

DOI:10.11918/202012021

采用核极限学习机的短期需水量预测模型

韩宏泉, 吴 珊, 侯本伟

(北京工业大学 城市建设学部, 北京 100124)

摘 要: 为满足给水系统日常管理对短期需水量预测时效的需求,建立了所需训练时间短的核极限学习机模型(kernel-based extreme learning machine, KELM);从提升预测精度的角度考虑,构造了以傅里叶级数为理论依据的残差修正模块(Fourier series, FS),利用该模块对需水量初始预测值与观测值之间的差值进行建模,完成对初始预测值的残差修正,将该模块叠加于KELM模型上形成了组合预测模型(KELM + FS)。通过实际数据对模型进行性能测试,结果表明:KELM模型能够与人工神经网络模型、支持向量回归模型产生相似的预测精度,但预测时间仅为二者平均值的5%左右;组合模型KELM + FS在未显著增加预测时间的前提下,比KELM模型相对预测精度提升了12%左右。在用于短期需水量预测时,无论单一模型KELM还是组合模型KELM + FS都能达到有效提升预测效率的目的。

关键词: 短期需水量预测;核极限学习机;组合模型;傅里叶级数;残差修正

中图分类号: TU991

文献标志码: A

文章编号: 0367-6234(2022)02-0017-08

Short-term water demand prediction model using kernel-based extreme learning machine

HAN Hongquan, WU Shan, HOU Benwei

(Faculty of Architecture, Civil and Transportation Engineering, Beijing University of Technology, Beijing 100124, China)

Abstract: To meet the requirements of the daily management of water supply systems for short-term water demand prediction timeliness, a kernel-based extreme learning machine model (KELM) was established, which requires short training time. From the perspective of improving the prediction accuracy, a residual correction module based on the Fourier series (FS) was constructed, which was used to model the difference between the initial predicted value and the observed value of water demand, and the residual correction of the initial predicted value was completed. The module was superimposed on the KELM model to form the hybrid prediction model (KELM + FS). The performance of the models was tested using real water demand data. Experimental results show that the KELM model could produce similar prediction accuracy as the artificial neural network model and the support vector regression model, but the prediction time was only about 5% of the average time of the two models. Compared with the KELM model, the relative prediction accuracy of the hybrid model KELM + FS was improved by about 12% without significantly increasing the prediction time. Therefore, when applied to short-term water demand prediction, both the single model KELM and the hybrid model KELM + FS could achieve the goal of improving the prediction efficiency.

Keywords: short-term water demand prediction; kernel-based extreme learning machine; hybrid model; Fourier series; residual correction

短期需水量预测模型在有限的预测区间(从1个月到1天不等)内进行需水量预测,预测时间步长通常控制在每天到每小时,其在给水系统的优化调度与日常管理中发挥着重要作用^[1-2]。以往的短期需水量预测模型一般面向整个城市^[3-6],预测步长多以小时和天为时间步长,主要应用于城市水量的供需平衡分析。近年来,随着硬件采集技术的发

展,收集步长为1~15 min的观测数据成为可能,物联网、云计算等新技术也不断融入水务行业,期望从供水系统大数据的研究中挖掘出更多工程应用价值。在此背景下,在线调度、压力分区调控、漏损控制、爆管监测等供水系统工程对短期需水量预测提出了新的速度和时间尺度需求。

很多学者对需水量预测方法进行了研究,早期传统的统计学方法因易于理解和实施得到了大量的应用,然而,由于自身的局限性很多时候并不能达到良好的预测效果^[2],例如,多元线性回归算法预设了特征与需水量之间的关系;指数平滑算法对其加

收稿日期: 2020-12-08

基金项目: 国家水体污染控制与治理科技重大专项(2017ZX07108-002);国家自然科学基金(51978023)

作者简介: 韩宏泉(1995—),女,硕士研究生

通信作者: 侯本伟, benweihou@bjut.edu.cn

权系数缺乏客观的识别方法;自回归移动平均算法要求时序数列是平稳的或者经差分处理后是平稳的,且本质上只能捕捉线性关系,而真实需水量的变化是非线性的,且具有随机性。近年来,机器学习方法被更多地应用于短期需水量预测中,机器学习方法是典型的数据驱动算法,通过历史数据来深入挖掘需水量与相关变量之间的关系,从而提高预测模型的稳定性和可靠性。其中,人工神经网络算法(artificial neural network, ANN)和支持向量回归算法(support vector regression, SVR)以及它们的扩展方法在文献中介绍最多^[3-5],研究中也多使用二者与新提出的机器学习模型进行性能比较,从而对新模型进行评估。

除 ANN 与 SVR 外,极限学习机算法(extreme learning machine, ELM)由于比传统机器学习方法拥有更快的训练速度,近来也被应用于预测模型中,但其实际应用效果表现出一定的差异性^[6-8]。Huang 等^[9]在 ELM 的基础上提出了核极限学习机(kernel-based extreme learning machine, KELM)算法,该算法继承了 ELM 较高的训练速度的优点的同时,被证明相比 ELM 具有更加稳定的性能并且有着与 SVR 相似的泛化能力^[10],但该方法目前尚未在短期需水量预测领域得到更多尝试。

研究人员除了不断追求单独预测模型效率的提高外,近来更多地倾向于组合模型的研究^[11-12],通过对两种或多种预测模型进行组合以形成优势互补,来提高整体预测效果。其中,傅里叶级数模型(Fourier series, FS)因实现简单且需要较少的训练数据而常用于与机器学习模型进行组合^[13-14]。本研究充分利用数据信息,将短期需水量的预测步长控制在 15 min,引入了 KELM 模型以提高短期需水量预测的速度,同时,提出了以 KELM 模型为初始预测模块叠加使用 FS 模块进行残差修正的组合模型(KELM + FS),以进一步提升预测精度。

1 方法原理

本研究提出的基于核极限学习机的短期需水量预测模型 KELM 通过对样本数据进行训练,建立起历史时刻需水量与目标时刻需水量之间的非线性关系,进而根据历史数据实现对目标时刻需水量的预测。模型为多维输入单维输出,输入特征为与目标时刻需水量数据高度相关的历史需水量数据,输出为目标时刻的需水量预测值。KELM + FS 模型中的残差修正模块 FS 则利用傅里叶级数对 KELM 的预测值与观测值之间的差值进行建模,识别出残差序列中隐含的周期性规律,从而完成对未来时刻残差

的推测,其输入为 T 个历史残差数据,输出为整个预测区间的残差序列。

1.1 极限学习机原理

ELM 是在单隐含层前馈神经网络的基础上提出的。与传统单隐含层前馈神经网络使用梯度下降法迭代求解网络参数不同,ELM 在训练过程中随机生成输入权值 ω 和偏置 b 且不再调整,将最优求解问题转化为简单的最小二乘求解问题,这使得 ELM 具有了更快的学习和运行速度^[15]。应用中 ELM 的训练机理可以简单描述如下^[6,15-16]:

假设给定训练样本 $\{x_i, y_i | x_i \in \mathbf{R}^D, i = 1, \dots, n\}$ (其中 x_i 表示第 i 个样本的 D 维输入向量 $x_i = [x_{i1}, \dots, x_{iD}]$, y_i 对应着第 i 个样本的观测数据),隐含层神经元数为 k ,根据 ELM 网络定义可以写出如下表达式:

$$\sum_{j=1}^k g(\omega_j, b_j, x_i) \beta_j = \hat{y}_i, \quad i = 1, 2, \dots, n, j = 1, 2, \dots, k \quad (1)$$

式中: $g(\cdot)$ 为显式激活函数; ω_j 为连接隐含层第 j 个神经元与输入层神经元的权向量, $\omega_j = [\omega_{j1}, \dots, \omega_{jD}]^T$; b_j 为隐含层第 j 个神经元的偏置,偏置向量 $b = [b_1, \dots, b_k]^T$; β_j 为隐含层第 j 个神经元与输出层神经元的连接权值,权值向量 $\beta = [\beta_1, \dots, \beta_k]$; $\hat{y}_i$ 为第 i 个样本的预测输出结果。

当 ELM 的输出结果能够以零误差逼近观测数据时,利用矩阵表达可以写为如下形式:

$$H\beta = Y \quad (2)$$

式中: H 为隐含层输出矩阵, $Y = [y_1, \dots, y_n]^T$ 为观测向量。

在 ELM 训练过程中, ω 和 b 采取随机生成的方法确定,为计算网络的输出结果, β 是唯一需要求解的参数^[17]。

以最小训练误差和最小输出权的范数为求解目标,根据线性代数知识,方程(2)的最小范数最小二乘解为^[18]

$$\beta = H^+ Y = (H^T H)^{-1} H^T Y \quad (3)$$

其中 H^+ 为隐含层输出矩阵 H 的 Moore-Penrose 广义逆矩阵。

1.2 核极限学习机原理

为了进一步简化 ELM 的应用,避免在确定隐含层神经元数时应用耗时的算法和随机分配权重的潜在缺点, Huang 等^[18]在 ELM 的基础上提出了 KELM, KELM 隐含层的映射不再采用 $g(\cdot)$ 这种简单的显式形式而是采用核矩阵确定,而且隐含层神经元数不再需要指定,在这种情况下 KELM 比 ELM 具有更稳定的性能,并且 KELM 与 SVR 具有相似的

泛化能力^[10]。KELM 的机制在 ELM 的基础上可以描述如下^[9,18-19]。

为增加模型的稳定性和泛化能力,在对 β 的求解中将引入一个正数 C ,并以最小训练误差和最小输出权的范数为求解目标^[20],则有

$$\min \frac{1}{2} \|\beta\|^2 + \frac{1}{2} C \sum_{i=1}^n \|e_i\|^2$$

$$\text{s. t. } \mathbf{h}(\mathbf{x}_i)\beta = \mathbf{y}_i^T - \mathbf{e}_i^T, i = 1, \dots, n \quad (4)$$

式中: e_i 为训练误差, $\mathbf{h}(\mathbf{x})$ 为隐含层特征映射函数。

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(\mathbf{x}_1) \\ \vdots \\ \mathbf{h}(\mathbf{x}_n) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \cdots & h_k(x_1) \\ \vdots & & \vdots \\ h_1(x_n) & \cdots & h_k(x_n) \end{bmatrix}_{n \times k} \quad (5)$$

此时 β 被计算为

$$\beta = \mathbf{H}^T (\mathbf{I}/C + \mathbf{H}\mathbf{H}^T)^{-1} \mathbf{Y} \quad (6)$$

对于 KELM 来说不必知道其映射函数的具体形式,而是构造隐式映射来代替,即构造核矩阵来代替 $\mathbf{H}^T \mathbf{H}$,则有

$$\Omega_{\text{ELM}} = \mathbf{H}\mathbf{H}^T$$

$$\Omega_{\text{ELM}_{i,j}} = \mathbf{h}(\mathbf{x}_i)\mathbf{h}(\mathbf{x}_j) = K(x_i, x_j) \quad (7)$$

最终网络输出结果可写为

$$y = f(x) = \mathbf{h}(\mathbf{x})\mathbf{H}^T (\mathbf{I}/C + \mathbf{H}\mathbf{H}^T)^{-1} \mathbf{Y} =$$

$$\begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_n) \end{bmatrix}^T (\mathbf{I}/C + \Omega_{\text{ELM}})^{-1} \mathbf{Y} \quad (8)$$

式中: Ω_{ELM} 为核矩阵, $K(x_i, x_j)$ 为核函数。在 KELM 中有多种核函数可供选择,比如线性核函数(linear)、多项式核函数(polynomial)、高斯径向基核函数(rbf)等^[19]。图 1 中给出了 ELM 与 KELM 多维输入单维输出的经典结构,其中 a 为 ELM 的映射函数, b 为 KELM 的映射函数。

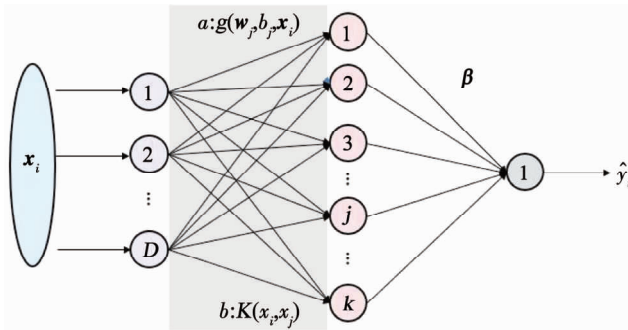


图 1 多维输入单维输出的 ELM 和 KELM 经典结构

Fig. 1 Typical structures of ELM and KELM with multidimensional input and one-dimensional output

1.3 傅里叶级数原理

傅里叶级数能够将周期信号或函数展开为一组不同频率三角函数的和,通过将周期内的历史时间序列数据拟合为傅里叶级数方程来实现对未来时刻

数据的推断。假设一个周期由 T 个时刻组成,将此周期在 $[0, 2\pi]$ 区间内进行标准化,则有

$$t_q = \frac{2\pi q}{T}, q = 1, \dots, T \quad (9)$$

此时傅里叶级数方程可以定义为

$$f^*(t_q) = a_0 + \sum_{p=1}^M [a_p \cos(pt_q) + b_p \sin(pt_q)] \quad (10)$$

式中: t_q 为经过标准化的时刻; M 为傅里叶多项式的级数, $M < T/2$; a_0 、 a_p 和 b_p 为可以调整的傅里叶级数方程的系数。

以观测值 $f(t_q)$ 与傅里叶级数方程估计值 $f^*(t_q)$ 之间差值的二范数作为目标函数,傅里叶级数方程的系数可以通过最小二乘法确定^[13]:

$$\begin{cases} a_0 = \frac{\sum_{q=0}^{T-1} f(t_q)}{T} \\ a_p = \frac{2 \sum_{q=0}^{T-1} f(t_q) \cos(pt_q)}{T} \\ b_p = \frac{2 \sum_{q=0}^{T-1} f(t_q) \sin(pt_q)}{T} \end{cases} \quad (11)$$

本研究中,首先选定残差周期,将一个残差周期内的历史时刻进行标准化处理,然后根据式(11)的系数表达式,将周期内的残差序列值 $f(t_q)$ 以及标准化时刻 t_q 依次代入式中,便可计算出相应的系数 a_0 、 a_p 和 b_p ,从而实现傅里叶级数方程的拟合,继而根据残差序列所蕴含的周期性特征,可使用拟合好的傅里叶级数方程推测未来时刻的残差值。

2 建模流程

本研究所提出的短期需水量预测模型是在预测区间为 24 h,预测步长为 15 min 的条件下建立的。根据监测条件,观测数据每 15 min 进行一次采样,每 24 h 进行一次传输,即历史数据集每 24 h 进行一次更新。历史数据集的大小始终保持不变,采用移动窗口更新,即当增添最近 96 个时刻的观测数据的同时会删除最早的 96 个时刻的数据。

2.1 输入特征选择

选择合适的输入特征对模型的建立十分重要。Guo^[2]、Bakker^[21]、Cutore 等^[22]在建立短期需水量预测模型时都仅选择历史需水量作为唯一输入特征,并且证明了在这一条件下可以实现可靠的预测。本研究同样使用历史需水量数据作为模型的唯一输入特征。

综合参考 Guo^[2]、Herrera^[3] 以及 Odan 等^[13] 的输入特征选择方案,并通过相关性分析和模型预测实验,最终选择 {当前时刻,前 1 d 的目标时刻,前 2 d 的目标时刻,前 1 周的目标时刻} 的用水需求数据作为模型的输入特征,即 $\{Q_t, Q_{t-95}, Q_{t-191}, Q_{t-671}\}$ 。预测期间 Q_{t-95}, Q_{t-191} 和 Q_{t-671} 使用对应的历史观测数据作为输入;在第 1 个目标时刻时, Q_t 使用对应的历史观测数据,从第 2 个目标时刻开始, Q_t 使用上一时刻的初始预测值作为下一时刻的特征输入。

2.2 KELM 模型建立步骤

1) 收集并通过数据清洗对历史需水量数据进行预处理。

2) 提取历史需水量数据的输入特征输入到预测模型中,通过 KELM 训练,建立历史需水量数据与下一时刻需水量数据之间的联系,从而实现 KELM 模型的建立。

3) 对建立好的模型输入第 s 个目标时刻($t+s$) 的特征,得到需水量预测值 $\hat{Q}_{t+s}$, t 表示当前时刻。

4) 重复步骤 3), 直到获得 96 个时刻的需水量预测值 $\hat{Q}_{t+s}, s=1, 2, \dots, 96$ 。

2.3 KELM + FS 模型建立步骤

KELM + FS 的建立主要包括建立需水量初始预测模块和建立残差修正模块两部分。图 2 展示了模型建立的流程。其中需水量初始预测模块的构建步骤与 KELM 模型相同,不再赘述。主要建模步骤如下:

1) 建立需水量初始预测模块。

2) 通过建模训练的历史需水量数据、初始预测模块对训练数据拟合(预测)结果,获得初始预测模型对建模数据的残差序列 $\mathbf{e} = [e_1, e_2, \dots, e_T]^T$ 。

$$e_i = \hat{Q}_i - Q_i, i=1, \dots, T \quad (12)$$

式中: e_i 为建模训练数据中第 i 时刻对应的模型拟合残差值, Q_i 为 i 时刻的需水量观测值, $\hat{Q}_i$ 为 i 时刻需水量的初始预测值, T 为用于拟合残差修正模块的历史数据(时刻)数量。

3) 以 T 个历史时刻的残差数据作为一个训练窗口拟合傅里叶级数系数,继而预测得到目标时刻对应的初始预测模块的残差值 $\hat{e}_{t+s} (s=1, 2, \dots, 96)$ 。

4) 计算目标时刻的最终需水量预测值:

$$\hat{Q}_{t+s}^* = \hat{Q}_{t+s} - \hat{e}_{t+s}, s=1, 2, \dots, 96 \quad (13)$$

式中: $\hat{Q}_{t+s}^*$ 为目标时刻需水量的最终预测值, $\hat{Q}_{t+s}$ 为目标时刻需水量的初始预测值。

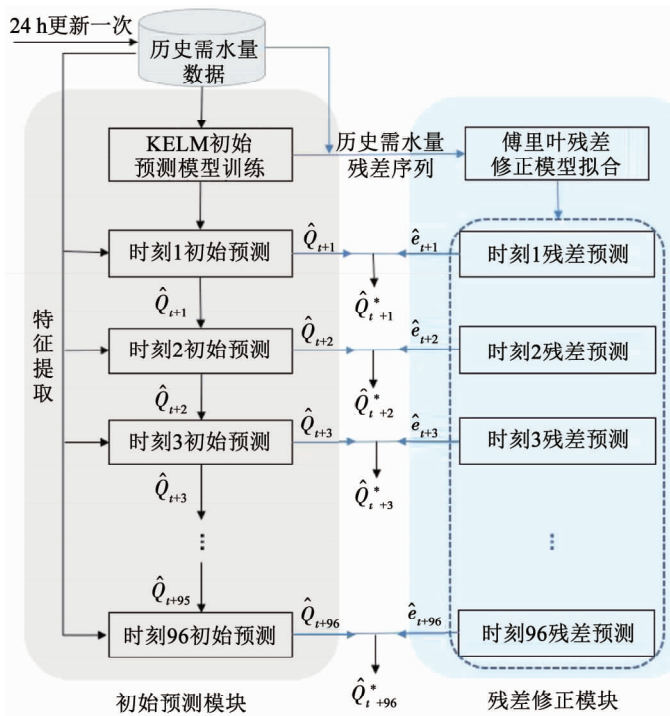


图 2 KELM + FS 模型建立流程

Fig. 2 Flow chart of KELM + FS model building

2.4 评价指标

目前已有许多方法应用于需水量预测模型的性能评价,本研究采用已广为接受的指标作为模型的评价指标,其中包括用于评估预测精度的拟合优度 (R^2)^[2,8,14]、平均绝对百分比误差 (MAPE, E_{MAP})^[2,14,23] 以及用于评价计算速度的模型计算时间。具体计算公式如下:

$$R^2 = 1 - \frac{\sum_{s=1}^{96} (Q_{t+s} - \hat{Q}_{t+s}^*)^2}{\sum_{s=1}^{96} (Q_{t+s} - \bar{Q}_{t+s})^2} \quad (14)$$

$$E_{\text{MAP}} = \frac{1}{96} \sum_{s=1}^{96} \frac{|Q_{t+s} - \hat{Q}_{t+s}^*|}{Q_{t+s}} \times 100\% \quad (15)$$

式中: $\hat{Q}_{t+s}^*$ 与 Q_{t+s} 分别代表第 s 个目标时刻需水量的预测值和实际观测值, $\bar{Q}_{t+s}$ 代表实际观测值的平均值。

3 实例分析

3.1 数据描述

采用某城市 3 个真实的独立计量区 (DMA) 历史需水量数据进行预测实验对提出的模型进行验证。这 3 个 DMA 的组成结构相似,大部分用户为居民用户,包含少量商业用户,属于典型的城市住宅区 DMA。获取数据的时间范围为 2019 年 6 月 1 日—2019 年 7 月 30 日,采样间隔为 15 min。表 1 给出收集数据时间范围内这 3 个 DMA 的基本数据特征,其

中 DMA1 的平均需水量最高,且相比具有更高的随机波动性;DMA2 平均需水量最低;DMA3 需水量数据变异系数最低。研究以 2019 年 7 月 26 日 96 个时刻的需水量数据作为测试数据,将其前 55 d 的数

据通过特征提取后作为模型的建立数据,其中 90% 作为训练集,10% 作为验证集。训练集与验证集用于寻找模型参数并建立模型,测试集则用于测试所建立模型的性能。

表 1 3 个 DMA 的基本需水量数据信息
Tab.1 Basic water demand data for three DMAs

编号	日期	最小需水量/ ($\text{m}^3 \cdot \text{h}^{-1}$)	最大需水量/ ($\text{m}^3 \cdot \text{h}^{-1}$)	平均需水量/ ($\text{m}^3 \cdot \text{h}^{-1}$)	需水量标准差/ ($\text{m}^3 \cdot \text{h}^{-1}$)	变异系数
DMA1	2019-06-01— 2019-07-30	38.00	168.00	97.63	31.60	0.32
DMA2		16.88	67.04	36.49	11.20	0.30
DMA3		36.00	108.00	65.96	15.38	0.23

3.2 模型参数

为了对所提出的模型进行客观评估,引入了已被广泛使用的 ANN 模型、SVR 模型,以及 KELM 的同源模型——ELM 模型和添加了残差修正模块的组合模型一起参与性能测试。所有模型的建立均以 MATLAB 2016b 为平台,所用计算机 CPU 型号为 Inter Core i7-10510U,内存为 16.0 GB。

其中 ANN 模型采用单层前馈架构构建,并使用广为接受的反向传播算法进行训练,其隐含层神经元 k 为该模型主要待定参数,采用 $\log n^{[24]}$ (n 为训练样本数)和 $2D+1^{[25]}$ (D 为输入特征维数)确定隐含层神经元的上下限,并通过试错法将产生最小泛化误差的神经元数确定为最优隐含层神经元数^[26]。

有关 SVR 的理论已有很多研究人员进行过详细叙述,具体见文献[4,27-28]。建立该模型需要确定的主要参数有核函数、不敏感系数 ε 、惩罚因子 C 、核函数宽度系数 γ 。研究中选择径向基函数(RBF)作为核函数。不敏感系数 ε 选取平台默认值 0.1, C 和 γ 的取值分别设为 $\{2^1, 2^2, \dots, 2^5\}$ 和 $\{2^{-5}, 2^{-4}, \dots, 2^{-1}\}^{[19]}$,于建模数据中使用网格搜索和5折交叉验证在取值范围内进行二者的组合寻优。

根据 1.1 节 ELM 基础原理,ELM 模型待定参数为隐含层神经元数 k 和激活函数类型。参考 Mouatadid 等^[6]对于 ELM 模型的训练方案,选取 Sig 作为激活函数,将 k 的取值设定为 $\{10, 20, \dots, 200\}$,使用试错法在泛化误差最小时获取 k 的最佳取值。

根据 1.2 节 KELM 基础原理可知,建立该模型需要确定的参数包括核函数、惩罚因子 C 。经过试验选定 linear 作为核函数,将惩罚因子 C 的取值设为 $\{2^{-20}, 2^{-19}, \dots, 2^{-10}\}^{[9]}$,使用试错法在泛化误差最小时获取最佳参数取值。

组合模型 * + FS 由初始预测模块和 FS 残差修正模块耦合而成,初始预测模块的参数确定方法与

前述单独模型的方法一致。对于 FS 模块,Odan 和 Reis^[13]在建立 FS 模型时指出按照每 7 d 为一个周期对傅里叶系数进行计算且每 24 h 进行一次更新是足够的,本研究采用与之相同的计算和更新策略,将 FS 模块一周期所包含的时刻数设为 $T=7 \times 96=672$,此时,傅里叶级数 M 应该满足 $M < T/2$ ($=336$),通过测试确定选用的 M 为 50。

表 2 给出了实验中各模型的主要参数取值。

表 2 各单独模型的主要参数取值
Tab.2 Main parameter values of each single model

编号	ANN	SVR		ELM	KELM	* + FS
	k	C	γ	k	C	M
DMA1	5	16	2^{-3}	130	2^{-16}	50
DMA2	3	32	2^{-5}	180	2^{-15}	50
DMA3	5	16	2^{-3}	190	2^{-14}	50

3.3 结果分析

表 3 给出不同模型对 3 个 DMA 进行连续 24 h 且步长为 15 min 的短期需水量预测评价结果。

3.3.1 KELM 模型性能分析

对于预测日,步长为 15 min 的观测值曲线和各单独模型预测值曲线见图 3。

KELM 模型能够产生与 SVR、ANN 模型相似的精度。根据表 3 统计数据可知,研究所引入的 KELM 模型在 3 个 DMA 的预测测试中都产生了与广为应用的 ANN 和 SVR 相近的较高精度。以指标 R^2 为例,ANN、SVR、KELM 模型在 3 个 DMA 测试中评价结果的均值分别为 0.96、0.96 和 0.97。观察图 3 也可以看出,KELM、ANN、SVR 的预测曲线非常接近并且与观测值曲线趋势一致,同时较好地捕捉了曲线峰谷变化。

与 ELM 模型相比,KELM 模型的表现更加稳定。ELM 模型在 3 个 DMA 短期需水量测试中表现出了较大差异,如对历史需水量数据变异系数最高

的 DMA1 进行预测时,其评价指标 R^2 仅为 0.76, E_{MAP} 也高达 12% 以上,而对其他两个 DMA 的测试中 R^2 均达到了 0.9 以上;KELM 模型则始终发挥稳定, R^2 保持在 0.95 以上。从图 3 中还可观察到相比其他单独模型,ELM 在对波峰波谷(时刻 30 ~ 35, 时刻 85 ~ 90)的识别中也明显不够准确。此外,从

表 3 数据可以看出,增加修正模块 FS 后,原本精度较低的 ELM 模型,得到的周期性残差补偿更为明显,这也可从侧面说明,KELM 模型相比 ELM 模型能够更多地识别周期性现象并将其耦合到自身的预测过程中。

表 3 各预测模型在测试集中的评价结果
Tab.3 Performance indicators of prediction models on test set

模型	R^2			$E_{\text{MAP}}/\%$			计算时间/s		
	DMA1	DMA2	DMA3	DMA1	DMA2	DMA3	DMA1	DMA2	DMA3
KELM	0.977	0.965	0.958	4.329	4.387	3.890	6.513	6.433	6.005
ANN	0.973	0.963	0.957	4.653	4.300	3.980	123.191	123.837	118.777
SVR	0.975	0.957	0.955	4.487	4.458	4.027	117.465	115.265	105.468
ELM	0.762	0.916	0.918	12.538	6.734	4.027	25.148	27.440	26.392
KELM + FS	0.980	0.974	0.962	3.809	3.633	3.565	10.928	11.121	10.621
ANN + FS	0.976	0.974	0.960	4.216	3.678	3.608	130.228	129.374	123.811
SVR + FS	0.978	0.972	0.958	4.024	3.744	3.698	122.472	120.025	111.181
ELM + FS	0.914	0.944	0.957	8.335	5.644	3.698	30.297	32.172	30.895

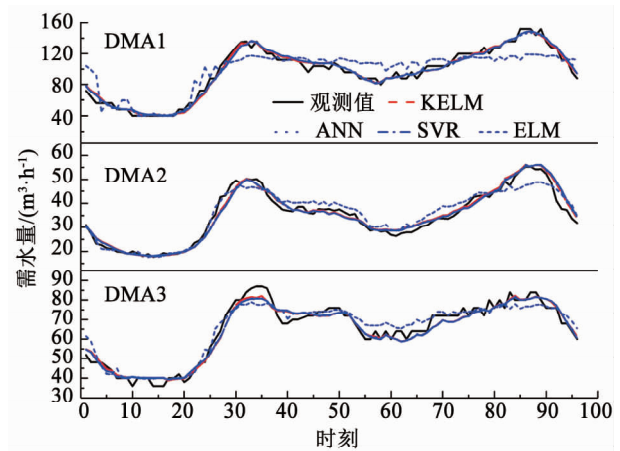


图 3 各单独模型预测值与观测值曲线
Fig.3 Curves of predicted and observed values of each single model

除了在计算精度和稳定性方面表现优良外,KELM 模型的更大优势体现在其计算速度上。KELM 与 ELM 一样,在训练过程中不再调整隐含层参数,使得模型训练所需时间明显下降,其在 3 个 DMA 的预测实验中均未超过 7 s,这也预示着其在

产生与 ANN、SVR 相似的精度时只需花费二者平均时间的 5% 左右。

为了测试 KELM 模型的泛化能力,使用 KELM 模型对 3 个 DMA 分别进行了连续 5 d(2019 年 7 月 26 日—2019 年 7 月 30 日)的需水量预测实验,其中 2019 年 7 月 26 日、29 日和 30 日为工作日,2019 年 7 月 27 日和 28 日为休息日,训练模型时使用每个预测日前 55 d 数据提取特征后的 90% 作为训练集,10% 作为验证集。表 4 给出了对应的评价结果,以 R^2 为例,KELM 模型在 3 个 DMA 中的平均精度为 0.97、0.95 和 0.95,表现出良好的精度效果。图 4 给出了连续预测中步长为 15 min 的预测值与观测值曲线,预测值始终表现出与观测值非常一致的趋势,也较为准确地捕捉到了峰谷变化。此外,3 个 DMA 工作日和休息日的 E_{MAP} 平均值分别为 4.47% 和 4.55%,对于工作日与休息日的预测精度并无较大差别。可见,KELM 模型在应对连续预测休息日与工作日的不同时能够表现出良好的泛化能力。

表 4 KELM 模型连续 5 d 预测的结果评价
Tab.4 Performance indicators of KELM prediction model on five consecutive days

日期	R^2			$E_{\text{MAP}}/\%$		
	DMA1	DMA2	DMA3	DMA1	DMA2	DMA3
2019-07-26	0.977	0.965	0.958	4.329	4.387	3.890
2019-07-27	0.951	0.951	0.947	5.135	5.193	3.766
2019-07-28	0.967	0.970	0.956	4.626	4.622	3.971
2019-07-29	0.964	0.941	0.942	4.978	5.097	4.073
2019-07-30	0.977	0.947	0.935	4.184	5.067	4.256

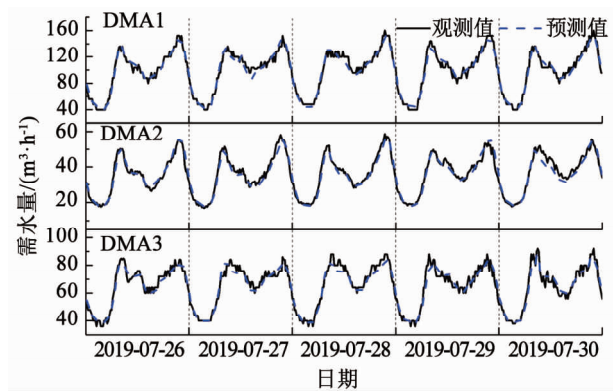


图4 KELM模型连续5 d的预测值与观测值曲线

Fig. 4 Curves of predicted and observed values of KELM model on five consecutive days

3.3.2 KELM + FS 模型性能分析

测试结果表明,为KELM模型增加修正模块FS后,其精度得到了一定程度的提升。表3的统计数据显示,KELM + FS模型在对3个DMA的需水量预测实验中在所有模型中精度表现最佳,相比KELM模型,以 E_{MAP} 指标为例,KELM + FS模型在3个DMA上的相对降幅分别为12.0%、17.2%和8.4%。事实上,增加修正模块后精度提高不仅限于KELM + FS模型,图5展示了所有单独模型及其组合模型的绝对百分比误差(APE, E_{AP})箱线图。可以看出,经过残差修正的模型相比对应的单独模型无论是在 E_{AP} 的数值范围、平均值,还是中位数方面都有所下降,这也在一定程度上验证了研究中所应用的修正模块的确可以捕捉残差序列中蕴含的周期性特征,并将其补偿到最终预测结果中,发挥进一步降低单独模型误差的作用。

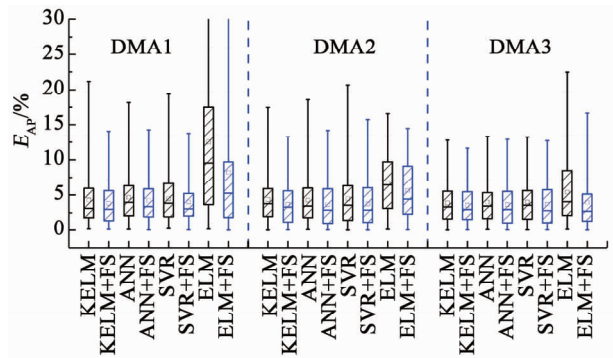


图5 各模型绝对百分比误差箱线图

Fig. 5 Boxplot of absolute percentage error of each model

综合来看,所提出的KELM + FS组合模型除拥有较高的预测精度外,在完成模型预测所花费的时间上同样极具优势,表现出高效预测性能。图6给出了所有模型在3个DMA中完成预测范围24 h,预测频率15 min的需水量预测的平均时间柱状图,可见作为耗时最短的KELM模型在添加修正模块后,

尽管其时间消耗有所提高,但KELM + FS仍然比目前广为应用的ANN和SVR模型快了大约10倍,这一点使得该模型在需要同时对大量用户进行预测的情况下仍然具有难以忽视的实际应用潜力。

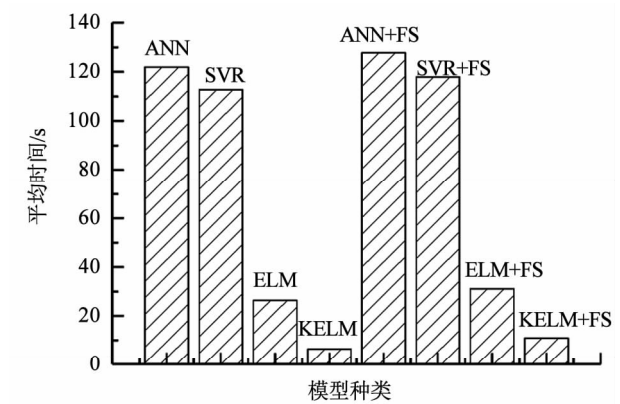


图6 各模型计算所需平均时间柱状图

Fig. 6 Bar chart of the average time required for each model

4 结 论

- 1)对短期需水量进行预测时,KELM模型计算精度与ANN模型、SVR模型相近(相对误差1.8%~5.5%),但KELM预测模型的建模时间为ANN、SVR模型建模时间的5%左右;相比极限学习机模型(ELM),KELM的预测性能更加稳定。
- 2)KELM模型具有较好的泛化能力,能够高效地实现连续日用水量预测,对工作日与休息日的用水量预测结果均达到较高的精度。
- 3)核极限学习机模型与傅里叶级数残差修正模块构成的组合预测模型KELM + FS,可以在不显著增加预测时间的前提下,提高KELM预测模型的精度。

参考文献

[1] PACCHIN E, GAGLIARDI F, ALVISI S, et al. A comparison of short-term water demand forecasting models[J]. Water Resources Management, 2019, 33(4): 1481. DOI:10.1007/s11269-019-02213-y

[2] GUO G, LIU S, WU Y, et al. Short-term water demand forecast based on deep learning method[J]. Journal of Water Resources Planning and Management, 2018, 144(12): 1. DOI:10.1061/(asce)wr.1943-5452.0000992

[3] HERRERA M, TORGO L, IZQUIERDO J, et al. Predictive models for forecasting hourly urban water demand[J]. Journal of Hydrology, 2010, 387(1/2): 141. DOI:10.1016/j.jhydrol.2010.04.005

[4] 陈磊, 张士乔. 基于最小二乘支持向量机的时用水量预测模型[J]. 哈尔滨工业大学学报, 2006, 38(9): 1528

CHEN Lei, ZHANG Tuqiao. Hourly water demand forecast model based on least squares support vector machine[J]. Journal of Harbin Institute of Technology, 2006, 38(9): 1528. DOI:10.3321/j.issn:0367-6234.2006.09.031

- [5] GHALEHKHOND I, ARDJMAND E, YOUNG W A, et al. Water demand forecasting: review of soft computing methods [J]. *Environmental Monitoring & Assessment*, 2017, 189 (7): 313. DOI:10.1007/s10661-017-6030-3
- [6] MOUATADID S, ADAMOWSKI J. Using extreme learning machines for short-term urban water demand forecasting [J]. *Urban Water Journal*, 2016, 14 (6): 630. DOI: 10.1080/1573062x.2016.1236133
- [7] 殷豪, 董联, 孟安波. 基于结合混沌纵横交叉的粒子群算法优化极限学习机的短期负荷预测[J]. *计算机应用研究*, 2018, 35 (7): 2088
YIN Hao, DONG Zhen, MENG Anbo. Short-term load forecasting based on extreme learning machine optimized by particle swarm optimization integrated with chaotic crisscross optimization [J]. *Application Research of Computers*, 2018, 35(7): 2088. DOI:10.3969/j.issn.1001-3695.2018.07.039
- [8] VIJAI P, BAGAVATHI S P. Performance comparison of techniques for water demand forecasting [J]. *Procedia Computer Science*, 2018, 143: 258. DOI:10.1016/j.procs.2018.10.394
- [9] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multiclass classification[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Part B, Cybernetics*, 2012, 42 (2): 513. DOI:10.1109/TSMCB.2011.2168604
- [10] FU H, VONG C M, WONG P K, et al. Fast detection of impact location using kernel extreme learning machine [J]. *Neural Computing and Applications*, 2014, 27 (1): 121. DOI:10.1007/s00521-014-1568-2
- [11] AL-ZAHRANI M A, ABO-MONASAR A. Urban residential water demand prediction based on artificial neural networks and time series models[J]. *Water Resources Management*, 2015, 29(10): 3651. DOI:10.1007/s11269-015-1021-z
- [12] ANELE A, HAMAM Y, ABU-MAHFOUZ A, et al. Overview, comparative assessment and recommendations of forecasting models for short-term water demand prediction[J]. *Water*, 2017, 9(11): 887. DOI:10.3390/w9110887
- [13] ODAN F K, REIS L F R. Hybrid water demand forecasting model associating artificial neural network with Fourier series[J]. *Journal of Water Resources Planning and Management*, 2012, 138 (3): 245. DOI:10.1061/(asce)wr.1943-5452.0000177
- [14] BRENTAN B M, LUVIZOTTO J E, HERRERA M, et al. Hybrid regression model for near real-time urban water demand forecasting [J]. *Journal of Computational and Applied Mathematics*, 2017, 309: 532. DOI:10.1016/j.cam.2016.02.009
- [15] CHEN X, DONG Z Y, MENG K, et al. Electricity price forecasting with extreme learning machine and bootstrapping [J]. *IEEE Transactions on Power Systems*, 2012, 27 (4): 2055. DOI:10.1109/tpwrs.2012.2190627
- [16] YASEEN Z M, SULAIMAN S O, DEO R C, et al. An enhanced extreme learning machine model for river flow forecasting: state-of-the-art, practical applications in water resource engineering area and future research direction[J]. *Journal of Hydrology*, 2019, 569:387. DOI:10.1016/j.jhydrol.2018.11.069
- [17] 郑晓芬, 钟旺, 李春祥. 基于核极限学习机的多变量非平稳脉动风速预测[J]. *振动与冲击*, 2017, 36(18): 223
ZHENG Xiaofen, ZHONG Wang, LI Chunxiang. Multivariate nonstationary fluctuating wind velocity prediction using kernel-based extreme learning machine [J]. *Journal of Vibration and Shock*, 2017, 36(18): 223. DOI:10.13465/j.cnki.jvs.2017.18.033
- [18] HUANG G B. An insight into extreme learning machines: random neurons, random features and kernels[J]. *Cognitive Computation*, 2014, 6(3): 376. DOI:10.1007/s12559-014-9255-2
- [19] SHAMSHIRBAND S, MOHAMMADI K, CHEN H L, et al. Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: a case study for city of Bandar Abass, Iran[J]. *Journal of Atmospheric & Solar Terrestrial Physics*, 2015, 134(105): 109. DOI:10.1016/j.jastp.2015.09.014
- [20] KONGSOROT Y, HORATA P, MUSIKAWAN P, et al. Kernel extreme learning machine based on fuzzy set theory for multi-label classification [J]. *International Journal of Machine Learning and Cybernetics*, 2017, 10(5): 979. DOI:10.1007/s13042-017-0776-3
- [21] BAKKER M, VREEBURG J H G, VAN SCHAGEN K M, et al. A fully adaptive forecasting model for short-term drinking water demand [J]. *Environmental Modelling & Software*, 2013, 48 (3): 141. DOI:10.1016/j.envsoft.2013.06.012
- [22] CUTORE P, CAMPISANO A, KAPELAN Z, et al. Probabilistic prediction of urban water consumption using the SCEM-UA algorithm [J]. *Urban Water Journal*, 2008, 5 (2): 125. DOI:10.1080/15730620701754434
- [23] 王亮, 张宏伟, 牛志广. 支持向量机在城市用水量短期预测中的应用[J]. *天津大学学报*, 2005, 38(11): 1021
WANG Liang, ZHANG Hongwei, NIU Zhiguang. Application of support vector machine in short-term prediction of urban water consumption[J]. *Journal of Tianjin University*, 2005, 38 (11): 1021. DOI:10.3969/j.issn.0493-2137.2005.11.017
- [24] WANAS N, AUDA G, KAMEL M S, et al. On the optimal number of hidden nodes in a neural network [C]//IEEE Canadian Conference on Electrical & Computer Engineering. Waterloo: IEEE, 1998: 918
- [25] MISHRA A K, DESAI V R. Drought forecasting using feed-forward recursive neural network [J]. *Ecological Modelling*, 2006, 198(1/2): 127. DOI:10.1016/j.ecolmodel.2006.04.017
- [26] JIA Y, CULVER T B. Bootstrapped artificial neural networks for synthetic flow generation with a small data sample[J]. *Journal of Hydrology*, 2006, 331 (3/4): 580. DOI: 10.1016/j.jhydrol.2006.06.005
- [27] 朱志宇, 姜长生, 张冰. 基于支持向量回归的混沌序列预测方法[J]. *电工技术学报*, 2005, 20(6): 57
ZHU Zhiyu, JIANG Changsheng, ZHANG Bing. Prediction of chaotic time series based on support vector regression [J]. *Transactions of China Electrotechnical Society*, 2005, 20(6): 57. DOI:10.3321/j.issn:1000-6753.2005.06.012
- [28] MIRANIAN A, ABDOLLAHZADE M. Developing a local least-squares support vector machines-based neuro-fuzzy model for nonlinear and chaotic time series prediction[J]. *IEEE Transactions Neural Network Learn Systems*, 2013, 24 (2): 207. DOI:10.1109/TNNLS.2012.2227148

(编辑 刘 彤)