

DOI:10.13476/j.cnki.nsbdtk.2022.0043

徐冬梅,夏王萍,王文川.基于黏菌算法优化 VMD-CNN-GRU 模型的年径流预测[J].南水北调与水利科技(中英文),2022,20(3):429-439. XU D M, XIA W P, WANG W C. Annual runoff prediction based on VMD-CNN-GRU model optimized by slime mould algorithm[J]. South-to-North Water Transfers and Water Science & Technology, 2022, 20(3): 429-439. (in Chinese)

# 基于黏菌算法优化 VMD-CNN-GRU 模型的年径流预测

徐冬梅,夏王萍,王文川

(华北水利水电大学水资源学院,郑州 450046)

**摘要:**为提高年径流预测精度,引入黏菌算法(slime mould algorithm, SMA)和变分模态分解算法(variational mode decomposition, VMD),提出一种基于卷积神经网络(convolutional neural network, CNN)和门控循环单元(gated recurrent unit, GRU)神经网络的组合预测模型(VMD-SMA-CNN-GRU)。利用 VMD 对径流数据进行分解;采用 SMA 优化 CNN-GRU 模型参数,构建模型对每个分量进行预测;各分量结果相加得到最终结果。以兰西水文站为例,将所建模型与 CEEMDAN (complete ensemble empirical mode decomposition with adaptive noise)-CNN-GRU、VMD-CNN-LSTM(long short-term memory)、VMD-LSTM、VMD-GRU、VMD-PSO (particle swarm optimization)-CNN-GRU、SMA-CNN-GRU 和 CNN-GRU 预测模型进行对比分析。结果表明:SMA 优化的 VMD-CNN-GRU 模型预测精度不仅高于上述 7 种模型,而且避免了人工试算确定 CNN-GRU 模型参数效率低的不足,为年径流预测提供了一种新方法。

**关键词:**变分模态分解算法;黏菌算法;卷积神经网络;门控循环单元神经网络;径流预测

中图分类号:TV214;P338 文献标志码:A 开放科学(资源服务)标志码(OSID):



中长期水文预报是水资源规划管理、防汛抗旱、水库优化调度的重要环节,其研究一直受到众多水文学者的广泛关注<sup>[1]</sup>。随着科学技术的不断发展,许多现代人工智能技术方法被应用于水文预报,例如 BP(back propagation)神经网络<sup>[2-3]</sup>、径向基神经网络<sup>[4-5]</sup>、Elman 神经网络<sup>[6-7]</sup>、支持向量机<sup>[8-9]</sup>、自适应模糊推理系统<sup>[10-11]</sup>、长短时记忆神经网络(long short-term memory, LSTM)<sup>[12-13]</sup>等。

卷积神经网络(convolutional neural network, CNN)是一种前馈式神经网络,能够充分挖掘数据之间的相关性,已被应用于物体识别<sup>[14]</sup>、故障诊断<sup>[15]</sup>、文本分类<sup>[16]</sup>等方面。门控循环单元神经网络(gated recurrent unit, GRU)是循环神经网络(re-

current neural network, RNN)衍生体系结构中的一种典型体系结构,常用于时间序列数据预测,能够解决梯度消失的问题,已在负荷预测<sup>[17]</sup>、气象要素预测<sup>[18]</sup>、风速预测<sup>[19]</sup>等领域得到应用。在径流预测方面的研究有:郭玉学等<sup>[20]</sup>利用多种递归神经网络对海岛水库进行径流预报,通过对比研究发现具有复杂神经元结构的长短时记忆神经网络和门控循环单元神经网络预报效果更好;李文武等<sup>[21]</sup>将相空间重构与变分模态分解(variational mode decomposition, VMD)和深度门控网络耦合,该模型对白山水库入库月径流量进行预测,在测试集上拟合效果较好;包苑村等<sup>[22]</sup>提出了一种 VMD-CNN-LSTM 组合模型,对渭河流域的 2 个水文站进行月径流预测,预

收稿日期:2021-11-04 修回日期:2022-03-11 网络出版时间:2022-03-17

网络出版地址:https://kns.cnki.net/kcms/detail/13.1430.TV.20220316.1009.002.html

基金项目:河南省重点研发与推广专项项目(202102310259;202102310588);河南省高校科技创新团队项目(18IRTSTHN009)

作者简介:徐冬梅(1977—),女,吉林双辽人,博士,教授,博士生导师,主要从事水资源系统分析等研究。E-mail:xudongmei@ncwu.edu.cn

通信作者:王文川(1976—),男,河南鹿邑人,博士,教授,博士生导师,主要从事工程水文与水资源利用等研究。E-mail:wangwen1621@

163.com

测精度对比其他模型更优且在预测峰值和谷值上优势更为明显。为了提高传统 CNN、GRU 单一模型预测精度,学者们已经提出卷积神经网络和门控循环单元神经网络的组合模型(CNN-GRU)并应用于多种领域。如:姚程文等<sup>[23]</sup>将 CNN-GRU 模型用于电力负荷数据预测,该模型在预测精度与预测效率方面有较好的结果;党建武等<sup>[24]</sup>将 CNN-GRU 模型用于股指预测,有效提高了股指预测的准确率;针对出菇房室内温湿度具有非线性等特点,赵全明等<sup>[25]</sup>利用 CNN-GRU 模型预测出菇房多点温湿度,相比较传统的 BP 神经网络等 CNN-GRU 模型预测精度更高。

针对多数 CNN-GRU 混合神经网络超参数的设置方法存在精度和效率偏低的问题,本文在引入 CNN-GRU 组合模型进行年径流预测工作中,采用黏菌优化算法(slime mould algorithm, SMA)来确定关键参数,提出基于 SMA 优化的 VMD-CNN-GRU 组合模型对兰西水文站年径流进行预测,并构建 CEEMDAN (complete ensemble empirical mode decomposition with adaptive noise)-CNN-GRU、VMD-CNN-LSTM、VMD-GRU、VMD-LSTM、VMD-PSO (particle swarm optimization)-CNN-GRU、SMA-CNN-GRU 和 CNN-GRU 作为对比模型,以验证本文提出模型的有效性。

## 1 研究方法

### 1.1 变分模态分解算法(VMD)

VMD 是 2014 年提出的一种新的处理信号的算法<sup>[26]</sup>,分解信号的同时还能够降低输入信号中存在的噪声,因此在处理非线性非平稳序列上具有一定的优势,相较于 2011 年由 Torres 等<sup>[27]</sup>提出的自适应噪声完整集成经验模态分解算法(CEEMDAN),VMD 可以有效地分离信号。VMD 是给实际输入信号寻找一组最优重构的模式集合,而每个模式都被约束在一个估计的中心频率上。当给序列确定恰当的模式分解个数后,VMD 就可以将非线性的原始序列分解成若干个具有不同频率并相对平稳的子序列。详细的 VMD 实现步骤请参阅文献<sup>[26]</sup>。

### 1.2 卷积神经网络(CNN)

卷积神经网络<sup>[28]</sup>是一种独特的深度网络。输入层、全连接层和输出层的结构与其他神经网络的基本相同,独特之处在于 CNN 拥有池化层和卷积层的结构。池化层能够筛选过滤掉多余的信息,不仅减少全连接层中的参数个数防止过拟合问题的出现,而且缩短了训练的时间。卷积层中的卷积核(过滤器)可以实现对输入矩阵自动提取特征,并且卷积

层的权值将在神经元之间共享。

### 1.3 门控循环单元神经网络(GRU)

门控循环单元神经网络是对长短时记忆神经网络<sup>[29]</sup>进行了改进。GRU 和 LSTM 都是通过“门”函数来进行计算的,不同之处在于 GRU 比 LSTM 少了一个“门”简化了模型结构。GRU 只有两个门,即更新门和重置门。GRU 与 LSTM 运算效果相差不多,但由于 GRU 参数少故计算速度更快。GRU 神经网络内部结构见图 1。

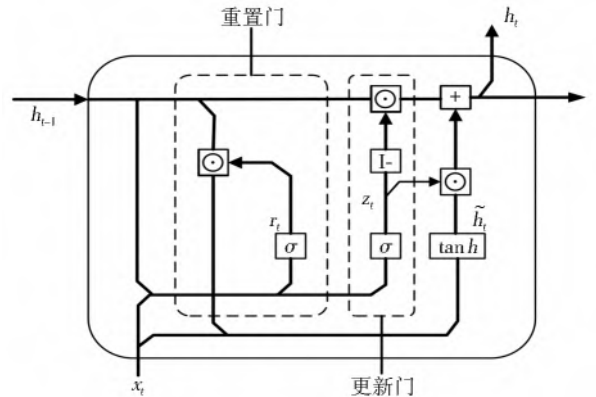


图 1 GRU 网络基本结构  
Fig. 1 Basic structure of GRU

GRU 的前向计算公式为

$$\text{更新门: } z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (1)$$

$$\text{重置门: } r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (2)$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t \cdot h_{t-1}, x_t]) \quad (3)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (4)$$

$$y_t = \sigma(W_o \cdot h_t) \quad (5)$$

式中: $x_t$  为输入向量; $W$  为权重参数; $h$  为隐藏层状态; $\cdot$  为两个向量点乘; $\sigma$  为 Sigmoid 激活函数; $\tilde{h}_t$  为隐藏状态的候选集; $y_t$  为输出向量。

### 1.4 黏菌算法(SMA)

黏菌优化算法由 Li 等<sup>[30]</sup>于 2020 年提出,根据黏菌多头绒泡菌在寻找食物过程中发生的一系列动作和身体上的变化来建立数学模型。文献<sup>[28]</sup>证明了该算法在实际问题中能够快速收敛并找到最优值。详细的数学模型如下。

阶段 1:接近食物。黏菌可以通过空气中的气味来寻找食物,具体用公式表示为

$$\vec{X}(t+1) = \begin{cases} \vec{X}_b(t) + \vec{v}_b \cdot (\vec{W} \cdot \vec{X}_A(t) - \vec{X}_B(t)), & r < p \\ \vec{v}_c \cdot \vec{X}(t), & r \geq p \end{cases} \quad (6)$$

式中: $\vec{v}_b$  的范围是  $[-a, a]$ ;  $\vec{v}_c$  的范围是  $[0, 1]$  并呈线性递减; $t$  为当前迭代次数; $\vec{X}_b$  为当前迭代中发现的气味浓度最高的个体的方位; $\vec{X}$  为黏菌的位置; $\vec{X}_A$  和  $\vec{X}_B$  为随机选择的两个黏菌的方位; $\vec{W}$  为权重。

其中,  $p, a, \vec{W}$  的值计算公式为

$$p = \tan h |S(i) - D_F|, i = 1, 2, \dots, n \quad (7)$$

式中:  $S(i)$  为每个黏菌的适应度;  $D_F$  为迭代过程中最佳适应度。

$$a = \arctan h \left( -\left( \frac{t}{\max_t} \right) + 1 \right) \quad (8)$$

$$\vec{W}(S_{\text{mellIndex}}) = \begin{cases} 1 + r \cdot \log \left( \frac{b_F - s(i)}{b_F - w_F} + 1 \right), C_{\text{condition}} \\ 1 - r \cdot \log \left( \frac{b_F - s(i)}{b_F - w_F} + 1 \right), O_{\text{thers}} \end{cases} \quad (9)$$

$$S_{\text{mellIndex}} = \text{sort}(S) \quad (10)$$

式中:  $C_{\text{condition}}$  表示适应度值排在种群数前 1/2 的个体;  $r$  为在  $[0, 1]$  上的随机数;  $\max_t$  为最大迭代次数;  $b_F$  (bestFitness) 和  $w_F$  (westFitness) 为当前迭代中的最优和最差适应度值;  $S_{\text{mellIndex}}$  为排序后的适应度值序列。

阶段 2: 包围食物。搜索食物时, 黏菌体内会受到食物浓度的影响产生一种信号。黏菌静脉接触的食物浓度越高, 体内产生的波越强, 细胞质流动越快, 静脉变得越粗。当食物浓度较高时, 该区域附近的权重较大; 当食物浓度较低时, 该区域的权重会降低。更新黏菌位置的计算公式为

$$\vec{W}^* = \begin{cases} r_{\text{and}} \cdot (U_B - L_B), r_{\text{and}} < z \\ \vec{X}_b(t) + \vec{v}_b \cdot (W \cdot \vec{X}_A(t) - \vec{X}_B(t)), r < p \\ \vec{v}_c \cdot \vec{X}(t), r \geq p \end{cases} \quad (11)$$

式中:  $L_B$  和  $U_B$  分别为搜索范围的上下边界值;  $r_{\text{and}}$  和  $r$  为  $[0, 1]$  中的随机数。

SMA 算法进行参数寻优的步骤如下。

步骤 1: 初始化 SMA 参数。设置种群规模  $n$ 、最大迭代次数  $M$ 、优化参数的个数及其范围, 令当前的迭代次数  $t = 1$ 。初始化个体, 随机生成  $n$  个黏菌个体的初始化位置。

步骤 2: 选用均方误差作为优化目标函数。将每个黏菌个体位置向量依次作为 CNN-GRU 模型的前提参数, 根据均方误差的公式分别计算出各自对应的适应度值, 然后对适应度值进行排序, 选出最优和最差适应度值。

步骤 3: 更新黏菌位置。利用式(9)计算  $W$  并更新最佳个体位置和适应度值, 然后利用式(8)更新  $a$  的值。若  $r_{\text{and}} < z$ , 则用式(11)①更新个体, 反之则更新  $p, \vec{v}_b, \vec{v}_c$ ; 若  $r < p$ , 则用式(11)②更新个体, 反之则用式(11)③进行计算。

步骤 4: 令  $t = t + 1$ 。判断  $t$  是否小于最大迭代次数  $M$ , 若是则重复步骤 2 和步骤 3, 反之则输出最优个体, 即算法的最优解, 算法结束。

## 1.5 组合模型构建

基于 SMA 优化的 VMD-CNN-GRU 模型的构建流程图见图 2, 具体的预测步骤总结如下。

采用 VMD 方法对原始的径流序列进行分解后得到若干子序列。

建立 SMA-CNN-GRU 模型。设置黏菌算法中的参数, 再利用黏菌算法对 CNN-GRU 模型的卷积层层数  $N$ 、GRU 层神经元个数  $H$ 、训练次数  $E$  和学习率  $\eta$  进行参数寻优。具体步骤见 1.4 节。

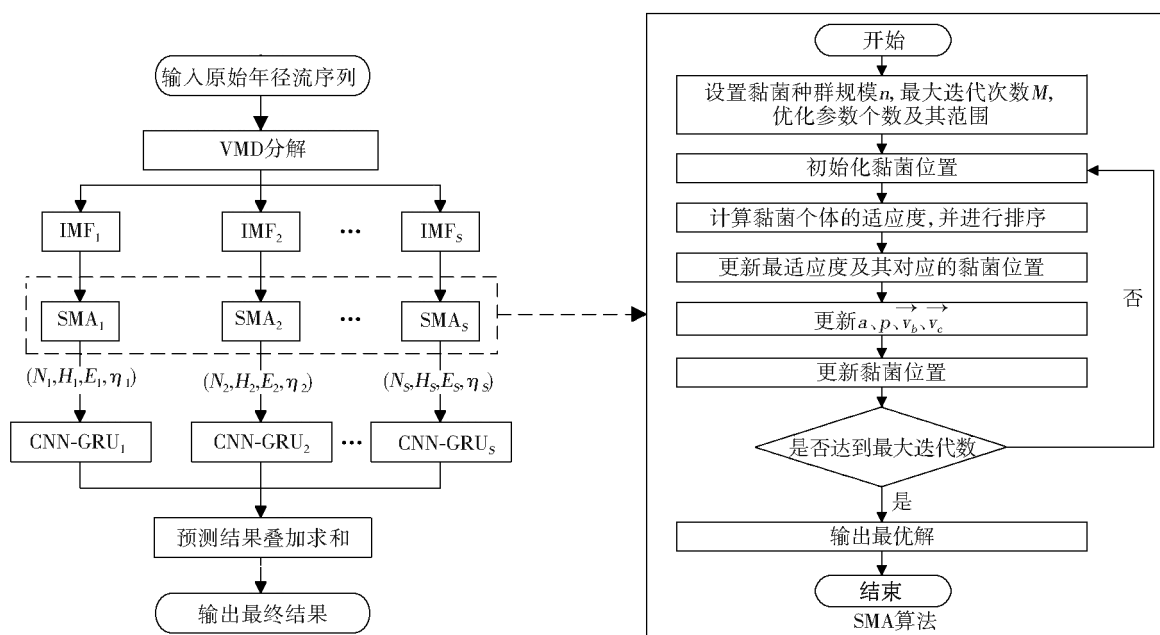


图 2 预测模型构建

Fig. 2 Prediction model construction

根据建立的 SMA-CNN-GRU 模型对预处理后的每个分量进行拟合得到预测值。

将所有预测后的子序列进行进一步累积得到最终结果。

### 1.6 模型验证

采用均方根误差 ( $E_{RMS}$ )、平均绝对误差 ( $E_{MA}$ )、平均绝对百分误差 ( $E_{MAP}$ ) 对模型进行评价。  $E_{RMS}$ 、 $E_{MA}$ 、 $E_{MAP}$  的计算公式为

$$E_{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

$$E_{MA} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

$$E_{MAP} = \frac{1}{N} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (14)$$

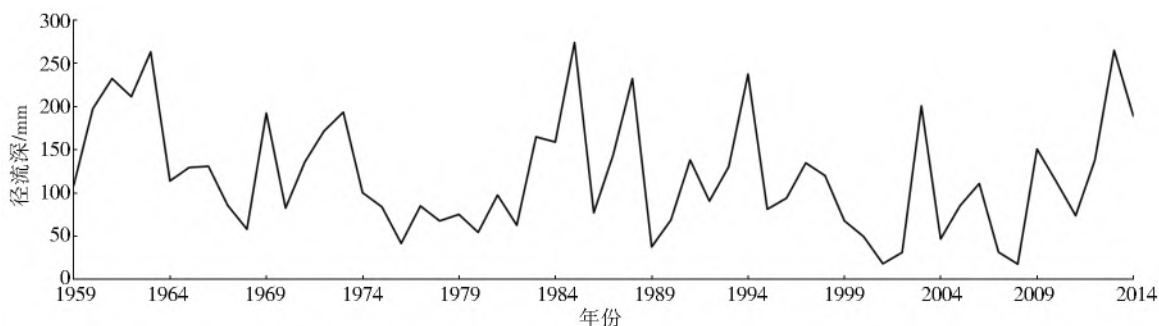


图 3 兰西水文站年径流序列

Fig. 3 Annual runoff series of Lanxi hydrological station

### 2.1 VMD 分解时序数据

VMD 分解径流序列的关键是提前设置合适的模态个数  $K$ ，不同分解个数会影响分解的结果，进而影响最终的预测结果。  $K$  值偏高可能导致模态混合或纯噪声模态，  $K$  值偏低可能导致模态的重复。通过比较各分量中心频率的方法来确定最终的模态个数  $K^{[31]}$ 。当选择不同的  $K$  值时，对兰西水文站年径流序列进行 VMD 分解后得到各子序列的中心频率，见表 1。从表 1 可以看出，在  $K=7$  时，出现的中心频率 3 156 Hz 和 3 492 Hz 相差不多，说明此时可能出现模态混叠现象，所以选  $K=6$  较为合适。

表 1 不同  $K$  值对应的中心频率

模态个数 $K$ /个	中心频率/Hz						
2	3	3 089					
3	3	1 745	3 219				
4	2	976	2 771	3 304			
5	2	945	1 883	3 196	3 986		
6	2	932	1 847	2 773	3 248	4 647	
7	1	508	1 056	1 901	3 156	3 492	4 704

式中： $y_i$  表示  $i$  时刻实际值； $\hat{y}_i$  表示  $i$  时刻模型预测值。

## 2 实例应用

兰西站是呼兰河下游的一个水文站，断面上集水面积为 27 736 km<sup>2</sup>，距河源 464 km。呼兰河是松花江左岸一级支流，全长 515 km，集水面积为 36 631 km<sup>2</sup>。以兰西水文站 1959—2014 年的年径流资料为研究对象，取 1959—2002 年年径流数据为训练集，2003—2014 年年径流数据为测试集。1959—2014 年兰西水文站年径流序列见图 3。

图 4 是分解后 6 个不同频率的子序列。对比图 3 原径流序列振幅变化没有明显的规律，由图 4 看出随着模态分量的增加，序列振幅变化呈现出周期性，序列越来越稳定，这个预处理使 CNN-GRU 模型可以更好地模拟数据中的特征信息并进行预测。

### 2.2 参数设置

在 CNN-GRU 模型中，根据一般经验，每层卷积核的数量为上一层的两倍。根据党建武等<sup>[24]</sup>的研究，再结合本文实例得到设置不同卷积层时各层对应的卷积核个数，具体数值见表 2。

表 2 不同卷积层对应的卷积核个数

卷积层层数/层	各层对应卷积核个数/个				
	第一层	第二层	第三层	第四层	第五层
2	32	64			
3	16	32	64		
4	16	32	64	128	
5	8	16	32	664	128

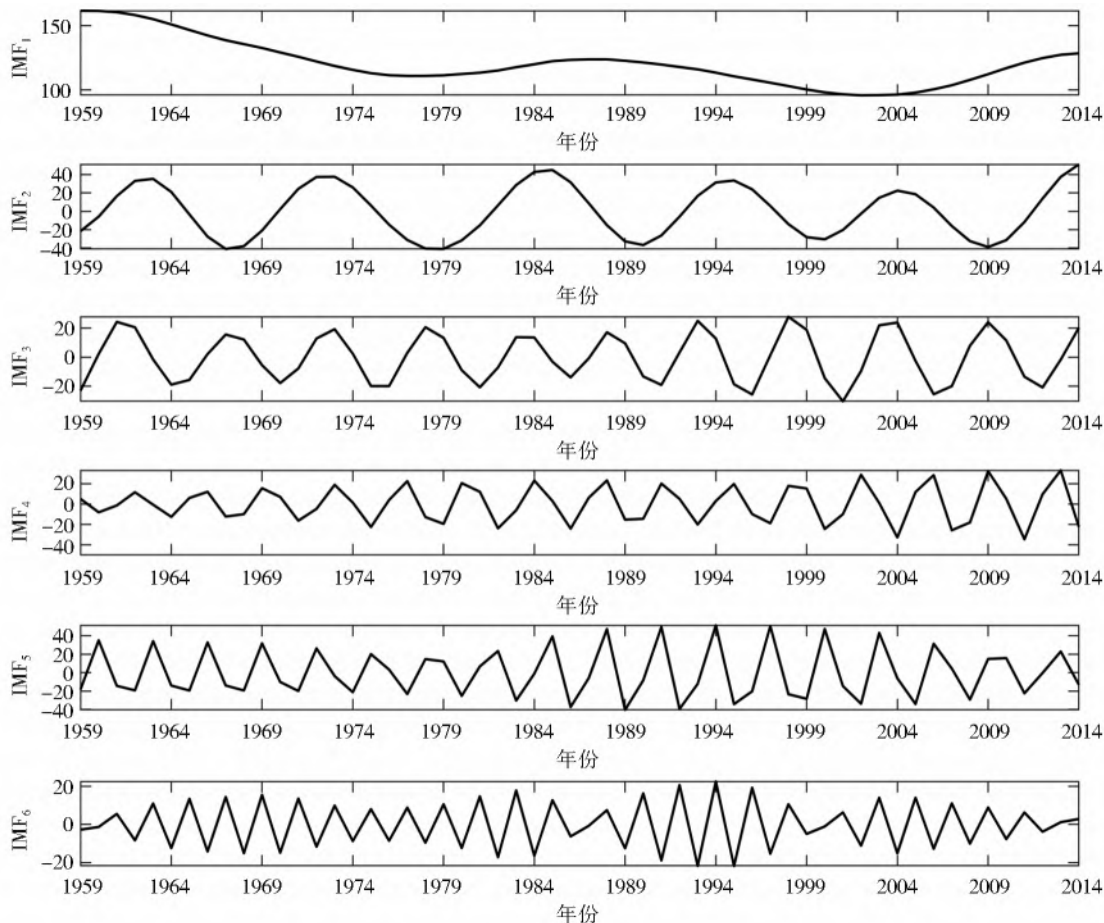


图4 兰西水文站年径流 VMD 分解图

Fig. 4 VMD decomposition diagrams of annual runoff at Lanxi hydrological station

然而对于设置卷积层层数、GRU 层神经元个数、训练次数、学习率的不确定性,本文提出采用黏菌优化算法来确定 CNN-GRU 模型卷积层层数等关键参数。设置 SMA 算法黏菌种群规模  $n=10$ ,最大迭代次数  $M=20$ ;卷积层层数  $N$ 、GRU 层神经元个数  $H$ 、训练次数  $E$ 、学习速率取值范围分别为  $[2,5]$ ,  $[100,200]$ ,  $[400,500]$ ,  $[0.005,0.010]$ 。为了确定合适的滞时来预测当前的径流,采用自相关函数 (autocorrelation function, ACF) 和偏自相关函数 (partial autocorrelation function, PACF) 确定 CNN-GRU 模型的输入步长。兰西水文站年径流序列的 ACF 图和 PACF 图见图 5。从图 5 可以看出,ACF 的估计值在 17 时达到峰值,并且在 17 个滞时以后其 PACF 的估计值均落在 95% 的置信带内。因此,时间步长取 17,即根据前 17 年的年径流来预测下一年的年径流。

由于原始序列经 VMD 分解后生成了 6 个子序列,针对不同序列该如何设置参数的问题,采用两种方案构建 VMD-SMA-CNN-GRU 的预测模型。方案 1 (VMD-SMA-CNN-GRU<sub>1</sub>) 采用对不同的序列设置不同的参数,方案 2 (VMD-SMA-CNN-GRU<sub>2</sub>)

采用对不同的序列设置统一的参数,不同方案下 SMA 对各子序列参数进行寻优后结果见表 3。除加入优化算法的模型以外,其他 6 种对比模型中,卷积层层数、GRU 层神经元个数、训练次数、学习率分别设置为各取值范围的中间值。因此,CNN 层数为 4,GRU 层和 LSTM 层神经元个数均为 150,训练次数为 450,学习率为 0.007 5。

### 2.3 结果分析

采用实测数据逐步分析验证。利用训练好的 CNN-GRU 模型先提取训练集中最后 17 年的数据,以预测第一个新值,接着取第二个所预测数据前面 17 年的实测数据来计算当前年的预测值,重复上述操作直至预测完验证期全部数据为止。选择 2003—2014 年的预报结果验证模型预测精度,即预测到第 12 年结束。

为了说明 SMA 优化的 VMD-CNN-GRU 模型的优势,使用 CEEMDAN-CNN-GRU 模型、VMD-CNN-LSTM 模型、VMD-LSTM 模型、VMD-GRU 模型、VMD-PSO-CNN-GRU 模型、SMA-CNN-GRU 模型和 CNN-GRU 模型作为基准进行对比。各模型预测后评价标准计算结果和对比见表 4 和图 6。

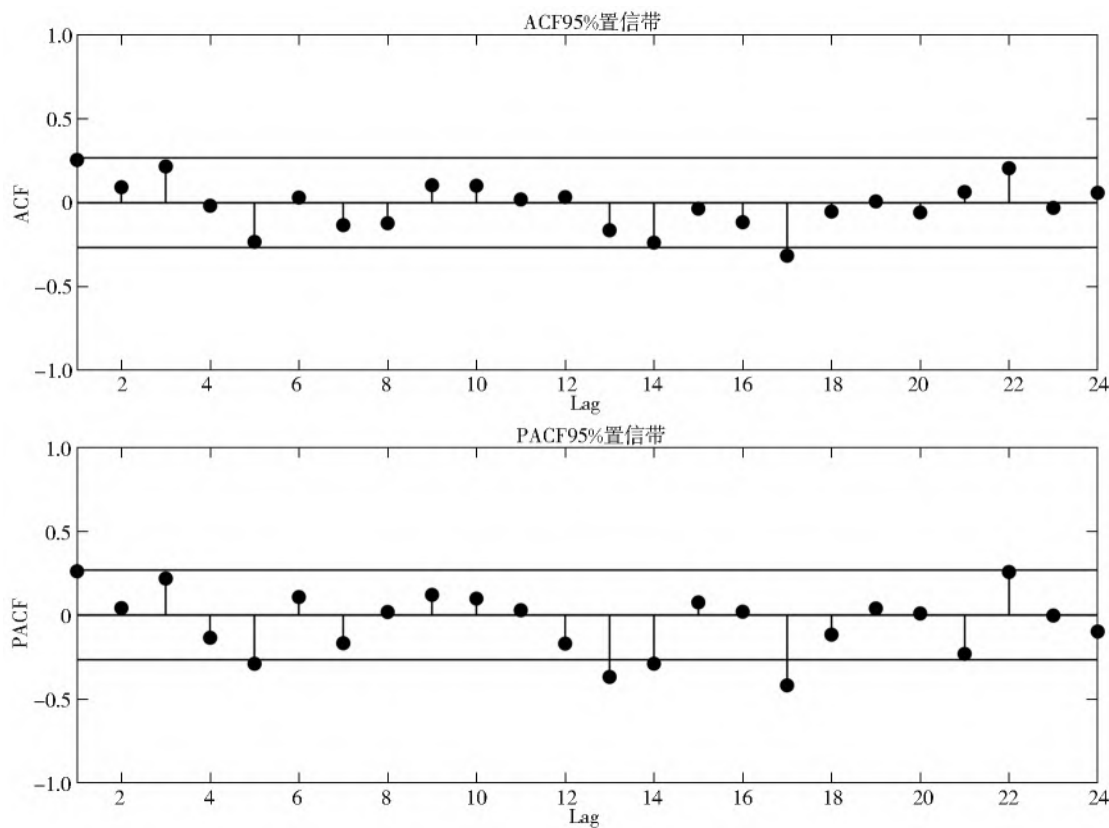


图 5 年径流序列 ACF 图和 PACF 图

Fig. 5 ACF and PACF plots of annual runoff series

表 3 不同方案下 SMA 对各子序列参数寻优结果

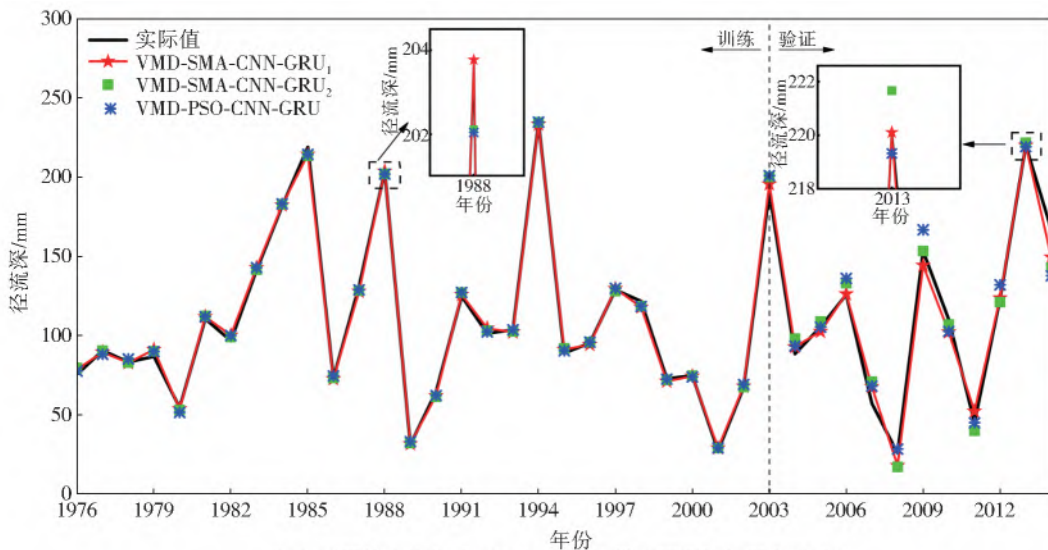
Tab. 3 Optimization results of sub-sequence parameters by SMA under different schemes

方案	子序列	参数寻优结果			
		CNN 层层数 $N$ /层	GRU 层神经元个数 $H$ /个	训练次数 $E$ /次	学习速率 $\eta$
1	IMF <sub>1</sub>	4	100	400	0.005 0
	IMF <sub>2</sub>	5	100	400	0.005 0
	IMF <sub>3</sub>	4	167	467	0.005 0
	IMF <sub>4</sub>	4	170	470	0.005 0
	IMF <sub>5</sub>	5	114	436	0.009 6
	IMF <sub>6</sub>	3	165	445	0.005 6
2	IMF <sub>1-6</sub>	5	200	400	0.005 0

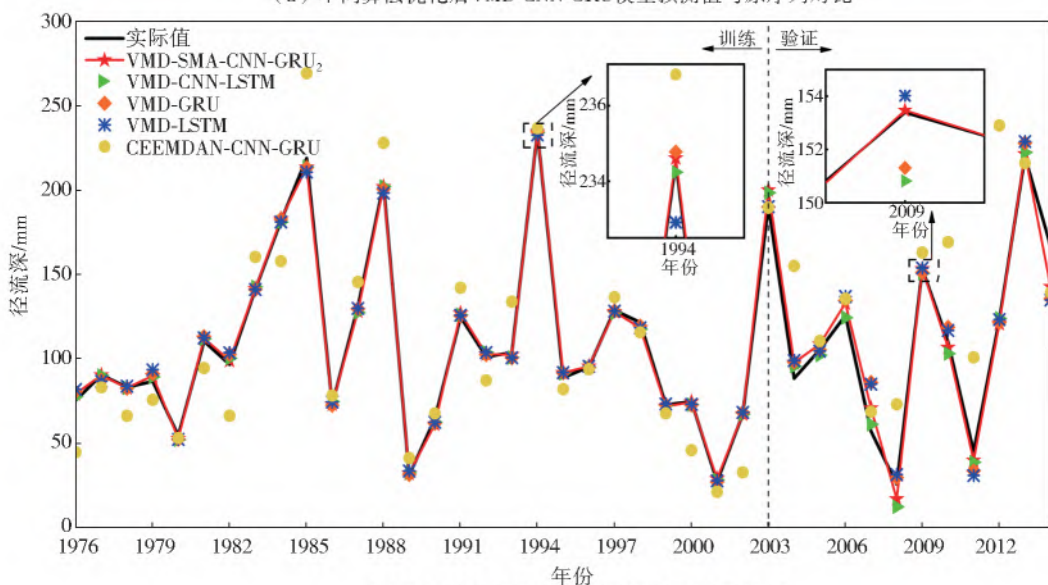
表 4 各模型评价标准计算结果

Tab. 4 Calculation results of each model evaluation standard

模型	训练期			验证期		
	$E_{RMS}/mm$	$E_{MA}/mm$	$E_{MAP}/\%$	$E_{RMS}/mm$	$E_{MA}/mm$	$E_{MAP}/\%$
VMD-SMA-CNN-GRU <sub>1</sub>	2.18	1.74	1.93	7.77	6.10	8.26
VMD-SMA-CNN-GRU <sub>2</sub>	2.13	1.64	1.85	9.71	7.35	9.47
VMD-PSO-CNN-GRU	2.10	1.83	2.18	11.29	8.28	7.47
VMD-CNN-LSTM	2.20	1.81	1.99	10.30	7.11	9.71
VMD-GRU	2.61	1.97	2.07	13.76	9.57	11.83
VMD-LSTM	3.39	2.53	2.75	14.14	9.96	12.84
CEEMDAN-CNN-GRU	19.87	15.72	16.43	48.10	34.45	49.02
SMA-CNN-GRU	20.59	16.02	16.93	48.87	42.53	62.48
CNN-GRU	20.09	15.33	15.97	65.64	57.30	72.80



(a) 不同算法优化后VMD-CNN-GRU模型预测值与原序列对比



(b) SMA优化前后各模型预测值与原序列对比

图6 各模型预测值与原序列对比

Fig. 6 Comparison between the predicted values of each model and the original sequence

从图 6(a)可以看出优化后的 VMD-CNN-GRU 模型拟合效果均很好,相比较之下 VMD-PSO-CNN-GRU 模型的效果略差,从局部放大图中可以明显看出 1988 年和 2013 年 VMD-SMA-CNN-GRU<sub>1</sub> 预测结果比其他 2 个模型更精确。从图 6(b)可以看出:VMD-SMA-CNN-GRU<sub>2</sub> 模型拟合的效果最好,尤其在预测峰值的情况下;CEEMDAN-CNN-GRU 模型拟合效果最差,只有曲线的大概趋势与原序列相同而具体数值与实际值相差较大。在训练期,除 CEEMDAN-CNN-GRU 模型以外,其他模型均达到较好的拟合效果,但通过局部放大 1994 年的预测结果后可以发现各模型之间还存在一定的差异,VMD-SMA-CNN-GRU<sub>2</sub> 模型预测精度更高。从验证期中可以看出,VMD-SMA-CNN-GRU<sub>2</sub> 和 VMD-CNN-LSTM 模型拟合程度优于 VMD-GRU

和 VMD-LSTM 模型。

由表 4 可以得出如下结果。

经 SMA 优化的 VMD-CNN-GRU 组合模型预测精度最高。对比 2 种不同方案下 VMD-SMA-CNN-GRU 的预测结果可知各子序列设置不同参数时可以提高预测精度。VMD-SMA-CNN-GRU<sub>1</sub> 与 VMD-PSO-CNN-GRU 模型相比, $E_{RMS}$ 、 $E_{MA}$  分别减少了 31.18%、26.33%,说明粒子群优化算法<sup>[32]</sup>(PSO)容易陷入局部最优无法找到最优解,而 SMA 算法能避免这一问题并有效的寻找优化参数。

VMD-SMA-CNN-GRU<sub>2</sub> 模型与 SMA-CNN-GRU 模型相比, $E_{RMS}$ 、 $E_{MA}$  和  $E_{MAP}$  分别减少了 80.13%、82.72%和 84.84%,可以看出分解技术和优化算法均能提高模型的预测精度,但通过比较发现分解技术的提高程度更大。

VMD-SMA-CNN-GRU<sub>2</sub> 模型与 CEEMDAN-CNN-GRU 模型相比,  $E_{RMS}$ 、 $E_{MA}$  和  $E_{MAP}$  分别减少了 79.81%、78.66% 和 80.68%; 与 CNN-GRU 模型相比,  $E_{RMS}$ 、 $E_{MA}$  和  $E_{MAP}$  分别减少了 85.21%、87.17% 和 86.99%。VMD-SMA-CNN-GRU<sub>2</sub> 模型预测精度明显高于 CEEMDAN-CNN-GRU 模型和 CNN-GRU 模型, 说明了 VMD 分解效果优于 CEEMDAN。

VMD-GRU 模型和 VMD-LSTM 模型相比预测精度相差不大, 但是 VMD-GRU 模型与 VMD-LSTM 模型训练时间分别为 38.55 s 和 57.03 s, 相比训练时间缩短了 32.40%, 通过对比两者模型的训练时间可以发现 GRU 模型效率更高。GRU 比 LSTM 训练时间短的原因是 GRU 模型只有两个门, 模型结构简单, 构建庞大的网络时更加有效。

VMD-SMA-CNN-GRU<sub>2</sub> 模型与 VMD-GRU 模型相比,  $E_{RMS}$ 、 $E_{MA}$  和  $E_{MAP}$  分别减少了 29.43%、23.20% 和 19.95%, 与 VMD-LSTM 模型相比  $E_{RMS}$ 、 $E_{MA}$  和  $E_{MAP}$  分别减少了 31.33%、26.20% 和 26.25%。加入 CNN 网络后模型预测精度明显提高, 这是因为 CNN 模型中卷积层中的卷积核(滤波器)发挥着提取数据特征作用, 通过卷积提取需要的特征然后传递给 GRU 层或者 LSTM 层。

综上所述, 在验证期各种模型对兰西水文站年径流预测精度存在的差异较大一些, 预测效果最好的方法依次为 VMD-SMA-CNN-GRU、VMD-PSO-CNN-GRU、VMD-CNN-LSTM、VMD-GRU、VMD-LSTM、CEEMDAN-CNN-GRU、SMA-CNN-GRU、CNN-GRU。年径流序列由于受多种因素的影响极其不平稳, VMD 可以自适应处理年径流序列, 对序列进行重构以达到去噪的效果, CNN 通过提取数据内部的特征使得 GRU 后续能够更加高效地进行预测, 而 SMA 算法为 CNN-GRU 模型确定了恰当的参数, 因此 VMD-SMA-CNN-GRU 模型具有良好的适应性和预测性能。

### 3 结 论

在 VMD-CNN-GRU 模型中, 隐含层神经元个数、训练次数等超参数的设置影响模型拟合效果, 利用 SMA 算法对 CNN-GRU 混合神经网络中超参数进行寻优改善了模型的性能, 避免了人工试算效率低的情况, 节约了运算时间成本同时提高了模型的预测精度。

分解技术和优化算法均能提高模型的预测精度, 并且分解技术的作用影响更大。径流序列具有

非线性和非平稳的特点, VMD 可以将原始序列分解成较为平稳的多个子序列, 增加数据的同时能够更好地进行拟合。VMD 与 CEEMDAN 相比, 在分解时 VMD 可以自主地设置合适的模态个数, 能够有效解决模态混叠的问题, 使得分解精度更高进而预测结果更准确。

VMD-CNN-GRU 模型同时具有 CNN 网络和 GRU 网络的优点。CNN 网络可以充分挖掘数据中的特征, GRU 网络适用于时间序列预测, 试验结果表明两者的结合可以有效的预测年径流, 为中长期径流预报提供了一种新途径。

### 参考文献(References):

- [1] 徐冬梅, 赵晓慎. 中长期水文预报方法研究综述[J]. 水利科技与经济, 2010, 16(1): 1-7. (XU D M, ZHAO X S. Review on study of mid and long-term hydrological forecasting technique[J]. Water Conservancy Science and Technology and Economy, 2010, 16(1): 1-7. (in Chinese)) DOI: 10.3969/j.issn.1006-7175.2010.01.001.
- [2] 王文川, 李文锦, 徐冬梅, 等. 基于马尔可夫链校正 GM-BP 模型的径流预测[J]. 南水北调与水利科技, 2019, 17(5): 44-49. (WANG W C, LI W J, XU D M, et al. Runoff prediction based on GM-BP model calibration against Markov chain[J]. South-to-North Water Transfers and Water Science & Technology, 2019, 17(5): 44-49. (in Chinese)) DOI: 10.13476/j.cnki.nsb-dqk.2019.0108.
- [3] 赵文刚, 刘晓群, 宋雯, 等. 基于 PSO-BP 神经网络的西洞庭湖南咀站径流预测[J]. 人民长江, 2019, 50(3): 124-130. (ZHAO W G, LIU X Q, SONG W, et al. Preliminary study on runoff forecast at Nanzui Station in West Dongting Lake based on PSO-BP neural network[J]. Yangtze River, 2019, 50(3): 124-130. (in Chinese)) DOI: 10.16232/j.cnki.1001-4179.2019.03.022.
- [4] 王丽学, 杨军, 孙靓, 等. 基于灰色系统与 RBF 神经网络的中长期水文预报[J]. 人民长江, 2015, 46(17): 15-17. (WANG L X, YANG J, SUN L, et al. Application of gray system and RBF neural network in the mid-long term hydrological forecasting[J]. Yangtze River, 2015, 46(17): 15-17. (in Chinese)) DOI: 10.16232/j.cnki.1001-4179.2015.17.004.
- [5] 刘蕊鑫, 纪昌明, 马源. 基于互信息的改进 RBF 神经网络预测模型及其应用[J]. 中国农村水利水电, 2016(10): 100-103, 115. (LIU R X, JI C M, MA Y. Mutual information-based amelioration of the nearest neighbor cluster RBF neural network prediction model[J]. China



- Rural Water and Hydropower, 2016 (10): 100-103, 115. (in Chinese)) DOI: 10. 3969/j. issn. 1007-2284. 2016. 10. 024.
- [6] 李继清,王爽,吴月秋,等. 径流预报的极点对称模态分解-Elman 网络模型[J]. 水力发电学报, 2021, 40(7): 13-22. (LI J Q, WANG S, WU Y Q, et al. Runoff forecasts using combined model of extreme-point symmetric mode decomposition and Elman neural network[J]. Journal of Hydroelectric Engineering, 2021, 40(7): 13-22. (in Chinese)) DOI: 10. 11660/slfdx. 20210702.
- [7] 李志新,赖志琴,龙云墨. 基于 GA-Elman 神经网络模型的年径流预测[J]. 水利水电技术, 2018, 49(8): 71-77. (LI Z X, LAI Z Q, LONG Y M. GA-Elman neural network model-based annual runoff prediction[J]. Water Resources and Hydropower Engineering, 2018, 49(8): 71-77. (in Chinese)) DOI: 10. 13928/j. cnki. wrahe. 2018. 08. 009.
- [8] 王文川,张洁铭,郑野,等. 基于马尔科夫链的 BP-SVM 模型的径流预测[J]. 水利水电技术, 2020, 51(11): 78-84. (WANG W C, ZHANG J M, ZHENG Y, et al. Runoff prediction based on BP-SVM model of Markov chain[J]. Water Resources and Hydropower Engineering, 2020, 51(11): 78-84. (in Chinese)) DOI: 10. 13928/j. cnki. wrahe. 2020. 11. 009.
- [9] 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. (JIN J, LIU Y, WANG Q H, et al. Mid-long term runoff forecasting based on wavelet de-noising and FA-SVM[J]. Yangtze River, 2020, 51(9): 67-72. (in Chinese)) DOI: 10. 16232/j. cnki. 1001-4179. 2020. 09. 013.
- [10] 谭乔凤,王旭,王浩,等. ANN、ANFIS 和 AR 模型在日径流时间序列预测中的应用比较[J]. 南水北调与水利科技, 2016, 14(6): 12-17, 26. (TAN Q F, WANG X, WANG H, et al. Comparative study of ANN, ANFIS and AR model for daily runoff time series prediction[J]. South-to-North Water Transfers and Water Science & Technology, 2016, 14(6): 12-17, 26. (in Chinese)) DOI: 10. 13476/j. cnki. nsbdqk. 2016. 06. 003.
- [11] 马盼盼,白涛,武连洲,等. 黄河源区河川径流短期预测的 ANFIS 模型[J]. 西北农林科技大学学报(自然科学版), 2018, 46(6): 145-154. (MA P P, BAI T, WU L Z, et al. Short-term runoff prediction by ANFIS model in source region of the Yellow River[J]. Journal of Northwest A&F University (Natural Science Edition), 2018, 46(6): 145-154. (in Chinese)) DOI: 10. 13207/j. cnki. jnwafu. 2018. 06. 019.
- [12] 殷兆凯,廖卫红,王若佳,等. 基于长短时记忆神经网络(LSTM)的降雨径流模拟及预报[J]. 南水北调与水利科技, 2019, 17(6): 1-9, 27. (YIN Z K, LIAO W H, WANG R J, et al. Rainfall-runoff modelling and forecasting based on Long Short-Term Memory (LSTM) [J]. South-to-North Water Transfers and Water Science & Technology, 2019, 17(6): 1-9, 27. (in Chinese)) DOI: 10. 13476/j. cnki. nsbdqk. 2019. 0129.
- [13] KRATZERT F, KLOTZ D, BRENNER C, et al. Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks[J]. Hydrology and Earth System Sciences, 2018, 22(11): 6005-6022. DOI: 10. 5194/hess-22-6005-2018.
- [14] 殷亚萍,柴文,凌毅德,等. 基于特征分析的卷积神经网络烟雾识别[J]. 无线电工程, 2021, 51(7): 526-533. (YIN Y P, CHAI W, LING Y D, et al. Convolutional neural network for smoke recognition based on feature analysis[J]. Radio Engineering, 2021, 51(7): 526-533. (in Chinese)) DOI: 10. 3969/j. issn. 1003-3106. 2021. 07. 002.
- [15] 陈剑,黄凯旋,吕伍佺,等. 基于 VMD 和卷积神经网络的变工况轴承故障诊断方法[J]. 计量学报, 2021, 42(7): 892-897. (CHEN J, HUANG K X, LYU W Y, et al. Bearing fault diagnosis method based on VMD and convolutional neural network undervarying operation conditions[J]. Acta Metrologica Sinica, 2021, 42(7): 892-897. (in Chinese)) DOI: 10. 3969/j. issn. 1000-1158. 2021. 07. 10.
- [16] 张昱,刘开峰,张全新,等. 基于组合-卷积神经网络的中文新闻文本分类[J]. 电子学报, 2021, 49(6): 1059-1067. (ZHANG Y, LIU K F, ZHANG Q X, et al. A combined convolutional neural network for Chinese news text classification[J]. Acta Electronica Sinica, 2021, 49(6): 1059-1067. (in Chinese)) DOI: 10. 12263/DZXB. 20200134.
- [17] 邓带雨,李坚,张真源,等. 基于 EEMD-GRU-MLR 的短期电力负荷预测[J]. 电网技术, 2020, 44(2): 593-602. (DENG D Y, LI J, ZHANG Z Y, et al. Short-term electric load forecasting based on EEMD-GRU-MLR[J]. Power System Technology, 2020, 44(2): 593-602. (in Chinese)) DOI: 10. 13335/j. 1000-3673. pst. 2019. 0113.
- [18] 王怡婷,陈曦,鞠兴旺,等. 基于 GRU 网络的气象要素预测算法[J]. 计算机仿真, 2021, 38(7): 419-423. (WANG Y T, CHEN X, JU X W, et al. A meteorological element prediction algorithm based on GRU network[J]. Computer Simulation, 2021, 38(7): 419-423. (in Chinese)) DOI: 10. 3969/j. issn. 1006-9348. 2021. 07. 089.
- [19] 金俊喆,武鹏,董祥祥,等. 基于 RF-GRU 风速预测的

- 风电 MPPT 控制[J]. 传感器与微系统, 2021, 40(5): 38-41. (JIN J Z, WU P, DONG X X, et al. Wind power MPPT control based on RF-GRU wind speed prediction[J]. Transducer and Microsystem Technologies, 2021, 40(5): 38-41. (in Chinese)) DOI: 10. 13873/J. 1000-9787(2021)05-0038-04.
- [20] 郭玉雪, 许月萍, 陈浩, 等. 基于多种递归神经网络的海岛水库径流预报[J]. 水力发电学报, 2021, 40(9): 14-26. (GUO Y X, XU Y P, CHEN H, et al. Reservoir inflow forecasting for island areas based on multiple recurrent neural networks[J]. Journal of Hydroelectric Engineering, 2021, 40(9): 14-26. (in Chinese)) DOI: 10. 11660/slfdbx. 20210902.
- [21] 李文武, 石强, 王凯, 等. 基于变分模态分解和深度门控网络的径流预测[J]. 水力发电学报, 2020, 39(3): 34-44. (LI W W, SHI Q, WANG K, et al. Runoff prediction based on variational mode decomposition and deep gated network[J]. Journal of Hydroelectric Engineering, 2020, 39(3): 34-44. (in Chinese)) DOI: 10. 11660/slfdbx. 20200304.
- [22] 包苑村, 解建仓, 罗军刚. 基于 VMD-CNN-LSTM 模型的渭河流域月径流预测[J]. 西安理工大学学报, 2021, 37(1): 1-8. (BAO Y C, XIE J C, LUO J G. Monthly runoff prediction of Weihe River basin based on VMD-CNN-LSTM model[J]. Journal of Xi' an University of Technology, 2021, 37(1): 1-8. (in Chinese)) DOI: 10. 19322/j. cnki. issn. 1006-4710. 2021. 01. 001.
- [23] 姚程文, 杨苹, 刘泽健. 基于 CNN-GRU 混合神经网络的负荷预测方法[J]. 电网技术, 2020, 44(9): 3416-3424. (YAO C W, YANG P, LIU Z J. Load forecasting method based on CNN-GRU hybrid neural network[J]. Power System Technology, 2020, 44(9): 3416-3424. (in Chinese)) DOI: 10. 13335/j. 1000-3673. pst. 2019. 2058.
- [24] 党建武, 从筱卿. 基于 CNN 和 GRU 的混合股指期货预测模型研究[J]. 计算机工程与应用, 2021, 57(16): 167-174. (DANG J W, CONG X Q. Research on hybrid stock index forecasting model based on CNN and GRU[J]. Computer Engineering and Applications, 2021, 57(16): 167-174. (in Chinese)) DOI: 10. 3778/j. issn. 1002-8331. 2004-0236.
- [25] 赵全明, 宋子涛, 李奇峰, 等. 基于 CNN-GRU 的菇房多点温湿度预测方法研究[J]. 农业机械学报, 2020, 51(9): 294-303. (ZHAO Q M, SONG Z T, LI Q F, et al. Multi-point prediction of temperature and humidity of mushroom based on CNN-GRU[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9): 294-303. (in Chinese)) DOI: 10. 6041/j. issn. 1000-1298. 2020. 09. 033.
- [26] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544. DOI: 10. 1109/tsp. 2013. 2288675.
- [27] TORRES M E, COLOMINAS M A, SCHLOTTHAUER G, et al. A complete ensemble empirical mode decomposition with adaptive noise[C]// 2011 IEEE International Conference on Acoustics, Prague, Czech Republic: IEEE, 2011: 4144-4147. DOI: 10. 1109/IC-ASSP. 2011. 5947265.
- [28] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324. DOI: 10. 1109/9780470544976. ch9.
- [29] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780. DOI: 10. 1162/neco. 1997. 9. 8. 1735.
- [30] LI S M, CHEN H L, WANG M J, et al. Slime mould algorithm: A new method for stochastic optimization [J]. Future Generation Computer Systems-the International Journal of Escience, 2020, 111: 300-323. DOI: 10. 1016/j. future. 2020. 03. 055.
- [31] 陈东宁, 张运东, 姚成玉, 等. 基于变分模态分解和多尺度排列熵的故障诊断[J]. 计算机集成制造系统, 2017, 23(12): 2604-2612. (CHEN D N, ZHANG Y D, YAO C Y, et al. Fault diagnosis method based on variational mode decomposition and multi-scale permutation entropy[J]. Computer Integrated Manufacturing Systems, 2017, 23(12): 2604-2612. (in Chinese)) DOI: 10. 13196/j. cims. 2017. 12. 005.
- [32] KENNEDY J, EBERHART R. Particle swarm optimization[C]// 1995 IEEE International Conference on Neural Networks Proceedings, Piscataway: IEEE, 1995: 1942-1948. DOI: 10. 2514/2. 2111.

### Annual runoff prediction based on VMD-CNN-GRU model optimized by slime mould algorithm

XU Dongmei, XIA Wangping, WANG Wenchuan

(College of Water Resources, North China University of Water Resources and Electric Power, Zhengzhou 450046, China)

**Abstract:** Medium and long-term hydrological forecasting is an essential link in management, optimization of water resources, flood control, drought relief, and reservoir optimal operation. With the rapid development of science and technology, many modern artificial intelligence (AI) models have been applied to hydrological forecasting, such as back-propagation artificial neural

network, support vector machine and long short-term memory neural network. Among the AI models, convolutional neural network (CNN) is a unique deep network, which can fully excavate the correlation between data. Gated recurrent unit neural network (GRU), a kind of the recurrent neural network, is a variant of long short-term memory neural network (LSTM). GRU is often used in time-series data prediction and can solve the problem of gradient disappearance. The combined model of convolutional neural network and gated recurrent unit neural network (CNN-GRU) was applied in various fields except runoff prediction. Additionally, for the setting of the parameters of CNN-GRU hybrid neural network, most people used the control variable method for trial calculation, which was not only low in efficiency and low in accuracy. Hence, a combined prediction model (VMD-SMA-CNN-GRU) based on convolutional neural network and gated recurrent unit neural network was proposed by introducing slime mould algorithm (SMA) and variational mode decomposition (VMD).

The four main steps of the present VMD-SMA-CNN-GRU forecasting model can be summarized as follows: The original runoff series was decomposed by VMD to obtain several intrinsic mode functions and a residual. The slime mould population size  $n$  and the maximum iteration  $M$  was set. Subsequently, SMA was used to optimize key parameters such as the number of convolution layers, the number of neurons in the hidden layer of GRU, training times and learning rate. The mean square error (MAE) was chosen as the objective function of the optimization algorithm. SMA-CNN-GRU model was used to predict all the subseries. The predicted values obtained above were accumulated to deduce the ultimate prediction results.

Lanxi hydrological station was selected as an example to illustrate the advantages of VMD-SMA-CNN-GRU model using annual runoff data from 1959 to 2014. The data from 1959 to 2002 was selected as a training set while from 2003 to 2014 took as a test set. The proposed hybrid model was compared with CEEMDAN (Complete ensemble empirical mode decomposition with adaptive noise)-CNN-GRU model, VMD-CNN-LSTM model, VMD-LSTM model, VMD-GRU model, VMD-PSO (Particle swarm optimization)-CNN-GRU model, SMA-CNN-GRU model, and CNN-GRU model. Additionally, three standard statistical performances measures, namely root mean squared error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) were employed to evaluate the performances of the six models mentioned above. In the training phase, all models except the CEEMDAN-CNN-GRU model achieved a better fit. In the validation phase, the VMD-SMA-CNN-GRU model predicted significantly better than several other models, especially in the case of predicted peaks. Results show that the SMA optimized VMD-CNN-GRU model has the best fitting effect, and the prediction accuracy is greater than that of the above seven comparison models.

According to the results, there are several conclusions as follows: Runoff series has the characteristics of nonlinearity and non-stationarity. Therefore, VMD can decompose the original sequence into several relatively stable sub-sequences, which can be better fitted while increasing data. Low efficiency of manual trial calculations is avoided by SMA to determine the parameters of CNN-GRU model. The CNN-GRU model has the advantage of utilizing two neural networks simultaneously. The VMD-SMA-CNN-GRU model established effectively improves the accuracy of runoff prediction and provides a new method for annual runoff prediction.

**Key words:** variational mode decomposition; slime mould algorithm; convolutional neural network; gated recurrent unit; runoff prediction