

基于小波去噪和 WNN-ARIMA 组合模型的年径流预测

赵文举¹, 刘 茜¹, 李宗礼^{1,2}, 王亚丽¹

(1. 兰州理工大学 能源与动力工程学院, 兰州 730050)

(2. 水利部水利水电规划设计总院, 北京 100120)

摘 要: 为解决单一的小波神经网络预测精度不高的问题, 提出一种新的基于小波去噪和 WNN-ARIMA 组合模型, 应用小波阈值去噪法对小波神经网络的输入值进行预处理, 同时对模型残差值进行 ARIMA 模型修正. 利用该组合模型对洮河流域下巴沟站年径流量进行预测, 预测趋势和预测值与原始实测数据吻合度高, 表明此组合模型可靠性强, 可以有效预测年径流量, 以期对洮河流域和其他流域的年径流量预测提供新方法, 为水利工程建设和水资源优化配置提供依据.

关键词: 小波神经网络; 小波消噪; ARIMA 时序模型; 组合预测模型

1 引言

流域年径流量预测对水资源优化配置和水利工程运行管理等有着重要意义. 由于受自然气候、流域地质因素和人类活动等的综合影响, 流域年径流变化趋势呈现出非线性、突变性和随机性等特点^[1]. 为此, 径流预测一直是国内外研究的热点和难点.

目前应用于径流预测的常用模型主要有神经网络、ARIMA 时序预测、灰色预测^[1-3]. Ouma 等^[4] 通过研究长短期记忆神经网络 (LSTM) 和小波神经网络 (WNN) 在流域径流时序预测中的应用, 得出 LSTM 更适合于气象和水文监测网络匮乏流域的径流预测. 然而单一的预测模型往往存在弊端, 近年来, 众多学者由以前的研究单一预测模型转向研究其组合模型. 刘育等^[5] 运用小波去噪法对径流序列进行数据预处理后, 利用 FA-SVM 模型与 BP 神经网络模型进行了中长期径流预报. Wang 等^[6] 采用 ARIMA 模型与集合经验模态分解 (EEMD) 相结合的方法对年径流时间序列进行预测, 发现 EEMD-ARIMA 组合模型可有效的提高径流预测精度. 随着组合模型的发展, 径流预测精度得到了极大的改进^[7-9], 但往往部分组合模型因追求预测精度而操作过于繁琐, 部分则不适合于人类活动因素过高产生的时序突变导致预测不准确的情况.

针对目前常用径流预测模型对突变性较强的水文时间序列资料的预测能力不足, 本文利用小波降噪对小波神经网络的输入值进行预处理, 并利用 ARIMA 模型对小波神经网络进行误差修正, 提出了一种基于小波去噪的 WNN-ARIMA 组合模型, 对洮河流域下巴沟站年径流量进行预测, 验证该组合模型的可靠性. 以期对流域水资源未来变化趋势预测提供科学参考, 对区域水资源合理配置和生态可持续发展具有重要指导意义.

收稿日期: 2021-06-28

资助项目: 国家自然科学基金 (51869010); 兰州理工大学红柳学科资助项目

2 研究方法

2.1 小波神经网络预报模型

小波神经网络 (wavelet neural network, WNN), 是将小波分析和人工神经网络结合的改进的误差反向传播神经网络. 小波神经网络一般由输入层、隐含层和输出层三层结构组成, 其隐含层的激活函数为小波函数, 输入层至隐含层的权值和隐含层的阈值为小波函数的伸缩因子和平移因子^[10-11], 其网络结构如图 1 所示.

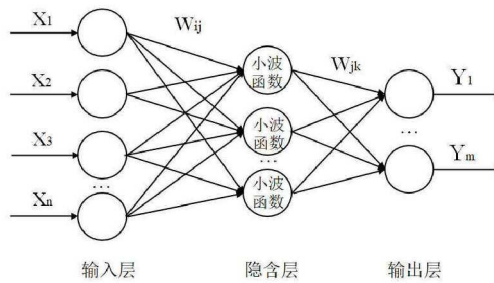


图 1 小波神经网络结构示意图

小波神经网络输入层序列为 $x_i (i = 1, 2, 3 \cdots n)$, 则其隐含层输出值可表示为:

$$\omega(j) = \omega_j \left(\frac{\sum_{i=1}^k W_{ij} x_i - b_j}{a_j} \right) \quad (j = 1, 2, \cdots, l) \quad (1)$$

式中, $\omega(j)$ 为隐含层中 j 节点的输出值; ω_j 为隐含层小波函数; W_{ij} 为输入层与隐含层各节点的连接权值; a_j 为小波函数伸缩因子; b_j 为小波函数平移因子.

本文小波神经网络隐含层小波基函数选取 Morlet 小波函数, 其表达式为:

$$\omega_j = \cos(1.75x) e^{-\frac{x^2}{2}} \quad (2)$$

其输出层计算公式为:

$$Y(k) = \sum_{j=1}^l W_{jk} h(j) \quad (k = 1, 2, \cdots, m) \quad (3)$$

式中, $Y(k)$ 为输出层中 k 节点的输出层; W_{jk} 为隐含层与输出层各节点的连接权值.

2.2 ARIMA 预报模型

ARIMA 模型 (Autoregressive Integrated Moving Average model), 差分整合移动平均自回归模型, 是将自回归模型 (AR)、移动平均模型 (MA) 和差分法结合的预测模型^[12].

其中, 自回归模型 (AR) 的表达式为:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-1} + \varepsilon_t \quad (4)$$

式中, μ 为常数项; p 为阶数; γ_i 为自相关系数; ε_t 为白噪声.

移动平均模型 (MA) 的表达式为:

$$y_t = \mu + \sum_{i=1}^q \theta_i \varepsilon_{t-1} + \varepsilon_t \quad (5)$$

式中, q 为阶数; θ_i 为回归系数.

由式 (4) 和式 (5) 得到 ARMA 模型表达式, 为:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-1} + \sum_{i=1}^q \theta_i \varepsilon_{t-1} + \varepsilon_t \quad (6)$$

于非平稳时间序列, 应使其先平稳化, 即对时间序列先进行 d 阶差分, 再利用式 (6) 进行模拟, 得到 ARIMA 模型, 记为 y'_t .

2.3 WNN-ARIMA 组合预报模型

利用小波神经网络模型的拟合序列 $Y = \{Y_1, Y_2, \dots, Y_t, \dots, Y_l\}$, 与原始的观测数据序列 $X = \{X_1, X_2, \dots, X_t, \dots, X_l\}$ 中的数据对应相减即可得到小波神经网络模型残差序列: $E = \{E_1, E_2, \dots, E_t, \dots, E_l\}$, 其中:

$$E_t = X_t - Y_t \quad (t = 1, 2, 3, \dots, l) \quad (7)$$

由式 (7) 可得:

$$X_t = Y_t + E_t \quad (t = 1, 2, 3, \dots, l) \quad (8)$$

以小波神经网络模型预测的残差序列 $E = \{E_1, E_2, \dots, E_t, \dots, E_l\}$ 为对象建立 ARIMA 模型, 设 ARIMA 模型预测值为 Y'_t , 则 WNN-ARIMA 组合预测模型为:

$$\hat{y}_t = Y_t + Y'_t \quad (9)$$

2.4 小波去噪原理

小波去噪 (wavelet denoising, WD) 是一种非线性去噪方法, 原始信号经小波分解后, 幅值较大的小波系数大多数为有用信号, 而幅值较小的系数一般都为噪声, 即可以认为有用信号的小波变换系数要大于噪声的小波变换系数^[5]. 本文选用小波阈值去噪法.

假设原始信号为:

$$F(t) = s(t) + e(t) \quad (10)$$

式中, $F(t)$ 为原始信号; $s(t)$ 为有用信号; $e(t)$ 为噪声.

对上式两边同时做小波变化得到:

$$WT_f(a, b) = WT_s(a, b) + WT_e(a, b) \quad (11)$$

式中, W 为小波基函数, a 为小波函数伸缩因子, b 为小波函数平移因子.

即实际测量信号的小波变换等于多个信号小波变换的和.

3 实例分析

为验证 WNN-ARIMA 模型的准确性和可靠性, 以洮河流域下巴沟站年径流量为例进行实例预测. 洮河位于甘肃省南部, 是黄河上游右岸第二大支流, 全长 673km, 多年平均河川径流量为 48.38 亿 m^3 . 洮河作为甘肃省中部干旱地区最重要的水源区, 其水资源的变化可牵动黄河流域的社会经济等的发展^[13]. 因此, 洮河径流量预测对优化调度各级水库、合理使用水资源以及水文信息的可靠性验证具有重要意义. 本文采用 WNN 模型、WD-WNN 模型和 WNN-ARIMA 组合模型三种模型对其年径流量进行预测, 并对比验证各模型精度.

3.1 小波神经网络年径流预报模型

利用 Matlab 建立小波神经网络, 对洮河流域下巴沟站年径流进行模拟预测. 将 1965–2016 年径流、输沙和降水资料分为两组, 将 1965–2011 年数据作为训练组, 2012–2016 年数据作为预测组. 其小波神经网络模型拟合结果如图 2 所示, 拟合值误差小于 20% 的有 36 年, 合格率为 76.6%. 2012–2016 年预测值见表 2, 5 年平均相对误差为 16.6%, 合格率为 60%. 其残差绝对平均值为 1.85 亿 m^3 , 平均相对误差为 14.0%. 小波神经网络模型的拟合结果与实际变化趋势大致相同, 但有部分年份数据差距较大, 且预测值与实际值差距较大, 不能良好的反应径流量未来趋势, 说明小波神经网络模型预测精度较低, 需对该模型进行修正.

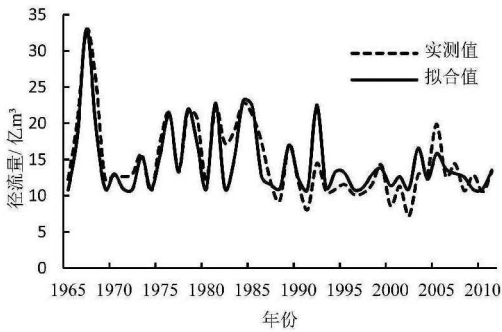


图 2 小波神经网络结构示意图

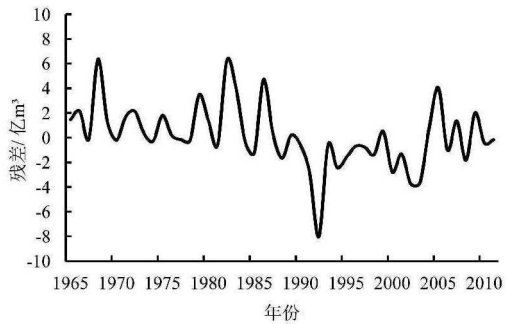
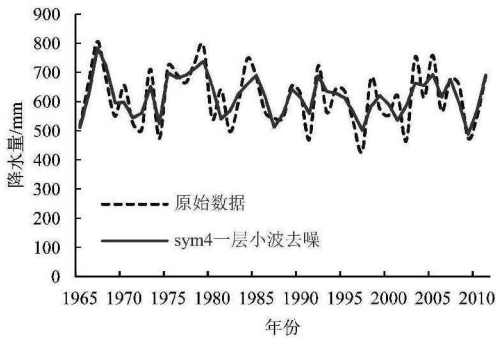
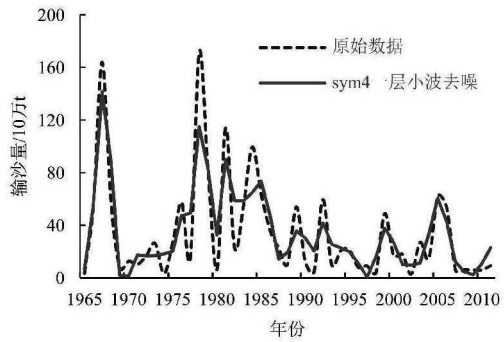


图 3 小波神经网络模型残差序列

3.2 WD-WNN 年径流量模型预测



(a) 年降水量序列小波去噪



(b) 年输沙量序列小波去噪

图 4 输入值小波去噪

由于小波神经网络是根据输入值和输出值之间的关系进行预测, 则输入值的平稳性对预测结果的影响巨大, 输入值的突变会导致拟合值与预测值的突变, 为此利用小波消噪对输入值进行去噪处理.

采用小波阈值法对输入年降水量和输沙量进行去噪, 选用 sym4 小波进行 1 层分解. 将图 4 中去噪后的年降水量和输沙量作为输入值重新建立小波神经网络, 其拟合结果见图 5.

利用小波消噪改进的小波神经网络预测模型拟合平均残差值为 1.92 亿 m^3 , 平均相对误差为 13.8%. 虽然其平均残差值有所增加, 但相对误差有一定的减少, 拟合合格率达到 83.0%, 较未消噪模型提高了 6.4%. 预测年份合格率达到 80%, 平均相对误差降低至 15.15%, 且预测

趋势与实际趋势基本一致.

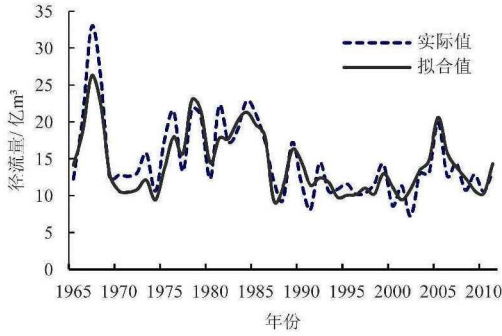


图5 WD-WNN 年径流量拟合

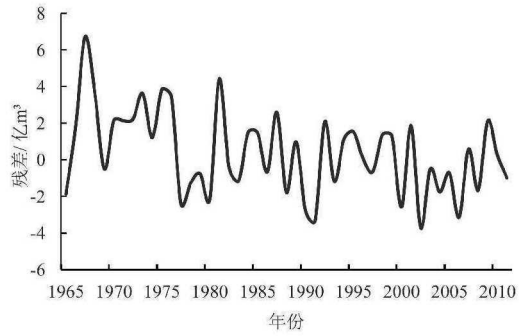


图6 WD-WNN 模型残差序列

表1 几种模型拟合精度对比

模型	残差绝对平均值 (亿 m ³)	相对误差绝对值 (%)	拟合合格率 (%)
WNN 模型	1.85	14.0	76.6
WD-WNN 模型	1.92	13.8	83.0
WNN-ARIMA 组合模型	1.22	7.8	95.7

3.3 基于 ARIMA 模型的预报残差修正

为进一步提高预测精度, 选择 ARIMA 模型进行残差修正. 根据 1965-2011 年小波去噪改进模型的残差序列, 利用 SPSS 软件进行 ARIMA 模型拟合. 经过 ARIMA 修正的预测模型拟合值与实测值高度吻合, 明显优于 WNN 模型和 WD-WNN 模型的拟合结果, 如图 7 所示, 拟合精度较其他两模型分别提高了 4.83% 和 3.34%, 其精度可达到 92.2%, 拟合合格率达 95.7%. 5 年预测值中, 仅 2012 年相对误差超过 20%, 其余四年皆为合格年, 预测精度达到 88.9%. WNN-ARIMA 组合模型可以更好的反应实测值趋势, 且在预测精度上有显著优势, 综合而言, WNN-ARIMA 模型更适合于流域年径流预测.

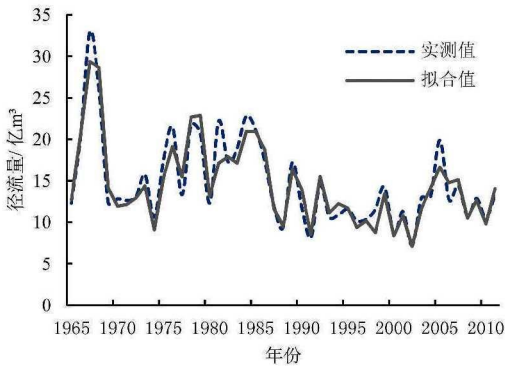


图7 WNN-ARIMA 组合模型拟合结果

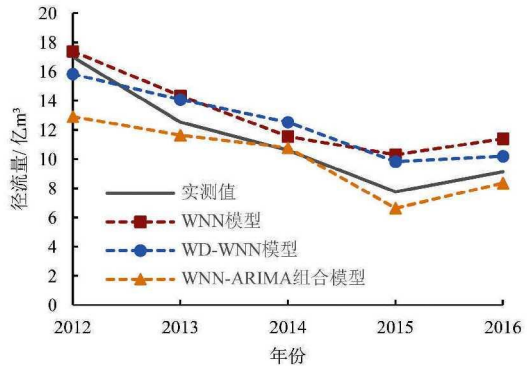


图8 各模型预测值与实测值对比图

表 2 几种模型预测结果与精度分析

模型	年份	实测年径流量 (亿 m^3)	预测年径流量 (亿 m^3)	相对误差绝对值 (%)	拟合合格率 (%)
WNN 模型	2012	17.000	17.379	2.23	60
	2013	12.530	14.332	14.38	
	2014	10.600	11.555	9.01	
	2015	7.758	10.291	32.65	
	2016	9.126	11.403	24.95	
	平均值	11.403	12.296	16.64	
WD-WNN 模型	2012	17.000	15.826	6.91	80
	2013	12.530	14.078	12.35	
	2014	10.600	12.529	18.20	
	2015	7.758	9.823	26.62	
	2016	9.126	10.191	11.67	
	平均值	11.403	12.296	16.64	
WNN-ARIMA 组合模型	2012	17.000	12.906	24.08	80
	2013	12.530	11.628	7.20	
	2014	10.600	11.779	11.12	
	2015	7.758	6.633	14.50	
	2016	9.126	9.321	2.14	
	平均值	11.403	10.453	11.81	

4 结论

小波神经网络模型在流域径流预测过程中由于预测值受输入值影响较大, 在输入值不稳定情况下, 会存在个别点预测精度不高的缺点. 为了能避免输入值突变, 提高模型预测精度, 本文在小波神经网络模型, 小波阈值消噪和 ARIMA 模型基础上提出了一种新型基于 ARIMA 残差修正的小波神经网络组合模型. 利用此模型对 1965-2011 年洮河流域下巴沟站径流数据进行了拟合, 结果表明模型拟合合格率达到 95.7%, 并预测了 2012-2016 年的径流量, 模型预测合格率达到 80%, 平均相对误差 11.81%, 与单一的预测模型相比具有更好的预测精度. 为此, WNN-ARIMA 组合模型可有效预测洮河流域年径流量, 同时为其他流域的年径流量预测提供新思路.

参考文献

- [1] 包丽娜, 唐德善, 胡晓波, 等. 基于小波分解及 Arima 误差修正的径流预测模型及应用 [J]. 长江科学院院报, 2018, 35(12): 18-21,33.
- [2] 王筱, 周维博. 基于优化的灰色-权马尔科夫模型的径流量预测 [J]. 数学的实践与认识, 2019, 49(22): 179-186.
- [3] 张建兴, 马孝义, 赵文举, 等. 生命轮回 -Markov 组合模型在年径流预报中的应用 [J]. 水力发电学报,

- 2008, 27(6): 32-36.
- [4] Ouma Y O, Cheruyot R, Wachera A N. Rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network with satellite-based meteorological data: case study of Nzoia hydrologic basin[J]. *Complex & Intelligent Systems*, 2021: 1-24.
- [5] 晋健, 刘育, 王琴慧, 等. 基于小波去噪和 FA-SVM 的中长期径流预报 [J]. *人民长江*, 2020, 51(9): 67-72.
- [6] Wang W C, Chau K W, Xu D W, et al. Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA Based on EEMD Decomposition[J]. *Water Resources Management*, 2015, 29(8): 2655-2675.
- [7] 赵泽谦, 黄强, 明波, 等. 基于多模型随机组合的水文集合预报方法研究 [J]. *水力发电学报*, 2021, 40(1): 76-87.
- [8] Zhang G P. Time series forecasting using a hybrid ARIMA and neural network model [J]. *Neuro-computing*, 2003, 50(1): 159-175.
- [9] 彭欣怡, 于国荣, 张代青. 不同组合小波神经网络模型对径流预测的适用性 [J]. *人民长江*, 2015, 46(24): 24-28.
- [10] 赵文举, 马孝义, 李军利, 等. 灰色时序组合模型及其在地下水埋深预测中的应用 [J]. *数学的实践与认识*, 2008, 48(18): 70-76.
- [11] 贾昊, 穆兴民, 赵广举, 等. 不同频率黄河上中游径流量变化特征及其趋势预测 [J]. *水土保持学报*, 2020, 34(6): 60-64+69.
- [12] 韩永贵, 韩磊, 黄晓宇, 等. 基于指数平滑和 ARIMA 模型的西北地区饱和和水汽压差预测 [J]. *干旱区研究*, 2021, 38(2): 303-313.
- [13] 牛最荣, 王启优, 孙栋元, 等. 基于径流还现的洮河流域径流变化特征研究 [J]. *干旱区地理*, 2021, 44(1): 149-157.

Annual Runoff Prediction Based on Wavelet Denoising and WNN-ARIMA Combined Model

ZHAO Wen-ju¹, LIU Xi¹, LI Zong-li^{1, 2}, WANG Ya-li¹

(1. College of Energy and Power Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

(2. General Institute for Water Resources and Hydropower Planning and Design, Ministry of Water Resources, Beijing 100120, China)

Abstract: In order to solve the problem of low prediction accuracy of a single wavelet neural network, this paper proposed a new combined model based on wavelet denoising and WNN-ARIMA. The wavelet threshold denoising method is used to preprocess the input value of the wavelet neural network, and the residual values of the model are corrected by the ARIMA model. The combined model is used to predict the annual runoff at the Xiabagou Station in the Taohe River Basin and the predicted trend and values are in good agreement with the original measured datas, which show that the combined model is highly reliable and can effectively predict the annual runoff. This combined model is expected to provide a new method for the annual runoff prediction of the Taohe River Basin and other basins, and provide a basis for the construction of water conservancy projects and the optimal allocation of water resources.

Keywords: wavelet neural network; wavelet denoising; ARIMA time series model; combined forecasting model