

DOI: 10.20040/j.cnki.1000-7709.2023.20222289

基于两阶段分解策略的月径流模拟模型研究

李 鑫,王双银,黄毓林,樊镕鑫,马雪燕

(西北农林科技大学水利与建筑工程学院, 陕西 杨陵 712100)

摘要: 中长期径流模拟可为水资源合理配置提供科学依据,对流域高质量发展具有重要意义。基于改进的自适应噪声完备集成经验模态分解算法(ICEEMDAN)和奇异谱分析(SSA)的两阶段分解策略,利用鲸鱼算法(WOA)优化的长短期记忆网络模型(LSTM),构建了月径流模拟模型 ICEEMDAN-SSA-WOA-LSTM。将其应用于石头河水库月径流模拟,并与单次分解的 ICEEMDAN-WOA-LSTM、SSA-WOA-LSTM 和未分解的 WOA-LSTM 模型进行对比分析。结果表明,ICEEMDAN-SSA-WOA-LSTM 模型模拟效果最佳,率定期和验证期 3 项评价指标均优于其他模型,验证期均方根误差为 $1.278 \text{ m}^3/\text{s}$ 、平均绝对误差为 $0.893 \text{ m}^3/\text{s}$ 、纳什效率系数为 0.985。两阶段分解策略月径流模拟模型可显著提高月径流模拟精度,可用于全年入库径流模拟。

关键词: 径流模拟; 二次分解; 鲸鱼优化算法; 长短期记忆网络; 石头河水库

中图分类号: TV121⁺.2; P338

文献标志码: A

文章编号: 1000-7709(2023)09-0006-05

1 引言

中长期径流模拟可为水资源合理配置提供有力支撑,对水资源可持续利用具有重要意义^[1]。传统分布式水文模型依据流域产汇流理论和参数化边界条件模拟径流,具有一定的局限性^[2]。近年来,以统计学和数据挖掘技术为基础的深度学习模型,以其优秀的径流演变信息挖掘能力,通过低成本输入却能获得较高精度的模拟结果而得到广泛应用^[3]。其中长短期记忆网络(LSTM)凭借特有的记忆属性,有效克服了模型训练过程中的梯度消失、维度灾难等问题,是当前水文模拟领域研究的热点^[4]。然而径流序列具有随机性和非线性特征,仅采用深度学习模型难以高效模拟整个径流过程。近年来将数据分解技术与深度学习模型相结合,提高了模型模拟精度^[5]。但这些方法分解产生的高频分量复杂程度较高,增大了模型模拟难度,影响径流序列整体模拟精度。为进一步挖掘高频分量中的有效信息,在改进自适应噪声完备集成经验模态分解技术(ICEEMDAN)的基础上,提出对高频分量进行再分解处理的两

阶段分解策略,将其应用于石头河流域,以期提高中长期径流模拟精度。

2 研究思路与方法

2.1 研究思路

采用 ICEEMDAN 对月径流序列进行第一阶段分解处理。ICEEMDAN 将完全集合经验模式分解(CEEMDAN)的模态整体均值计算方式改进为局部均值计算,对于月径流序列中包含的不同尺度特征成分具有更好的提取能力^[6]。第一阶段分解得到多个本征模态分量(IMF),各分量对应频率不同,其中高频分量掺杂了较多噪声信息,复杂程度较高,影响模型的整体模拟精度。因此采用样本熵^[7]衡量各 IMF 分量的复杂程度,判定其中的高频分量,并提出对高频分量进行第二阶段分解的策略,以期提高径流的整体模拟精度。

第二阶段选用奇异谱分析(SSA)分解高频分量。SSA 可从包含噪声的序列中提取较多可靠信息,有效进行周期分量重建,且重建后的序列较为平滑,这一特性可显著降低模型模拟难度^[8]。因此本文引入应用奇异谱分析的第二阶段分解策

收稿日期: 2022-10-25, **修回日期:** 2022-11-25

基金项目: 陕西省水资源与河库调度中心关中地区灌溉水源水质状况调查研究项目(SXRCZB2021-ZC-CS1008)

作者简介: 李鑫(1997-),男,硕士研究生,研究方向为径流预测,E-mail:531743572@qq.com

通讯作者: 王双银(1969-),男,博士、副教授、硕导,研究方向为水文预报和水资源配置,E-mail:wshy0806@nwsuaf.edu.cn

略分解高频分量,得到多个重构分量(RC)。

对于第一阶段分解得到的低频分量和第二阶段对高频分量分解得到的重构分量,应用基于鲸鱼优化算法(WOA)^[9]的长短期记忆网络分别进行模拟,得到模拟序列 M_i 、 N_i ($i=1,2,\dots,n$),将所有模拟序列线性叠加,得到最终月径流模拟结果。具体研究思路见图 1。

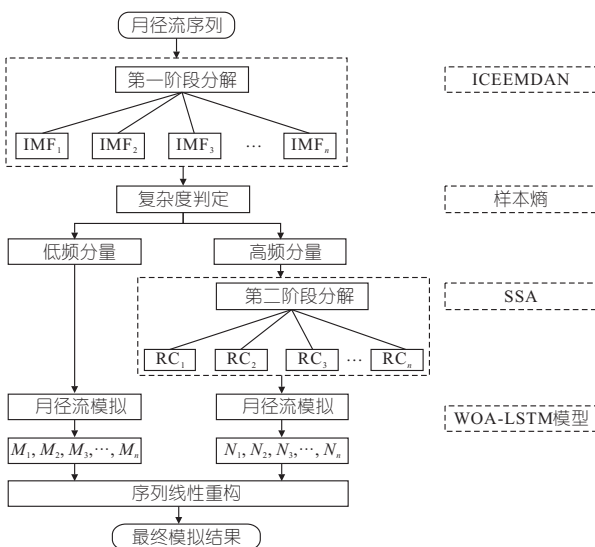


图 1 基于两阶段分解策略的月径流模拟研究流程图

Fig. 1 Flow chart of monthly runoff simulation study based on two-phase decomposition strategy

2.2 研究方法

2.2.1 改进的自适应噪声完备集成经验模态分解

改进的自适应噪声完备集成经验模态分解(ICEEMDAN)基本原理如下。首先向原始月径流序列中加入白噪声,构建新序列 d_1 ,并通过计算新序列 d_1 的局部均值得到第 1 组残差 r_1 ,新序列 d_1 与第 1 组残差 r_1 的差值即为第 1 个分量 IMF_1 ;而后继续向第 1 组残差 r_1 添加白噪声得到新序列 d_2 ,依然对 d_2 采用局部均值计算得到第 2 组残差 r_2 , r_1 与 r_2 的差值即为第 2 个分量 IMF_2 ;以此类推直到残差 r_i ($i=1,2,\dots,n$) 不可再分解或达到最高迭代次数,得到后续的分量 IMF_i ($i=1,2,\dots,n$)。具体计算方法详见文献[6]。

2.2.2 样本熵

通过计算样本熵衡量各 IMF 分量的复杂程度,区分高、低频分量。样本熵为条件概率的负自然对数,其计算简洁高效,对分量长度依赖性低。样本熵值越大,分量的复杂度越大,反之则越小。具体计算方式为:

$$E(m, r, N) = \ln B^m(r) - \ln B^{m+1}(r) \quad (1)$$

式中, E 为样本熵; m 为嵌入维度; r 为相似容限; N 为分量序列长度; B 为分量序列的自相似度概率。

2.2.3 奇异谱分析

奇异谱分析(SSA)是一种处理非线性时间序列的分析技术,能较好地提取序列关键信息的同时无需先验信息,具有稳定识别和强化周期信号的优点。SSA 的基本思路是将原序列转换为 $L \times K$ 维的轨迹矩阵,并对轨迹矩阵进行奇异值分解,得到多个不同周期的分量,并根据分量间相关性将其组合,得到最终的重构分量 R_{RC_i} ($i=1,2,\dots,n$)。具体原理详见文献[8]。

2.2.4 基于鲸鱼优化算法的长短期记忆网络

长短期记忆网络通过输入门、遗忘门和输出门 3 个门结构有选择性地确定从当前单元状态删除、添加或存储的信息,3 个门结构在输入和反馈间建立滞后连接,并强制在循环神经元内维持连续流动误差,最终缓解导数乘法引起的梯度消失问题。具体计算公式详见文献[4]。其神经网络元结构见图 2。

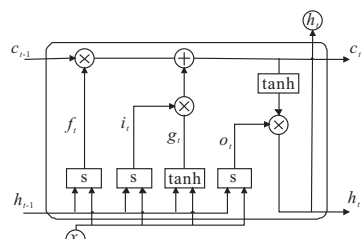


图 2 LSTM 神经网络结构

Fig. 2 Structure of LSTM

LSTM 模型参数大多依赖于经验选取,针对不同序列往往难以达到最佳性能,通过引入智能优化算法可有效解决这一问题。已有研究表明,鲸鱼优化算法(WOA)比粒子群优化算法(PSO)具有更好的寻优精度和全局搜索能力^[10]。因此,采用 WOA 对 LSTM 模型中隐藏层神经元数目及学习率两个参数进行寻优确定。WOA 具体原理见文献[9]。

2.2.5 评价指标

采用均方根误差(R_{RMSE})、平均绝对误差(M_{MAE})及纳什效率系数(N_{NS})3 个指标作为模型模拟精度的评价标准。 N_{NS} 值越接近 1, R_{RMSE} 、 M_{MAE} 值越接近 0,表示模拟精度越高。

3 模型应用

3.1 石头河流域概况

石头河水库位于秦岭北麓渭河一级支流石头河干流上,控制流域面积 673 km^2 ,总库容 $1.47 \times 10^8 \text{ m}^3$,是一座集防洪、供水、发电、灌溉功能于一体的综合利用水利工程。水库承担西安、咸阳、宝

鸡、杨凌等城市的供水任务,在保障流域社会经济快速稳定发展方面发挥重要作用。石头河干流先后设立斜峪关水文站(1950~1983 年数据)和鸚鸽水文站(1976 年至今的数据),控制流域面积分别为 686、507 km²。将黄河流域水文年鉴中刊布的两水文站逐月平均流量采用水文比拟法统一折算至水库坝址断面,形成 1950~2020 年共 71 年的石头河水库月径流序列。石头河流域示意图见图 3。

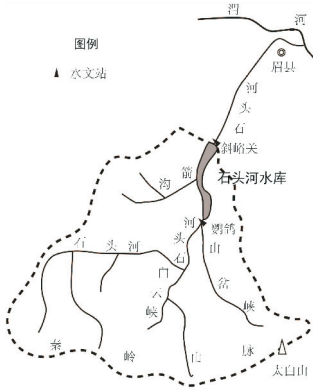


图 3 石头河流域示意图

Fig. 3 Schematic diagram of the Stone River

3.2 序列分解

对石头河水库 1950~2020 年共 71 年月径流序列进行 ICEEMDAN 分解,共得到 9 个频率不同的本征模态分量 IMF₁~IMF₉ 和 1 个趋势分量(Res)。各 IMF 分量频率由高到低排列,高、低频分量间振幅差距明显,表明原始径流序列波动性较强;趋势分量 Res 呈平缓—加速—平缓的下降趋势,表明原始径流序列总体呈下降趋势且在各时段下降速度不同。

利用样本熵进一步判断 IMF₁~IMF₉ 中的高、低频分量,计算结果见图 4。IMF₁、IMF₂、IMF₃ 的样本熵值基本浮动在 2.0,而 IMF₄ 样本熵值仅达到 1.3,缩减了近 1/3,因此将 IMF₁、IMF₂、IMF₃ 划定为高频分量。需应用奇异谱分

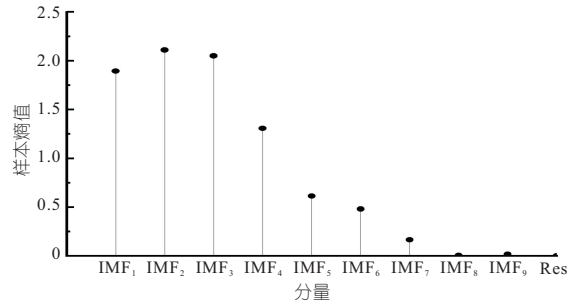


图 4 各分量样本熵值

Fig. 4 Entropy of each subsequence

析(SSA)进一步挖掘其中的有效信息。结合 ICEEMDAN 的分解层数,最终设定 SSA 的分解层数为 10,将 IMF₁、IMF₂、IMF₃ 各分解为 10 个重构分量。

3.3 模型构建

径流的形成与自身不同时段的前期径流具有相关性。针对两阶段分解得到的若干分量,采用偏自相关函数(PACF)来确定输入因子。首先计算时间序列的 PACF 值,然后找出前 12 个月中通过 95%置信检验的前期径流。由于不同分量筛选得到的前期径流数目不同,最少为 3 个,最多为 10 个,为保证后续模型输入维度一致,以各分量同时满足的最多前期径流数目为标准,最终选用 PACF 值最高的 3 个前期径流作为模型的输入因子,结果见表 1。

搭建 LSTM 模型需确定相应参数,根据输入因子的选取结果,LSTM 模型的输入层数设置为 3。模型隐藏层神经元数目及学习率两个参数利用 WOA 优化确定,其中神经元数目设置的搜索边界为[8,400],学习率设置的搜索边界为[0.01, 0.1]。训练过程采用 Adam 算法,训练次数设置为 100 次。WOA 算法优化过程采用均方误差作为模型适应度,设置种群数量为 10,最大迭代次数为 20 次。利用前 64 年共计 768 个月径流数据为率定期样本率定模型,后 7 年共计 84 个月径流

表 1 各分量输入因子选取结果
Tab. 1 Selection results of input factors of each component

| ICEEMDAN | | SSA | | ICEEMDAN-SSA-IMF ₁ | | ICEEMDAN-SSA-IMF ₂ | | ICEEMDAN-SSA-IMF ₃ | |
|------------------|-----------------------------|------------------|-----------------------------|-------------------------------|-----------------------------|-------------------------------|-----------------------------|-------------------------------|------------------------------|
| 分量 | 输入因子 | 分量 | 输入因子 | 分量 | 输入因子 | 分量 | 输入因子 | 分量 | 输入因子 |
| IMF ₁ | $x_{t-1}, x_{t-2}, x_{t-6}$ | RC ₁ | $x_{t-1}, x_{t-2}, x_{t-4}$ | RC ₁ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₁ | $x_{t-1}, x_{t-2}, x_{t-6}$ | RC ₁ | $x_{t-1}, x_{t-2}, x_{t-4}$ |
| IMF ₂ | $x_{t-1}, x_{t-5}, x_{t-6}$ | RC ₂ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₂ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₂ | $x_{t-1}, x_{t-2}, x_{t-6}$ | RC ₂ | $x_{t-1}, x_{t-2}, x_{t-3}$ |
| IMF ₃ | $x_{t-1}, x_{t-2}, x_{t-7}$ | RC ₃ | $x_{t-1}, x_{t-2}, x_{t-8}$ | RC ₃ | $x_{t-1}, x_{t-2}, x_{t-4}$ | RC ₃ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₃ | $x_{t-1}, x_{t-2}, x_{t-3}$ |
| IMF ₄ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₄ | $x_{t-1}, x_{t-2}, x_{t-4}$ | RC ₄ | $x_{t-2}, x_{t-4}, x_{t-6}$ | RC ₄ | $x_{t-1}, x_{t-2}, x_{t-5}$ | RC ₄ | $x_{t-1}, x_{t-2}, x_{t-11}$ |
| IMF ₅ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₅ | $x_{t-2}, x_{t-4}, x_{t-5}$ | RC ₅ | $x_{t-2}, x_{t-4}, x_{t-6}$ | RC ₅ | $x_{t-2}, x_{t-3}, x_{t-6}$ | RC ₅ | $x_{t-1}, x_{t-2}, x_{t-3}$ |
| IMF ₆ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₆ | $x_{t-2}, x_{t-3}, x_{t-4}$ | RC ₆ | $x_{t-2}, x_{t-4}, x_{t-6}$ | RC ₆ | $x_{t-2}, x_{t-4}, x_{t-5}$ | RC ₆ | $x_{t-2}, x_{t-4}, x_{t-7}$ |
| IMF ₇ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₇ | $x_{t-1}, x_{t-2}, x_{t-4}$ | RC ₇ | $x_{t-1}, x_{t-2}, x_{t-4}$ | RC ₇ | $x_{t-2}, x_{t-4}, x_{t-6}$ | RC ₇ | $x_{t-2}, x_{t-4}, x_{t-6}$ |
| IMF ₈ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₈ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₈ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₈ | $x_{t-1}, x_{t-2}, x_{t-4}$ | RC ₈ | $x_{t-2}, x_{t-4}, x_{t-6}$ |
| IMF ₉ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₉ | $x_{t-1}, x_{t-2}, x_{t-4}$ | RC ₉ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₉ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₉ | $x_{t-1}, x_{t-2}, x_{t-4}$ |
| Res | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₁₀ | $x_{t-2}, x_{t-3}, x_{t-4}$ | RC ₁₀ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₁₀ | $x_{t-1}, x_{t-2}, x_{t-3}$ | RC ₁₀ | $x_{t-1}, x_{t-2}, x_{t-3}$ |

注: $x_{t-1}, x_{t-2}, x_{t-6}$ 分别为前 1、2、6 月的径流信息。如模拟 7 月的径流,需用到 6、5、1 月的径流信息。

数据作为验证期样本进行验证。

3.4 模型性能评价

为进一步检验 ICEEMDAN-SSA-WOA-LSTM 模型的模拟性能,选用未经分解处理的 WOA-LSTM 模型及单一分解的 ICEEMDAN-WOA-LSTM、SSA-WOA-LSTM 模型分别模拟月径流序列,具体模拟结果见图 5、6。

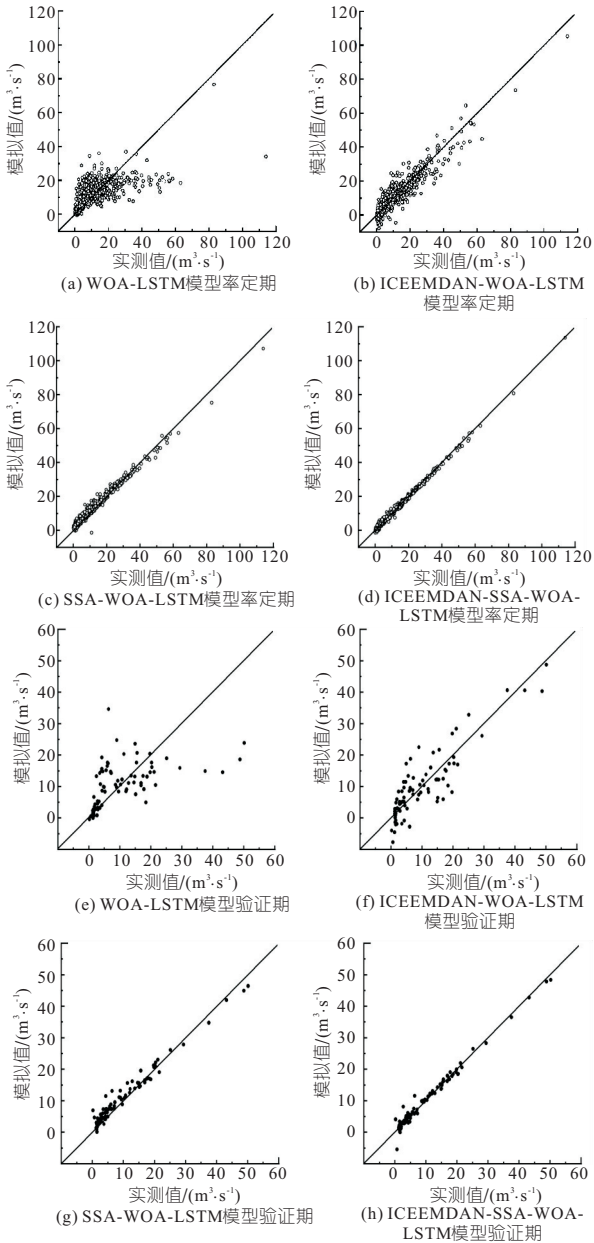


图 5 4 种模型率定期和验证期 Q-Q 图

Fig. 5 Q-Q chart of four model rates in periodic and validation periods

对比 4 种模型在率定期和验证期的应用效果,未经分解的 WOA-LSTM 模型表现较差,而加入分解技术的耦合模型均能较好地模拟径流。从模拟效果来看,ICEEMDAN-SSA-WOA-LSTM 模型 > SSA-WOA-LSTM 模型 > ICEEMDAN-WOA-LSTM 模型 > WOA-LSTM 模型 (“>”表示优于)。

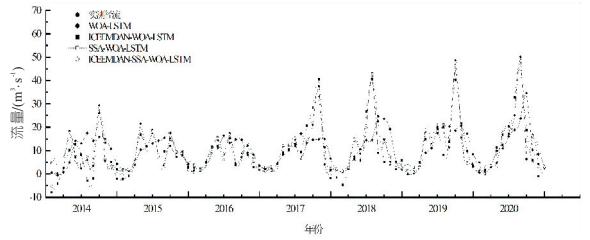


图 6 4 种模型验证期模拟过程线图

Fig. 6 Simulation process line diagram for 4 model validation periods

在率定期,WOA-LSTM 模型的模拟能力并不稳定,模拟值大多仅集中在 $0 \sim 40 \text{ m}^3/\text{s}$,且实测值大于 $25 \text{ m}^3/\text{s}$ 以后的模拟结果明显偏离 45° 线,模拟误差与实测值呈正相关关系。而加入分解技术后的模型模拟效果均有较大改善,在单一分解的模型中,ICEEMDAN-WOA-LSTM 模型的模拟精度较 WOA-LSTM 模型显著提升,但结果出现较多负值,这与 ICEEMDAN 分解过程中加入了高斯白噪声有关,模型模拟过程低估了分量中的负值,导致序列重构误差不断累积,最终得到不合理的结果;SSA-WOA-LSTM 模型对径流序列的模拟能力较强,但出现高流量模拟结果偏低、低流量模拟结果偏高的均化现象。两阶段分解的 ICEEMDAN-SSA-WOA-LSTM 模型则在保证高模拟精度的同时克服了此缺陷。

在验证期,WOA-LSTM 模型的模拟结果出现明显偏差。加入分解技术的耦合模型均准确模拟了径流的丰枯情况,SSA-WOA-LSTM 模型对于高峰现的捕捉能力较强,模拟结果更接近实测值,但对于低值的拟合能力较差,结果偏高,均化现象较率定期更明显;ICEEMDAN-WOA-LSTM 模型的结果相反,对于高峰现的捕捉能力弱于 SSA-WOA-LSTM 模型,对于枯水流量则表现更好,但依然出现了较多负值;两阶段分解的 ICEEMDAN-SSA-WOA-LSTM 模型模拟结果中绝大部分点据分布在 45° 线附近,模拟值与实测值之间保持良好的一致性,无论在峰值捕捉还是整体模拟精度方面均优于其他模型。

相较于单次分解的 ICEEMDAN-WOA-LSTM、SSA-WOA-LSTM 模型和未分解的 WOA-LSTM 模型,ICEEMDAN-SSA-WOA-LSTM 模型通过对原始序列进行极大限度的分解处理,有效挖掘其中的周期、趋势信息,削弱了序列的波动范围,得到一系列规律明显的平滑分量,显著降低了后续深度学习模型的模拟难度,提升了序列的模拟精度。

4 种模拟模型率定期及验证期的评价结果见

表 2。结果表明,无论是率定期还是验证期,组合模型 ICEEMDAN-SSA-WOA-LSTM、SSA-WOA-LSTM、ICEEMDAN-WOA-LSTM 的评价指标远优于单一的 WOA-LSTM 模型,其中 ICEEMDAN-SSA-WOA-LSTM 模型的模拟效果最佳,验证期均方根误差为 1.278 m³/s、平均绝对误差为 0.893 m³/s、纳什效率系数达到 0.985。对比单一分解的 ICEEMDAN-WOA-LSTM、SSA-WOA-LSTM 模型,两阶段分解的 ICEEMDAN-SSA-WOA-LSTM 模型将验证期纳什效率系数分别提高了 24.37%、4.5%,均方根误差和平均绝对误差分别减少了 73.20%、49.37% 和 74.67%、56.82%,对于复杂的非线性径流数据具有更好的模拟效果。

表 2 4 种模型对应的模拟精度指标

Tab. 2 The simulation accuracy index corresponding to the four models

| 模型 | 率定期 | | | 验证期 | | |
|-----------------------|------------|-----------|----------|------------|-----------|----------|
| | R_{RMSE} | M_{MAE} | N_{NS} | R_{RMSE} | M_{MAE} | N_{NS} |
| WOA-LSTM | 8.451 | 4.922 | 0.455 | 8.950 | 5.846 | 0.266 |
| ICEEMDAN-WOA-LSTM | 4.041 | 2.786 | 0.875 | 4.769 | 3.526 | 0.792 |
| SSA-WOA-LSTM | 2.291 | 1.935 | 0.960 | 2.524 | 2.068 | 0.942 |
| ICEEMDAN-SSA-WOA-LSTM | 0.882 | 0.657 | 0.994 | 1.278 | 0.893 | 0.985 |

注: R_{RMSE} 、 M_{MAE} 单位均为 m³/s。

4 结论

a. 相较于单一的 WOA-LSTM 模型,分解技术能够有效提取序列中的特征信息,显著提升模型对序列随机成分的捕捉能力。

b. SSA 对 ICEEMDAN 分解产生的高频分量具有显著的平滑处理能力,明显改善了高频分量模拟过程中出现的负值现象。

c. 两阶段分解策略能够对原始序列进行极大限度的再处理,提出的耦合模型有效提高了径流序列的模拟精度。

参考文献:

- [1] 刘柏君,权锦,杨超慧,等. 基于 SWAT 模型的青海省主要流域径流模拟[J]. 水资源保护,2016,32(6): 39-44.
- [2] 练继建,孙萧仲,马超,等. 基于 EEMD-AR 模型的丹江口水库年径流随机模拟与预报[J]. 水利水电科技进展,2017,37(5):16-21.
- [3] GREER B HUMPHREY, MATTHEW S GIBBS, GRAEME C DANDY, et al. A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network[J]. Journal of hydrology, 2016, 540: 623-640.
- [4] 熊一橙,徐炜,张锐,等. 基于 LSTM 网络的长江上游流域径流模拟研究[J]. 水电能源科学,2021,39(9):22-24,40.
- [5] 梁浩,黄生志,孟二浩,等. 基于多种混合模型的径流预测研究[J]. 水利学报,2020,51(1):112-125.
- [6] COLOMINAS M A, SCHLOTTHAUER G, TORRES M E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing[J]. Biomedical signal processing & control, 2014, 14(1): 19-29.
- [7] STEVE P. Approximate entropy (ApEn) as a complexity measure[J]. Chaos: A interdisciplinary journal of nonlinear science, 1995, 5(1):110-117.
- [8] 徐克红,程鹏飞,文汉江. 太阳黑子数时间序列的奇异谱分析和小波分析[J]. 测绘科学,2007,32(6): 35-38,205.
- [9] MIRJALILI S, ANDREW L. The whale optimization algorithm[J]. Advances in engineering software, 2016, 95:51-67.
- [10] 李祿德,崔东文. 基于小波包分解与相空间重构的 SSA-ELM 水文时间序列预报模型[J]. 人民珠江, 2022, 43(8):100-108.

Study of Monthly Runoff Simulation Model Based on Two-stage Decomposition Strategy

LI Xin, WANG Shuang-yin, HUANG Yu-lin, FAN Rong-xin, MA Xue-yan

(College of Water Resources and Architectural Engineering, Northwest A & F University, Yangling 712100, China)

Abstract: The medium and long-term runoff simulation can provide a scientific basis for the rational allocation of water resources, which is of great significance to the high-quality development of the basin. Based on the improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) and singular spectrum analysis (SSA) two-stage decomposition strategy, the monthly runoff simulation model ICEEMDAN-SSA-WOA-LSTM was constructed by using the long short-term memory network model (LSTM) optimized by the whale optimization algorithm (WOA). It was applied to the simulation of monthly runoff in Stone River Reservoir and compared with the single decomposed ICEEMDAN-WOA-LSTM, SSA-WOA-LSTM and the undecomposed WOA-LSTM models. The results show that the ICEEMDAN-SSA-WOA-LSTM model has the best simulation effect, and the three evaluation indexes in the calibration period and validation period are better than other models, with the root mean square error of 1.278 m³/s, the average absolute error of 0.893 m³/s, and the Nash efficiency coefficient of 0.985 in the validation period. The two-stage decomposition strategy model can significantly improve the accuracy of monthly runoff simulation and can be used for year-round incoming runoff simulation.

Key words: runoff simulation; secondary decomposition; whale optimization algorithm; long short-term memory neural network; Stone River Reservoir