



Article

# Application and Comparison of Different Models for Quantifying the Aquatic Community in a Dam-Controlled River

Jing Liu <sup>1,2</sup>, Chao Zang <sup>3,\*</sup>, Qiting Zuo <sup>4</sup> , Chunhui Han <sup>1</sup> and Stefan Krause <sup>2,5</sup>

<sup>1</sup> College of Water Resources, North China University of Water Resources and Electric Power, Zhengzhou 450001, China

<sup>2</sup> School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham B15 2TT, UK

<sup>3</sup> Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba 305-8571, Japan

<sup>4</sup> School of Water Conservancy Engineering, Zhengzhou University, Zhengzhou 450001, China

<sup>5</sup> LEHNA—Laboratoire d'Ecologie des Hydrosystemes Naturels et Anthropises, University of Lyon, 69622 Villeurbanne, France

\* Correspondence: zangchaodl@163.com

**Abstract:** In order to develop a better model for quantifying aquatic community using environmental factors that are easy to get, we construct quantitative aquatic community models that utilize the different relationships between water environmental impact factors and aquatic biodiversity as follows: a multi-factor linear-based (MLE) model and a black box-based 'Genetic algorithm-BP artificial neural networks' (GA-BP) model. A comparison of the model efficiency and their outputs is conducted by applying the models to real-life cases, referring to the 49 groups of seasonal data observed over seven field sampling campaigns in Shaying River, China, and then performing model to reproduce the seasonal and inter-annual variation of the water ecological characteristics in the Huaidian (HD) site over 10 years. The results show that (1) the MLE and GA-BP models constructed in this paper are effective in quantifying aquatic communities in dam-controlled rivers; and (2) the performance of GA-BP models based on black-box relationships in predicting the aquatic community is better, more stable, and reliable; (3) reproducing the seasonal and inter-annual aquatic biodiversity in the HD site of Shaying River shows that the seasonal variation of species diversity for phytoplankton, zooplankton, and zoobenthos are inconsistent, and the inter-annual levels of diversity are low due to the negative impact of dam control. Our models can be used as a tool for aquatic community prediction and can become a contribution to showing how quantitative models in other dam-controlled rivers to assisting in dam management strategies.

**Keywords:** aquatic community simulation; water ecosystem; GA-BP; modeling comparison; dam-controlled river



**Citation:** Liu, J.; Zang, C.; Zuo, Q.; Han, C.; Krause, S. Application and Comparison of Different Models for Quantifying the Aquatic Community in a Dam-Controlled River. *Int. J. Environ. Res. Public Health* **2023**, *20*, 4148. <https://doi.org/10.3390/ijerph20054148>

Academic Editors: Ying Zhu, Wei Chen, Xinli Xing and Peiyue Li

Received: 16 November 2022

Revised: 17 February 2023

Accepted: 22 February 2023

Published: 25 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Rivers represent some of the world's most biodiverse ecosystems [1] and provide critical ecosystem services to society and the environment [2]. Many rivers globally are dammed, and their flow is regulated for energy, water storage, irrigation, and protection against floods and are fragmented by artificial barriers to block free flow [3,4]. The number of river dams in Europe exceeds one million [5], in the United States at least 90,000 are over six-feet tall, and in China reaches 100,000 with a flow of more than 5 m<sup>3</sup>/s. While dams are generally designed to control with the purpose of stabilizing low flows and reducing peak flows for ensuring water security [6], human survival, and economic development, they also have significant ecological impacts, e.g., by impeding the flow of essential nutrients [7,8], altering thermal regimes [9,10], modifying sediment transport [11], disrupting growth cycles of aquatic organisms [12], or affecting ecosystem structure [13,14] and functions [15] along river networks, all of which cause the reduction of biodiversity of aquatic

communities [16–18], and the more the deviation from the natural regime, the greater the loss of those ecosystems [19]. Aquatic biodiversity refers to the groups of organisms that determine the ecological status and functioning of rivers. The ecological status of a river environment is strongly affected by the composition and condition of the populations of organisms that inhabit it, known as aquatic communities, which include phytoplankton, zooplankton, and benthos. Their composition and functional status provide useful information and quantifiable indicators for the ecological health of freshwater systems [20]. Water environmental factors, including hydrodynamics and water quality, are regarded as one of the most essential determinants of the river ecosystem [21,22]. Therefore, modeling the response relationship of aquatic communities to water environmental factors is an efficient way to simulate water ecosystem features.

As awareness of aquatic community-related water eco-environment problems increases, model quantification of the water ecosystem in rivers remains one of the most serious challenges. Many studies have been conducted on the ecological models to quantify aquatic communities [23,24], with a main focus on the uptake model [25,26], bioenergetic model [27], evolution model of aquatic organisms [28–30], species interactions model [31,32], and numerical models that quantify relations between aquatic densities and diversity [33]. Those types of models are currently becoming more comprehensive, allowing for the modeling and prediction of an increasing number of aquatic features and processes [34]. However, as the complexity of the model increases, the number of data samples required, the number of parameters, and the difficulty of fitting the model all increase accordingly. In addition, ecology models generally have restricted access to source code and limited model flexibility [23]. For dam-controlled rivers, many modeling approaches are applied to evaluate the changes in hydrological regime [35], water temperature [36], or water quality [37] caused by dam building or operation. However, for water ecosystems that are regularly disturbed by dams, many evaluations have been carried out by typically using weighted bio-indicators or standard test species to analyze as evaluation endpoints [38], while model techniques are mainly focused on the eco-impact of monomer (unit) or group, e.g., fish [39], macroinvertebrates [40,41], algae [42] and vegetation [43], instead of the diversity alteration of community in the biological realm.

Although these ecological models offer numerous advantages and benefits to practices in a wide range of scales so far, they also have weaknesses, such as the model structure that is overly complicated, constrained model flexibility, and contains a significant amount of hard-to-get biotic response data (e.g., abundance, survival of aquatic organisms). Computation time may grow with the complexity of the modeling approach and implementation, and specifically, highly over-parameterized models are intrinsically data-hungry and require substantial amounts of data and information that do not necessarily exist, creating a need for parameter estimation without much background information to constrain parameter spaces. In general, obtaining biotic response data by field experiments is more expensive and time-consuming than physical and chemical water environmental factors (e.g., hydrodynamics and water quality). Therefore, new approaches and simple models that are computationally less costly and produce outputs that are easier to interpret are needed to improve predictions of how the existence and operation of dams will influence aquatic communities, helping to frame a strategy to support healthy and sustainable water ecosystem development.

An alternative to advancing this field could be through a model imputing easy-to-get physical and chemical water environmental factors to simulate aquatic community impact by the dam to overcome the limitations of previous studies. Our group previously constructed a multivariate nonlinear regression model (MNLE model) to quantify the interactive relationship between environmental factors and aquatic communities in a dam-controlled river [44]. However, the accuracy of the model is not excellent because the nonlinear regression functions cannot truthfully describe the complex interrelated processes of aquatic communities driven by environmental factors.

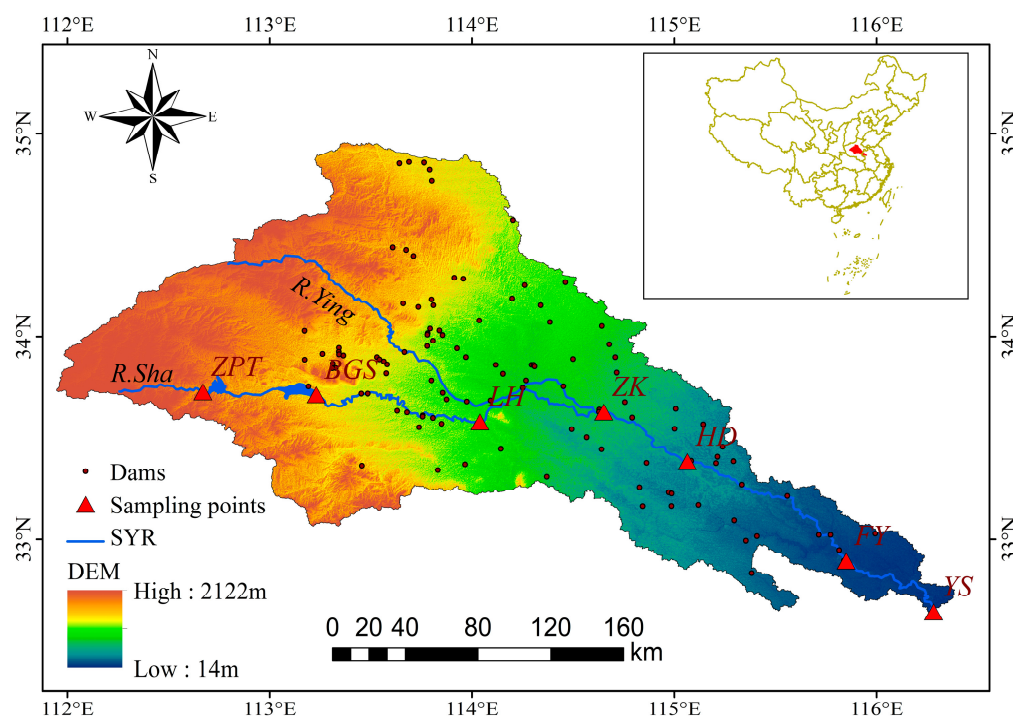
The black-box model, which does not require any assumptions about the system but can frequently precisely capture all phenomena that are properly represented by data [45], can improve model accuracy without capturing complex interrelated processes. Backpropagation (BP) neural network, which is a forward learning algorithm and an error BP neural network, is one of the typical black box models successfully developed and widely used for estimating and predicting, in particular, modeling non-linear water ecosystems and the water science field. The BP neural network has substantial self-learning and nonlinear mapping capabilities [46]. However, due to limitations such as slow convergence speed, sensitivity to weight initialization, and tendency to slip into local extremes [47], using BP neural networks directly in prediction may not achieve yield satisfactory outcomes. Genetic algorithm (GA) is a stochastic search method that can be effective in optimizing the BP neural networks to overcome its inherent limitations. GA mimics the process of natural biological evolution by employing the principle of survival of the fittest to demonstrate excellent generalization capacity and good performance with higher precision [48].

Prediction models that quantify the aquatic community's response to water environmental factors in dam-controlled rivers have shown to be helpful tools in identifying the impact of the water ecosystem by the dam, as modeling aquatic communities simply, efficiently, and economically can be a challenge. In order to develop a better model for quantifying the water ecological environment using environmental factors that are easy to get and readily available, in this study, we expanded on previous research in this field by specifically focusing on (1) analyzing the linear and black-box response between aquatic communities and water environmental factors to develop the linear regression model (MLE model) and GA-BP model, respectively; (2) comparing the models in order to find the optimal model with the best performance; (3) applying the model to reproduce the inter-annual and seasonal variations of the aquatic community of the Huaidian (HD) site in the Shaying River (SYR), China, from 2005 to 2014. This study can become a contribution to showing how quantitative models in other dam-controlled rivers can be applied to the aquatic community to help determine management actions in the dam-controlled river to better fulfill the water eco-environment preservation goal.

## 2. Materials and Methods

### 2.1. Study Area and Data Collection

The Shaying River (SYR) is the largest and perhaps most polluted tributary of the Huaihe River Basin [49], which is regulated by more than 115 dams, reservoirs, and sluices. The river flow regulation through those dams and sluices has a significant impact on the flow regime of the SYR [50], causing ecological changes [51,52]. The SYR is located between 32°31'~34°59' of north latitude and 111°56'~116°31' of east longitude, which belongs to the warm temperate semi-humid continental monsoon climate zone. The annual average temperature ranges from 14 °C to 16 °C, and the annual precipitation is 753.43 mm, with four distinct seasons [53]. Rainfall varies sharply during the year; the main flood season is June to August, with peak flows usually occurring in July. Forest coverage in the headwaters exceeds 80%, and cultivated land is mainly planted with wheat in the middle stream and downstream. The SYR flows through more than 40 cities and counties, which are characterized by frequent water pollution accidents and prominent contradictions in flood control and pollution prevention. Figure 1 shows the geographical location of the research area, the topographic gradients from its sources to the outlet, and the distribution of sampling points.



**Figure 1.** The geographical location of the research area, the topographic gradients from its sources to the outlet, and the distribution of sampling points.

We divide the whole year into the dry season and the wet season. The wet season lasts from June to October, and its runoff accounts for more than 70% of the total annual runoff, while the dry season is from November to May. The SYR was selected to carry out monitoring experiments on the river ecology and environment during the dry season and the wet season to obtain seasonal data on the species and density of phytoplankton, zooplankton, and benthos. Seven sampling points (see Figure 1) were set up, and from upstream to downstream of the SYR were Zhaopingtai (ZPT), Baiguishan (BGS), Luohe (LH), Zhoukou (ZK), Huaidian (HD), Fuyang (FY), and Yingshang (YS). We carried out one field investigation in the SYR every six months to obtain the river's ecological and environmental samples. Since December 2012, seven field sampling campaigns have been carried out, and 49 groups of seasonal data have been obtained. Both datasets of hydrodynamic and water quality are provided by the Huaihe River Commission of the Ministry of Water Resources, P.R.C., and China National Environmental Monitoring Centre (CNEMC). The seasonal water quality datasets include dissolved oxygen (*DO*), *pH*, total dissolved solids (*TDS*), ammonia nitrogen ( $\text{NH}_3\text{-N}$ ), five-day biochemical oxygen demand ( $\text{BOD}_5$ ), and permanganate index ( $\text{COD}_{Mn}$ ), total nitrogen (*TN*), total phosphorus (*TP*). The seasonal hydrodynamic factor includes water flow (*Q*) and temperature (*T*). In addition, habitat quality assessments were conducted for each monitoring section to determine the spatial variability of habitats in the river ecosystem [54].

## 2.2. Methodology

### 2.2.1. Output Variables

Aquatic biodiversity was selected as the output variable of the model. Changes in the river environment have the potential to directly cause changes to the community composition and diversity of aquatic plants and animals. In general, the more complex the community structure, the higher the biodiversity formed. The Shannon-Wiener diversity indexes based on biological data can be used to quantify the diversity of aquatic communities. The species diversity of phytoplankton (*P-SWI*), zooplankton (*Z-SWI*), and zoobenthos (*B-SWI*) were used to present the aquatic community in this research; thus, the output variables of the model are denoted as  $f_{P-SWI}$ ,  $f_{Z-SWI}$ , and  $f_{B-SWI}$ , respectively. The species and

density of phytoplankton, zooplankton, and benthos need to be transformed into species diversity through the Shannon-Wiener index [55], and the larger the index, the higher the complexity of the community and the greater the aquatic biodiversity. The Shannon-Wiener index was calculated here as follows:

$$f_{SWI} = - \sum_{i=1}^m p_i \log_2 p_i \tag{1}$$

$$p_i = \frac{n_i}{N} \tag{2}$$

where  $f_{SWI}$  was the Shannon-Wiener index;  $i$  was the number of species,  $i = 1, 2, 3 \dots m$ ;  $n_i$  was the density of the  $i$ -th species;  $N$  was the sum of the densities of all species; and  $p_i$  was the proportion of the  $i$ -th species [56].

### 2.2.2. Input Variables

The key factors that affect the aquatic community from the water quality and hydrodynamic factors were the input variables of the model. We obtained the key factors from our previous research, which identified the key influencing factors using Redundancy Analysis and Monte Carlo tests based on 49 groups of the seasonal water quality and hydrodynamic factors collected from the seven field sampling campaigns on the SYR, and found that  $Q$ ,  $DO$ ,  $TP$ , and  $TN$  were key impact factors for  $f_{P-SWI}$ ;  $DO$ ,  $Q$ , and  $TN$  for  $f_{Z-SWI}$ ;  $DO$ ,  $Q$ , and  $COD_{Mn}$  for  $f_{B-SWI}$  [44]. The input and output variables required for model calculation are shown in Table 1.

**Table 1.** The input and output variables of the model.

Input Variables	Output Variables
$Q, DO, TP, \text{ and } TN$	$f_{P-SWI}$
$DO, Q, \text{ and } TN$	$f_{Z-SWI}$
$DO, Q, \text{ and } COD_{Mn}$	$f_{B-SWI}$

### 2.2.3. MLE Model Construction

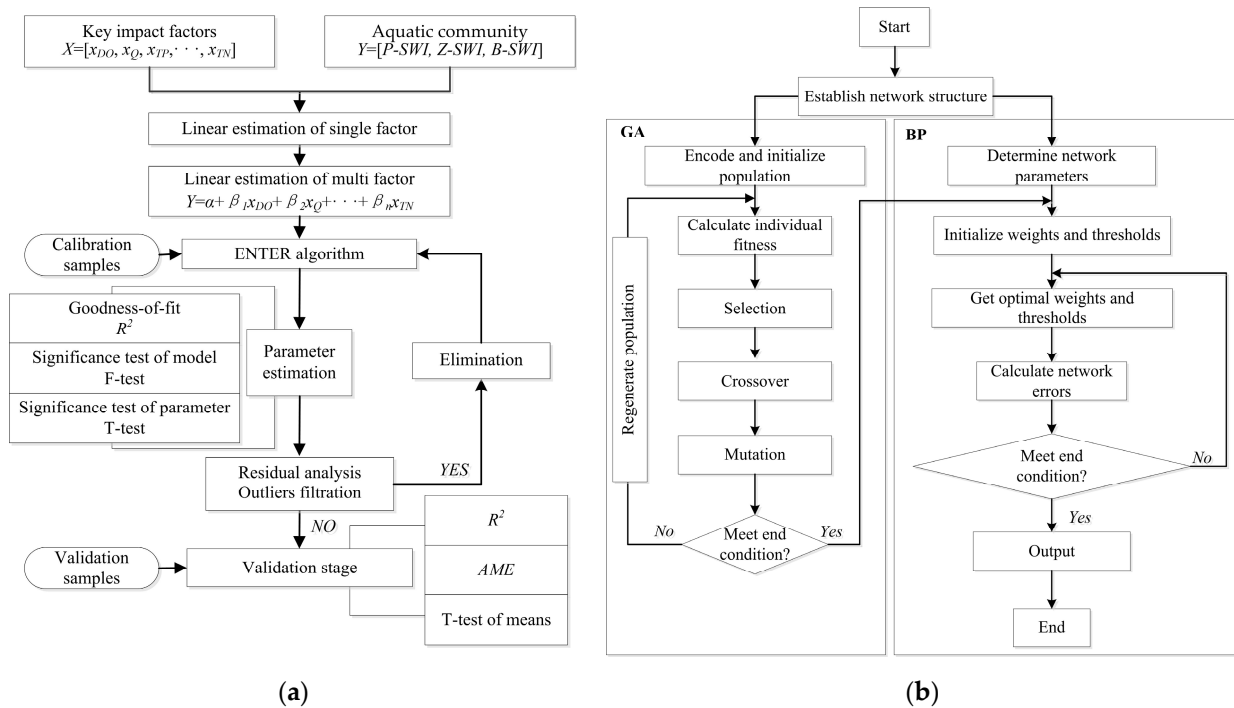
The building process of the MLE model was shown in Figure 2a, which includes significance tests of the parameters for optimal correlation between output and input variables, the estimation of goodness-of-fit, residual analysis, and elimination of outliers when the residuals fall outside the range  $(-2, 2)$ , etc. The specific operation flow can refer to a similar description by Zuo et al. (2019) [44]. In this paper, based on the 49 groups of data obtained from the seven monitoring experiments carried out in the SYR, 40 groups of the data (80% of total data) were randomly generated as the calibration samples, and 9 groups of the data (20% of the total data) were generated as the validation samples. The structures of the MLE models are as follows:

$$f_{P-SWI} = a_P + \beta_{P1}x_{DO} + \beta_{P2}x_Q + \beta_{P3}x_{TP} + \beta_{P4}x_{TN} \tag{3}$$

$$f_{Z-SWI} = \alpha_Z + \beta_{Z1}x_{DO} + \beta_{Z2}x_Q + \beta_{Z3}x_{TN} \tag{4}$$

$$f_{B-SWI} = \alpha_B + \beta_{B1}x_{DO} + \beta_{B2}x_Q + \beta_{B3}x_{COD_{Mn}} \tag{5}$$

where  $f$  is the Shannon-Wiener index/output variable;  $\beta$  is the coefficients;  $\alpha_0$  is a constant;  $x$  is the key impact factor/input variable.



**Figure 2.** Flow diagram indicating the model conceptualization and building process of the MLE and GA-BP models. (a) MLE model, (b) GA-BP model.

2.2.4. GA-BP Model Construction

BP artificial neural networks as a member of the machine learning model, is a kind of multi-layer feedforward network trained according to an error backpropagation algorithm. One of its great advantages is that it can learn and store the mapping relationship between the input and output patterns of the network without clearly describing the mathematical relationship equation in advance. Generally, the BP neural network with a single hidden layer including enough neurons can approach any nonlinear function by adjusting its connection weight and transfer function. GA is a pseudo-biological optimization algorithm, and nature is searching for the optimal solution by simulating the evolutionary process of “survival of the fittest” for the population composed of the feasible solutions to the problem and combining natural selection and genetic phenomena. The feasible solutions are first encoded as chromosomes or individuals, and then excellent individuals with high fitness are selected for genetic operation. Genetic operation mainly includes three kinds of genetic operators: selection, crossover, and mutation. Among them, the selection and crossover operators realize the search function, and the mutation operator improves the optimization ability of the algorithm. The combination of the above two makes the model have global searchability, improves the efficiency and convergence speed, and then further improves the accuracy of the model [48]. The building process of the BP optimized by GA was shown in Figure 2b.

In this machine learning model, we constructed the BP with a three-layer network topology composed of the input layer, hidden layer, and output layer for this study. The number of neurons in the input layer was determined by the number of key impact factors, and the number of neurons in the output layer was determined by the number of the characteristic factors of the aquatic community (output variable). For the hidden layer, the determination of the number of neurons is relatively complex, and there is no mature theory at present. If the number of neurons is too small, the convergence speed and the training accuracy of the network will be low; if the number is too large, the network structure and the amount of iterative calculation will be huge and even overfitting. We used the trial-and-error method to train the network several times to determine the number of hidden layer neurons and finally obtained the best topological structure of the BP, and

MATLAB was used to write the optimized BP algorithm to obtain the best solution. The initial values of neurons in the hidden layer can be obtained by the empirical formula as follows:

$$K = 2m + 1 \quad (6)$$

where  $K$  was the number of the hidden layer neurons;  $m$  was the number of the input layer neurons.

A Tan-Sigmoid function was chosen as the excitation function to activate neurons. To meet the range requirements of the activation function and avoid the network paralysis caused by too-large data series, the sample data should be normalized to be  $[-1, 1]$ . The normalization equation was as follows:

$$y = 2 ( x - x_{min} ) / ( x_{max} - x_{min} ) - 1 \quad (7)$$

where  $x$  was the original data;  $y$  was the normalized data;  $x_{min}$  was the minimum of the original data series;  $x_{max}$  was the maximum of the original data series.

### 2.2.5. Statistical Analyses

A total of 49 groups of data were obtained from the seven field sampling campaigns in SYH. The mean absolute error (*MAE*) and mean relative error (*MRE*) allowed the identification of outliers and estimation of the accuracy of the prediction models. The coefficient of determination (*R*-squared or  $R^2$ ) was used for correlations between predicted and observed outcomes. SPSS 22.0 was the statistical calculation software.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - f_i| \quad (8)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - f_i|}{y_i} \times 100\% \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - f_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

where  $y_i$  is the true value,  $f_i$  is the predictive value, and  $n$  is the number of predictive values. The smaller the error is and the closer  $R^2$  is to 1.0, the higher the simulation accuracy of the model is.

The *t*-test (two-tailed) was used to compare the means of the simulation and the observation groups to test whether there was a significant difference. The null hypothesis ( $H_0$ ) is that the true difference between the two-group means is zero, and the alternate hypothesis ( $H_a$ ) is that the true difference is different from zero. *Z*-score value was confirmed by the following formula:

$$Z - score = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left( S^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \right)}} \quad (11)$$

where  $\bar{x}_1$  and  $\bar{x}_2$  are the mean values of the simulation group and observation group respectively.  $S^2$  is pooled standard error of the two groups.  $n_1$  and  $n_2$  are the sample numbers. The null hypothesis is accepted if  $|Z - score| \leq 1.96$ .

### 3. Results

#### 3.1. Model Structure and Parameter Setting

##### 3.1.1. MLE Model

After deleting the sets of standardized residual exception data that fall outside the region  $(-2, 2)$ , the “ENTER algorithm” was used to create the ideal fitting model from the randomly screened out 40 groups of sample data. According to the quantization procedure of the MLE model (Figure 2a) and the “ENTER algorithm” based on SPSS 22.0, Equations (3)~(5) can be rewritten as follows:

$$f_{P-SWI} = 2.1809 + 0.1228x_{DO} + 0.0043x_Q - 0.7810x_{TP} - 0.0598x_{TN} \tag{12}$$

$$f_{Z-SWI} = -0.0553 + 0.1894x_{DO} + 0.0062x_Q - 0.0048x_{TN} \tag{13}$$

$$f_{B-SWI} = 0.5772 + 0.1275x_{DO} + 0.0138x_Q - 0.2294x_{COD_{Mn}} \tag{14}$$

Table 2 displays the results of the goodness-of-fit and F tests for the MLE models. The  $R^2$  for  $P$ -SWI,  $Z$ -SWI, and  $B$ -SWI is 0.561, 0.678, and 0.621, respectively, and the F tests for all of the models are significant at the  $p < 0.05$  level, indicating that the MLE models for  $P$ -SWI,  $Z$ -SWI, and  $B$ -SWI have passed the 95% level of a significance test. The results of the t-test used to find the optimal parameters in the MLE models are shown in Table 3. Except for  $B$ -SWI, the majority of other models’ parameters and constant coefficients failed the 95% level of a significance test, implying that while MLE models are significant, not all key impact factors have a significant linear relationship with aquatic communities.

**Table 2.** Goodness-of-fit and F tests of the MLE models.

Model	$R^2$	Significance Test	
		F	p
$f_{P-SWI}$	0.561	10.538	0.000 *
$f_{Z-SWI}$	0.678	23.903	0.000 *
$f_{B-SWI}$	0.621	19.155	0.000 *

\*: significance at 0.05 level.

**Table 3.** Structure and parameter settings of MEL models.

Model	Function Item	Parameter		Significance Test	
		Name	Value	t	p
$f_{P-SWI}$	constant	$\alpha_p$	2.1809	4.725	0.000 *
	$x_{DO}$	$\beta_{p1}$	0.1228	3.317	0.002 *
	$x_Q$	$\beta_{p2}$	0.0043	1.421	0.165
	$x_{TP}$	$\beta_{p3}$	-0.7810	-0.927	0.361
	$x_{TN}$	$\beta_{p4}$	-0.0598	-1.293	0.205
$f_{Z-SWI}$	constant	$\alpha_z$	-0.0553	-0.1650	0.870
	$x_{DO}$	$\beta_{z1}$	0.1894	7.0842	0.000 *
	$x_Q$	$\beta_{z2}$	0.0062	2.6388	0.013
	$x_{TP}$	$\beta_{z3}$	-0.0048	-0.1433	0.887
$f_{B-SWI}$	constant	$\alpha_b$	0.5772	1.0376	0.307
	$x_{DO}$	$\beta_{b1}$	0.1275	3.2575	0.003 *
	$x_Q$	$\beta_{b2}$	0.0138	4.3158	0.000 *
	$x_{COD_{Mn}}$	$\beta_{b3}$	-0.2294	-2.8501	0.007 *

\*: significance at 0.05 level.

##### 3.1.2. GA-BP Model

The key impact factors are used as input variables of the models. The model parameters include transfer function, training function, learning rate ( $v$ ), training times ( $epochs$ ), population ( $Mp$ ), iteration times ( $Ts$ ), crossover probability ( $pc$ ), and mutation probability



( $pm$ ). After several model tests based on the trial-and-error method, the best number of hidden layer neurons for  $P$ -SWI,  $Z$ -SWI, and  $B$ -SWI is 9, 3, and 4, respectively. The model structures and parameters were shown in Table 4.

**Table 4.** Structure and parameter settings of GA-BP models.

Model		GA-BP		
		$P$ -SWI	$Z$ -SWI	$B$ -SWI
variables	input output	$DO, Q, TN, TP$ $f_{P-SWI}$	$Q, DO, TN$ $f_{Z-SWI}$	$Q, DO, COD_{Mn}$ $f_{B-SWI}$
layer nodes	input layer neurons, $m_1$	4	3	3
	hidden layer neurons, $m_2$	9	3	4
	output layer neurons, $m_3$	1	1	1
parameters	transfer function training function		tansig and purelin trainlm	
	learning rate, $v$	0.1	0.1	0.1
	training times, $epochs$	100	100	50
	population, $Mp$	30	20	20
	iteration times, $Ts$	80	100	50
	crossover probability, $pc$	0.3	0.3	0.3
	mutation probability, $pm$	0.1	0.1	0.1

### 3.2. Model Calibration and Validation

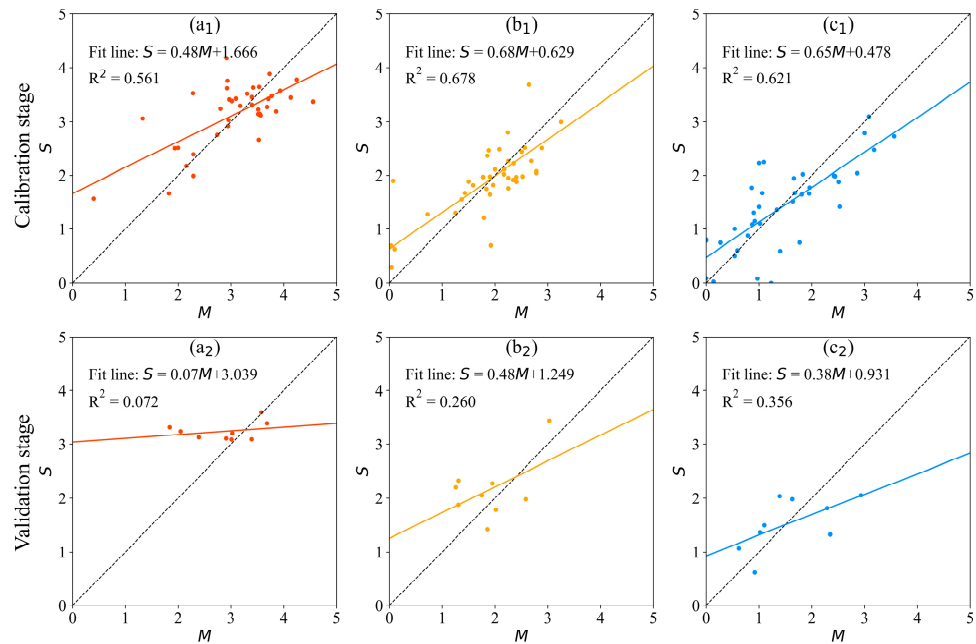
#### 3.2.1. MLE Model

After the optimal parameters of the MLE models were obtained, the performance of the MLE models at the calibration and validation stage is presented in Figure 3. At the calibration stage, the mean absolute error (MAE) of the MLE models for  $P$ -SWI,  $Z$ -SWI, and  $B$ -SWI were 0.432, 0.355, and 0.525, respectively, and the  $R^2$  are 0.561, 0.678, and 0.621, respectively. At the validation stage, the MAE of the MLE models for  $P$ -SWI,  $Z$ -SWI, and  $B$ -SWI were 0.498, 0.539, and 0.599, respectively, and the  $R^2$  was 0.072, 0.260, and 0.351, respectively. It can be seen in Figure 3 that the performance of the MLE models is reasonable in the calibration stage but not ideal in the validation stage, particularly for  $P$ -SWI, where the fitting accuracy is the worst.

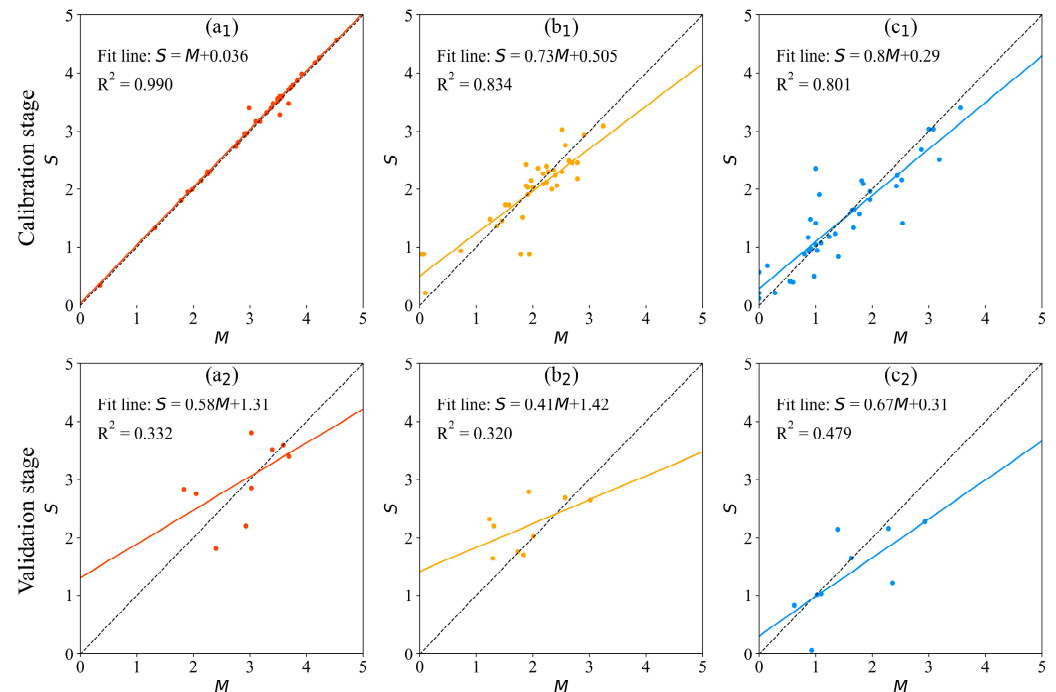
#### 3.2.2. GA-BP Model

GA-BP models were established to predict the characteristics of the aquatic community. A total of 49 groups of data are selected to model the prediction model, and the numbers of calibration samples and validation samples are the same as in the MLE models. For the model of  $P$ -SWI, each group of sample data in the data set contains four related parameters and one original parameter, so the input matrix and output matrix of the model training sample set are  $40 \times 4$  dimensions and  $40 \times 1$  dimensions, respectively, and the input matrix and output matrix of the model test sample set are  $9 \times 4$  dimensions and  $9 \times 1$  dimension, respectively. For the models of  $Z$ -SWI and  $B$ -SWI, each group of sample data in the data set contains three related parameters and one original parameter, so the input matrix and output matrix of the model training sample set are  $40 \times 3$  dimensions and  $40 \times 1$  dimensions, respectively, and the input matrix and output matrix of the model test sample set are  $9 \times 3$  dimensions and  $9 \times 1$  dimension, respectively. Figure 4 shows the performance of the GA-BP models. At the calibration stage, the MAE of the GA-BP models for  $P$ -SWI,  $Z$ -SWI, and  $B$ -SWI were 0.051, 0.296, and 0.286, respectively, and the  $R^2$  was 0.990, 0.834, and 0.801, respectively. At the validation stage, The MAE of the GA-BP

models for *P-SWI*, *Z-SWI*, and *B-SWI* were 0.486, 0.426, and 0.427, respectively, and the  $R^2$  was 0.332, 0.320, and 0.479, respectively.



**Figure 3.** The performance of the MLE models. *M*: measured values; *S*: simulated values. Subgraphs (a–c) represent *P-SWI*, *Z-SWI*, and *B-SWI*, respectively. Subscripts 1 and 2 represent the calibration stage and the validation stage, respectively.



**Figure 4.** The performance of the GA-BP models. *M*: measured values; *S*: simulated values. Subgraphs (a–c) represent *P-SWI*, *Z-SWI*, and *B-SWI*, respectively. Subscripts 1 and 2 represent the calibration stage and the validation stage, respectively.

We performed *t*-tests on the means of simulated and observed outputs to further verify the reliability and rationality of the models, and the results are shown in Table 5. Both the MLE and GA-BP models for *P-SWI*, *Z-SWI*, and *B-SWI* were not statistically different

between simulation and observation at the 0.05 level, showing that the models constructed herein of *P-SWI*, *Z-SWI*, and *B-SWI* are valid.

**Table 5.** The results of the *t*-test on the means of simulated and observed outputs.

Model	The Mean Value of Outputs		Calibration Stage			Validation Stage		
			$f_{P-SWI}$	$f_{Z-SWI}$	$f_{B-SWI}$	$f_{P-SWI}$	$f_{Z-SWI}$	$f_{B-SWI}$
MLE	Simulation		3.144	1.923	1.442	3.241	2.151	1.377
	Observation		3.055	1.916	1.440	2.880	1.881	1.586
	<i>t</i> -test (2-tailed)	H <sub>0</sub> Z-score results	0.55	0.10	−0.22	1.59	0.99	−0.19
mean simulation = mean observation no significant difference at the 0.05 level								
GA-BP	Simulation		3.082	1.908	1.442	2.971	2.195	1.377
	Observation		3.055	1.916	1.440	2.880	1.881	1.586
	<i>t</i> -test (2-tailed)	H <sub>0</sub> Z-score results	0.14	−0.05	0.01	0.29	1.26	−0.58
mean simulation = mean observation no significant difference at the 0.05 level								

### 3.3. Model Comparison

The performance of the tested models is demonstrated in Table 6 by a comparison of the MLE models, MNLE models completed by our team previously [44], and GA-BP models. At the calibration stage, the performance order of the models is GA-BP model > MNLE model > MLE model. The *R*<sup>2</sup> of the GA-BP models is the highest, and the *MAE* is the lowest, revealing that the GA-BP models have the best accuracy. At the validation stage, although the *R*<sup>2</sup> of the GA-BP model for *P-SWI* is slightly lower than that of the MNLE model, *MAE* is better. For *B-SWI* and *Z-SWI*, both *R*<sup>2</sup> and *MAE* of the GA-BP models are better than that of the MLE and MNLE models.

**Table 6.** Comparison of the MLE, MNLE, and GA-BP models.

Model	Model Equation/Structure	Calibration			Validation		
		<i>MRE</i>	<i>MAE</i>	<i>R</i> <sup>2</sup>	<i>MRE</i>	<i>MAE</i>	<i>R</i> <sup>2</sup>
GA-BP	$f_{P-SWI} : m_1 = 4, m_2 = 9, m_3 = 1,$ $epochs = 100, Mp = 30, Ts = 80$	1.6%	0.0338	0.990	20%	0.4833	0.332
	$f_{Z-SWI} : m_1 = 3, m_2 = 3, m_3 = 1,$ $epochs = 100, Mp = 20, Ts = 100$	13%	0.3075	0.834	28.2%	0.4217	0.320
	$f_{B-SWI} : m_1 = 3, m_2 = 4, m_3 = 1,$ $epochs = 50, Mp = 20, Ts = 50$	18.5%	0.2862	0.801	28.1%	0.4523	0.479
MLE	$f_{P-SWI} = 2.1809 + 0.1228x_{DO} + 0.0043x_Q$ $- 0.7810x_{TP} - 0.0598x_{TN}$	20.7%	0.4323	0.561	22.5%	0.4979	0.072
	$f_{Z-SWI} = -0.0553 + 0.1894x_{DO}$ $+ 0.0062x_Q - 0.0048x_{TN}$	17%	0.3551	0.678	33.7%	0.5394	0.260
	$f_{B-SWI} = 0.5772 + 0.1275x_{DO}$ $+ 0.0138x_Q - 0.2294x_{COD_{Mn}}$	41.5%	0.5246	0.621	37.5%	0.5985	0.356
MNLE	$f_{P-SWI} = -0.2371 - 0.0219x_{DO}^2 - 0.0491x_{TN}^2$ $+ 0.5080x_{DO} + 0.0136x_Q + 0.4750x_{TN}$	9.1%	0.2309	0.811	22%	0.5136	0.357
	$- 0.0045x_Qx_{TN} - 0.1014x_{TN}x_{TP}$						
	$f_{Z-SWI} = -1.6626 + 0.0013x_Qx_{TN} + 1.5803 \ln x_{DO}$	15.1%	0.3443	0.723	34.5%	0.4587	0.305
	$f_{B-SWI} = 0.5450 - 0.0262x_Q + 0.0044x_{DO}x_Q$ $+ 0.7714/x_{COD_{Mn}}$	42%	0.4262	0.657	29.5%	0.5179	0.407

In addition, at the calibration stage, the *MRE* of GA-BP models for *P-SWI*, *Z-SWI*, and *B-SWI* were 1.6%, 13%, and 18.5%, respectively, which was much lower than that of MLE and MNLE models. At the validation stage, the *MRE* of GA-BP models for *P-SWI*, *Z-SWI*, and *B-SWI* increased to 20%, 28.2%, and 28.1%, respectively, but it remained

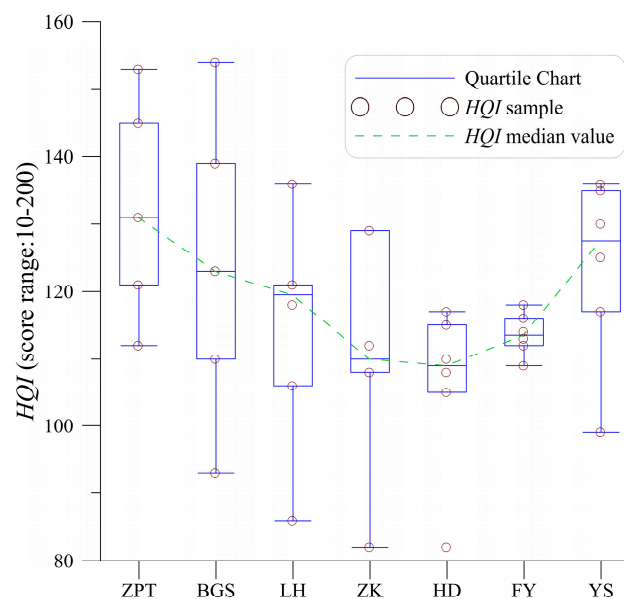
the lowest when compared to the MLE and MNLE models. As a result, the simulation performance of the GA-BP models is better than that of the MLE and MNLE models and is more stable and reliable in reproducing the river ecological environment. It can be concluded that the relationship between water environment factors and aquatic community in the dam-controlled river is exceedingly complicated and cannot be explained just by linear correlations.

## 4. Discussion

### 4.1. Variability of Habitat Quality

The adaptability of the constructed model to all sampling sites in the study area is a major concern that cannot be ignored. Given the variability of habitats, directly quantifying the aquatic communities under different habitat qualities will not result in the same values each time or for each sampling site. Thus, it is important to assess the variability in habitat quality that can cause heterogeneity in aquatic communities. In this study, the habitat assessment index (*HQI*) from Wei et al. (2009) [54] was used to identify the variability of habitat quality, which included ten parameters (substrate, habitat complexity, velocity-depth combination, bank stability, bank conservation, vegetation cover, vegetation diversity, the intensity of human activities, water cognition, and riverside land use).

Figure 5 shows the habitat quality assessment results for each site of the SYR from 2012. The average *HQI*s of the SYR range between 110 and 130 (the threshold range for habitat quality is 10–200), indicating that there is little spatial variability in habitat quality. Therefore, the constructed model can be applied to all sampling sites in SYR.

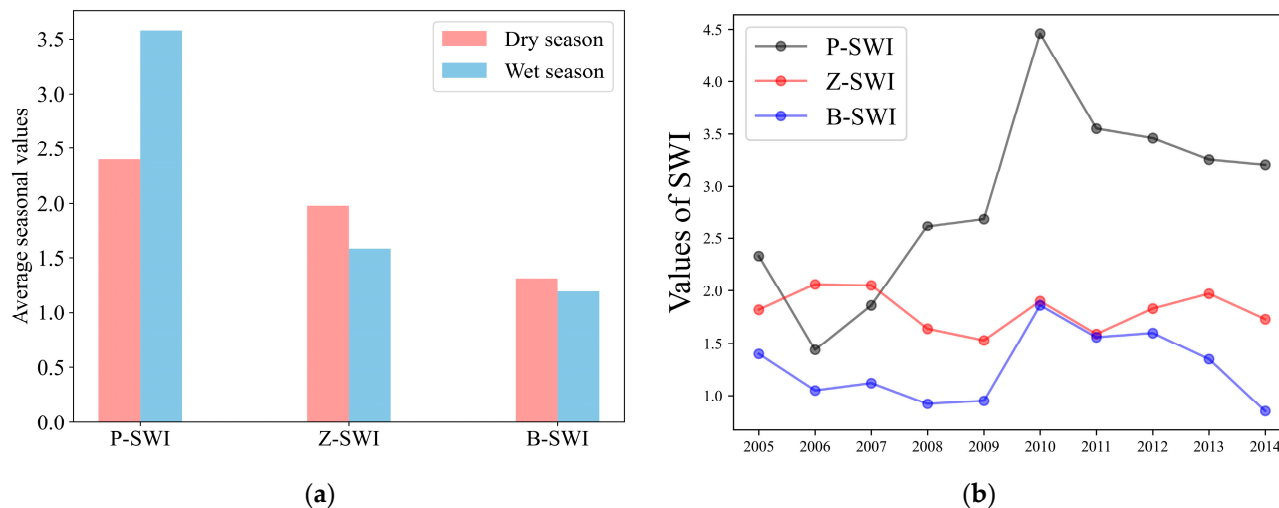


**Figure 5.** The assessment results of the habitat quality index (*HQI*).

### 4.2. Aquatic Community Characteristics

The GA-BP models were applied to reproduce the aquatic community characteristics of the HD site in the SYR in the wet season and the dry season from 2005 to 2014. Figure 6a shows the ten-year average seasonal values of the *P-SWI*, *Z-SWI*, and *B-SWI*. Both in the dry season and the wet season, *P-SWI* is at the maximum, and *B-SWI* is at the minimum. For *P-SWI*, the value in the wet season is higher than that in the dry season, but for *Z-SWI* and *B-SWI*, the result is the opposite, indicating that the seasonal variation of the *P-SWI*, *Z-SWI*, and *B-SWI* are inconsistent. In natural rivers, seasonal changes of *P-SWI*, *Z-SWI*, and *B-SWI* should be consistent because the interaction of the bottom-up effect (i.e., the density, biomass, and species richness of the lower trophic class determine the population structure of the higher trophic class) and the top-down effect (i.e., the higher trophic class

control and affect the community structure of the lower trophic class through predation) balances the evolution of biological community in the river ecosystem [57]. The inconsistent seasonal variation of the *P-SWI*, *Z-SWI*, and *B-SWI* in the HD site of the SYR indicates that the disturbance of dam control has a negative impact on the aquatic community of the SYR [50,52].



**Figure 6.** The variation of the *P-SWI*, *Z-SWI*, and *B-SWI*. (a) Seasonal variation. (b) Inter-annual variation.

Figure 6b shows the inter-annual variation of the *P-SWI*, *Z-SWI*, and *B-SWI*. From 2005 to 2014, *P-SWI* showed a significant growth trend, *Z-SWI* changed smoothly, *B-SWI* fluctuated, but the overall trend was increasing. Generally, the growth period of phytoplankton is shorter and faster than that of zooplankton and zoobenthos, and that of zoobenthos is the slowest and the longest. Therefore, it can be inferred that the aquatic community in the HD site of the SYR has improved from 2005 to 2014. However, in addition to the values of *P-SWI* of more than 3.0, *Z-SWI* and *B-SWI* are still at a low level, especially *B-SWI*, which did not exceed 2.0 from 2005 to 2014, revealing that the ecological diversity level is still low, and the aquatic community still needs to be further repaired and protected. Perhaps dam operation based on environmental flows will be one of the effective ways to protect the disturbed-river ecosystems [58,59].

#### 4.3. Model Constraints

Even while the GA-BP aquatic community model performs better than the other two models, it inevitably still has certain flaws: (1) The accuracy of the model is affected by the number of the sample database. Specifically, the more monitoring samples there are, the better the simulation accuracy of the GA-BP model. (2) The GA-BP model is a “black box” model that does not account for the physical, chemical, and biological interactions between the aquatic community and its influencing factors. (3) The model parameters cannot be directly applied to other rivers; they can only be used as a reference for modeling aquatic communities in other rivers. (4) The contribution of dam effects on the aquatic community model has yet to be quantified in the current model.

## 5. Conclusions

We provide different models of considering the actual demand to simulate the aquatic community and already present useful information about the pros and cons of the different types and structures of the models. The models are compared to real-life cases in SRY, China. The main findings of this study are as follows.

(1) MLE and GA-BP models, which we established respectively based on the linear and black-box relationships between aquatic biodiversity and water environmental factors, were effective in the simulating aquatic community in the dam-controlled river.

(2) Comparison of the MLE, MNLE, and GA-BP models, the GA-BP model performed the best in the context of this study. GA-BP model is more efficient and accurate for predicting aquatic communities by just inputting key impact factors such as water quality and hydrodynamics, which are easy to get.

(3) Applying the GA-BP model to reproduce the aquatic community characteristics of the HD site in the SYR from 2005 to 2014 revealed that the seasonal variation of the *P-SWI*, *Z-SWI*, and *B-SWI* in the HD site of the SYR was inconsistent, and the inter-annual levels of diversity were low. Dam control has a negative impact on the aquatic community, and the operation based on environmental flows will be one of the effective ways to protect the SYR's ecosystems.

Our models can be used as a tool for aquatic community simulation and prediction; as an economical, efficient, and referential approach in other dam-controlled rivers; and as part of water eco-environment preservation by assisting in dam planning and management strategies. The next phase of our work will focus on collecting more detailed sample data in the study area to further improve the accuracy of the model and develop physical mechanism-based models.

**Author Contributions:** J.L.: Methodology, Formal analysis, Writing—original draft, Resources. C.Z.: Experiment, Modeling, Writing—edit figures & tables, Validation. Q.Z.: Supervision, Project administration, Writing—review & editing. C.H.: Experiment. S.K.: Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China [Grant No. 51909091 and 51979107] and the China Scholarship Council [Grant No. 202108410234].

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** The authors are grateful to the editors and anonymous reviewers for their insightful comments and helpful suggestions.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Reid, A.J.; Carlson, A.K.; Creed, I.F.; Eliason, E.J.; Gell, P.A.; Johnson, P.T.; Kidd, K.A.; MacCormack, T.J.; Olden, J.D.; Ormerod, S.J.; et al. Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biol. Rev. Camb. Phil. Soc.* **2019**, *94*, 849–873. [[CrossRef](#)] [[PubMed](#)]
2. Grizzetti, B.; Liquele, C.; Pistocchi, A.; Vigiak, O.; Zulian, G.; Bouraoui, F.; De Roo, A.; Cardoso, A.C. Relationship between ecological condition and ecosystem services in European rivers, lakes and coastal waters. *Sci. Total Environ.* **2019**, *671*, 452–465. [[CrossRef](#)] [[PubMed](#)]
3. Bouska, K.L.; Houser, J.N.; De Jager, N.R.; Van Appledorn, M.; Rogala, J.T. Applying concepts of general resilience to large river ecosystems: A case study from the Upper Mississippi and Illinois rivers. *Ecol. Indic.* **2019**, *101*, 1094–1110. [[CrossRef](#)]
4. Grill, G.; Lehner, B.; Thieme, M.; Geenen, B.; Tickner, D.; Antonelli, F.; Babu, S.; Borrelli, P.; Cheng, L.; Crochetiere, H.; et al. Mapping the world's free-flowing rivers. *Nature* **2019**, *569*, 215–221. [[CrossRef](#)]
5. Belletti, B.; Garcia de Leaniz, C.; Jones, J.; Bizzi, S.; Börger, L.; Segura, G.; Castelletti, A.; Van de Bund, W.; Aarestrup, K.; Barry, J.; et al. More than one million barriers fragment Europe's rivers. *Nature* **2020**, *588*, 436–441. [[CrossRef](#)]
6. Grigg, N.S. Global water infrastructure: State of the art review. *Int. J. Water Resour. Dev.* **2019**, *35*, 181–205. [[CrossRef](#)]
7. Van Cappellen, P.; Maavara, T. Rivers in the Anthropocene: Global scale modifications of riverine nutrient fluxes by damming. *Ecohydrol. Hydrobiol.* **2016**, *16*, 106–111. [[CrossRef](#)]
8. Winton, R.S.; Calamita, E.; Wehrli, B. Reviews and syntheses: Dams, water quality and tropical reservoir stratification. *Biogeosciences* **2019**, *16*, 1657–1671. [[CrossRef](#)]
9. Poole, G.C.; Berman, C.H. An ecological perspective on in-stream temperature: Natural heat dynamics and mechanisms of human-caused thermal degradation. *Environ. Manage.* **2001**, *27*, 787–802. [[CrossRef](#)]
10. McCartney, M. Living with dams: Managing the environmental impacts. *Water Policy* **2009**, *11*, 121–139. [[CrossRef](#)]

11. Bizzi, S.; Dinh, Q.; Bernardi, D.; Denaro, S.; Schippa, L.; Soncini-Sessa, R. On the control of riverbed incision induced by run-of-river power plant. *Resour. Res.* **2015**, *51*, 5023–5040. [[CrossRef](#)]
12. Poff, N.L.; Allen, J.D.; Bain, M.B.; Karr, J.R.; Prestegard, K.L.; Richter, B.D.; Sparks, R.E.; Stromberg, J.C. The natural flow regime—A paradigm for river conservation and restoration. *Bioscience* **1997**, *47*, 769–784. [[CrossRef](#)]
13. Maavara, T.; Chen, Q.; Van Meter, K.; Brown, L.E.; Zhang, J.; Ni, J.; Zarfl, C. River dam impacts on biogeochemical cycling. *Nat. Rev. Earth Environ.* **2020**, *1*, 103–116. [[CrossRef](#)]
14. Carpenter-Bundhoo, L.; Butler, G.L.; Bond, N.R.; Bunn, S.E.; Reinfelds, I.V.; Kennard, M.J. Effects of a low-head weir on multi-scaled movement and behavior of three riverine fish species. *Sci. Rep.* **2020**, *10*, 6817. [[CrossRef](#)]
15. Dudgeon, D.; Arthington, A.H.; Gessner, M.O.; Kawabata, Z.; Knowler, D.J.; Lévêque, C.; Naiman, R.J.; Prieur-Richard, A.H.; Soto, D.; Stiassny, M.L.; et al. Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biol. Rev. Cambridge Philos. Soc.* **2006**, *81*, 163–182. [[CrossRef](#)]
16. Cardinale, B.J.; Duffy, J.E.; Gonzalez, A.; Hooper, D.U.; Perrings, C.; Venail, P.; Narwani, A.; Mace, G.M.; Tilman, D.; Wardle, D.A.; et al. Biodiversity loss and its impact on humanity. *Nature* **2012**, *486*, 59–67. [[CrossRef](#)]
17. Liermann, C.R.; Nilsson, C.; Robertson, J.; Ng, R.Y. Implications of dam obstruction for global freshwater fish diversity. *Bioscience* **2012**, *62*, 539–548. [[CrossRef](#)]
18. Poff, N.L.; Olden, J.D.; Merritt, D.M.; Pepin, D.M. Homogenization of regional river dynamics by dams and global biodiversity implications. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 5732–5737. [[CrossRef](#)]
19. Poff, N.L.; Richter, B.D.; Arthington, A.H.; Bunn, S.E.; Naiman, R.J.; Kendy, E.; Acreman, M.; Apse, C.; Bledsoe, B.P.; Freeman, M.C.; et al. The ecological limits of hydrologic alteration (ELOHA): A new framework for developing regional environmental flow standards. *Freshw. Biol.* **2010**, *55*, 147–170. [[CrossRef](#)]
20. Rameshkumar, S.; Radhakrishnan, K.; Aanand, S.; Rajaram, R. Influence of physicochemical water quality on aquatic macrophyte diversity in seasonal wetlands. *Appl. Water Sci.* **2019**, *9*, 12. [[CrossRef](#)]
21. O'Hare, M.T.; Aguiar, F.C.; Asaeda, T.; Bakker, E.S.; Chambers, P.A.; Clayton, J.S.; Elger, A.; Ferreira, T.M.; Gross, E.M.; Gunn, I.D.; et al. Plants in aquatic ecosystems: Current trends and future directions. *Hydrobiologia* **2018**, *812*, 1–11. [[CrossRef](#)]
22. Acreman, M. Environmental flows—Basics for novices. *Wiley Interdiscip. Rev. Water* **2016**, *3*, 622–628. [[CrossRef](#)]
23. Trolle, D.; Hamilton, D.P.; Hipsey, M.R.; Bolding, K.; Bruggeman, J.; Mooij, W.M.; Janse, J.H.; Nielsen, A.; Jeppesen, E.; Elliott, J.A.; et al. A community-based framework for aquatic ecosystem models. *Hydrobiologia* **2012**, *683*, 25–34. [[CrossRef](#)]
24. Even, S.; Poulin, M.; Garnier, J.; Billen, G.; Servais, P.; Chesterikoff, A.; Coste, M. River ecosystem modelling: Application of the PROSE model to the Seine river (France). *Hydrobiologia* **1998**, *373*, 27–45. [[CrossRef](#)]
25. Gao, S.; Shen, A.; Jiang, J.; Wang, H.; Yuan, S. Kinetics of phosphate uptake in the dinoflagellate *Karenia mikimotoi* in response to phosphate stress and temperature. *Ecol. Model.* **2022**, *468*, 109909. [[CrossRef](#)]
26. Lee, E.; Jalalizadeh, M.; Zhang, Q. Growth kinetic models for microalgae cultivation: A review. *Algal. Res.* **2015**, *12*, 497–512. [[CrossRef](#)]
27. Van Poorten, B.T.; Walters, C.J.; Taylor, N.G. A field-based bioenergetics model for estimating time-varying food consumption and growth. *Trans. Am. Fish. Soc.* **2012**, *141*, 943–961. [[CrossRef](#)]
28. Stokes, M.F.; Perron, J.T. Modeling the evolution of aquatic organisms in dynamic river basins. *J. Geophys. Res. Earth Surf.* **2020**, *125*, e2020JF005652. [[CrossRef](#)]
29. Albert, J.S.; Schoolmaster, D.R., Jr.; Tagliacollo, V.; Duke-Sylvester, S.M. Barrier displacement on a neutral landscape: Toward a theory of continental biogeography. *Syst. Biol.* **2017**, *66*, 167–182. [[CrossRef](#)]
30. Bicudo, T.C.; Sacek, V.; De Almeida, R.P.; Bates, J.M.; Ribas, C.C. Andean tectonics and mantle dynamics as a pervasive influence on Amazonian ecosystem. *Sci. Rep.* **2019**, *9*, 16879. [[CrossRef](#)]
31. Priyadarshi, A.; Chandra, R.; Kishi, M.J.; Smith, S.L.; Yamazaki, H. Understanding plankton ecosystem dynamics under realistic micro-scale variability requires modeling at least three trophic levels. *Ecol. Model.* **2022**, *467*, 109936. [[CrossRef](#)]
32. Sarker, S.; Al-Noman, M.; Basak, S.C.; Islam, M.M. Do biotic interactions explain zooplankton diversity differences in the Meghna River estuary ecosystems of Bangladesh? *Estuar. Coast. Shelf Sci.* **2018**, *212*, 146–152. [[CrossRef](#)]
33. Larson, D.M.; DeJong, D.; Anteau, M.J.; Fitzpatrick, M.J.; Keith, B.; Schilling, E.G.; Thoele, B. High abundance of a single taxon (amphipods) predicts aquatic macrophyte biodiversity in prairie wetlands. *Biodivers. Conserv.* **2022**, *31*, 1073–1093. [[CrossRef](#)]
34. De Moura, R.S.T.; Henry-Silva, G.G. Food web and ecological models used to assess aquatic ecosystems submitted to aqua-culture activities. *Cienc. Rural.* **2019**, *49*, e20180050. [[CrossRef](#)]
35. Jardim, P.F.; Melo, M.M.M.; Ribeiro, L.D.C.; Collischonn, W.; Paz, A.R.D. A modeling assessment of large-scale hydrologic alteration in south American pantanal due to upstream dam operation. *Front. Environ. Sci.* **2020**, *8*, 567450. [[CrossRef](#)]
36. Vishwakarma, D.K.; Ali, R.; Bhat, S.A.; Elbeltagi, A.; Kushwaha, N.L.; Kumar, R.; Rajput, J.; Heddarn, S.; Kuriqi, A. Pre- and post-dam river water temperature alteration prediction using advanced machine learning models. *Environ. Sci. Pollut. Res.* **2022**, *29*, 83321–83346. [[CrossRef](#)]
37. Zhang, Y.; Xia, J.; Chen, J.; Zhang, M. Water quantity and quality optimization modeling of dams operation based on SWAT in Wenyu River Catchment, China. *Environ. Monit. Assess.* **2011**, *173*, 409–430. [[CrossRef](#)]
38. Campbell, D.E. Proposal for including what is valuable to ecosystems in environmental assessments. *Environ. Sci. Technol.* **2001**, *35*, 2867–2873. [[CrossRef](#)]

39. Garcia, A.; Jorde, K.; Habit, E.; Caamano, D.; Parra, O. Downstream Environmental Effects of Dam Operations: Changes in Habitat Quality for Native Fish Species. *River Res. Appl.* **2011**, *27*, 312–327. [[CrossRef](#)]
40. Marchant, R.; Hehir, G. The use of AUSRIVAS predictive models to assess the response of lotic macroinvertebrates to dams in south-east Australia. *Freshw. Biol.* **2002**, *47*, 1033–1050. [[CrossRef](#)]
41. Chen, S.; Chen, B.; Fath, B.D. Assessing the cumulative environmental impact of hydropower construction on river systems based on energy network model. *Renew. Sust. Energ. Rev.* **2015**, *42*, 78–92. [[CrossRef](#)]
42. Wu, N.; Tang, T.; Fu, X.; Jiang, W.; Li, F.; Zhou, S.; Cai, Q.; Fohrer, N. Impacts of cascade run-of-river dams on benthic diatoms in the Xiangxi River, China. *Aquat. Sci.* **2010**, *72*, 117–125. [[CrossRef](#)]
43. Egger, G.; Politti, E.; Woo, H.; Cho, K.-H.; Park, M.; Cho, H.; Benjankar, R.; Lee, N.-J.; Lee, H. Dynamic vegetation model as a tool for ecological impact assessments of dam operation. *J. Hydro-Environ. Res.* **2012**, *6*, 151–161. [[CrossRef](#)]
44. Zuo, Q.; Han, C.; Liu, J.; Li, J.; Li, W. Quantitative research on the water ecological environment of dam-controlled rivers: Case study of the Shaying River, China. *Hydrol. Sci. J.* **2019**, *64*, 2129–2140. [[CrossRef](#)]
45. Zhao, Q.; Hastie, T. Causal interpretations of black-box models. *J. Bus. Econ. Stat.* **2021**, *39*, 272–281. [[CrossRef](#)]
46. Zeng, Y.; Zeng, Y.; Choi, B.; Wang, L. Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network. *Energy* **2017**, *127*, 381–396. [[CrossRef](#)]
47. Chau, K.W. Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. *J. Hydrol.* **2006**, *329*, 363–367. [[CrossRef](#)]
48. Meng, Z.; Long, X.; Gai, H.; Dang, Z.L.; Wang, S.; Xu, P. Identification of the shear parameters for lunar regolith based on a GA-BP neural network. *J. Terramechanics* **2020**, *89*, 22–29. [[CrossRef](#)]
49. Chen, H.; Zuo, Q.T.; Zhang, Y.Y. Impact factor analysis of aquatic species diversity in the Huai River Basin, China. *Water Supply* **2019**, *19*, 2061–2071. [[CrossRef](#)]
50. Luo, Z.; Zhang, S.; Liu, H.; Wang, L.; Wang, S.; Wang, L. Assessment of multiple dam- and sluice-induced alterations in hydrologic regime and ecological flow. *J. Hydrol.* **2023**, *617*, 128960. [[CrossRef](#)]
51. Luo, Z.; Shao, Q.; Liu, H. Comparative evaluation of river water quality and ecological changes at upstream and downstream sites of dams/sluices in different regulation scenarios. *J. Hydrol.* **2021**, *597*, 126290. [[CrossRef](#)]
52. Chen, H.; Li, W.; Zuo, Q.; Zhang, Y.; Liang, S. Evaluation of aquatic ecological health of sluice-controlled rivers in Huai River Basin (China) using evaluation index system. *Environ. Sci. Pollut. Res.* **2022**, *29*, 65128–65143. [[CrossRef](#)]
53. Luo, Z.; Shao, Q.; Zuo, Q.; Cui, Y. Impact of land use and urbanization on river water quality and ecology in a dam dominated basin. *J. Hydrol.* **2020**, *584*, 124655. [[CrossRef](#)]
54. Meng, W.; Zhang, N.; Zhang, Y. Integrated assessment of river health based on water quality, aquatic life and physical habitat. *J. Environ. Sci.* **2009**, *21*, 1017–1027. [[CrossRef](#)]
55. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press, Urbana-Champaign: Champaign, IL, USA, 1964.
56. McIntosh, R.P. An index of diversity and the relation of certain concepts to diversity. *Ecology* **1967**, *48*, 392–404. [[CrossRef](#)]
57. Ibáñez, C.; Alcaraz, C.; Caiola, N.; Rovira, A.; Trobajo, R.; Alonso, M.; Duran, C.; Jiménez, P.J.; Munné, A.; Prat, N. Regime shift from phytoplankton to macrophyte dominance in a large river: Top-down versus bottom-up effects. *Sci. Total Environ.* **2012**, *416*, 314–322. [[CrossRef](#)]
58. Chen, W.; Olden, J.D. Designing flows to resolve human and environmental water needs in a dam-regulated river. *Nat. Commun.* **2017**, *8*, 2158. [[CrossRef](#)]
59. Alvarez-Garreton, C.; Boisier, J.P.; Billi, M.; Lefort, I.; Marinao, R.; Barría, P. Protecting environmental flows to achieve long-term water security. *J. Environ. Manag.* **2023**, *328*, 116914. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.