

DOI: 10.20040/j.cnki.1000-7709.2023.20221932

一种基于物理核函数高斯过程回归的月径流 预报模型及其应用

孙娜^{1a}, 张楠^{1b}, 张帅², 彭甜^{1a}, 周建中³, 张海荣⁴

(1. 淮阴工学院 a. 自动化学院; b. 机械与材料工程学院, 江苏 淮安 223003;
2. 西安数峰科技信息有限公司, 陕西 西安 710054; 3. 华中科技大学土木与水利工程学院, 湖北 武汉 430074;
4. 中国长江电力股份有限公司, 湖北 宜昌 443133)

摘要: 鉴于传统的单一径流预报模型很难描述径流未来变化规律, 将自适应变分模态分解(AVMD)与基于组合物理核函数的高斯过程回归(GPR-CK)相结合, 构建了 AVMD-GPR-CK 预报模型, 该模型采用 AVMD 将实测径流分解为多个子序列, 对子序列依据其自身特点分别建模, 子序列预报结果叠加重构即为最终预报结果。模型应用于金沙江流域向家坝站未来 1~12 个月的径流预报的结果表明, 所有预见期 AVMD-GPR-CK 模型的不确定性系数均大于 0.94, 平均绝对百分比误差(M_{MAPE})在 $\pm 17\%$ 以内, 预见期在 10 个月以内时, M_{MAPE} 在 $\pm 10\%$ 以内; 预报精度明显优于常见的 BP、GRNN、RBF、RELM 模型。

关键词: 月径流预报; 变分模态分解; 高斯过程回归; 组合核函数

中图分类号: TV124; P338

文献标志码: A

文章编号: 1000-7709(2023)04-0039-05

1 引言

高精度长预见期的径流预报可为流域防洪抗旱、梯级水库联调、水资源开发与管理等提供重要依据^[1]。由于径流影响因子众多, 且受气候变化和人类活动双重胁迫, 径流常呈较强的非线性、非平稳性、多尺度及随机性等特征, 这导致传统的单一模型很难描述其未来变化规律, 难以获得理想的预报精度。近年来, 结合两种及以上方法优点的耦合预报模型受到广泛关注^[2]。蔡文静等^[3]分别将经验模态分解(EMD)、变分模态分解(VMD)、离散小波变换(DWT)时频分析方法与 LSTM 模型组合, 通过实例应用说明 VMD 更适合径流预报; 周靖楠等^[4]将输入变量选择方法与高斯过程回归(GPR)相结合取得了较理想的预报结果。为提高月径流预报精度, 针对径流的非平稳性及其蕴含的周期性、趋势性和随机性, 本文设计了一种新的物理信息驱动的组合核函数

(CK), 构建了 AVMD-GPR-CK 模型, 将其应用于金沙江流域开展预报试验, 并与常用模型 BP、GRNN、RELM、RBF 模型进行对比, 验证其预报性能, 以期为中长期径流预报提供新方法。

2 研究方法

2.1 自适应变分模态分解

VMD 为一种准正交、自适应时间序列分解方法。将观测径流 $Q(t)$ 分解为 d 个有限带宽本征模态函数 $u_k(t)$ ($k=1, 2, \dots, d$) 的主要步骤如下。

步骤 1 利用 Hilbert 变换计算每个 $u_k(t)$ 对应的解析信号, 从而得到其单侧频谱 $[\delta(t) + j/(\pi t)] * u_k(t)$, 其中, $\delta(t)$ 为狄拉克函数; j 为旋转因子, 表示相位移动 $\pi/2$ 。

步骤 2 将每个 $u_k(t)$ 与其对应中心频率 ω_k 的指数项混叠, 调制到对应的基频带 $\{[\delta(t) + j/(\pi t)]u_k(t)\}e^{-j\omega_k t}$ 。

步骤 3 计算解调信号的高斯平滑度估计带

收稿日期: 2022-09-17, **修回日期:** 2022-10-19

基金项目: 江苏省高校自然科学基金项目(20KJD480003); 江苏省双创计划(JSSCBS(2020)31035); 江苏省自然科学基金(BK20201069); 国家自然科学基金项目(91547208, 51909010)

作者简介: 孙娜(1992-), 女, 讲师、硕导, 研究方向为水文水资源、人工智能应用, E-mail: sunna1347@126.com

通讯作者: 张楠(1991-), 男, 讲师、硕导, 研究方向为水电、风电等清洁能源优化运行与控制、人工智能应用, E-mail: zhangnanhust@163.com

宽,求解带约束的变分问题,其目标函数为:

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \|\partial_t [\delta(t) + j/(\pi t)] \cdot u_k(t)\} e^{-j\omega_k t} \right\|_2^2 \\ \text{s. t. } \sum_k u_k = Q(t) \end{cases} \quad (1)$$

步骤 4 引入二次惩罚因子 α 及拉格朗日乘子 λ , 将约束变分问题(式(1))转化为无约束问题,即:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \|\partial_t [\delta(t) + j/(\pi t)] \cdot u_k(t)\} e^{-j\omega_k t} \right\|_2^2 + \langle \lambda, Q(t) - \sum_k u_k \rangle \quad (2)$$

其中, α 能有效降低高斯噪声的干扰, λ 能增强约束的严格性。

步骤 5 使用 ADMM 求解式(2), 交替更新 u_k^{n+1} 、 ω_k^{n+1} 、 λ^{n+1} 以搜索函数式(2)的“鞍点”, 更新公式为:

$$\begin{cases} \hat{u}_k^{n+1} = \left[\hat{Q}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right] / [1 + 2\alpha(\omega - \omega_k)^2] \\ \omega_k^{n+1} = \int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega / \int_0^\infty |\hat{u}_k(\omega)|^2 d\omega \\ \hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau(\hat{Q}(\omega) - \sum_{i \neq k} \hat{u}_i^{n+1}(\omega)) \end{cases} \quad (3)$$

式中, $\hat{Q}(\omega)$ 、 $\hat{u}_i(\omega)$ 、 $\hat{\lambda}(\omega)$ 分别为 $Q(t)$ 、 $u_i(t)$ 、 $\lambda(t)$ 的傅里叶变换; ω 为角频率。

步骤 6 停止判断。满足下式时停止迭代:

$$\sum_{k=1}^d \|u_k^{n+1} - u_k^n\|_2^2 / \|u_k^n\|_2^2 < \epsilon \quad (4)$$

研究表明^[5], VMD 应用于非线性非平稳序列预报时, 预报性能受模态分量个数 d 影响。为此, 定义了分解性能因子 (Q_{QFD}) 以自动确定 VMD 的最佳模态分解数 d 实现自适应 VMD (AVMD)。 Q_{QFD} 求解如下: 首先, 计算标准 VMD 分解得到的 d 个模态分量的样本熵, 记为 $\{S_i\} = \{S_1, S_2, \dots, S_d\}$; 基于此, 计算样本熵的变差系数 C_V 刻画各模态分量的信息差异。此外, 计算 $Q(t)$ 与重构 $\hat{Q}(t)$ 的总体误差 e_{all} , 则 Q_{QFD} 计算式为:

$$Q_{\text{QFD}} = C_V(S_i) / e_{\text{all}} = \frac{C_V(S_i)}{\sum_{i=1}^N |Q(t) - \hat{Q}(t)|} = \frac{C_V(S_i)}{\sum_{i=1}^N |Q(t) - (u_1 + \dots + u_d)|} \quad (5)$$

Q_{QFD} 值越大, 分解效果越好。

2.2 高斯过程回归

2.2.1 基本原理

GPR 兼具贝叶斯的灵活归纳推理能力和机器学习的非线性拟合能力, 在处理非线性、高维度和少样本回归问题中具有明显优势。基于机器学习的中长期径流预报可表述为:

$$y = f(X) + \epsilon, \epsilon \sim N(0, \sigma_n^2 \mathbf{I}_n) \quad (6)$$

式中, y 为受噪声污染的观测径流; X 为与径流相关的输入因子, 如区域降水、气温和/或遥相关因子; f 为隐式函数, 表示 X 与待预测径流 y 之间的映射关系; ϵ 为相互独立且服从高斯分布的观测噪声; σ_n^2 为噪声方差; \mathbf{I}_n 为单位阵。

高斯过程为一系列随机变量的集合, 可由均值函数 $m(X)$ 和协方差函数 $k(X, X)$ 唯一确定, 定义为:

$$\begin{cases} f(X) \sim \text{GP}(m(X), k(X, X)) \\ m(X) = E[f(X)] \\ k(X, X) = E[(f(X) - m(X)) \cdot (f(X) - m(X))^T] \end{cases} \quad (7)$$

式中, GP 为高斯过程; $m(X)$ 为均值函数。

为简化推导, $m(X)$ 经预处理使之为零, 则 y 的先验分布为:

$$y \sim N(0, \mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n) \quad (8)$$

式中, $\mathbf{K}(X, X)$ 为训练核矩阵。

考虑噪声加入, 根据高斯过程的性质可知, 训练输入样本目标输出 y 与测试样本的输出 y_{test} 服从高斯联合分布:

$$\begin{bmatrix} y \\ y_{\text{test}} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n & \mathbf{K}(X, x_{\text{test}}) \\ \mathbf{K}(x_{\text{test}}, X) & k(x_{\text{test}}, x_{\text{test}}) \end{bmatrix}\right) \quad (9)$$

其中 $\mathbf{K}(x_{\text{test}}, X) = \mathbf{K}(X, x_{\text{test}})^T$ 式中, $\mathbf{K}(x_{\text{test}}, X)$ 为测试输入集 x_* 与训练输入变量 X 的 $n \times 1$ 阶协方差函数矩阵; $k(x_{\text{test}}, x_{\text{test}})$ 为测试输入变量 x_{test} 自身的协方差。

X 、 y 组成的观测数据集 D 给定条件下, 新输入 x_{test} , 根据 Bayesian 后验概率可推断预测值 y_{test} 的后验分布为:

$$p(y_{\text{test}} | X, y, x_{\text{test}}) \sim N(E(y_{\text{test}}), \mathbf{K}(y_{\text{test}}, y_{\text{test}})) \quad (10)$$

y_{test} 的均值 $E(y_{\text{test}})$ 、后验方差 $\text{Cov}(y_{\text{test}})$ 分别为:

$$E(y_{\text{test}}) = \mathbf{K}(x_{\text{test}}, X) [\mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n]^{-1} y \quad (11)$$

$$\text{Cov}(y_{\text{test}}) = k(x_{\text{test}}, x_{\text{test}}) -$$

$$\mathbf{K}(x_{\text{test}}, X) (\mathbf{K}(X, X) + \sigma_n^2 \mathbf{I}_n)^{-1} \mathbf{K}(X, x_{\text{test}}) \quad (12)$$

2.2.2 组合核函数及超参数求解

GPR 的 $\mathbf{K}(X, X)$ 可衡量实测值与预测值的拟合程度, 由式(10)可知, $\mathbf{K}(X, X)$ 对 GPR 预测

性能起决定性作用。常用的核函数 $\mathbf{K}(X, X)$ 有平方指数核函数(SE)、Matern 核函数(Ma)、有理平方核函数(RQ)、线性核函数(LIN)和随机噪声核函数(RN)^[6]。考虑到径流影响因素较多且过程复杂,单一核函数只能学习样本某一方面的特征,难以表征复杂径流过程的未来变化规律,立足径流过程自身蕴含的周期性、趋势性及随机性,本文将 Ma、LIN、RN 叠加构造了一种新的组合核函数(CK),以增强核函数的非线性表达能力及细节信息捕捉能力。定义为:

$$k_{\text{new}} = k_{\text{Ma}} + k_{\text{LIN}} + k_{\text{RN}} = \sigma_{\text{IMa}}^2 \left(1 + \frac{\sqrt{5} \|x - x'\|}{l} + \frac{5 \|x - x'\|^2}{3l^2} \right) \cdot \exp\left(-\frac{\sqrt{5} \|x - x'\|}{l}\right) + \sigma_{\text{ILIN}}^2 (X - c)(X' - c) + \sigma_{\text{IRN}}^2 \delta_{ij} \quad (13)$$

式中, σ_{IMa}^2 、 l 均为 Ma 核函数的超参数,分别为函数与其平均值的平均偏差、定长参数; x 、 x' 均为样本点; σ_{ILIN}^2 为 LIN 核函数的超参数,表示函数与其平均值的平均偏差; c 决定了后验中所有线经过的点的 x 坐标; σ_{IRN}^2 为 RN 核函数的超参数,表示噪声水平; δ_{ij} 为克罗内克 δ 函数。

超参数 $\theta = \{\sigma_{\text{IMa}}^2, l, \sigma_{\text{ILIN}}^2, \sigma_{\text{IRN}}^2\}$ 为组合核函数中未知参数的集合,通过对训练样本对数似然函数 $L(\theta) = -\log p(y | X, \theta)$ 求极值,利用共轭梯度法(CG)即可获得最优 θ 。 $L(\theta)$ 的数学表达式为:

$$L(\theta) = \frac{1}{2} y^T C^{-1} y + \frac{1}{2} \log |C| + \frac{n}{2} \log 2\pi \quad (14)$$

式中, C 为正定矩阵。

$L(\theta)$ 关于 θ 的偏导数表达式为:

$$\frac{\partial L(\theta)}{\partial \theta_i} = \frac{1}{2} \text{tr} \left(((C^{-1} y)(C^{-1} y)^T - C^{-1}) \frac{\partial C}{\partial \theta_i} \right)$$

$$C = K_n + \sigma_n^2 I_n \quad (15)$$

获得最优超参数 θ 后,利用式(11)、(12)即可获得样本 x_{test} 预测均值 $\hat{y}_* = E(y_{\text{test}})$ 及其均方误差 $\sigma = \text{Cov}(y_{\text{test}})$ 。由 3σ 法则可知,预测值在置信水平 $1 - \alpha$ 的区间^[7]为:

$$\left(\hat{y}_* - \frac{\sigma z_{\alpha/2}}{\sqrt{n}}, \hat{y}_* + \frac{\sigma z_{\alpha/2}}{\sqrt{n}} \right) \quad (16)$$

式中, σ 为方差; σ/\sqrt{n} 为标准差; $z_{\alpha/2}$ 为标准正态分布的分位数。

2.3 其他机器学习模型

为了说明 GPR-CK 模型进行月径流预报的优势,选择四种广泛应用的单一机器学习作为对比模型,分别为 BP、RBF、GRNN、RELM 神经网络。四种模型与 GPR-CK 模型的输入一致,最佳

输入因子采用偏自相关系数法(PACF)确定。

2.4 AVMD-GPR-CK 模型预测流程

步骤 1 搜集相关数据,基于分解性能因子确定 AVMD 的最佳模态分解个数 d 。

步骤 2 采用 AVMD 分解原始降雨径流数据得到 d 个相对平稳的子序列。

步骤 3 采用 PACF 确定各子序列的输入输出向量,并归一化,按 3 : 1 划为训练和检验样本。

步骤 4 训练样本用于构建 GPR-CK 模型,采用 CG 法优化组合核函数的超参数。

步骤 5 检验样本驱动优化后的 GPR-CK 模型,得到各子序列的预报值。

步骤 6 将各预报值反归一化,叠加重构后得到最终预测结果,使用常用的均方根误差 R_{RMSE} 、平均绝对误差 M_{MAE} 、平均绝对百分比误差 M_{MAPE} 、确定性系数 N_{NSE} 评价模型的预测性能,与常用模型对比验证 AVMD-GPR-CK 的预报效果。

评价指标定义式见文献[1],预报流程见图 1。

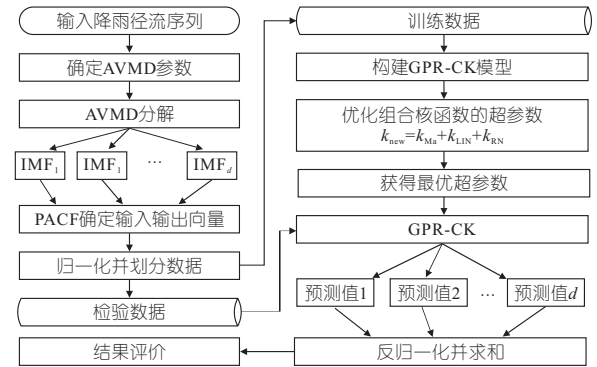


图 1 AVMD-GPR-CK 预测模型流程

Fig. 1 Flowchart of AVMD-GPR-CK

3 实例应用

3.1 研究区域和数据

以金沙江流域向家坝水文站为例验证模型性能。研究数据包括 1961 年 1 月~2008 年 12 月实测径流及同期面平均降水,面降水采用泰森多边形法计算,站点信息见文献[8]。逐月径流数据来源于长江水利委员会水文局,降雨源于中国气象数据网(<https://data.cma.cn/>)。将相关面降雨和前期径流作为主要输入变量,建立月径流预报模型,预测未来 1~12 个月径流,其中前 36 年为率定期,其余为检验期。

3.2 模型参数设置

AVMD-GPR-CK 的超参数初始值为 $l = 0.9$ 、 $\sigma_{\text{IMa}}^2 = 4$ 、 $\sigma_{\text{ILIN}}^2 = \text{rand}(1)$ 、 $\sigma_{\text{IRN}}^2 = 0.01$;模态分解数 d 的变化范围为 3~20,径流分解序列见图 2。对比模型的参数设置见文献[6,8]。

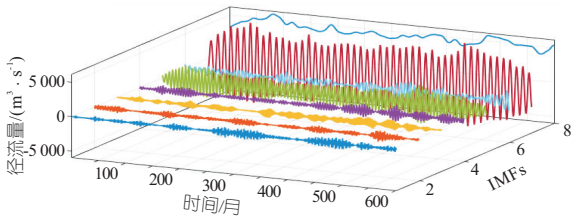


图 2 AVMD 分解 3D 图

Fig. 2 Decomposition results of AVMD

3.3 结果与分析

3.3.1 核函数的影响

图 3 为基于不同核函数 GPR 的预报结果。由图 3 可知,基于不同类型核函数 GPR 的模拟效果存在差异。单核 GPR 模型中,GPR-Ma 预报效果最佳,GPR-RQ 较差,GPR-SE 居中。GPR-CK 的预报效果优于三种单核 GPR 模型。如 GPR-CK 模型所有预见期 N_{NSE} 值均大于 0.75,其中预见期 L 为 1、10、11、12 个月时, N_{NSE} 值超过 0.8,表明预报效果良好;预见期越长,预报效果反而越好,出现这一现象的原因为 $L=10\sim 12$ 个月时,正值枯水期,枯水期径流过程不像汛期径流变化剧烈,相对平稳,与局部气象因子相关性更强,更利于模型捕捉其演变特征。总之,对于向家坝水文站 GPR-CK 模型在预见期 L 不超过 10 个月时均能达到不错的预报效果, L 为 10~12 个月时,GPR-CK 的 M_{MAPE} 指标略差于 GPR-Ma。

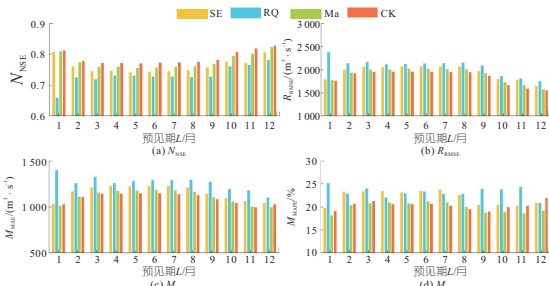


图 3 基于不同核函数的高斯过程回归模型预报结果

Fig. 3 $N_{NSE}, R_{RMSE}, M_{MAE}, M_{MAPE}$ of GPR with different kernel functions

3.3.2 GPR-CK 与常用机器学习预报结果对比

图 4 为 $L=1$ 时 GPR-CK 和其对比较模型检验期预报结果。由图 4 可看出,GPR-CK 的 N_{NSE} 值最大, $R_{RMSE}, M_{MAE}, M_{MAPE}$ 值最小,分别为 $0.813, 1\ 763\ m^3/s, 1\ 030\ m^3/s, 19\%$ 。证明了 GPR-CK 适合作为金沙江流域径流预报模型。

3.3.3 AVMD-GPR-CK 与 GPR-CK 预报结果对比

图 5 为 AVMD-GPR-CK 和 GPR-CK 不同预见期检验期预报结果。由图 5 可看出:① AVMD-GPR-CK 模型所有预见期的 N_{NSE} 值均大于 0.94,显著高于 GPR-CK 最优预报值,即预见期 $L=1$ 个月时 GPR-CK 的 N_{NSE} 值最优为 0.813。

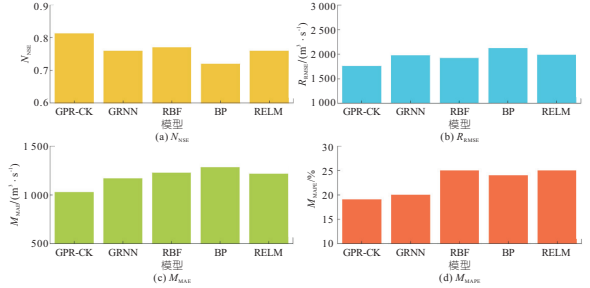


图 4 GPR-CK 与常见机器学习模型预报结果对比

Fig. 4 Comparison of $N_{NSE}, R_{RMSE}, M_{MAE}, M_{MAPE}$ of GPR-CK and commonly used ANNs

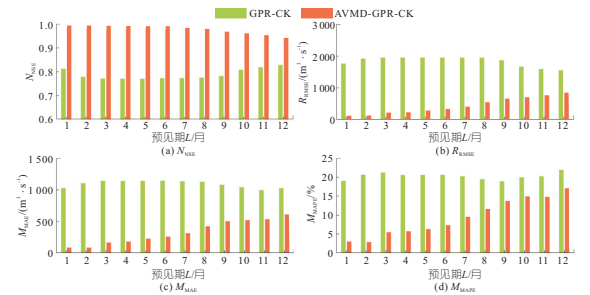


图 5 AVMD-GPR-CK 和 GPR-CK 模型预报结果

Fig. 5 Forecast results of AVMD-GPR-CK and GPR-CK

②随着预见期的增加,AVMD-GPR-CK 的误差逐渐增大,当 $L=12$ 个月时, R_{RMSE} 值达到最大 $849\ m^3/s$;对于 GPR-CK,其 R_{RMSE} 值随预见期的增大呈先增大后减小的趋势,可能的原因为当 L 为 10~12 个月时,预报月份处于枯水期,径流的波动性相比汛期的波动小,便于模型捕捉径流变化规律;GPR-CK 模型 R_{RMSE} 最小值为 $1\ 555\ m^3/s$,此时 $L=12$ 个月,显然大于 AVMD-GPR-CK 模型 R_{RMSE} 的最大值。③当 $L < 8$ 个月时,AVMD-GPR-CK 的 M_{MAPE} 值在 $\pm 10\%$ 内;当 $L < 11$ 个月时, M_{MAPE} 值小于 15% ;当 $L=12$ 个月时, M_{MAPE} 值达到最大 $\pm 17\%$;GPR-CK 模型 M_{MAPE} 值在 $19\% \sim 22\%$ 左右,仍高于 AVMD-GPR-CK 模型 M_{MAPE} 最大值。

为了直观表示模型的预测效果,分别绘制了预报效果最好 ($L=1$)、最差月份 ($L=12$) 的径流过程,见图 6。由图 6 可看出,当 $L=1$ 个月时,预报径流过程能精确捕捉观测径流的涨落,几乎完全重合,不确定区间很窄,说明预报置信水平高、效果好;当 $L=12$ 个月时,径流峰值和低流量不确定区间比中流量区间大,但均在 90% 置信区间内。可见,预报结果不确定区间基本能覆盖径流波动区域,且随着预见期的增加,不确定性增加但也在可接受范围内。总之,对比结果均可证明所提 AVMD-GPR-CK 的预报优势。

3.3.4 AVMD 与 CK 贡献分析

为量化分析 AVMD 和 CK 对耦合模型的贡

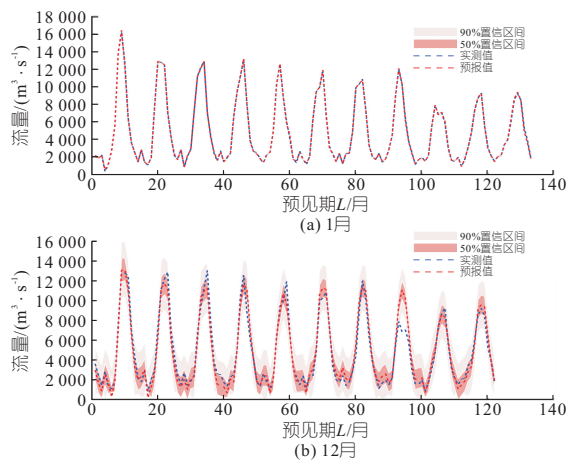


图 6 AVMD-GPR-CK 实测与预报径流过程对比

Fig. 6 Comparison of observed runoff and forecasted runoff produced by AVMD-GPR-CK

献,将 GPR-SE、GPR-Ma 和 GPR-RQ 作为基准模型,分析 AVMD-GPR-CK 和 GPR-CK 相对于这三种模型 N_{NSE} 的提升率。 N_{NSE} 提升率计算式为 $P_{NSE} = [(N_{NSE} - N_{NSE_{base}}) / N_{NSE_{base}}] \times 100\%$ ($N_{NSE_{base}}$ 为基准模型的 N_{NSE} 值)。由计算结果可知,GPR-CK 相对于 GPR-RQ、GPR-SE、GPR-Ma 的 P_{NSE} 均值分别为 3.40%、7.90%、1.49%,相对于 GPR-Ma 的 P_{NSE} 并不明显,可能原因是组合核函数超参数增多,增加了迭代寻优的难度。AVMD-GPR-CK 相对于 GPR-SE、GPR-RQ、GPR-Ma、GPR-CK 的 P_{NSE} 分别为 28.72%、34.32%、26.36%、24.50%。因此,AVMD 对提升 AVMD-GPR-CK 模型的 N_{NSE} 值贡献率为 13.8%~29.17%,CK 对耦合模型的贡献率为 0.34%~2.13%;且采用 AVMD 分解更有助于组合核函数发挥作用。

4 结论

a. 核函数对 GPR 预报性能影响明显,基于物理信息的组合核函数能提高 GPR 预报性能,

Monthly Runoff Prediction Model and Its Application Based on GPR with Physically Composite Kernel

SUN Na^{1a}, ZHANG Nan^{1b}, ZHANG Shuai², PENG Tian^{1a}, ZHOU Jian-zhong³, ZHANG Hai-rong⁴
 (1a. Faculty of Automation; 1b. Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology, Huaian 223003, China; 2. Xi'an Shufeng Technological Information Ltd., Xi'an 710054, China;
 3. School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China;
 4. China Yangtze Power Co., Ltd., Yichang 443133, China)

Abstract: In light of the difficulty of traditional single runoff prediction models to describe future variation in runoff, a monthly runoff prediction model named AVMD-GPR-CK based on adaptive variational modal decomposition (AVMD) and Gaussian process regression (GPR-CK) with physically composite kernel was proposed. In the proposed model, the runoff series was decomposed into several subseries using AVMD. Then subseries were separately modeled according to their own characteristics, and the final prediction result was the superposition of the subsequence prediction results. The AVMD-GPR-CK was applied to forecast the future 1-12 months runoff at Xiangjiaba station in the Jinsha River basin. The results show that the deterministic coefficient of the AVMD-GPR-CK model is greater than 0.94, and the mean absolute percentage error (M_{MAPE}) is within $\pm 17\%$ for all leading times, and the M_{MAPE} is inside $\pm 10\%$ for leading times within 10 months. Furthermore, the accuracy of the AVMD-GPR-CK is significantly better than those of the commonly used BP, GRNN, RBF, and RELM models.

Key words: monthly runoff forecasting; variational mode decomposition; Gaussian process regression; composite kernel function

GPR-CK 优于常见神经网络模型。

b. AVMD 能充分挖掘有助于提高径流峰值预报精度的局部信息,贡献率为 13.8%~29.17%,显著大于 CK 的贡献率,说明数据预处理比模型参数对预报结果影响更大。

c. AVMD-GPR-CK 的 M_{MAPE} 均在 $\pm 17\%$ 以内,说明该模型有效提升了预报精度,并延长了预见期,给缺资料地区径流预报提供了一种新方法。

参考文献:

- [1] SUN N, ZHANG S, PENG T, et al. Multi-variables-driven model based on random forest and gaussian process regression for monthly streamflow forecasting[J]. Water, 2022, 14(11):1828.
- [2] 王忠义, 崔东文. 基于小波包分解-非洲秃鹫优化算法-深度极限学习机的水文预报模型及其应用[J]. 水电能源科学, 2022, 40(8):26-31.
- [3] 蔡文静, 陈伏龙, 何朝飞, 等. 基于时频分析的 LSTM 组合模型径流预测[J]. 干旱区地理, 2021, 44(6):1696-1706.
- [4] 周靖楠, 刘振男, 陆之洋, 等. 基于 MI-KPCA 与高斯回归过程的北汝河中长期径流预测[J]. 水电能源科学, 2021, 39(5):42-45.
- [5] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. IEEE transactions on signal processing, 2014, 62(3):531-544.
- [6] SUN N, ZHANG S, PENG T, et al. A composite uncertainty forecasting model for unstable time series: application of wind speed and streamflow forecasting[J]. IEEE access, 2020, 8: 209251-209266.
- [7] 彭虹桥, 顾洁, 胡玉, 等. 基于混沌粒子群-高斯过程回归的饱和负荷概率预测模型[J]. 电力系统自动化, 2017, 41(21): 25-32, 155.
- [8] 孙娜, 周建中, 朱双, 等. 基于小波分析的两种神经网络耦合模型在月径流预测中的应用[J]. 水电能源科学, 2018, 36(4):14-17, 32.