

## 基于小波—神经网络耦合模型对云南星云湖富营养化气象驱动因子的分析<sup>\*</sup>

潘佳敏<sup>1</sup>, 冯禹昊<sup>1</sup>, 谢平<sup>2</sup>, 方精云<sup>1\*\*</sup>

(1: 北京大学城市与环境学院, 北京 100871)

(2: 云南大学生态与环境学院高原湖泊生态与治理研究院, 昆明 650091)

**摘要:** 气象因子是影响湖泊富营养化的重要因素, 而湖泊富营养化对人群健康、生态系统和社会经济等均有负面影响. 本文基于统计资料及遥感数据, 结合 Morlet 小波分析和 BP 多层前馈神经网络 (BP 神经网络) 构建了不同时间尺度下的小波—神经网络耦合模型, 分析了 1986—2011 年云南星云湖水华强度变化与月降雨量、月平均气温、月平均风速、月日照时数变化之间的关系, 探究了影响湖泊富营养化的主导气象因子. 结果表明: 气象因子的波动周期是影响湖泊年内水华强度变化的重要因素; 小波—神经网络耦合模型能有效提高数据拟合的精度, 最优小波—神经网络耦合模型的拟合优度为 0.605, 高于 BP 神经网络的拟合优度 0.292; 小波—神经网络耦合模型能更有效地对星云湖富营养化程度进行分析和描述, 其均方误差和相关系数均优于 BP 神经网络; 根据最优小波—神经网络耦合模型下的各气象因子的平均影响值, 可知月平均气温是影响星云湖富营养化的主导气象因子, 其次是月降水率、月平均风速, 最后是月日照时数. 综上, 小波—神经网络耦合模型相比 BP 神经网络对样本数据具有更好的适应性, 拟合精度更高, 能为星云湖的保护与富营养化的治理提供参考依据.

**关键词:** 富营养化; 神经网络; 小波分析; 耦合模型; 气象因子; 星云湖

## Meteorological driving factor of eutrophication in Lake Xingyun, Yunnan based on the Coupling Model of Wavelet analysis and Neural Network<sup>\*</sup>

Pan Jiamin<sup>1</sup>, Feng Yuhao<sup>1</sup>, Xie Ping<sup>2</sup> & Fang Jingyun<sup>1\*\*</sup>

(1: College of Urban and Environmental Sciences, Peking University, Beijing 100871, P.R.China)

(2: Institute for Ecological Research and Pollution Control of Plateau Lakes, School of Ecology and Environmental Science, Yunnan University, Kunming 650091, P.R.China)

**Abstract:** There is a link between changes of climate and the degree of eutrophication of lakes. Eutrophication of lakes has a negative impact on human health, ecosystems, and socioeconomics. Based on statistical data and remote sensing data, coupling models of Morlet wavelet analysis and BP neural network in different time scales were used to analyze the eutrophication trends of Lake Xingyun in Yunnan from 1986 to 2011 and their multi-scale relationships with the climatic variables including monthly precipitation, monthly average temperature, monthly average wind speed, and monthly sunshine duration. The results show that the fluctuation period of climatic factors are important for affecting the monthly variation in the bloom intensity. The coupling models of Morlet wavelet analysis and BP neural network can effectively improve the accuracy of data fitting. Goodness of fit of the optimal coupling model is 0.605 which is higher than the BP neural network's goodness of fit (0.292). The optimal coupling model can analysis and describe eutrophication better than the BP neural network. The mean square error and the correlation coefficients of the optimal coupling model were better than those by the BP neural network. The monthly average temperature was the dominant climatic factor affecting the eutrophication of Lake Xingyun, followed by the monthly precipitation, monthly average wind speed and monthly sunshine duration in the optimal coupling model. In conclusion, we show that the coupling models of Morlet wavelet analysis and BP

<sup>\*</sup> 2020-02-27 收稿; 2020-06-29 收修改稿.

国家自然科学基金项目 (31988102) 资助.

<sup>\*\*</sup> 通信作者; E-mail: jyfang@urban.pku.edu.cn.

neural network, which has better adaptability to periodically changing sample data and higher prediction accuracy than the BP neural network, can provide reliable reference for the protection and eutrophication control of Lake Xingyun.

**Keywords:** Eutrophication; neural network; wavelet analysis; coupling of models; meteorological factor; Lake Xingyun

水体富营养化是全球性的问题<sup>[1-2]</sup>,1980s 以来,全球 68% 的湖泊有富营养化恶化的趋势<sup>[3]</sup>,并且一些地区的湖泊出现有害蓝藻水华<sup>[4]</sup>,给生态环境、人类健康造成极大的负面影响<sup>[5]</sup>,且引起大量的经济损失并带来高昂的治理费用<sup>[6]</sup>.我国云南湖泊富营养化严重,星云湖作为云南湖泊中水质恶化最为迅速的小型湖泊,易受到外界因素的干扰,现已重度富营养化,并常年暴发蓝藻水华<sup>[7]</sup>.为治理星云湖的水体污染,政府在“十二五”期间及“十三五”期间投入资金分别高达 6 亿元和 27 亿元以上<sup>[8-9]</sup>.过去对星云湖水质的研究主要集中在面源污染如地表覆盖变化的影响<sup>[10-13]</sup>,也有研究探究了内源污染沉积物<sup>[14]</sup>、营养盐浓度<sup>[15]</sup>和人类活动<sup>[16]</sup>对星云湖水质的影响.而探究气候因子对星云湖水质影响的研究较少,且仅停留在描述性分析方面<sup>[17]</sup>.气象因子是影响小型湖泊富营养化程度的重要因素,利用现存气象数据探究该湖泊类型富营养化的驱动因素是具有现实意义的.

湖泊富营养化是外界环境因子的综合影响下的复杂结果.由人类活动导致的湖泊中氮、磷等污染物超标是水体富营养化程度加剧的主要驱动因素<sup>[18-19]</sup>,但在营养盐条件充足的条件下,气象因子是环境因子中影响湖泊富营养化程度的主要限制因子.湖泊富营养化程度和气象因子均具有周期型波动的特点,且不同气象因子影响湖泊的富营养程度的机制不同:降雨量的增加会促进磷的迁移加剧富营养化<sup>[20-21]</sup>;气候变暖会改变湖泊中氮、磷浓度,进而改变湖泊中植物群落的组成<sup>[22]</sup>;水温、风速、太阳辐射是影响湖泊浮游植物丰度和分布的重要因素<sup>[23]</sup>等.诸多学者探究了气象因子对湖泊富营养化程度的影响大小:罗晓春等采用随机森林算法计算了导致太湖富营养化发生的各主导气象因子的贡献率<sup>[24]</sup>;Li 等将气象因子与其他环境因子作为解释变量建立多元回归模型,探讨了中国巢湖浮游生物量年内、年际变化<sup>[25]</sup>;类似的研究还在湖北长湖<sup>[26]</sup>、滇池<sup>[27]</sup>、千岛湖<sup>[28]</sup>等湖泊开展.这些研究就湖泊富营养化的发生机制和影响因素进行了全面的探索,并提供了相应的研究范式,但仍存在以下两点待改善之处:其一,对湖泊特定气候条件下导致的富营养化程度的讨论不充分.以上研究更多侧重的是年际变化的趋势,较少根据月度数据进行讨论.且针对降水量、温度、风速和日照时间 4 类主要气象因子进行讨论的研究较少.其二,多采用传统的线性统计分析方法,但变量之间多存在非线性关系,难以描述多因素特别是气象因子对富营养化这个复杂过程的影响,因此寻找新的方法对该过程进行度量是非常必要的.

BP 神经网络(又称多层前馈神经网络或反向传播网络)是目前应用最为广泛的一类人工神经网络,已用于水体的水质评价<sup>[29]</sup>、海岸的藻类水华模拟<sup>[30]</sup>、湖泊叶绿素 *a* 预测<sup>[31]</sup>等方面.小波变换是多分辨率分析方法并且具有良好的局部检测功能,适合表征数据变换的瞬态和奇异点特征,且可以保持频率特性品质因数恒定,常用于噪声祛除、多尺度趋势分析<sup>[32]</sup>.将小波分析和神经网络两种方法进行耦合,可以结合两者优势构建更好的分析模型:一方面是能优化模型的输入数据,即通过小波分析后的数据训练神经网络模型,可剔除或降低遥感数据和气象站点数据本身存在的误差、噪声或异常值的干扰,保留所需尺度下的数据信息,进而降低过拟合风险,增强模型的可信度和泛化能力;另一方面能构建的多个模型进行评判筛选,即将小波分析提供的大量不同时间尺度的数据作为输入数据,构建多个模型,并从中选出合适的时间尺度下的最优模型,能有效提升模型性能.

本文将云南星云湖的水华强度指数作为输出数据,将月降水量、月平均温度、月平均风速、月日照时数 4 个主要气象因子作为输入数据,构建并筛选最优小波—神经网络耦合模型,并将其与 BP 神经网络对比分析.探讨了气象因子对星云湖富营养化程度和变化周期的影响,为星云湖的保护与富营养化的治理提供参考依据.

## 1 资料与方法

### 1.1 数据来源

1.1.1 气象因子数据来源 气象数据来自国家气象科学数据中心(<http://data.cma.cn/>),选取 1986—2011 年

玉溪基本站(区站号:56875)20—20时月降水量、月平均气温、月平均风速和月日照时数4个气象因子作为研究变量.图1为4个气象因子年内变化情况箱式图,“×”表示月均值,可看出星云湖流域各月度气象因子的变化特征,其具有热带季风性、高原山地气候的特点.

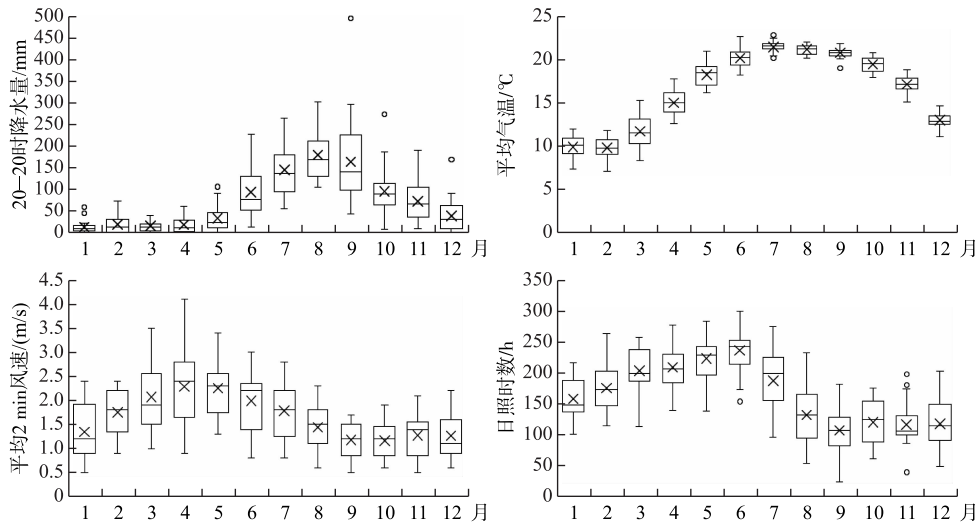


图1 1986—2011年气象因子年内变化情况

Fig.1 Monthly variation of meteorological factors from 1986 to 2011

1.1.2 水华强度计算方法及其处理 本文选取的是美国陆地卫星系列第五颗卫星(Landsat 5)TM成像传感器的大气表观反射率(TOA)产品,Landsat 5的重返周期为16 d,其中6个无量纲大气表观反射波段的空间分辨率均为30 m.整个运算过程在Google Earth Engine(GEE,<https://developers.google.com/earth-engine/>)平台进行,根据星云湖所处位置,我们对1986—2011年间每个月的遥感数据进行搜索、运算、取月平均值、裁剪,最终获得共计211个月的数据.具体的计算过程如下,在根据Landsat 5 TM TOA中BQA波段数据提供的信息进行除云之后,利用H<sub>0</sub>等提出的水华监测算法计算逐月的平均近地面水华强度<sup>[3]</sup>:

$$B = F_c(\text{Band4} - 1.03 \cdot \text{Band5}) \quad (1)$$

式中, $B$ 为水华强度,每个像素点的取值范围介于0~0.1之间;Band4是L5 TM的第4波段大气表观反射率(无量纲);Band5是L5 TM的第5波段大气表观反射率(无量纲).而 $F_c$ 为绿度滤波器,筛选出色调低于阈值 $H$ 的像元,其计算公式如下:

$$F_c = \begin{cases} 1 & (\text{当 } H < 1.6 \text{ 时}) \\ 0 & (\text{当 } H > 1.6 \text{ 时}) \end{cases} \quad (2)$$

判断条件中的阈值 $H$ 的计算方式如下所示:

$$H = \begin{cases} \frac{\text{Band2} - \text{Band1}}{\text{Band3} + \text{Band2} - 2 \cdot \text{Band1}} & (\text{当 } \text{Band1} = \min(\text{Band1}, \text{Band2}, \text{Band3}) \text{ 时}) \\ \frac{\text{Band3} - \text{Band2}}{\text{Band3} + \text{Band1} - 2 \cdot \text{Band2}} + 2 & (\text{当 } \text{Band2} = \min(\text{Band1}, \text{Band2}, \text{Band3}) \text{ 时}) \\ \frac{\text{Band1} - \text{Band3}}{\text{Band2} + \text{Band1} - 2 \cdot \text{Band3}} + 1 & (\text{当 } \text{Band3} = \min(\text{Band1}, \text{Band2}, \text{Band3}) \text{ 时}) \end{cases} \quad (3)$$

式中,Band1是L5 TM的第1波段大气表观反射率(无量纲);Band2是L5 TM的第2波段大气表观反射率(无量纲);Band3是L5 TM的第3波段大气表观反射率(无量纲).

根据上述算法,得到星云湖211幅月均水华强度 $B$ 的遥感图像,其每幅图像的栅格数据情况如图2所示,横坐标表示图像中存在的有效数据的栅格数量与掩膜栅格总数的比值(有效栅格占比),纵坐标表示图像数量,选取有效栅格占比0.8作为图像的筛选阈值,大于0.8的遥感图像是符合条件的,最终有193个月的

有效图像数据. 再分别对每幅月均水华强度  $B$  的图像中所有栅格的数值取平均值, 作为当月星云湖的平均水华强度  $B$ , 各月份数据存在情况如表 1 所示.

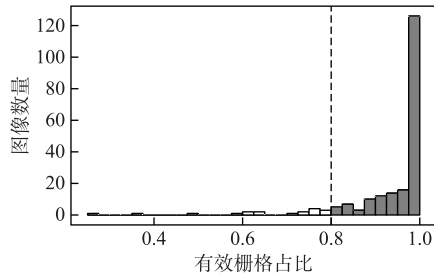


图 2 水华强度  $B$  图像的有效栅格占比频数分布直方图

Fig.2 Frequency distribution histogram of proportion of effective pixels of bloom intensity  $B$  images

表 1 各月份的水华强度  $B$

Tab.1 Summary of bloom intensity  $B$  in each month

	1 月	2 月	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月	总数
有效图像数量	17	18	21	21	23	12	9	11	15	14	18	14	193
$B$ 的平均值( $\times 10^{-2}$ )	1.48	1.78	1.88	1.84	1.55	2.23	3.35	2.57	2.49	1.99	1.50	1.50	1.91
$B$ 的标准差( $\times 10^{-2}$ )	0.64	1.21	1.16	2.20	0.78	1.60	1.65	0.73	1.16	1.15	0.44	0.59	1.27

1.2 小波-神经网络耦合模型基本原理

1.2.1 小波分析 小波分析把时间和频率作为独立变量, 将一维原始信号在时间和频率两个方向上展开, 从而清楚地了解时间序列的不同时域尺度上的频率特征, 实现对时间序列数据的多时间尺度分解. 具体地, 连续小波变换可以定义为<sup>[33]</sup>:

$$W(a, \tau) = \frac{1}{\sqrt{a}} \int x(t) \psi^* \left( \frac{t - \tau}{a} \right) dt \quad (4)$$

式中,  $x(t)$  为原始信号;  $\psi(t)$  为小波母函数;  $a$  为尺度因子;  $\tau$  为位移因子;  $*$  表示共轭; 积分结果  $W(a, \tau)$  为小波系数, 反映了尺度为  $a$ 、位移为  $\tau$  时的小波函数与原始信号间相关性的强弱.

Morlet 小波是一种单频复正弦调制的高斯波, 具有很好的时域和频域局部性, 常用于信号的分解和时频分析中, 且能消除用实小波变换系数作为判据产生的虚假振荡, 其小波母函数的时域形式如下<sup>[33-34]</sup>:

$$\psi(t) = \frac{1}{\sqrt{\pi f_b}} e^{-i^2/f_b} e^{(2\pi f_c t)i} \quad (5)$$

式中,  $f_b$  为小波带宽,  $f_c$  为小波中心频率,  $i$  为虚数符号.

1.2.2 神经网络基本原理 80%~90% 的人工神经网络模型采用的是 BP 神经网络或它的变形. BP 神经网络由一个输入层, 一个或多个隐藏层和一个输出层构成, 每一层由一定数量的神经元构成<sup>[35]</sup>. BP 算法是一种监督式的学习算法, 由两部分组成: 信息的正向传播与误差的反向传播. 在正向传播过程中, 输入信息从输入端经隐藏层逐层计算传向输出层, 每一层神经元的状态只影响下一层神经元的状态. 如果在输出层没有得到期望的输出, 则计算输出层的误差变化值, 然后转向反向传播. 通过网络将误差信号沿原来的连接通路反传回来修改各层神经元的权值直至达到期望目标. 其主要思想是: 对于  $q$  个输入学习样本:  $P_1, P_2, \dots, P_q$ , 已知其对应的输出样本为:  $T_1, T_2, \dots, T_q$ . 学习的目的是用网络的实际输出  $A_1, A_2, \dots, A_q$  与目标矢量  $T_1, T_2, \dots, T_q$  之间的误差来修改其权值, 使  $A_m (m=1, 2, \dots, q)$  与期望的  $T_m$ , 尽可能接近. 即使网络的输出层的误差平方和达到最小. 它是通过连续不断地在相对于误差函数斜率下降方向上计算网络权值和偏差的变化而逐渐逼近目标的. 每一次权值和偏差的变化都与误差的大小呈正比, 并以反向传播的方式传递到每一层<sup>[36]</sup>.

**1.2.3 小波-神经网络耦合模型构建** 整个模型的耦合包括:(1)对气象因子的小波分析;(2)基于小波分析输出数据,构建并筛选多个 BP 神经网络. 图 3 为两种模型的耦合过程图,具体如下:首先利用 Matlab R2018b 软件自带的小波工具包对各气象因子原始信号进行 Morlet 小波分析分解,将接近于且略大于有效数据 193 的数值作为最大时间尺度,其设定为 200 月. 以 1 个月为单位,输出得到 200 个不同时间尺度下各气象因子的小波系数. 对每个时间尺度下,删除仅存在气象因子小波系数但不存在水华强度  $B$  的月份数据,再对剩下的气象因子小波系数进行数据标准化,将标准化后的  $193 \times 4$  的气象因子矩阵的小波系数作为 BP 神经网络的输入数据和对应月份的水华强度  $B$  作为 BP 神经网络的输出数据,其中设定 BP 神经网络随机选取 70% 的数据作为训练集,15% 的数据作为测试集,15% 的数据作为验证集. 本文将 BP 神经网络的隐藏层个数设置为 1 层,而对于隐藏层神经元个数  $N$  的确定,目前尚无理论依据和有效方法,本文的解决方法为对 BP 神经网络分别设定 2~12 个隐藏层神经元个数进行测试,由此得到 200 个不同时间尺度下,具有 2~12 个隐藏层神经元个数的 BP 神经网络. 最后根据各小波-BP 神经网络耦合模型的拟合优度 ( $R^2$ ) 最大的原则筛选出最优小波-BP 神经网络耦合模型. 其中,  $R^2$  的计算公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^{193} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{193} (y_i - \bar{y})^2} \quad (6)$$

式中,  $y_i$  表示真实的观测值,文中为水华强度  $B$  值的大小;  $\bar{y}$  表示真实观测值的平均值,即所有水华强度  $B$  值的平均值;  $\hat{y}$  表示模型的拟合值,文中为神经网络拟合的水华强度  $B$  值大小.

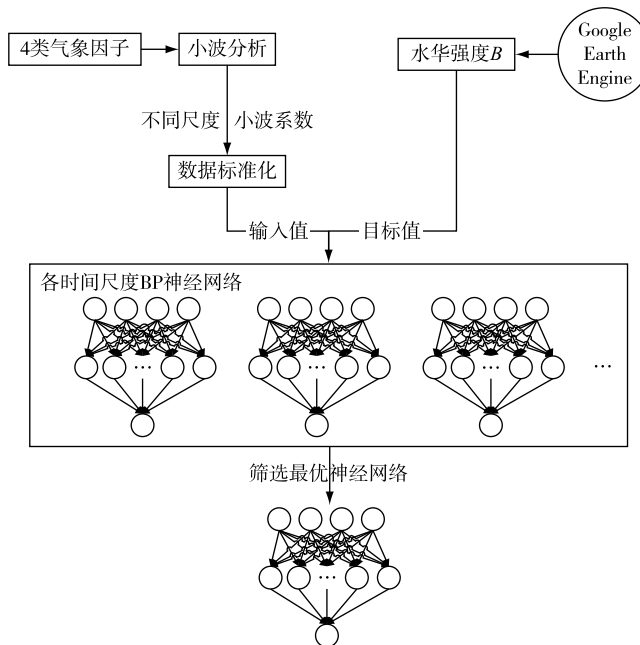


图 3 小波-神经网络耦合模型构建过程

Fig.3 Construction of wavelet-neural network coupling model

## 2 结果与分析

### 2.1 最优小波-神经网络模型的筛选结果

图 4 为在 200 个不同时间尺度下,具有 2~12 个不同隐藏层神经元个数  $N$  的小波-BP 神经网络耦合模型的拟合优度 ( $R^2$ ) 峰峦图. 通过计算模型的  $R^2$  来评价模型对原始数据拟合程度,  $R^2$  值越接近 1,说明模型对

观测值的拟合程度越好;反之, $R^2$ 值越小,说明模型对观测值的拟合程度越差. 根据图 4 所展示的计算结果, $R^2$ 大多分布在 0~0.4 之间. 当隐藏层神经元个数  $N$  为 9,时间尺度为 11 月的情况下,神经网络的  $R^2$  达到最高,为 0.605,此时耦合模型效用最高. 同时也说明通过选取合适的隐藏层神经元个数,可以有效提升小波-神经网络耦合模型性能.

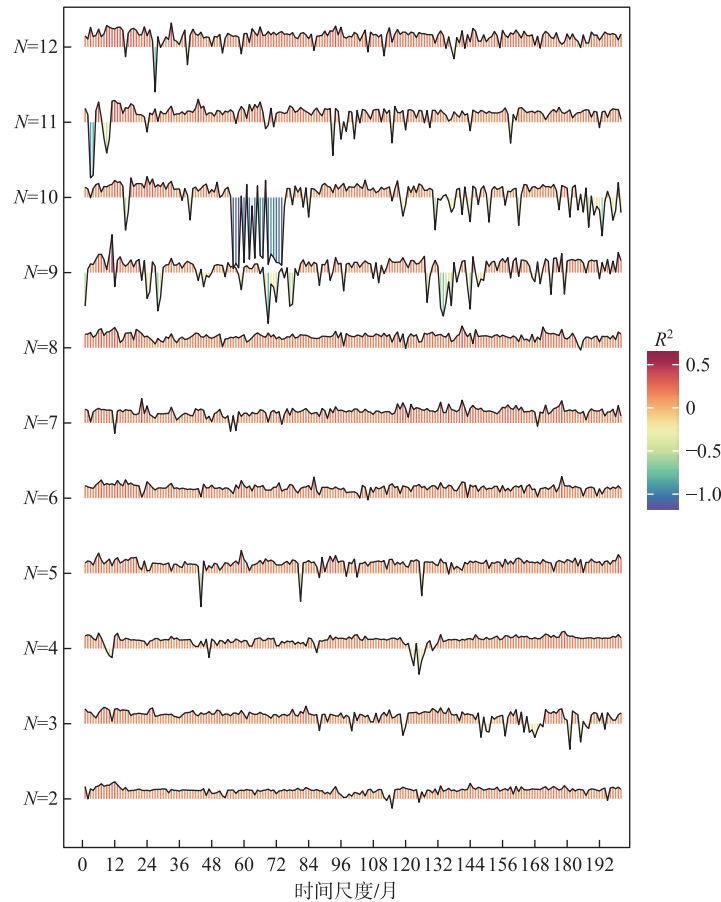


图 4 不同隐藏层神经元个数( $N$ )下 200 个不同尺度小波-神经网络耦合模型的拟合优度( $R^2$ )

Fig.4 Fitting coefficient ( $R^2$ ) of wavelet-neural network coupling model at 200 different scales with different number of hidden layer neurons ( $N$ )

表 2 为设定在不同隐藏层神经元个数( $N=2, 3, \dots, 12$ )的情况下,小波-神经网络耦合模型的最优时间尺度和与其对应的性能参数情况. 隐藏层神经元个数不同时,最优模型对应的时间尺度不同,但大多接近于 12 或者接近于 12 的倍数. 这与输入数据气象因子的波动周期一致,说明最优模型对应的时间尺度受气象因子的波动周期影响,也证实了气象因子的波动周期也是影响水华强度  $B$  变化的重要因素. 模型性能参数包括均方误差 ( $MSE$ ) 及回归系数 ( $R$ ),回归系数中的 Validation  $R$  反映的是模型的泛化性能,Total  $R$  反映的是模型对所有数据的拟合情况,当  $MSE$  越小, $R$  越接近 1 时,模型的效用越高. 由表 2 可知,当  $N$  为 7 时, $MSE$  最小;当  $N$  为 10 时,Validation  $R$  最大;当  $N$  为 9 时,Total  $R$  最大. 综合考虑,本文选择隐藏层神经元数为 9 的小波-神经网络模型作为最优神经网络.

2.2 不同模型的拟合效果及气象因子的平均影响值比较

为了与以上最优小波-神经网络耦合模型进行比较研究耦合效果,本研究另构建了隐藏层为 1,具有不同隐藏层神经元个数( $N=2, 3, \dots, 12$ )的 BP 神经网络,并用标准化后的气象原始数据和水华强度  $B$  分别

表 2 不同隐藏层神经元个数( $N$ )下小波-神经网络耦合模型的最优时间尺度及该尺度下的模型性能参数  
Tab.2 Optimal time scale and model performance parameters of the wavelet-neural network coupling models with different number of hidden layer neurons ( $N$ )

模型性能参数	隐藏层神经元个数( $N$ )										
	2	3	4	5	6	7	8	9	10	11	12
最优时间尺度/月	12	83	179	59	178	22	141	11	24	43	34
$MSE$	0.55	0.74	1.49	1.73	0.89	0.53	0.95	1.85	1.81	1.46	1.21
Validation $R$	0.52	0.07	0.30	0.39	0.36	0.19	0.22	0.55	0.63	0.47	0.20
Total $R$	0.40	0.42	0.40	0.53	0.50	0.54	0.51	0.74	0.49	0.52	0.09

作为输入输出数据训练网络,拟合参数结果如表 3 所示. 并以  $R^2$  最大的原则,最终确定利用隐藏层神经元  $N$  为 9 的 BP 神经网络与最优小波-神经网络耦合模型进行比较.

表 3 不同隐藏层神经元个数下的 BP 神经网络  
Tab.3 BP neural network model fitting parameters with different number of hidden layer neurons

模型拟合参数	隐藏层神经元个数										
	2	3	4	5	6	7	8	9	10	11	12
$R^2$	0.19	0.24	-0.031	0.19	0.19	0.16	0.25	0.29	0.22	0.21	0.28
$MSE$	0.56	0.63	1.68	2.14	0.96	0.27	1.09	2.43	2.81	1.76	1.21
Validation $R$	0.51	0.28	0.03	0.07	0.24	0.63	0.12	0.39	0.20	0.27	0.34
Total $R$	0.28	0.36	0.00	0.31	0.31	0.28	0.39	0.43	0.34	0.31	0.43

图 5 展示的是两种模型数据拟合的具体情况,空心点表示水华强度  $B$  标准化后的值,实线为模型输出值和目标值的拟合直线,虚线为 1:1 线. 结果显示在训练集、测试集、验证集和总体数据中,最优小波-神经网络耦合模型的  $R$  均高于 BP 神经网络模型中的  $R$ ,也就是说在加入小波分析进行模型耦合并筛选后,模型

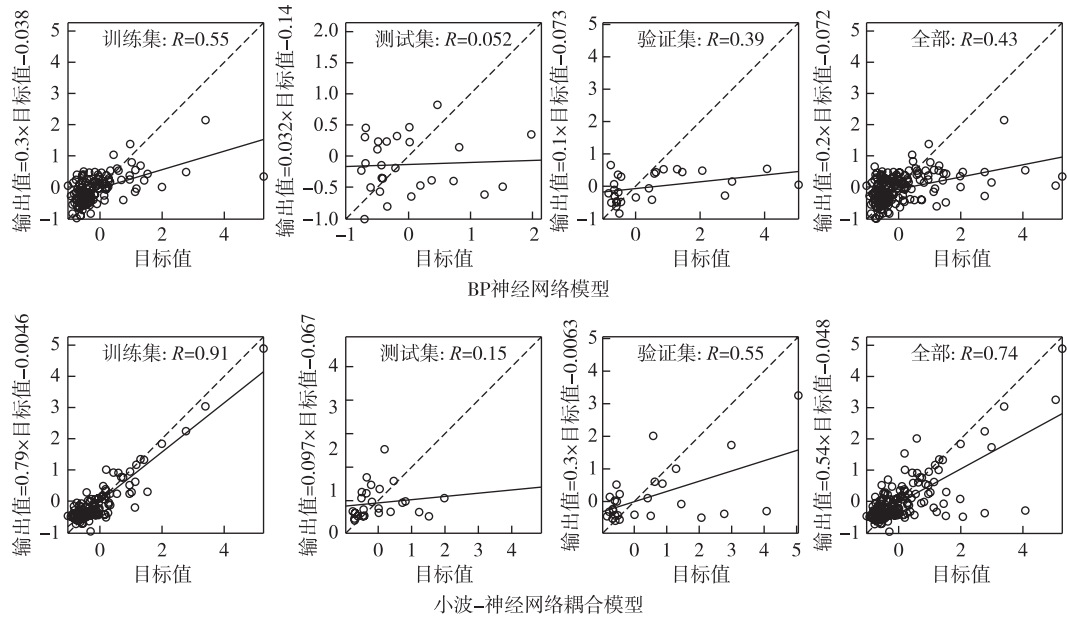


图 5 两类模型数据拟合情况对比  
Fig.5 Comparison of two models in data fitting



对数据拟合效果有明显的提升. 但当水华强度  $B$  处于较高水平时, 所有的拟合线均低于 1:1 线, 说明在水华强度较高的情况下, 两类模型的拟合效果表现较差, 这与缺少高水平水华强度  $B$  的数据有关(表 1).

BP 神经网络的平均影响值法(简称  $MIV$  算法)被认为是在神经网络中进行敏感变量筛选和评价变量重要性的最好方法之一. 在本文中评价各气象因子影响大小的具体过程是: 将训练数据的每个气象因子增加 10% 或者减小 10% 得到的两组新训练数据的自变量, 代入用原始训练数据训练好的 BP 神经网络, 分别得到两组数据的估计结果, 而这两组估计结果的差值被称为  $IV$ , 对结果取平均为平均影响值( $MIV$ ), 依次算出各个自变量的  $MIV$ , 就能确定各个输入变量对网络结果的影响程度<sup>[37]</sup>. 图 6 显示了各气象因子在 BP 神经网络模型和最优小波-神经网络耦合模型中  $MIV$  的大小, 各气象因子的  $MIV$  反映对模型结果的影响程度, 正负值分别代表促进作用和抑制作用. 对于 BP 神经网络模型来说, 影响估计值大小的主要因子是月降水量, 其次是月平均风速和月平均气温抑制作用, 而月日照时数促进作用较小, 这与过去该 4 个气象因子对湖泊富营养化程度起到促进作用的研究结果相差甚远<sup>[20-24]</sup>; 而对于最优小波-神经网络耦合模型来说, 4 个气象因子均为促进作用, 月平均气温对估计值的影响最大, 其次是月降水量和月平均风速, 月日照时数影响最小. 对比两种模型各气象因子的  $MIV$  大小可知, 最优小波-神经网络耦合模型相比 BP 神经网络对气候因子的变化更为敏感. 且根据以往的研究结果, 在最优小波-神经网络耦合模型中, 各气象因子对湖泊富营养化的发生机制的解释更合理.

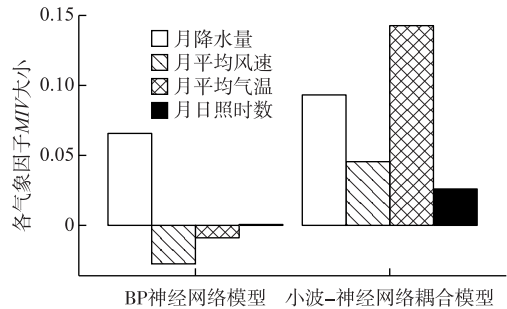


图 6 两种模型各气象因子平均影响值  $MIV$  大小

Fig.6 Mean Impact Value ( $MIV$ ) of each meteorological factor of the two models

### 3 讨论

1980s 以来, 云南星云湖流域工农业发展迅速, 工业废水、生活污水的排放导致星云湖的生态环境逐步恶化、富营养程度加深<sup>[38]</sup>, 甚至暴发有害蓝藻水华<sup>[15]</sup>. 很多研究讨论了面源污染、内源污染、人类活动和营养盐对星云湖富营养化程度的影响<sup>[10-16]</sup>, 但尚不多见探究气象因子对星云湖富营养化程度影响的相关研究. 而诸多文献表明气象因子是影响湖泊富营养化程度的关键因素<sup>[20-23]</sup>. 罗晓春等的分析结果表明气温是影响太湖蓝藻水华综合指数的主导因子, 其次是风速、降水, 最后是日照时间<sup>[24]</sup>, 这与本文的研究结果基本一致. Brias 等研究表明年内降水的变化是导致湖泊富营养化程度加剧的主要驱动因素<sup>[21]</sup>, 而本文进一步补充发现气象因子的波动周期也是影响年内水华强度变化的重要因素.

本文根据气象站点月度气象因子及遥感图像数据所算出的水华指数, 构建并筛选合适的模型用于描述星云湖富营养化程度变化并探讨各气象因子对富营养化程度的影响大小是极具意义的. 郭志海等发现星云湖水水质状况与植被指数  $NDVI$  具有关联性<sup>[39]</sup>, 这进一步证实了遥感数据在星云湖应用的有效性. 由于遥感数据和气象站点的数据不可避免出现误差和异常值, 直接利用这些数据进行拟合分析会使构建的机器学习模型产生精度方面的问题. 于家斌等分析表明对叶绿素  $a$  浓度数据平滑后再利用神经网络能更好地预测蓝藻水华的暴发<sup>[40]</sup>, 这与本文所采取方法的本质相似. 小波分析方法本身能对周期性变化的数据进行分析, 本文发现其能降低数据异常和数据误差带来的影响, 将小波分析与 BP 神经网络耦合, 能有效提升 BP 神经网络对数据拟合效果和模型的泛化效果. 本文最终将隐藏层神经元数为 9、时间尺度为 11 月的小波-神经网络耦合模型作为最优模型, 并将此与无耦合 BP 神经网络进行比较, 发现前者对于星云湖富营养化程度的拟合更加精确, 也验证了在小型湖泊用气象数据变化拟合富营养化程度变化的可行性. 本文显示了小波-神经网络耦合模型的优越性, 该研究方法可以延申至有关于富营养化其他类似小型湖泊的研究.

本文仍然存在一定局限性, 首先, 样本量数据不大, 尤其是高水华强度值的样本量较少, 导致模型对高水华强度值的拟合效果低, 进而影响小波-神经网络耦合模型整体的拟合优度. 其次, 气象因子只能解释部分星云湖富营养化程度变化的波动情况, 并没有将营养盐浓度、面源污染、内源污染等人为因素和其他因子



纳入模型,难以精确拟合实际的湖泊富营养化程度. 这些不足可在今后的研究中进行补充或探讨.

## 4 结论

1)在对最优小波—神经网络模型进行筛选的过程中,不同隐藏层神经元下的最优小波时间尺度与气象因子的波动周期或波动周期的倍数一致或者接近,证实了气象因子的波动周期是影响年内水华强度变化的重要因素.

2)将最优小波—神经网络耦合模型与无耦合 BP 神经网络模型进行比较发现,小波—神经网络耦合模型能有效提高数据拟合的精度. 最优小波—神经网络耦合模型的拟合优度为 0.605,高于 BP 神经网络的拟合优度 0.292. 且小波—神经网络耦合模型的均方误差  $MSE$  和相关系数  $R$  均优于 BP 神经网络,能更有效地对星云湖富营养化程度进行分析和描述.

3)分别用最优小波—神经网络耦合模型和无耦合 BP 神经网络模型计算各气象因子的平均影响值,发现最优小波—神经网络耦合模型的结果更加符合湖泊富营养化发生的机制. 根据最优小波—神经网络耦合模型的气象因子的平均影响值,得到月平均气温是影响星云湖富营养化的主导气象因子,其次是月降水量、月平均风速,最后是月日照时数.

## 5 参考文献

- [1] Fahrenkamp-Uppenbrink J. Watch out for river nutrient imbalances. *Science*, 2019, **365**(6454): 652-654.
- [2] Jeppesen E, Søndergaard M, Liu ZW. Lake restoration and management in a climate change perspective: An introduction. *Water*, 2017, **9**(2): 122. DOI: 10.3390/w9020122.
- [3] Ho JC, Michalak AM, Pahlevan N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature*, 2019, **574**(7780): 667-670. DOI: 10.1038/s41586-019-1648-7.
- [4] Clark JM, Schaeffer BA, Darling JA *et al.* Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. *Ecological Indicators*, 2017, **80**: 84-95. DOI: 10.1016/j.ecolind.2017.04.046.
- [5] Su JQ, Wang X, Yang ZF. Lake eutrophication modeling in considering climatic factors change: A review. *Chinese Journal of Applied Ecology*, 2012, **23**(11): 3197-3206. [苏洁琼, 王烜, 杨志峰. 考虑气候因子变化的湖泊富营养化模型研究进展. 应用生态学报, 2012, **23**(11): 3197-3206.]
- [6] Carpenter SR, Bolgrien D, Lathrop RC *et al.* Ecological and economic analysis of lake eutrophication by nonpoint pollution. *Austral Ecology*, 1998, **23**(1): 68-79. DOI: 10.1111/j.1442-9993.1998.tb00706.x.
- [7] Chen XF, Chuai XM, Yang LY. Status quo, historical evolution and causes of eutrophication in lakes in typical lake regions of China. *Rural Eco-Environment*, 2014, **30**(4): 438-443. [陈小锋, 揣小明, 杨柳燕. 中国典型湖区湖泊富营养化现状、历史演变趋势及成因分析. 生态与农村环境学报, 2014, **30**(4): 438-443.]
- [8] Yuxi Environmental Protection Bureau ed. '13th Five-Year Plan' for Water Environmental Protection in the Xingyun Lake Basin (2016-2020), 2016. [玉溪市环保局. 星云湖流域水环境保护治理“十三五”规划(2016—2020年), 2016.]
- [9] Yuxi Environmental Protection Bureau ed. '12th Five-Year Plan' for Water Environmental Protection in the Xingyun Lake Basin, 2011. [玉溪市环保局. 星云湖流域水污染综合防治“十二五”规划, 2011.]
- [10] Yi D, Gan S, Yu ZR. Correlation analysis of surface cover changes and lake environmental conditions in Lake Xingyu. *Journal of Zhejiang Agricultural Sciences*, 2019, **60**(3): 459-461, 527. [易达, 甘淑, 余章蓉. 星云湖流域地表覆被变化与湖泊环境状况关联分析. 浙江农业科学, 2019, **60**(3): 459-461, 527.]
- [11] Hou J, Zhao XH. Agricultural non-point source pollution prevention and countermeasure research of plateau lakes in Yunnan Province. *Yunnan Environmental Science*, 2018, **37**(6): 63-65. [侯娟, 赵祥华. 云南省高原湖泊农业面源污染防治和对策研究. 环境科学导刊, 2018, **37**(6): 63-65.]
- [12] Zheng TT, Zhao YQ, Lu FF *et al.* Analysis on the driving force of farming non-point source pollution in the Xingyun Lake Basin of Yunnan Province. *Journal of Ecology and Rural Environment*, 2019, **35**(6): 730-737. [郑田甜, 赵筱青, 卢飞飞等. 云南星云湖流域种植业面源污染驱动力分析. 生态与农村环境学报, 2019, **35**(6): 730-737.]
- [13] Zhao ZJ, Zheng TT, Zhao XQ *et al.* Non-point source pollutants loss of planting industry in the Yunnan plateau lake basin, China. *Agro-Environment and Development*, 2018, (1): 40-47. [赵祖军, 郑田甜, 赵筱青等. 云南高原湖泊流域种植

- 业面源污染物的流失特征分析. 农业资源与环境学报, 2018, (1): 40-47.]
- [14] Gao CL, Yu JQ, Min XY *et al.* Heavy metal concentrations in sediments from Xingyun Lake, southwestern China; Implications for environmental changes and human activities. *Environmental Earth Sciences*, 2018, **77**(19): 666. DOI: 10.1007/s12665-018-7840-5.
- [15] Huang SF, Liu W, Wang XT *et al.* Neural network modeling of the eutrophication and strategy of pollution control in Lake Xingyun. *Journal of South China Normal University: Natural Science Edition*, 2013, **45**(5): 93-97. [黄少峰, 刘威, 王旭涛等. 星云湖富营养化进程的神经网络模拟及污染控制对策. 华南师范大学学报: 自然科学版, 2013, **45**(5): 93-97.]
- [16] Zheng TT, Zhao ZJ, Zhao XQ *et al.* Water quality change and humanities driving force in Lake Xingyun, Yunnan Province. *J Lake Sci*, 2018, **30**(1): 79-90. DOI: 10.18307/2018.0108. [郑田甜, 赵祖军, 赵筱青等. 云南星云湖水水质变化及其人文因素驱动力分析. 湖泊科学, 2018, **30**(1): 79-90.]
- [17] Zhang YY, Xie ML, Liu SS *et al.* The analysis of water quality change trend for Lake Fuxian & Lake Xingyun. *Geographical Science Research*, 2015, **4**(3): 81-94. [张云鹰, 谢曼平, 刘姗姗等. 抚仙湖和星云湖水水质变化趋势分析. 地理科学研究, 2015, **4**(3): 81-94.]
- [18] Zhang HL, Li SJ, Feng QL *et al.* Environmental change and human activities during the 20th century reconstructed from the sediment of Xingyun Lake, Yunnan Province, China. *Quaternary International*, 2010, **212**(1): 14-20.
- [19] Sinha E, Michalak AM, Calvin KV *et al.* Societal decisions about climate mitigation will have dramatic impacts on eutrophication in the 21st century. *Nature Communications*, 2019, **10**: 939. DOI: 10.1038/s41467-019-08884-w.
- [20] Sinha E, Michalak AM, Balaji V. Eutrophication will increase during the 21st century as a result of precipitation changes. *Science*, 2017, **357**(6349): 405-408. DOI: 10.1126/science.aan2409.
- [21] Brias A, Mathias JD, Deffuant G. Inter-annual rainfall variability may foster lake regime shifts; An example from Lake Bourget in France. *Ecological Modelling*, 2018, **389**: 11-18. DOI: 10.1016/j.ecolmodel.2018.10.004.
- [22] Feuchtmayr H, Moran R, Hatton K *et al.* Global warming and eutrophication; Effects on water chemistry and autotrophic communities in experimental hypertrophic shallow lake mesocosms. *Journal of Applied Ecology*, 2009, **46**(3): 713-723. DOI: 10.1111/j.1365-2664.2009.01644.x.
- [23] Guo CX, Zhu GW, Qin BQ *et al.* Climate exerts a greater modulating effect on the phytoplankton community after 2007 in eutrophic Lake Taihu, China: Evidence from 25 years of recordings. *Ecological Indicators*, 2019, **105**: 82-91. DOI: 10.1016/j.ecolind.2019.05.034.
- [24] Luo XC, Hang X, Cao Y *et al.* Dominant meteorological factors affecting cyanobacterial blooms under eutrophication in Lake Taihu. *J Lake Sci*, 2019, **31**(5): 1248-1258. DOI: 10.18307/2019.0512. [罗晓春, 杭鑫, 曹云等. 太湖富营养化条件下影响蓝藻水华的主导气象因子. 湖泊科学, 2019, **31**(5): 1248-1258.]
- [25] Li J, Ma RH, Xue K *et al.* Drivers to spatial and temporal dynamics of column integrated phytoplankton biomass in the shallow lake of Chaohu, China. *Ecological Indicators*, 2020, **109**: 105812. DOI: 10.1016/j.ecolind.2019.105812.
- [26] Liu T, Xia ZH, Zhu L. Influence of meteorological factors on temporal and spatial variation of water quality in Changhu Lake, Hubei Province. *Advances in Meteorological Science and Technology*, 2018, **8**(5): 78-80. [刘韬, 夏智宏, 朱浪. 气象条件对湖北长湖水质的影响研究. 气象科技进展, 2018, **8**(5): 78-80.]
- [27] Wang JH, He LQS, Yang C *et al.* Comparison of algal bloom related meteorological and water quality factors and algal bloom conditions among Lakes Taihu, Chaohu, and Dianchi (1981-2015). *J Lake Sci*, 2018, **30**(4): 897-906. DOI: 10.18307/2018.0403. [王菁晗, 何吕奇姝, 杨成等. 太湖、巢湖、滇池水华与相关气象、水质因子及其响应的比较 (1981-2015 年). 湖泊科学, 2018, **30**(4): 897-906.]
- [28] Da WY. Study on spatio-temporal dynamics and influencing factors of aquatic environment in Lake Qiandaohu [Dissertation]. Nanchong: China West Normal University, 2019. [笱文怡. 千岛湖水环境时空变化特征及其影响因素[学位论文]. 南充: 西华师范大学, 2019.]
- [29] Zhao L, Dai TJ, Chen L *et al.* Prediction of water quality in Taolinkou Reservoir based on BP Neural Network Model. *Journal of Anhui Agricultural Sciences*, 2019, **47**(24): 73-77, 81. [赵林, 戴天骄, 陈亮等. 基于 BP 神经网络的桃林口水库水质预测. 安徽农业科学, 2019, **47**(24): 73-77, 81.]
- [30] Lee JHW, Huang Y, Dickman M *et al.* Neural network modelling of coastal algal blooms. *Ecological Modelling*, 2003, **159**(2/3): 179-201. DOI: 10.1016/s0304-3800(02)00281-8.

- [31] Pei HP, Luo NN, Jiang Y. Applications of back propagation neural network for predicting the concentration of chlorophyll-a in West Lake. *Acta Ecologica Sinica*, 2004, **24**(2): 246-251. [裴洪平, 罗妮娜, 蒋勇. 利用 BP 神经网络方法预测西湖叶绿素 a 的浓度. *生态学报*, 2004, **24**(2): 246-251.]
- [32] Heng T. Wavelet analysis and its application [Dissertation]. Chengdu: Sichuan University, 2003. [衡彤. 小波分析及其应用研究[学位论文]. 成都: 四川大学, 2003.]
- [33] Feng YH, Zhu JL. Analysis on runoff change and the driving force of the Liaohe River basin based on morlet wavelet. *Research of Soil and Water Conservation*, 2019, **26**(2): 208-215. [冯禹昊, 朱江玲. 基于 Morlet 小波的辽河径流量变化及其成因分析. *水土保持研究*, 2019, **26**(2): 208-215.]
- [34] Qi SJ, Chen HR. Multiple time scale analysis on yearly rainfall with morlet wavelet analysis. *South-to-North Water Transfers and Water Science & Technology*, 2010, **8**(3): 79-82. [祁顺杰, 陈皓锐. Morlet 小波在降雨的多时间尺度分析中的应用. *南水北调与水利科技*, 2010, **8**(3): 79-82.]
- [35] Fan ZY. BP Neural Network Model and Learning Algorithm. *Software Guide*, 2011, **10**(7): 66-68. [樊振宇. BP 神经网络模型与学习算法. *软件导刊*, 2011, **10**(7): 66-68.]
- [36] Liu TS. The research and application on bp nerual network improvement[Dissertation]. Harbin: Northeast Agricultural University, 2011. [刘天舒. BP 神经网络的改进研究及应用[学位论文]. 哈尔滨: 东北农业大学, 2011.]
- [37] Zhou Y. Research on Ball Bearing Fault Diagnosis Technology based on MIV Algorithm Selection and BP Neural Network [Dissertation]. Beijing: Beijing Jiaotong University, 2011. [周莹. 基于 MIV 特征筛选和 BP 神经网络的滚动轴承故障诊断技术研究[学位论文]. 北京: 北京交通大学, 2011.]
- [38] Zhang LY, Zhang HL, Zhang ST *et al.* Environmental evolution in the past 200 years of Xingyun Lake, central Yunnan Province, southwestern China. *Geological Science and Technology Information*, 2006, **25**(3): 8-12. [张立原, 张宏亮, 张世涛等. 滇中星云湖地区 200 年来的环境演化. *地质科技情报*, 2006, **25**(3): 8-12.]
- [39] Guo ZH, Yuan XP, Gan S *et al.* Relationship between water quality and vegetation index in Lake Xingyun. *Journal of Zhejiang Agricultural Sciences*, 2017, **58**(3): 499-503. [郭志海, 袁希平, 甘淑等. 星云湖水质状况与植被指数的关联研究. *浙江农业科学*, 2017, **58**(3): 499-503.]
- [40] Yu JB, Shang FF, Wang XY *et al.* Cyanobacterial bloom forecast method based on GF-LSTM network. *Journal of Computer Applications*, 2020, (1): 1-6. [于家斌, 尚方方, 王小艺等. 基于 GF-LSTM 网络的蓝藻水华预测方法. *计算机应用*, 2020, (1): 1-6.]