

引文格式:徐定建.一种基于主成分分析改进的 PSO-BP 神经网络赤潮预测模型[J].测绘通报,2021(S2):234-240. DOI:10.13474/j.cnki.11-2246.2021.0628.

# 一种基于主成分分析改进的 PSO-BP 神经网络赤潮预测模型

徐定建

(重庆市勘测院,重庆 401121)

**摘要:**针对赤潮爆发受多因素影响且影响因素间存在相关性的特性,提出了一种基于主成分分析改进的 PSO-BP 神经网络预测模型(PCA-PSO-BP)。在充分利用原始观测信息的前提下,采用主成分分析消除变量间的相关性,从而减少网络输入节点数,简化网络模型,同时结合粒子群优化算法初始化网络初始权值和阈值,建立高精度 PCA-PSO-BP 神经网络预测模型。结合赤潮监测实例数据,分别与传统 BP 神经网络、PCA-BP 神经网络、PSO-BP 神经网络预测模型进行对比,结果表明采用 PCA-PSO-BP 神经网络预测赤潮具有一定的可行性,可以提高预测模型精度。

**关键词:**赤潮爆发;主成分分析;粒子群优化算法;神经网络;预测

中图分类号:P237

文献标识码:A

文章编号:0494-0911(2021)S2-0234-07

## An improved PSO-BP neural network red tide prediction model based on principal component analysis

XU Dingjian

(Chongqing Survey Institute, Chongqing 401121, China)

**Abstract:** Aiming at the characteristics that red tide outbreak was affected by many factors and there was correlation among them, PSO-BP prediction model based on principal component analysis (PCA-PSO-BP) is proposed in this paper. Under the premise of making full use of the original information, the principal component analysis is used to eliminate the correlation between variables. Then the network model is simplified by reducing the number of input nodes. Finally, the initial weights and thresholds of the network are initialized by particle swarm optimization algorithm to form a high-precision PCA-PSO-BP neural network prediction model. Combining with the actual data of red tide monitoring, the prediction models of BP neural network, PCA-BP neural network and PSO-BP neural network are compared. The results showed that the prediction model of red tide using PCA-PSO-BP neural network is feasible, which improves the accuracy of prediction model.

**Key words:** red tide outbreak; principal component analysis; particle swarm optimization algorithm; neural network; prediction

近年来,赤潮灾害爆发频率越来越高<sup>[1]</sup>,严重威胁沿海地区的旅游业、浅海养殖业发展。赤潮爆发将造成巨大的经济损失,因此对沿海地区赤潮监测预测十分重要。目前学者对赤潮预测主要从分类预测和数值预测两个方面开展研究。文献[2]提出一种分组属性加权聚类的 CSFCM(Clustering Object on Subsets of Attributes algorithm and the fuzzy C-means algorithm, CSFCM)算法定性评价赤潮灾害,缺乏精确的预测模型。文献[3—4]提出不同形式的神经网络预测模型,将透明度、pH 值、水温、盐度、化学需氧量(Chemical Oxygen Demand, COD)等因素<sup>[5]</sup>作为网络输入,实现赤潮数值预报,然而由于影响因素间存在相关性,会造成网络模型参数冗余,影响模型预测精度。

本文基于 BP(Back Propagation, BP)神经网络

数值预测模型,采用粒子群优化(Particle Swarm Optimization, PSO)算法克服传统 BP 神经网络收敛速度慢、易陷入局部最小值<sup>[6-7]</sup>的缺陷,同时顾及模型输入变量间的相关性,在保留有效原始数据信息的前提下采用主成分分析(Principal Component Analysis, PCA)进行数据降维,从而消除变量间相关性干扰,进一步提高网络模型预测精度。本文结合部分赤潮监测数据,分别与传统 BP 神经网络、PCA-BP 神经网络、PSO-BP 神经网络预测模型的预测结果进行对比,验证了本文算法的有效性和可行性。

### 1 主成分分析原理

主成分分析是一种研究多个具有相关性变量的多元统计方法<sup>[8-9]</sup>,在研究复杂系统问题时充分利用原始观测信息,通过数据降维操作将多个相关变

作者简介:徐定建(1985—),男,中级工程师,研究方向为变形监测分析与预报。E-mail:406828533@qq.com

量简化成互不相关的主成分变量,可以有效地解决信息冗余问题,提高算法效率。主要包括数据标准化、确定主成分、计算主成分方差贡献率和累计方差贡献率、选取主成分等过程。

(1)数据标准化:在某一特定复杂系统中,由  $m$  个变量  $[X_1 X_2 \cdots X_m]$  影响的  $n$  个样本数据组成原始观测信息矩阵

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} = [X_1 X_2 \cdots X_m] \quad (1)$$

考虑各变量间采用标准的差异,方差较大的变量会对主成分分析造成较大的干扰,需标准化原始观测数据,计算公式如下

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_{x_j}} \quad (2)$$

式中,  $\bar{x}_j, \sigma_{x_j}$  分别为  $X_j$  的均值和标准差,标准化观测信息矩阵为  $X' = (x'_{ij})_{n \times m}$ 。

(2)主成分确定:标准化的观测信息矩阵  $X'$  包含了全部原始观测信息,对变量线性组合即可获得  $m$  个互不相关新变量,即  $m$  个主成分,公式为

$$\left. \begin{aligned} y_1 &= c_{11}X_1 + c_{12}X_2 + \cdots + c_{1m}X_m \\ y_2 &= c_{21}X_1 + c_{22}X_2 + \cdots + c_{2m}X_m \\ &\vdots \\ y_m &= c_{m1}X_1 + c_{m2}X_2 + \cdots + c_{mm}X_m \end{aligned} \right\} \quad (3)$$

式中,  $y_1, y_2, \cdots, y_m$  分别是原变量  $X_1, X_2, \cdots, X_m$  的第 1、第 2、……、第  $m$  主成分,  $c_{ij} (i = 1, 2, \cdots, m; j = 1, 2, \cdots, m)$  是原有变量  $X_i (i = 1, 2, \cdots, m)$  位于主成分  $y_i (i = 1, 2, \cdots, m)$  上的荷载,也是相关系数矩阵各特征根对应的特征向量,且满足  $c_{i1}^2 + c_{i2}^2 + \cdots + c_{im}^2 = 1$ 。

(3)主成分方差贡献率和累计方差贡献率计算:相关系数矩阵的特征根等于对应主成分的方差<sup>[10]</sup>,反映了对应主成分包含信息量在原始观测信息中的占比。本文依据特征根的大小计算主成分方差贡献率和累计方差贡献率,见式(4)和式(5)

$$\lambda_i / \sum_{i=1}^m \lambda_i \quad (4)$$

$$\sum_{i=1}^k \lambda_i / \sum_{i=1}^m \lambda_i \quad (5)$$

式中,  $\lambda_i (i = 1, 2, \cdots, m)$  是相关系数矩阵的特征根。由式(3)可以看出,变量综合信息的表达能力与方差贡献率成正比,结合式(4),如果前  $i$  个主成

分的累计方差贡献率越大,前  $i$  个主成分包含原始观测的信息量就越大。

(4)主成分选取:本文采取主成分分析的主要目的是充分利用原始观测信息,通过降低数据维度解决网络参数冗余问题,因此在保证累计方差贡献率的前提下提取主成分十分必要。一般情况下,选取的主成分个数不超过 5 个或 6 个,累计方差贡献率不低于 80%<sup>[11]</sup>。

## 2 PCA-PSO-BP 神经网络模型

### 2.1 BP 神经网络基本原理

BP 神经网络是一种信号前向传递,误差反向传播的多层前馈网络,包括输入层、隐含层和输出层<sup>[12-13]</sup>。因其误差函数梯度下降的特性,可不断调整隐含层的连接权值来优化学习网络结构。BP 神经网络因其自学习、自适应的特性适用于非线性预测<sup>[14]</sup>,在数据预测领域应用十分广泛。

### 2.2 粒子群算法原理

粒子群优化算法是一种群体智能优化算法<sup>[15]</sup>,每个粒子表示算法中不同的适应度值,代表了优化问题的一个可行解,通过动态调整自身和其他粒子的速度在解空间中寻找最优解<sup>[16]</sup>。对于  $D$  维解空间,包含  $n$  个粒子的种群  $Y = (Y_1, Y_2, \cdots, Y_n)$ ,第  $i$  个粒子的位置为  $X_i = [x_{i1} x_{i2} \cdots x_{iD}]^T$ ,最小化每个粒子在位置  $X_i$  对应的适应度值,从而确定最优粒子。第  $i$  个粒子速度为  $V_i = [V_{i1} V_{i2} \cdots V_{iD}]^T$ ,个体极值为  $P_i = [P_{i1} P_{i2} \cdots P_{iD}]^T$ ,种群的群体极值为  $P_g = [P_{g1} P_{g2} \cdots P_{gD}]^T$ 。粒子每次迭代调整自身速度和位置见式(6)和式(7)

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k) \quad (6)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (7)$$

式中,  $d = 1, 2, \cdots, D; i = 1, 2, \cdots, n; \omega$  为惯性权重;  $k$  为当前迭代次数;  $c_1, c_2$  分别为调整粒子走向自身最优位置和群体最优位置步长,称为加速因子;  $r_1, r_2$  为  $[0, 1]$  的随机数。一般将粒子的位置和速度分别限制在区间  $[-X_{\max}, X_{\max}]$ 、 $[0, V_{\max}]$ ,以加快粒子的搜索速度。

### 2.3 PSO-BP 神经网络模型

由于 BP 神经网络对初始权值和学习速率等网络参数十分敏感,易陷入局部最优解;PSO 算法基于群体智能优化,动态调整粒子的速度和位置寻找全局最优化解,具有较强的全局搜索能力<sup>[17]</sup>,因此本文采用 PSO 算法初始化网络连接权值和阈值,进

而优化 BP 神经网络结构。主要过程如下:

(1) BP 神经网络构建:网络隐含层神经元节点数依据经验公式<sup>[18]</sup>如下

$$\sqrt{n+l}+a \quad (8)$$

式中,  $n$  为输入层节点数;  $l$  为输出层节点数;  $a$  是区间  $[1, 10]$  的整数。利用 PSO 算法初始化网络权值和阈值。

(2) 粒子群初始化:粒子群数量、学习因子、惯性权重等参数可设定为经验值,粒子维数由输入层、隐含层和输出层节点数确定,公式为

$$H = h_i + h_i \times h_j + h_j \times h_k + h_k \quad (9)$$

式中,  $h_i$ 、 $h_j$ 、 $h_k$  分别为输入层、隐含层和输出层节点数。

(3) 适应度函数确定:适应度函数值是表征粒子好坏的重要参数,可通过最小化适应度函数值达到目标函数最小的目的,本文采取 BP 神经网络的均方误差作为适应度函数,公式为

$$\text{Fitness} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M (Y_{ij} - T_{ij})^2 \quad (10)$$

式中,  $N$  为输入样本总数;  $M$  为粒子维数;  $Y_{ij}$  为第  $i$  个样本实际输出;  $T_{ij}$  为第  $i$  个样本期望输出。由式 (10) 可知,适应度函数值越小的网络误差越小,粒子适应性越好。

(4) 最优粒子确定:通过动态调整粒子的速度和位置,最小化适应度函数即可确定最优粒子。

## 2.4 PCA-PSO-BP 神经网络模型构建

虽然粒子群算法能够优化 BP 神经网络寻求全局最优解,但当网络输入参数较多并且参数之间存在很强的相关性时,如果直接采用 PSO-BP 神经网络模型进行预测,会导致网络输入过于复杂,出现过拟合的现象,而且有可能引入模型干扰导致模型精度不高。采用 PCA 算法可以有效地处理变量间的

相关性,在保留足够原始观测信息的前提下,提取主要因素,降低数据维度,从而减少网络输入,达到简化网络的目的。具体过程如图 1 所示。

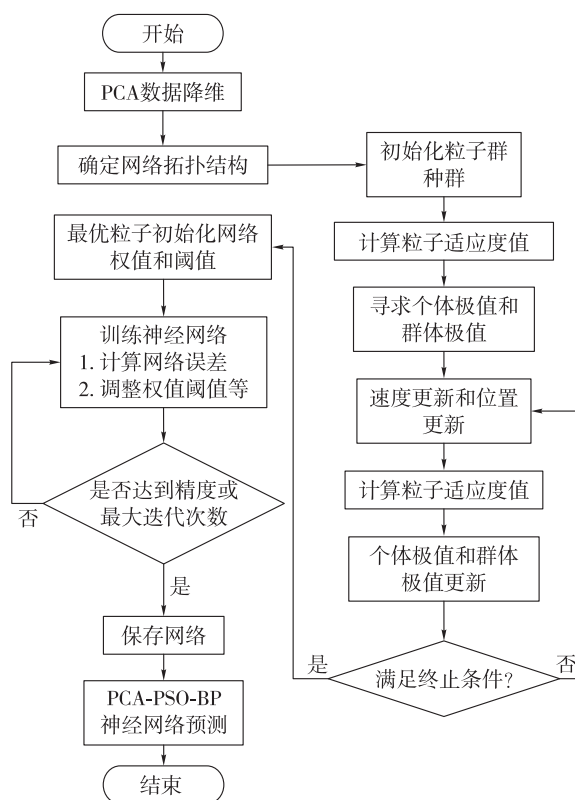


图 1 PCA-PSO-BP 神经网络模型流程

## 3 赤潮预测实例

### 3.1 试验数据

赤潮爆发对我国浅海养殖业损失巨大,对赤潮进行实时监测预测十分必要。本文数据选取 2007 年 7 月、8 月间烟台市四十里湾海域赤潮 15 组监测数据<sup>[19]</sup>,将浮游植物密度作为决定性因素,即网络输出,实时监测数据见表 1。

表 1 烟台市四十里湾部分监测数据

样本	水温 /(℃)	透明度 /m	pH	盐度 /(%)	COD /(mg/L)	DO /(mg/L)	磷酸盐 /(mg/L)	总氮 /(mg/L)	叶绿素 a /(mg/L)	浮游植物密度 /(10 <sup>5</sup> cell/L)
1	26.47	1.5	8.54	30.000	2.76	14.22	1.900 00	171.24	4.19	16.50
2	26.22	1.2	8.63	29.200	3.25	16.81	3.300 00	179.21	11.20	71.74
3	24.47	1.2	8.34	29.950	1.83	8.90	11.100 00	157.34	38.87	4.90
4	24.44	1.2	8.55	29.730	3.91	12.41	6.400 00	238.51	4.94	0.88
5	25.15	1.2	8.07	30.620	4.99	8.31	5.300 00	153.10	78.38	16.88
6	24.20	3.5	8.23	31.365	2.03	9.08	0.001 79	0.054 069	0.002	0.349 96
7	24.20	2.5	8.30	31.292	2.36	11.80	0.002 08	0.026 613	0.002 05	0.512 62
8	23.70	2.2	8.10	30.718	2.14	8.85	0.006 91	0.111 62	0.002 25	0.499 40
9	24.50	3.0	8.29	31.070	2.15	9.38	0.002 36	0.027 88	0.001 46	0.181 53

续表1										
样本	水温 /(℃)	透明度 /m	pH	盐度 /(%)	COD /(mg/L)	DO /(mg/L)	磷酸盐 /(mg/L)	总氮 /(mg/L)	叶绿素 a /(mg/L)	浮游植物密度 /(10 <sup>5</sup> cell/L)
10	24.40	4.0	8.28	31.238	2.00	8.56	0.002 50	0.030 53	0.000 447	0.305 73
11	24.20	3.5	8.26	31.430	2.13	8.29	0.002 64	0.037 40	0.001 26	0.228 67
12	26.70	1.1	8.21	22.912	2.87	11.40	0.145 00	1.389 40	0.014 1	6.370 00
13	26.60	0.9	7.53	25.016	1.64	13.90	0.101 00	1.378 10	0.027 1	19.00
14	26.50	0.4	7.28	26.896	6.03	17.00	0.094 70	0.873 72	0.220	41.32
15	26.50	0.5	8.00	30.057	4.04	14.11	0.005 62	0.294 60	0.032 9	30.80

3.2 主成分分析处理

由表 1 可知,实测数据中浮游植物密度受水温、透明度、pH、盐度等 9 个理化因子影响,其影响因子的相关系数矩阵见表 2。由表 2 可知,影响因子间相关系数中有 10 个绝对值大于 0.5,由此可见因子间存在一定的相关性,预测模型的精度必然会受到一定影响。对 15 组影响因子数据进行主成分分析,其主成分方差贡献率见表 3。由表 3 可知,前 5 个主成分的累计贡献率已达到 94.838 4%,表明前 5 个主成分已包含原始数据大部分信息。求得前 5 个主成分的荷载矩阵见表 4。

表 2 各影响因子间相关矩阵										
	水温 /(℃)	透明度 /m	pH	盐度 /(%)	COD /(mg/L)	DO /(mg/L)	磷酸盐 /(mg/L)	总氮 /(mg/L)	叶绿素 a /(mg/L)	
水温	1	<b>-0.716 8</b>	-0.357 1	<b>-0.725 3</b>	0.415 9	<b>0.765 6</b>	-0.143 1	0.061 4	-0.062 3	
透明度	-0.716 8	1	0.348 4	<b>0.590 6</b>	<b>-0.596 0</b>	<b>-0.669 5</b>	-0.337 0	-0.372 8	-0.254 0	
pH	-0.357 1	0.348 4	1	0.449 0	-0.371 5	-0.257 0	0.344 2	<b>0.536 6</b>	0.060 2	
盐度	-0.725 3	0.590 6	0.449 0	1	-0.164 3	-0.454 1	0.107 8	0.117 5	0.148 2	
COD	0.415 9	-0.596 0	-0.371 5	-0.164 3	1	0.477 5	0.093 1	0.246 0	0.323 6	
DO	0.765 6	-0.669 5	-0.257 0	-0.454 1	0.477 5	1	-0.122 9	0.185 1	-0.299 5	
磷酸盐	-0.143 1	-0.337 0	0.344 2	0.107 8	0.093 1	-0.122 9	1	<b>0.794 7</b>	<b>0.661 4</b>	
总氮	0.061 4	-0.372 8	0.536 6	0.117 5	0.246 0	0.185 1	0.794 7	1	<b>0.505 9</b>	
叶绿素 a	-0.062 3	-0.254 0	0.060 2	0.148 2	0.323 6	-0.299 5	0.661 4	0.505 9	1	

表 3 主成分方差贡献率							
主成分	特征值	方差贡献率 /(%)	累计方差贡献率 /(%)	主成分	特征值	方差贡献率 /(%)	累计方差贡献率 /(%)
1	3.535 4	39.282 6	39.282 6	6	0.226 7	2.519 4	97.357 8
2	2.721 8	30.241 8	69.524 3	7	0.118 8	1.320 1	98.677 9
3	1.129 7	12.552 7	82.077 0	8	0.068 4	0.760 1	99.438 0
4	0.797 0	8.855 3	90.932 3	9	0.050 6	0.562 0	100
5	0.351 5	3.906 1	94.838 4				

表 4 主成分荷载矩阵									
主成分	水温 /(℃)	透明度 /m	pH	盐度 /(%)	COD /(mg/L)	DO /(mg/L)	磷酸盐 /(mg/L)	总氮 /(mg/L)	叶绿素 a /(mg/L)
第一主成分	0.895 1	-0.889 2	-0.529 8	-0.739 0	0.651 6	0.821 0	-0.004 0	0.132 6	0.022 9
第二主成分	-0.109 4	-0.324 0	0.468 1	0.239 8	0.262 0	-0.085 1	0.915 5	0.905 8	0.770 6
第三主成分	0.194 3	0.017 1	0.626 7	-0.078 5	-0.376 9	0.401 4	-0.003 9	0.328 7	-0.530 6
第四主成分	-0.100 5	0.089 7	0.062 0	0.554 0	0.535 3	0.310 3	-0.245 2	0.079 7	-0.136 9
第五主成分	0.294 4	0.150 8	0.262 1	-0.069 1	0.090 6	-0.128 9	-0.272 6	-0.014 6	0.263 4



由表4和式(3)可得主成分的线性表达式如式(11)所示。原网络输入由9个理化因子降低至只有5个,从而减少网络输入,提高算法效率和精度。

$$\left. \begin{aligned} y_1 &= 0.8951X_1 - 0.8892X_2 - 0.5298X_3 - \\ &\quad 0.7390X_4 + 0.6516X_5 + 0.8210X_6 - \\ &\quad 0.0040X_7 + 0.1326X_8 + 0.0229X_9 \\ y_2 &= -0.1094X_1 - 0.3240X_2 + 0.4681X_3 + \\ &\quad 0.2398X_4 + 0.2620X_5 - 0.0851X_6 + \\ &\quad 0.9155X_7 + 0.9059X_8 + 0.7706X_9 \\ y_3 &= 0.1943X_1 + 0.0171X_2 + 0.6267X_3 - \\ &\quad 0.0785X_4 - 0.3769X_5 + 0.4014X_6 - \\ &\quad 0.0039X_7 + 0.3287X_8 - 0.5306X_9 \\ y_4 &= -0.1005X_1 + 0.0897X_2 + 0.0620X_3 + \\ &\quad 0.5540X_4 + 0.5353X_5 + 0.3103X_6 - \\ &\quad 0.2452X_7 + 0.0797X_8 - 0.1369X_9 \\ y_5 &= 0.2944X_1 + 0.1508X_2 + 0.2621X_3 - \\ &\quad 0.0691X_4 + 0.0906X_5 - 0.1289X_6 - \\ &\quad 0.2726X_7 - 0.0146X_8 + 0.2634X_9 \end{aligned} \right\} \quad (11)$$

### 3.3 PCA-PSO-BP神经网络赤潮预测

本文利用降维后的前13组数据作为 PSO-BP

神经网络的训练样本,后2组数据作为预测数据,用平均绝对误差(mean absolute error, MAE)、平均绝对百分误差(absolute percentage error, MAPE)、误差均方差(mean squared error, MSE)<sup>[20]</sup>进行精度评定。

为降低网络复杂度,本文采用单隐含层的神经网络结构。依据式(8)得隐含层节点数为3到13间的整数,采用基于L-M优化算法进行多次网络训练,最终确定隐含层节点数目为5时网络拟合效果最佳。

为验证本文算法的有效性,本文算法分别与 PSO-BP、PCA-BP、BP 神经网络预测模型进行对比,训练次数均设为1000,目标误差均设为0.00001,学习速率均设为0.1,PCA-PSO-BP 和 PSO-BP 神经网络预测模型粒子种群均设为50,最大允许迭代次数设为30,加速因子 $c_1$ 、 $c_2$ 均设为1.49445,最大限制速度 $v_{\max}$ 设为1,PCA-PSO-BP、PCA-BP 神经网络输入层数为5,而 PSO-BP 和 BP 神经网络输入层数为9。训练后的4种网络性能见图2,可以发现采用粒子群优化的神经网络能够更快收敛,同时采用主成分分析网络模型精度更高。同时将预测结果列于表5,各模型预测值、精度比较分别列于表5—表6。

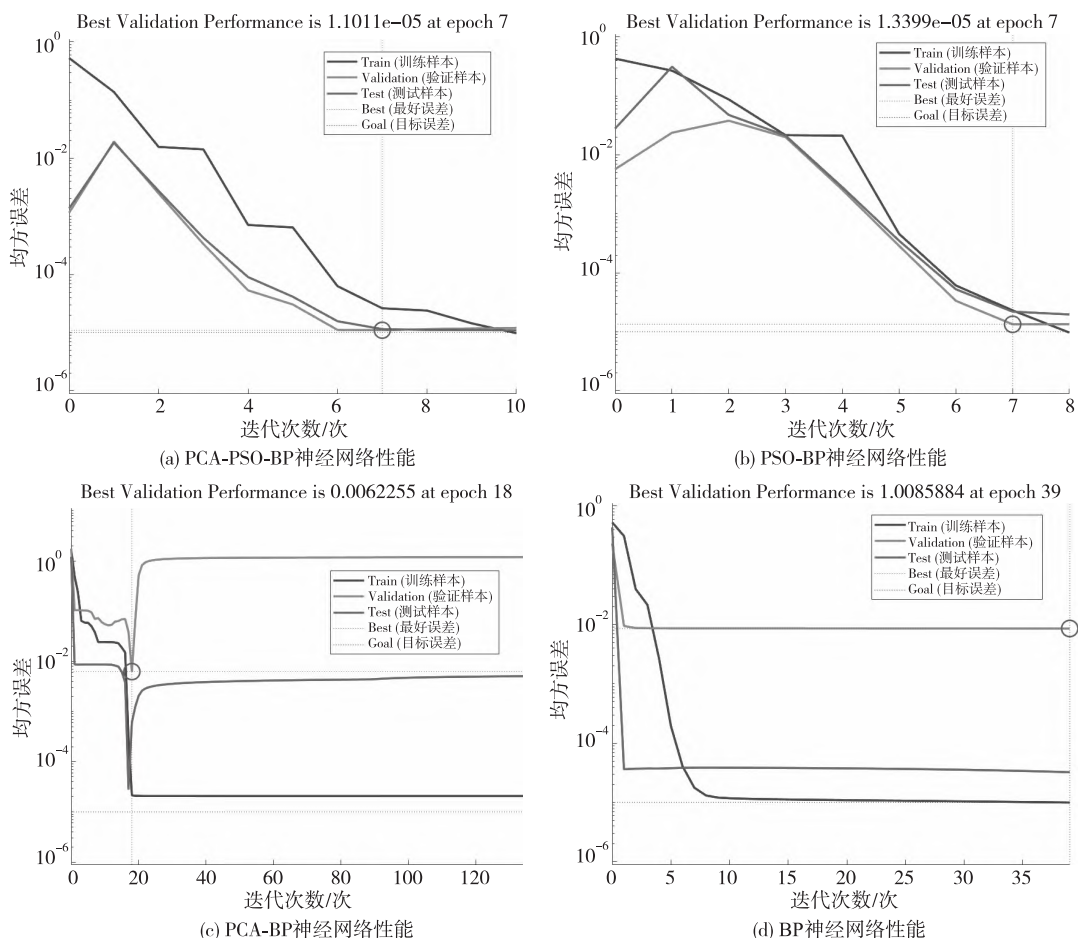


图2 神经网络性能

表 5 各预测模型赤潮预测值

监测 期数	实测值/(10 <sup>5</sup> cell/L)	BP 神经网络预测模型		PCA-BP 神经网络预测模型	
		预测值/(10 <sup>5</sup> cell/L)	绝对误差/(10 <sup>5</sup> cell/L)	预测值/(10 <sup>5</sup> cell/L)	绝对误差/(10 <sup>5</sup> cell/L)
14	41.32	43.410 3	2.090 3	39.850 4	1.469 6
15	30.80	27.542 5	3.257 5	32.459 3	1.659 3

监测 期数	实测值/(10 <sup>5</sup> cell/L)	PSO-BP 神经网络预测模型		PCA-PSO-BP 神经网络预测模型	
		预测值/(10 <sup>5</sup> cell/L)	绝对误差/(10 <sup>5</sup> cell/L)	预测值/(10 <sup>5</sup> cell/L)	绝对误差/(10 <sup>5</sup> cell/L)
14	41.32	40.642 5	0.677 5	41.868 3	<b>0.548 3</b>
15	30.80	31.715 2	0.915 2	29.957 6	<b>0.842 4</b>

表 6 各预测模型预测精度比较

神经网络预测模型	MAE/(10 <sup>5</sup> cell/L)	MSE/(10 <sup>5</sup> cell/L)	MAPE/(%)
BP	2.673 9	7.149 7	7.817 6
PCA-BP	1.564 5	2.447 5	4.472 0
PSO-BP	0.796 4	0.634 2	2.305 5
PCA-PSO-BP	<b>0.695 4</b>	<b>0.483 5</b>	<b>2.031 0</b>

对比图 2(a)和(b)分析,经过主成分分析降低了网络输入,网络目标函数收敛值更小,有效地解决了输入变量间的相关性及输入数据过多的缺陷;综合图 2 可以发现,PCA-PSO-BP 神经网络收敛速度最快、网络拟合整体性能最好,而单纯 BP 神经网络 39 次迭代的目标函数收敛误差最大,拟合效果最差,再次验证了粒子群算法可以提高预测模型精度,再结合主成分分析,提取主成分消除输入变量间的相关性,预测模型的精度可进一步提升。结合表 6 可以发现,传统 BP 神经网络模型的预测误差最大,可见传统 BP 神经网络模型出现过拟合,导致泛化能力差。再结合表 7 可以发现,PCA-PSO-BP 神经网络预测模型在 MAE、MSE、MAPE 三项指标上较 PSO-BP 神经网络预测模型分别提高 12.682 1%、23.762 2%、11.906 3%,较传统 BP 神经网络分别提高 73.993 0%、93.237 5%、74.020 2%;PCA-BP 神经网络在三项指标上较传统 BP 神经网络分别提高 41.490 0%、65.767 8%、42.795 7%,预测精度显著提升,因此本文结合粒子群和主成分分析算法在收敛速度和预测精度上较其他三种模型都有一定提升,验证了本文算法的有效性和可行性。

4 结 语

为了解决复杂系统中变量相关性影响赤潮预测模型精度这一问题,本文采用主成分分析方法,在保留原始观测信息的前提下对输入数据降维,简化模型输入,建立了一种高精度赤潮预测模型。针对赤潮实例数据进行验证,通过与 PSO-BP、PCA-BP、BP 神经网络预测模型进行对比分析,PCA-PSO-BP 神

经网络预测模型收敛速度快、精度高,验证了本文算法在赤潮预测领域的有效性和可行性。

参考文献:

[1] 张青田. 中国海域赤潮发生趋势的年际变化[J]. 中国环境监测,2013,29(5):98-102.

[2] 王兴强,刘长兴,刘国伟,等. 改进的 CSFCM 聚类算法及其在赤潮监测中的应用[J]. 计算机工程与应用,2012,48(8):233-235,245.

[3] GOKARAJU B,DURHHA S S. Ensemble methodology using multistage learning for improved detection of harmful algal blooms[J]. IEEE Geoscience and Remote Sensing Letters,2012,9(5):827-835.

[4] KIM B M, SEO K B, KOO J S. Measurement and analysis of three-dimensional shape variations of microalgae based on a digital holographic microscopy for prediction of the spread point of red tide [J]. International Multidisciplinary Microscopy Congress, 2014,154:269-274.

[5] 张健,杨翼,杨璐,等. 东海近岸海域赤潮与环境因子的关系[J]. 广东海洋大学学报,2019,39(1):66-70.

[6] 刘汉丽,周成虎,朱阿兴,等. 多子群遗传神经网络模型用于路口短时交通流量预测[J]. 测绘学报,2009,38(4):363-368.

[7] 许雄,钟燕飞,张良培,等. 基于空间自相关 BP 神经网络的遥感影像亚像元定位[J]. 测绘学报,2011,40(3):307-311.

[8] 吴翊,李永生,胡庆军. 应用数理统计[M]. 长沙:国防科技大学出版社,1995.

[9] 姜双城,林培梅,蔡建堤,等. 基于 PCA 法的沉积物质量评价模型构建及其应用[J]. 应用海洋学学报,

- 2020,39(4):501-510.
- [10] 宋传峰,秘金钟,党亚民.主成分分析在GPS水准数据处理中的应用[J].测绘科学,2014,39(6):90-93.
- [11] 郭书良,陈建宏,杨珊.基于主成分分析与BP网络的锚杆支护方案优选[J].工程设计学报,2012,19(2):150-155.
- [12] 李海涛,邵泽东.基于头脑风暴优化算法与BP神经网络的海水水质评价模型研究[J].应用海洋学学报,2020,39(1):57-62.
- [13] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报,2017,40(6):1229-1251.
- [14] 朱贵重,胡松.基于LSTM-RNN的海水面温度模型研究[J].应用海洋学学报,2019,38(2):191-197.
- [15] KENNEDY J, EBERHART R. Particle swarm optimization[C]//Proceedings of IEEE International Conference on Neural Networks. Perth:Neural Networks, 1995:1942-1948.
- [16] 尹怀仙,王凯,张铁柱,等.基于PSO-BP神经网络的城轨列车转向架轮对轴箱故障预测[J].复杂系统与复杂性科学,2015,12(4):97-103.
- [17] 郁磊. MATLAB智能算法30个案例分析[M].北京:北京航空航天大学出版社,2011.
- [18] 邓传军,欧阳斌,陈艳红.一种基于PSO-BP神经网络的建筑物沉降预测模型[J].测绘科学,2018,43(6):27-31,38.
- [19] 胡婧,姬鹏,张承慧.基于LMBP的烟台四十里湾赤潮预测模型[J].系统仿真学报,2009,21(19):6250-6255.
- [20] 冯绍权,花向红,陶武勇,等.一种基于GA-BP-MC神经网络的高铁桥墩沉降预测模型[J].测绘通报,2019(7):50-53,82.

(上接第230页)

撑,最终为政府实现信息掌握到位、治安防控到位、问题解决到位、便民服务到位,并不断提升居民群众生活的幸福感和获得感。

#### 参考文献:

- [1] 自然资源部.智慧城市时空大数据平台建设技术大纲(2019版)[EB/OL].[http://gi.mnr.gov.cn/201902/t20190218\\_2395831.html](http://gi.mnr.gov.cn/201902/t20190218_2395831.html).
- [2] 许镇,吴莹莹,郝新田,等.CIM研究综述[J].土木工程信息技术,2020,12(3):1-7.
- [3] 汪科,杨柳忠,季珏.新时期我国推进智慧城市和CIM工作的认识和思考[J].建设科技,2020(18):9-12.
- [4] 陈才,张育雄.加快构建CIM平台,助力数字孪生城市建设[J].信息通信技术与政策,2020(11):14-17.
- [5] 王喜,杨华,范况生.城市网格化管理系统的关键技术及示范应用研究[J].测绘科学,2006(4):117-119.
- [6] 陈曙凌,王新南,褚洪岩.智慧社区运营管理平台建设[J].智能建筑与智慧城市,2017(12):87-89.
- [7] 王斌,贺嘉钰.试析信息技术在基层社会管理中的应用:以社区网格化为例[J].国际新闻界,2013,35(9):6-14.
- [8] 和世开.社区地图在社区网格化管理中的应用[J].城市勘测,2019(1):111-113.
- [9] 陈和.城市社区网格化管理模式与实践探析——以温州社区为例[J].温州大学学报(社会科学版),2015,28,(4):79-86.
- [10] 张鹏程,何华贵,杨梅.智慧广州时空云平台在“四标四实”中的应用与实践[J].测绘与空间地理信息,2019,42(3):24-25,30.
- [11] 丁占峰.基于GIS的人口网格化管理系统的设计与实现[D].上海:东华理工大学,2015.
- [12] 时静.基于三维GIS的社区网格化管理系统研究[D].唐山:华北理工大学,2016.
- [13] 李长富.网格化社会综治管理系统的设计及实现[D].昆明:昆明理工大学,2016.
- [14] 王诚.基于GIS的智慧城市综合治理系统的设计与实现[J].电子世界,2017(22):127-128.
- [15] 黄文涛.基于无人机航测技术的房产测绘实践[J].资源信息与工程,2019,34(1):141-143.
- [16] 杨宇青.二三维GIS技术在贵阳市社区网格化管理中的应用[J].科技风,2018(36):80.
- [17] 罗康,张驰月.基于GIS的城市网格化管理关键技术研究[J].科技资讯,2011(25):44-45.
- [18] 张希磊.基于GIS的社会综治管理系统建设及应用[D].济南:山东大学,2018.