

张贻婷, 李天宏. 基于长短时记忆神经网络的河流水质预测研究[J]. 环境科学与技术, 2021, 44(8): 163–169. Zhang Yiting, Li Tianhong. River water quality prediction based on long short-term memory neural network[J]. Environmental Science & Technology, 2021, 44(8): 163–169.

基于长短时记忆神经网络的河流水质预测研究

张贻婷^{1,2}, 李天宏^{1,2*}

(1. 北京大学环境科学与工程学院, 北京 100871;

2. 国家环境保护河流全物质通量重点实验室, 北京 100871)

摘要: 准确高效地预测河流水质变化趋势对河流环境治理与保护具有重要意义。该文利用广州市白坭河上自动监测站每2 h的水质数据, 从单测站数据时序之间的相关性和上游测站的影响两方面, 分别建立长短时记忆网络(LSTM)河流水质预测的循环神经网络模型。模型以氨氮浓度为输出变量, 比较了不同输入变量下的模型预测效果, 并以最优模型和常用的深度学习算法支持向量机(SVM)进行了比较。结果表明: 单测站 LSTM 模型经输入变量特征选择后的预测结果比仅使用氨氮浓度单变量的时间序列预测更接近真实数值; 对加入上游监测站的双测站 LSTM 模型, 输入的变量经过特征选择时, 模型预测效果优于全部水质变量作为输入的预测结果, 也优于单测站 LSTM 模型; 但不进行特征选择时, 输入变量增加, 模型学习到噪声而使精度下降; 和 SVM 模型相比, 最优特征组合的 LSTM 模型具有更好的预测效果。研究也表明, 对输入变量进行特征选择后, LSTM 模型是一种有潜力的河流水质预测方法。

关键词: 水质预测; 长短时记忆网络; 循环神经网络; 深度学习; 特征选择

中图分类号: TP183; X832 **文献标志码:** A **doi:** 10.19672/j.cnki.1003-6504.0018.21.338 **文章编号:** 1003-6504(2021)08-0163-07

River Water Quality Prediction Based on Long Short-term Memory Neural Network

ZHANG Yiting^{1,2}, LI Tianhong^{1,2*}

(1. College of Environmental Science and Engineering, Peking University, Beijing 100871, China;

2. State Environmental Protection Key Laboratory of All Material Fluxes in River Ecosystems, Beijing 100871, China)

Abstract: Predicting water quality in rivers with high accuracy and efficiency is of great importance for water environment management and protection. Long short-term memory neural network (LSTM) was used to predict water quality fed with data gauged by every two hours from two automatic monitoring stations at Baini River in Guangzhou, China. Correlations between ammonia nitrogen concentration and other environmental variables in one station were analyzed, spatial response relationships between two monitoring stations was also taken into consideration in correlation analysis. LSTM models were developed to forecast ammonia nitrogen concentration with several input and output combinations; and the prediction results by different model inputs were compared. In addition, LSTM model with the best performance was also compared with support vector machine (SVM), a commonly used data-mining algorithm. The results showed that feature selection improved LSTM predictive performance compared with time series forecasting of ammonia nitrogen concentration when using data from single monitoring station. LSTM model could be improved by adding data of the upstream monitoring station, but using all environmental variables from two stations as inputs decreased model performance. LSTM model illustrated the highest prediction accuracy when feature selection was conducted on input variables and the geographic relationship were considered at the same time. In this case, LSTM exceeded SVM in its highest performance situations. Also, the results suggested that LSTM would be a potential water quality forecasting method for rivers while input variables are reasonably selected.

Key words: water quality prediction; long short-term memory networks; RNN; deep learning; feature selection

可靠的水质模型对预测水体水质变化趋势、采取治理措施和建立决策预警机制具有重要意义^[1-3]。河流水质在宏观时间尺度上表现出季节性及周期性等

特点, 具有渐变性、非线性和不确定性。河流水质参数不仅受外在因素的影响, 还受到自变量历史值和随机扰动影响, 滞后项与相关变量的选取是影响预测精

《环境科学与技术》编辑部: (网址) <http://fjks.chinajournal.net.cn> (电话) 027-87643502 (电子信箱) hjksxyjs@vip.126.com

收稿日期: 2021-03-05; 修回 2021-06-17

基金项目: 国家自然科学基金(51721006)

作者简介: 张贻婷(1996-), 女, 硕士研究生, 主要从事工程环境模拟与信息系统方面的研究, (电子信箱) inty@pku.edu.cn; *通讯作者, 男, 研究员, 博士, (电子信箱) lth@pku.edu.cn。

度的因素之一^[4,5]。传统建模方法在处理大量未知或未指定输入数据的数据集时耗时耗力。近年来,机器学习算法因其在非线性模拟方面的良好特性,在水环境建模中越来越多地得以应用,其中以人工神经网络算法尤为典型^[6-10]。此外,改进的机器学习模型或结合不同模型特点的混合模型^[11-15]也用于水质预测。由于传统神经网络的神经元结构不适合处理有长期依赖关系的数据,有研究者引入长期短期记忆网络(long short-term memory, LSTM)算法,通过在输入和输出之间建立专门记忆较长时间间隔的结构对神经网络算法进行改进并应用于水质预测^[16]。利用LSTM算法模拟EFDC水质模型的研究^[17]也表明LSTM可以达到EFDC模型的预测能力。当前基于LSTM的水质预测较少进行长时间序列预测,未充分发挥深度学习在大数据量方面优势。深度学习在训练过程中能学习到与目标变量相关的特征,然其过程未知,无法从原理方面解释,因此模型输入前进行特征工程的必要性一直存在讨论空间,探讨模型输入变量选择对预测结果的影响,也有助于对模型过程做出解释。本研究在通过数据挖掘获得水质数据之间关联性的基础上,选择不同的输入变量,构建基于LSTM的水质预测模型,以提高河流水质预测的精度和效率。

1 材料与方法

1.1 研究区域

白坭河位于珠江三角洲北端,流域面积约1 493 km²,主要流经广州市的花都区与白云区,上游与北江相通,向下汇入珠江。流经区域属亚热带季风气候

区,季风环流盛行。本研究所用的数据为广州市白坭河支流的2个水质自动监测站的水质数据,上游为支流兴华涌站,下游为天马河站。2个站的位置及其汇水范围如图1所示。



图1 监测站位置及汇水区

Fig.1 The location of monitoring stations and catchment areas

1.2 数据来源和预处理

兴华涌站和天马河站2个自动监测站的水质数据采集时间均为2017年7月-2019年12月,每2 h采集一次。监测的水质指标包括透明度、溶解氧、ORP、水位、气温、环境湿度、氨氮等,采集及分析方法均参照《地表水环境质量标准》(GB 3838-2002)进行。对数据的缺失值采用2种处理方式:如果一次监测中仅缺失一个指标,即使用前一次的监测值来补上;如果某个监测值连续缺失,为避免人为填补造成较大误差则删除这些监测时刻的数据。最终取前9 588组作为训练数据集,后732组数据作为检验数据集。表1是监测数据的描述性统计情况。

表1 水质时间序列的描述性统计情况

Table 1 Descriptive statistics of water quality parameters

	透明度/cm	溶解氧/(mg·L ⁻¹)	ORP/mV	氨氮/(mg·L ⁻¹)	水位/m	气温/℃	环境湿度/%
数量	9 939	10 015	10 016	10 016	9 864	10 015	10 015
均值	23.8	3.4	367.6	3.8	2.3	22.8	79.5
标准差	6.4	2.1	117.3	3.0	2.0	2.8	15.6

平均完整率98.99%;平均有效率95.67%;平均故障率4.33%。

长短时记忆网络在输入数据前需先将数据转为三维张量,并选择回溯时间步长与每批次数量,对应将输入数据映射为三维形式,输入模型的训练数据集,这一过程采用图2滑动窗口法实现,预测下一时刻水质。

此外,机器学习需将输入数据标准化,降低模型陷入局部最优的可能性,提升模型收敛速度。本研究选用离差标准化方法,即把每个输入数据映射到[0,1]区间,计算公式如式(1),其中 m 为原始数据标准化后的数据值, x 为原始数据, x_{\min} 和 x_{\max} 分别是数据集的最小值和最大值。

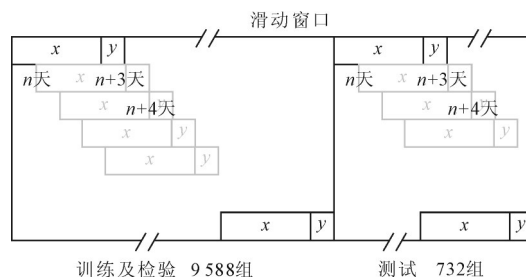


图2 滑动窗口示意图

Fig.2 The developed rolling data set

$$m = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})} \quad (1)$$

1.3 水质相关性分析

时间序列预测的本质是对历史数据进行挖掘和分析,根据历史数据之间的关系来预测未来的变化过程^[18]。长期短期记忆网络不仅能进行单变量时间序列预测,也可以进行多维变量输入的预测,本研究透过上下游点位之间的关系,考虑地理空间位置影响,提升模型预测能力。

通过特征选择可以减少模型学习时间,提升模型效率,使模型泛化能力更强,也有助于增加模型的可解释性^[19,20]。本研究以相关性分析作为特征选择的依据,即通过皮尔逊相关系数及斯皮尔曼等级相关系数综合评估氨氮浓度与其他监测数据之间的相关性,并据此进行特征选择,确定模型的输入变量。

1.4 LSTM 水质预测模型

循环神经网络(RNN)无法学习长期的记忆,循环多次后发生梯度消失。目前主流使用由Hochreiter和Schmidhuber^[21]提出的LSTM来克服RNN无法记忆的问题。和RNN相比,LSTM增加了一个专门用来记忆的单元(\tilde{C}_t)来记忆上一轮的结果,其内部结构如图3所示。

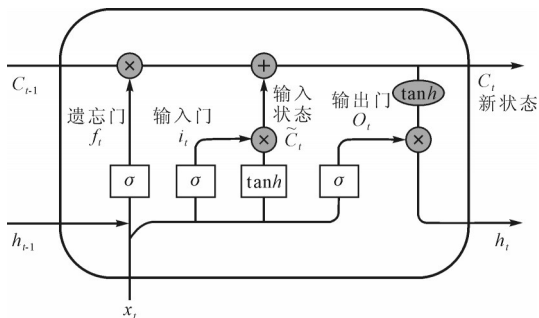


图3 LSTM内部结构示意图
Fig.3 LSTM internal architecture

LSTM是4个神经层交互形成,第1层由遗忘门来确定输入单元状态,公式如式(2),用来决定多少保留,多少信息传递到下一轮,通过sigmoid转换决定是否保留。

$$f_t = \sigma(W_f \cdot [h_{t-1} \cdot x_t] + b_f) \quad (2)$$

下一步决定存储新信息,包含两部分,包含sigmoid层和tanh层,前者决定对哪些数值进行修改,后者会生成新的向量,输入门公式如式(3),新单元的状态公式如式(4):

$$i_t = \sigma(W_i \cdot [h_{t-1} \cdot x_t] + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1} \cdot x_t] + b_c) \quad (4)$$

接着更新旧单元状态,将旧单元状态乘以遗忘门加上新候选值。最后经由输出门过滤新增并决定输出内容。输出门公式如式(5),最后输出公式如式(6):

$$o_t = \sigma(W_o \cdot [h_{t-1} \cdot x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

模型结构和主要参数如表2所示,第1层为300个神经元的LSTM训练模型,根据上述原理,模型基于过往的拟合结果进行学习,利用LSTM具有记忆的特性优化水质预测能力。第2层的目的是防止发生过拟合情况,让所有神经元更好的学习,之后连接到第3层用来对上一层的神经元进行全连接,最终在第4层经过激活函数后输出。

表2 模型结构
Table 2 Network structure

神经元层	模型	输出张量	参数量
1	LSTM	(None, 300)	369 600
2	Dropout	(None, 300)	0
3	Dense	(None, 1)	301
4	Activation	(None, 1)	0

1.5 模型比较与定量评估方法

为了评价本文基于深度学习和LSTM的水质预测方法的可靠性,选择常用的机器学习算法支持向量机回归(SVR)作为对比。支持向量回归机通过“核映射”把输入样本空间映射到高维的特征空间,在高维特征空间中进行线性回归来实现非线性处理^[22],在小样本数据的回归和预测具有良好的逼近能力和泛化能力^[23]。SVM基本型以拟合一回归函数并向外延伸出间隔范围,为了最大化间隔使最大化(见式(7)),使数据点到此范围的距离最小,函数尽量逼近每个数据,于是转化为一个凸二次规划问题,其形式为式(8),使回归模型的优化目标函数为 $\frac{1}{2} \|\omega\|^2$,如式(7)~(8)。 ξ_i 与 ξ_i^* 表示数据到超平面两侧的距离, C 为惩罚系数,其作用是平衡结构风险。

$$\max \frac{2}{\|\omega\|} + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (7)$$

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (8)$$

本研究采用平均绝对误差(MAE),均方误差(MSE)和均方根误差(RMSE)来定量评估模型预测效果(式(9)~(11))。为减少算法的随机性误差,在实验过程中设置了一个随机种子(random seed),以确保运算结果的一致性。

$$MAE = \frac{\sum_{t=1}^n |A_t - F_t|}{n} \quad (9)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2 \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}} \quad (11)$$

式(9)~(11)中, A_t 为实测值; F_t 为预测值。

2 结果与分析

2.1 模型输入特征选择

氨氮浓度在滞后2 h、1 d和7 d的自相关散布矩阵

情形如图4,间隔2 h的氨氮浓度呈现明显的线性关系,表示水质变化并非随机过程,而是呈现时间序列

变化。时间间距1 d的水质情形仍呈现线性关系,7 d后线性趋势降低,范围逐渐变宽,离群点增多。

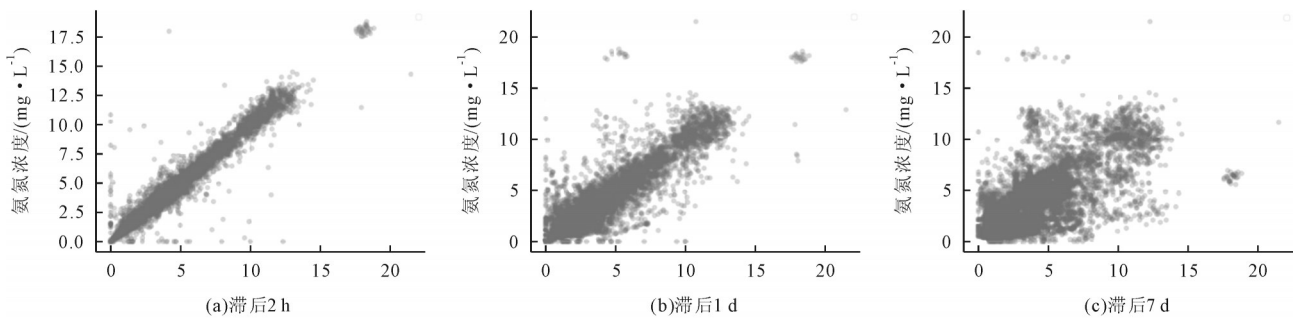


图4 氨氮水质自相关散布矩阵
Fig.4 Correlation matrix of ammonia nitrogen

本研究考虑流域上下游站点氨氮浓度在空间上的关联性。图5(a)散布矩阵显示了上下游氨氮浓度间的相关性,从中可以看出部分数据点成线性关系,两者之间分布多数集中在下游氨氮浓度较低的部份。

图5(b)及图5(c)对应上游两测站的氨氮浓度分布,可以看出兴华涌氨氮较集中分布在低浓度区间,下游天马河的氨氮浓度变化范围较大,数值范围变化相较上游的兴华涌剧烈。

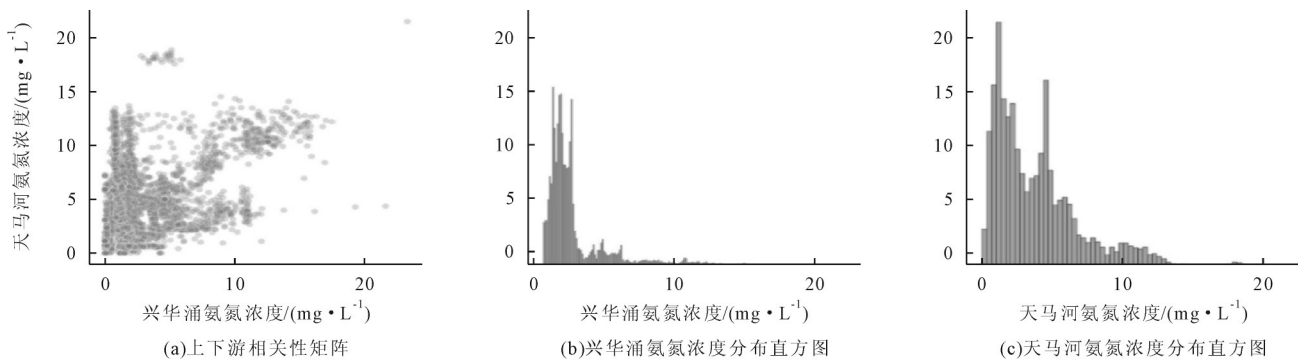


图5 氨氮水质浓度分布
Fig.5 Distribution of ammonia nitrogen

以皮尔逊相关系数获得相关先验氨氮浓度、气温和其他水质参数之间的关系。水质预测利用预处理后的数据及其相关信息,构建基于深度LSTM的模型。假设输入与输出样本关系,根据研究背景环境的复杂性,适应水质变化的动态特性提高预测精度,根据所选用的模型对水质数据建模与优化,提高水质

预测的效率和精度,最终选择模型输入的特征选择见表3。模型根据数据来源分为2类,一类是单一监测站以不同特征训练水质预测模型,第二类是考虑上游测站地理空间响应关系的预测模型。最后讨论两者不同输入变量的预测效果,并和常用的深度学习算法SVM模型进行比较。

表3 LSTM模型输入变量
Table 3 LSTM input variables

数据来源	选择方法	输入变量
单一测站	所有特征	透明度、溶解氧、ORP、水位、气温、环境湿度、氨氮
	特征选择	气温、水位、氨氮
	时间序列	氨氮
上下游测站	特征选择 时间序列	上游:水位、氨氮;下游:气温、水位、氨氮 氨氮

2.2 单变量时间序列预测

使用LSTM构建模型,分析不同输入变量情形下模型预测效果。位于下游的天马测站氨氮浓度变化趋势如图6(a)。图6(b)显示了分别使用单变量时间序列、特征选择、全部变量作为模型输入变量时的预

测结果,表4比较上述几种情形预测的评价指标。结果表明,单变量作为模型输入的时间序列模型能达到一定预测效果,加入水位及气温进行预测,MSE从0.679 3降至0.542 2。从图6(b)观察预测的趋势线,氨氮浓度高时,仅使用单变量作为输入模拟数值偏低,

以全部特征作为输入时预测精度较高,但当后期浓度下降后,经过特征选择后的模型(蓝线)能更適切贴近真实值,证实通过特征选择后再进行训练,模型预测效果提升。使用所有特征进行训练(绿线)在浓度高时,更贴近实际监测站数值,但当氨氮浓度下降后预测值仍偏高,可能是由于模型训练时与其他变量间的差异并无相应降低,导致模型没有学习到相应的变

化。从图6(b)和表4也可以看到,预测7 d的结果更能显现出特征选择后的模型预测效果优于其他两者。因此基于相关性分析后输入变量可提升模型预测效果,优于仅靠氨氮进行时间序列的预测;而将测站所有变量都用于模型输入时,模型预测结果的整体精度相较单一变量有提高,但在预测后期氨氮浓度较低时预测值偏大,预测效果下降。

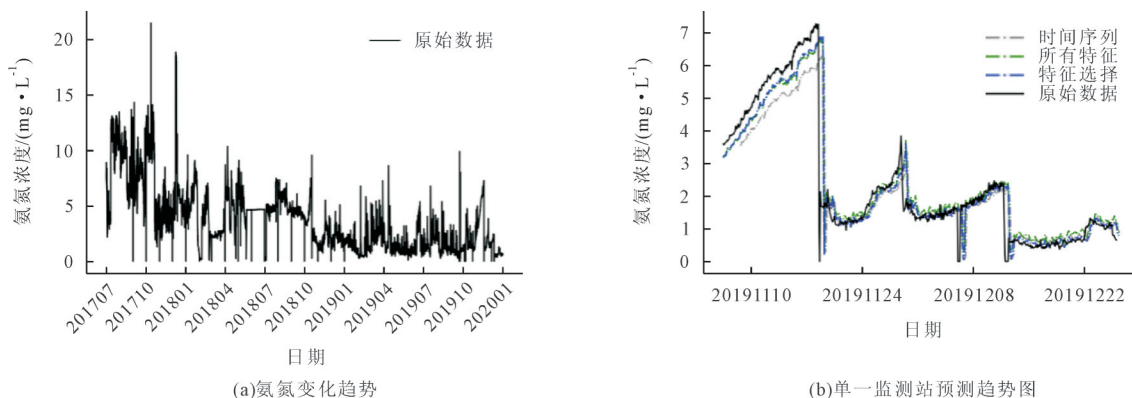


图6 天马监测站氨氮趋势
Fig.6 Graphs of ammonia nitrogen variation trend of Tianma

表4 LSTM单一监测站训练评价指标

Table 4 Accuracy assessment of the models with single-station monitoring data set

名称	实验参数	输出	评价指标			评价指标(预测7 d)		
			MSE	MAE	RMSE	MSE	MAE	RMSE
特征选择	气温、水位、氨氮	氨氮	0.542 2	0.338 1	0.736 4	0.040 1	0.152	0.200 2
时间序列	氨氮	氨氮	0.679 3	0.468 1	0.824 2	0.066 4	0.212 3	0.257 8
所有特征	透明度、溶解氧、ORP、水位、气温、环境湿度、氨氮	氨氮	0.610 2	0.404 2	0.781 2	0.091 4	0.255 5	0.302 4

2.3 考虑上游测站影响的预测模型

考虑上游测站影响的几种模型的预测结果和实

测值的对比如图7所示,定量评价指标的对比结果如表5所示。

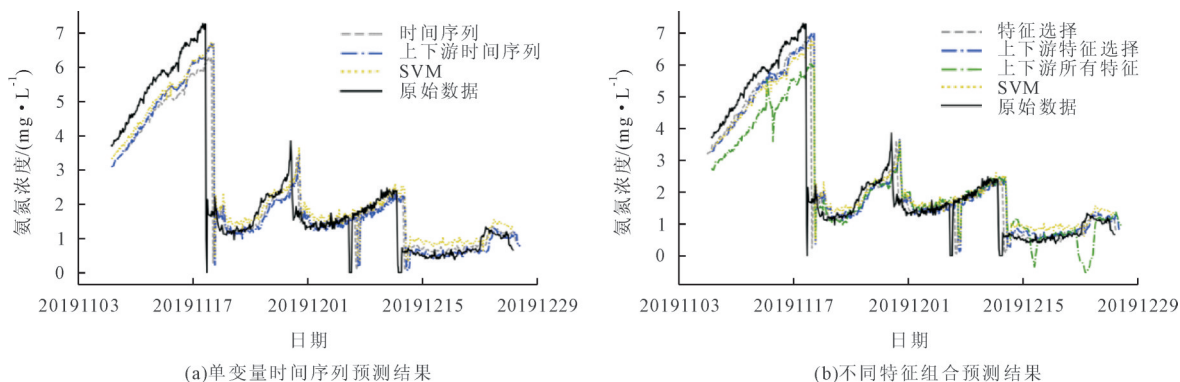


图7 加入上游数据集预测结果
Fig.7 Predictions of the LSTM model adding upstream monitoring station

表5 上下游测站特征预测评价

Table 5 Accuracy assessment of the models

模型	实验参数	评价指标		
		MSE	MAE	RMSE
上下游特征选择	上游:水位、氨氮;下游:气温、氨氮	0.488 6	0.317 0	0.699 0
上下游时间序列	上下游氨氮	0.503 7	0.363 1	0.709 7
特征选择	水位、气温、氨氮	0.542 2	0.338 1	0.736 4
时间序列	氨氮	0.679 3	0.468 1	0.824 2
上下游所有特征	上下游所有特征	1.091 4	0.678 0	1.044 7
SVM	上游:水位、氨氮;下游:气温、氨氮	0.866 7	0.527 2	0.931 0

结果表明,相比仅使用氨氮浓度这一目标变量进行时间序列的预测,加入上游测站数据后,MSE、MAE和RMSE分别降至0.503 7、0.363 1、0.709 7,图7(a)蓝线表现出上下游时间序列能反映氨氮变化趋势。模型采用特征选择后的输入变量时,MSE、MAE和RMSE分别降为0.488 6、0.317 0、0.699 0,模型精度提升。从图7(b)趋势图可以看出,加入上游测站的输入变量后,初期氨氮浓度偏高时更加符合真实数据,且随时间变化浓度下降后仍能维持预测能力,反映真实变化。最后将所有指标直接输入到模型中训练看是否大量的数据能否提高驯良效果,结果却显示预测精度最差,且从预测的趋势变化来看,表现也不稳定,不符合真实情形。双测站LSTM模型使用所有指标作为输入的结论与单站LSTM略有不同,说明当使用超过一定量的输入变量不仅无法提升模型效果,模型可能学到额外噪声使得学习效果变差。总体而言,加入上游数据集后的模型预测效果要明显优于仅使用单一测站数据的模型。当使用所有特征训练时,无关或冗余特征会干扰学习,导致精度下降。因此,在特征变量较多时,先采取特征工程对输入数据进行处理,对变量进行相关性分析后再决定输入维度,不仅可以降低模型复杂程度和节约计算成本,而且能适配预测结果,提高模型预测精度。

图7和表5也可以看出,选定LSTM模型效果最优的参数组合输入SVM模型进行预测,其MSE、MAE、RMSE分别为0.866 7、0.527 2、0.931 0,说明考虑上游站点影响的LSTM模型的预测精度也优于SVM模型。

3 讨论

基于对数据的挖掘与分析,对所有变量进行特征选择后确定模型输入,水质预测的精度得以提升,而加入所有特征后效果并未优化,说明即使增加模型输入维度不会使预测效果提升,反而可能因为输入不相关的变量而降低了预测精度。

对比上下游水质数据集与单一监测站水质数据进行预测的结果。加入地理位置上游的数据集相较于单一监测站进行数据挖掘更能使模型精度上升,但若使用所有变量进行训练,结果不如仅使用相关性高的输入变量。相较所有训练模型,使用空间关系并经由特征选择后的输入变量表现最接近真实值,本研究探讨对深度学习模型输入维度的选择,发现以同一监测站的水质数据进行挖掘,使用较多的维度虽然能提升精度,但过度的选择参数反而使降低效能,此外,应用空间上的关系,利用上游支流信息协同模型训练,反

而较单一测站数据挖掘的表现更接近真实情况。

此外,模型训练过程中SVM需要大量非支持向量参与训练过程,从而进行了大量的二次规划计算,导致速度慢。LSTM在输入变量较多时比SVM优势明显,计算精度高。秦文虎等^[24]探讨缺失值填补对水质模型的效果,使用LSTM与SVM,结果表示相较SVM模型,LSTM预测精度更高,结论与本研究一致。

4 结论

使用广州白坭河上2个自动监测站2017年7月-2019年底每2 h的水质数据,建立了深度学习的循环神经网络LSTM进行水质预测,比较不同输入情形下模型的预测效果,并与支持向量机模型对比了预测精度。结果表明在数据量大时深度学习能达到一定预测水平;通过不同输入维度的数据对模型进行优化,证实了利用上下游空间关系构建深度学习模型比单一监测站数据更能反映河流水质变化;输入变量经特征选择后,LSTM预测结果优于支持向量机,表明了LSTM对于河流水质预测的可靠性和应用潜力。

使用非机理模型进行水质预测可以降低水质预测成本,利用深度学习可以提高水质预测效果。然而如何获得更大量的数据作为模型输入以提升模型性能仍然是问题,当前主流方法是使用同一监测站其他水质指标进行预测。本研究考虑了上下游测站之间的联系后,水质预测精度明显提升。本研究也证实,加入更多维度的数据并未提升模型预测效能,因此特征选择在数量大且变量多时仍是建立高效非机理水质预测模型中一项重要任务。

[参考文献]

- [1] 周彦辰,胡铁松,陈进,等. 耦合动态方程的神经网络模型在水质预测中的应用[J]. 长江科学院院报, 2017,34(9): 1-5.
Zhou Yanchen, Hu Tiesong, Chen Jin, et al. Application of neural network model coupled with dynamic equation in water quality prediction[J]. Journal of Yangtze River Scientific Research Institute, 2017,34(9):1-5.
- [2] 李如忠. 水质预测理论模式研究进展与趋势分析[J]. 合肥工业大学学报:自然科学版, 2006,29(1):26-30.
Li Ruzhong. Advance and trend analysis of theoretical methodology for water quality forecast[J]. Journal of Hefei University of Technology: Natural Science Edition, 2006,29(1): 26-30.
- [3] 杜娟,宋鹏程. 中国地表水水质自动监测系统应用研究[J]. 环境科学与管理, 2020,45(2):115-119.

Du Juan, Song Pengcheng. Application and key questions of

- automatic water quality monitoring system in China[J]. Environmental Science and Management, 2020,45(2):115-119.
- [4] 罗学科,何云霄,刘鹏,等. ARIMA-SVR组合方法在水质预测中的应用[J]. 长江科学院院报, 2020,37(10):21-27.
Luo Xueke, He Yunxiao, Liu Peng, et al. Water quality prediction using an ARIMA-SVR hybrid model[J]. Journal of Yangtze River Scientific Research Institute, 2020, 37(10): 21-27.
- [5] 周志青,邹国防,王磊,等. 基于ARIMA/RBF-NN的时间序列水质预测模型研究[J]. 科技通报, 2017,33(9):236-240.
Zhou Zhiqing, Zou Guofang, Wang Lei, et al. Time-series prediction model of water quality based on ARIMA/RBF-NN[J]. Bulletin of Science and Technology, 2017, 33(9): 236-240.
- [6] Hu Z H, Zhang Y R, Zhao Y C, et al. A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture[J]. Sensors, 2019, 19(6): 1420.
- [7] Wang Y Y, Zhou J, Chen K J, et al. Water Quality Prediction Method Based on LSTM Neural Network[C]//Nanjing, China: 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). 2017:1-5.
- [8] Huang C J, Kuo P H. A deep CNN-LSTM model for particulate matter (PM_{2.5}) forecasting in smart cities[J]. Sensors, 2018,18(7):2220.
- [9] Dogan E, Sengorur B, Koklu R. Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique[J]. Journal of Environmental Management, 2009,90(2):1229-1235.
- [10] 赵军,张祯宇,谢哲宇,等. 基于BP人工神经网络的闽江口水厂水质模拟[J]. 环境科学与技术, 2020,43(S1):198-203.
Zhao Jun, Zhang Zhenyu, Xie Zheyu, et al. Water quality modeling for water works in Minjiang River estuary based on BP neural network model[J]. Environmental Science & Technology, 2020,43(S1):198-203.
- [11] 张颖,高倩倩. 基于灰色模型和模糊神经网络的综合水质预测模型研究[J]. 环境工程学报, 2015,9(2):537-545.
Zhang Ying, Gao Qianqian. Comprehensive prediction model of water quality based on grey model and fuzzy neural network[J]. Chinese Journal of Environmental Engineering, 2015,9(2):537-545.
- [12] Xu L Q, Liu S Y. Study of short-term water quality prediction model based on wavelet neural network[J]. Mathematical and Computer Modelling, 2013,58(3/4):807-813.
- [13] 冯民权,邢肖鹏,薛鹏松. BP网络马尔可夫模型的水质预测研究:基于灰色关联分析[J]. 自然灾害学报, 2011,20(5):169-175.
Feng Minquan, Xing Xiaopeng, Xue Pengsong. Water quality prediction using BP network and Markov model: based on grey incidence analysis[J]. Journal of Natural Disasters, 2011,20(5):169-175.
- [14] Ömer Faruk D. A hybrid neural network and ARIMA model for water quality time series prediction[J]. Engineering Applications of Artificial Intelligence, 2010,23(4):586-594.
- [15] 翟伟,毛静,孟雅丹,等. 结合灰色理论的人工神经网络方法在水质预测中的应用[J]. 南水北调与水利科技, 2020,18(1):138-143.
Zhai Wei, Mao Jing, Meng Yadan, et al. Study on the prediction of water quality based on artificial neural network combined with grey theory[J]. South-to-North Water Transfers and Water Science & Technology, 2020,18(1):138-143.
- [16] 涂吉昌,陈超波,王景成,等. 基于深度学习的水质预测模型研究[J]. 自动化与仪表, 2019,34(6):96-100.
Tu Jichang, Chen Chaobo, Wang Jingcheng, et al. Research on water quality prediction model based on deep learning[J]. Automation & Instrumentation, 2019,34(6):96-100.
- [17] Liang Z Y, Zou R, Chen X, et al. Simulating the forecast capacity of a complicated water quality model using the long short-term memory approach[J]. Journal of Hydrology, 2020,581:124432.
- [18] 王勇,张新政,高向军. 时序规则挖掘[J]. 计算机工程, 2005, 31(23):61-62.
Wang Yong, Zhang Xinzheng, Gao Xiangjun. Time-series rules mining[J]. Computer Engineering, 2005,31(23):61-62.
- [19] 崔鸿雁,徐帅,张利锋,等. 机器学习中的特征选择方法研究及展望[J]. 北京邮电大学学报, 2018,41(1):1-12.
Cui Hongyan, Xu Shuai, Zhang Lifeng, et al. The key techniques and future vision of feature selection in machine learning[J]. Journal of Beijing University of Posts and Telecommunications, 2018,41(1):1-12.
- [20] Kuhn M, Johnson K. Applied Predictive Modeling[M]. 1st ed., New York:Springer, 2013.
- [21] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997,9(8):1735-1780.
- [22] Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods [M]. Cambridge:Cambridge University Press, 2000.
- [23] 武国正,徐宗学,李畅游. 支持向量回归机在水质预测中的应用与验证[J]. 中国农村水利水电, 2012(1):25-29.
Wu Guozheng, Xu Zongxue, Li Changyou. Water quality prediction and validation by using support vector regression machines[J]. China Rural Water and Hydropower, 2012(1): 25-29.
- [24] 秦文虎,陈溪莹. 基于长短时记忆网络的水质预测模型研究[J]. 安全与环境学报, 2020,20(1):328-334.
Qin Wenhui, Chen Xiying. Water quality forecast and prediction model based on long- and short-term memory network [J]. Journal of Safety and Environment, 2020,20(1):328-334.