

停车场泊位占有率预测方法评价

唐克双, 郝兆康, 衣谢博闻, 刘冰清

(同济大学 道路与交通工程教育部重点实验室, 上海 201804)

摘要: 采用上海市五角场地区的停车泊位检测数据, 分析了商业、办公和体育场 3 种不同类型停车场泊位占有率 (parking occupancy rate, POR) 的时变特征, 并评价了 ARIMA (autoregressive integrated moving average)、卡尔曼滤波和 BP (back propagation) 神经网络等 3 种常用方法在 POR 预测中的适用性。结果表明, ARIMA 和 BP 神经网络的预测精度总体优于卡尔曼滤波, BP 神经网络在商业和办公停车场的短时预测中有较好的精度; 3 种方法的预测精度均随预测时间步长的增加而逐渐降低; 不同类型停车场的 POR 预测精度存在较大差异, 工作日的预测精度一般高于非工作日, 且模型具有较好的自适应性。

关键词: 停车泊位占有率预测; ARIMA 模型; 卡尔曼滤波; BP 神经网络模型

中图分类号: U491.1

文献标志码: A

Evaluation of Prediction Methods for Parking Occupancy Rate

TANG Keshuang, HAO Zhaokang, YIXIE Bowen, LIU Bingqing

(Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China)

Abstract: This paper analyzed temporary characteristics of parking occupancy rate (POR) at three different types of parking lots, i. e., shopping mall, office building and stadium, and evaluated the applicability of autoregressive integrated moving average method (ARIMA), Kalman filter and BP neural networks on the prediction of POR, based on parking lot detection data at Wujiaochang District, Shanghai. The results show that ARIMA and BP neural networks can achieve higher prediction accuracies as compared with the Kalman filter method, and the BP neural networks performs best for the short-term prediction of shopping mall and office building. The prediction accuracy of the three methods decreases as the forecasting time step increases. Different prediction accuracies exist for different types of parking lots,

and the prediction accuracy for weekdays is higher than that for weekends. And the model has good adaptability. This paper can provide reference for the selection of prediction methods for different types of parking lots.

Key words: parking occupancy rate prediction; autoregressive integrated moving average model (ARIMA); Kalman filter; back propagation neural networks

近年来, 停车难已成为各国城市特别是大都市亟待解决的交通问题。调查表明, 城市路网中有 30% 的交通量是由寻找停车位造成的^[1]。停车诱导系统是现阶段解决停车难的重要手段^[2]。目前我国停车诱导系统的工作方式主要是通过停车诱导屏发布实时的空余停车位信息^[3]。但驾驶员从看到诱导屏上的信息到行驶至停车场需要花费一定的时间, 如果诱导屏只发布实时的可用泊位数, 可能会导致驾驶员到达停车场时的可用泊位数和在诱导屏上看到的有差异, 驾驶员可能要重新寻找新的停车场, 从而降低驾驶员对诱导系统的信任。为了提高停车诱导系统的可靠性, 诱导屏应发布一定时间后的可用停车位信息^[4-5], 避免或减少上述情况的发生, 减少不必要的绕行^[6]。

因此, 停车可用泊位数预测的准确性对于停车诱导系统的运行效果很重要。然而, 准确的停车泊位预测并不是一件简单的事情。首先, 不同类型停车场的停车需求特征和供应水平不同, 导致停车位的时间变化趋势有差异, 对所有类型的停车场采用同一种方法进行预测可能会造成较大的误差^[7-8]; 其次, 从驾驶员看到诱导屏到行驶至停车场的时随道路交通状态的波动而动态变化, 不同的预测时间跨度会对预测精度产生较大的影响^[6]; 另外, 不同的停车泊位预测方法对可用车位变化时间序列特征的描述存在各自的优势和劣势, 具有不同的适应性^[4,9]。

收稿日期: 2016-04-30

基金项目: “十二五”国家科技支撑计划(2014BAG03B02)

第一作者: 唐克双(1980—), 男, 副教授, 博士生导师, 工学博士, 主要研究方向为智能交通系统。E-mail: tang@tongji.edu.cn

针对以上问题,本文采用上海市五角场区域 2014 年 5—9 月的停车泊位检测数据对常用的停车场泊位占有率预测方法进行了系统评价.首先分析了商业、办公和体育场 3 种类型停车场的泊位占有率(parking occupancy rate, POR)时变特征,然后分别采用自回归移动平均(autoressive integrated moving average, ARIMA)、卡尔曼滤波(Kalman filter)和 BP 神经网络模型(back propagation neural networks, BPNN)对 POR 进行预测,对比分析了 5, 10, 20 和 30 min 等不同预测时间跨度下的预测精度.本研究的结论可为不同类型停车场泊位预测方法的选择提供参考.

1 文献综述

国内外已经有许多学者对停车场泊位预测问题进行了研究,其中常用的预测方法可以分为两类:一类是传统的时间序列预测方法,如指数平滑法、ARIMA 模型、马尔科夫法、卡尔曼滤波等;另一类是一些启发式算法,如神经网络模型、遗传算法、模拟退火算法、粒子滤波算法等^[10].

传统的时间序列预测方法能揭示停车位变化趋势的规律,可以根据这些规律,利用历史或实时停车位数据预测未来的空余停车位.例如, Dunning^[11]采用 ARIMA 模型预测停车可用泊位数,发现在停车场到达率较低时预测精度很好,但当停车场到达率增大时,预测精度降低. Liu 等^[6]采用混沌时间序列中的加权一阶局域法对医院附近的停车可用泊位数进行预测,发现预测精度会随着预测时间跨度的增加而逐渐降低. Tamrazian 等^[5]采用了离线和在线预测两种时间序列预测方法对校园附近停车场的泊位占有率进行预测,发现在线预测方法的精度更高. Rajabioun 等^[12]考虑了某区域中各停车场的泊位变化规律和时空相关关系,采用多元自回归模型对该区域内停车场的可用泊位数进行预测. Okutani 等^[13]和 Ji 等^[14]采用卡尔曼滤波方法对交通状态进行预测,但将其用于停车场泊位预测的研究较少.

启发式算法指那些具有自学习功能,可利用部分信息对计算产生推理的算法,包括一些模仿自然界或生命体行为模式的算法.在停车泊位预测中,最常用的启发式算法是神经网络模型.例如 Yang 等^[15]和 Sun 等^[16]采用神经网络模型预测停车场可用泊位数,发现通过改变神经元个数、学习速率等可以优化神经网络的结构,从而有效提高神经网络模

型的预测精度. Vlahogianni 等^[17]采用神经网络模型对路边停车场泊位占有率进行了预测,并提出由于文中所用数据量较少,不能反映停车泊位占有率的月份和季节变化特征,更大数据量的使用会使结果更令人信服. Ji 等^[18]分别采用小波神经网络模型和最大 Lyapunov 指数法对停车可用泊位数进行预测,发现小波神经网络模型比后者有更好的预测效果.

综上所述,以往研究中常用的停车泊位时间序列预测方法和启发式算法主要有 ARIMA 模型和神经网络模型.同时,大多研究局限于用某一种方法对特定类型的停车泊位进行预测,针对不同方法在不同类型停车场的适用性评价研究很少.另外,虽然卡尔曼滤波被广泛应用于交通状态的预测,但将其用于停车场泊位预测的研究还较少.因此,本文的研究目的是评价上述 3 种方法针对不同的停车场类型、预测时间步长以及工作日/非工作日停车泊位占有率预测的适用性.

2 停车泊位占有率时变特征分析

为研究不同类型停车场泊位占有率的变化特性,选择了位于上海市百联又一城、创智天地和江湾体育场等 3 个停车场的泊位数据进行分析.这 3 个停车场均位于上海五角场城市副中心.其中,百联又一城是该地区的购物中心,创智天地是以信息产业为主的高新技术产业办公楼,江湾体育场是该区域的体育活动中心.这 3 个停车场可以分别代表城市商业、办公、体育场等 3 种停车场类型.本研究所用数据是上述 3 个停车场在 2014 年 5 月 1 日—9 月 7 日的 5 min 间隔停车泊位检测数据,具体包括停车场名称,总泊位数,可用泊位数,进、出车辆数,数据采集时刻等.

由于神经网络模型在应用前通常需要对数据进行量纲一化处理^[16],为方便各个方法之间的比较,故将原始可用泊位数数据结合停车场总泊位数转化为 POR 数据.以往的研究表明,停车场泊位变化在工作日和非工作日之间一般存在显著差异^[18],故将 POR 数据分为工作日和非工作日来讨论.

2.1 停车泊位占有率的日变化特征

图 1 展示了 3 个停车场典型工作日和非工作日的时变规律.由图 1 可知,商业、办公和体育场 3 种类型的 POR 日变化趋势和幅度差异很大,其中商业停车场经常出现停满情况,办公和体育场停车场则偶尔出现停满情况.体育场停车场会有少数曲线

波动较大,这是由于体育场有时举行重大活动,吸引大量车流到附近停车.对于每个停车场,工作日和非工作日的变化特征都有明显差别,如商业停车场在工作日有2个高峰期,在非工作日则只有1个高峰期,且非工作日高峰期中几乎一直处于停满状态.出现这种差别是因为人们在非工作日的出行购物行为

一般多于工作日,故非工作日商业停车场 POR 较高.由于停车场晚上关闭,3个停车场在 23:00 至次日 6:00 之间 POR 通常接近于 0 且保持不变,为了减少数据存储量和运算量,本文只分析停车场 1 d 内 6:00~23:00 时段的停车泊位占有率数据.

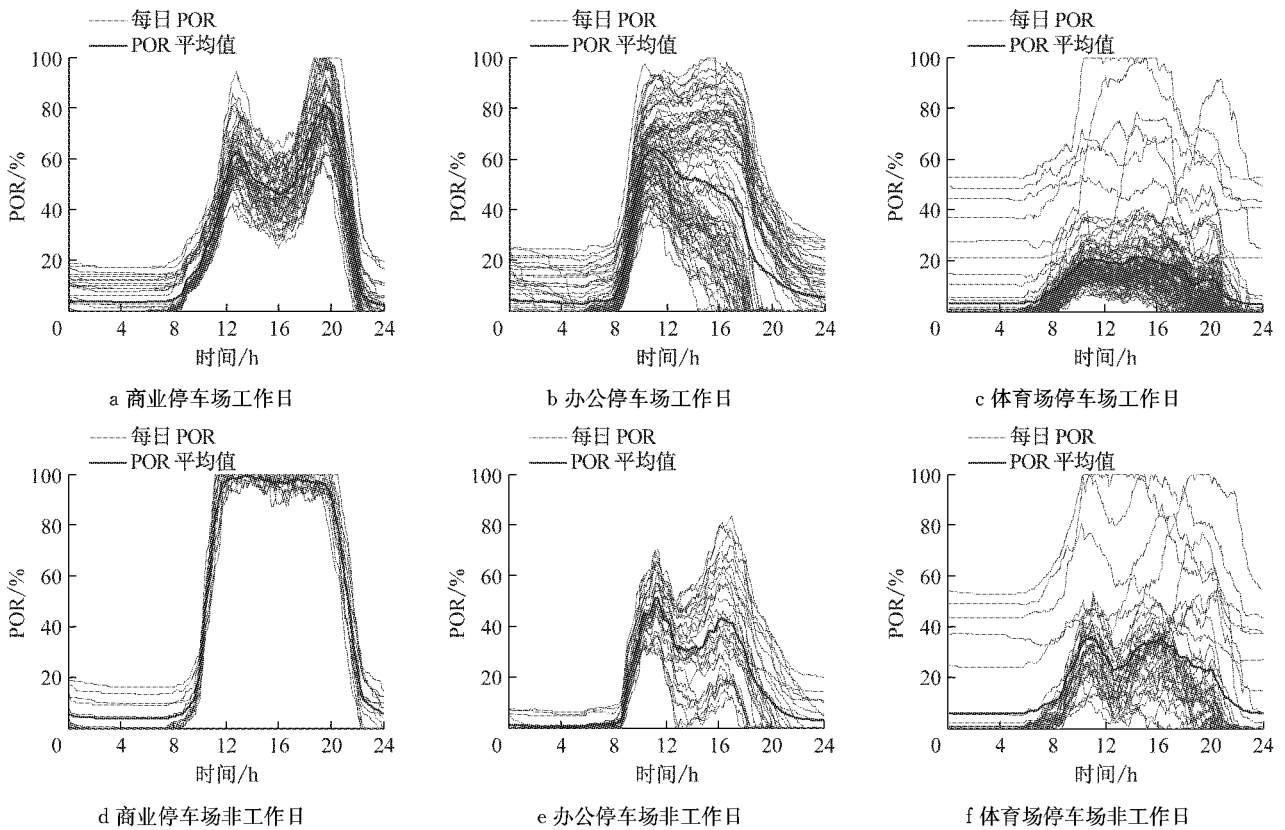


图1 3个停车场工作日及非工作日的POR日变化曲线

Fig.1 POR fluctuations on typical weekdays and weekends at three types of parking lots

2.2 停车泊位占有率的总体变化特征

3种停车场在2014年5月1日—9月7日的POR总体变化趋势见图2,图2中横坐标的时间对应的是停车位数据采集时间.由图2可知,3种停车场POR总体变化趋势和幅度差异很大,商业停车场POR日平均值浮动范围是0.3~1.0,整体高于另外2个停车场.办公停车场在5月—7月的POR明显高于其余几个月.体育场停车场POR日平均值通常小于0.4,但偶尔会有很大的波动,和体育场举行重大活动有关.此外,3个停车场通常每隔1周会有较大的周期性波动,体现出一定的周变特征.

3 模型选择

由图1和图2可知,3个停车场的POR变化具有不同的周期性和趋势性特征,在时间序列中属于

不平稳序列.如前所述,ARIMA模型、卡尔曼滤波模型和BP神经网络模型均可用于对不平稳时间序列进行预测,因此本文选取这3种模型进行对比分析.其中,ARIMA模型在建模时不必深究时间序列产生的背景,只需要有限的样本时间序列,就可以建立起较高精度的预测模型,但其存在低阶模型预测精度低、高阶模型参数估计难度大的不足;卡尔曼滤波模型具有动态修改预测权值的优点,但存在建立卡尔曼状态方程和测量方程较困难的不足^[19];BP神经网络模型在预测中具有较强的学习和数据处理能力,但存在容易陷入局部最小等缺点^[9].上述3种模型分别从不同角度对数据的时间序列特征进行建模,在预测中都有各自的优势和局限性.

3.1 ARIMA模型

ARIMA预测模型具体表达式如下:

$$\varphi(B)(1-B)^d x_t = c + \theta(B)\varepsilon_t \quad (1)$$

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad (2)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (3)$$

式(1)~(3)中: x_t 为 t 时刻 POR 的值; ε_t 为 t 时刻 POR 时间序列的随机误差, 是相互独立的白噪声序列, 它们服从均值为 0, 方差为常量 σ^2 的正态分布; $\varphi_i (i=1, 2, 3, \dots, p)$ 和 $\theta_j (j=1, 2, 3, \dots, q)$ 为模型的待估计参数; c 为待估计的常数项; B 为后移差分算子; d 为差分次数; p 和 q 为模型的阶. 上述模型可记作 ARIMA(p, d, q).

在用 ARIMA 模型对 POR 进行预测时, 先对 POR 时间序列进行平稳化处理, 如果序列是非平稳的, 可以通过差分变化使其满足平稳性条件; 然后根据数据的自相关函数 (autocorrelation function, ACF) 和偏自相关函数 (partial autocorrelation function, PACF) 进行模型识别, 并选择 Akaike 信息准则统计量 (Akaike information criterion, AIC) 较小的模型进行定阶, 即确定 p 和 q 的值; 接着估计模型中 φ_i 和 θ_j 的值, 并进行显著性检验; 最后根据估计的模型进行预测.

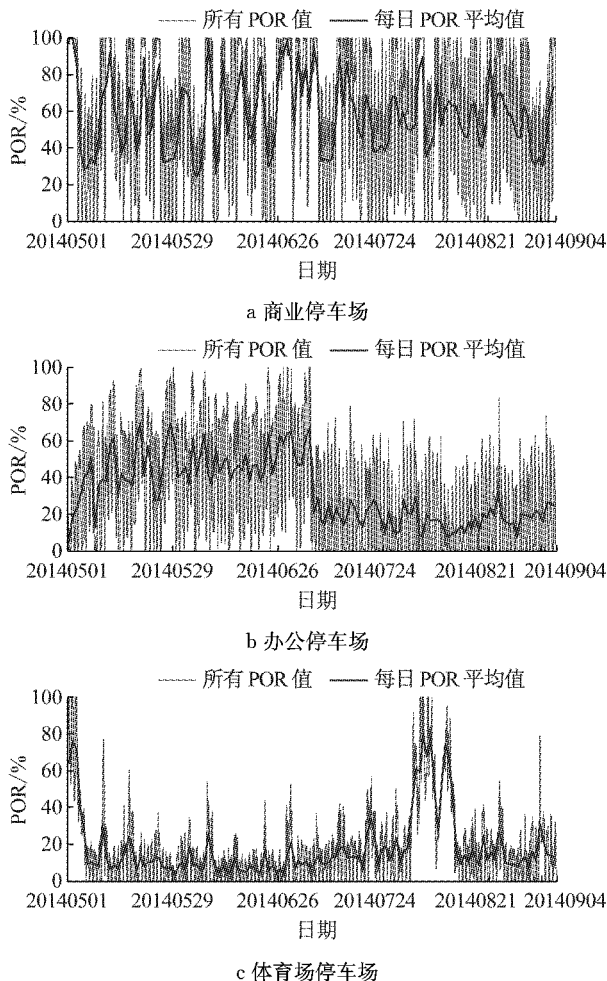


图 2 3 个停车场的 POR 总体变化曲线

Fig.2 Overall POR fluctuations at the three parking lots

3.2 卡尔曼滤波

卡尔曼滤波预测过程中的状态方程可表示为

$$\mathbf{X}_t = \mathbf{A}_{t-1} \mathbf{X}_{t-1} + \mathbf{W}_t \quad (4)$$

式中: \mathbf{X}_t 为 t 时刻的 POR 状态向量; \mathbf{A}_{t-1} 为从 $t-1$ 时刻到 t 时刻 POR 的状态转移矩阵; \mathbf{W}_t 为 t 时刻的系统噪声向量.

卡尔曼滤波预测过程中的观测方程可表示为

$$\mathbf{Z}_t = \mathbf{H}_t \mathbf{X}_t + \mathbf{V}_t \quad (5)$$

式中: \mathbf{Z}_t 为 t 时刻的 POR 观测向量; \mathbf{H}_t 为 t 时刻的预测输出转移矩阵; \mathbf{V}_t 为 t 时刻的观测噪声向量. 卡尔曼滤波中具体的预测递推方程如下:

$$\hat{\mathbf{X}}_t^- = \mathbf{A}_{t-1} \hat{\mathbf{X}}_{t-1} \quad (6)$$

$$\mathbf{P}_t^- = \mathbf{A}_{t-1} \mathbf{P}_{t-1} \mathbf{A}_{t-1}^T + \hat{\mathbf{Q}}_{t-1} \quad (7)$$

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t^- \mathbf{H}_t^T + \hat{\mathbf{R}}_{t-1})^{-1} \quad (8)$$

$$\hat{\mathbf{X}}_t = \hat{\mathbf{X}}_t^- + \mathbf{K}_t (\mathbf{Z}_t - \mathbf{H}_t \hat{\mathbf{X}}_t^-) \quad (9)$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_t^- \quad (10)$$

式中: $\hat{\mathbf{X}}_t$ 为 t 时刻 POR 的预测向量; $\hat{\mathbf{X}}_t^-$ 为 $t-1$ 时刻到 t 时刻 POR 的先验估计向量; \mathbf{P}_t^- 和 \mathbf{P}_t 分别为 t 时刻状态向量误差协方差矩阵的前估计和后估计; $\hat{\mathbf{Q}}_t$ 和 $\hat{\mathbf{R}}_t$ 分别为 t 时刻的系统噪声协方差和观测噪声协方差的估计; \mathbf{I} 为 $n \times n$ 标识矩阵; \mathbf{K}_t 为 $n \times m$ 矩阵, 被称为卡尔曼增益.

式(6)~(10)是卡尔曼滤波的 5 个基本方程, 其中式(6)~(7)是时间更新方程, 式(8)~(10)是状态更新方程. 预测前, 需要先确定卡尔曼滤波预测模型中 POR 的状态方程和观测方程, 然后根据式(6)~(10)的预测递推方程对 POR 进行预测.

3.3 BP 神经网络模型

BP 神经网络的结构包含输入层、隐含层和输出层 3 部分. 现有理论证明, 包含单个隐含层的 BP 神经网络, 通过增加隐含层的神经元个数, 就可以实现对任意非线性的映射^[20], 故本文采取单隐层的神经网络. BP 神经网络具体结构见图 3.

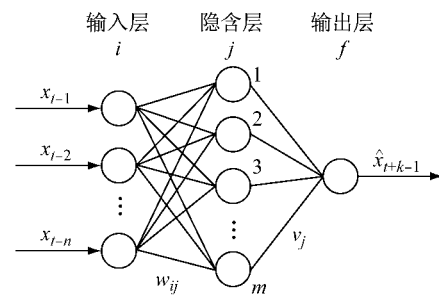


图 3 BP 神经网络模型结构

Fig.3 Structure of BP neural networks

在用BP神经网络模型对POR进行预测时,主要通过BP算法,用历史数据来训练神经网络,并最终确定合适的模型参数. BP算法基本思想是:先赋予网络初始权值和阈值,通过层间前向信息传递计算出网络的泊位占有率的输出值,根据比较网络的输出值和实际的泊位占有率的值之间的误差来修改网络的权值和阈值,通过不断的反复训练和比较使得实际输出和期望输出之间的误差达到最小. 在得到满足精度要求的连接权值 w_{ij} 和阈值 v_j 后,可输入某时刻前POR的时间序列值 x_1, x_2, \dots, x_n ,然后对未来的POR值 $x_{n+1}, x_{n+2}, \dots, x_{n+k}$ 进行预测. 其中 n 表示输入变量的个数,即输入层的节点数, k 表示预测的步长,如预测时间跨度是10 min时,预测步长 $k=2$.

本文先对神经网络的一些基本参数进行设定. 包括:①学习速率取0.01;②最大迭代次数取300;③所用激活函数为双曲正切S形函数,即“tansig”函数;④所用训练函数为梯度下降自适应学习率训练函数,即“traingdx”函数. 对于输入层和隐含层节点数的确定,本文采用试凑法. 根据以往的研究,本文输入层节点取2~20,隐含层节点取2~20,通过比较不同节点组合下的训练精度来得到最优的节点组合,即模型的最优结构^[18,20].

4 模型参数标定

本文采用2014年5月—8月6:00—23:00的POR作为历史数据来建立POR预测的ARIMA模型、卡尔曼滤波模型和BP神经网络模型,并根据工作日或非工作日,不同的预测时间跨度(即不同的预测步数)选择不同的模型参数. 其中ARIMA模型和卡尔曼滤波模型在进行多步预测(预测时间跨度超过5 min)时,都是将前一步预测值作为下一步预测的历史数据,所用的预测模型和单步预测时相同,故模型参数标定时不用考虑预测时间跨度;而BP神经网络模型进行多步预测时,不同预测步数的预测模型相互较为独立,故在进行模型参数标定时需要考虑预测时间跨度的不同.

4.1 ARIMA模型参数的确定

为建立ARIMA预测模型,需确定模型中的参数 p, d, q . 由图2可知,3个停车场的POR序列均为非平稳时间序列,为了消除POR序列中的不平稳成分,需对POR序列进行差分. 以商业停车场工作日的POR数据为例,记POR序列在2阶差分后为

$\{Y_t\}$ 序列,在单位根检验(ADF)结果中, $\{Y_t\}$ 序列的检验值为-51.908 86,小于1%,5%,10%显著性水平的临界值,说明 $\{Y_t\}$ 序列是平稳时间序列,故确定差分次数 $d=1$.

通过观察 $\{Y_t\}$ 序列的ACF图和PACF图,初步确定参数 p 和 q 的范围为: $p=14\sim 18; q=1\sim 5$. 为对模型进行定阶,需对多组可行阶数进行参数估计,然后选择AIC较小的模型作为最终模型. 商业停车场预测模型在各组可行阶数下的AIC值见表1. 由表1可知,当 $p=15, q=1$ 时AIC值最小为-7.503 76,故确定最终预测模型为ARIMA(15, 2, 1).

表1 模型各组可行阶数下的AIC值

Tab.1 AIC value with different feasible orders of the model

p	不同 q 时的AIC值				
	1	2	3	4	5
14	-7.499 7	-7.499 6	-7.498 9	-7.499 8	-7.502 7
15	-7.503 8	-7.498 8	-7.500 4	-7.502 7	-7.503 6
16	-7.497 0	-7.190 8	-7.499 1	-7.501 5	-7.484 5
17	-7.495 1	-7.497 6	-7.496 8	-7.492 1	-7.500 1
18	-7.482 9	-7.481 5	-7.480 3	-7.495 6	-7.497 7

对ARIMA定阶后,再对ARIMA预测模型的参数进行估计,模型的各个参数都通过检验. 在单位根检验结果中,模型残差序列的ADF检验值为-133.164 2,小于1%,5%,10%显著性水平的临界值,说明预测模型的残差序列是平稳随机序列,表明ARIMA预测模型是合适的. 按照上述方法,所有场景中ARIMA模型的定阶结果见表2.

表2 ARIMA最优定阶结果

Tab.2 Results of optimal orders of ARIMA model

停车场类型	是否工作日	p	d	q
商业	工作日	15	2	1
	非工作日	12	2	1
办公	工作日	8	2	1
	非工作日	12	1	1
体育场	工作日	6	1	6
	非工作日	8	1	14

对所有场景下的ARIMA预测模型的系数和残差序列进行检验,模型都通过检验,可进行预测.

4.2 卡尔曼滤波模型方程的确定

为实现卡尔曼预测,首先要推导出状态方程和观测方程. 由式(2)可知,状态转移矩阵描述了 $t-1$ 时刻到 t 时刻的变化关系,它是确定状态方程的关键. 本文通过对 $t-1$ 时刻和 t 时刻的POR进行一元回归得到转移矩阵. 以商业停车场工作日的POR数

据为例, $t-1$ 时刻和 t 时刻 POR 的关系见图 4.

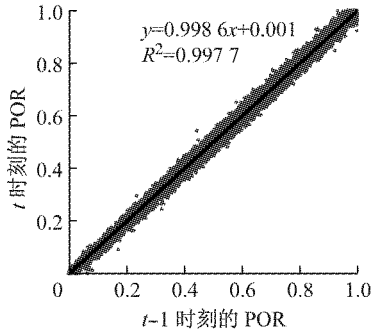


图 4 $t-1$ 时刻和 t 时刻的 POR

Fig.4 POR at time $t-1$ and POR at time t

由图 4 可知, $t-1$ 时刻和 t 时刻的 POR 较接近于线性关系. 故将 $t-1$ 时刻的 POR 作为自变量, t 时刻的 POR 作为因变量进行一元线性回归, 以商业停车场工作日为例, 得到结果如下:

$$x_t = 0.9986x_{t-1} + 0.0010 \quad (11)$$

式中: x_t, x_{t-1} 分别为 t 时刻和 $t-1$ 时刻的 POR. 回归结果中斜率和常数项的显著值都小于 0.05, 拟合的 R^2 值为 0.9977, 拟合效果较好.

由图 4 可知, 散点在趋势线附近仍有较大波动, 将这些波动视为系统噪声, 则式(11)可以转化为矩阵形式

$$\begin{pmatrix} x_t \\ 0.0010 \end{pmatrix} = \begin{pmatrix} 0.9986 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ 0.0010 \end{pmatrix} + W_t \quad (12)$$

再令

$$\mathbf{X}_t = \begin{pmatrix} x_t \\ 0.0010 \end{pmatrix} \quad \mathbf{A}_{t-1} = \begin{pmatrix} 0.9986 & 1 \\ 0 & 1 \end{pmatrix} \\ \mathbf{X}_{t-1} = \begin{pmatrix} x_{t-1} \\ 0.0010 \end{pmatrix} \quad (13)$$

则式(12)可以看作卡尔曼滤波预测过程中的状态方程. POR 预测的观测方程为

$$\mathbf{Z}_t = (1 \ 0) \begin{pmatrix} x_t \\ 0.0010 \end{pmatrix} + V_t \quad (14)$$

在确定了状态方程和观测方程之后, 可通过式(6)~(10)进行预测. 所有场景中卡尔曼滤波参数一元回归估计结果见表 3. 按照上述方法, 根据表 3 可得到所有场景中的状态方程和观测方程, 然后通过式(6)~(10)进行预测.

4.3 BP 神经网络模型结构的确定

神经网络模型结构通常根据均方误差 e 确定^[16,20], 表达式如下:

$$e = \sum_{t=1}^N (\hat{x}_t - x_t)^2 / N \quad (15)$$

式中: x_t 为 t 时刻 POR 的实际值; \hat{x}_t 为 t 时刻 POR 的预测值; N 为预测样本总数.

表 3 参数的一元回归估计结果

Tab.3 Estimated parameters of simple regression				
停车场类型	是否工作日	斜率 a	截距 b	R^2
商业	工作日	0.9986	0.0010	0.9977
	非工作日	0.9968	0.0022	0.9931
办公	工作日	0.9990	0.0003	0.9980
	非工作日	0.9983	0.0005	0.9965
体育场	工作日	0.9984	0.0002	0.9968
	非工作日	0.9981	0.0004	0.9961

由式(15)可知, e 在 0~1 之间, 且 e 值越小, 表示历史数据的训练效果越好. 以商业停车场工作日为例, 输入层节点数 n 取 2~20, 隐含层节点数 m 取 2~20, 历史数据的训练结果见图 5.

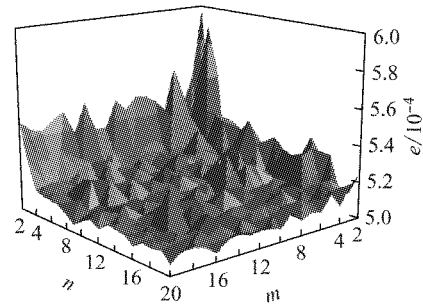


图 5 均方误差的 BP 神经网络训练结果

Fig.5 MSEs of training results of BP neural networks

由图 5 可知, 当输入层和隐含层节点数 n 小于 10 时, e 随着节点数的增加有逐渐减小的趋势; 当输入层 n 和隐含层节点数 m 在 10~20 范围时, e 波动较小, 不再有明显的下降趋势. 经计算, 当输入层节点数 n 为 19, 隐含层节点数 m 为 15 时, e 值最小为 5.02×10^{-4} , 可见 BPNN 最佳结构是 19-15-1. 同理可得其他场景中的最优结构见表 4. 可根据最优结构输入层节点数和隐含层节点数来进行预测.

5 预测结果比较

根据本文建立的模型, 采用 2014 年 9 月第 1 个星期六:00~23:00 的 POR 数据来对各个模型的预测精度进行测试. 以商业停车场为例, 图 6 中横坐标的时间对应的是 POR 的预测时间, 图 6 显示了商业停车场在预测时间跨度分别是 5, 10, 30 min(预测步长 $k=1, 2, 6$) 时, ARIMA 模型、卡尔曼滤波和 BP 神经网络模型 3 种方法的预测精度.

由图 6 可知, 当预测时间跨度较小时, 3 种方法的预测精度都很好. 随着预测时间跨度的增加, 3 种

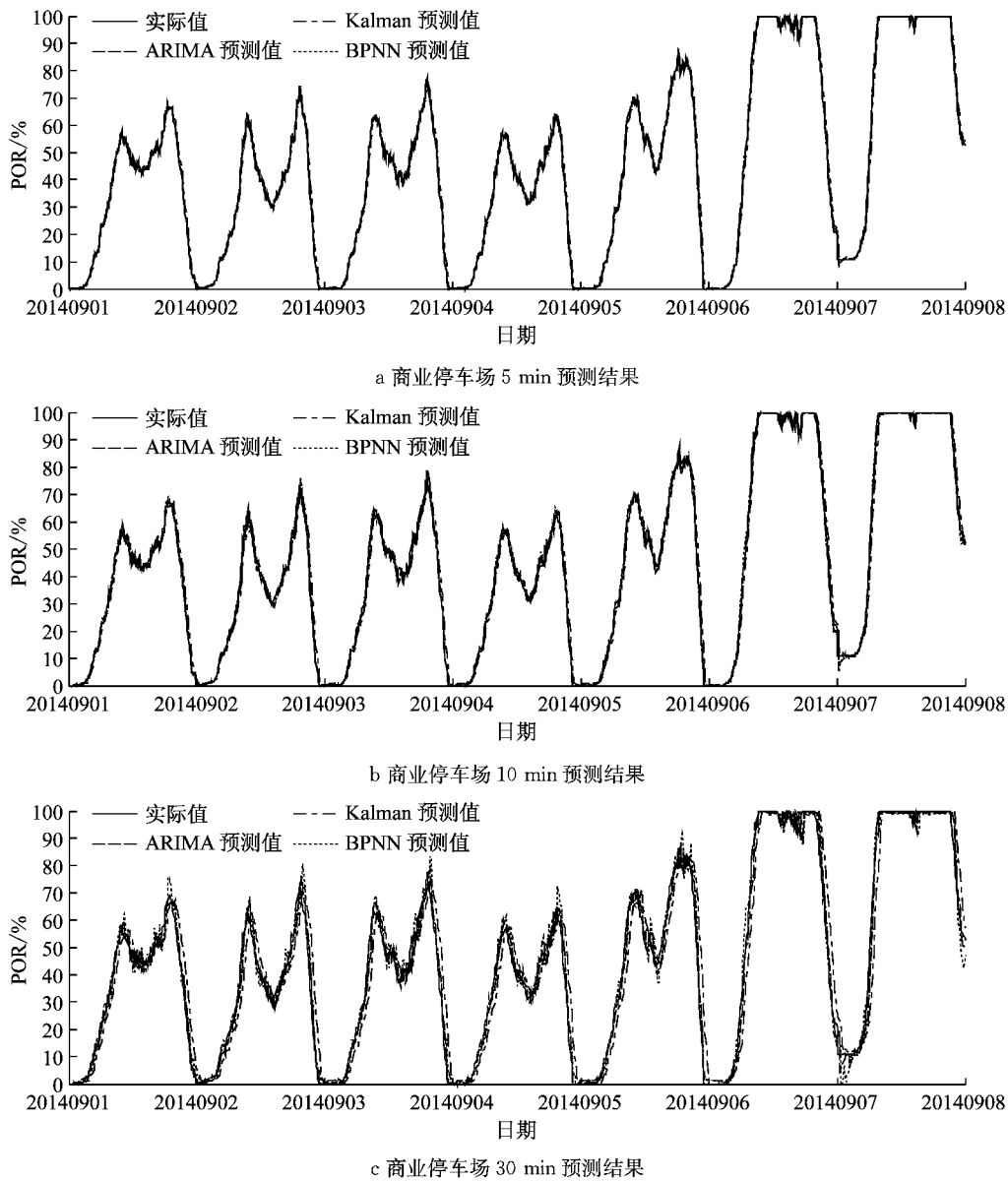


图 6 商业停车场 3 种方法的预测结果

Fig.6 Forecasting results of three methods for the parking lot in shopping mall

方法的预测精度不同程度地逐渐降低. ARIMA 模型和卡尔曼滤波的预测曲线和实际值相比出现了不同程度的滞后,其中卡尔曼滤波的滞后更加明显;BP 神经网络模型的预测曲线没有明显的滞后现象,但出现了较大的波动. 总体看来,随着预测时间跨度的增加,卡尔曼滤波的预测曲线和实际偏离较为严重,ARIMA 模型和 BP 神经网络模型的预测精度相对较高.

为了更全面地描述不同方法的预测精度,运用平均绝对误差 E 和均方根误差 r 来描述预测精度. 其中 E 和 r 的表达式为

$$E = \frac{1}{N} \sum_{i=1}^N |\hat{x}_i - x_i| \quad (16)$$

$$r = \sqrt{\frac{\sum_{i=1}^N (\hat{x}_i - x_i)^2}{N}} \quad (17)$$

式(16)~(17)中: x_t 为 t 时刻 POR 的实际值; \hat{x}_t 为 t 时刻 POR 的预测值; N 表示预测样本总数. E 和 r 的值都在 $0 \sim 1$ 之间,且值越小表示预测精度越高.

得到 POR 数据的 E 和 r 值后,可根据停车场的总泊位数将其转化为相应的可用泊位数的 E 和 r 值. 2014 年 9 月第 1 个星期 6:00~23:00 的 POR 数据的具体预测结果见表 5. 由表 5 可知:

(1) 在不同场景中,如果选择合适的方法,都可以获得较好的预测精度. 当预测时间跨度为 5 min 时,若选择合适的方法,POR 的 E 值都在 1% 以下, r 值在 1.2% 以下;随着预测时间跨度的增加,各方法

表 4 各场景下 BP 神经网络模型的最优结构

Tab. 4 Optimal structures of BP neural networks for different cases

停车场类型	是否工作日	预测时间跨度/min	n	m
商业	工作日	5	19	15
		10	16	19
		20	18	17
		30	16	17
	非工作日	5	19	14
		10	18	16
		20	19	16
		30	19	14
办公	工作日	5	18	14
		10	19	17
		20	17	14
		30	20	19
	非工作日	5	19	14
		10	18	12
		20	17	11
		30	20	17
体育场	工作日	5	15	15
		10	15	12
		20	17	15
		30	19	19
	非工作日	5	18	14
		10	18	17
		20	17	19
		30	18	18

的预测精度都逐渐降低. 但当预测时间跨度为 30 min 时, 若选择合适的方法, POR 的 E 值仍可保持在 2.6% 以下, r 值在 3.9% 以下.

(2) 对于办公和体育场停车场, 无论哪种方法, 工作日的预测精度都比非工作日高; 对于商业停车场, 工作日的 r 值比非工作日的小, 但 E 值比非工作日的大的.

(3) 不同类型的停车场预测差异较大. 如在 POR 的预测结果中, 对于 ARIMA 模型和 BP 神经网络模型, 在不同场景中一直是办公停车场的预测精度最高, 体育场停车场次之, 商业停车场最低. 这种差别源于不同类型的停车场 POR 的变化曲线有较大差异.

(4) 不同方法的预测精度差别很大. 如在商业和办公停车场中, 在预测时间跨度小于 10 min 时, BP 神经网络模型有最高的预测精度; 在预测时间跨度大于 10 min 时, ARIMA 模型有最高的预测精度. 无论哪种场景下, 卡尔曼滤波预测精度都最低.

(5) POR 的预测精度和可用泊位数的预测精度不完全一致. 如对于 BP 神经网络模型, 商业停车场

POR 的 r 值都大于办公停车场, 但泊位的 r 值都等于或小于办公停车场, 这是由于商业停车场比办公停车场的总泊位数少很多.

随着时间的推移、客观条件的变化, 模型质量可能下降, 模型参数可能要调整, 为了检验模型的自适应性, 采用 2015 年 4 月的 POR 数据对各个模型的预测精度进行测试, 结果见表 6. 由表 5 和表 6 的对比可知, 模型在 2015 年 4 月数据的预测精度多数低于 2014 年 9 月第 1 个星期的预测精度, 但相差不多. 可见随着时间的推移, 模型的预测精度有所下降, 但模型仍能保持较好的自适应性. 对训练数据进行实时更新, 调整相应的模型参数可能会使模型的预测精度进一步提高, 这也是下一步的研究方向之一.

6 结论

本研究分析了商业、办公和体育场地区 3 种类型停车场的 POR 变化特征, 并根据数据特征采用 ARIMA 模型、卡尔曼滤波模型和 BP 神经网络模型 3 种方法对 3 种停车场的 POR 进行预测和比较. 结果发现: ①3 种预测模型中, ARIMA 模型和 BP 神经网络模型有较高的预测精度, 卡尔曼滤波模型的预测精度较低, BP 神经网络模型在一些场景中的短时预测中具有优势; ②随着预测时间跨度的增加, 3 种方法的预测精度都逐渐降低; ③不同类型的停车场的 POR 预测精度存在差异, 对于同一种类型的停车场, 不同预测方法也会导致预测精度的不同; ④工作日预测精度一般比非工作日高; ⑤模型的预测精度会随着时间的推移而有所下降, 但模型仍能保持较好的自适应性.

停车泊位数预测的准确性直接影响停车诱导系统的运行效果, 本研究可为不同类型停车场泊位预测方法的选择提供参考. 但本研究仍存在一些不足, 例如, 所选用的停车场类型和预测方法有限, 且没有考虑各停车场之间的相关性以及天气、节假日等因素对预测的影响等. 下一步的研究将针对上述不足, 采用更多的数据量, 对不同的预测方法进行研究比较. 同时, 可考虑开发融合的模型来进行预测, 还可以根据停车诱导屏具体位置和停车场周围的交通状态, 开发考虑停车成功可靠性的预测方法等.

表 5 3 种方法的预测结果比较(测试数据:2014 年 9 月第 1 个星期)

Tab.5 Comparison of the three prediction methods(Test data: the first week of September 2014)

停车场类型	总泊位数/个	是否工作日	预测时间跨度/min	E					
				ARIMA		Kalman		BPNN	
				POR/%	可用泊位数	POR/%	可用泊位数	POR/%	可用泊位数
商业	306	工作日	5	0.9	3	1.7	5	0.9	3
			10	1.5	5	2.6	8	1.4	4
			20	2.1	6	4.4	13	2.2	7
			30	2.6	8	6.1	19	3.1	10
		非工作日	5	0.8	2	1.5	5	0.8	2
			10	1.4	4	2.4	7	1.3	4
			20	1.9	6	4.2	13	1.9	6
			30	2.6	8	5.9	18	3.0	9
办公	490	工作日	5	0.6	3	1.1	5	0.6	3
			10	0.9	4	1.7	8	0.8	4
			20	1.2	6	2.9	14	1.3	7
			30	1.6	8	4.1	20	1.9	9
		非工作日	5	0.7	4	1.3	6	0.7	4
			10	1.2	6	2.0	10	1.2	6
			20	1.7	8	3.3	16	1.9	9
			30	2.2	11	4.6	23	2.7	13
体育场	264	工作日	5	0.6	2	0.9	2	0.7	2
			10	1.0	3	1.3	3	1.1	3
			20	1.3	4	2.0	5	1.8	5
			30	1.6	4	2.7	7	2.4	6
		非工作日	5	0.9	2	1.2	3	0.9	2
			10	1.4	4	1.7	5	1.4	4
			20	1.8	5	2.7	7	2.3	6
			30	2.2	6	3.6	10	3.0	8
停车场类型	总泊位数/个	是否工作日	预测时间跨度/min	r					
				ARIMA		Kalman		BPNN	
				POR/%	可用泊位数	POR/%	可用泊位数	POR/%	可用泊位数
商业	306	工作日	5	1.2	4	2.3	7	1.2	4
			10	2.0	6	3.6	11	1.8	6
			20	2.7	8	6.0	18	2.9	9
			30	3.4	10	8.4	26	4.2	13
		非工作日	5	1.3	4	2.5	8	1.2	4
			10	2.2	7	4.0	12	1.9	6
			20	3.0	9	6.8	21	3.0	9
			30	3.9	12	9.6	29	4.4	13
办公	490	工作日	5	0.8	4	1.7	8	0.8	4
			10	1.3	6	2.7	13	1.2	6
			20	1.8	9	4.6	22	1.9	9
			30	2.3	11	6.4	31	2.6	13
		非工作日	5	1.0	5	1.8	9	1.0	5
			10	1.6	8	2.7	13	1.6	8
			20	2.2	11	4.6	22	2.5	12
			30	2.8	14	6.3	31	3.6	18
体育场	264	工作日	5	1.0	3	1.4	4	1.0	3
			10	1.6	4	2.1	6	1.6	4
			20	2.1	6	3.4	9	2.6	7
			30	2.6	7	4.6	12	3.7	10
		非工作日	5	1.2	3	1.6	4	1.3	3
			10	1.8	5	2.3	6	1.9	5
			20	2.4	6	3.6	10	3.0	8
			30	2.9	8	4.8	13	4.0	11

表 6 3种方法的预测结果比较(测试数据:2015年4月)

Tab.6 Comparison of the three prediction methods(Test data: April 2015)

停车场 类型	总泊位 数/个	是否 工作日	预测时间 跨度/min	<i>E</i>					
				ARIMA		Kalman		BPNN	
				POR/%	可用 泊位数	POR/%	可用 泊位数	POR/%	可用 泊位数
商业	306	工作日	5	1.0	3	1.8	6	0.9	3
			10	1.6	5	2.8	9	1.4	4
			20	2.2	7	4.8	15	2.4	7
			30	2.8	9	6.8	21	3.3	10
		非工作日	5	1.2	4	2.2	7	1.2	4
			10	1.9	6	3.3	10	1.8	6
			20	2.6	8	5.4	16	2.7	8
			30	3.4	10	7.4	23	3.8	12
办公	490	工作日	5	0.6	3	1.1	5	0.6	3
			10	1.0	5	1.7	9	0.9	5
			20	1.4	7	3.0	14	1.6	8
			30	1.8	9	4.1	20	2.2	11
		非工作日	5	0.7	3	1.2	6	0.7	3
			10	1.2	6	1.9	9	1.1	6
			20	1.7	8	3.2	16	2.0	10
			30	2.2	11	4.4	21	2.9	14
体育场	264	工作日	5	0.6	2	0.7	2	0.6	2
			10	0.8	2	1.0	3	0.9	2
			20	1.1	3	1.5	4	1.3	4
			30	1.3	3	2.0	5	1.8	5
		非工作日	5	0.8	2	1.0	3	0.8	2
			10	1.3	3	1.5	4	1.3	3
			20	1.7	4	2.3	6	2.0	5
			30	2.0	5	3.0	8	2.7	7
停车场 类型	总泊位 数/个	是否 工作日	预测时间 跨度/min	<i>r</i>					
				ARIMA		Kalman		BPNN	
				POR/%	可用 泊位数	POR/%	可用 泊位数	POR/%	可用 泊位数
商业	306	工作日	5	1.4	4	2.6	8	1.3	4
			10	2.1	7	4.0	12	2.0	6
			20	2.9	9	6.7	21	3.3	10
			30	3.7	11	9.4	29	4.5	14
		非工作日	5	1.7	5	3.1	10	1.6	5
			10	2.6	8	4.9	15	2.4	7
			20	3.6	11	8.3	25	3.6	11
			30	4.6	14	11.5	35	5.3	16
办公	490	工作日	5	1.1	5	1.9	9	1.1	5
			10	1.8	9	2.9	14	1.7	8
			20	2.4	12	4.8	23	2.7	13
			30	3.0	15	6.5	32	3.6	18
		非工作日	5	1.1	5	1.9	9	1.1	5
			10	1.8	9	2.9	14	1.8	9
			20	2.5	12	4.8	24	2.9	14
			30	3.2	16	6.5	32	4.0	19
体育场	264	工作日	5	0.8	2	1.0	3	0.8	2
			10	1.2	3	1.5	4	1.2	3
			20	1.6	4	2.3	6	1.9	5
			30	1.9	5	3.1	8	2.5	7
		非工作日	5	1.2	3	1.5	4	1.2	3
			10	1.8	5	2.1	6	1.8	5
			20	2.3	6	3.3	9	2.8	7
			30	2.8	7	4.2	11	3.7	10

参考文献:

- [1] SHOUP D C. Cruising for parking[J]. *Transport Policy*, 2006, 13(6): 479.
- [2] TENG H, FALCOCCHIO J C, QI Y, *et al.* Parking difficulty and parking information needs for off-street spaces in the CBD [C/CD] // *Transportation Research Board 80th Annual Meeting*. Washington D C: TRB, 2001.
- [3] 郭树明. 智能停车场管理系统中车位信息预测子系统的设计与实现[D]. 长春:吉林大学, 2012.
GUO Shuming. Design and implementation of parking information forecasting subsystem in intelligent parking management system[D]. Changchun: Jilin University, 2012.
- [4] BLYTHE P T. Forecasting available parking space with largest Lyapunov exponents method[J]. *Journal of Central South University*, 2014, 21(4): 1624.
- [5] TAMRAZIAN A, QIAN Z S, RAJAGOPAL R. Where is my parking spot? On-line and off-line prediction of time-varying parking occupancy [C/CD] // *Transportation Research Board 94th Annual Meeting*. Washington D C: TRB, 2015.
- [6] LIU S, GUAN H, YAN H, *et al.* Unoccupied parking space prediction of chaotic time series [C]//ICCTP 2010. *Integrated Transportation Systems: Green, Intelligent, Reliable*. Beijing: ICCTP, 2010: 2122-2131.
- [7] RICHTER F, DI MARTINO S, MATTFELD D C. Temporal and spatial clustering for a parking prediction service[C]// *Tools with Artificial Intelligence (ICTAI)*. 2014 IEEE 26th International Conference on. Limassol: IEEE, 2014: 278-282.
- [8] 李祎, 孙静怡, 张旭东. 城市停车状况的模糊聚类分析[J]. *昆明理工大学学报(理工版)*, 2008, 33(2): 62.
LI Yi, SUN Jingyi, ZHANG Xudong. Fuzzy cluster analysis of city parking status [J]. *Journal of Kunming University of Science and Technology (Science and Technology)*, 2008, 33(2): 62.
- [9] 熊志斌. 基于 ARIMA 与神经网络集成的 GDP 时间序列预测研究[J]. *数理统计与管理*, 2011, 30(2): 306.
XIONG Zhibin. Research on GDP time series forecasting based on integrating ARIMA with neural networks [J]. *Journal of Applied Statistics and Management*, 2011, 30(2): 306.
- [10] VLAHOIANNI E I, GOLIAS J C, KARLAFTIS M G. Short-term traffic forecasting: overview of objectives and methods [J]. *Transport Reviews*, 2004, 24(5): 533.
- [11] DUNNING A E. Method and system for projecting dynamic parking availability based on an ongoing survey for remote lots with high demand: US, Patent 7,049,979[P]. 2006-5-23.
- [12] RAJABIOUN T, IOANNOU P. On-street and off-street parking availability prediction using multivariate spatiotemporal models [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(5): 2913.
- [13] OKUTANI I, STEPHANEDES Y J. Dynamic prediction of traffic volume through Kalman filtering theory [J]. *Transportation Research Part B: Methodological*, 1984, 18(1): 1.
- [14] JI H, XU A, SUI X, *et al.* The applied research of Kalman in the dynamic travel time prediction[C]//*Geoinformatics. 2010 18th International Conference on*. Kyoto: IEEE, 2010: 1-5.
- [15] YANG Z, LIU H, WANG X. The research on the key technologies for improving efficiency of parking guidance system[C] // *Intelligent Transportation Systems*. St Louis: IEEE, 2003, 2: 1177-1182.
- [16] SUN Yong, LIU Chunping, WANG Yihuai, *et al.* A forecasting model for parking guidance system[C]//*Computer Science and Information Engineering. 2009 WRI World Congress on*. Los Angeles: IEEE, 2009, 3: 607-611.
- [17] VLAHOIANNI E I, KEPAPTSOGLU K, TSETOS V, *et al.* A real-time parking prediction system for smart cities[J]. *Journal of Intelligent Transportation Systems*, 2016, 20(2): 192.
- [18] JI Y, TANG D, BLYTHE P, *et al.* Short-term forecasting of available parking space using wavelet neural network model [J]. *Intelligent Transport Systems*, 2015, 9(2): 202.
- [19] 潘迪夫, 刘辉, 李燕飞. 基于时间序列分析和卡尔曼滤波算法的风电场风速预测优化模型[J]. *电网技术*, 2008, 32(7): 82.
PAN Difu, LIU Hui, LI Yanfei. A wind speed forecasting optimization model for wind farms based on time series analysis and Kalman filter algorithm [J]. *Power System Technology*, 2008, 32(7): 82.
- [20] FADARE D A. The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria [J]. *Applied Energy*, 2010, 87(3): 934.