

# 基于改进的 GRNN-Markov 水质预测模型研究及应用

徐睿<sup>1</sup>, 张晓斌<sup>2\*</sup>, 薛鹏松<sup>3</sup>

(1.运城市水利勘测设计研究院有限公司, 山西 运城 044000;

2.运城学院 应用化学系, 山西 运城 044000; 3.陕西锦科环保工程有限公司, 西安 710119)

**摘要:**【目的】针对水环境系统的复杂性, 结合 GRNN 网络预测模型和马尔科夫理论, 构建了改进 GRNN-Markov 水质预测模型, 为模拟不确定性、复杂多变的河流水质变化趋势提供可靠的方法依据。【方法】以汾河入黄口主要污染物 COD<sub>Cr</sub> 为研究指标, 采用 3 $\sigma$  准则对水质监测资料进行预处理, 利用灰色关联分析 (GRA) 确定输入节点, 解决 GRNN 网络对关键参数不能自动寻优的缺点, 同时采用改进 GRNN 模型对水质数据进行模拟预测, 针对水质预测数据的随机波动性, 通过 Markov 模型修正误差残值, 以达到更好的预测结果, 为水环境保护与治理提供新思路、新方法。【结果】改进 GRNN-Markov 水质预测模型, 可以提高水质预测结果的精度, 使相对误差从-38.27%~15.71% 提高到-25.77%~5.25%, 修正结果更加接近实测值。【结论】验证了组合模型在小样本水质预测中的可行性, 为水环境管理提供了科学依据。

**关键词:** 灰色关联分析 (GRA); 改进 GRNN-Markov 模型; 汾河入黄口; 水质预测

中图分类号: X832; TP183

文献标志码: A

doi: 10.13522/j.cnki.gggs.2022152

徐睿, 张晓斌, 薛鹏松. 基于改进的 GRNN-Markov 水质预测模型研究及应用[J]. 灌溉排水学报, 2022, 41(Supp.1): 104-110.

XU Rui, ZHANG Xiaobin, XUE Pengsong. Research and Application on Improved GRNN-Markov Water Quality Prediction Model[J]. Journal of Irrigation and Drainage, 2022, 41(Supp.1): 104-110.

## 0 引言

水环境系统具有复杂性及多变性的特点, 实现对河流水质的模拟和准确预测, 将为水环境保护与规划提供科学依据。河流水质预测的主要方法有 ARIMA 模型法 (时间序列法)<sup>[1-2]</sup>、Multiple Regression 模型法 (多元回归法)<sup>[3]</sup>、模糊逻辑法<sup>[4]</sup>、Artificial Neural Network 简称 ANN (人工神经网络法) 等<sup>[5-6]</sup>。国内外学者在水质、水量预测研究方面已取得一些研究成果。刘东君等<sup>[8]</sup>采用最优加权组合模型对再生水补给的永定河水质 (溶解氧 DO) 进行了预测, 拟合结果表明, 组合模型相对于单一模型而言, 可以平衡偏差, 提高预测精度, 使模型适用性更好。刘洁等<sup>[9]</sup>为实现河流水质的实时预警预测, 构建了基于 IGA-BPNN (遗传-神经网络) 的水质预测模型, 预测结果表明, 该模型可以有效地识别长时间的水质异

常情况或显著的瞬时水质变化。张文丽等<sup>[10]</sup>利用 GM (1, 1) 贫数据建模和 GRNN 非线性适应能力强优点, 构建了灰色广义回归神经网络预测模型, 并将该模型应用于工业废水排放量预测中, 研究显示, 对比 GRNN 模型和改进 GM 模型, 该模型能更准确地对废水排放量进行模拟。赵雪花等<sup>[11]</sup>针对径流序列的非线性和噪声特性, 构建了基于 CEEMD-GRNN 组合的月径流预测模型, 该模型能够有效提高预测精度, 为径流预测提供一种新的有效方法。王雪冬等<sup>[12]</sup>采用广义回归神经网络对矿井涌水量进行了预测, 研究表明 GRNN 模型较 BP 网络在训练样本数量较少时预测精度更高, 对矿井涌水量预测结果准确、可靠。陶志勇等<sup>[13]</sup>针对复杂的水环境, 建立了 FOA-GRNN 水质预测模型, 并以太湖水水质为例进行了预测, 并与其他模型相比, 研究表明其仿真结果稳定性强、精度高, 论证了该模型在水质预测方面的优越性, 为水环境决策管理提供了科学依据。王昱文等<sup>[14]</sup>为解决传统水质预测方法的不足, 针对长江流域水质现状, 提出了基于复合神经网络的多元水质预测 (MWQP) 模型, 经对比试验, 该模型具有预测精度高、适用性强等优点。

由以上研究发现, 单一的 BP 神经网络模型在进

收稿日期: 2022-03-23

基金项目: 国家自然科学基金项目 (51479215); 山西省水利厅科技项目 (TZ2019026); 运城学院学科经费资助项目

作者简介: 徐睿 (1970-), 男, 山西运城人。高级工程师, 主要从事水利勘测设计、流域生态修复。E-mail: sjyxurui@163.com

通信作者: 张晓斌 (1977-), 男, 山西运城人。教授, 博士, 主要从事流域生态修复、环境模拟及污染控制。E-mail: sxzxb88@126.com

行预测时存在学习速度慢、易陷入局部极值，预测精度不高等缺点，相对于 BP 神经网络模型，GRNN 网络模型具有收敛速度快、非线性逼近能力强、预测精度较好等优点，但也存在计算复杂度高、空间复杂度高、且模型精度依赖于光滑因子(Spread 值)等缺点。采用组合模型(其他算法改进神经网络模型)进行预测时，相对于单一模型而言，其适用性更强，预测精度更高，但对于受外界干扰较大，易发生波动的随机数据，其预测精度会降低。本研究在前人研究的基础上，针对河流水环境系统的复杂性和多变性，结合 GRNN 网络较强的非线性映射能力和高精度逼近能力，通过构建改进 GRNN-Markov 水质预测模型，先利用 GRNN 网络所需样本数量少、对解决非线性问题有着很好实用性的优点，对历史水质数据进行模拟、得出较好的预测结果，再针对水质预测数据的随机波动性，通过 Markov 模型修正误差残值，达到更好的预测精度，从而实现 GRNN 网络与 Markov 模型的优势互补，为模拟不确定性、复杂多变的河流水质变化趋势提供可靠的方法和依据，为水环境保护与治理提供新思路、新方向<sup>[15]</sup>。

## 1 基本原理及方法

### 1.1 广义回归神经网络—GRNN

GRNN 广义神经网络是高度并行径向基网络的一个分支，其进行非线性回归分析根据的是数理统计的原理<sup>[16]</sup>。GRNN 网络具有很强的非线性映射能力和柔性网络结构以及高度的容错性和鲁棒性，适用于解决非线性问题<sup>[17]</sup>。GRNN 网络在结构上与 RBF 网络较为相似，其网络结构如图 1 所示。GRNN 网络在逼近能力、分类能力和学习速度方面具有较强优势，网络最后收敛于样本量积聚最多的优化回归面，并且在数据缺乏时效果也较好<sup>[18]</sup>。

设随机变量  $x$  和随机变量  $y$ ，则 GRNN 模型  $y$  的预测输出  $\hat{y}$  为：

$$\hat{y} = E[y|x] = \frac{\int_{-\infty}^{\infty} yf(x, y)dy}{\int_{-\infty}^{\infty} f(x, y)dy}, \quad (1)$$

式中： $\hat{y}$  为输入是  $x$  时， $y$  的预测输出； $E[y|x]$  为输出向量  $y$  的期望值； $f(x, y)$  为随机变量  $x$  和随机变量  $y$  的联合概率密度函数。

采用 Parzen 非参数估算得到 GRNN 网络的概率密度函数  $f(x, y)$ ：

$$f(x, y) = \frac{1}{n(2\pi)^{(p+1)/2} \sigma^{p+1}} \sum_{i=1}^n \exp\left[-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}\right] \exp\left[-\frac{(x-y_i)}{2\sigma^2}\right], \quad (2)$$

经过积分变换即可得到 GRNN 模型的网络输出为  $\hat{y}$ ：

$$\hat{y} = \frac{\sum_{i=1}^n y_i \exp\left[-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}\right]}{\sum_{i=1}^n \exp\left[-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}\right]}, \quad (3)$$

式中： $x_i, y_i (i=1, 2, 3, \dots, n)$  分别为随机变量  $x, y$  的样本观测值； $p$  为随机变量  $x$  的维数； $\sigma$  为光滑因子。

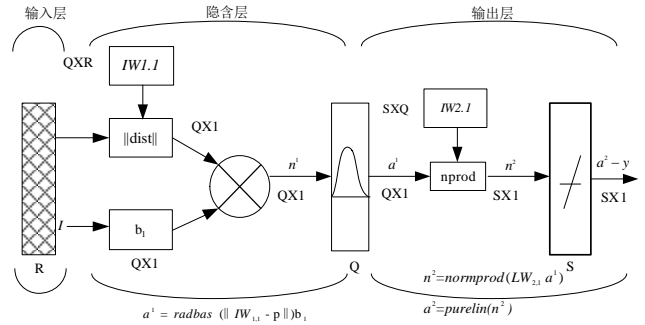


图 1 广义回归神经网络结构图

Fig.1 Generalized regression network structure diagram

### 1.2 灰色关联分析法

将灰色关联分析<sup>[19]</sup>(GRA)与 GRNN 网络模型相结合，可解决 GRNN 网络建模过程中输入节点无法自动寻优的缺点。关联度计算步骤如下：

1) 首先，对参考序列  $X_0$  和比较序列  $X_i$  进行无量纲化处理。

$$X_0(k), k=1, 2, 3, \dots, n, \quad (4)$$

$$X_i = \{X_i(k) | k=1, 2, \dots, n\} \quad i=1, 2, \dots, m. \quad (5)$$

2) 其次，计算参考序列与比较序列的关联系数  $\xi_i(k)$ 。

$$\xi_i(k) = \frac{\min_i \min_k |X_0(k) - X_i(k)| + \rho \max_i \max_k |X_0(k) - X_i(k)|}{|X_0(k) - X_i(k)| + \rho \max_i \max_k |X_0(k) - X_i(k)|}, \quad (6)$$

式中： $\rho$  为分辨系数， $0 < \rho < 1$ ，一般情况下  $\rho=0.5$ 。

3) 最后，根据计算所得的关联系数求出关联度。

$$r_i = \frac{1}{N} \sum_k \xi_i(k), \quad (7)$$

式中： $r_i$  为关联度，其值越大，则比较序列  $X_i$  与确定序列  $X_0$  关系越密切。

### 1.3 马尔科夫模型

马尔科夫预测法，是一种关于事件发生的概率预测方法，它将时间序列看成一个随机过程，通过事物不同状态的初始概率与转移概率，确定未来状态的发展<sup>[20]</sup>。其定义为：如果对任何一列状态  $i_0, i_1, \dots, i_{n-1}, i, j$ ，及对任何  $n \geq 0$ ，随机过程  $\{X_n, n \geq 0\}$  满足：

$P\{X_{n+1}=j | X_0=i_0, \dots, X_{n-1}=i_{n-1}, X_n=i\} = P\{X_{n+1}=j | X_n=i\}$ ，则随机过程  $\{X_n, n \geq 0\}$  为离散时间马尔科夫链<sup>[21]</sup>。

本研究在进行水质预测时，先将 GRA—GRNN

模型预测值与实测值的相对误差作为一个随机序列，采用黄金分割率对马尔科夫链的状态区间进行划分，计算状态转移概率和状态转移矩阵，确定水质预测值状态，从而对误差进行修正，进而提高水质预测精度。改进 GRNN-Markov 模型算法流程如图 2 所示。

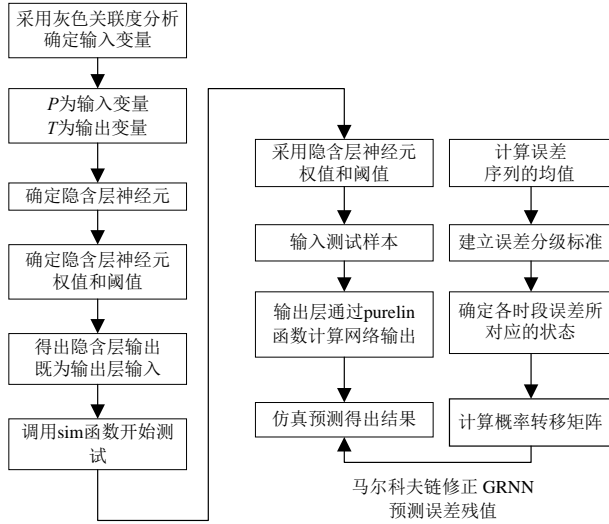


图 2 改进 GRNN-Markov 模型算法流程图

Fig.2 Algorithm flowchart of the improved GRNN-Markov model

### 1.3.1 马尔科夫链的状态划分

应用式 (8) 黄金分割率对马尔科夫链的状态区间进行划分，由于状态区间分级序列须为正值，先采用式 (9) 将水质监测数据归一化到[0, 1]之间，再利用式 (10) 将划分区间还原<sup>[22]</sup>。

$$\lambda_r = \Omega^s \bar{x} \quad |s| < t, r = 1, 2, \dots, t, \quad (8)$$

$$x = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad (9)$$

$$x = x_{\min} + x(x_{\max} - x_{\min})。 \quad (10)$$

### 1.3.2 状态转移概率与状态转移矩阵

假设序列  $x_1, x_2, \dots, x_n$  含  $k$  个状态，构成状态空间  $T = \{t_1, t_2, \dots, t_k\}$ 。设  $f_{ij}$  为状态  $t_i$  经过 1 次转移到达状态  $t_j$  的频数，其中  $t_i, t_j \in T, i, j \in I = \{1, 2, \dots, k\}$  则由  $f_{ij}$  组成的矩阵  $(f_{ij})_{i, j \in I}$  就是相应的状态转移频数矩阵，则转移概率  $P_{ij}$  为：

$$P_{ij} = \frac{f_{ij}}{\sum_{j=1}^k f_{ij}}。 \quad (11)$$

由  $P_{ij}$  组成的矩阵  $(P_{ij})_{i, j \in I}$  叫作一步状态转移概率矩阵，记作  $P_1$ 。

$$P_1 = (P_{ij})_{i, j \in I}。 \quad (12)$$

### 1.3.3 Markov 修正 GRA-GRNN 网络预测结果

采用式 (13) 对 GRNN 网络预测结果进行修正。

$$F = F_g / (1 - q), \quad (13)$$

式中： $F_g$  为预测值； $q$  为原状态区间的分界值。

## 2 改进 GRNN-Markov 水质预测模型的建立

本研究在 GRNN 预测模型和马尔科夫理论的基础上，构建改进 GRNN-Markov 水质预测模型。该组合模型由 3 个部分组成，首先对水质监测数据进行处理，剔除异常数据，利用灰色关联分析 (GRA) 确定 GRNN 模型的输入节点。其次，采用训练数据对 GRNN 模型进行泛化，利用预测样本对改进 GRNN 进行验证。最后，采用 Markov 模型修正误差残值，以得到更为准确的预测结果。

改进 GRNN-Markov 模型算法流程如图 2 所示。

## 3 应用实例

### 3.1 预测因子及 GRNN 网络输入节点的确定

近年来，汾河入黄口水质污染严重，其中化学需氧量 (COD<sub>cr</sub>)、氨氮 (NH<sub>3</sub>-N) 及挥发酚 3 项指标尤为严重，本研究将污染最严重的指标 COD<sub>cr</sub> 作为研究对象，进行水质预测模拟分析。汾河入黄口图如图 3 所示。



图 3 汾河入黄口图

Fig.3 Schematic of Fen River's estuary to Yellow River

本研究收集到汾河入黄口近年来枯水期水质监测资料，共计 27 个监测样本，采用 3σ 准则 (拉依达) 对水质监测资料进行预处理，剔除异常测值。根据灰色关联分析 (GRA) 确定与 COD<sub>cr</sub> 关联度的密切的比较序列，作为广义回归神经网络 (GRNN) 的输入节点，通过关联度计算后，选取与确定序列关联密切的 NH<sub>3</sub>-N、挥发酚、水温 (T)、生化需氧量 (BOD<sub>5</sub>) 及化学需氧量 (COD<sub>cr</sub>) 自身作为 GRNN 模型的输入节点，建立 GRA-GRNN 改进模型，解决 GRNN 模型不能自动寻优输入节点的缺点，提高网络模型的可靠性。COD<sub>cr</sub> 关联度计算结果如图 4 所示。各监测因子月平均水质监测数据如图 5 所示。

### 3.2 GRA-GRNN 水质预测模型

GRA-GRNN 模型在结构上与 RBF 网络较为相似，且具有逼近能力强和学习速度快等优点，对于小样本数据也能够精确预测，由于长系列水质监测数据不易

获取，因此本研究采用 GRA-GRNN 改进模型对水质进行预测。

### 3.2.1 GRA-GRNN 模型的建立与泛化

利用训练数据对改进 GRNN 网络进行参数的确定和网络拓扑结构的训练，以保证预测模型得到最佳的泛化程度，其本质上是一个优化问题，即最小化改进 GRNN 网络水质预测误差。本研究训练样本 20 个，样本编号 1#~样本编号 20#，预测样本为 7 个，样本编号 21#~样本编号 27#。

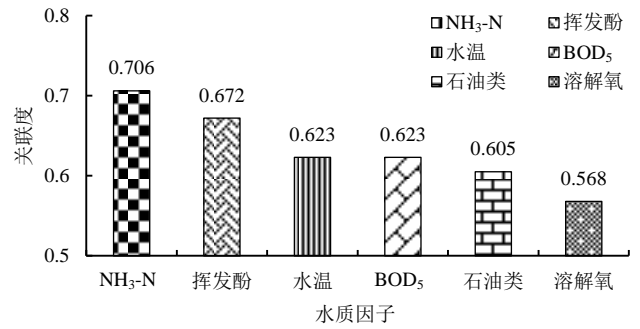
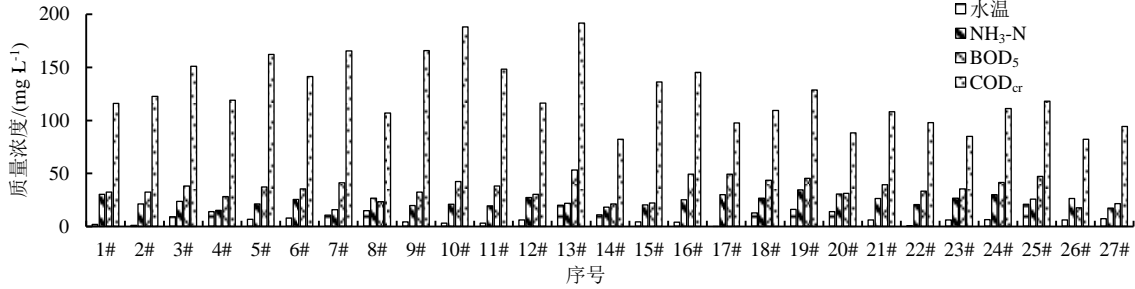
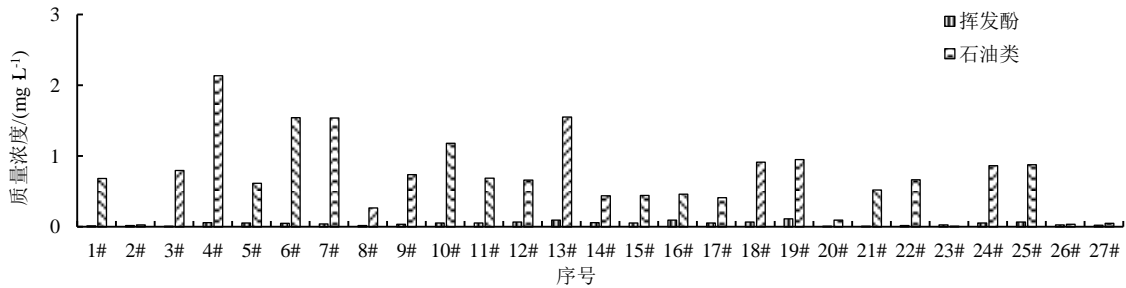


图 4 COD<sub>cr</sub> 关联度计算结果图

Fig.4 Correlation calculation result graph of COD<sub>cr</sub>



(a) 水温、NH<sub>3</sub>-N、BOD<sub>5</sub>、COD<sub>cr</sub> 月平均水质监测数据



(b) 挥发酚、石油类月平均水质监测数据

图 5 监测断面各监测因子月平均水质监测数据图

Fig.5 Water quality monitoring data of Monitoring section

GRNN 网络中 SPREAD 为径向基函数的扩展速度，默认值为 1，为了选取合适的 SPREAD 值，采取循环训练的方法，从而达到最好的模拟效果。SPREAD 值越小，网络对样本的逼近就越强；SPREAD 值越大，网络对样本数据的逼近过程就越平滑，但误差也相应增大。本研究经过反复程序运行，确定 SPREAD 值为 0.8 时，所选数据的拟合结果较好。GRNN 训练样本实测值和拟合值对比如图 6 所示。

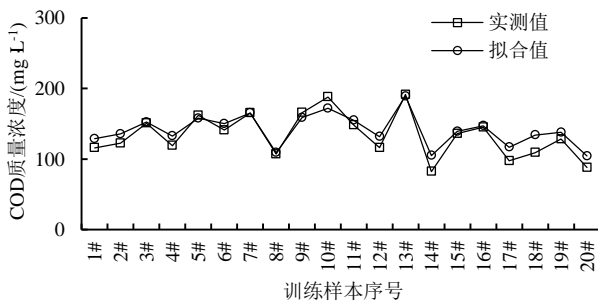


图 6 GRNN 训练样本实测值和拟合值对比图

Fig.6 Comparison diagram of GRNN training samples measured and fitted values

### 3.2.2 GRA-GRNN 模型的验证

利用预测样本（样本编号 21#~样本编号 27#）对训练后的 GRNN 河流水质预测模型的性能和泛化能力进行验证。GRA-GRNN 水质模型预测结果、相对误差如表 1 和图 7 所示。

表 1 GRA-GRNN 模型预测值与实测值的相对误差

Table 1 Relative error of GRA-GRNN predicted values and measured values

样本序号	COD <sub>cr</sub> 质量浓度 实测值/(mg L <sup>-1</sup> )	COD <sub>cr</sub> 质量浓度 预测值/(mg L <sup>-1</sup> )	相对误差/%
21#	108.00	124.97	-15.71
22#	97.60	131.91	-35.15
23#	84.90	116.75	-37.51
24#	111.00	128.88	-16.11
25#	118.00	139.98	-18.63
26#	82.00	113.38	-38.27
27#	94.00	116.12	-23.53

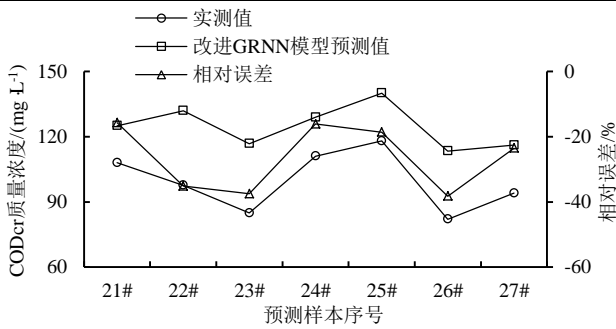


图 7 改进 GRNN 模型预测值与实测值对比图

Fig.7 Comparison diagram of improved GRNN predictive values and measured values

由表 1 和图 7 可知:

1) GRA-GRNN 预测模型可以对汾河入黄口小样本水质进行预测, 水质预测结果介于 113.38~139.98 mg/L 之间, 相对误差介于-38.27%~-15.71%之间。其中 22#、23#、26#样本相对误差较大, 分别为-35.15%、-37.51%、-38.27%。

2) 由 GRA-GRNN 预测模型对汾河入黄口水质预测结果及相对误差可知, 该模型可用于河流水质预测, 但也存在一定的局限性, 当河流水质变化比较平稳时, 模拟与验证精度较好, 当水质浓度变化较大时, GRA-GRNN 模型预测误差会增大, 分析其原因主要是水环境系统的复杂性、多变性, 导致水质数据变化的非线性特点, 尽管水质数据经过归一化处理、转换, 但仍无法满足平稳、正态、零均值这一前提条件, 从而导致预测结果精度降低。

表 2 相对误差状态分布

Table 2 Relative error state distribution

序号	COD <sub>cr</sub> 质量浓度 实测值/(mg L <sup>-1</sup> )	COD <sub>cr</sub> 质量浓度 预测值/(mg L <sup>-1</sup> )	归一化 相对误差	状态
1#	115.75	128.781 3	0.450 2	E <sub>2</sub>
2#	122.51	135.490 5	0.468 5	E <sub>2</sub>
3#	150.73	152.339 1	0.732 3	E <sub>2</sub>
4#	118.92	132.566 7	0.444 2	E <sub>2</sub>
5#	162.00	157.868 5	0.832 4	E <sub>2</sub>
6#	141.12	150.062 8	0.586 4	E <sub>2</sub>
7#	165.22	165.407 5	0.758 7	E <sub>2</sub>
8#	106.93	109.006 8	0.708 1	E <sub>2</sub>
9#	165.67	158.629 7	0.879 5	E <sub>2</sub>
10#	188.00	171.825 9	1.000 0	E <sub>3</sub>
11#	148.00	154.788 4	0.634 9	E <sub>2</sub>
12#	116.00	131.692 8	0.387 3	E <sub>2</sub>
13#	191.48	189.052 9	0.796 9	E <sub>2</sub>
14#	82.22	104.846 5	0.000 0	E <sub>1</sub>
15#	136.00	139.163 7	0.697 4	E <sub>2</sub>
16#	145.02	147.145 7	0.721 3	E <sub>2</sub>
17#	97.54	116.821 1	0.214 6	E <sub>1</sub>
18#	109.07	134.083 3	0.127 0	E <sub>1</sub>
19#	128.37	137.822 5	0.558 0	E <sub>2</sub>
20#	88.00	104.353 1	0.247 4	E <sub>1</sub>

3) Markov 模型适用于预测随机变化性强的数据, 为了提高 GRA-GRNN 水质模型的预测精度, 达到预

期的预测结果, 本研究将 Markov 引入 GRA-GRNN 模型, 弥补 GRA-GRNN 模型对多变性和波动性数据预测精度低的不足。

### 3.3 采用 Markov 修正 GRA-GRNN 模型预测误差

GRA-GRNN 网络水质预测模型训练样本所得的相对误差状态分布如表 2 所示, 本研究采用黄金分割率对马尔科夫链的状态区间进行划分, 划分结果如表 3 所示。

表 3 马尔科夫状态区间划分结果

Table 3 Division result Markov state interval

状态	归一化相对误差界限	相对误差界限
E <sub>1</sub>	[0, 0.347 5]	[-0.275 2, -0.149 7]
E <sub>2</sub>	(0.347 5, 0.909 8)	(-0.149 7, 0.053 4)
E <sub>3</sub>	[0.909 8, 1]	[0.053 4, 0.086 0]

根据划分的状态集, 确定状态转移概率, 计算状态转移矩阵。状态转移矩阵 P<sup>(1)</sup>~P<sup>(7)</sup>计算结果。

$$P^{(1)} = \begin{bmatrix} 0.333 3 & 0.666 7 & 0.000 0 \\ 0.200 0 & 0.733 3 & 0.066 7 \\ 0.000 0 & 1.000 0 & 0.000 0 \end{bmatrix}$$

$$P^{(2)} = \begin{bmatrix} 0.244 4 & 0.711 1 & 0.044 4 \\ 0.213 3 & 0.737 8 & 0.048 9 \\ 0.200 0 & 0.733 3 & 0.066 7 \end{bmatrix}$$

$$P^{(3)} = \begin{bmatrix} 0.223 7 & 0.728 9 & 0.047 4 \\ 0.218 7 & 0.732 1 & 0.049 2 \\ 0.213 3 & 0.737 8 & 0.048 9 \end{bmatrix}$$

$$P^{(4)} = \begin{bmatrix} 0.220 3 & 0.731 1 & 0.048 6 \\ 0.219 3 & 0.731 9 & 0.048 8 \\ 0.218 7 & 0.732 1 & 0.049 2 \end{bmatrix}$$

$$P^{(5)} = \begin{bmatrix} 0.219 7 & 0.731 6 & 0.048 7 \\ 0.219 5 & 0.731 7 & 0.048 8 \\ 0.219 3 & 0.731 9 & 0.048 8 \end{bmatrix}$$

$$P^{(6)} = \begin{bmatrix} 0.219 5 & 0.731 7 & 0.048 8 \\ 0.219 5 & 0.731 7 & 0.048 8 \\ 0.219 5 & 0.731 7 & 0.048 8 \end{bmatrix}$$

$$P^{(7)} = \begin{bmatrix} 0.219 5 & 0.731 7 & 0.048 8 \\ 0.219 5 & 0.731 7 & 0.048 8 \\ 0.219 5 & 0.731 7 & 0.048 8 \end{bmatrix}$$

表 4 为 GRA-GRNN 预测值及改进 GRNN-Markov 修正模型预测值对比表, 图 8 为实测值、预测值与 Markov 修正值对比图, 图 9 为改进 GRNN 模型与 Markov 修正模型相对误差对比图。

表 4 GRA-GRNN 预测值及改进

GRNN-Markov 修正模型预测值对比表

Table 4 GRA-GRNN predictions and GRNN-Markov corrections

样本 序号	质量浓度 实测值/ (mg·L <sup>-1</sup> )	改进 GRNN 模型		改进 GRNN-Markov 修正模型	
		质量浓度 预测值/ (mg·L <sup>-1</sup> )	相对 误差/%	质量浓度 预测值/(mg·L <sup>-1</sup> )	相对 误差/%
21#	108.00	124.97	-15.71	113.67	-5.25
22#	97.60	131.91	-35.15	119.97	-22.92
23#	84.90	116.75	-37.51	106.19	-25.07
24#	111.00	128.88	-16.11	117.23	-5.61
25#	118.00	139.98	-18.63	127.32	-7.90
26#	82.00	113.38	-38.27	103.13	-25.77
27#	94.00	116.12	-23.53	105.62	-12.36

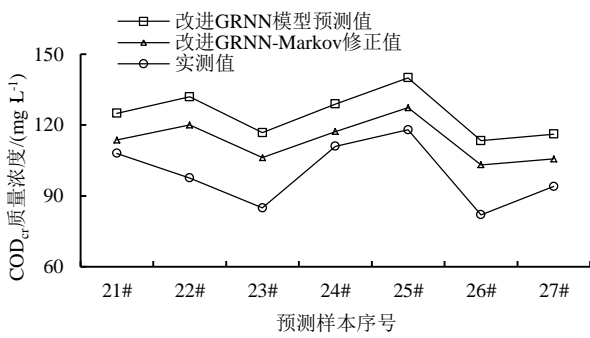


图8 实测值、预测值与 Markov 修正值对比图

Fig.8 Comparison diagram of measured values predictive values and Markov corrections

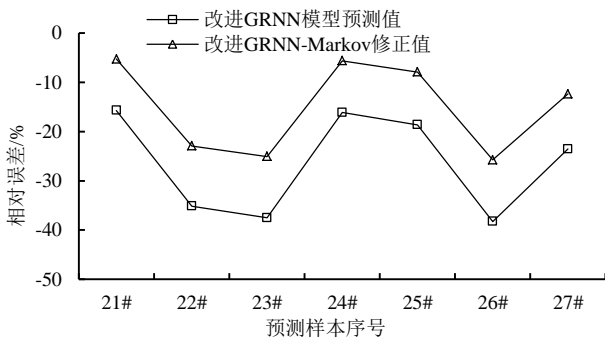


图9 改进 GRNN 模型与 Markov 修正模型相对误差对比图

Fig.9 Relative error comparison chart of improved GRNN predicted and Markov corrections

由图8—图9、表4可知：

1) 改进 GRNN-Markov 模型水质预测结果介于 103.13~127.32 mg/L 之间，相对于 GRA-GRNN 模型，22#、23#、26# 样本预测结果相对误差分别从 -35.15%、-37.51%、-38.27% 提高到了 -22.92%、-25.07%、25.77%。由此可知，GRA-GRNN 模型预测结果存在一定程度的波动导致稳定性较差，且预测精度低于改进 GRNN-Markov 模型。

2) 马尔科夫模型适用于预测随机波动大的动态水质变化过程，本研究将马尔科夫理论引入 GRA-GRNN 模型，将 GRA-GRNN 模型预测值与实测值的相对误差作为随机序列，划分马尔科夫状态区间，计算概率转移矩阵，修正误差残值，提高预测精度。经过马尔科夫模型的修正，水质预测结果的相对误差从 -38.27%~-15.71% 提高到 -25.77%~-5.25%，修正结果更加接近实测值，从而实现了 GRA-GRNN 模型与 Markov 模型的优势互补，为模拟不确定性、复杂多变的河流水质变化趋势提供了可靠的方法依据，为水环境保护与治理提供了新思路、新方向。

3) 改进 GRNN-Markov 模型相较于单一的 GRNN 模型，水质预测精度更高，可见组合模型是水质预测模型研究的发展方向。亦可将该组合模型应用于水库、湖泊水质预测研究，探讨其适用性。

4) 值得注意的是，GRNN 模型的稳定性与预测

精度主要取决于光滑因子  $\sigma$ ，今后还可进一步对 GRNN 模型进行优化及算法处理，寻求 Spread 最优值，最终确定合适的光滑因子，优化 GRNN 模型，再将其与 Markov 模型相结合，可为水质预测模型提供新的研究方向。

## 4 结论

1) GRA-GRNN 预测模型及改进 GRNN-Markov 模型均可以对汾河入黄口水质进行预测。改进 GRNN-Markov 模型水质预测结果介于 103.13~127.32 mg/L 之间，相对于 GRA-GRNN 模型，22#、23#、26# 样本预测结果相对误差分别从 -35.15%、-37.51%、-38.27% 提高到了 -22.92%、-25.07%、25.77%。由此可知，GRA-GRNN 模型预测结果存在一定程度的波动导致稳定性较差，且预测精度低于改进 GRNN-Markov 模型。

2) 马尔科夫模型适用于预测随机波动大的动态水质变化过程，本研究将马尔科夫理论引入 GRA-GRNN 模型，将 GRA-GRNN 模型预测值与实测值的相对误差作为随机序列，划分马尔科夫状态区间，计算概率转移矩阵，修正误差残值，提高预测精度。经过马尔科夫模型的修正，水质预测结果的相对误差从 -38.27%~-15.71% 提高到 -25.77%~-5.25%，修正结果更加接近实测值，从而实现了 GRA-GRNN 模型与 Markov 模型的优势互补，为模拟不确定性、复杂多变的河流水质变化趋势提供了可靠的方法依据，为水环境保护与治理提供了新思路、新方向。

3) 对于不确定性、复杂多变的河流水环境系统、且样本数量有限时，可构建组合模型（改进 GRNN-Markov 模型、GM-Markov 模型等）提高水质预测精度，是水质预测模型研究的发展方向。亦可将该模型应用于水库、湖泊水质预测研究，探讨其适用性。今后还可进一步对模型进行优化及算法处理，以期达到更为精确的水质预测结果。

## 参考文献：

- [1] 董泉汐. 基于深度学习的水环境时间序列预测方法研究[D]. 北京: 北京工业大学, 2020.
- [2] 石月. 基于时间序列分析的松花江流域水质预测[D]. 哈尔滨: 哈尔滨师范大学, 2015.
- [3] 陈昭明, 王伟, 赵迎, 等. 改进主成分分析与多元回归融合的汉丰湖水水质评估及预测[J]. 环境监测管理与技术, 2020, 32(4): 15-19.
- [4] 朱冰, 赵兰兰, 李萌. T-S-K 模糊逻辑算法在抚河水文预报中的应用[J]. 水文, 2015, 35(3): 53-58.
- [5] 周彦辰, 胡铁松, 陈进, 等. 耦合动态方程的神经网络模型在水质预测中的应用[J]. 长江科学院院报, 2017, 34(9): 1-5.
- [6] OSTAD-ALI-ASKARI K, SHAYANNEJAD M,

- GHORBANIZADEH-KHARAZI H. Artificial neural network for modeling nitrate pollution of groundwater in marginal area of Zayandeh-rood River, Isfahan, Iran[J]. KSCE Journal of Civil Engineering, 2017, 21(1): 134-140.
- [7] 崔东文. 几种神经网络模型在湖库富营养化程度评价中的应用[J]. 水资源保护, 2012, 28(6): 12-18.
- [8] 刘东君, 邹志红. 最优加权组合预测法在水质预测中的应用研究[J]. 环境科学学报, 2012, 32(12): 3 128-3 132.
- [9] 刘洁, 祝榕婕, 姜德迅, 等. 基于遗传-神经网络的实时水质预测模型[J]. 南水北调与水利科技(中英文), 2020, 18(6): 93-100.
- [10] 张文丽, 路金喜, 宋双虎, 等. 基于灰色广义回归神经网络的工业废水排放量预测[J]. 水资源与水工程学报, 2007, 18(1): 64-67.
- [11] 赵雪花, 桑宇婷, 祝雪萍. 基于 CEEMD-GRNN 组合模型的月径流预测方法[J]. 人民长江, 2019, 50(4): 117-123, 141.
- [12] 王雪冬, 董慧. 基于广义回归神经网络的矿井涌水量预测[J]. 中国安全生产科学技术, 2014, 10(11): 90-93.
- [13] 陶志勇, 曹琦, 徐光宪. 基于 FOA-GRNN 模型的太湖水质预测研究[J]. 电子测量与仪器学报, 2021, 35(11): 83-90.
- [14] 王昱文, 杜震洪, 戴震, 等. 基于复合神经网络的多元水质指标预测模型[J]. 浙江大学学报(理学版), 2022, 49(3): 354-362, 375.
- [15] XIN Z H, YE L, ZHANG C. Application of export coefficient model and QUAL2K for water environmental management in a rural watershed[J]. Sustainability, 2019, 11(21): 6022.
- [16] 张柏发. 基于 LOADEST 和 PAG/RNN 模型的非点源污染河流营养物质负荷量模拟与预测研究[D]. 杭州: 浙江大学, 2015.
- [17] 裴浩鹏. GRNN 及 PNN 分类算法性能分析[D]. 开封: 河南大学, 2017.
- [18] 施珮, 袁永明, 张红燕, 等. GRNN 和 Elman 神经网络在水体溶解氧预测中的应用[J]. 江苏农业科学, 2017, 45(23): 217-221.
- [19] YAN F, QIAO D Y, QIAN B, et al. Improvement of CCME WQI using grey relational method[J]. Journal of Hydrology, 2016, 543: 316-323.
- [20] 林岩, 陈帅, 陈燕, 等. 道路交通事故的灰色马尔科夫预测模型与算法[J]. 武汉理工大学学报(交通科学与工程版), 2013, 37(5): 924-928.
- [21] 蒋艳辉. 基于指数平滑法与马尔科夫链的道路交通事故预测研究[D]. 哈尔滨: 哈尔滨工业大学, 2018.
- [22] 邢肖鹏. 汾河下游及入黄口水环境模拟与污染物总量控制研究[D]. 西安: 西安理工大学, 2012.

## Research and Application on Improved GRNN-Markov Water Quality Prediction Model

XU Rui<sup>1</sup>, ZHANG Xiaobin<sup>2\*</sup>, XUE Pengsong<sup>3</sup>

(1. Yuncheng Water Conservancy Survey and Design Research Institute Co., Ltd., Yuncheng 044000, China;

2. Yuncheng University, Department of Applied Chemistry, Yuncheng 044000, China;

3. Shaanxi Jinke Environmental Protection Engineering Co., Ltd., Xi'an 710119, China)

**Abstract:** 【Objective】 In view of the complexity of the water environment system, combined with the GRNN network prediction model and Markov theory, an improved GRNN-Markov water quality prediction model was constructed. This model, which provides a reliable method for simulating uncertain, complex and variable river water quality trends.【Method】 According to actual situation of water quality in Fen River's estuary to Yellow River, taking the main pollutant COD<sub>Cr</sub> as the research index, the water quality monitoring data were preprocessed using the Laida criteria. Using Grey Relational Analysis to determine input nodes of GRNN Network which solved the problem that the GRNN network unable to auto select and optimize input nodes. At the same time, the improved GRNN model was used to simulate and predict the water quality data. In view of the random fluctuation of the water quality prediction data, the Markov model was used to correct the error residual value to achieve better prediction results and provide new ideas and methods for water environmental protection and governance. 【Result】 The combined model provides new ideas and methods for water environmental protection and governance. The research show that the improved GRNN-Markov water quality prediction model can improve the accuracy of water quality prediction results. Relative error was -38.27%~-15.71% based on GRA-GRNN model, Relative error was -25.77%~-5.25% based on improved GRNN-Markov water quality prediction model, the correction result was close to the measured value. 【Conclusion】 This combination model can be used in water quality prediction based on small sample data. This study provides a scientific basis for water environment management.

**Key words:** grey relational analysis (GRA); improved GRNN-Markov model; Fen River's estuary to Yellow River; water quality prediction

责任编辑: 赵宇龙