文章编号:1005-0523(2022)05-0078-08

基于自相关分析的交通流预测输入步长研究

王 爽,黄海超,石宝存,陈景雅

(河海大学土木与交通学院,江苏 南京 210024)

摘要:目前的交通流预测研究中,输入步长主要取决于人为的选择,容易受到干扰,并且缺少从理论方面选择输入步长的方法。为了能够自适应地选取输入步长,基于交通流历史时间序列的自相关分析,以机器学习中典型的最小二乘支持向量机(LSSVM)、随机森林(RF)以及长短记忆(LSTM)3种算法进行多输入步长的交通流预测,探究以自相关系数值选取最佳输入步长的方法可行性。实验结果表明,在输入步长的自相关系数为0.80~0.91时,LSSVM能获得较优的预测精度,当自相关系数为0.47~0.51时,LSTM能有较好的预测精度,而RF交通流预测的最低误差对应的输入步长自相关程度较低,自相关分析方法可能并不适用。

关键词:交通流预测;自相关分析;机器学习;输入步长

中图分类号:U293.5 **文献标志码**:A

本文引用格式:王爽,黄海超,石宝存,等. 基于自相关分析的交通流预测输入步长研究[J]. 华东交通大学学报,2022,39(5):78-85.

Research on Time Step of Traffic Flow Prediction Based on Autocorrelation Analysis

Wang Shuang, Huang Haichao, Shi Baocun, Chen Jingya

(College of Civil and Transportation Engineering, Hohai University, Nanjing 210024, China)

Abstract: In the current research of traffic flow prediction, the time step mainly depends on the artificial selection which is easy to be disturbed, and there is a lack of methods to select the time step from the theoretical aspect. In order to adaptively select the time step, three typical algorithms of least squares support vector machine (LSSVM), random forest(RF) and long short-term memory(LSTM) in machine learning are used to predict traffic flow with multiple time step based on the autocorrelation analysis of historical time series for traffic flow and explore the feasibility of selecting the best time step by the value of autocorrelation coefficient. The experimental results show that when the autocorrelation coefficient of the time step is 0.83~0.91, LSSVM can obtain better prediction accuracy while the autocorrelation coefficient is 0.47~0.51, LSTM can have better prediction accuracy. However, due to the low autocorrelation degree of the time step corresponding to the lowest error of traffic flow prediction, the autocorrelation analysis method may not be applicable to RF.

Key words: traffic flow prediction; autocorrelation analysis; machine learning; time step

Citation format: WANG S, HUANG H C, SHI B C, et al. Research on time step of traffic flow prediction based on autocorrelation analysis[J]. Journal of East China Jiaotong University, 2022, 39(5): 78-85.

收稿日期:2022-03-11

基金项目:国家自然科学基金项目(52078190);教育部人文社会科学研究规划基金项目(18YJAZH119)

随着城市的发展,汽车保有量急剧增加,现有的城市道路交通难以满足日益增长的交通需求,两者之间的供需不平衡导致了交通拥堵、环境污染、交通事故等一系列问题,为了解决这些交通问题,智能交通系统(intelligent transportation system,ITS)应运而生[1]。交通流预测作为 ITS 的核心内容,通过计算机技术、大数据挖掘等提前对城市道路交通流及其状态进行预测,从而有助于交通管理者进行车流的合理规划,达到缓解交通拥堵和提高道路通行能力等目的。

目前的交通流时间序列预测模型主要分为基于数理统计的传统预测模型和基于机器学习的预测模型,通过对历史时间序列进行建模以及特征学习,预测未来时刻的交通流量。自回归积分滑动模型(autoregressive integrated moving average model,ARIMA)[2] 作为典型的基于数理统计的交通流预测模型,在线性数据处理上具有优势,但在大规模非线性的交通数据处理上存在不足。Lu 等[3]将 ARIMA 模型与 LSTM(long short-term memory,LSTM)相结合,分别利用 ARIMA 处理线性数据的优势以及 LSTM 处理非线性数据的优势进行交通流预测,验证了该组合模型的多功能性。随着人工智能和大数据的发展,机器学习模型在处理非线性交通数据上的能力,使其逐渐成为研究者们青睐的交通流预测模型。机器学习模型中,常用的交通流预测模型有支持向量机(support vector machines,SVM)[4]、随机森林(random forests,RF)[5]、人工神经网络(artificial neural network,ANN)[6]。SVM 模型在 1995 年被 Cortes 和 Vapnik 首次提出,童林等[7],闫贺等[8]对 SVM 进行改进,提高了短时交通流预测的精度。针对 SVM 求解二次规划问题时训练时间过长等问题,Suyken 等提出了最小二乘支持向量机(least square SVM, LSSVM),有效地减短了运行时长。谷远利等[9]通过免疫算法优化了 LSSVM 的超参数,优化模型训练速度的同时,提高了预测精度。戴丽珍等[10]通过灰狼算法对 LSSVM 的参数进行优化,增强了模型的泛化能力。Hamner[11],熊亭等[12]采用了 RF 进行交通流预测并取得了良好的效果。LSTM 具有捕捉非线性数据中长期依赖的能力,常被广泛运用在交通流时间序列预测中。2015 年,Ma 等[13]使用 LSTM 对速度进行了预测,并证明了相比较于 RNN,SVM,ARIMA 模型,LSTM 在捕捉交通流数据非线性特征以及长时间依赖的优势。

交通流预测的研究主要集中于模型优化,针对输入步长的讨论较少。林培群等[14]在进行高速公路的行程时间预测时,选择 12 步步长作为不同预测模型的输入,研究预测步长增加后不同模型的预测性能。王博文等[15]针对 LSTM 预测步长增加误差迅速累积的问题对模型进行改进,选取了 8 步作为模型输入对未来的 1~12 步的时刻进行预测。Huang 等[16]为了验证提出的时间序列分解方法和相应混合预测模型的有效性,选用 12 步作为输入步长进行了多步的交通流预测。高华睿等[17]在双向长短时记忆网络预测交通流时,采用了滑动窗口模型,以 15 步步长作为滑动窗口长度对交通流数据进行重采样,学习交通流数据随时间变化的趋势变化特征。金玮[18]利用傅里叶级数提取交通流量的周期性分量特征,分离残差项,选取 1、3、6、9、12 作为输入步长分别进行预测,分析预测结果用来判断模型基于周期性分量的预测效果。熊亭等[19]将扩散卷积算子应用于 GRU(gate recurrent unit)模型进行交通流预测,通过分析 1~8 步输入步长对 MAPE 值的影响,选择 5步输入步长进行特征学习和模型训练。

综上所述,在选择输入步长的方法方面研究较少,大多直接选择较小的固定步长作为模型的输入,也有研究通过一些递减的规律来选择输入步长,但这种方法仍依赖于人为的选择。为了能够自适应地选取输入步长,研究基于交通流时间序列的自相关分析,将输入步长与本身的数据特征相结合,采用机器学习模型中常用的 LSSVM,RF,LSTM 3 种模型,探究自相关系数与最佳输入步长之间的关系,最终提出一种基于时间序列分析能够自适应选取最佳输入步长的方法。

1　相关理论

1.1　自相关分析

在时间序列分析中,自相关函数常用来分析时间序列与其自身 k 阶滞后序列之间的相关关系,以此度量历史数据对 t 时刻产生的影响。在交通流数据的自相关分析中,可以有效地识别时间序列中的周期性特征,也可以通过相关程度粗略的判断输入数据与预测数据之间的关系[20]。其中,以 $k=0,1,2,\cdots,n$ 表示时间序列的滞后阶数,给定

目标预测路段的时间序列$\{y_t\}$及滞后序列$\{y_{t+k}\}$,自相关函数定义[21]如下

$$\rho_k=\frac{\mathrm{cov}(y_t,y_{t+k})}{\sqrt{\sigma_{y_t}^{\ 2}}\sqrt{\sigma_{y_{t+k}}^{\ 2}}} \tag{1}$$

式中:k为滞后步长;$\mathrm{cov}(y_t,y_{t+k})$为$y_t$与$y_{t+k}$的协方差;$\sigma_{y_t}$为$y_t$的均方差。

1.2 最小二乘支持向量机

最小二乘支持向量机作为支持向量机的改进模型,不仅可以有效的处理强非线性数据,而且通过简化求解二次规划问题减少了训练时间,极大地增强了LSSVM的适用性。与SVM相同,LSSVM将交通流数据通过非线性变换映射到高维空间[22-23]

$$f(\boldsymbol{x})=\boldsymbol{\omega}^{\mathrm{T}}\phi(\boldsymbol{x})+c \tag{2}$$

式中:$\boldsymbol{\omega}$为超平面的权值;c为常数;$\phi(\cdot)$为空间转换函数。

最小二乘支持向量机的目标函数定义如下

$$\begin{cases}\min J=\dfrac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}+\gamma\sum\limits_{i=1}^{l}e_i^2\\ \text{s.t. } y_i=\boldsymbol{\omega}^{\mathrm{T}}\phi(\boldsymbol{x})+c+e_i\end{cases} \tag{3}$$

式中:e_i为误差;γ为惩罚因子。

构造拉格朗日函数如下

$$L(\boldsymbol{\omega},b,\boldsymbol{e},a_i)=\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}+\gamma\sum_{i=1}^{l}e_i^2-\sum_{i=1}^{l}a_i(\boldsymbol{\omega}^{\mathrm{T}}\phi(\boldsymbol{x})+c-y_i+e_i) \tag{4}$$

式中:a_i为拉格朗日乘子。根据KKT条件可得

$$\begin{cases}\dfrac{\partial L}{\partial\boldsymbol{\omega}}=0\rightarrow\sum\limits_{i=1}^{l}a_i\phi(x_i)=0\\ \dfrac{\partial L}{\partial c}=0\rightarrow\sum\limits_{i=1}^{l}a_i=0\\ \dfrac{\partial L}{\partial e_i}=0\rightarrow a_i+\gamma e_i=0\\ \dfrac{\partial L}{\partial a_i}=0\rightarrow\boldsymbol{\omega}^{\mathrm{T}}\phi(x_i)+c-y_i+e_i=0\end{cases} \tag{5}$$

求解式(5),可得LSSVM数学模型

$$f(x)=\sum_{i=1}^{l}a_iK(x,x_i)+c \tag{6}$$

式中:$K(\cdot)$为模型核函数,通常采用RBF核函数,即

$$K_{\mathrm{RBF}}=\exp\left(-\frac{\|x-x_i\|^2}{2\varphi^2}\right) \tag{7}$$

式中:φ为核函数参数。

1.3 随机森林

随机森林是一种有监督的Bagging(bootstrap aggregating)类集成学习方法,具有强大的数据挖掘和学习能力,在众多领域均有广泛的应用。在交通流预测研究中,主要采用随机森林的回归算法,由Bootstrap法重采样随机形成决策树,大量相互独立的决策树组成森林,将多个决策树预测结果取平均值得到最终的结果,收敛速度快并且具有较高的精确性。而且随机森林能很好的处理交通流数据中的噪声和异常值,数据随机以及特征随机的特点能够避免训练数据出现过拟合的现象,在大规模的交通流数据中依然可以有效地运行[24]。

1.4 长短记忆神经网络

长短记忆神经网络是一种递归神经网络结构,作为循环神经网络(recurrent neural network,RNN)的一个重要变体[25],它解决了RNN中存在的长期依赖和梯度弥散的问题,通过记忆细胞长时间记忆某个值,可以更好地捕捉长时间步中时间序列中的规律。典型的LSTM在RNN的基础上引入了输入门、遗忘门、输出门来控制信息传递与丢弃。其中,输入门获取输入数据并对其进行更新,遗忘门给记忆细胞选择权是否维持旧的值,输出门计算并输出结果。

本文将输入时间序列表示为$\boldsymbol{X}=[x_1,x_2,\cdots,x_n]$,对输入$t$时刻数据进行更新

$$\tilde{c}_t=\tan h(\boldsymbol{W}_c[a_{t-1},x_t]+b_c) \tag{8}$$

通过更新门和遗忘门对记忆细胞中存储的值进行更新为

$$\Gamma_{\mathrm{u}}=\boldsymbol{W}_{\mathrm{u}}[a_{t-1},x_t]+b_{\mathrm{u}} \tag{9}$$

$$\Gamma_{\mathrm{f}}=\varepsilon(\boldsymbol{W}_{\mathrm{f}}[a_{t-1},x_t]+b_{\mathrm{f}}) \tag{10}$$

输出门输出结果为

$$\Gamma_{\mathrm{o}}=\varepsilon(\boldsymbol{W}_{\mathrm{o}}[a_{t-1},x_t]+b_{\mathrm{o}}) \tag{11}$$

$$c_t=\Gamma_{\mathrm{u}}\tilde{c}_t+\Gamma_f\tilde{c}_{t-1} \tag{12}$$

$$a_t=\Gamma_{\mathrm{o}}c_t \tag{13}$$

式中:x_t为t时刻的输入;$\tilde{c}_t$为t时刻的候选记忆细胞;c_t为t时刻记忆细胞;a_{t-1}为$t-1$时刻的隐藏状态;$\Gamma_{\mathrm{u}},\Gamma_{\mathrm{f}},\Gamma_{\mathrm{o}}$分别为$t$时刻的更新门、遗忘门、输出门的输出;$\boldsymbol{W}_c,\boldsymbol{W}_{\mathrm{u}},\boldsymbol{W}_{\mathrm{f}},\boldsymbol{W}_{\mathrm{o}}$分别为各门的权重参数;$b_c,b_{\mathrm{u}},b_{\mathrm{f}},b_{\mathrm{o}}$为偏置参数;$\tan h,\varepsilon$为激活函数。

1.5 交通流预测方法

LSSVM,RF,LSTM分别作为机器学习模型中二分类、集成学习、神经网络中常用的预测模型,选用这3种模型研究通过自相关分析选取最佳输入步长的方法。将车流量时间序列作为预测模型的输

入,以不同的输入步长分别进行训练,预测未来 15 min 的车流量,取 5 次预测误差的均值作为最终的预测结果进行分析。

输入步长的长度以 n 来表示,交通流预测的整体流程图如图 1 所示。

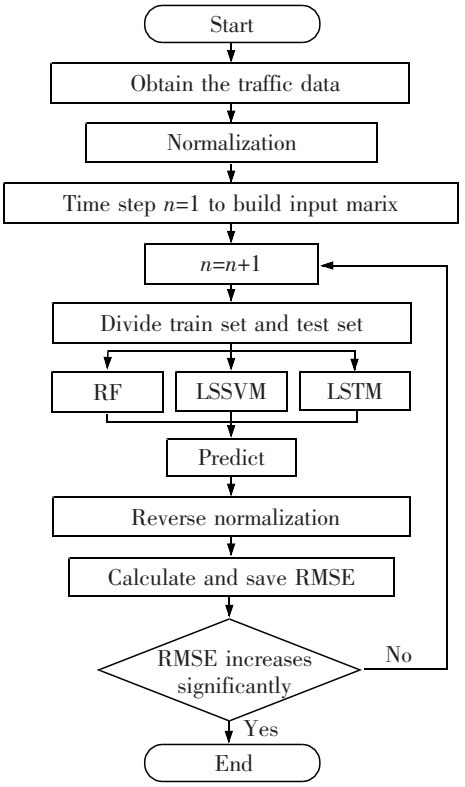


图 1　交通流预测流程图

Fig.1　Flow chart of traffic flow prediction

2　实例应用

2.1　数据来源及预处理

数据集来源于英国曼彻斯特郡的某个路段的车流量数据，选取编号为 A9094、B26053、B9117、A9116、A9070、A5103 的 6 个观测点作为研究对象,为了方便说明,将 6 个观测点的编号分别简化为序号 1~6。每个观测点数据集的时间跨度为 2020 年 1 月 6 日—2020 年 2 月 9 日,共 35 d,每组数据的采样时间间隔为 15 min,样本总量为 20 160。

为了能够获得较好的拟合效果,提高模型的收敛速度,本文采用 Min-Max 函数对数据进行归一化处理。按 8:2 的比例划分训练集和测试集，其中 2020 年 1 月 6 日—2020 年 2 月 2 日共 28 d 作为训练集,2020 年 2 月 3 日—2020 年 2 月 9 日共 7 d 的数据为测试集。

2.2　模型参数设置及评价指标选取

参数寻优过程保存在网站 https://github.com/Team-Cheng/ACF-parameter-optimization-.git 中,寻找各模型的局部最优参数,具体设置如下。LSSVM 模型选用 RBF 核函数,模型的性能主要取决于 γ 和 σ 两个参数，每个观测点的 γ 和 σ 分别设置为 10、0.1。RF 模型中主要确定最佳叶子数以及树数,在模型训练中每个观测点设置了 200 棵决策树,每棵决策树的最佳叶子数为 20。LSTM 模型采用 2 层 LSTM 结构,2 层隐藏层的节点数分别设为 32 和 16，添加了 Dropout 层来增加模型的泛化能力,设置丢弃概率为 0.1。选用 ReLU 作为激活函数，优化函数选用 Adam 优化函数,初始学习率设置为 0.005,进行 100 轮迭代训练,梯度阈值设置为 1 以防止梯度爆炸。

采用均方根误差(root mean square error, RMSE)作为指标对预测结果的精确度进行评价,它通过计算预测值和真实值之间的标准差来反映预测结果的准确性,计算式为

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{14}$$

式中:y_i 为目标预测路段的实际交通流量值;$\hat{y}_i$ 为目标预测路段的预测交通流量值。

2.3　实验结果与分析

首先,确定 3 个交通流预测模型输入步长的阈值。随机选取一个观测点,以输入步长 n=1,2,3,…,i 分别进行预测，观察 RMSE 值的趋势变化。3 个模型不同输入步长的 RMSE 值如图 2 所示。从对 RMSE 进行拟合的曲线图中可以看出，随着输入步长的增加,LSSVM 模型的预测误差呈现出先减小后逐渐增大的趋势;LSTM 预测结果波动较大,但总体而言,呈逐渐增加的趋势,在 48 步之后,有明显增加的趋势；而 RF 模型的 RMSE 值随着输入步长的增加呈逐渐降低的趋势。综合 LSSVM 以及 LSTM 模型的预测结果分析,选择 48 步输入步长作为上限。

2.3.1　LSSVM 预测结果

不同输入步长下 LSSVM 的预测误差以及自相关系数值如图 3 所示。从对 RMSE 进行拟合的曲线图中可以看出，随着输入步长的增加,LSSVM 的预测误差呈先减小后增大的趋势,当输入步长超过 10 步后,RMSE 值增加迅速，输入步长对 LSSVM 模型的预测结果影响较大。结合 6 个观测点交通流时间序列的自相关分析,当输入步长的自相关系数值在

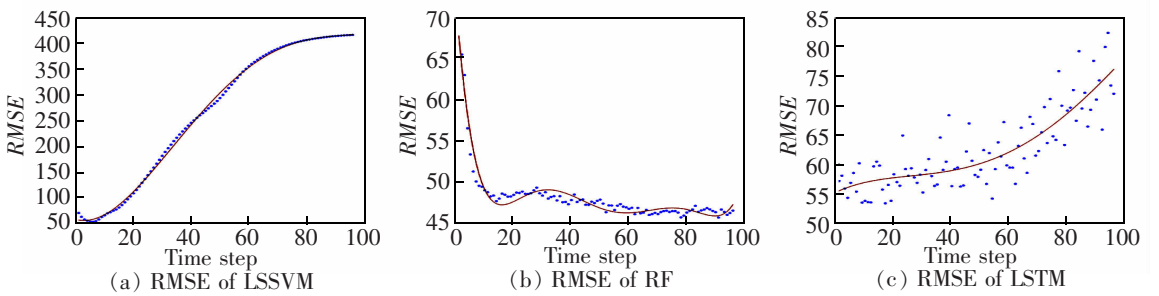


图 2 各模型不同输入步长的 RMSE

Fig.2 RMSE of different time step of each model

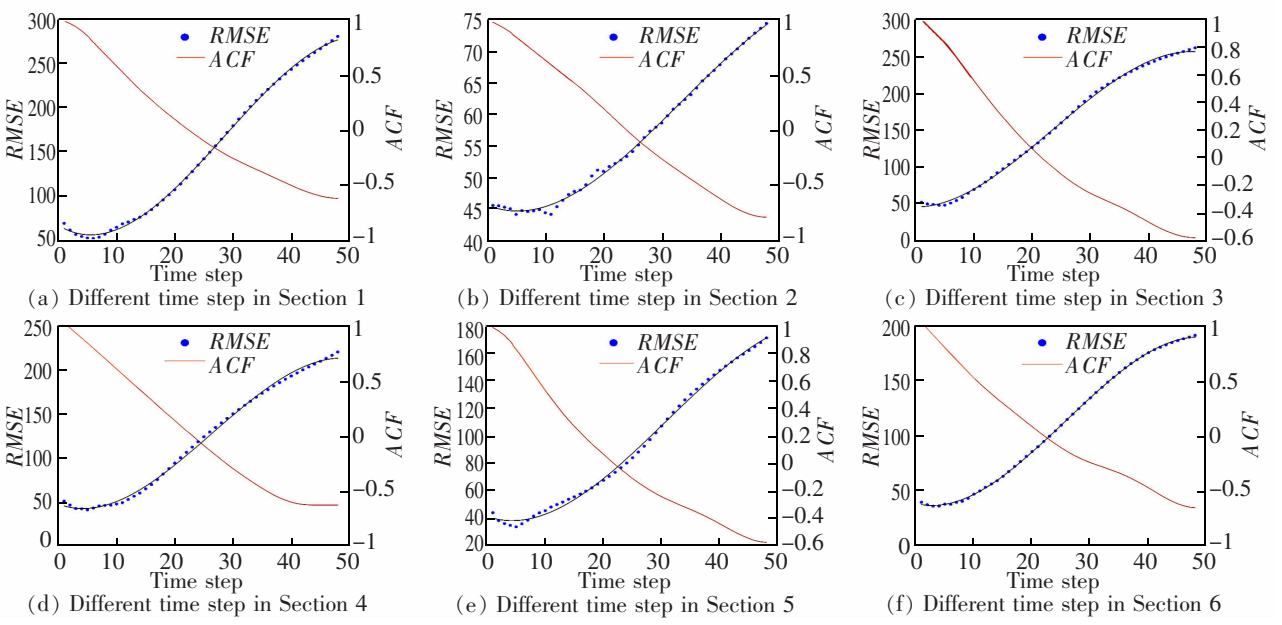


图 3 LSSVM 不同输入步长 RMSE 及 ACF

Fig.3 RMSE and ACF of different time step of LSSVM

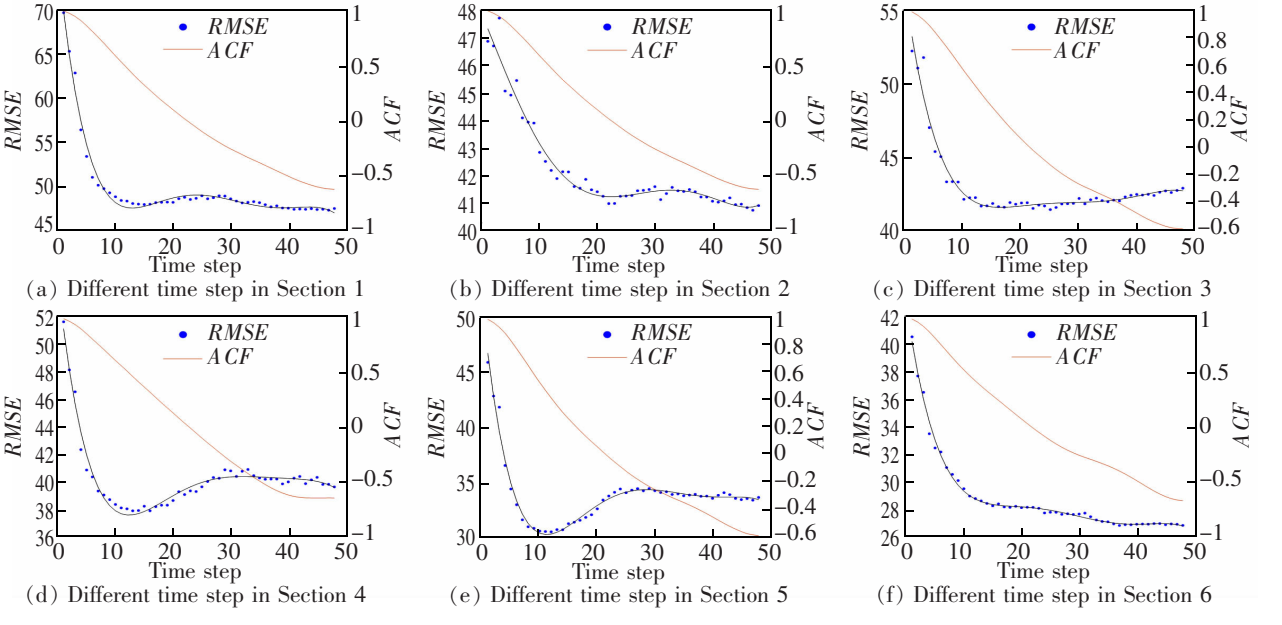


图 4 RF 不同输入步长 RMSE 及 ACF

Fig.4 RMSE and ACF of different time step of RF

0.80~0.91 时,LSSVM 会有最小的 RMSE 值。

2.3.2　RF 预测结果

不同输入步长下 RF 模型的 RMSE 值以及自相关系数值如图 4 所示。从对 RMSE 进行拟合的曲线图可以看出,随着输入步长的增加,RF 模型的 RMSE 值逐渐减小,有可能存在最小的预测误差,也有可能随着输入步长的增加误差继续减小。原因可能是 RF 模型在训练过程中,输入步长跨度内的交通流数据作为 RF 中特征的一部分,可能受到特征随机性选择或者特征数量对模型预测的影响。而综合 6 个观测点交通流数据的自相关分析,难以找到相同的规律,通过时间序列的自相关分析寻找预测误差较小的输入步长的方法可能并不适用于 RF 模型。

2.3.3　LSTM 预测结果

基于 6 个观测点的车流量数据,对各观测点从 1 步到 48 步的不同输入步长分别进行预测,取 5 次 RMSE 平均值进行结果分析。LSTM 模型各观测点不同输入步长 RMSE 以及相应的自相关系数值如图 5 所示。从对 RMSE 进行拟合的曲线图可以看出,LSTM 模型的预测结果相比较于 LSSVM 以及 RF 模型,波动程度较大。但从 RMSE 值曲线整体而言,LSTM 预测模型的误差会先呈下降趋势,之后随着步长的增加,RMSE 值呈增加趋势。综合观察 6 个观测点达到最小误差的输入步长及相应的自相关系数值,可以发现,当输入步长的自相关系数在 0.47~0.51 之间时,LSTM 可以获得较好的拟合效果和较好的精度。

2.3.4　结果分析

通过对交通流时间序列进行自相关分析,结合 3 种模型不同输入步长的预测结果,3 种模型最佳输入步长的自相关系数值如表 1 所示。其中,LSSVM 以及 LSTM 最佳输入步长的自相关系数有

表 1　各模型最佳输入步长及自相关系数值

Tab.1　Optimal time step and autocorrelation coefficient of each model

Model	Number	Input step	*RMSE*	*ACF*
LSSVM	1	6	51.57	0.80
	2	5	44.03	0.85
	3	4	47.10	0.88
	4	5	40.29	0.85
	5	5	33.50	0.83
	6	3	35.56	0.91
RF	1	–	–	–
	2	–	–	–
	3	25	41.44	0.18
	4	13	38.00	0.46
	5	12	30.61	0.42
	6	–	–	–
LSTM	1	12	53.72	0.47
	2	13	43.24	0.51
	3	11	46.95	0.50
	4	12	42.50	0.51
	5	11	32.47	0.47
	6	11	31.96	0.47

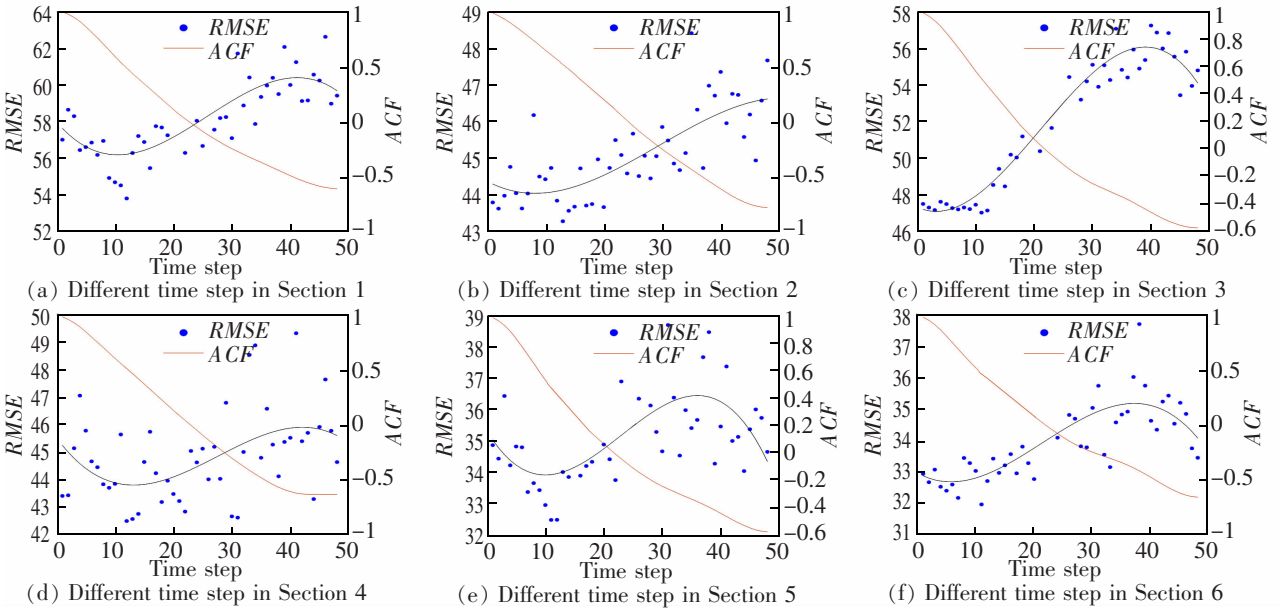


图 5　LSTM 不同输入步长 RMSE 及 ACF

Fig.5　RMSE and ACF of different time step of LSTM

一定的规律，通过时间序列的自相关分析，LSSVM可以选取自相关系数在0.80~0.91的输入步长，LSTM则可以选取自相关系数在0.47~0.51的输入步长，而RF预测模型中精度较好的输入步长与其相应的自相关系数之间的规律难以找到，数据自相关分析的方法可能并不适用。

3 结论

基于交通流时间序列的自相关分析，选用机器学习中常用的LSSVM、RF、LSTM模型进行预测，提出一种自适应选取最佳输入步长的方法。

1）当采用LSSVM模型时，建议选用自相关系数值在0.80~0.91的输入步长，RMSE值可能有最小值，预测精度较高。

2）采用LSTM模型进行交通流预测时，可以选取自相关系数值在0.47~0.51的输入步长，预测误差较低。

3）选用RF模型进行交通流预测时，预测误差最小的输入步长较大或者自相关程度较低，通过时间序列的自相关分析选择输入步长的方法可能并不适用，针对RF模型的最佳输入步长选择方法需要进一步的研究。

参考文献：

[1] 赵娜，袁家斌，徐晗. 智能交通系统综述[J]. 计算机科学，2014，41(11)：7-11.
ZHAO N，YUAN N，XU H. Overview of intelligent transportation system[J]. Computer Science，2014，41(11)：7-11.

[2] WILLIAMS B M，HOEL L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process：Theoretical basis and empirical results[J]. Journal of transportation engineering，2003，129(6)：664-672.

[3] LU S Q，ZHANG Q Y，CHEN G S，et al. A combined method for short-term traffic flow prediction based on recurrent neural network[J]. Alexandria Engineering Journal，2021，60(1)：87-94.

[4] DINH T T，HUNG T V. Support vector machine for shortterm traffic flow prediction and improvement of its model training using nearest neighbor approach[J]. Transportation Research Record，2021，2675(4)：362-373.

[5] OU J S，XIA J X，WU Y J，et al. Shortterm traffic flow forecasting for urban roads using datadriven feature selection strategy and biascorrected random forests[J]. Transportation Research Record，2017，2645(1)：157-167.

[6] MA C X，TAN L M，XU X C. Shortterm traffic flow prediction based on genetic artificial neural network and exponential smoothing[J]. Promet-Traffic & Transportation，2020，32(6)：747-760.

[7] 童林，官铮. 改进鲸鱼优化支持向量机的交通流量模糊粒化预测[J]. 计算机应用，2021，41(10)：2919-2927.
TONG L，GUAN Z. Fuzzy granulation prediction of traffic flow based on improved whale optimization support vector machine[J]. Journal of Computer Applications，2021，41(10)：2919-2927.

[8] 闫贺，朱丽，戚湧. 基于改进SVM的城市道路短时交通状态预测[J]. 南京师大学报(自然科学版)，2019，42(3)：129-137.
YAN H，ZHU L，QI Y. Shortterm traffic condition prediction of urban road based on improved SVM[J]. Journal of Nanjing Normal University(Nature Science Edition)，2019，42(3)：129-137.

[9] 谷远利，张源，芮小平，等. 基于免疫算法优化LSSVM的短时交通流预测[J]. 吉林大学学报(工学版)，2019，49(6)：1852-1857.
GU Y L，ZHANG Y，RUI X P，et al. Shortterm traffic flow prediction based on LSSVM optimized by immune algorithm[J]. Journal of Jilin University(Engineering and Technology Edition)，2019，49(6)：1852-1857.

[10] 戴丽珍，付涛，杨刚，等. 一种改进灰狼算法优化LSSVM的交通流量预测[J]. 小型微型计算机系统，2020，41(12)：2672-2676.
DAI L Z，FU T，YANG G，et al. Traffic flow prediction method by LSSVM optimized by an improved GWO[J]. Journal of Chinese Computer Systems，2020，41(12)：2672-2676.

[11] HAMNER B. Predicting travel times with contextdependent random forests by modeling local and aggregate traffic flow[C]//Sydney：Proceedings of 2010 IEEE International Conference on Data Mining Workshops，2010.

[12] 熊亭，戚湧，张伟斌，等. 基于时空相关性的短时交通流预测模型[J]. 计算机工程与设计，2019，40(2)：501-507.
XIONG T，QI Y，ZHANG W B，et al. Shortterm traffic flow forecasting model based on temporalspatial correlation[J]. Computer Engineering and Design，2019，40(2)：501-507.

[13] MA X L，TAO Z M，WANG Y H，et al. Long shortterm memory neural network for traffic speed prediction using remote microwave sensor data[J]. Transportation Research Part C：Emerging Technologies，2015，54：187-197.

[14] 林培群，夏雨，周楚昊. 引入时空特征的高速公路行程时

间预测方法[J]. 华南理工大学学报(自然科学版),2021,49(8):1-11.
LIN P Q,XIA Y,ZHOU C H. Freeway travel time prediction based on spatial and temporal characteristics of road networks[J]. Journal of South China University of Technology (Nature Science Edition),2021,49(8):1-11.

[15] 王博文,王景升,王统一,等. 基于长短时记忆神经网络的 Encoder-Decoder 多步交通流预测模型[J]. 重庆大学学报,2021,44(11):71-80.
WANG B W,WANG J S,WANG T Y,et al. An Encoder-Decoder multistep traffic flow prediction model based on long short-time memory network[J]. Journal of Chongqing University,2021,44(11):71-80.

[16] HUANG H C,CHEN J Y,SUN R,et al. Shortterm traffic prediction based on time series decomposition[J]. Physica A:Statistical Mechanics and its Applications,2022,585:126441.

[17] 高华睿,郝龙,王明明,等. 基于 Att-Bi-LSTM 的高速公路短时交通流预测研究[J]. 武汉理工大学学报,2020,42(9):59-64.
GAO H R,HAO L,WANG M M,et al. Research on Short-term traffic flow forecast of expressway based on Att-Bi-LSTM[J]. Journal of Wuhan University of Technology,2020,42(9):59-64.

[18] 金玮. 基于周期性分量提取的城市快速路交通流短时预测理论与方法研究[D]. 北京:北京交通大学,2018.
JIN W. Theoretical research for shortterm traffic flow prediction on urban expressway based on periodic component extraction[D]. Beijing:Beijing Jiatong University,2018.

[19] 熊亭,戚湧,张伟斌. 基于 DCGRU-RF 模型的路网短时交通流预测[J]. 计算机科学,2020,47(5):84-89.
XIONG T,QI Y,ZHANG W B. Shortterm traffic flow prediction based on DCGRU-RF model for road network [J]. Computer Science,2020,47(5):84-89.

[20] 刑珊珊. 基于时空关联的城市快速路短时交通流预测方法研究[D]. 北京:北京交通大学,2018.
XING S S. Study on shortterm traffic flow forecasting method of urban expressway based on spatialtemporal correlation[D]. Beijing:Beijing Jiaotong University,2018.

[21] WU C L,HE H D,SONG R F,et al. Prediction of air pollutants on roadside of the elevated roads with combination of pollutants periodicity and deep learning method[J]. Building and Environment,2022,207(Part B):108436.

[22] 彭延峰,彭志华,刘燕飞. 基于 ASNBD-CMFE 特征信息提取的短时交通流预测[J]. 北京交通大学学报,2021,45(2):28-35.
PENG Y F,PENG Z H,LIU Y F. Shortterm traffic flow prediction based on the feature information extraction applying ASNBD-CMFE[J]. Journal of Beijing Jiaotong University,2021,45(2):28-35.

[23] 胡松,成卫,李艾. 一种改进鲸鱼算法及其在短时交通流预测中的应用研究[J]. 小型微型计算机系统,2021,42(8):1627-1632.
HU S,CHENG W,LI A. Improved whale algorithm and its application in short-term traffic flow prediction[J]. Journal of Chinese Computer Systems,2021,42(8):1627-1632.

[24] 吕红燕,冯倩. 随机森林算法研究综述[J]. 河北省科学院学报,2019,36(3):37-41.
LYU H Y,FENG Q. A review of random forests algorithm [J]. Journal of the Hebei Academy of Sciences,2019,36(3):37-41.

[25] 黄海超,陈景雅,王爽,等. 多因素轨道交通客流量预测模型研究[J]. 华东交通大学学报,2021,38(3):61-66.
HUANG H C,CHEN J Y,WANG S,et al. Multifactor rail transit passenger flow prediction model[J]. Journal of East China Jiaotong University,2021,38(3):61-66.



第一作者：王爽(1998—)，女，硕士研究生，研究方向为机器学习在智能交通领域的应用。E-mail:201304060010@hhu.edu.cn。



通信作者：陈景雅(1967—)，女，教授，硕士研究生导师，研究方向为互通立交安全评估。E-mail:13805151397@139.com。

(责任编辑：吴海燕　姜红贵)