (Kumar et al., 2015a; Kumar et al., 2015b). Another paper on AHTP prediction used random forest (RF) approaches (Win et al., 2018). However, there is great uncertainty in the relationship between the AAS of antihypertensive peptides and its corresponding AHTPA-EC₅₀. So far, the mapping relationship between AHTPA-EC₅₀ and its corresponding AAS has not been reported. The existing published data show that AHTPA-EC₅₀ has multi-scale characteristics. It is difficult to establish a deterministic model between the AAS and AHTPA-EC₅₀ directly.

Fractal phenomena generally exist in nature. Fractal data have the characteristics of instability, self-similarity, and multi-scale (Ruderman, 1996; Ghosh and Somvanshi, 2008; Al-Hamdan, et al, 2010; Al-Hamdan et al, 2012). The spectrum of fractal data is consistent (Pentland, 1984; Nill and Bouzas, 1992; Wornell and Oppenheim, 1992). These characteristics can be used to describe physical phenomena with statistical fractal. Fractional Brownian motion (FBM) (Chow, 2011; Kim and Kim, 2004; Fouché and Mukeru, 2013) is more universal than ordinary Brownian motion, and it can better describe the fractal phenomena in nature. FBM can be modeled and described by the time series of dynamic system, and time-series analysis is an important method of system identification and analysis. Yule first proposed the autoregressive (AR) model to predict the law of market change in 1927. In the 1960s, time-series analysis made a great progress in spectral analysis and estimation. The research of linear time-series model has been greatly developed from the AR model to autoregressive moving average (ARMA) modeling theory.

Engle and Granger developed estimation procedures, tests, and empirical examples for the relationship between co-integration and error correction models (Engle and Granger 1987), and Hannan and Deistler proposed the multivariable VARMA model and VARMAX model (Hannan and Deistler, 1988). However, Moran proposed the limitations of linear model in the 1950s (Moran, 1953). The non-linear time-series model follows to become an attracting research topic until the late 1970s and early 1980s. These research studies include the threshold autoregressive model, exponential autoregressive model, bilinear model, non-linear autoregressive model, and state-dependent model. Tong et al. gave the threshold autoregressive model (Tong, 1983), and Ozaki proposed an exponential autoregressive model (Ozaki, 1980). The system identification is generally based on the complete clarity of input–output causality. In practical application, the system output can be measured, but the input of some specific systems is difficult to observe and measure. In that situation, it is not easy to determine the causal relationship between input and output. In that case, the traditional system identification method is difficult to apply. Although the system's input cannot always be determined, it is certain that there is a relationship between some known parameters or data and the system output. These known parameters or data can directly or indirectly affect the system output. If the relevant data are also regarded as the system input, then the time-series model with exogenous input is determined. Tong analyzed the non-linear time series with exogenous input, established the relationship between non-linear time series and non-linear dynamic system (chaos), and studied the prediction based on non-linear time series (Tong, 1990).

In this paper, a kind of time series construction method on AHTPA-EC₅₀ and its corresponding AAS is proposed firstly. We can find a lot of fractal characteristics from the two group time series. Then, the two groups of constructed series are modeled as two different NARX time-series models. Furthermore, two NARXNNs are used to perform the



proposed model. And then we further proposed a prediction method for AHTPA-EC₅₀ based on two NARXNNs and ML classification algorithms. The model and prediction method are useful and meaningful on antihypertensive active peptide research, drug design, and industrial production.

2 MATERIALS AND METHODS

2.1 Analysis of AHTPA-EC₅₀ and Its Corresponding AAS

2.1.1 Statistical Analysis of AHTPA-EC₅₀

559 group AHTPA-EC₅₀ data and their corresponding AAS are shown in **Figure 1**. Due to the difficulty of display,

TABLE 1 | Numerical definitions of amino acids.

Amino acids	A	C	D	E	F	G	H	I	K	L
Numerical definitions	1	2	3	4	5	6	7	8	9	10
Amino acids	M	N	P	Q	R	S	T	V	W	Y
Numerical definitions	11	12	13	14	15	16	17	18	19	20

Supplementary Material marks the corresponding AAS every four EC₅₀ values (interval = 3). The statistical histogram is analyzed, and histogram analysis of AHTPA-EC₅₀ is shown in **Figure 2A**. We can see that AHTPA-EC₅₀ is concentrated on the right side of the longitudinal axis of the coordinate and there is some very large AHTPA-EC₅₀ value in these data. The characteristics of large distribution span and asymmetry appear in AHTPA-EC₅₀ data. Comparing with the normal distribution data with the same mean and variance, it can be seen that AHTPA-EC₅₀ data deviate very far from the normal distribution. In order to reduce the scale of AHTPA-EC₅₀, the natural logarithm of AHTPA-EC₅₀ data is calculated. The distribution of natural logarithm of AHTPA-EC₅₀ is further analyzed, and the histogram distribution is shown in **Figure 2B**. Compared with the normal distribution of the same mean and variance, the natural logarithm histogram of AHTPA-EC₅₀ cut off more slowly in the tail, and it shows the characteristics of a long tail. This is an important feature of fractal data.

2.1.2 Encoding for AAS

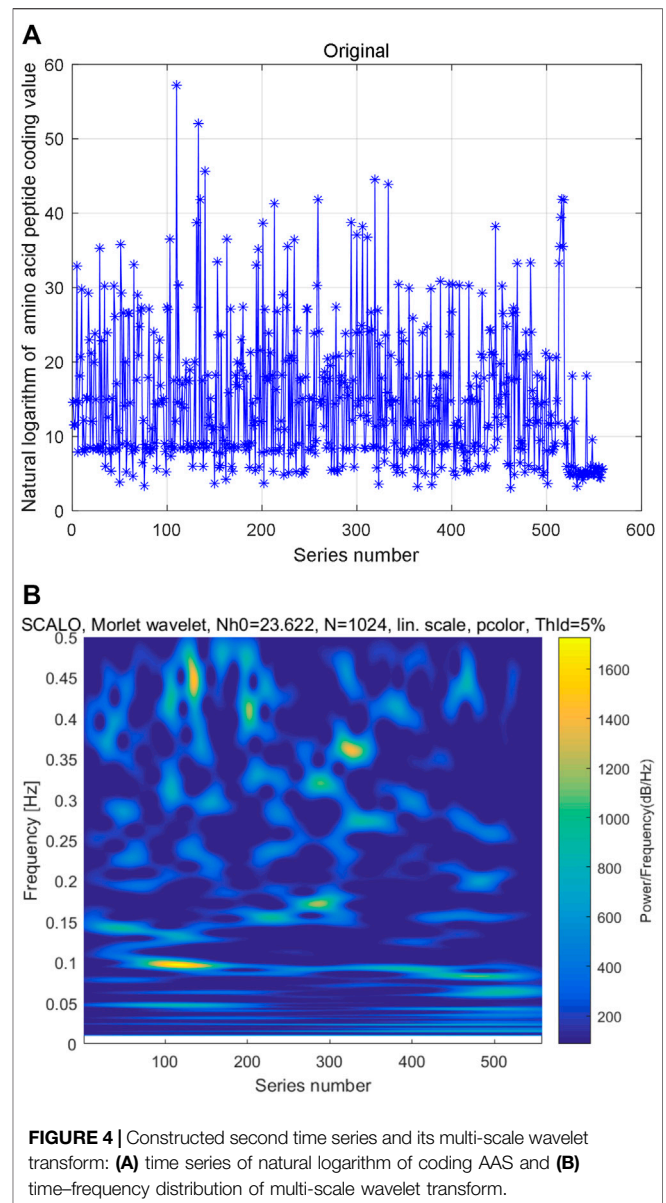
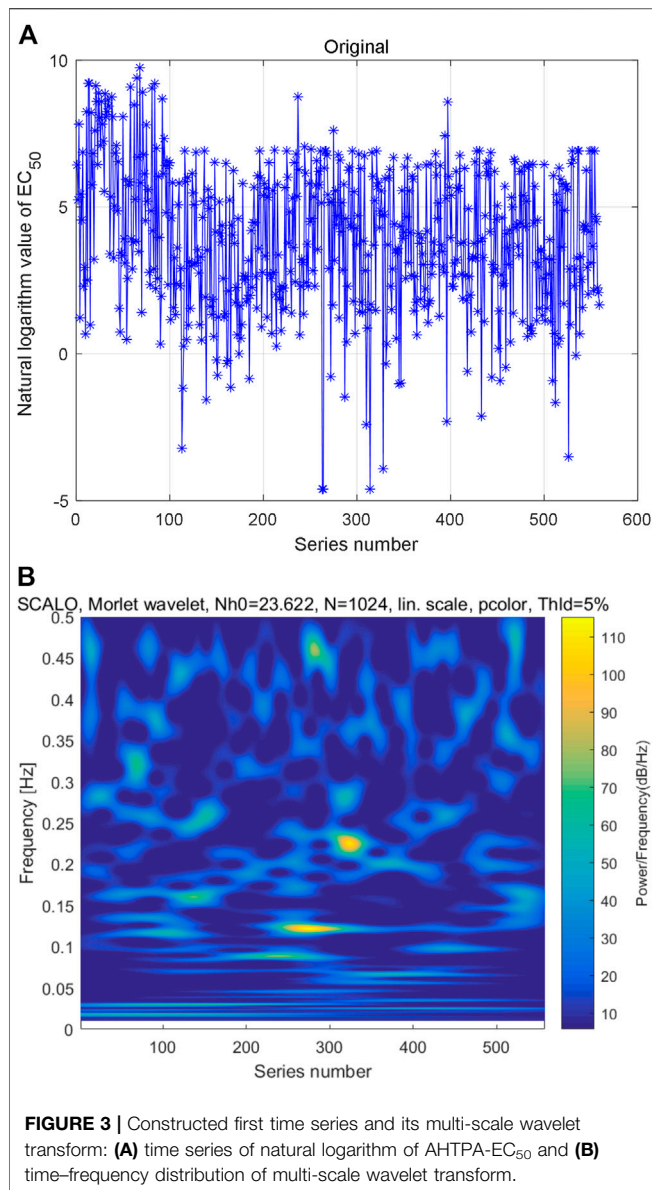
The expression of amino acid is different from the digital number, and it is a symbolic quantity that cannot be directly quantified. In order to analyze the relationship between the AAS and its corresponding AHTPA-EC₅₀, it is necessary to encode for the AAS. The numerical definitions of different amino acids are shown in **Table 1**. The AAS is digitally encoded in a 21 base system. Because the number 0 cannot appear in the first place of the combined code, the number 0 is not defined here.

2.1.3 Constructed Time Series and Its Time-Frequency Characteristics

(1) Constructed time series based on sorting code of AAS

As mentioned above, the AAS can be converted to decimal digit by numerical definitions of amino acids. After sorting the natural logarithm of coding numbers from small to large, the natural logarithm of AHTPA-EC₅₀ can be constructed. The constructed time series is shown in **Figure 3A**. Multi-scale wavelet transform is performed to the constructed AHTPA-EC₅₀ time series, and the time-frequency distribution is shown in **Figure 3B**. There is also no obvious law between high-energy data and series number and frequency in **Figure 3B**, and different time-frequency relationships show similar patterns.

(2) Constructed time series based on sorting AHTPA-EC₅₀

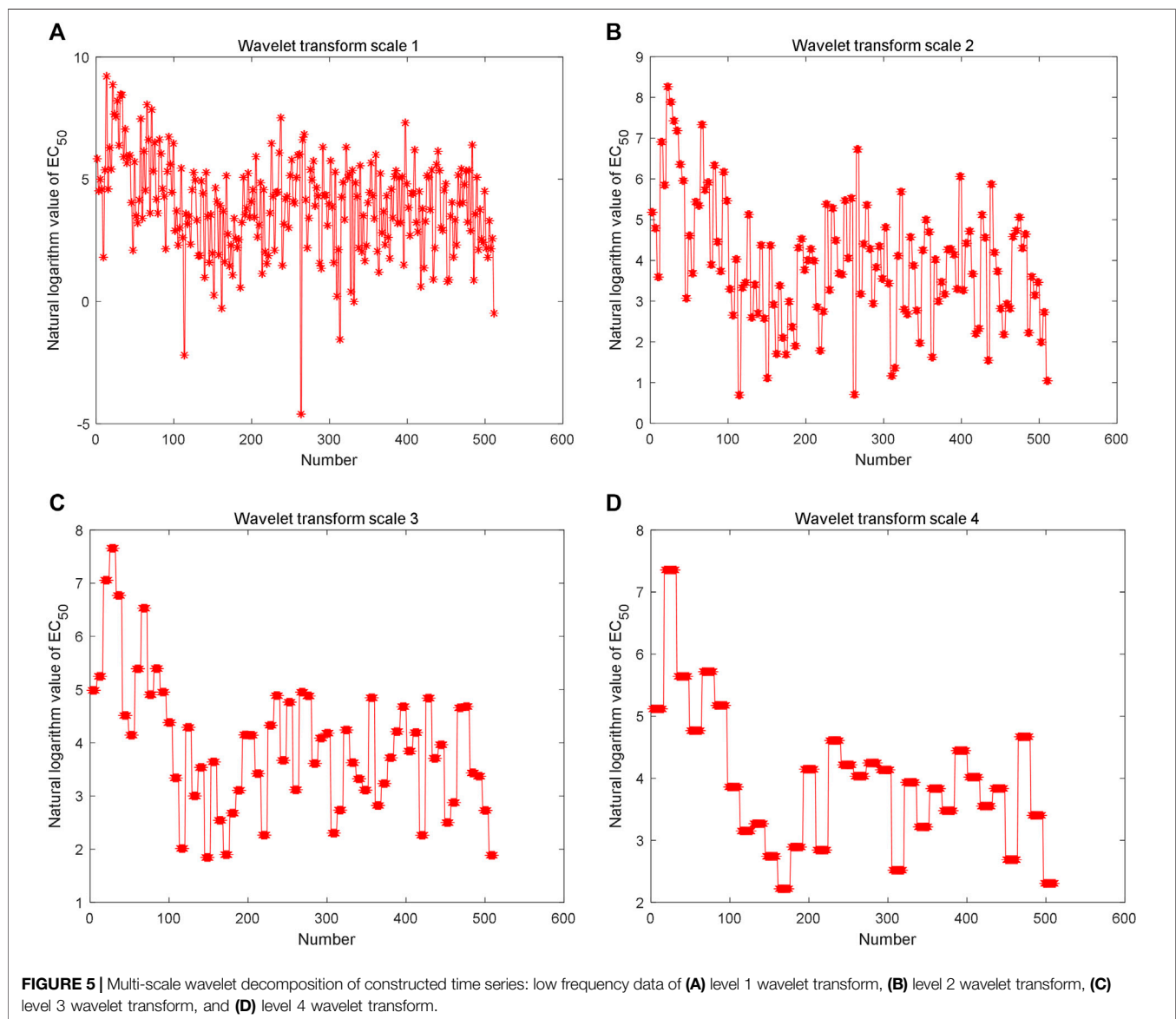


We also constructed natural logarithm of AHTPA- EC_{50} time series by sorting the data from small to large. The AAS is converted to decimal digit by numerical definitions of amino acids. After sorting the natural logarithm of AHTPA- EC_{50} from small to large, the time series of natural logarithm of coding value of AAS is also constructed. The constructed time series is shown in **Figure 4A**. Multi-scale wavelet transform is performed to the natural logarithm of coding value of AAS. The constructed time series of AAS and its time-frequency distribution are shown in **Figure 4B**. And there is no obvious law between high-energy data and series number and frequency. However, different time-frequency relationships show similar patterns.

In summary, the relationship between the natural logarithm of AHTPA- EC_{50} and its corresponding natural logarithm of coding

AAS is special. If one of the series is sorted, the other will be a random number series. We deduce that there is not a direct regression modeling for their relationship.

The Haar wavelet is further used to decompose the reconstructed time series to analyze fractal characteristics (data in **Figure 3A**) in multiple scales. The low-frequency data of different scales are shown in **Figures 5A,B,C,D**. The Hurst index of the time series is estimated by multi-scale wavelet transform data, as shown in **Figure 6A**, in which the wavelet transform scales are 1–9. The estimated Hurst index is used to generate FBM, and the empirical probability distribution of the generated FBM data is shown in **Figure 6B**. 10,000 FBM data are generated by the Monte Carlo method here. The probability distribution data corresponding to the constructed natural logarithm of



AHTPA- EC_{50} are represented in red, and the curve closest to the constructed natural logarithm of AHTPA- EC_{50} is shown in blue. It can be seen that the constructed AHTPA- EC_{50} is very close to the FBM time series.

2.2 Non-Linear Autoregressive Time-Series Modeling and Its Implementation

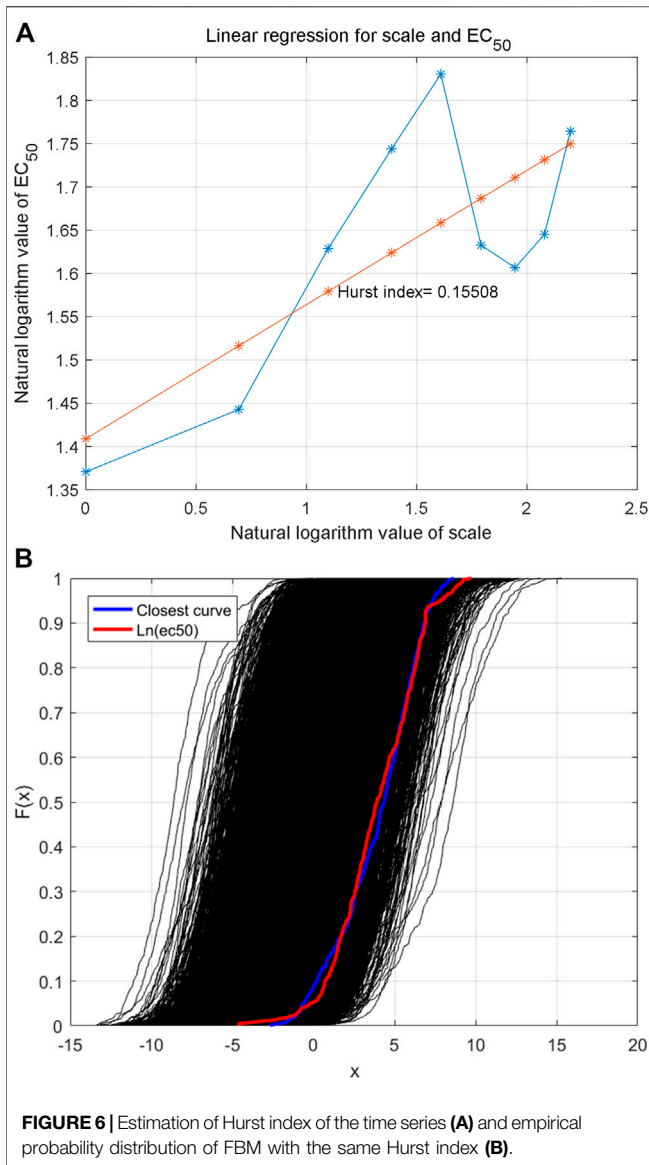
2.1.4 Correlation Analysis

Although the constructed series shows fractal characteristics, the relationship between the natural logarithm of coding value of AAS and its corresponding natural logarithm of AHTPA- EC_{50} still needs to be analyzed. **Figure 7A** shows the cross-correlation analysis for the first group of constructed time series, and it shows weak correlation between the two time series. **Figure 7B** shows the autocorrelation analysis for sorting natural logarithm of

AHTPA- EC_{50} . We can see that the sorting natural logarithm of AHTPA- EC_{50} showed weak autocorrelation. **Figure 8A** shows the cross-correlation analysis for the second group of time series, and it shows weak correlation between the two time series. **Figure 8B** shows the autocorrelation analysis for constructed natural logarithm of AHTPA- EC_{50} , and the natural logarithm of AHTPA- EC_{50} based on the coding value AAS showed obvious autocorrelation.

2.1.5 Non-Linear Autoregressive Model With Exogenous Input

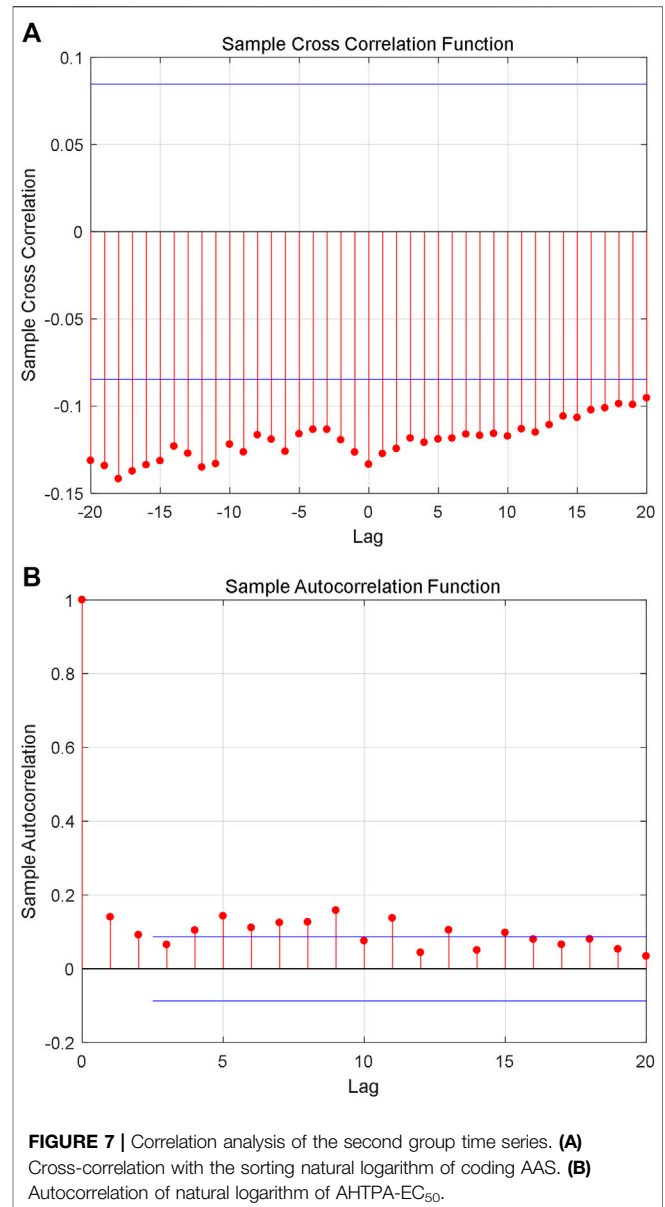
According to the above analysis, the two groups' constructed AHTPA- EC_{50} data are modeled as an autoregressive time series, and the natural logarithm of coding AAS is used as the exogenous input parameter. The non-linear autoregressive model with



exogenous input is established to describe the relationship between the AAS and its corresponding AHTPA-EC₅₀, and this relationship is described as

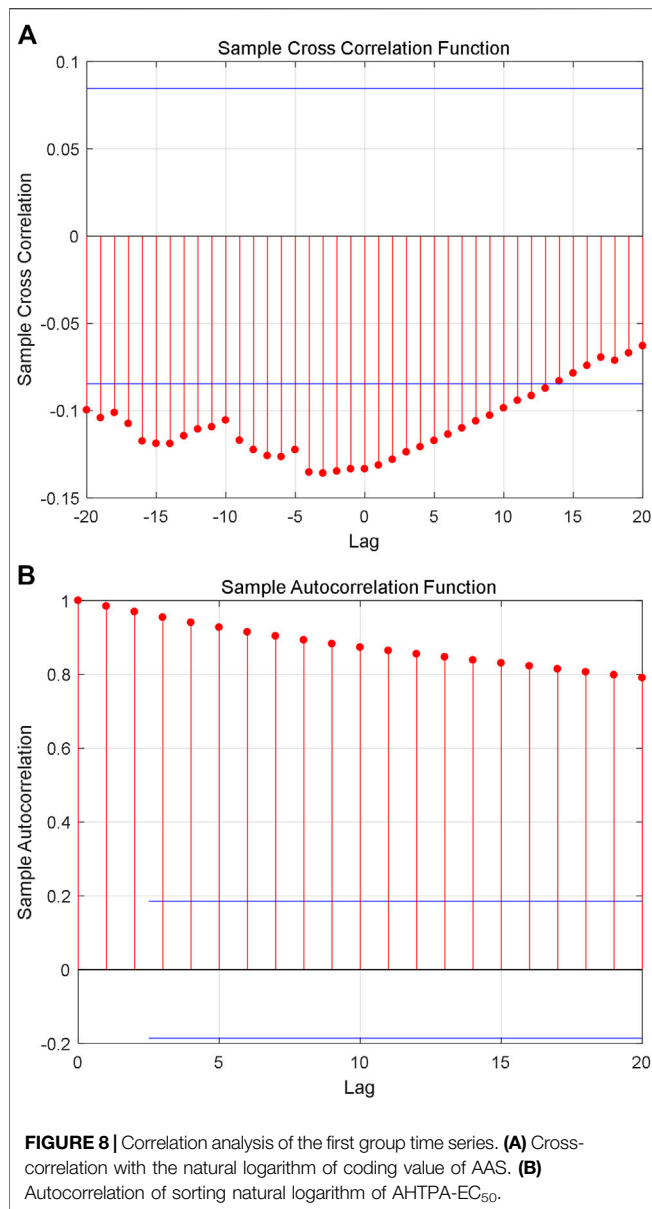
$$y(t) = f \left[\begin{matrix} y(t-1), y(t-2), \dots, y(t-n_y) \\ u(t-1), u(t-2), \dots, u(t-n_u) \end{matrix} \right], \quad (1)$$

where $y(t), y(t-1), y(t-2), \dots, y(t-n_y)$ represent time series at different time and $u(t-1), u(t-2), \dots, u(t-n_u)$ represent exogenous inputs at different time, y denotes the natural logarithm of AHTPA-EC₅₀, and u denotes the natural logarithm of coding AAS value. According to the characteristics of AAS and its corresponding AHTPA-EC₅₀, the AAS is defined as the input parameter affecting AHTPA-EC₅₀ here.



2.1.6 Neural Network Implementation of Model

The NARX model of AHTPA-EC₅₀ and AAS was realized by the NARXNN. This neural network was performed in Matlab. The two neural network structures are shown in **Figure 9**. The mean square error (MSE) is selected as the performance function of NARXNN. The Levenberg–Marquardt algorithm is used for net training. The division ratio of training set, verification set, and test set in neural network learning samples is 0.7:0.15:0.15. The delay corresponding to the two constructed series is 1:3 and 1:2, respectively, and the hidden layer has 10 neurons.



2.1.7 Prediction Method for AHTPA-EC₅₀

We further proposed a method for AHTPA-EC₅₀ prediction. This method includes two parts: classification and AHTPA-EC₅₀ prediction. The ML algorithm is used to classify the AAS. The classification corresponds to different digital segments of AHTPA-EC₅₀. The feature representation is necessary in this process. This prediction method is described in **Figure 10**. Support vector machine (SVM) is used for classification in this research.

3 Results

3.1 Prediction Results of the Proposed Model

As mentioned above, there are 559 groups of samples in total. However, these data include different antihypertensive

peptides, whose length is from 2 to 20. We select the samples of AAS, whose length is fixed. There are 231 samples of dipeptides and tripeptides in our dataset. They are larger than other peptides. These data are used to verify the proposed model and prediction model. We also constructed two series according to the above method. The first 200 groups in the first series of samples are used for training, and the last 31 data are used for validation and testing. The training results of the constructed series are shown in **Figure 11**.

For the first NARXNN corresponding to the first group series, the training error is 4.895193, the validation error is 4.636605, and the testing error is 3.546904. For the second NARXNN corresponding to the second group series, the training error is 0.001881, the validation error is 0.124045, and the testing error is 0.010165. The second NARXNN has high accuracy; however, it needs the sorting number, and it cannot be used for prediction alone. The classification of the proposed prediction method can provide a rough location in the series. The first NARXNN also gives an original estimation value of AHTPA-EC₅₀. The AHTPA-EC₅₀ will be predicted in the segment of the second series, and two known term AASs help in prediction. The known AASs are selected by the rough location and original estimation value. The second NARXNN is trained every time; therefore, the output will be changed slightly. The first and second NARXNNs are trained in **Figures 11A,B**.

The AHTPA-EC₅₀ of dipeptides and tripeptides is used to verify the prediction method. The first 200 groups in the first series of samples are used for training, and the last 31 data in the first series are used for testing. The proposed method demands classification, and we assume that the classification is correct here; thus, we input the AAS in segments. And the classification is designed as three classification. AHTPA-EC₅₀ = 1, and median values of the series are designed as segment points. The results of prediction are shown in **Figure 12**. Therefore, when the classification is correct, the MSE is 0.5589. We also designed a backpropagation neural network (BPNN) for comparison. The network structure is designed as 3–10–1. The mean square error (MSE) is selected as a performance function. The Levenberg–Marquardt algorithm is used for net training. The logsig function is set as the input function, and the pure linear function is used in the second layer. The number of iterations is set to 1000, the learning rate is 0.1, and the learning target is 0.00001. The results are shown in **Figure 13**, where test samples are randomly selected 100 times. The results reveal that the proposed method has better accuracy than the BPNN.

3.2 Classification of AAS for AHTPA-EC₅₀

As mentioned above, the proposed prediction method demands a rough position which is used in NARX2 prediction. Two classification and three classification are designed for the proposed prediction method here. SVM is used for the classification of AHTPA-EC₅₀ and its corresponding AAS here. We classify the AAS whose length is less than three amino acids. 231 samples of dipeptides and

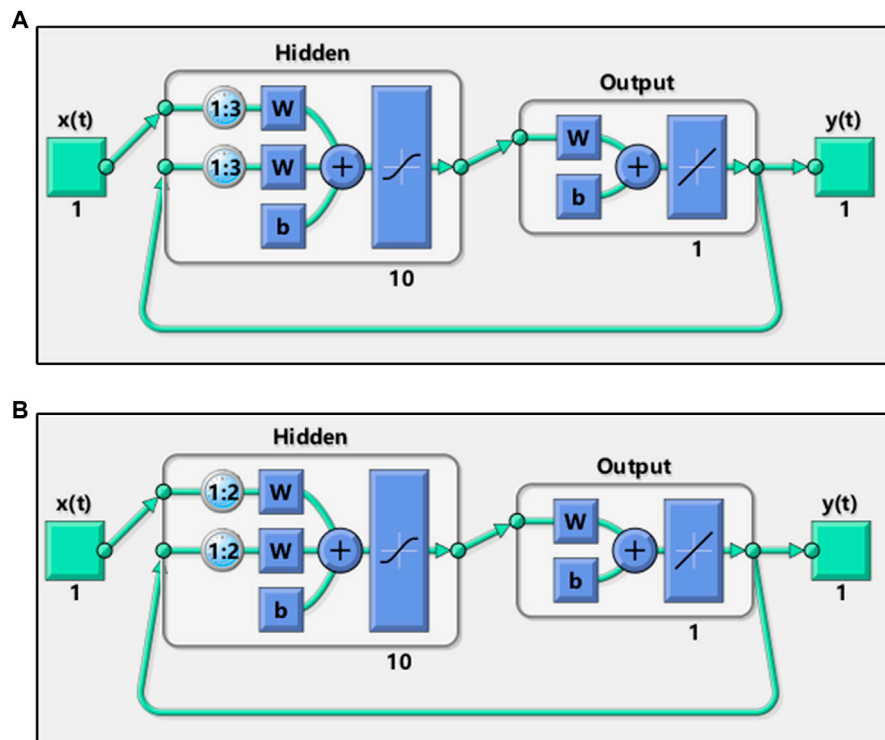


FIGURE 9 | Structures of the neural network for the (A) first series and (B) second series.

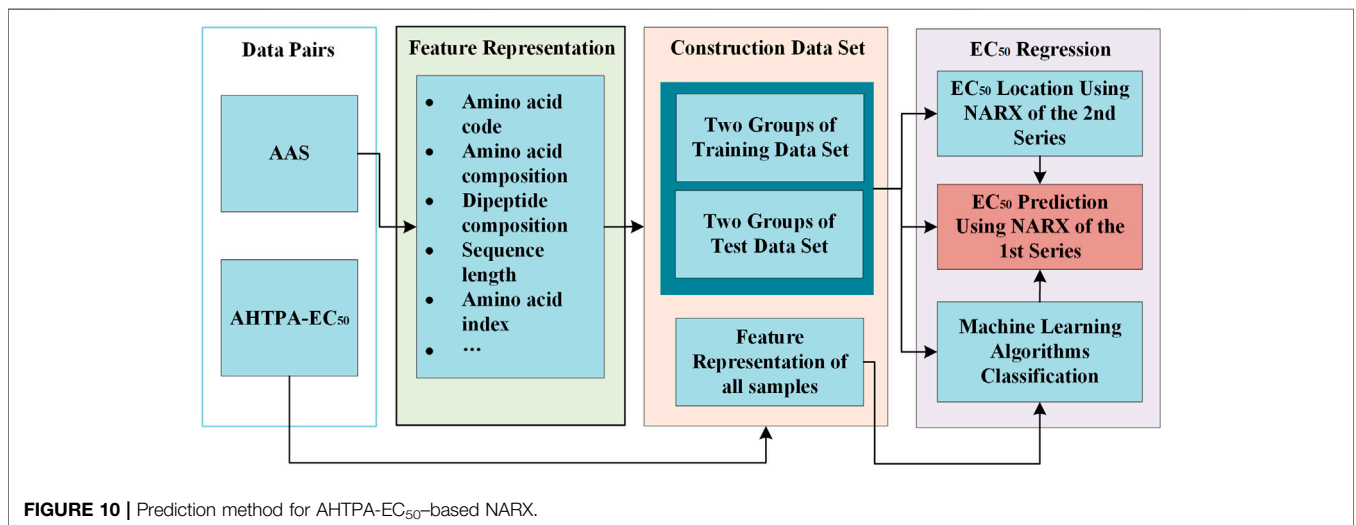
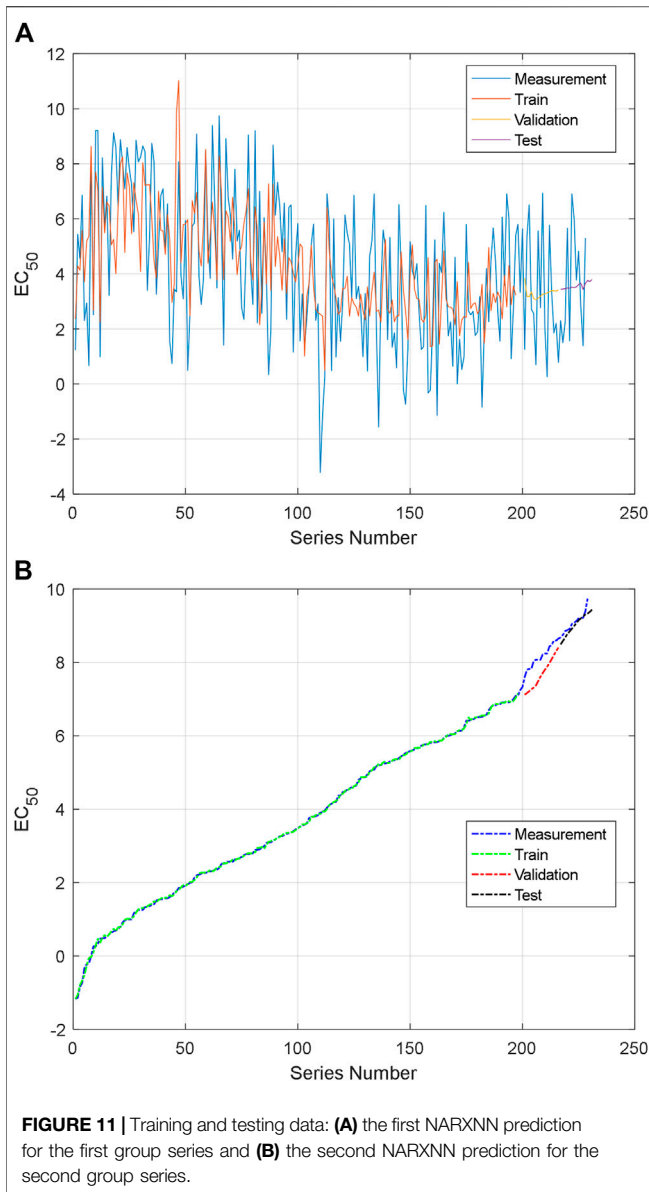


FIGURE 10 | Prediction method for AHTPA-EC₅₀-based NARX.

tripeptides are classified here. For three classification, AHTPA-EC₅₀ = 1 and median values of the series are designed as segment points. For two classification, the median value of the series is designed as the segment point. The label design is shown in Figure 14.

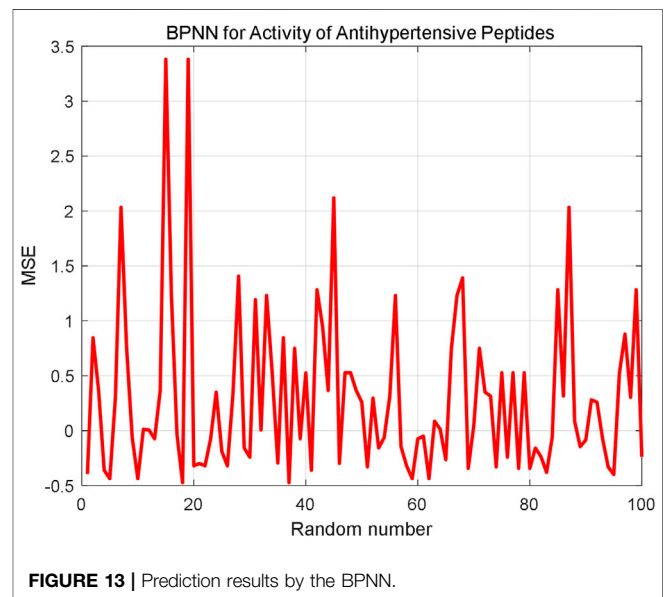
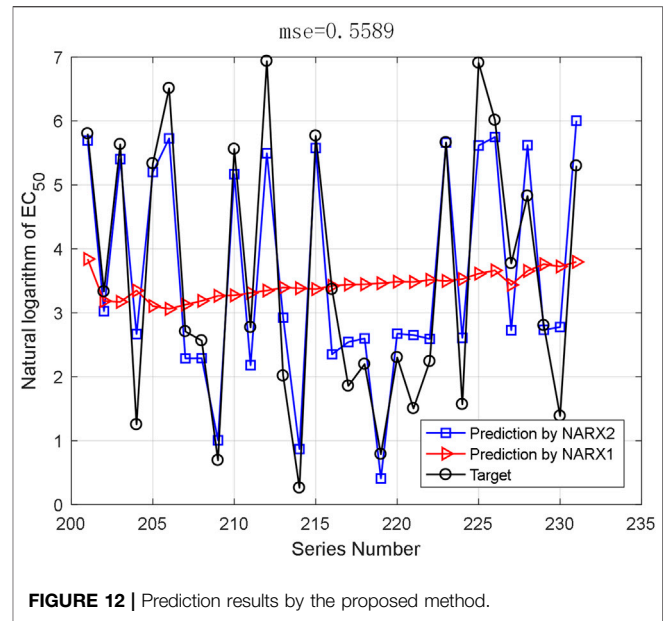
For two classification, there are 161 training data pairs and 70 testing data pairs which are used for classification. And eight feature descriptors are extracted from the peptide sequence. They are the amino acid composition, the digital description of AAS, the peptide sequence code, and the length of peptide sequence.



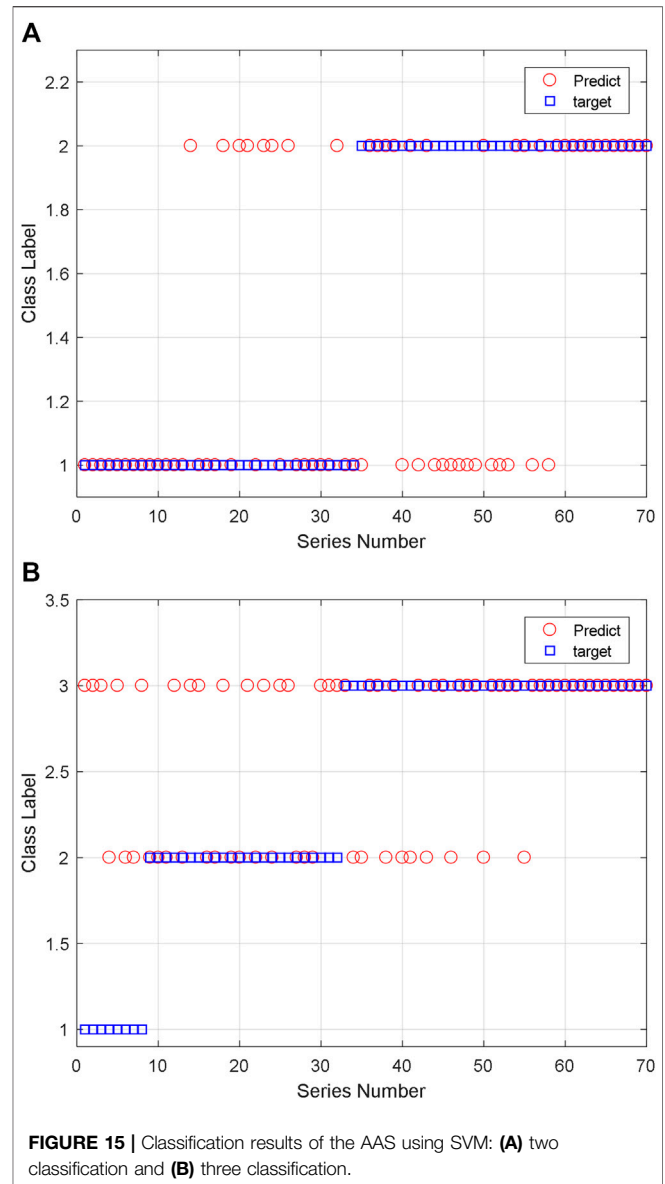
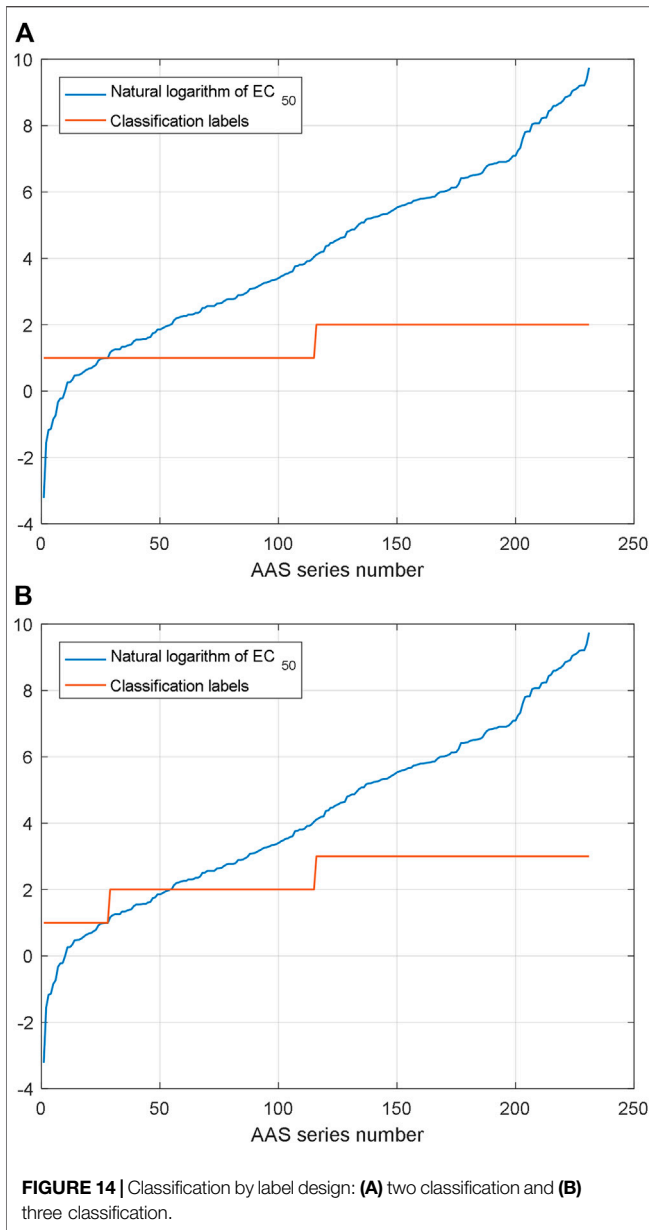
The classification results are shown in **Figure 15**. We can see that the two classification accuracy is 68.57% and the three classification accuracy is 60.00%. Due to the limitations in training, the effect of three classification is not very good. If the quantity of training sample increases and other ML algorithms are also used, we think the accuracy can be improved.

4 Conclusion

In this paper, the statistical distribution of AHTPA-EC₅₀ is analyzed. Two group time series are constructed between AHTPA-EC₅₀ and its corresponding AAS. According to the characteristics of constructed time series, AHTPA-EC₅₀ is modeled by the NARX model. Then, a prediction method of AHTPA-EC₅₀ is proposed. Dipeptides and tripeptides are used



to verify the proposed model and prediction method. The results show that the MSE is 0.5589 when the classification is correct. Finally, we tried to classify the dipeptide and tripeptide data by SVM. Although the accuracy of classification is not very high, it is still feasible. The proposed model and prediction method provide a solution for AHTPA-EC₅₀ prediction, and they are useful and meaningful on antihypertensive active peptide research, drug design, and industrial production (Chen et al., 2020; Granger and Joyeux 1980).



DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, and further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XX and CZ designed the algorithm. DW and MD proposed the problem, pointed research direction, and provided the dataset. XX wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 31730069 and 31371805).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.801728/full#supplementary-material>

REFERENCES

- Al-Hamdan, M., Cruise, J., Rickman, D., and Quattrochi, D. (2010). Effects of Spatial and Spectral Resolutions on Fractal Dimensions in Forested Landscapes. *Remote Sensing* 2, 611–640. doi:10.3390/rs2030611
- Al-Hamdan, M. Z., Cruise, J. F., Rickman, D. L., and Quattrochi, D. A. (2012). Characterization of Forested Landscapes from Remotely Sensed Data Using Fractals and Spatial Autocorrelation. *Adv. Civil Eng.* 2012, 1–14. doi:10.1155/2012/945613
- Baldi, P., Brunak, S., and Bach, F. (2001). *Bioinformatics: The Machine Learning Approach*. Cambridge: Mass MIT Press.
- Chen, Z., Pang, M., Zhao, Z., Li, S., Miao, R., Zhang, Y., et al. (2020). Feature Selection May Improve Deep Neural Networks for the Bioinformatics Problems. *Bioinformatics* 36 (5), 1542–1552. doi:10.1093/bioinformatics/btz763
- Chow, W. C. (2011). Fractal (Fractional) Brownian Motion. *Wires Comp. Stat.* 3 (2), 149–162. doi:10.1002/wics.142
- Engle, R. F., and Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55 (2), 251–276. doi:10.2307/1913253/10.2307/1913236
- Fouché, W. L., and Mukeru, S. (2013). On the Fourier Structure of the Zero Set of Fractional Brownian Motion. *Stat. Probab. Lett.* 83, 459–466. doi:10.1016/j.spl.2012.10.015
- Ghosh, J. K., and Somvanshi, A. (2008). Fractal-based Dimensionality Reduction of Hyperspectral Images. *J. Indian Soc. Remote Sens* 36, 235–241. doi:10.1007/s12524-008-0024-0
- Granger, C. W. J., and Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *J. Time Ser. Anal.* 1, 15–29. doi:10.1017/CBO9780511753978.01810.1111/j.1467-9892.1980.tb00297.x
- Hannan, E. J., and Deistler, M. (1988). *The Statistical Theory of Linear Systems*. New York: Wiley, 5–48.
- Kim, T. S., and Kim, S. (2004). Singularity Spectra of Fractional Brownian Motions as a Multi-Fractal. *Chaos, Solitons & Fractals* 19, 613–619. doi:10.1016/S0960-0779(03)00187-5
- Kumar, R., Chaudhary, K., Sharma, M., Nagpal, G., Chauhan, J. S., Singh, S., et al. (2015a). AHTPDB: a Comprehensive Platform for Analysis and Presentation of Antihypertensive Peptides. *Nucleic Acids Res.* 43, D956–D962. doi:10.1093/nar/gku1141
- Kumar, R., Chaudhary, K., Singh Chauhan, J., Nagpal, G., Kumar, R., Sharma, M., et al. (2015b). An In Silico Platform for Predicting, Screening and Designing of Antihypertensive Peptides. *Sci. Rep.* 5, 12512. doi:10.1038/srep12512
- Libbrecht, M. W., and Noble, W. S. (2015). Machine Learning Applications in Genetics and Genomics. *Nat. Rev. Genet.* 16 (6), 321–332. doi:10.1038/nrg3920
- Majumder, K., and Wu, J. (2010). A New Approach for Identification of Novel Antihypertensive Peptides from Egg Proteins by QSAR and Bioinformatics. *Food Res. Int.* 43, 1371–1378. doi:10.1016/j.foodres.2010.04.027
- Manavalan, B., Basith, S., Shin, T. H., Wei, L., and Lee, G. (2019). mAHTPred: a Sequence-Based Meta-Predictor for Improving the Prediction of Antihypertensive Peptides Using Effective Feature Representation. *Bioinformatics* 35 (16), 2757–2765. doi:10.1093/bioinformatics/bty1047
- Moran, P. (1953). The Statistical Analysis of the Canadian lynx Cycle. *Aust. J. Zool.* 1, 291–298. doi:10.1071/ZO9530291
- Nill, N. B., and Bouzas, B. H. (1992). Objective Image Quality Measure Derived from Digital Image Power Spectra. *Opt. Eng.* 31, 813–825. doi:10.1117/12.56114
- Ozaki, T. (1980). Non-linear Time Series Models for Non-linear Random Vibrations. *J. Appl. Probab.* 17, 84–93. doi:10.1017/S0021900200046829
- Pentland, A. P. (1984). Fractal-Based Description of Natural Scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-6, 661–674. doi:10.1109/TPAMI.1984.4767591
- Ruderman, D. L. (1996). Origins of Scaling in Natural Images. *Vis. Res.* 37, 3385–3398. doi:10.1117/12.238707
- Stone, M. E. (20182018). *Kaplan's Essentials of Cardiac Anesthesia*. ISBN 978-0-323-49798-5. Elsevier. Mechanisms of Action: ACE inhibitors act by inhibiting one of several proteases responsible for cleaving the decapeptide Ang I to form the octapeptide Ang II. doi:10.1016/c2012-0-06151-0
- Tong, H. (1990). *Nonlinear Time Series: A Dynamical Systems Approach*. Oxford: Oxford University Press, 14–37.
- Tong, H. (1983). *Threshold Models in Non-linear Time Series Analysis*. New York: Springer-Verlag, 7–34.
- Tong, J., Liu, S., Zhou, P., Wu, B., and Li, Z. (2008). A Novel Descriptor of Amino Acids and its Application in Peptide QSAR. *J. Theor. Biol.* 253, 90–97. doi:10.1016/j.jtbi.2008.02.030
- Tu, M., Cheng, S., Lu, W., and Du, M. (2018a). Advancement and Prospects of Bioinformatics Analysis for Studying Bioactive Peptides from Food-Derived Protein: Sequence, Structure, and Functions. *Trac Trends Anal. Chem.* 105, 7–17. doi:10.1016/j.trac.2018.04.005
- Tu, M., Wang, C., Chen, C., Zhang, R., Liu, H., Lu, W., et al. (2018b). Identification of a Novel ACE-Inhibitory Peptide from Casein and Evaluation of the Inhibitory Mechanisms. *Food Chem.* 256, 98–104. doi:10.1016/j.foodchem.2018.02.107
- Win, T. S., Schaduangrat, N., Prachayasittikul, V., Nantasenamat, C., and Shoombuatong, W. (2018). PAAP: a Web Server for Predicting Antihypertensive Activity of Peptides. *Future Med. Chem.* 10, 1749–1767. doi:10.4155/fmc-2017-0300
- Wornell, G. W., and Oppenheim, A. V. (1992). Estimation of Fractal Signals from Noisy Measurements Using Wavelets. *IEEE Trans. Signal. Process.* 40, 611–623. doi:10.1109/78.120804
- Wu, D., Tu, M., Wang, Z., Wu, C., Yu, C., Battino, M., et al. (2020). Biological and Conventional Food Processing Modifications on Food Proteins: Structure, Functionality, and Bioactivity. *Biotechnol. Adv.* 40, 107491. doi:10.1016/j.biotechadv.2019.107491
- Yang, Z., Wang, J., Yang, J., Qi, Z., He, J., and He, J. (2020). Recognizing Proteins with Binding Function in Elymus Nutans Based on Machine Learning Methods. *Comb. Chem. High Throughput Screen.* 23 (6), 554–562. doi:10.2174/1386207323666200330120154
- Zhang, R.-z., Xu, X.-h., Chen, T.-b., Li, L., and Rao, P.-f. (2000). An Assay for Angiotensin-Converting Enzyme Using Capillary Zone Electrophoresis. *Anal. Biochem.* 280 (2), 286–290. doi:10.1006/abio.2000.4535
- Zhang, Z., Wang, J., and Liu, J. (2021). DeepRTCP: Predicting ATP-Binding Cassette Transporters Based on 1-Dimensional Convolutional Network. *Front. Cel. Dev. Biol.* 8, 614080. doi:10.3389/fcell.2020.614080
- Zou, Q., and Liu, Q. (2019). Advanced Machine Learning Techniques for Bioinformatics. *Ieee/acm Trans. Comput. Biol. Bioinf.* 16 (4), 1182–1183. doi:10.1109/TCBB.2019.2919039

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xie, Zhu, Wu and Du. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.