

文章编号:1671-6833(2024)06-0032-08

基于二次分解和支持向量机的月径流预测方法

甘容^{1,2}, 马超鑫^{1,2}, 高勇³, 郭林³, 侯晓丽⁴, 路学永⁵

(1. 郑州大学 水利与交通学院, 河南 郑州 450001; 2. 河南省地下水污染防治与修复重点实验室, 河南 郑州 450001; 3. 河南省地质研究院, 河南 郑州 450001; 4. 河南省豫东水利保障中心, 河南 开封 475000; 5. 中国南水北调集团中线有限公司渠首分公司, 河南 南阳 473000)

摘要: 针对径流序列的非线性和非平稳性特征, 提出了一种基于加权回归的季节趋势分解(STL)和变分模态分解(VMD)组合的二次分解, 结合支持向量机(SVM)的月径流预测模型 STL-VMD-SVM。该模型利用 STL 将原始径流序列分解为不同频率的季节项、趋势项和残差项, 并通过 VMD 将残差项分解为 IMFs。建立 SVM 模型预测季节项、趋势项和 IMFs, 所有 IMFs 的预测值之和为残差项的预测值, 季节项、趋势项和残差项之积为原始径流序列的最终预测值。基于伊洛河流域黑石关站及黄河干流高村站的月径流时间序列进行了实例应用及普适性评价, 并与 BP 神经网络模型和长短期记忆神经网络模型(LSTM)进行对比。结果表明: 对于伊洛河黑石关站径流预测, 所提模型验证期的 NSE 、 $MAPE$ 、 $RMSE$ 、 R 分别为 0.977, 13.705%, 0.327, 0.991, 其预测精度均优于单一模型和一次分解模型, STL-VMD 二次分解可以有效提高模型预测精度; 在黄河干流高村站径流预测中验证期的 NSE 、 $MAPE$ 、 $RMSE$ 、 R 分别为 0.979, 8.509%, 3.263, 0.989, 也达到了很好的预测效果。

关键词: 月径流预测; 二次分解; STL; VMD; SVM; 神经网络

中图分类号: P338; TV121

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2024.06.003

近年来, 旱涝极端水文事件给各国带来了巨大的经济损失和人员伤亡^[1], 其重要原因之一是缺乏准确的旱涝预警信息。另外, 随着人类活动的加剧^[2], 流域径流受到间歇性径流、地下水过度开采和水质恶化等问题的困扰, 对流域生态环境和可持续发展带来了巨大挑战^[3]。因此, 探究径流演变模式并进行精准预测势在必行。

由于径流序列预测周期长, 影响因素多, 且具有较强的非线性和非平稳性, 使得传统方法或是机器学习对于原始序列的预测效果不佳。通过适当的分解技术对原始数据进行预处理, 可以过滤干扰和混合噪声, 提高模型预测精度^[4]。雷庆文等^[5]使用基于加权回归的季节趋势分解(seasonal-trend decomposition using loess, STL)和 CNN(convolutional neural networks)-LSTM(long short term memory)的组合模

型对月径流量进行了预测, 发现经过分解预处理的数据预测效果显著提高。

一次分解技术虽然可以提高模型的预测能力, 但其对数据的随机性、规则性和非平稳性不能很好地处理, 特别是对数据序列中较大异常值的预测效果不够好。因此, 有学者提出二次分解技术。王文川等^[6]基于 CEEMD(complementary ensemble empirical mode decomposition)-VMD(variational mode decomposition)的组合分解方法, 结合 BP 神经网络(back propagation neural network)模型研究了月尺度径流量的预测, 发现加入 VMD 方法后的二次分解模型的预测精度高于一次分解模型。除了分解方法的选择及其不同的组合, 使用不同的基础预测模型也会对径流序列的预测精度产生影响。Yang 等^[7]采用基于二次分解的回声状态网络(echo state net-

收稿日期: 2024-05-10; 修订日期: 2024-06-05

基金项目: 河南省重点研发与推广专项(232102320026, 232102320032); 河南省自然资源厅科研项目(202361001); 国家自然科学基金资助项目(51509222, 51909091)

作者简介: 甘容(1983—), 女, 湖北随州人, 郑州大学副教授, 博士, 主要从事水文学及水资源研究, E-mail: ganrong168@163.com。

通信作者: 郭林(1987—), 男, 河南南阳人, 河南省地质研究院高级工程师, 主要从事水-工-环研究, E-mail: guolin_cug@126.com。

引用本文: 甘容, 马超鑫, 高勇, 等. 基于二次分解和支持向量机的月径流预测方法[J]. 郑州大学学报(工学版), 2024, 45(6): 32-39. (GAN R, MA C X, GAO Y, et al. Monthly runoff prediction method based on secondary decomposition and support vector machine[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(6): 32-39.)

work,ESN)研究日径流预测,其 3 个水文站点径流预测的 NSE 均在 0.99 以上。

本研究基于二次分解技术的思路,针对径流序列的季节特征,采用 STL-VMD 的二次分解方法,与支持向量机 (support vector machine, SVM) 相结合组成 STL-VMD-SVM 模型,开展伊洛河流域黑石关站月径流预测,并在黄河干流高村站验证该模型的普适性。

1 研究区概况及数据

伊洛河流域面积约为 18 800 km²,跨越陕西和河南两省 21 个县市^[8],流域上游以土石山区为主,下游以低山丘陵、河谷平原为主,分别占流域总面积的约 50%,40%和 10%。伊洛河地处湿润半湿润地区,是黄河流域重要的水源涵养区之一,流域年平均气温为 12~14 ℃,年平均降雨量约为 680 mm,降雨主要集中在 5—10 月份,占全年降水量的 80%以上,年平均日照时数为 2 291.6 h^[9]。图 1 为所绘制的伊洛河流域分布概况图。

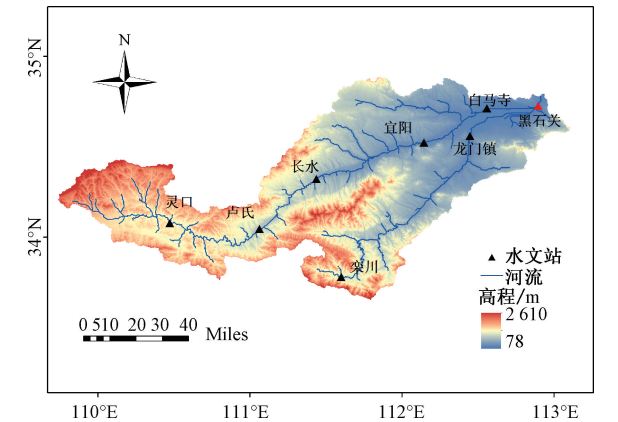


图 1 伊洛河流域概况图

Figure 1 Overview of the Yiluo River Basin

本文以黑石关水文站 1962 年 1 月至 2022 年 12 月实测月径流量资料为研究对象,其中 1962 年 1 月至 2012 年 12 月为训练期,2013 年 1 月至 2022 年 12 月为验证期。

2 研究方法

2.1 自相关分析

前期历史径流输入数的选取会直接影响模型预测结果的准确度。输入量太少,会造成模型模拟准

确度不高;输入量太多,会导致模型模拟过拟合^[10]。通过对径流时间序列的自相关分析可以粗略判定时间序列中与下一时刻的观测值具有较强相关关系的历史观测数据个数^[11]。因此,本文运用 SPSS 软件,对径流序列进行自相关分析,ACF 为自相关值,结果如图 2 所示。

由图 2 可知,某一径流量与其前推第 1,2,12 个相关性最大,受其影响最深。径流预测需要一定的历史资料数据,本文选择输入历史资料个数为 12(即前 12 个数据作为输入,第 13 个数据作为输出)。

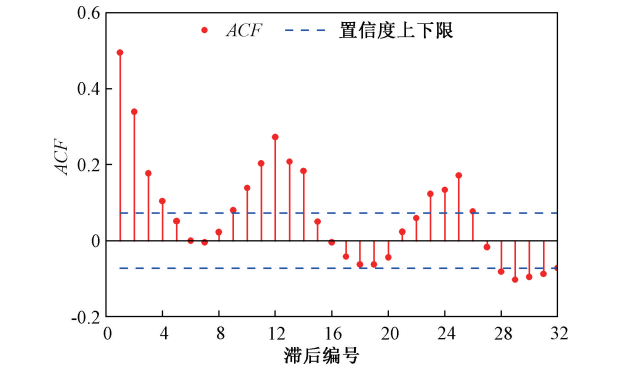


图 2 黑石关站实测月径流序列自相关性分析

Figure 2 Autocorrelation analysis of monthly runoff series measured at Heishiguan station

2.2 STL

基于局部多项式拟合回归的 STL 计算效率高,对趋势项和季节项的鲁棒性估计能力强^[12]。其有加法和乘法 2 种模式^[7],分解模式分别为

$$Y_t = S_t + T_t + R_t; \tag{1}$$

$$Y_t = S_t T_t R_t. \tag{2}$$

式中: Y_t 为原始时间序列; S_t 为季节分量; T_t 为趋势分量; R_t 为残差分量。

分别使用加法和乘法建立组合模型,得到的预测精度如表 1 所示,虽然加法模型的 $RMSE$ 比乘法模型小,但是其 $MAPE$ 过大, NSE 也不如乘法模型高,所以认为整体预测精度不如乘法模型。因此,使用式(2)对时间序列进行分解。

2.3 VMD

VMD 由 Dragomiretskiy 等^[13]提出,其将原始信号分解为一组离散模态分量,具有高度自适应性和较强的鲁棒性,能够有效去噪^[14]。具体步骤如下。

表 1 STL 两种分解模式预测精度对比

Table 1 Comparison of prediction accuracy between two decomposition modes of STL

STL 分解 模式	训练期				验证期			
	R	$RMSE$	$MAPE/\%$	NSE	R	$RMSE$	$MAPE/\%$	NSE
加法	0.983	0.171	61.149	0.966	0.974	0.168	71.025	0.949
乘法	0.992	0.361	15.857	0.983	0.991	0.327	13.705	0.977

步骤 1 利用希尔伯特变换计算各子序列的关联解析信号 $u_k(t)$, 得到其单边频谱。

步骤 2 将各子序列及其对应的中心频率 ω_k 调整到相应的基带。

步骤 3 通过解调信号的高斯平滑度估计频率带宽。带宽和中心频率有条件限制,即满足每个子序列的估计带宽之和最小化,因此,约束变分问题^[15]如式(3)所示:

$$\begin{cases} \min_{|u_k|, |\omega_k|} \left[\sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right] \\ \text{s. t.} \quad \sum_{k=1}^K u_k = f(t). \end{cases}; \quad (3)$$

式中: $f(t)$ 为原始序列; u_k 为 $f(t)$ 的第 k 个分量; ω_k 为 u_k 的中心频率; $\delta(t)$ 为狄拉克分布; K 为模态数; t 为时间; $\|\cdot\|$ 表示向量的范数; ∂_t 表示带宽; j 为虚数单位,用于表示复数形式的函数。

通过添加二次惩罚因子 α 和拉格朗日乘子 λ 可将式(3)转化为无约束问题,然后利用交替方向乘子法对其进行求解。 u_k 和 ω_k 的解表示为

$$\hat{u}_k^{n+1} = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2}; \quad (4)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega}. \quad (5)$$

式中: n 为迭代次数; $\hat{f}(\omega)$ 、 $\hat{u}_i(\omega)$ 、 $\hat{\lambda}(\omega)$ 和 \hat{u}_k^{n+1} 分别为 $f(t)$ 、 $u_i(t)$ 、 $\lambda(t)$ 和 u_k^{n+1} 的傅里叶变换。

2.4 SVM

SVM 在解决非线性问题方面具有独特优势^[16],与传统的人工神经网络等方法相比,SVM 在很大程度上解决了过拟合、非线性、高维度和局部极小值问题^[17],广泛应用于径流预测领域。它通过核函数实现非线性向线性问题的变换,并将支持向量机的学习过程转化为易计算的凸优化^[18]问题:

$$\begin{cases} \min \left[\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \right]; \\ \text{s. t.} \quad y_i - \mathbf{w}\Phi(\mathbf{x}_i) - b \leq \varepsilon + \xi_i; \\ \quad -y_i + \mathbf{w}\Phi(\mathbf{x}_i) + b \leq \varepsilon + \xi_i^*; \\ \quad i = 1, 2, \dots, l, \xi_i \geq 0, \xi_i^* \geq 0. \end{cases} \quad (6)$$

最终回归函数为

$$f(\mathbf{x}) = \sum_{i=1}^l (a_i - a_i^*) K(\mathbf{x}_i, \mathbf{x}_j) + b. \quad (7)$$

式中: $\Phi(x)$ 为非线性映射函数; \mathbf{w} 为方向向量; b

为回归函数截距; ε 为不敏感系数; $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_j)$ 为满足 Mercer 条件的核函数; \mathbf{x}_i 与 \mathbf{x}_j 为样本向量; y_i 代表训练样本的类别。 ξ_i^* 、 ξ_i 为松弛变量; a_i^* 和 a_i 为二次规划中的拉格朗日乘子。

SVM 的性能受核函数、惩罚因子 C 和核参数 g 的影响。径向基核函数(radial basis function, RBF)参数相对较少又可以很好地处理高维复杂样本问题^[19]。其表达式为

$$K(\mathbf{x}, \mathbf{x}_i) = e^{-g \|\mathbf{x} - \mathbf{x}_i\|^2}. \quad (8)$$

2.5 评价指标

为了衡量模型的预测效果,本文选取均方根误差 $RMSE$ 、纳什系数 NSE 、平均绝对百分比误差 $MAPE$ 和相关系数 R 来评价模型的预测精度。 $RMSE$ 、 NSE 、 $MAPE$ 和 R 的计算公式^[6]如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_1(i) - y_0(i))^2}; \quad (9)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (y_1(i) - y_0(i))^2}{\sum_{i=1}^n (y_0(i) - \bar{y}_0)^2}; \quad (10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_1(i) - y_0(i)}{y_1(i)} \right| \times 100\%; \quad (11)$$

$$R = \frac{\sum_{i=1}^n (y_0(i) - \bar{y}_0)(y_1(i) - \bar{y}_1)}{\sqrt{\sum_{i=1}^n (y_0(i) - \bar{y}_0)^2 \sum_{i=1}^n (y_1(i) - \bar{y}_1)^2}}. \quad (12)$$

式中: $y_1(i)$ 和 $y_0(i)$ 分别为预测值和实测值; \bar{y}_1 和 \bar{y}_0 分别为两者的平均值; n 为样本总数。

2.6 组合模型预测流程

预测模型流程如图 3 所示。

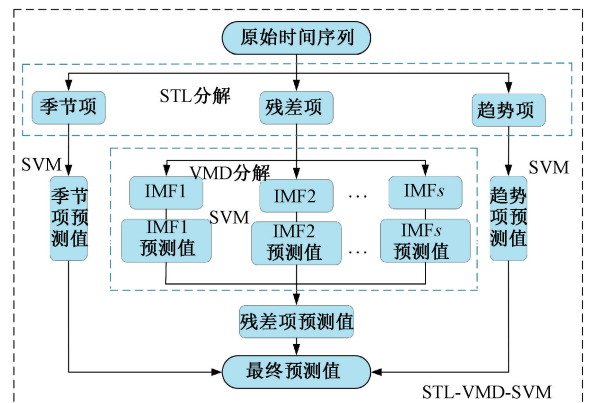


图 3 组合预测模型流程

Figure 3 Combined prediction model process

首先,利用 STL 将原始径流序列分解为不同频率的季节项、趋势项和残差项。其次,通过 VMD 将

STL 分解出的残差项进一步分解为一系列分量 IMF_s。再次,建立 SVM 模型预测季节项、趋势项和 IMF_s。最后,所有 IMF_s 的预测值之和为残差项的预测值,季节项、趋势项和残差项之积为原始径流序列的最终预测值。

3 分解及预测结果分析

3.1 STL 分解结果

首先对径流时间序列进行 STL 分解,分解结果如图 4 所示。趋势项与原始径流序列的变化规律呈现出总体一致性,但趋势项的变化更加平滑,残差项波动呈现随机性,季节项呈现出明显的变化规律。

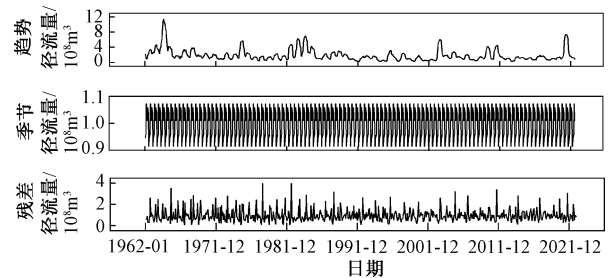


图 4 径流时间序列 STL 分解

Figure 4 STL decomposition of runoff time series

3.2 VMD 分解结果

对 STL 分解出的残差项进行 VMD 分解。分解时,模态数的选择对于预测效果有较大影响,整体的预测效果主要由分解出的残差量的预测效果决定(除残差分量外其他分量预测效果差别不大),表 2

表 2 不同模态数残差分量预测的 NSE

Table 2 NSE prediction of residual components for different modal numbers		
K	NSE	
	训练期	验证期
2	0.65	0.26
3	0.70	0.39
4	0.72	0.39
5	0.61	0.47
6	0.70	0.55
7	0.72	0.55
8	0.53	0.33
9	0.56	0.44
10	0.50	0.35

表 3 3 种模型的预测评价结果

Table 3 Prediction and evaluation results of three models

预测模型	训练期				验证期			
	R	RMSE	MAPE%	NSE	R	RMSE	MAPE%	NSE
SVM	0.609	2.125	51.600	0.300	0.465	2.174	48.440	0.168
STL-SVM	0.870	1.253	34.653	0.756	0.886	1.141	35.621	0.771
STL-VMD-SVM	0.992	0.361	15.857	0.983	0.991	0.327	13.705	0.977

是模态数 K 由 2 到 10 分解出的残差量预测的 NSE 。

由表 2 可知,当模态数 K 为 7 时训练期和验证期的预测效果最好。所以利用 VMD 将 STL 的残差项分解为 7 个 IMF_s 和 1 个残差分量,如图 5 所示。

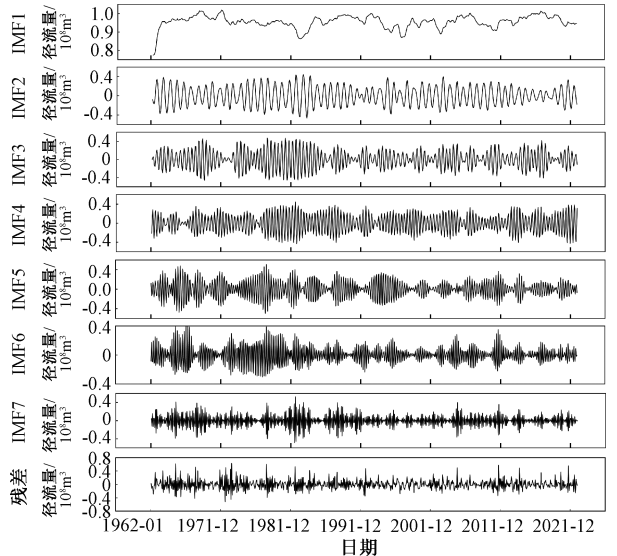


图 5 VMD 分解结果

Figure 5 VMD decomposition results

3.3 模型预测结果

未分解、一次分解和二次分解 3 种模型的性能指标值见表 3。由表 3 可知,STL-VMD-SVM 组合模型训练期的 R 、 $RMSE$ 、 $MAPE$ 和 NSE 分别为 0.992、0.361、15.857% 和 0.983;验证期分别为 0.991、0.327、13.705% 和 0.977,4 项评价指标均达到预测标准且有较高的精度。与 SVM 和 STL-SVM 模型相比,STL-VMD-SVM 组合模型的所有评价指标均为最优,一次分解模型的性能均优于未分解的模型。

表 4 为经过多次调节试验得出的 SVM 模型预测各个分量时的惩罚因子 C 和核参数 g 的最佳值。

如图 6 所示,模型预测值与实测值对比表明,STL-VMD-SVM 的组合模型能够较好地学习径流序列的波动规律和极值,对径流极值的模拟效果较好,因此,可以认为基于 STL-VMD-SVM 的组合模型能够较好地模拟研究区水文站点的径流过程。

为了进一步验证组合模型的有效性,将组合模型的预测结果与 SVM 模型以及 STL-SVM 模型的预测结果进行对比。所有模型采用相同的预测输入、

表 4 惩罚因子和核参数

Table 4 Penalty factors and kernel parameters					
分量	<i>C</i>	<i>g</i>	分量	<i>C</i>	<i>g</i>
季节	55	0.6	IMF4	46	0.5
趋势	18	0.6	IMF5	42	0.6
IMF1	55	0.6	IMF6	41	0.6
IMF2	55	0.5	IMF7	56	0.9
IMF3	55	0.1	残差	8	0.6

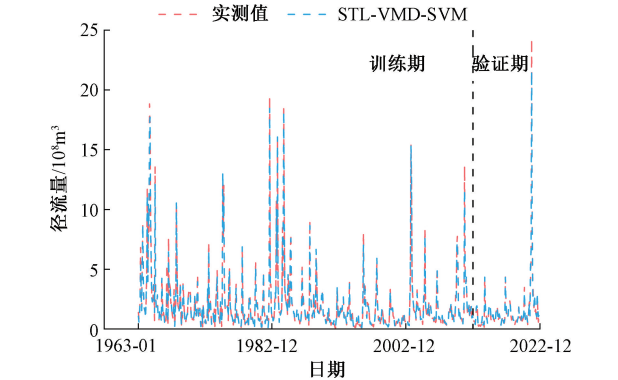


图 6 实测值与预测值对比

Figure 6 Comparison between measured and predicted values

参数、训练期以及验证期。

图 7 展示了验证期 3 种模型的预测效果。从图 7 可以看出,相较于 SVM 模型和 STL-SVM 模型,STL-VMD-SVM 组合模型能更精准地捕获原始径流量的变化特征,大幅度提升了模型预测能力,拟合效果最优。

从图 8 可以看出,STL-VMD-SVM 组合模型预测的最大值、最小值、中位数和异常值与实测值最接近。

本文基于二次分解思路的 STL-VMD 组合分解方法适用于分解月径流时间序列,采用此方法预处理后的月径流数据作为预测模型的输入,可以有效地消除数据的非平稳特征,有助于提高模型的预测性能。

3.4 与 BP、LSTM 模型的对比及普适性分析

为了进一步证明本文的组合模型在径流预测中的优势,选择了 2 个常用的径流预测模型:BP 神经网络和长短期记忆神经网络 (LSTM) 模型,并与 STL-VMD 的二次分解方法组合成 STL-VMD-BP 和 STL-VMD-LSTM 模型作为对比。

表 5 和图 9 显示了这 3 种模型的性能,可以看

表 5 3 种组合模型的预测评价指标

预测模型	训练期				验证期			
	<i>R</i>	<i>RMSE</i>	<i>MAPE%</i>	<i>NSE</i>	<i>R</i>	<i>RMSE</i>	<i>MAPE%</i>	<i>NSE</i>
STL-VMD-BP	0.985	0.45	21.213	0.969	0.986	0.416	21.649	0.970
STL-VMD-LSTM	0.988	0.401	18.315	0.974	0.991	0.388	18.185	0.973
STL-VMD-SVM	0.992	0.361	15.857	0.983	0.991	0.327	13.705	0.977

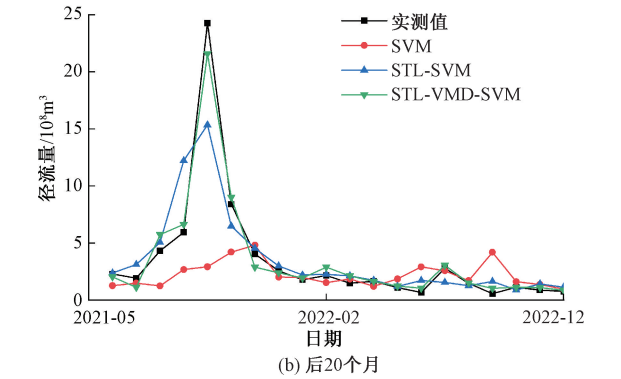
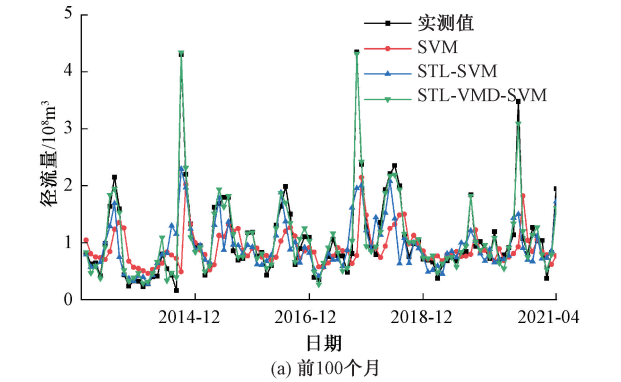


图 7 3 种模型验证期预测结果对比

Figure 7 Comparison of prediction results of three model validation periods

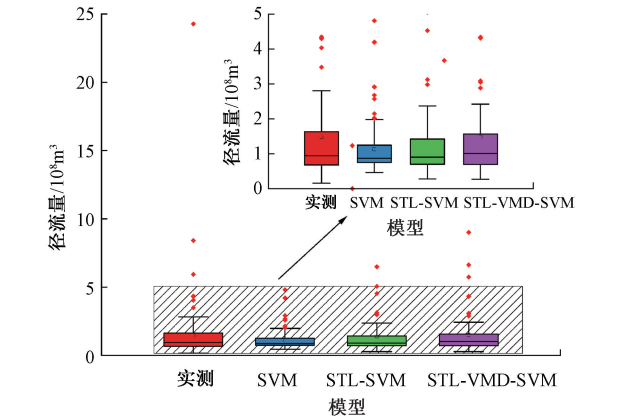


图 8 验证期预测结果箱线图

Figure 8 Box plot of validation period prediction results

出,本文的 STL-VMD-SVM 模型在所有评价指标上略优于其他 2 种模型。

考虑到不同的流域及站点径流序列具有不同的波动性,为了验证模型的普适性,选取黄河干流高村

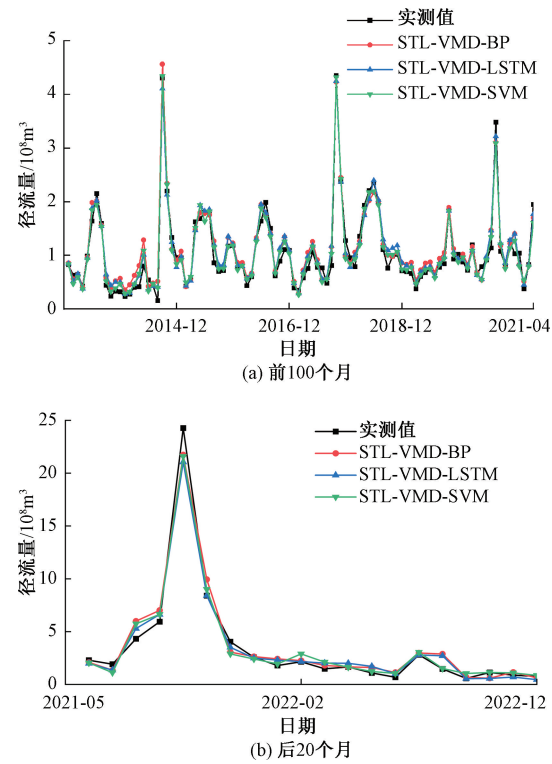


图 9 常用模型验证期预测结果对比

Figure 9 Comparison of prediction results for validation period of common models

站 2007—2022 年月径流序列,利用上文所述相同方法和步骤测试此模型,测试结果如表 6 和图 10 所示。

表 6 高村站径流序列预测评价结果

Table 6 Prediction and evaluation results of runoff sequence at Gaocun station				
模拟期	<i>R</i>	<i>RMSE</i>	<i>MAPE</i> %	<i>NSE</i>
训练期	0.993	1.650	5.927	0.986
验证期	0.989	3.263	8.509	0.979

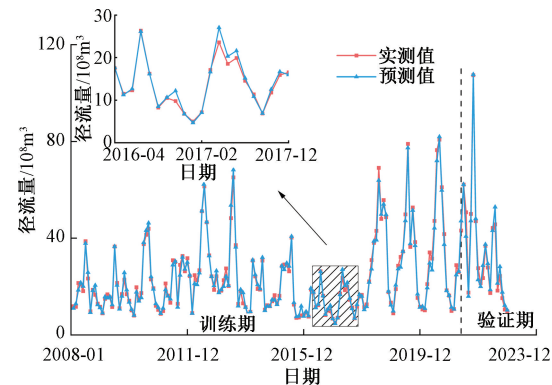


图 10 高村站径流序列实测值与预测值对比

Figure 10 Comparison of measured and predicted runoff sequences at Gaocun station

由表 6 和图 10 可知,该模型在高村站径流序列预测的应用性能仍比较优越,具有一定的普适性。STL-VMD-SVM 组合模型使月径流时间序列预测结果达到了更高的准确度。

4 结论

径流序列具有高度的非稳态和强季节性特征,为提高其预测精度,本文基于二次分解思路,将 STL-VMD 组合分解技术与 SVM 模型相结合,并将其应用于伊洛河流域黑石关站的月径流预测。主要研究结论如下。

- (1)分解增强了 SVM 的预测能力。STL 分解法能够有效去除径流序列的季节成分,基于二次分解的模型比一次分解的模型具有更好的性能。
- (2)基于二次分解后与 SVM 相结合的模型在 *R*、*RMSE*、*MAPE* 和 *NSE* 这 4 项评价指标中,略高于与常用的 BP 和 LSTM 模型相结合。STL-VMD-SVM 组合模型月径流预测具有优势,对黄河干流高村站径流序列的预测也有较好的效果,证明该模型具有一定的普适性。

使用 SVM 模型进行预测时,是利用手动调节其惩罚因子与核参数,工作量较大。在下一步研究中,考虑采用适当的优化算法进行调参,例如粒子群优化算法或麻雀搜索算法,以减少工作量并进一步提高预测精度。

参考文献:

[1] HIRABAYASHI Y, MAHENDRAN R, KOIRALA S, et al. Global flood risk under climate change[J]. Nature Climate Change, 2013, 3(9): 816–821.

[2] DEY P, MISHRA A. Separating the impacts of climate change and human activities on streamflow: a review of methodologies and critical assumptions[J]. Journal of Hydrology, 2017, 548:278–290.

[3] KOTTA J, HERKÜL K, JAAGUS J, et al. Linking atmospheric, terrestrial and aquatic environments: regime shifts in the Estonian climate over the past 50 years[J]. PLoS One, 2018, 13(12): e0209568.

[4] HE X X, LUO J G, LI P, et al. A hybrid model based on variational mode decomposition and gradient boosting regression tree for monthly runoff forecasting[J]. Water Resources Management, 2020, 34(2): 865–884.

[5] 雷庆文, 高培强, 李建林. 时序分解和 CNN-LSTM 相融合的月径流预报模型[J]. 长江科学院院报, 2023, 40(6): 49–54.

LEI Q W, GAO P Q, LI J L. A monthly runoff forecast model combining time series decomposition and CNN-LSTM[J]. Journal of Changjiang River Scientific Research Institute, 2023, 40(6): 49–54.

[6] 王文川, 杜玉瑾, 和吉, 等. 基于 CEEMDAN-VMD-BP 模型的月径流量预测研究[J]. 华北水利水电大学学

报(自然科学版), 2023, 44(1): 32-40, 48.

WANG W C, DU Y J, HE J, et al. Research on monthly runoff prediction based on CEEMDAN-VMD-BP model [J]. Journal of North China University of Water Resources and Electric Power (Natural Science Edition), 2023, 44(1): 32-40, 48.

[7] YANG H, LI W D. Data decomposition, seasonal adjustment method and machine learning combined for runoff prediction; a case study[J]. Water Resources Management, 2023, 37(1): 557-581.

[8] 朱得胜, 吕锡芝, 倪用鑫, 等. 2001—2018 年伊洛河流域产水系数演变分析[J]. 中国农村水利水电, 2022(9): 139-145.

ZHU D S, LYU X Z, NI Y X, et al. An analysis of the changes in the water yield coefficient in the Yiluo River Basin from 2001 to 2018[J]. China Rural Water and Hydropower, 2022(9): 139-145.

[9] HOU J, QIN T L, YAN D H, et al. Evaluation of water-land resources regulation potential in the Yiluo River Basin, China[J]. Ecological Indicators, 2023, 153: 110410.

[10] 张艺馨. 基于改进 BP 神经网络模型的采空区特殊下垫面条件下径流预报研究[D]. 太原: 太原理工大学, 2021.

ZHANG Y X. Study on runoff prediction of goaf with special underlying surface based on improved BP neural network model[D]. Taiyuan: Taiyuan University of Technology, 2021.

[11] 邢珊珊. 基于时空关联的城市快速路短时交通流预测方法研究[D]. 北京: 北京交通大学, 2017.

XING S S. Study on short-term traffic flow forecasting method of urban expressway based on spatial-temporal correlation[D]. Beijing: Beijing Jiaotong University, 2017.

[12] 薛联青, 周天文, 刘远洪, 等. 基于两阶段分解和可解释机器学习的中长期径流预测[J]. 中国农村水利水电, 2023(7): 1-7, 18.

XUE L Q, ZHOU T W, LIU Y H, et al. Medium and long-term runoff forecasting based on two-stage decomposition and interpretable machine learning[J]. China Rural Water and Hydropower, 2023(7): 1-7, 18.

[13] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.

[14] SHARMA V, PAREY A. Extraction of weak fault transients using variational mode decomposition for fault diagnosis of gearbox under varying speed[J]. Engineering Failure Analysis, 2020, 107: 104204.

[15] 廖晓辉, 陈川川. 改进的 VMD-HT 在电能质量扰动检测中的应用[J]. 郑州大学学报(工学版), 2021, 42(1): 21-27.

LIAO X H, CHEN C C. Application of improved VMD-HT in power quality disturbance detection[J]. Journal of Zhengzhou University (Engineering Science), 2021, 42(1): 21-27.

[16] TOMAR D, AGARWAL S. Twin support vector machine: a review from 2007 to 2014[J]. Egyptian Informatics Journal, 2015, 16(1): 55-69.

[17] 张颖超, 成金杰, 邓华, 等. 基于相似日和特征提取的短期风电功率预测[J]. 郑州大学学报(工学版), 2020, 41(5): 44-49.

ZHANG Y C, CHENG J J, DENG H, et al. Short-term wind power prediction based on similar day and feature extraction[J]. Journal of Zhengzhou University (Engineering Science), 2020, 41(5): 44-49.

[18] CHEN S, REN M M, SUN W. Combining two-stage decomposition based machine learning methods for annual runoff forecasting [J]. Journal of Hydrology, 2021, 603: 126945.

[19] 纪昌明, 周婷, 向腾飞, 等. 基于网格搜索和交叉验证的支持向量机在梯级水电系统隐随机调度中的应用[J]. 电力自动化设备, 2014, 34(3): 125-131.

JI C M, ZHOU T, XIANG T F, et al. Application of support vector machine based on grid search and cross validation in implicit stochastic dispatch of cascaded hydropower stations[J]. Electric Power Automation Equipment, 2014, 34(3): 125-131.

Monthly Runoff Prediction Method Based on Secondary Decomposition and Support Vector Machine

GAN Rong^{1,2}, MA Chaoxin^{1,2}, GAO Yong³, GUO Lin³, HOU Xiaoli⁴, LU Xueyong⁵

(1. School of Water Resources and Transportation, Zhengzhou University, Zhengzhou 450001, China; 2. Henan Key Laboratory of Groundwater Pollution Prevention and Rehabilitation, Zhengzhou 450001, China; 3. Henan Provincial Geological Research Institute, Zhengzhou 450001, China; 4. Henan Province Yudong Water Resources Guarantee Center, Kaifeng 475000, China; 5. Canal Head Branch Company of China South-to-North Water Diversion Middle Route Corporation Limited, Nanyang 473000, China)

Abstract: A monthly runoff prediction model(STL-VMD-SVM) based on a secondary decomposition using loess (STL) and variational mode decomposition (VMD) combined with a support vector machine(SVM) was proposed

to address the nonlinear and non-stationary characteristics of runoff sequences. This model utilized STL to decompose the original runoff sequence into seasonal, trend, and residual terms of different frequencies and decomposed the residual term into IMFs through VMD. An SVM model was established to predict seasonal, trend, and IMFs. The sum of the predicted values of all IMFs was the predicted value of the residual term, and the product of seasonal, trend, and residual terms was the final predicted value of the original runoff series. Based on the monthly runoff time series of Heishiguan Station and Gaocun Station on the mainstream of the Yellow River in the Yiluo River Basin, an example application and universality evaluation were conducted, and compared with the BP neural network model and the long shortterm memory neural network model(LSTM). The results showed that for the runoff prediction of Heishiguan Station in the Yiluo River Basin, the *NSE*, *MAPE*, *RMSE*, and *R* in the validation period of the proposed model were 0.977, 13.705%, 0.327 and 0.991, respectively, and their prediction accuracy was better than that of the single model and the primary decomposition model. The secondary decomposition of STL-VMD could effectively improve the prediction accuracy of the model. The *NSE*, *MAPE*, *RMSE*, and *R* during the validation period in the runoff prediction at Gaocun Station on the mainstream of the Yellow River were 0.979, 8.509%, 3.263, and 0.989, respectively, which also achieved good prediction results.

Keywords: monthly runoff prediction; secondary decomposition; STL; VMD; SVM; neural network

(上接第 31 页)

Preparation and Performance Analysis of Bio-asphalt from Liquefaction and Resinification of Straw

DING Zhan^{1,2}, AN Linyu^{1,2}, LI Huifeng³, TIAN Chenxi^{1,2}, ZHOU Chunyu^{1,2}, LIU Fengkai^{1,2}

(1. College of Water Conservancy and Environment, Chang'an University, Xi'an 710054, China; 2. Key Laboratory of Groundwater Hydrology and Ecological Effects in Arid Areas, Ministry of Education, Chang'an University, Xi'an 710054, China; 3. College of Highway, Chang'an University, Xi'an 710064, China)

Abstract: Waste wood, crop residues, livestock manure and waste oil, were widely used to produce bio-oil. And then partially or completely replace the petroleum-based asphalt. But the road performance of this asphalt was insufficient. In this study, the development of green and sustainable bio-asphalt to partially replace petroleum-based asphalt with wood-based phenolic resin made by straw liquefaction products was proposed. Firstly, the separation of straw components was carried out to extract cellulose and lignin. Straw and its primary components were liquefied to explore the main factors influencing the liquefaction of straw. Then the liquefied products were combined with formaldehyde to synthesize lignin-based phenolic resin WPR. And the phenolic resin PR, which was obtained from the reaction with phenol, was compared and analyzed. Finally, the resin was put into matrix asphalt with different ratios to prepare bio-asphalt, and the performance of the bio-asphalt was analyzed by the three major indexes of each bio-asphalt and viscosity. The results showed that the liquefaction rate of straw mainly depended on the liquefaction rate of cellulose, and its liquefaction reaction was complex, with a reaction kinetic level of 1.71. From the resin yield and liquefaction products, WPR, PR FT-IR spectral analysis showed that the WPR resin synthesis rate was higher and had a better reactivity than the PR. Only when mixed with the liquefaction products, the performance of the bio-asphalt prepared was poorer. However, when mixed with WPR resin, the bio-asphalt's high temperature stability, low-temperature cracking resistance, deformation resistance, and temperature stability were better than matrix asphalt.

Keywords: liquefied resinate; wood-based resins; bio-asphalt; viscosity