

基于 VMD-LSTM 的淮河上游月径流预测应用研究

鲁志杰, 吴琪

淮河水利委员会水文局(信息中心), 蚌埠市, 安徽省, 中国

摘要: 针对淮河流域水资源监测体系不完善, 径流量预测精度低、水资源管理与调配智能化程度无法满足“四预”业务应用需求等问题, 基于信号分解方法与长短期记忆神经网络构建一种组合模型(VMD-LSTM)。该模型在淮河王家坝以上流域重要断面径流预测应用中, 相比 BP、LSTM 神经网络模型的预测结果, 在均方根误差(RRMSE)、平均绝对误差(MMAE)以及确定性系数(DDC)评价指标上均有显著提升。该模型能够有效解决径流序列的非平稳特性, 学习长期依赖性, 提高月径流预测精度, 为流域水资源管理与调配提供科学、智能的决策支持。

关键词: 径流预测; 变分模态分解; 长短期记忆神经网络

Application Research on Monthly Runoff

Prediction of the Huai River Based on VMD-LSTM

Zhijie Lu, Qi Wu

Hydrological bureau (Information Center) of Huaihe River Commission, Bengbu City, Anhui Province, China

Abstract: A combined model (VMD-LSTM) is constructed based on Variational Mode Decomposition (VMD) and Long Short Term Memory Neural Network (LSTM) to address the problems of incomplete water resources monitoring system, low accuracy of runoff prediction, and insufficient intelligence in water resources management and allocation that cannot meet the requirements of the "Four Predictions" business application in the Huaihe River Basin. In the application of this model for important section runoff prediction in the Wangjiaba and above watershed of the Huai River, compared with the prediction results of BP and LSTM neural network models, there is a significant improvement in the evaluation indicators of root mean square error (RRMSE), mean absolute error (MMAE), and coefficient of certainty (DDC). This model can effectively address the non-stationary characteristics of runoff sequences, learn long-term dependencies, improve monthly runoff prediction accuracy, and provide scientific and intelligent decision support for watershed water resource management and allocation.

Keywords: Runoff prediction; Variational mode decomposition; Long Short Term Memory Neural Network

1 引言

对照智慧水利建设目标和要求, 淮河流域目前尚存在水资源监测体系不完善, 感知覆盖范围不全面, 信息传输不及时等问题, 水资源管理与调配智能化程度, 尚无法满足“四预”业务应用需求[1]。

径流预测是水量分配和水资源调度管理的基础, 精确的径流预测能为淮河流域加强水资源统一调度管理, 规范调度行为等提供重要依据[2]。

近年来, 许多学者的研究表明实测的径流序列存在显著的非平稳性特征, 导致现有预测模型与方法的适用性面临巨大挑战[3, 4]。月径流量序列的非平稳性, 造成使用单一模型进行月径流量预测的时候, 不能充分识别序列的非平稳性特征[5]。信号分解可以将径流量序列分解为多个相对平稳的子序列, 使得预测模型能够更好地捕捉序列非线性特征[6, 7]。信号分解在径流预测问题上有很多应用, 邵骏等[8]基于 Bayes 框架和信号分解构建了径流预报模型, 结果表明, 信号分解

能有效提高预报模型的适应性。

长短期记忆神经网络 (LSTM) 在循环神经网络的基础上, 增加了门控制单元结构, 对于长期依赖性的时间序列预测有较好的应用效果[9]。胡庆芳等[10]在汉江流域基于长短期记忆神经网络构建日径流序列预测方法, 结果表明长短期记忆神经网络模型对于径流序列峰值的拟合更加精确。

为提升中长期径流预测的精度, 适应水文序列的非线性变化, 通过结合信号分解与神经网络方法, 构建基于 VMD-LSTM 的组合预测模型。将该模型应用于淮河王家坝以上流域重要断面径流预测, VMD 用于分解原始径流序列, LSTM 可以捕捉子序列的长期依赖关系, 在具备更好适应性的同时提高径流预测的精度。

2 基于 VMD-LSTM 的月径流预测模型

2.1 变分模态分解

Dragomiretskiy 等[11]提出了变分模态分解 (VMD), 其利用非递归和变分法能够将水文序列分解为多个相对稳定的分量。

VMD 将非线性序列 $h(t)$ 分解为 r 个分量序列具体步骤为: ①通过希尔伯特变换对各分量 $h_r(t)$ 相应的解析信号进行计算; ②调整各分量 $h_r(t)$ 到对应的基频带; ③各分量 $h_r(t)$ 通过估计频率带宽, 得到:

$$\min_{\{h_r\}, \{Z_r\}} \left\{ \sum_r \left\| \partial_t [(g(t) + j/\pi) h_r(t)] e^{-jZ_r t} \right\|_2^2 \right\} \quad (1)$$

$$\text{s.t. } \sum_r h_r = f \quad (2)$$

式中, h_r 为第 r 个模态函数; Z_r 是与模态函数对应为中心频率; $g(t)$ 是 Dirac 分布, 为复平面上模态函数中心频率的相量描述。文献[11]给出了 VMD 分解详细的计算方法。

2.2 长短期记忆神经网络

长短期记忆神经网络 (LSTM) 在循环神经网络的基础上增加了门控制单元结构[12, 13], 其针对循环神经网络在具有长程依赖关系的时间序列预测上的问题进行了优化。与循环神经网络相比, 增加了内部自循环, 可关联长时间信息, 学习远距离依赖关系, 达到优化数据, 提高模型性能的效果。一个 LSTM 单元由遗忘门 F_t 、输入门

I_t 和输出门 O_t 组成, 其中 I_t 判断上层输入是否进入隐藏层, O_t 选择该单元的输出是否进入下一单元, F_t 表示是否将该单元的历史信息进行存储。 C_{t-1} 、 C_{t-2} 分别为前一单元和本单元的信息状态, σ 为 sigmoid 函数。LSTM 单元结构见图 1。

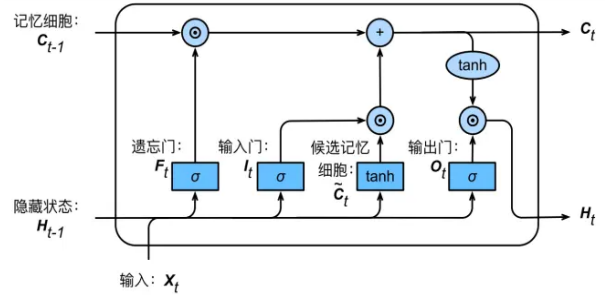


图 1. LSTM 模型单元结构图

2.3 VMD-LSTM 月径流预测模型

河川径流的形成过程机理复杂, 从降雨到形成出口断面流量的径流过程受多重因素影响, 单一影响因子序列作为 LSTM 模型输入的拟合效果不好, 很难达到满意的训练效果。因此, 通过结合信号分解与神经网络方法, 构建基于 VMD-LSTM 的组合预测模型。该模型的径流预测步骤为:

- 1) 利用 VMD 将月径流序列分解为多个子序列;
- 2) 对分解后的多个子序列进行归一化处理;
- 3) 模型输入: 每个子序列的前 1 年 (12 个月) 径流量。

4) 模型输出: 下一个月的径流量。

模型输入输出如下式所示:

$$\text{input} = [S1_{(t-12,t)}, S2_{(t-12,t)}, \dots, Sn_{(t-12,t)}, R_{(t-12,t)}] \quad (3)$$

$$\text{output} = L_{t+1} \quad (4)$$

式中, $t > 12$; L 为原始月径流序列。

4) 将选择好的输入输出放入 LSTM 模型训练, 进行预测, 其流程见图 2 所示。

2.4 模型评价指标

采用《水文情报预报规范》[14]中的均方误差 (RRMSE)、平均绝对误差 (MMAE) 及确定性系数 (DDC) 作为评价指标, 对 VMD-LSTM 组合模型的预测精度进行评估。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_0)^2} \quad (5)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y_0| \quad (6)$$

$$DC = 1 - \frac{\sum_{i=1}^n [y_i - y_0]^2}{\sum_{i=1}^n [y_0 - y]^2} \quad (7)$$

式中, y_i 为 i 时刻的预测径流量; y_0 为 i 时刻的实测径流量; y 为径流量序列的均值。

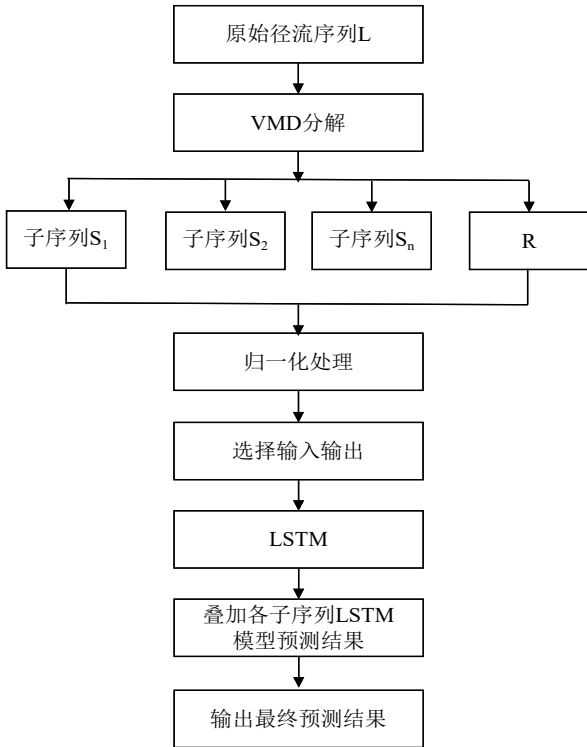


图2. VMD-LSTM 模型流程图

3 讨论

3.1 资料情况

淮河干流洪河口以上为上游, 本次研究选取淮河王家坝以上干支流4个主要控制站作为代表站进行月径流预测应用研究, 4个代表站分别为淮河干流息县站、王家坝站, 淮北支流洪汝河班台站, 淮南支流潢河潢川站。站点资料系列采用1980~2020年的492个实测月径流资料, 模型训练集选择各站前350个月数据, 其余作为测试集。淮河上游站点分布见图3。

3.2 径流序列 VMD 分解

为提高 VMD-LSTM 模型训练效率和拟合精度, 利用变分模态分解 (VMD) 方法对月径流量序列进行分解, 充分识别序列非线性特征。对月径流量序列进行信号分解, 确定 VMD 模态数 r 。通过多次预实验, 发现 VMD 分解层数大于 8 时,

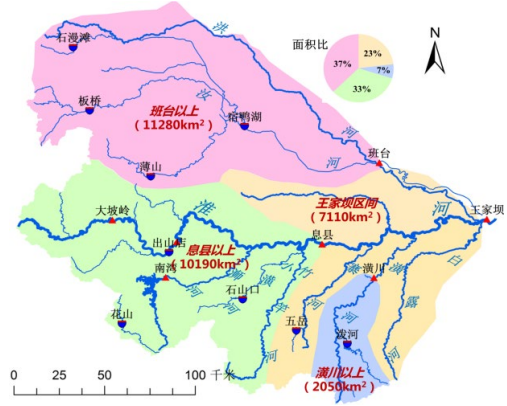


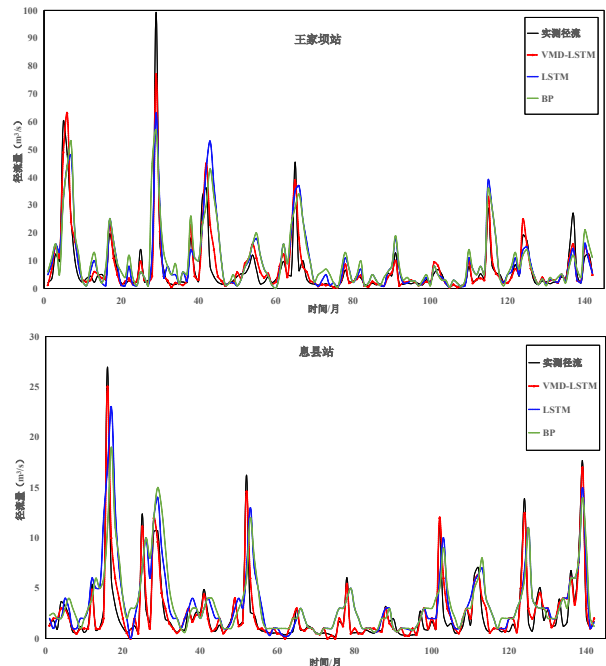
图3. 淮河上游站点分布图

中心频率开始出现混叠现象, 即不同频率的尺度未完全分离 [15], 隐藏设置 $r=8$, 即通过 VMD 分解将 1980-2020 年月径流序列分解为 9 个有限带宽的子序列。

3.3 对比试验及结果

预报因子的针对性选择是影响径流预测模型的重要因素, 选取前 12 个月径流序列作为模型主要输入, 为提高模型预测精度, 根据关联性分析同时选择上游水文站前 6 个月径流数据及区间预测面雨量等因子作为输入, 并作归一化处理。对 VMD 分解好的子序列分别建立神经网络模型, 最后在叠加各项预测值的基础上得到月径流量预测值。

在相同的预测流程下选择单一神经网络模型 (BP、LSTM) 和 VMD-LSTM 组合模型进行对比分析, 利用训练集的数据训练模型, 用测试集验证模型, 不同模型的测试集预测结果见图 4。



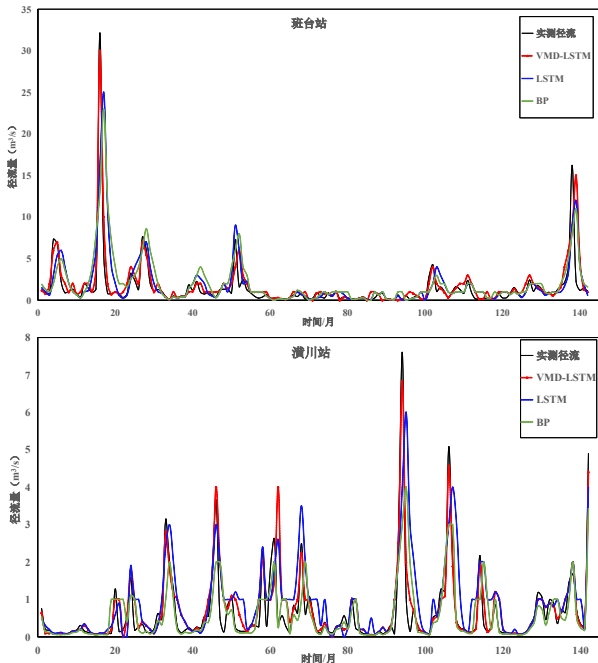


图4 淮河上游4站不同模型测试集预测结果图

由表1不同模型的统计参数可知，VMD-LSTM组合模型相比于BP、LSTM单一的预测模型在各项评价指标上均有显著提升，与其他两种模型结果相比，径流预测过程与实测过程拟合度最好。而且组合模型对峰值的拟合更为准确，说明对原始径流序列进行合理分解能有效提高预测精度，同时通过VMD分解为径流子序列，进一步提高了月径流预测模型的精度。

表1 3种模型预测结果统计对比

站点	模型	训练期			检验期		
		R_{RMS}	M_{MA}	D_{DC}	R_{RMS}	M_{MA}	D_{DC}
息县	BP	5.63	2.63	0.79	6.05	2.93	0.76
	LSTM	4.52	2.18	0.86	4.91	2.34	0.84
	VMD-LST	3.49	1.96	0.90	3.78	2.05	0.89
王家坝	BP	5.53	2.59	0.81	6.02	2.87	0.77
	LSTM	4.47	2.15	0.88	4.88	2.31	0.85
	VMD-LST	3.46	1.93	0.92	3.75	2.02	0.91
班台	BP	5.57	2.53	0.82	5.82	2.77	0.78
	LSTM	4.41	2.10	0.89	4.76	2.26	0.86
	VMD-LST	3.39	1.88	0.93	3.68	1.95	0.92
潢川	BP	5.67	2.66	0.78	6.07	2.96	0.75
	LSTM	4.56	2.12	0.85	4.93	2.35	0.83
	VMD-LST	3.52	1.98	0.89	3.80	2.06	0.88

由图5散点图可知，单一神经网络模型（BP、LSTM）预测径流量点据分散，VMD-LSTM组合模型预测径流量点据集中在45°线附近，最接近实测值。

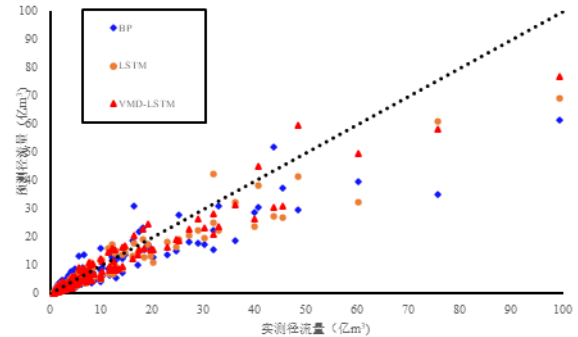


图5 几种模型预测与实测月径流散点图

由此可知，单一神经网络预测模型中，LSTM相比于BP模型有更好的应用效果；基于信号分解组合模型能充分识别月径流量序列局部特征，从而提高模型预测精度。本次研究构建的VMD-LSTM组合预测模型能同时发挥信号分解（VMD）和神经网络（LSTM）的双重优势，表现出明显优于单一神经网络模型的预测性能和优势，可为提高月径流量的预测精度提供新的思路 and 参考。

4 结论

为了提高月径流预测的精度，有效识别月径流序列中的局部特征信息，本文提出了基于VMD-LSTM的组合预测模型，并应用于淮河王家坝以上流域重要断面月径流预测，得出的结论见下：

(1) 对于月径流序列的非线性特征，通过VMD将原始径流资料分解为多个子序列，为模型输入提供了较好的数据基础。

(2) VMD-LSTM组合预测模型较于单一模型能更好地提取序列局部特性，且对于序列峰值的拟合更为精确。本文将VMD-LSTM模型应用于淮河王家坝以上流域重要断面月径流预测，相较于单一的BP、LSTM模型能取得更高的精度，未来研究可进一步应用到全流域旬、月、年等不同尺度的径流预测，为淮河的径流预测提供了一种新的选择。

致谢

本文的研究工作得到水利部重大科技项目

(SKR-2022031, SKS-2022013, SKR-2022032) 的支持, 谨此致谢!

参考文献

- [1] 王凯,徐时进,刘昌军,等.数字孪生淮河防洪“四预”系统关键技术[J].水利信息化,2023,(06):1-4+31.
- [2] 常春晓,周宇,吴晨晨.淮河流域水资源调度实践与经验[J].治淮,2023,(09):4-6+9.
- [3] 秦毅,李时.应对水文序列非一致性变化影响的溯源重构法研究[J].水利学报,2021,52(07):807-818.
- [4] 谢平,陈广才,夏军.变化环境下非一致性年径流序列的水文频率计算原理[J].武汉大学学报(工学版), 2005, (06):6-9+15.
- [5] 杨丽洁.数据驱动模型在洪水预报中的应用及其发展趋势[J].电脑知识与技术,2018,14(17):275-277.
- [6] 张洪波,余荧皓,孙文博,等.面向 EMD 分解的径流分量重构方法对比研究[J].南水北调与水利科技, 2017, 15(01):60-66+166.
- [7] 孙国梁,李保健,徐冬梅,等.基于 VMD-SSA-LSTM 的月径流预测模型及应用[J].水电能源科学,2022,40(05): 18-21.
- [8] 邵骏,袁鹏,张文江,等.基于贝叶斯框架的 LS-SVM 中长期径流预报模型研究[J].水力发电学报,2010,29(05): 178-182.
- [9] 陶思铭,梁忠民,陈在妮,等.长短期记忆网络在中长期径流预报中的应用[J].武汉大学学报(工学版), 2021, 54(01):21-27.
- [10] 胡庆芳,曹士圻,杨辉斌,等.汉江流域安康站日径流预测的 LSTM 模型初步研究[J].地理科学进展,2020,39(4): 636-642.
- [11] DRAGOMIRETSKIY K,ZOSSO D. Variational Mode Decomposition[J]. IEEE Transactions on Signal Processing,2014,62(3):531-544.
- [12] 孙望良,周建中,彭利鸿,等.DFA_VMD_LSTM 组合日径流预测模型研究[J].水电能源科学,2021,39(03): 12-15.
- [13] HOCHREITER S, SCHMIDHUBER JURGEN. Long short-term memory[J]. Neural computation,1997,9(8): 1735-1780.
- [14] 水利部水文局.水文情报预报规范: GB/T22482-2008[S]. 北京:中国标准出版社,2008.
- [15] 付华,刘超,张松,等.VMD-SWT 联合算法在故障行波检测中的应用[J].电力系统及其自动化学报,2020, 32(05):55-61.

Copyright © 2024 by author(s) and Global Science Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access