



Chaotic Modeling of Stream Nitrate Concentration and Transportation via IFPA-ESN and Turning Point Analyses

Yandong Tang¹, Jiahao Deng^{2*}, Cuiping Zang¹ and Qihong Wu³

¹Sichuan Engineering Technical College, Deyang, China, ²College of Computing and Digital Media, DePaul University, Chicago, IL, United States, ³School of Architecture and Civil Engineering, Chengdu University, Chengdu, China

OPEN ACCESS

Edited by:

Jingren Zhou,
Sichuan University, China

Reviewed by:

Ge Gao,
McGill University, Canada
Tinghui Ouyang,
National Institute of Advanced
Industrial Science and Technology
(AIST), Japan

*Correspondence:

Jiahao Deng
jdeng5@depaul.edu

Specialty section:

This article was submitted to
Environmental Informatics and Remote
Sensing,
a section of the journal
Frontiers in Environmental Science

Received: 15 January 2022

Accepted: 04 February 2022

Published: 04 March 2022

Citation:

Tang Y, Deng J, Zang C and Wu Q
(2022) Chaotic Modeling of Stream
Nitrate Concentration and
Transportation via IFPA-ESN and
Turning Point Analyses.
Front. Environ. Sci. 10:855694.
doi: 10.3389/fenvs.2022.855694

Increased concentrations of nitrogenous compounds in stream networks are detrimental to the health of both humans and ecosystems. Monitoring, modeling, and forecasting nitrate concentration in the temporal domain are essential for an in-depth understanding of nitrate dynamics and transformation within stream networks. In this study, an advanced chaotic modeling and forecasting approach integrated with turning point analysis is proposed. First, the time-series daily nitrate concentrations in the form of nitrate-nitrite were reconstructed based on the chaotic characteristics and then input into the forecasting models. Second, an echo state network (ESN) was developed for one-day-ahead nitrate concentration forecasting, and the hyperparameters were optimized through an improved flower pollination algorithm (IFPA) to achieve a high efficiency. Furthermore, turning point analysis was performed to quantify the relationship between discharge and peak nitrate concentration. The Ricker function was fitted, and the parameters were estimated for turning points using the forecasted nitrate concentration and measured discharge. Field data, including daily stream nitrate concentration and information on discharge collected from eight different monitoring sites in the southern Sichuan Basin, China, were utilized for case studies. A comparative analysis was performed under three modeling scenarios, viz. conventional time-series modeling, temporal signal decomposition, and data reconstruction and embedding with chaotic characteristics. Four benchmark time-series forecasting algorithms were compared against the proposed IFPA-ESN in the above-mentioned scenarios. For each site, parameters of the Ricker functions were estimated, and turning points were computed based on the forecasted nitrate concentration and discharge. Computational results validated the superiority of the proposed approach in improving the accuracy of stream nitrate concentration prediction. The limitations to the supply and transportation of nitrogenous compounds were quantified, which would be valuable for pollution mitigation in the future.

Keywords: nitrate concentration, chaotic system, improved flower pollination optimization, echo state network, turning point analysis

INTRODUCTION

In view of recent environmental challenges, river pollution is one of the most widespread problems due to industrial, economic, and agricultural growth (Domangue and Mortazavi 2018). Inorganic nitrogen, which is usually in the form of nitrate-nitrite ($\text{NO}_x\text{-N}$), is one of the major pollutants in stream networks. According to recent monitoring data, surface water systems are increasingly being polluted by nitrogen, with an increase in nitrogen concentration annually (Sajedi-Hosseini et al., 2018). Furthermore, the high concentrations of $\text{NO}_x\text{-N}$ in drinking water pose increased risks of various diseases (e.g., cancer) to humans. However, water treatment to reduce nitrogen compounds is challenging owing to the high dynamics of both pollutant emissions and stream discharge impacted by precipitation and groundwater in the temporal domain (Zhang et al., 2018; Shi et al., 2019). Therefore, the most practical and effective approach to prevent and control river nitrogen pollution is to forecast daily variation and investigate its transportation mechanism quantitatively.

Nitrate normally exists in nature and is the end product of the aerobic decomposition of organic nitrogenous matter and microorganisms. In stream networks with clean surface water, most water samples contain less than 1 mg L^{-1} of $\text{NO}_x\text{-N}$; according to Stamenković et al. (2020), it is seldom observed to surpass 5 mg L^{-1} . However, in some areas, several agricultural landscapes are artificially saturated with excessive nitrogen compounds to maximize crop yield (Blesh and Drinkwater, 2013). This inevitably results in the accumulation of immense amounts of this nutrient in the soil (Jones et al., 2017). Subsequently, artificial drainage or excessive precipitation transports the soluble $\text{NO}_x\text{-N}$ into the aquifers, lakes, and streams near agricultural landscapes. As a result, the river systems encounter a nearly unlimited supply of $\text{NO}_x\text{-N}$, and the concentration is largely driven by stream discharge (Villarini et al., 2016).

The challenge is evident in the recent and ongoing rise in the nitrate pollution problem, i.e., how to estimate $\text{NO}_x\text{-N}$ concentration in the stream and quantify the relationship between discharge and concentration. Improved forecasts could aid in: 1) accurate estimation of nitrate concentration in real time on a daily basis and 2) adequate quantitative modeling of the nonlinear relationship between stream discharge and concentration.

Two major types of methods for modeling and forecasting of nitrate pollution in stream networks are conventional water quality modeling methods and statistical methods. In the last few decades, water quality modeling has served as the foundation for estimating the water nitrate concentration. David et al. (1997) discovered that high soil mineralization rates, fertilization, and tile drainage contribute notably to nitrate transport to rivers. Belitz et al. (2015) studied the transportation mechanism of nitrate pollutants in underground water systems using water quality models. Keupers and Willems (2017) developed advection-dispersion models to investigate the transportation of nitrate pollutants in stream systems with respect to stream discharge, water height, and velocity. All these authors used

mathematical relations to model the nitrate concentrations in surface waters, which constitute the basis for both forecasting and inference.

In the literature, machine-learning models developed on the basis of statistical methods have gained more attention in recent studies. Nolan et al. (2014) constructed random forest models to predict log nitrate for domestic and public supply wells in California. Ransom et al. (2017) further improved the prediction accuracy by developing a boosting regression tree algorithm to predict nitrate concentration in the Central Valley aquifer, California, United States. Ostad-Ali-Askari et al. (2017) developed artificial neural networks to model nitrate concentrations in groundwater in the temporal domain. Stamenković et al. (2020) constructed a multilayer perceptron and predicted nitrate concentrations using 26 water quality parameters as inputs. In summary, the highly nonlinear relationships between the inputs and nitrate concentrations were “learned” and “trained” with field data using the machine-learning models. All machine-learning approaches have greatly improved the accuracy of nitrate concentration forecasts in streams and offer insights into data patterns in the spatial and temporal domains.

Deep learning approaches have become a more promising approach for forecasting tasks in the temporal domain. Basic deep learning algorithms include deep belief networks (Ouyang et al., 2019a), convolutional neural network (LeCun and Bengio 1995), deep neural network (DNN) (Hu et al., 2016), long short-term memory recurrent neural network (LSTM-RNN) (Kong et al., 2017), and stacked extreme learning machine (Huang et al., 2011). The applications of deep learning approaches in forecasting tasks include deterministic forecasting methods, deep-learned feature extraction, error post-processing, and network structure optimization (He et al., 2017a; Li et al., 2018; Li et al., 2022; Ouyang et al., 2018; Ouyang et al., 2019b; Xu et al., 2019; Ahmad et al., 2021; Hrnjica et al., 2021; Tang et al., 2021). The main challenge in using deep learning techniques is to construct the most fitting prediction model for a particular dataset, such as the nitrate concentration in this study.

With an increase in nitrate pollution in stream networks, nitrate concentrations in stream networks have become a complicated system. Forecasting with high precision is the key factor in both the long and short term, and the computational cost is also important. Conventional short-term forecasting tasks in the temporal domain can be classified into two approaches, namely time-series modeling, and data decomposition. The time-series approaches use an auto-correction function and seasonality analysis to determine the temporal dependency patterns in the dataset and construct forecasting models. On the other hand, decomposition approaches decompose the time-series data into several basis signals and build sub-models one-by-one to perform forecasting. Then, the final prediction is made by combining the forecasting results from the sub-models.

In this study, in addition to the two major types of conventional approaches, we proposed a novel framework using a chaotic modeling approach to forecast the nitrate concentrations in stream networks. First, we considered that the daily nitrate concentration is generated from a chaotic

system, and we can compute the chaotic characteristics. The reconstructed data based on chaotic characteristics can contain both numerical features and structural information from the original physical systems. Second, we proposed forecasting models using an echo state network (ESN), wherein the hyperparameters were optimized using an improved flower pollination algorithm (IFPA). Furthermore, a comparative analysis was performed against conventional forecasting approaches and other benchmarking deep learning forecasting algorithms. Third, to quantify the relationship between daily nitrate concentrations and stream discharge, we performed a turning point analysis. The parametric Ricker function was constructed to depict the nonlinear relationship, and the parameters were estimated using the delta method. Finally, based on the computed turning points, we estimated the transportation limitations of nitrate pollutants in streams.

The main contributions of this study are as follows:

- This study proposes a deep learning approach, integrated with chaotic models, to forecast short-term nitrate concentrations in stream networks. Onsite data were collected from a closed source at eight monitoring locations spread across three different regions.
- The ESN was first applied for the prediction of nitrate concentration, and its hyperparameters were optimized using the IFPA algorithm.
- The forecasting accuracy was compared with other forecasting approaches and other benchmarking deep learning algorithms.
- Turning point analysis was performed to quantify the relationship between stream discharge and nitrate concentration. The transportation limit of nitrate pollutants was also investigated.

The remainder of this paper is organized as follows. **Section 2** describes the details of the dataset and the case study area. **Section 3** introduces the chaotic modeling approach for forecasting nitrate concentration in stream networks. The forecasting algorithm ESN optimized with IFPA, as well as other benchmarking deep learning algorithms, are discussed. **Section 4** presents the detailed procedures of turning point analysis and the investigation of transportation limits. Experimental results are provided and discussed in **Section 5**. Finally, the conclusions are presented in **Section 6**.

DATA COLLECTION AND PRE-PROCESSING

Data Collection and Description

The daily nitrate concentration data and discharge data has been collected from our case study area located in south Sichuan Basin, southwestern China. It is one of the most heavily populated and intensely industrialized areas in China which faces serious environmental pollution problems (Cui et al., 2021; Zhou et al., 2021). This area contains approximately about 10% of the national population and many big cities including Chengdu,

Leshan, Yibin, Luzhou, Jiangjin, and Chongqing are located in this area. Two big rivers namely Minjiang River and Yangtze Rivers flow across this region and many stream networks are affiliated with the two rivers which connects the cities.

The emission of nitrogen compounds especially like NO_x-N nitrates is mainly from the agricultural activities, residents, and industry. The deposition of NO_x-N degrades the quality of both groundwater and surface water in regional stream networks. In order to measure the pollutant, the nitrate concentrations in 8 monitoring points in the streams locate in the suburbs of major cities along the Yangtze River are measured on daily basis during the year of 2020. The locations of major cities and the monitoring points which are in the suburbs of these cities are displayed in **Figure 1** below.

The dataset collected for this study includes daily contrite concentrations from 8 monitoring points. Meanwhile, the daily discharge volume from surrounding area is also provided. The summary of the dataset has been presented in **Table 1**. It summarizes the basic statistics including the mean, median, maximum value, minimum value, and standard deviation of the daily nitrate concentrations as well as stream discharge. The unit for the measured nitrate concentration is milligrams per liter (mg L⁻¹) and the unit for the discharge is millimeter (mm) which is mainly from precipitation or groundwater supply.

According to **Table 1**, daily concentrations of nitrate concentration (NO_x-N) ranged from 7.05 to 35.00 mg L⁻¹. Concentrations that exceeded the drinking water standards of 10 mg L⁻¹ for most of the time for almost all monitoring points. This may be exaggerated by the fact that the records in many summer flood seasons are not included for the safety issues during the data collection process. Meanwhile, for 8 monitoring sites, the average daily discharge on days when NO_x-N was measured ranged between 1.32 and 1.51 mm during the period of the study. These data also exclude the flood seasons as the observations may be inaccurate and meaningless. Each year, the nitrate concentration varied with discharge during the period of study. Each year, concentrations tend to increase with discharge in a nonlinear fashion.

Abnormal Data Detection

In general, data abnormality is a rarely observed phenomena when some human or systematic errors occur during the measuring process (Gao et al., 2020; Gao et al., 2021). To increase the reliability of the models in our study, we proposed to use “six-sigma” criterion from statistical analysis in our data pre-processing part to detect data abnormality. For a daily nitrate concentration time-series, assumed as { x_n }, the mean and standard deviation can be defined as \bar{x} and σ respectively. Then, we can use the “six-sigma” to detect whether a new coming data sample $x(t)$ locates in the abnormal region or not. In detail, the “six-sigma” criterion can be expressed in **Eq. 1** as follows:

$$\bar{x}(t-1) - k\sigma(t-1) < x(t) < \bar{x}(t-1) + k\sigma(t-1) \quad (1)$$

where $\bar{x}(t-1)$ denotes the mean of all former measured nitrate concentration or discharge; $\sigma(t-1)$ is the standard deviation of the historic data; k is the parameter for the scale of the criterion

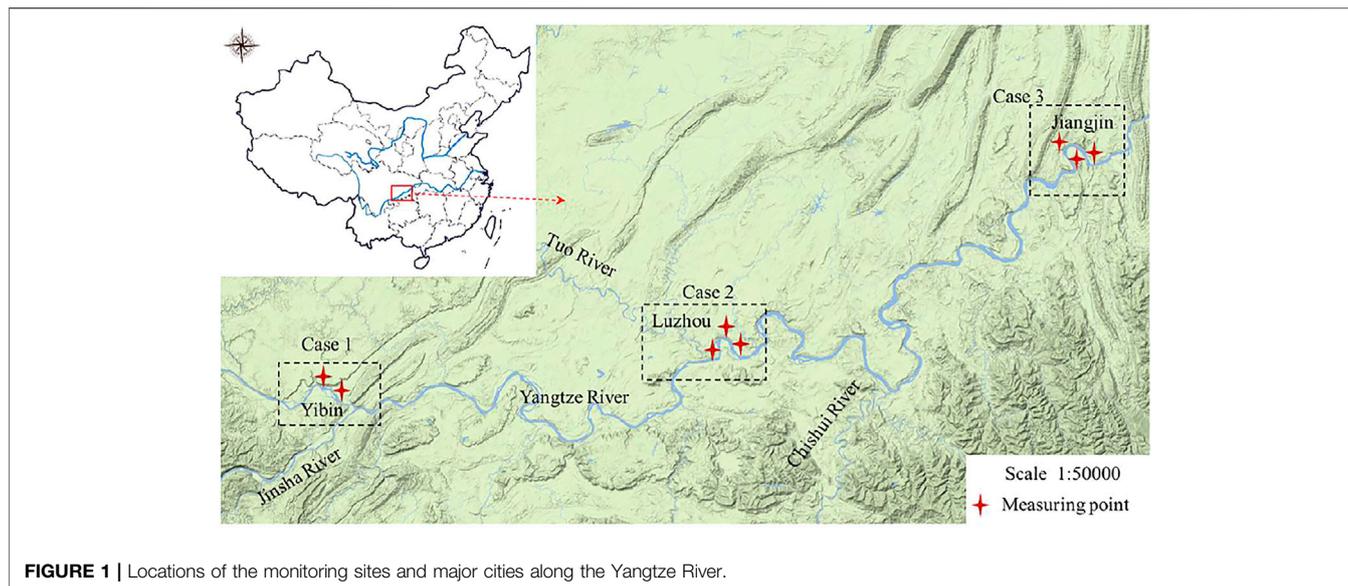


TABLE 1 | Summary of the dataset including daily nitrate concentration and discharge.

Location	n (days)	NO _x -N (mg L ⁻¹)					Discharge (mm)				
		Mean	Median	Max	Min	Std	Mean	Median	Max	Min	Std
Yibin-1	320	16.81	15.86	35.00	4.00	6.25	1.52	1.51	3.35	0.03	0.42
Yibin-2	297	7.05	6.38	21.00	0.01	4.57	1.36	1.34	3.36	0.02	0.46
Luzhou-1	274	18.76	17.56	35.00	4.00	7.67	1.41	1.32	3.33	0.05	0.32
Luzhou-2	298	13.58	12.60	25.00	2.00	6.05	1.39	1.33	3.35	0.01	0.37
Luzhou-3	301	12.57	11.76	22.00	3.00	5.00	1.32	1.37	3.34	0.04	0.33
Jiangjin-1	303	15.92	14.91	35.00	4.00	6.95	1.38	1.39	3.23	0.03	0.15
Jiangjin-2	311	12.99	12.29	27.00	4.00	5.23	1.50	1.65	3.36	0.02	0.53
Jiangjin-3	299	8.42	7.85	20.00	1.00	4.32	1.51	1.57	3.36	0.04	0.42

which is usually set as 3. If the new measured data is falling outside of the range defined in Eq. 1, we can simply delete it and replaced with revised values.

CHAOTIC MODELING OF STREAM NITRATE CONCENTRATION

Dataset Reconstruction

Time-series data representing the daily nitrate concentration usually include dynamic information of the original system. With a certain level of dynamics, the original system is considered as a chaotic system. If the time-series nitrate data are embedded into a higher space, its intrinsic data structure can be reproduced (Ouyang et al., 2020). Hence, reconstruction of the phase space is a valuable solution for modeling chaotic nitrate concentration time series. In general, given the original nitrate concentration time series $\{x_n\}$ for phase space reconstruction, the data in the new higher space can be expressed by Eq. 2 as follows:

$$\begin{aligned} x_n &= (x_n, x_{n+\tau}, \dots, x_{n+(m-1)\tau}) \in R^m, n = 1, 2, \dots, N_0 \\ &= N - (m - 1)\tau \end{aligned} \quad (2)$$

where x_n denotes the reconstructed nitrate concentration time series in the higher space R^m ; N denotes the number of total data points in the original nitrate concentration time series; N_0 denotes the number of data points in the reconstructed time-series in the higher space, and τ and m are two reconstruction parameters that represent the time delay and embedded dimension, respectively (Rosenstein et al., 1994). As explained by Ouyang et al. (2016) in the embedding theory, the data features in both the original and reconstructed space will be equally dynamic and diffeomorphic.

Temporal Dependency

The time delay τ , described in Section 3.1, represents the temporal dependency between the current time series and historic lagged ones. An appropriate time delay parameter value is essential for unfolding and organizing the time series in the phase reconstruction process. In this study, mutual information was introduced and applied to compute the temporal dependency of the current nitrate concentration time series and historic data.

With two discrete nitrate concentration time series defined as $x(t)$ and $x(t + \tau)$, the temporal dependency between them can be computed using Eq. 3 as follows:

$$I(x(t), x(t + \tau)) = H(x(t)) + H(x(t + \tau)) - H(x(t), x(t + \tau)) \quad (3)$$

where $I(x(t), x(t + \tau))$ denotes the mutual information entropy between the two series, where a larger value indicates a strong temporal correlation and vice versa, and the function $H(*)$ computes the information entropy, which can be expressed by Eq. 4 as follows:

$$H(X) = - \sum P(X) \log(P(X)) \quad (4)$$

where $P(X)$ denotes the probability of the variable X , and $H(X)$ is the information entropy. Hence, for convenience, the mutual information entropy between the time series $x(t)$ and $x(t + \tau)$ can be expressed by Eq. 5 as follows:

$$\begin{aligned} I(x(t), x(t + \tau)) = & - \sum P(x(t)) \log(P(x(t))) \\ & - \sum P(x(t + \tau)) \log(P(x(t + \tau))) \\ & + \sum P(x(t), x(t + \tau)) \log(P(x(t), x(t + \tau))) \\ & \times \end{aligned} \quad (5)$$

where $P(x(t), x(t + \tau))$ denotes the joint probability distribution of the two nitrate concentration series.

Embedded Dimension

In addition to the time delay, the embedded dimension is another factor affecting the size of the reconstructed phase space. An unsuitable embedded dimension would cause the track of the time series to fold or intersect in the reconstructed phase space. This reflects the number of independent factors that govern the dynamics of the underlying physics system. Only a suitable embedded dimension can ensure that the reconstructed system contains all the independent factors needed to unfold the original system.

In this study, the false nearest neighbors (FNN) method (Bazine & Mabrouki 2019) was adopted to search for the optimal setting for the number of embedded dimensions. It detects the number of neighboring points for different settings of the embedded dimensions repeatedly. If there is no sharp increase with an increase in the number of false neighbors, the embedded dimension increases from m to $m+1$. Intuitively, the calculations describe variances among false neighbors.

Given two sets of nearest points $x_n(i)$ and $x'_n(i)$, the computation of false nearest neighbors can be achieved following the three steps given below:

Step 1: Compute the distance between the two nearest points using Eq. (6)

$$R_m(i) = x_n(i) - x'_n(i) \quad (6)$$

where $R_m(i)$ is the distance between two points in the m -dimensional phase space. Here, we use the Euclidean distance, which is calculated using the L2-norm.

Step 2: Increase the dimension from m to $m+1$, and the new distance in the new space can be expressed by Eq. 7 as follows:

$$R_{m+1}(i) = \sqrt{R_m^2(i) + x_n(i + m\tau) - x'_n(i + m\tau)^2} \quad (7)$$

Step 3: If the distance between the two nearest points in the new space $x_n(i + m\tau) - x'_n(i + m\tau)^2$ increases significantly, the two points will be classified as false nearest neighbors and will be separated in a higher dimension space. The classification follows the criterion in Eq. 8, as follows:

$$a(i, m) = \frac{x_n(i + m\tau) - x'_n(i + m\tau)}{R_m(i)} * 100\% \quad (8)$$

where $a(i, m)$ is the criterion for judging the false nearest neighbors. The two given samples are classified as false nearest neighbors if $a(i, m) > a_{th}$. Here, a_{th} is the threshold that is usually selected in the range of 0.1 and 0.5.

Lyapunov Exponent

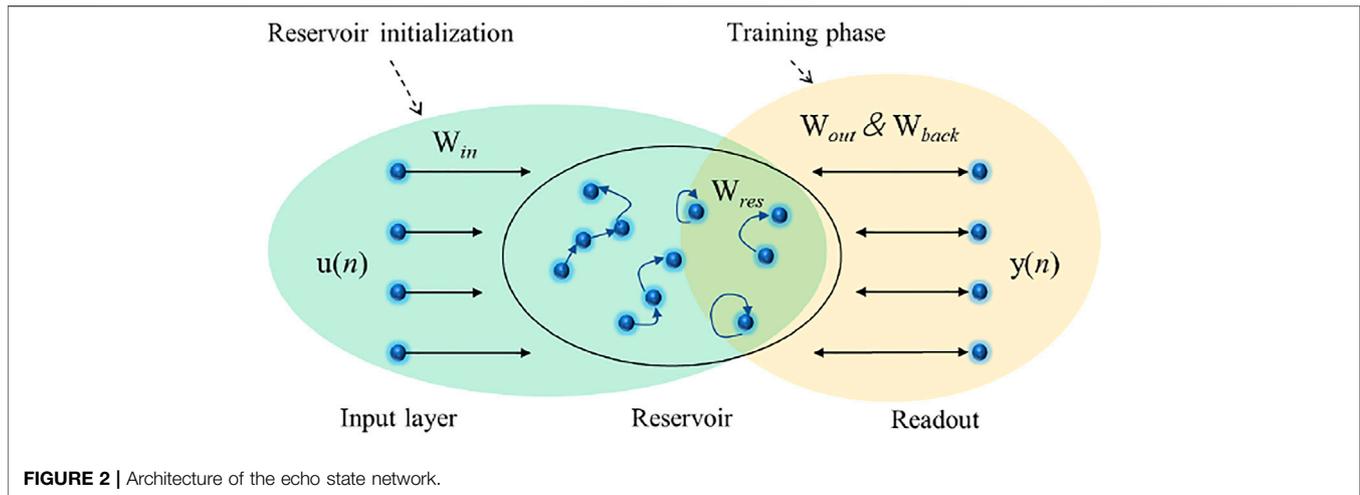
It is seen that the reconstructed data in the new phase space not only retains the original information but also provides more details such as data structures. Hence, studying the chaotic characteristics of the nitrate concentration before data reconstruction is essential.

The Lyapunov exponent is a widely used metric to test whether a system is chaotic or not. It represents the average exponential rates of divergence (expansion) or convergence (contraction) of nearest points in the embedded phase space (Jiang et al., 2021). In general, positive Lyapunov exponent value indicates the presence of chaos and the data structure is unfolded in the reconstructed phase space. If it equals to zero, the system is considered to have bifurcation points or seasonality patterns. Negative values refer to the data points are from stable physics system. Therefore, a positive value of the Lyapunov exponent is needed to demonstrate a phase space is suitable and the reconstructed data is effective to study the original system.

The computation of the Lyapunov exponent can be achieved by applying the approach proposed by Wolf et al. (1985). The process is presented as follows. First, given a start point $x_n(i)$ and its nearest neighbor point $x'_n(i)$, the distance between the two points can be computed as $L_i = x_n(i) - x'_n(i)$. Next, with the evolution of a given time T , two new points $x_n(j)$ and $x'_n(j)$ as nearest neighbors can be obtained and the new distance can be computed as $L_j = x_n(j) - x'_n(j)$. Hence, after compute it repeatedly for k -steps, the value of the Lyapunov exponent can be obtained in Eq. 9 as follows:

$$\lambda = \frac{1}{kT} \sum_{i=1}^k \ln \frac{L_i}{L_{i-1}} \quad (9)$$

In most scenarios, the value of L_i is very small. The Lyapunov exponent can be regarded as the slope coefficient of the linear function between T and $\ln L_i$ (He and Kusiak 2018).



Echo State Network

Using the reconstructed data in the new embedded dimension, the day-ahead nitrate concentrations in stream networks were forecasted in this study. Here, the ESN (Chouikhi et al., 2017), which is a deep recurrent neural network, was selected for forecasting. An ESN comprises three layers, namely the input layer, dynamic reservoir, and output payer. Compared with other conventional deep neural networks (DNNs), the ESN contains a dynamic reservoir consisting of several sparsely connected neurons instead of interconnected hidden layers. This structure enables the ESN to have an improved information-processing ability. The architecture of the ESN is shown in **Figure 2**. Since 2001, ESN has been successfully applied in various fields, including time-series forecasting and classification.

As illustrated in **Figure 2**, the ESN contains n units in the input layer u , N neurons in the reservoir that form the internal states, and n units in the output layer y . Since this study focuses on day-ahead forecasting of nitrate concentration, the number of units in the output layer y is 1. Assume at time step i , the input state matrix is defined as $u(i) = [u_1(i), u_2(i), u_3(i), \dots, u_n(i)]^T$. The reservoir state matrix can be defined as $h(i) = [h_1(i), h_2(i), h_3(i), \dots, h_n(i)]^T$ while the output matrix can be defined as $y(i) = [y_1(i), y_2(i), y_3(i), \dots, y_n(i)]^T$. Then, during iteration, we can formulate the updating rules in **Eq. 10** and **11** as follows:

$$h(i+1) = f(W_{in} * u(i+1) + W_{res} * h(i) + W_{back} * y(i)) \quad (10)$$

$$y(i+1) = g(W_{out} * h(i+1)) \quad (11)$$

where i denotes the number of iterations; W_{in} is the weight matrix that connects the input layer and dynamic reservoir in the size of $N \times n$; W_{out} is the weight matrix that connects the dynamic reservoir and output layer in the size of $1 \times N$; the W_{out} of $N \times 1$ connects the output state and internal reservoir and W_{res} in the size of $N \times N$ enables interconnection of reservoir internal states; $f()$ and $g()$ are activation functions which can be linear or nonlinear. In most scenarios, the linear activation function is selected and so it is in this research.

Improved Flower Pollination Algorithm

The setting of hyperparameters of ESN such as reserve layer nodes, reserve layer connection rate, learning rate, and spectral radius all impact the forecasting performance. The improved flower pollination algorithm (IFPA) (Lei et al., 2018), which is based on the principle of biological pollination, has attracted considerable amount of attention in optimizing hyperparameters of deep learning algorithms.

The IFPA algorithm is a swarm intelligence optimization algorithm which contains three major components: cross-pollination, abiotic-pollination and cross-conversion (Li et al., 2021a; Li et al., 2021b). The cross-pollination is a global search strategy which simulates the insects and birds flying under the Levy distribution as in **Eq. 12** and **13**:

$$L(\lambda) \sim \frac{\lambda \Gamma(\lambda) \sin(\pi\lambda/2)}{\pi} * \frac{1}{S^{1+\lambda}} \quad (12)$$

$$s = \frac{U}{|V|^{1/\lambda}} \quad (13)$$

where U obeys the Gaussian distribution $N(0, \sigma^2)$, and V follows the standard normal distribution $N(0, 1)$. The σ is calculated in **Eq. 14** as follows:

$$\sigma^2 = \left\{ \frac{\Gamma(1+\lambda) \sin(\pi\lambda/2)}{\lambda \Gamma\left[\frac{1+\lambda}{2}\right] 2^{(\lambda-1)/2}} \right\}^{1/\lambda} \quad (14)$$

Hence, the updating rules in cross-pollination can be expressed in **Eq. 15** as:

$$z_i^{iter+1} = z_i^{iter} + \gamma L(z_i^{iter} - z_i^{best}) \quad (15)$$

Furthermore, the abiotic-pollination is a local searching process and the updating rules is expressed in **Eq. 16** as follows:

$$z_i^{iter+1} = z_i^{iter} + \varepsilon(z_j^{iter} - z_k^{iter}) \quad (16)$$

where z_j^{iter} and z_k^{iter} denote different pollens (local solutions) other than z_i^{iter} ; and ε is a constant that obeys the uniform distribution of (0,1).

Third, a conversion probability P is introduced to mimic the principle of biological pollination. It controls the conversion of pollen between global pollination and local pollination pollen within iterations. In each iteration, the probability of having global pollination is P and thus the probability of having local pollination is $1-P$ (Lei et al., 2018; Tang and Zhang, 2019).

Conventional Approaches and Benchmarking Algorithms

To demonstrate the accuracy and effectiveness of the proposed chaotic modeling approach integrated with IFPA-ESN, the comparative analysis is performed in this research against other conventional modeling approach as well as the other benchmarking deep-learning algorithms.

The conventional time-series forecasting approach is based on the autoregression and dependency between residual error and lagged observations. The computation of autocorrelation function (ACF) and PACF (PACF) (Li et al., 2021c) forms the foundations of constructing time-series forecasting approach. The ACF calculates the degree of interaction between current observation and historic lagged observations. PACF helps to determine with its preceding values the correctness degree of current variables while retaining certain constant values. The temporal dependencies as well as the seasonal patterns can be effectively extracted from the dataset. The inputs of the forecasting algorithm can be determined via the computed ACFs and PACFs.

Besides, the temporal signal decomposition gains more popularity in recent publications in short-term forecasting tasks. The main idea is to use temporal signal decomposition algorithms to “break down” a complicated time-series data into simpler and similar series rapidly and efficiently. Then, individual forecasting models can be constructed on each sub-series with improved forecasting accuracy. The Empirical Mode Decomposition (EMD) (Rilling et al., 2003), the Hilbert Vibration Decomposition (HVD) (Sharma and Sharma 2015), and the Variational Mode Decomposition (VMD) (Dragomiretskiy and Zosso 2013) are deemed to be the most representative signal decomposition algorithms used in short-term time-series forecasting tasks. In this research, the VMD is selected for the comparative analysis and it decomposed the complicated nitrate concentration time-series into a finite number of simplified intrinsic mode functions (IMFs). Then the forecasting algorithms are developed based in these IMFs to forecast the short-term values. In the final step, the forecasting results are aggregated to provide the final prediction outcome.

For the benchmarking deep-learning algorithms, the deep neural network (DNN) (Ryu et al., 2017), deep belief network (DBN) (Li et al., 2020), and long short-term memory recurrent neural network (LSTM-RNN) (Kong et al., 2017) are selected to perform the short-term nitrate concentration forecasting tasks with the three modeling approaches as the comparative analysis against the proposed IFPA-ESN. The DNN uses a cascade of multiple layers with nonlinear processing units. It is

a fully connected and feedforward network that allows computational models to be composed of multiple processing layers to learn the deep temporal representations in the time-series data. The DBN is consisted of multiple unsupervised restricted Boltzmann machines (RBMs) and a supervised regression layer stacked on the top. Each RBM has a layer of input neurons and a single hidden layer with hidden-to-all-visible connections. Layer-wise training is implemented across multiple RBMs to ensure promising forecasting accuracy. The LSTM-RNN servers as the most widely used state-of-art time-series forecasting algorithms. It is a special type of RNN that is capable to learn long-term and short-term temporal dependencies. Besides, its LSTM blocks contains three gates namely forget gate, input gate, and output gate. These gates make the LSTM blocks smarter than classical neurons and enable them to memorize recent sequences. All selected benchmarking algorithms have demonstrated superior performance in short-term forecasting tasks (He et al., 2017b).

Measurement Metrics

Four measurement metrics are selected to evaluate the performance of the chaotic modeling approach using the IFPA-ESN algorithm including mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), and hit rate (HR). The definition of the matrices as well as the formulas to compute them are illustrated in **Table 2** respectively.

TURNING POINT ANALYSIS

Parametric Ricker Function

Besides the pollutant concentration forecasting, we also aim to quantify how the discharge regime in the agricultural watershed reflected catchment hydrology and the delivery pathway of nitrate pollutant. The transportation and supply limitations of the nitrogen compounds varies by different locations also (Gao and Meguid 2021). Hence, the turning point analysis is performed in this research to quantify the relation between predicted daily nitrate concentration and stream discharge. In addition, the turning point will be utilized to determine the supply and transport limitations for the case study locations respectively.

Here, the parametric Ricker function (Kirchner et al., 2004) is utilized to model the relationship between daily discharge and predicted nitrate concentration. The Ricker function can be expressed as :

$$f(x; a, b, c) = ax^b e^{cx} \quad (17)$$

where $a > 0$, $b > 0$, and $c < 0$ in most of the cases. It is commonly used to model the hump-shaped data that are skewed to the right such as the discharge-concentration relationship. Since there are measurement errors and the fitted values can vary by measuring time and location, we can further revise the form as **Eq. 18** to incorporate our study as follows:

$$Y_i = aX_i^b e^{cX_i} + \varepsilon_i \quad (18)$$

TABLE 2 | Measurement metrics.

Name	Definition	Equation
MAE	Mean absolute error	$\frac{\sum_{i=1}^N \hat{y}_i - y_i }{N}$
MAPE	Mean absolute percentage error	$\frac{\sum_{i=1}^N \frac{ \hat{y}_i - y_i }{y_i}}{N}$
RMSE	Root-mean square error	$\sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}}$
Hit Rate	Percentages of predicted values fall within the 5% intervals of the actual value	$\frac{1}{N} \sum_{i=1}^N I(i) = \begin{cases} 1 & \text{if } \left \frac{\hat{y}_i - y_i}{y_i} \right \leq 0.05 \\ 0 & \text{if } \left \frac{\hat{y}_i - y_i}{y_i} \right > 0.05 \end{cases}$

where X_i is the discharge at time i ; Y_i is the predicted nitrate concentration at time i ; a , b , and c are unknown model parameters; and ε_i denotes the random error at measurement which has mean zero.

Delta Method

In many practical applications, fitting the mean structure of the parametric curve is insufficient. First, in the long-term, the mean structure may vary which brings challenge using only one single parametric function. Second, a slight shift of data points caused by measurement error or extremely high/low discharge would significantly impact the fitted mean structure of the fitted curve. Last, the random errors during the data collection process is inevitable which may complicate the statistical analysis (Toyoda and Wu, 2019; Wu et al., 2020). Hence, in this research, the quantile regression is adopted and the standard errors are also estimated for the parameters as well as the turning points.

The delta method (Jones et al., 2017; Le et al., 2021) is selected to derive the standard errors for the estimated parameters and turning points. First, the natural log is taken on both sides of the Ricker function and it can be expressed in:

$$\ln y = \ln a + b \ln x + cx \tag{19}$$

Then, the actual discharge data x and predicted nitrate concentration y both can be input into Eq. 19 to derive the estimates of a , b , and c . Then, the turning point is defined to be the x -value for which $\frac{\partial f(x;\theta)}{\partial \theta} = 0$ where $\theta = (a, b, c)^T$. By trivial derivation, the turning point for the Ricker function is $-b/c$. Next, we suppose the estimated parameters $\hat{\theta} \sim N(\theta_0, \Omega)$ and the turning point has the expression as $h(\theta)$. By delta's method, we can derive that the turning points has the standard error following Eq. 20 as:

$$h\left(\hat{\theta}_n\right) \sim N\left(h\left(\theta_0\right), \Delta\right) \tag{20}$$

where, Δ can be computed following Eq. 21 as follows:

$$\Delta = \frac{\partial h(\theta)}{\partial \theta^T} \Big|_{\hat{\theta}_n} \left(\frac{\Omega}{n} \right) \frac{\partial h(\theta)}{\partial \theta} \Big|_{\hat{\theta}_n} \tag{21}$$

In this study, we can derive that for turning point $h(\theta) = -b/c$ and $\frac{\partial h(\theta)}{\partial \theta} = \left(0, -\frac{1}{c}, \frac{b}{c^2}\right)^T$ respectively.

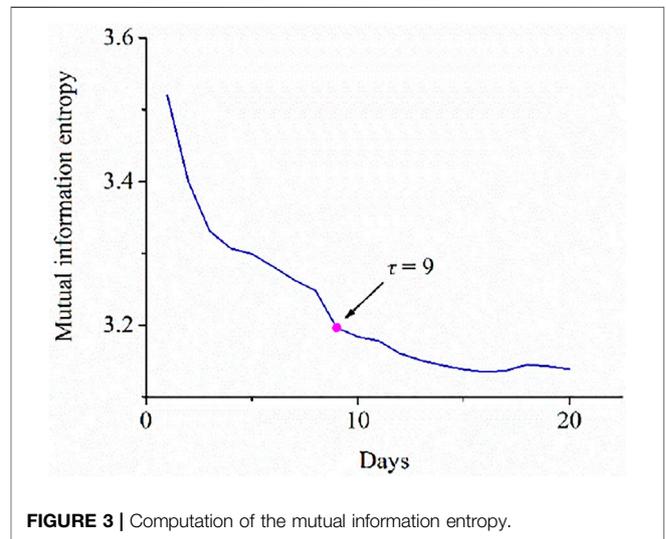


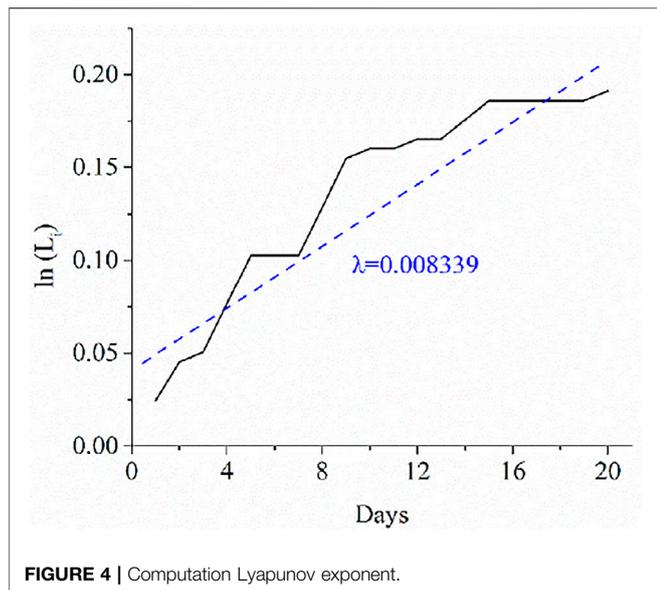
FIGURE 3 | Computation of the mutual information entropy.

EXPERIMENTAL RESULTS AND DISCUSSION

Data Reconstruction and Modeling

Considering the data homogeneity of nitrate concentration in streams, the daily measured concentrations in 2020 at the eight monitoring points were used for training, validation, and testing. As described in section 3, data reconstruction is helpful for modeling. The mutual information entropy and FNN were applied simultaneously to obtain the optimal parameter settings for data reconstruction. Finally, the Lyapunov exponents, as described in Section 3.4, were utilized to examine whether the reconstructed data using the parameter setting represent a chaotic system.

First, the mutual information entropy values were computed for each monitoring point; for example, one of the monitoring points, namely “Yibin-1,” has been illustrated in Figure 3 with respect to the computation of mutual information entropy. The curve depicts the mutual information entropy with respect to time delay in days. It shows an exponentially decaying pattern, which indicates that the relevance of the two series weakens with an increase in the delay times. For convenience, the optimal time



delay is usually selected when the value of the mutual information reaches its first local minimum value. Hence, in the example shown in **Figure 3**, the delay time of the nitrate concentration is set as 9. The time delay τ values for all monitoring points are computed and summarized in **Table 3**.

Meanwhile, the minimum embedded dimensions m for the monitoring points were computed using the FNN method while taking the computed time delay τ values as a prerequisite. Here, we set the value of the threshold to $a_{th} = 0.15$. When the percentage of false nearest neighbors is close to 0%, it implies that the structure of the system is completely unfolded in the phase space with dimension m . The computed minimum dimension m values for all the eight monitoring points are listed in **Table 3**.

In addition, the Lyapunov exponents were computed to determine whether the reconstructed time-series system is chaotic. As shown in **Figure 4**, the reconstructed nitrate concentration data in “Yibing-1” location was measured for the Lyapunov exponent λ . It was calculated using the slope coefficient between $\ln L$ and T . If λ is positive (i.e., $\lambda > 0$ in **Figure 4**), the system is chaotic, and the reconstructed time series is unfolded, as described in **Sections 3.3** and **3.4**. The computed λ values for all the eight monitoring points are summarized in **Table 3**.

Prediction Based on Chaotic Modeling

Furthermore, a comparative analysis was performed against the traditional time-series forecasting approach and the temporal

TABLE 4 | Summary of the global performance measures of four benchmark functions.

Sphere	Rastrigin		Ackley		Schaffer		
MIN	230	MIN	560	MIN	423	MIN	656
MAX	2000	MAX	1980	MAX	1,675	MAX	1778
Average	1,251	Average	1,423	Average	1,137	Average	1,052
CR	68.8%	CR	76.2%	CR	72.5%	CR	83.3%

signal decomposition forecasting approach. In each approach, the proposed IFPA-ESN algorithm was compared with the other benchmarking deep learning algorithms listed in **Section 3.7** above.

First, the IFPA optimization algorithm was used to optimize the hyperparameters of the ESN network. The reserve layer nodes, reserved layer connection rate, learning rate, and spectral radius were optimized simultaneously. Four benchmark functions, namely sphere, Rastrigin, Ackley, and Schaffer, were selected and compared. The benchmarking functions are all “landscape generator” class, and the differences among the functions relate to the number of minima within the search space, general convergence rate to the global optimal, and additional harmonic potential outside the search space to force the solution to stay within the boundaries. For each function, we performed a maximum of 2000 iterations for each experiment, and a total of 20 experiments were conducted. The performances of the four functions are summarized in **Table 4**.

As shown in **Table 4**, MIN, MAX, and Average denote the minimum, maximum, and average MAPEs, respectively, per iteration for the 20 experiments performed. CR refers to the convergence rate, which is the overall probability that the algorithm can converge to the default minimum value using benchmark functions. From **Table 4**, it can be inferred that the Schaffer function has the smallest mean iterations of convergence compared with the others. In addition, the convergence rate of the Schaffer function is the highest, which implies the largest probability of reaching the default minimum value of the loss function. Hence, the Schaffer function was selected as the default function for the IFPA algorithm. Using the Schaffer function, the mean MAPEs and the upper/lower bounds of the 95% confidence interval were computed in each iteration. The convergence measured using the MAPE for each monitoring point is illustrated in **Figure 5**.

Next, to demonstrate the performance of the proposed chaotic modeling approach using IFPA-ESN, two other modeling approaches, conventional time-series forecasting and temporal signal decomposition, were applied and compared. The time-

TABLE 3 | Summary of the data reconstruction.

Metric	Location								
	Yibin-1	Yibin-2	Luzhou-1	Luzhou-2	Luzhou-3	Jiangjin-1	Jiangjin-2	Jiangjin-3	
τ	6	7	7	4	8	5	7	5	
m	7	5	8	6	6	5	9	7	
λ	0.0083	0.0025	0.0061	0.0089	0.0059	0.0096	0.0061	0.0043	

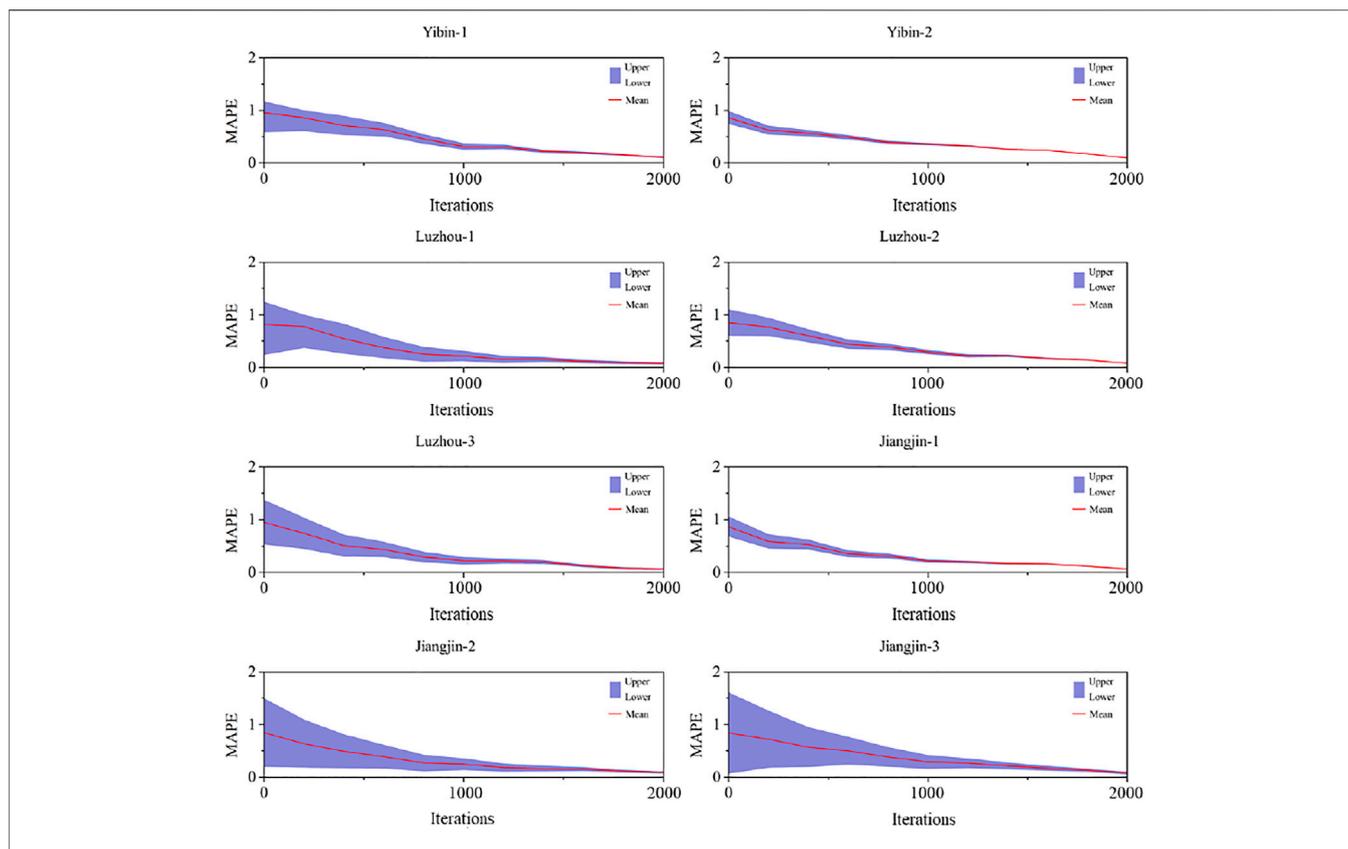


FIGURE 5 | Convergence of MAPE by iterations using IFPA algorithm.

TABLE 5 | Nitrate concentration prediction with conventional time-series modeling.

Location	Metric	DNN	DBN	LSTM	IFPA-ESN	Location	Metric	DNN	DBN	LSTM	IFPA-ESN
Yibin-1	MAE	2.86	2.29	2.26	1.84	Yibin-2	MAE	2.86	2.23	2.20	1.60
	MAPE	0.19	0.16	0.14	0.12		MAPE	0.20	0.16	0.16	0.11
	RMSE	3.99	3.31	3.09	2.29		RMSE	4.50	3.86	3.75	2.51
	HR	0.63	0.66	0.68	0.73		HR	0.67	0.72	0.71	0.77
Luzhou-1	MAE	3.49	2.51	2.42	1.93	Luzhou-2	MAE	2.56	1.84	1.75	1.41
	MAPE	0.17	0.12	0.11	0.09		MAPE	0.18	0.12	0.10	0.09
	RMSE	3.87	3.49	3.27	2.44		RMSE	2.72	2.44	2.30	1.72
	HR	0.71	0.76	0.77	0.80		HR	0.72	0.75	0.77	0.80
Luzhou-3	MAE	2.38	1.87	1.80	1.23	Jiangjin-1	MAE	2.33	1.82	1.73	1.54
	MAPE	0.18	0.12	0.11	0.08		MAPE	0.13	0.11	0.10	0.08
	RMSE	3.34	2.41	2.33	1.48		RMSE	3.56	2.69	2.61	2.25
	HR	0.69	0.77	0.76	0.82		HR	0.74	0.79	0.79	0.82
Jiangjin-2	MAE	2.65	2.19	2.12	1.31	Jiangjin-3	MAE	2.06	1.92	1.76	1.15
	MAPE	0.17	0.11	0.11	0.09		MAPE	0.16	0.11	0.11	0.09
	RMSE	3.10	2.25	2.17	1.52		RMSE	3.30	2.11	2.02	1.49
	HR	0.73	0.77	0.77	0.81		HR	0.72	0.77	0.77	0.79

series forecasting approach was based on computation of the ACF and PACF values for each monitoring point. The temporal dependencies and seasonality were considered in the modeling approach. In the signal decomposition approach, the time-series nitrate concentrations were decomposed into nine simplified intrinsic mode functions (IMFs). For each IMF, a short-term forecasting sub-model was constructed. The final forecasting

results were the aggregation of all forecasting outcomes produced by the nine sub-models.

Table 5 summarizes the forecasting performance over the testing dataset of all the deep learning algorithms tested, as described in Section 3.7. Using a conventional time-series forecasting approach, the proposed IFPA-ESN outperformed all the other benchmarking algorithms tested. The forecasting

TABLE 6 | Nitrate concentration prediction with decomposition approach.

Location	Matric	DNN	DBN	LSTM	IFPA-ESN	Location	Matric	DNN	DBN	LSTM	IFPA-ESN
Yibin-1	MAE	2.70	2.18	2.13	1.74	Yibin-2	MAE	2.73	2.12	2.09	1.51
	MAPE	0.18	0.15	0.14	0.12		MAPE	0.19	0.15	0.15	0.11
	RMSE	3.78	3.12	2.94	2.17		RMSE	4.33	3.65	3.54	2.37
	HR	0.68	0.72	0.73	0.79		HR	0.71	0.75	0.74	0.81
Luzhou-1	MAE	3.32	2.37	2.28	1.83	Luzhou-2	MAE	2.45	1.76	1.66	1.34
	MAPE	0.16	0.12	0.11	0.08		MAPE	0.17	0.12	0.10	0.09
	RMSE	3.66	3.29	3.09	2.30		RMSE	2.62	2.34	2.20	1.63
	HR	0.77	0.82	0.83	0.87		HR	0.76	0.79	0.81	0.85
Luzhou-3	MAE	2.25	1.77	1.71	1.16	Jiangjin-1	MAE	2.23	1.73	1.63	1.46
	MAPE	0.17	0.12	0.11	0.07		MAPE	0.13	0.11	0.10	0.07
	RMSE	3.15	2.29	2.21	1.40		RMSE	3.30	2.51	2.41	2.10
	HR	0.74	0.83	0.83	0.88		HR	0.78	0.84	0.83	0.86
Jiangjin-2	MAE	2.51	2.08	2.01	1.24	Jiangjin-3	MAE	1.96	1.84	1.67	1.10
	MAPE	0.16	0.11	0.11	0.08		MAPE	0.15	0.11	0.11	0.08
	RMSE	2.94	2.12	2.04	1.44		RMSE	3.15	2.00	1.95	1.42
	HR	0.79	0.83	0.84	0.88		HR	0.77	0.81	0.82	0.84

TABLE 7 | Nitrate concentration prediction with chaotic modeling approach.

Location	Matric	DNN	DBN	LSTM	IFPA-ESN	Location	Matric	DNN	DBN	LSTM	IFPA-ESN
Yibin-1	MAE	2.56	2.06	2.03	1.65	Yibin-2	MAE	2.58	2.00	1.98	1.43
	MAPE	0.17	0.14	0.13	0.11		MAPE	0.18	0.14	0.14	0.10
	RMSE	3.59	2.96	2.78	2.05		RMSE	3.92	3.33	3.23	2.16
	HR	0.71	0.75	0.77	0.83		HR	0.75	0.80	0.79	0.86
Luzhou-1	MAE	3.14	2.25	2.16	1.73	Luzhou-2	MAE	2.31	1.65	1.57	1.27
	MAPE	0.15	0.11	0.10	0.08		MAPE	0.16	0.11	0.09	0.06
	RMSE	3.48	3.12	2.93	2.18		RMSE	2.46	2.20	2.07	1.54
	HR	0.81	0.86	0.87	0.91		HR	0.80	0.84	0.86	0.90
Luzhou-3	MAE	2.14	1.68	1.62	1.10	Jiangjin-1	MAE	2.09	1.63	1.55	1.38
	MAPE	0.16	0.11	0.10	0.07		MAPE	0.12	0.10	0.09	0.07
	RMSE	2.99	2.17	2.09	1.33		RMSE	2.79	2.12	2.05	1.78
	HR	0.78	0.87	0.87	0.92		HR	0.83	0.88	0.88	0.91
Jiangjin-2	MAE	2.26	1.86	1.81	1.12	Jiangjin-3	MAE	1.85	1.74	1.58	1.03
	MAPE	0.15	0.10	0.10	0.09		MAPE	0.14	0.10	0.10	0.08
	RMSE	2.92	2.12	2.04	1.44		RMSE	2.98	1.90	1.82	1.34
	HR	0.83	0.87	0.88	0.92		HR	0.81	0.86	0.86	0.88

performances of all the algorithms tested using the signal decomposition approach are presented in **Table 6**. The proposed IFPA-ESN provided the smallest MAE, MAPE, and RMSE values. Meanwhile, it provided the highest values of HR, which demonstrates its superior power in short-term nitrate concentration forecasting.

The global performance of short-term nitrate concentrations obtained using the chaotic modeling approach is summarized in **Table 7**. IFPA-ESN still outperformed all the algorithms tested in this scenario. In addition, for all the tested algorithms, the chaotic modeling approach had fewer forecasting errors, which validated its superior power in discovering the real data structure. The forecasting outcomes of the testing dataset and the observed actual nitrate concentrations are shown in **Figure 6**.

Discussion of the Turning Points

The forecasted nitrate concentration in the testing dataset (100 days) integrated with the measured stream discharge during the same period was utilized for turning point analysis. Here, the

forecasted daily concentration versus discharge data at the same monitoring location is provided. An examination of the turning point analysis, as described in **Section 4**, was performed, and the results are summarized and illustrated in **Table 8** and **Figure 7**.

Table 8 presents results of the fitted Ricker function using the quantile regression. The turning points were also computed with the standard errors provided. As shown in **Table 8**, the turning points varied from 0.79 to 2.27 mm depending on the measured stream location, which is impacted by drought. The standard errors were computed using the Delta's method, and the values varied between 0.05 and 0.14 mm. Considering the estimated turning points, the standard errors were within a small range, which indicated the reliability of the estimation.

The fitted Ricker function is plotted in **Figure 7**. It depicts the nonlinear relationship between the forecasted nitrate concentrations and the stream discharges. The turning points are labeled in red in **Figure 7**, and the fitted Ricker function is plotted as a blue curve. At each monitoring point, the nitrate concentration started to decline as the discharge

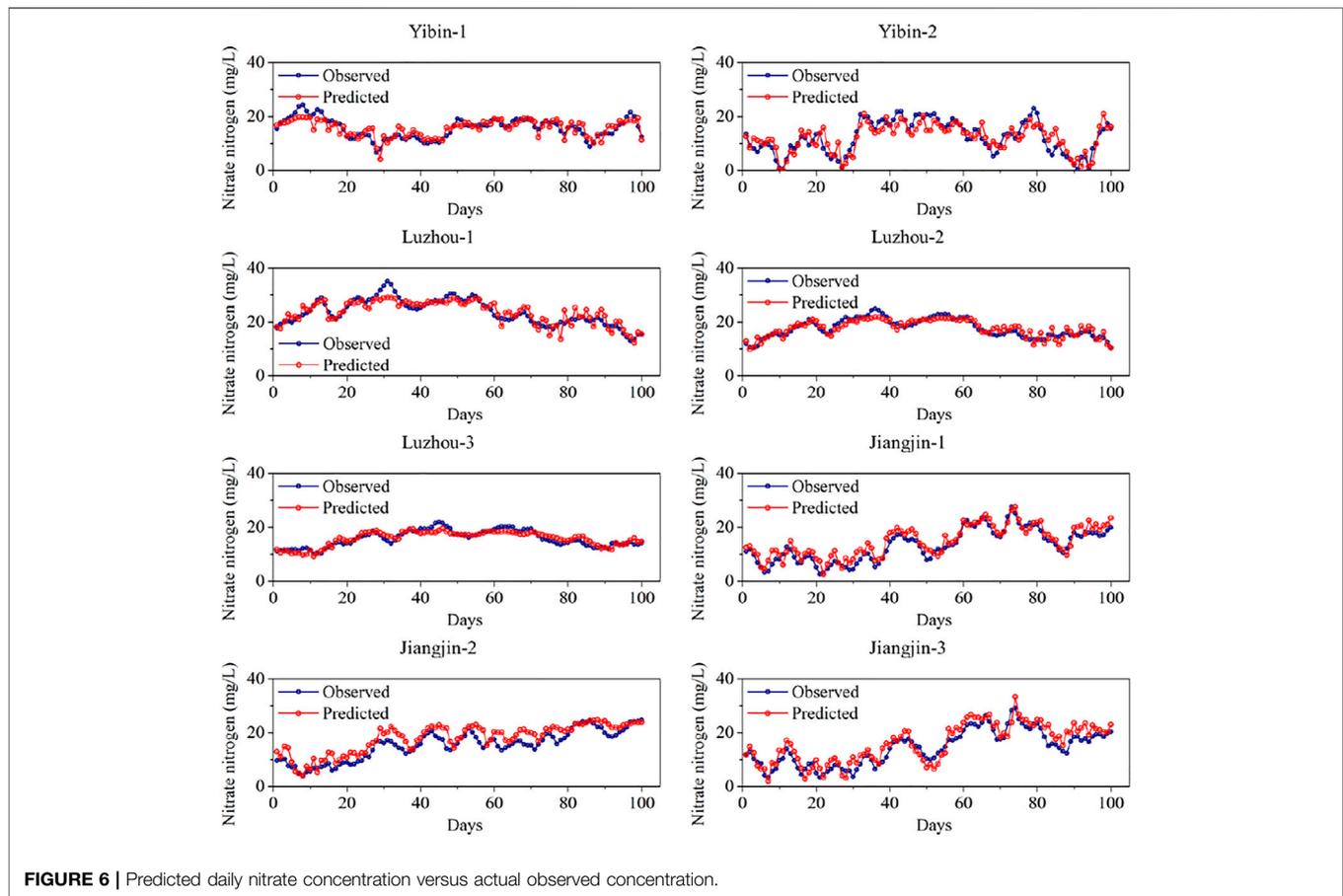


FIGURE 6 | Predicted daily nitrate concentration versus actual observed concentration.

TABLE 8 | Fitting of turning point.

Location	a	se(a)	b	se(b)	c	se(c)	TP (mm)	se (TP, mm)
Yibin-1	25.57	3.06	0.58	0.05	-0.27	0.06	2.18	0.14
Yibin-2	24.69	1.89	0.71	0.03	-0.34	0.06	2.06	0.12
Luzhou-1	36.72	0.69	0.49	0.05	-0.32	0.07	1.52	0.12
Luzhou-2	43.37	4.06	0.30	0.05	-0.20	0.07	1.49	0.08
Luzhou-3	25.53	2.67	0.17	0.01	-0.21	0.05	0.79	0.13
Jiangjin-1	37.13	1.11	0.62	0.05	-0.33	0.08	1.86	0.05
Jiangjin-2	27.14	4.92	0.29	0.04	-0.21	0.07	1.41	0.13
Jiangjin-3	31.36	1.45	0.67	0.04	-0.29	0.08	2.27	0.06

exceeded the turning point. This supports the idea that drought impacts stream nitrate concentrations in the temporal domain.

Table 9 summarizes the delineation between transport- and supply-limited days. Based on the percentages of loads before the turning point, the locations such as Yibin-1, Yibin-2, Luzhou-1, Luzhou-2, Luzhou-3, and Jiangjin-1 are transport limited. Only two locations, namely Jiangjin-2 and Jiangjin-3, are supply limited. The proportion of loads before the turning points varied between 53 and 81% under the transport-limited condition. It is noteworthy that the proportion of loads before the turning points was between 41 and 42% under supply-limited conditions. These conditions indicate that supply

considerations are likely to control the flux of $\text{NO}_x\text{-N}$ in stream networks.

CONCLUSION

In this study, insights are provided for experts, academicians, and officers of local environment protection agencies via chaotic modeling of stream nitrate concentrations using IFPA-ESN. This study also provides a turning point analysis to quantify nitrate concentration-discharge relations. In addition, the limitations of pollutant transport and supply patterns were estimated in all case studies.

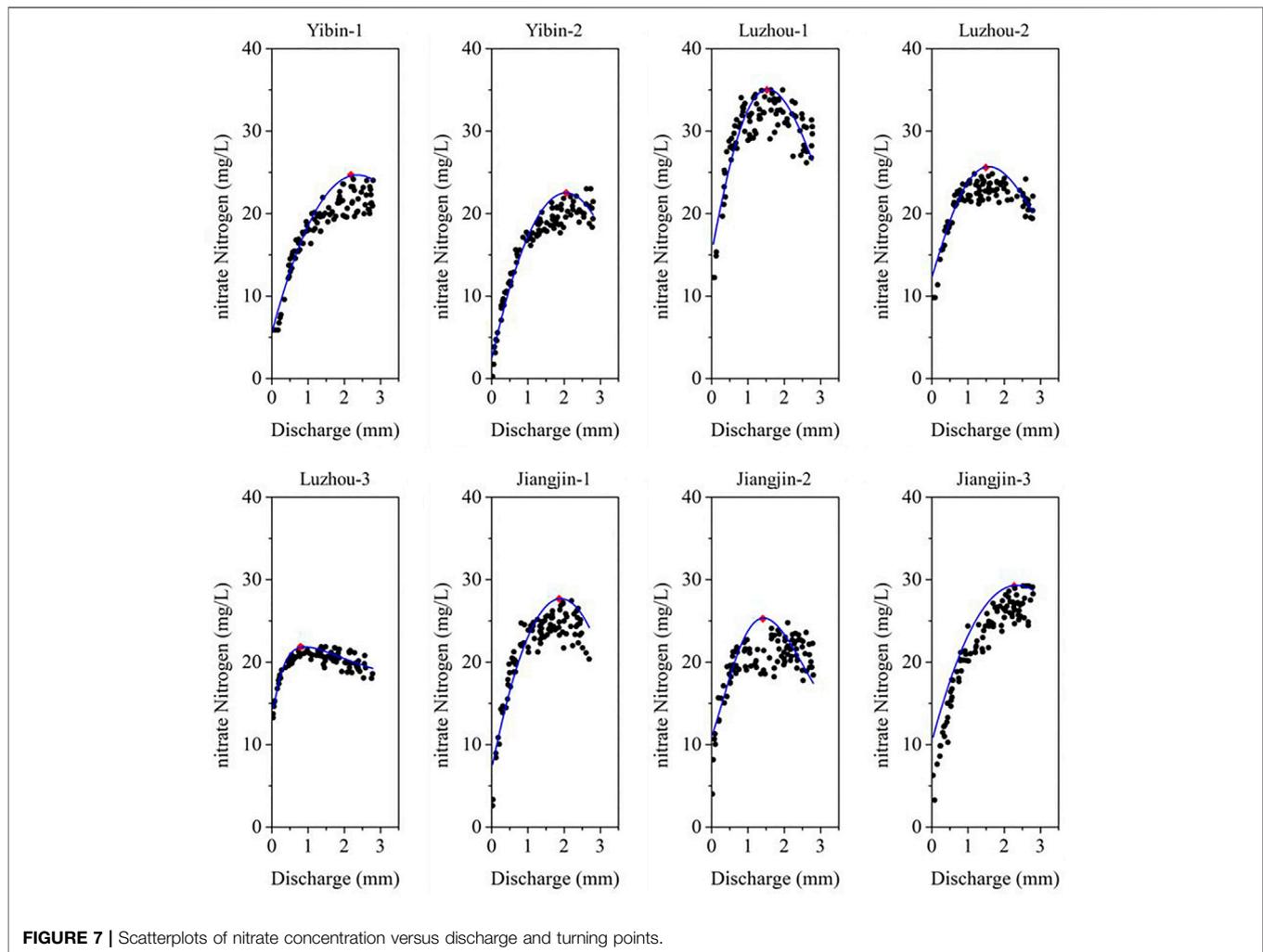


FIGURE 7 | Scatterplots of nitrate concentration versus discharge and turning points.

TABLE 9 | Limitation analysis of transport and supply.

Location	Load before turning point (%)	Load after turning point (%)	Limitation
Yibin-1	53	47	Transport
Yibin-2	81	19	Transport
Luzhou-1	60	40	Transport
Luzhou-2	69	31	Transport
Luzhou-3	70	30	Transport
Jiangjin-1	64	36	Transport
Jiangjin-2	41	59	Supply
Jiangjin-3	42	58	Supply

The computational results revealed that IFPA-ESN integrated with a chaotic modeling strategy achieved a satisfactory level of accuracy for forecasting the daily nitrate concentration in stream networks using historical time-series data. Compared with other time-series modeling scenarios, the chaotic modeling strategy demonstrated superior performance in terms of forecasting accuracy. Meanwhile, the hyperparameters of the ESN were carefully tuned or searched to reach the optimal setting using the IFPA algorithm. The

IFPA-ESN provided peak forecasting performances comparable to those of other benchmarking machine-learning and deep learning algorithms in our eight case study locations. On the other hand, the turning point analysis using parametric Ricker functions accurately fitted the trend of nonlinear relations between nitrate concentration and discharge in eight case study locations. Furthermore, the delineation between transport and supply limitations was discussed to gain insights into nitrate pollution in stream systems.

Future recommendations include incorporating various climate-related external factors into the decision-making process in the light of further reduction in forecasting volatility; these external factors include precipitation, ambient temperature, wind speed, air humidity, and soil-water relationships. Moreover, future studies should aim to develop federated learning systems to incorporate the deep learning-based chaotic modeling strategy to forecast nitrate pollution simultaneously for multiple locations while protecting data privacy.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

REFERENCES

- Ahmad, T., Zhang, D., and Huang, C. (2021). Methodological Framework for Short-And Medium-Term Energy, Solar and Wind Power Forecasting with Stochastic-Based Machine Learning Approach to Monetary and Energy Policy Applications. *Energy* 231, 120911. doi:10.1016/j.energy.2021.120911
- Bazine, H., and Mabrouki, M. (2019). Chaotic Dynamics Applied in Time Prediction of Photovoltaic Production. *Renew. Energ.* 136, 1255–1265. doi:10.1016/j.renene.2018.09.098
- Belitz, K., Fram, M. S., and Johnson, T. D. (2015). Metrics for Assessing the Quality of Groundwater Used for Public Supply, CA, USA: Equivalent-Population and Area. *Environ. Sci. Technol.* 49 (14), 8330–8338. doi:10.1021/acs.est.5b00265
- Blesh, J., and Drinkwater, L. E. (2013). The Impact of Nitrogen Source and Crop Rotation on Nitrogen Mass Balances in the Mississippi River Basin. *Ecol. Appl.* 23 (5), 1017–1035. doi:10.1890/12-0132.1
- Chouikhi, N., Ammar, B., Rokbani, N., and Alimi, A. M. (2017). PSO-based Analysis of Echo State Network Parameters for Time Series Forecasting. *Appl. Soft Comput.* 55, 211–225. doi:10.1016/j.asoc.2017.01.049
- Cui, S., Pei, X., Jiang, Y., Wang, G., Fan, X., Yang, Q., et al. (2021). Liquefaction within a Bedding Fault: Understanding the Initiation and Movement of the Daguangbao Landslide Triggered by the 2008 Wenchuan Earthquake (Ms = 8.0). *Eng. Geology.* 295, 106455. doi:10.1016/j.enggeo.2021.106455
- David, M. B., Gentry, L. E., Kovacic, D. A., and Smith, K. M. (1997). Nitrogen Balance in and export from an Agricultural Watershed. *Am. Soc. Agron. Crop Sci. Soc. America, Soil Sci. Soc. America* 26 (No. 4), 1038–1048. doi:10.2134/jeq1997.00472425002600040015x
- Domangue, R. J., and Mortazavi, B. (2018). Nitrate Reduction Pathways in the Presence of Excess Nitrogen in a Shallow Eutrophic Estuary. *Environ. Pollut.* 238, 599–606. doi:10.1016/j.envpol.2018.03.033
- Dragomiretskiy, K., and Zosso, D. (2013). Variational Mode Decomposition. *IEEE Transactions Signal. Processing* 62 (3), 531–544.
- Gao, G., and Meguid, M. A. (2021). On the Role of Joint Roughness on the Micromechanics of Rock Fracturing Process: a Numerical Study. *Acta Geotechnica*, 1–26. doi:10.1007/s11440-021-01401-8
- Gao, G., Meguid, M. A., Chouinard, L. E., and Xu, C. (2020). Insights into the Transport and Fragmentation Characteristics of Earthquake-Induced Rock Avalanche: Numerical Study. *Int. J. Geomech.* 20 (9), 04020157. doi:10.1061/(asce)gm.1943-5622.0001800
- Gao, G., Meguid, M. A., Chouinard, L. E., and Zhan, W. (2021). Dynamic Disintegration Processes Accompanying Transport of an Earthquake-Induced Landslide. *Landslides* 18 (3), 909–933. doi:10.1007/s10346-020-01508-1
- He, Y., Deng, J., and Li, H. (2017b). “Short-term Power Load Forecasting with Deep Belief Network and Copula Models,” in 9th International conference on intelligent human-machine systems and cybernetics (IHMSC), IEEE, 1, 191–194. doi:10.1109/ihmsc.2017.50

AUTHOR CONTRIBUTIONS

YT conceptualized the study, contributed to the study methodology, and wrote the original draft. JD contributed to the study methodology, data curation and investigation. CZ contributed to software and formal analysis. QW contributed to editing and data curation. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by the Regional Innovation Cooperation Programs of Sichuan province (2021YFQ0050).

- He, Y., Kusiak, A., Ouyang, T., and Teng, W. (2017a). Data-driven Modeling of Truck Engine Exhaust Valve Failures: a Case Study. *J. Mech. Sci. Technol.* 31 (6), 2747–2757. doi:10.1007/s12206-017-0518-1
- He, Y., and Kusiak, A. (2018). Performance Assessment of Wind Turbines: Data-Derived Quantitative Metrics. *IEEE Trans. Sustain. Energ.* 9 (1), 65–73. doi:10.1109/tste.2017.2715061
- Hrnjica, B., Mehr, A. D., Jakupović, E., Crnkčić, A., and Hasanagić, R. (2021). “Application of Deep Learning Neural Networks for Nitrate Prediction in the Klokot River, Bosnia and Herzegovina,” in 2021 7th International Conference on Control, Instrumentation and Automation (ICCIA) (IEEE), 1–6. February.
- Hu, Q., Zhang, R., and Zhou, Y. (2016). Transfer Learning for Short-Term Wind Speed Prediction with Deep Neural Networks. *Renew. Energ.* 85, 83–95. doi:10.1016/j.renene.2015.06.034
- Huang, G.-B., Wang, D. H., and Lan, Y. (2011). Extreme Learning Machines: a Survey. *Int. J. Mach. Learn. Cyber.* 2 (2), 107–122. doi:10.1007/s13042-011-0019-y
- Jiang, J., Tang, S., Liu, R., Sivakumar, B., Wu, X., and Pang, T. (2021). A Hybrid Wavelet-Lyapunov Exponent Model for River Water Quality Forecast. *J. Hydroinformatics* 23 (4), 864–878. doi:10.2166/hydro.2021.023
- Jones, C. S., Wang, B., Schilling, K. E., and Chan, K.-s. (2017). Nitrate Transport and Supply Limitations Quantified Using High-Frequency Stream Monitoring and Turning point Analysis. *J. Hydrol.* 549, 581–591. doi:10.1016/j.jhydrol.2017.04.041
- Keupers, I., and Willems, P. (2017). Development and Testing of a Fast Conceptual River Water Quality Model. *Water Res.* 113, 62–71. doi:10.1016/j.watres.2017.01.054
- Kirchner, J. W., Feng, X., Neal, C., and Robson, A. J. (2004). The fine Structure of Water-Quality Dynamics: The (high-Frequency) Wave of the Future. *Hydrol. Process.* 18 (7), 1353–1359. doi:10.1002/hyp.5537
- Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., and Zhang, Y. (2017). Short-term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Trans. Smart Grid* 10 (1), 841–851.
- Le, S., Wu, Y., Guo, Y., and Del Vecchio, C. (2021). “Game Theoretic Approach for a Service Function Chain Routing in NFV with Coupled Constraints,” in *IEEE Transactions on Circuits and Systems II: Express Briefs*. doi:10.1109/tcsii.2021.3070025
- LeCun, Y., and Bengio, Y. (1995). Convolutional Networks for Images, Speech, and Time Series. *The handbook Brain Theor. Neural networks* 3361 (10), 1995.
- Lei, X., Fang, M., Wu, F. X., and Chen, L. (2018). Improved Flower Pollination Algorithm for Identifying Essential Proteins. *BMC Syst. Biol.* 12 (4), 46–140. doi:10.1186/s12918-018-0573-y
- Li, H., Deng, J., Feng, P., Pu, C., Arachchige, D. D. K., and Cheng, Q. (2021a). Short-Term Nacelle Orientation Forecasting Using Bilinear Transformation and ICEEMDAN Framework. *Front. Energ. Res.* 9, 780928. doi:10.3389/ferg.2021.780928
- Li, H., Deng, J., Yuan, S., Feng, P., and Arachchige, D. D. K. (2021b). Monitoring and Identifying Wind Turbine Generator Bearing Faults Using Deep Belief Network and EWMA Control Charts. *Front. Energ. Res.* 9, 799039. doi:10.3389/ferg.2021.799039

- Li, H., He, Y., Xu, Q., Deng, J., Li, W., and Wei, Y. (2022). Detection and Segmentation of Loess Landslides via Satellite Images: a Two-phase Framework. *Landslides*, 1–14. doi:10.1007/s10346-021-01789-0
- Li, H., Xu, Q., He, Y., and Deng, J. (2018). Prediction of Landslide Displacement with an Ensemble-Based Extreme Learning Machine and Copula Models. *Landslides* 15 (10), 2047–2059. doi:10.1007/s10346-018-1020-2
- Li, H., Xu, Q., He, Y., Fan, X., and Li, S. (2020). Modeling and Predicting Reservoir Landslide Displacement with Deep Belief Network and EWMA Control Charts: a Case Study in Three Gorges Reservoir. *Landslides* 17 (3), 693–707. doi:10.1007/s10346-019-01312-6
- Li, H., Xu, Q., He, Y., Fan, X., Yang, H., and Li, S. (2021c). Temporal Detection of Sharp Landslide Deformation with Ensemble-Based LSTM-RNNs and Hurst Exponent. *Geomatics, Nat. Hazards Risk* 12 (1), 3089–3113. doi:10.1080/19475705.2021.1994474
- Nolan, B. T., Gronberg, J. M., Faunt, C. C., Eberts, S. M., and Belitz, K. (2014). Modeling Nitrate at Domestic and Public-Supply Well Depths in the Central Valley, California. *Environ. Sci. Technol.* 48 (10), 5643–5651. doi:10.1021/es405452q
- Ostad-Ali-Askari, K., Shayannejad, M., and Ghorbanizadeh-Kharazi, H. (2017). Artificial Neural Network for Modeling Nitrate Pollution of Groundwater in Marginal Area of Zayandeh-Rood River, Isfahan, Iran. *KSCE J. Civ Eng.* 21 (1), 134–140. doi:10.1007/s12205-016-0572-8
- Ouyang, T., Cha, X., and Qin, L. (2016). Medium-or Long-Term Wind Power Prediction with Combined Models of Meteorological Multi-Variables. *Power Syst. Tech.* 40 (03), 847–852.
- Ouyang, T., He, Y., and Huang, H. (2018). Monitoring Wind Turbines' Unhealthy Status: a Data-Driven Approach. *IEEE Trans. Emerging Top. Comput. Intelligence* 3 (2), 163–172.
- Ouyang, T., He, Y., Li, H., Sun, Z., and Baek, S. (2019a). Modeling and Forecasting Short-Term Power Load with Copula Model and Deep Belief Network. *IEEE Trans. Emerg. Top. Comput. Intell.* 3 (2), 127–136. doi:10.1109/tetci.2018.2880511
- Ouyang, T., Huang, H., He, Y., and Tang, Z. (2020). Chaotic Wind Power Time Series Prediction via Switching Data-Driven Modes. *Renew. Energ.* 145, 270–281. doi:10.1016/j.renene.2019.06.047
- Ouyang, T., Zha, X., Qin, L., He, Y., and Tang, Z. (2019b). Prediction of Wind Power Ramp Events Based on Residual Correction. *Renew. Energ.* 136, 781–792. doi:10.1016/j.renene.2019.01.049
- Ransom, K. M., Nolan, B. T., A. Traum, J., Faunt, C. C., Bell, A. M., Gronberg, J. A. M., Wheeler, D. C., Z. Rosecrans, C., Jurgens, B., Schwarz, G. E., Belitz, K., M. Eberts, S., Kourakos, G., and Harter, T. (2017). A Hybrid Machine Learning Model to Predict and Visualize Nitrate Concentration throughout the Central Valley Aquifer, California, USA. *Sci. Total Environ.* 601–602, 1160–1172. doi:10.1016/j.scitotenv.2017.05.192
- Rilling, G., Flandrin, P., and Goncalves, P. (2003). "June). On Empirical Mode Decomposition and its Algorithms," in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, 3, 8–11. NSIP-03, Grado (1).3
- Rosenstein, M. T., Collins, J. J., and De Luca, C. J. (1994). Reconstruction Expansion as a Geometry-Based Framework for Choosing Proper Delay Times. *Physica D: Nonlinear Phenomena* 73 (1-2), 82–98. doi:10.1016/0167-2789(94)90226-7
- Ryu, S., Noh, J., and Kim, H. (2017). Deep Neural Network Based Demand Side Short Term Load Forecasting. *Energies* 10 (1), 3.
- Sajedi-Hosseini, F., Malekian, A., Choubin, B., Rahmati, O., Cipullo, S., Coulon, F., et al. (2018). A Novel Machine Learning-Based Approach for the Risk Assessment of Nitrate Groundwater Contamination. *Sci. total Environ.* 644, 954–962. doi:10.1016/j.scitotenv.2018.07.054
- Sharma, H., and Sharma, K. K. (2015). Baseline Wander Removal of ECG Signals Using Hilbert Vibration Decomposition. *Electron. Lett.* 51 (6), 447–449. doi:10.1049/el.2014.4076
- Shi, P., Zhang, Y., Song, J., Li, P., Wang, Y., Zhang, X., Li, Z., Bi, Z., Zhang, X., Qin, Y., and Zhu, T. (2019). Response of Nitrogen Pollution in Surface Water to Land Use and Social-Economic Factors in the Weihe River Watershed, Northwest China. *Sustain. Cities Soc.* 50, 101658. doi:10.1016/j.scs.2019.101658
- Stamenković, L. J., Mrazovac Kurilić, S., and Presburger Ulniković, V. (2020). Prediction of Nitrate Concentration in Danube River Water by Using Artificial Neural Networks. *Water Supply* 20 (6), 2119–2132.
- Tang, Z., and Zhang, Z. (2019). The Multi-Objective Optimization of Combustion System Operations Based on Deep Data-Driven Models. *Energy* 182, 37–47. doi:10.1016/j.energy.2019.06.051
- Tang, Z., Zhao, G., and Ouyang, T. (2021). Two-phase Deep Learning Model for Short-Term Wind Direction Forecasting. *Renew. Energ.* 173, 1005–1016. doi:10.1016/j.renene.2021.04.041
- Toyoda, M., and Wu, Y. (2019). Mayer-type Optimal Control of Probabilistic Boolean Control Network with Uncertain Selection Probabilities. *IEEE Trans. cybernetics* 51 (6), 3079–3092. doi:10.1109/tcyb.2019.2954849
- Villarini, G., Jones, C. S., and Schilling, K. E. (2016). Soybean Area and Baseflow Driving Nitrate in Iowa's Raccoon River. *J. Environ. Qual.* 45 (6), 1949–1959. doi:10.2134/jeq2016.05.0180
- Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). Determining Lyapunov Exponents from a Time Series. *Physica D: nonlinear phenomena* 16 (3), 285–317. doi:10.1016/0167-2789(85)90011-9
- Wu, Y., Guo, Y., and Toyoda, M. (2020). "Policy Iteration Approach to the Infinite Horizon Average Optimal Control of Probabilistic Boolean Networks," in *IEEE Transactions on Neural Networks and Learning Systems*.
- Xu, Q., Li, H., He, Y., Liu, F., and Peng, D. (2019). Comparison of Data-Driven Models of Loess Landslide Runout Distance Estimation. *Bull. Eng. Geol. Environ.* 78 (2), 1281–1294. doi:10.1007/s10064-017-1176-3
- Zhang, Y., Shi, P., Li, F., Wei, A., Song, J., and Ma, J. (2018). Quantification of Nitrate Sources and Fates in Rivers in an Irrigated Agricultural Area Using Environmental Isotopes and a Bayesian Isotope Mixing Model. *Chemosphere* 208, 493–501. doi:10.1016/j.chemosphere.2018.05.164
- Zhou, J., Wei, J., Yang, T., Zhang, P., Liu, F., and Chen, J. (2021). Seepage Channel Development in the crown Pillar: Insights from Induced Microseismicity. *Int. J. Rock Mech. Mining Sci.* 145, 104851. doi:10.1016/j.ijrmms.2021.104851

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 Tang, Deng, Zang and Wu. 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.