

# 基于奇异谱分析和 CKF-LSSVM 的短时交通流量预测

商 强<sup>1</sup>, 杨兆升<sup>1,2,3</sup>, 张 伟<sup>1,2,4</sup>, 邴其春<sup>1</sup>, 周熙阳<sup>1</sup>

(1. 吉林大学 交通学院, 长春 130022; 2. 吉林大学 汽车仿真与控制国家重点实验室, 长春 130022; 3. 吉林大学 吉林省道路交通重点实验室, 长春 130022; 4. 山东高速公路股份有限公司, 济南 250014)

**摘要:** 为了提高短时交通流预测的精度, 提出了基于奇异谱分析和组合核函数最小二乘支持向量机(CKF-LSSVM)的短时交通流预测模型。首先, 采用奇异谱分析方法, 滤除交通流序列的噪声成分。然后, 使用降噪后的交通流数据训练 CKF-LSSVM, 并通过粒子群优化算法确定模型参数。最后, 以厦门市的实测数据为基础, 对预测模型进行实验验证和对比分析。结果表明: 本文所构建模型具有较好的预测效果, 能够有效提高短时交通流预测精度。

**关键词:** 交通运输系统工程; 短时交通流预测; 奇异谱分析; 支持向量机; 组合核函数

**中图分类号:** U491 **文献标志码:** A **文章编号:** 1671-5497(2016)06-1792-07

**DOI:** 10.13229/j.cnki.jdxbgxb201606006

## Short-term traffic flow prediction based on singular spectrum analysis and CKF-LSSVM

SHANG Qiang<sup>1</sup>, YANG Zhao-sheng<sup>1,2,3</sup>, ZHANG Wei<sup>1,2,4</sup>, Bing Qi-chun<sup>1</sup>, ZHOU Xi-yang<sup>1</sup>

(1. College of Transportation, Jilin University, Changchun 130022, China; 2. State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China; 3. Jilin Province Key Laboratory of Road Traffic, Jilin University, Changchun 130022, China; 4. Shandong High-speed Co., Ltd., Ji'nan 250014, China)

**Abstract:** In order to improve the accuracy of short-time traffic flow prediction, a short-term traffic flow prediction model based on singular spectrum analysis and Combined Kernel Function (CKF) Least Square Support Vector Machine (LSSVM) is proposed. Singular spectrum analysis technology is used to filter out the noise of traffic time series. Then, the processed traffic flow data are used to train the CKF-LSSVM, and the parameters of the model are determined by particle swarm optimization algorithm. Finally, validation and comparative analysis of the model are carried out using the measured in Xiamen, China. Experimental results indicate that the proposed model has good prediction performance and can effectively improve the accuracy of short-time traffic flow prediction.

**Key words:** engineering of communication and transportation system; short-term traffic flow prediction; singular spectrum analysis; support vector machine; combined kernel function

**收稿日期:** 2015-11-23.

**基金项目:** 国家科技支撑计划项目(2014BAG03B03); 国家自然科学基金项目(51308248, 51408257, 51308249); 山东省省管企业科技创新项目(20122150251-1).

**作者简介:** 商强(1987-), 男, 博士研究生. 研究方向: 智能交通系统理论与技术.

E-mail: shangqiang14@mails.jlu.edu.cn

**通信作者:** 张伟(1978-), 男, 高级工程师, 在站博士后. 研究方向: 智能交通系统理论与技术. E-mail: z\_wei@126.com

## 0 引 言

短时交通流预测是智能交通系统(ITS)研究的重要组成部分。实时、准确的短时交通流预测能够直接应用于 ITS 子系统中<sup>[1]</sup>,如先进的交通管理系统(ATMS)和先进的出行者信息系统(ATIS)。因此,短时交通流预测受到了广泛的关注。已有文献提出了大量的短时交通流预测模型和方法,大致可以分为 3 类:①基于交通流模拟的预测方法,通过模拟交通流运行状态,预测其未来的发展趋势。宏观交通流模拟是将交通流比拟成水流,借鉴流体力学的相关理论分析交通流。微观交通流模拟则考虑的是单车行为以及路网中车与车之间的相互作用<sup>[2,3]</sup>。②基于统计学的预测模型。通过数据集的回归拟合和参数的选择优化,以实现交通预测。例如:局部线性回归模型、ARMA 模型和 Kalman 滤波模型<sup>[4-6]</sup>。③基于数据挖掘的预测方法。这类方法一般没有固定的形式,而是以数据驱动,通过智能算法寻找历史数据的变化趋势,进而预测未来数据。例如人工神经网络和支持向量机<sup>[7,8]</sup>。与前两类方法相比,基于数据挖掘的预测方法没有对模型做出诸多假设(例如假设观测时间序列是平稳的,残差服从正态分布,预先给定模型的结构),因而具有更好的适用性。最小二乘支持向量机(Least squares support vector machine, LSSVM)是支持向量机的一种改进形式,学习速度更快<sup>[9]</sup>。然而,在数据采集和传输过程中,不可避免地受到交通检测器自身软(硬)件故障、外界环境异常(如电磁干扰)等因素的影响,导致采集数据中含有噪声成分。如果将观测数据直接用于 LSSVM 的训练,则一定程度上降低了模型的泛化能力。奇异谱分析(Singular spectrum analysis, SSA)是一种有效的降噪方法,已在水文预测、电力负荷预测和天气预测等方面有了相关应用<sup>[10]</sup>。文献[11]使用多种方法处理 GPS 坐标数据,结果表明 SSA 降噪效果优于小波和经验模态分解方法。

本文提出一种基于 SSA 和组合核函数最小二乘支持向量机(Combined kernel function least squares support vector machine, CKF-LSSVM)的短时交通流预测模型(SSA-CKF-LSSVM)。首先,通过 SSA 将交通流序列分解和重构,得到原序列的趋势成分、振荡成分和噪声成分。然后,使用除去噪声成分的交通流数据训练 CKF-

LSSVM,并通过粒子群优化算法确定模型参数组合。最后,利用实测交通流数据,将新建模型与其他模型的预测效果进行对比。

## 1 奇异谱分析

将长度为  $N$  的一维时间序列  $T = (x_1, x_2, \dots, x_N)$  转化为一系列的  $L$  维向量:  $\mathbf{X}_i = (x_i, x_{i+1}, \dots, x_{i+L-1})^T$ 。其中,  $L(1 < L < N)$  称为窗口长度(Window length)。由  $K$  个向量  $\mathbf{X}_i(i = 1, \dots, K = N - L + 1)$  构成轨迹矩阵  $\mathbf{X}(L \times K)$ :

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{bmatrix} \quad (1)$$

将  $\mathbf{X}$  进行奇异值分解:

$$\mathbf{X} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^T \quad (2)$$

式中:  $\mathbf{S}$  为矩阵  $\mathbf{X}$  的奇异值,等价于矩阵  $\mathbf{X}\mathbf{X}^T$  特征值的平方根;  $\mathbf{U}$  为  $\mathbf{X}$  的左奇异向量,等价于矩阵  $\mathbf{X}\mathbf{X}^T$  的特征向量;  $\mathbf{V}$  为矩阵  $\mathbf{X}$  的右奇异向量,等价于  $\mathbf{X}^T\mathbf{X}$  的特征向量。

将时间序列  $T$  的元素由奇异谱分析展开:

$$x_{i+j} = \sum_{k=1}^M a_{ik} \mathbf{E}_{kj} \quad (3)$$

式中:  $i = 1, \dots, K; j = 1, \dots, M; \mathbf{E}_{kj} = \mathbf{U}$  为  $\mathbf{X}\mathbf{X}^T$  的特征向量,  $\mathbf{E}$  称为时间经验正交函数;  $a_{ik}$  为通过式(4)求得的时间主成分:

$$a_{ik} = \sum_{j=1}^M x_{i+j} \mathbf{E}'_{kj} \quad (4)$$

$x_i$  的第  $k$  个主成分记作  $\mathbf{X}_i^k$ , 通过第  $k$  个时间经验正交函数和时间主成分重构得到,如式(5)所示:

$$\mathbf{X}_i^k = \begin{cases} \left( \sum_{j=1}^L a_{ij,k} \mathbf{E}_{kj} \right) / L, & L \leq i \leq K \\ \left( \sum_{j=1}^i a_{ij,k} \mathbf{E}_{kj} \right) / i, & 1 \leq i \leq L-1 \\ \frac{1}{N-i+1} \sum_{j=i-n+1}^L a_{ij,k} \mathbf{E}_{kj}, & N-L+2 \leq i \leq N \end{cases} \quad (5)$$

根据式(6)选择对应较大奇异值的  $m$  个成分相加,所得到  $\tilde{\mathbf{x}}_i$  近似代表  $\mathbf{x}_i$ :

$$\tilde{\mathbf{x}}_i = \sum_{k=1}^m \mathbf{X}_i^k, \quad i = 1, \dots, N \quad (6)$$

因此,重构过程很大程度上滤除那些对应较小奇异值的噪声成分。

## 2 组合核函数 LSSVM

设  $D = (x_i, y_i), i = 1, \dots, N$  为样本集,  $x_i \in \mathbf{R}^m$  为输入;  $y_i \in \mathbf{R}$  为相应输出。回归函数为:

$$f(x) = \omega^T \varphi(x) + b \quad (7)$$

式中:  $\varphi(\cdot)$  为非线性函数;  $\omega$  为权向量;  $b$  为偏置量。

LSSVM 回归的优化目标函数为:

$$\begin{cases} \min_{\omega, b, e} J_1(\omega, e) = \mu E_w + \zeta E_D = \frac{1}{2} \mu \omega^T \omega + \frac{1}{2} \zeta \sum_{i=1}^N e_i^2 \\ \text{s. t. } y_i = \omega^T \varphi(x) + b + e_i \end{cases} \quad (8)$$

式中:  $e_i$  为误差变量;  $\mu$  和  $\zeta$  为可变参数;  $E_w = \frac{1}{2} \omega^T \omega$ ;  $E_D = \frac{1}{2} \sum_{i=1}^N e_i^2 = \frac{1}{2} \sum_{i=1}^N [y_i - \omega^T \varphi(x) + b]^2$ 。

为了求解上述优化问题,因此构建 Lagrange 函数:

$$L(\omega, b, e, \alpha) = J_1(\omega, e) - \sum_{i=1}^N \alpha_i \{ \omega^T \varphi(x) + b + e_i - y_i \} \quad (9)$$

式中:  $\alpha_i$  为 Lagrange 乘子,由 KKT 条件可得:

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \rightarrow \omega = \sum_{i=1}^N \alpha_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^N \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \rightarrow \omega^T \varphi(x_i) + b + e_i - y_i = 0 \end{cases} \quad (10)$$

式中:惩罚系数  $\gamma = \zeta/\mu$ 。消除  $\omega$  和  $e_i$  后,原优化问题变为:

$$\begin{bmatrix} \mathbf{0} & \mathbf{l}^T \\ \mathbf{l} & \mathbf{\Omega}_{ij} + \frac{1}{\gamma} \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{Y} \end{bmatrix} \quad (11)$$

式中:  $\mathbf{\Omega}_{ij} = \varphi(x_i)^T \varphi(x_j) = K(x_i, x_j)$ ;  $\mathbf{l} = [1, \dots, 1]^T$ ;  $\mathbf{a} = [\alpha_1, \dots, \alpha_N]^T$ ;  $\mathbf{Y} = [y_1, \dots, y_N]^T$ 。

通过式(11)求出  $\mathbf{a}$  和  $\mathbf{b}$ , 那么 LSSVM 回归模型为:

$$f(x) = \sum_{i=1}^N \alpha_i K(x, x_i) + b \quad (12)$$

式中:  $K(x, x_i)$  为核函数,需符合 Mercer 条件,如 RBF 核函数、线性核函数和多项式核函数等。其中,RBF 核函数为:

$$K(x, x_i) = \exp[-\|x - x_i\|^2 / (2\sigma^2)], \sigma > 0 \quad (13)$$

多项式核函数为:

$$K(x, x_i) = (\mathbf{x}x_i + 1)^q \quad (14)$$

式中:  $q$  为参数。

RBF 核函数在实际应用中最为广泛,是一种典型局部核函数,学习能力较强,泛化能力较弱;而全局核函数与之相反<sup>[12]</sup>。多项式核函数是常用的全局核函数。这两种核函数的有机组合能够提高 LSSVM 模型预测精度。因此,本文模型采用多项式函数( $K_{\text{poly}}$ )和 RBF 核函数( $K_{\text{RBF}}$ )的线性组合核函数  $K_{\text{com}}$ ,即:

$$K_{\text{com}} = \rho K_{\text{RBF}} + (1 - \rho) K_{\text{poly}} \quad (15)$$

式中:  $\rho$  为参数。

## 3 优化模型参数

组合核函数 LSSVM 参数包括:惩罚系数  $\gamma$ 、RBF 核函数参数  $\sigma$ 、多项式核函数参数  $q$  和组合核函数系数  $\rho$ 。通过粒子群寻优算法(Particle swarm optimization, PSO)确定模型参数组合。PSO 是一种启发式全局寻优算法,其基本原理是:假设  $d$  维搜索空间中的第  $i$  个微粒的速度  $v_i = (v_{i1}, v_{i2}, \dots, v_{id})$ , 位置  $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$ 。通过评价各微粒的适应度,确定  $t$  时刻每个微粒所经过的最佳位置  $p_i = (p_{i1}, p_{i2}, \dots, p_{id})$  和种群中已找到的最优位置  $p_g = (p_{g1}, p_{g2}, \dots, p_{gd})$ 。根据式(16)更新微粒的速度,根据式(17)更微粒的位置:

$$v_{i,j}(t+1) = wv_{i,j}(t) + c_1 \text{rand}(p_{i,j} - x_{i,j}(t)) + c_2 \text{rand}(p_{g,i} - x_{i,j}(t)) \quad (16)$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \quad (17)$$

$$j = 1, 2, \dots, d$$

式中:  $w$  为惯性权值;  $c_1$  和  $c_2$  为加速因子;rand 为 0 到 1 之间的随机数,通过设置微粒的速度区间  $[v_{\min}, v_{\max}]$  和位置范围  $[x_{\min}, x_{\max}]$  可适当控制微粒移动。

惯性权值  $w$  代表粒子前一次迭代速度对当前迭代速度的影响,较大的  $w$  全局搜索能力较好,而较小的  $w$  局部搜索能力较好。不断调整惯性权值,符合算法进化特性。因此,使用线性递减的惯性权值公式:

$$w = w_{\text{start}} - \frac{w_{\text{start}} - w_{\text{end}}}{T_{\text{max}}} T \quad (18)$$

式中:  $w_{\text{start}}$  为起始惯性权值;  $w_{\text{end}}$  为终止惯性权值;  $T$  为当前迭代次数;  $T_{\text{max}}$  为最大迭代次数。

SSA-CKF-LSSVM 模型构建和优化流程为:

Step1 采用奇异谱分析,滤除原交通流时间序列的噪声成分。

Step1.1 嵌入:确定延迟窗口长度  $L$ 。

Step1.2 奇异值分解:获得  $L$  个奇异值。

Step1.3 分组:将奇异值从大到小排序,后面的奇异值较小,且变化不大,对应噪声成分。

Step1.4 对角平均:根据奇异值分组,重构降噪交通流量序列。

Step2 使用降噪后数据,构建训练样本集。

Step3 粒子群算法优化模型参数。

Step3.1 模型参数初始化,粒子群参数初始化。

Step3.2 确定适应度函数。

Step3.3 计算并比较适应度,更新微粒速度和位置。

Step3.4 满足停止条件,输出模型参数。

## 4 实例分析

数据来源于厦门市莲前西路的地磁检测器,检测器统计时间间隔为 5 min,采集连续 5 个工作日(2015 年 1 月 5 日~2015 年 1 月 9 日)的交通流量数据。编号为 DC00004963 的检测器和编号为 DC00004964 的检测器分别采集由西向东和由东向西的交通流量数据,所得交通流量时间序列如图 1 所示。前 4 天交通流量用于 SSA-CKF-LSSVM 模型的构建,第 5 天交通流量用于模型预测效果的验证和对比分析。下面将以 DC00004963 号检测器数据为例详细说明模型构

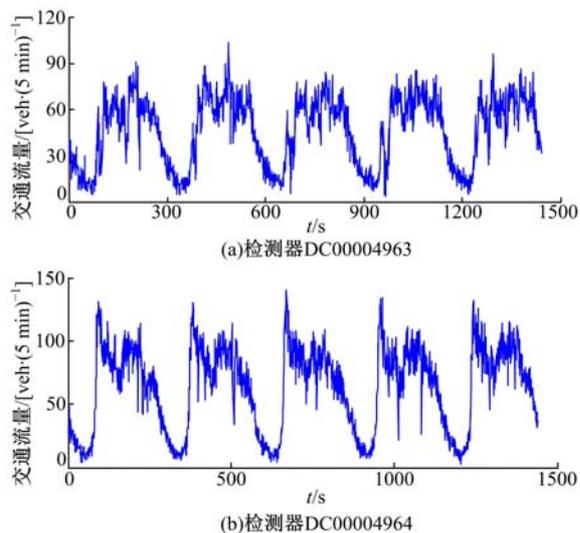


图 1 交通流量时间序列

Fig.1 Traffic flow time series

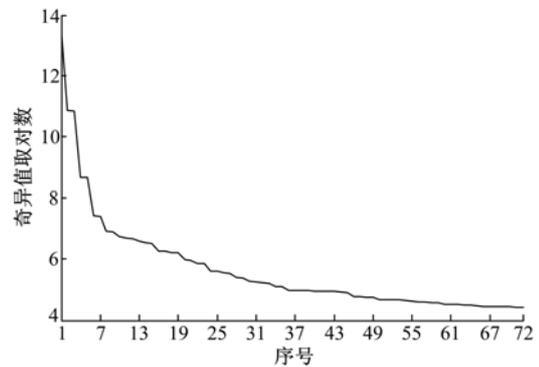


图 2 交通流量序列的奇异谱

Fig.2 Singular spectrum of traffic flow time series 建和优化的过程。

### 4.1 奇异谱分析

在奇异谱分析中,窗口长度  $L$  是唯一需要确定的参数。此处  $L$  取 288,即一天交通流量序列的长度。所得  $L$  个奇异值按从大到小顺序排列称为奇异谱。前 72 个(总共 288 个)奇异值按从大到小顺序排列的曲线如图 2 所示。从第 36 个奇异值开始,后面奇异值变化缓慢,因此 36~288 个奇异值对应着噪声成分。最大奇异值对应着趋势成分,2~35 个奇异值对应着振荡成分,例如(2,3)、(4,5)、(6,7)和(8,9)等,它们成对存在,奇异值大小相同或相近,也说明它们对应着振荡成分。

图 3 为前 9 个奇异值(特征三元组)对应的重构序列及其贡献率。第 1~35 个奇异值对应重构的序列如图 4 所示,累积贡献率为 98.390%。图 5 为第 36~288 个奇异值对应的重构序列,即噪声序列。由图 3~图 5 可知:第 1 个奇异值对应的重构序列为趋势成分,贡献率为 81.104%;第 2~35 个奇异值对应的重构序列为振荡成分,累积贡献率为 17.286%;第 36~288 个奇异值对应的重构序列为噪声成分,具有较高的频率,累积贡献率仅为 1.610%。

### 4.2 组合核函数 LSSVM 的参数优化

将预测值作为模型输出,预测值前两小时的交通流量序列作为模型输入,以此实现交通流量的单步预测,即使用历史数据滚动预测下一时刻交通流量。使用前 4 天 SSA 降噪交通流量数据构建训练集,包括了 1128 组输入-输出关系。

采用 PSO 算法优化模型参数,包括惩罚系数  $\gamma$ 、RBF 核函数参数  $\sigma$ 、多项式核函数参数  $q$  和组合核函数系数  $\rho$ ,共计 4 个参数,即粒子群维数  $d = 4$ 。设置种群数量为 20;加速因子  $c_1 = 1.5$ ;

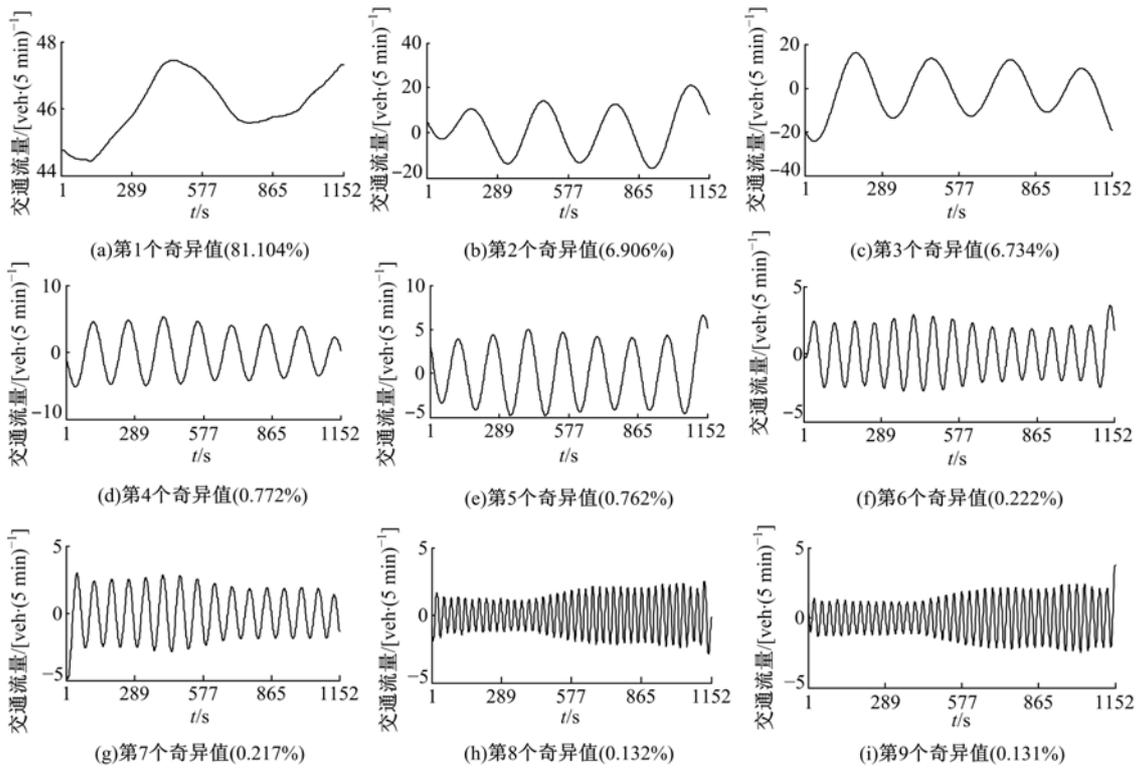


图3 前9个奇异值对应的重构序列及其贡献率

Fig. 3 Reconstructed traffic series of first 9 singular value

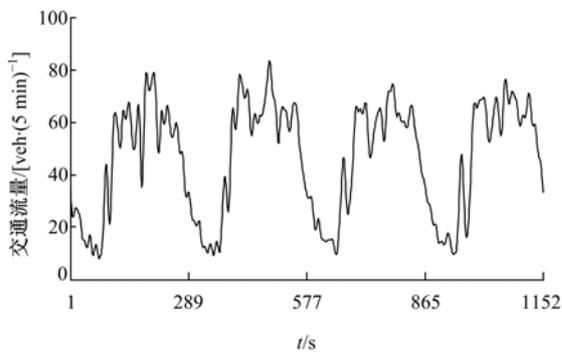


图4 重构交通流量序列

Fig. 4 Reconstructed traffic time series

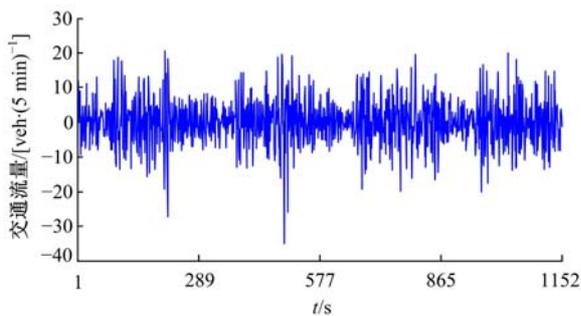


图5 噪声序列

Fig. 5 Noisy time series

$c_2 = 1.7$ ; 最大迭代次数为  $T_{max} = 100$ ; 最大速度  $v_{max} = 2$ ; 起始惯性权重  $\omega_{start} = 0.9$ ; 终止惯性权重  $\omega_{end} = 0.4$ 。模型参数:  $\gamma \in (0, 1000]$ ,  $\sigma \in (0, 100]$ ,  $\rho \in [0, 1]$ ,  $q \in (0, 1, \dots, 10)$ 。适应度函数 fitness 为平均百分比误差(MAPE):

$$fitness = MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (19)$$

式中:  $\hat{y}_i$  为第  $i$  时刻的预测值;  $y_i$  为第  $i$  时刻的真实值;  $n$  为样本量。

为了更具有统计意义,采用5折交叉验证,完成模型训练和参数优化。图6为适应度变化曲线

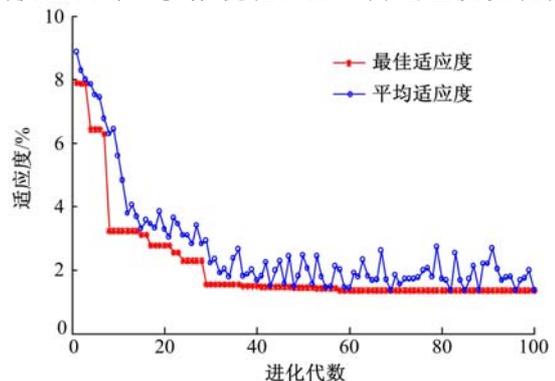


图6 PSO寻优适应度曲线

Fig. 6 Fitness curve of PSO

线,模型的优化参数组合:  $\gamma = 76.63, \sigma = 8.57, \rho = 0.83, q = 3$ 。

### 4.3 预测结果及分析

为了更好地分析 SSA-CKF-LSSVM 模型的预测效果,引入 ARMA 模型、CKF-LSSVM 模型和 SSA-BPNN 模型作对比。其中,ARMA 模型使用前提是预测平稳时间序列。对于非平稳序时间序列,须要通过差分将其转化为平稳时间序列。其他几个模型均属于智能学习模型,需要较多的训练样本。使用原始交通流量数据构建 ARMA 模型和 CKF-LSSVM 模型。使用 SSA 降噪后交通流量数据构建 SSA-BPNN 模型。其中,CKF-LSSVM 模型参数的优化同样使用 PSO 算法。预测效果评价指标包括:平均绝对误差(MAE)、平均相对百分比误差(MAPE)、均方误差(MSE)和均等系数(EC),计算公式如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (20)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (21)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (22)$$

$$EC = 1 - \frac{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\sqrt{\sum_{i=1}^n y_i^2} + \sqrt{\sum_{i=1}^n \hat{y}_i^2}} \quad (23)$$

MAE、MAPE、MSE 越小,说明预测误差越小,预测效果越好。均等系数 EC 越接近于 1,说明预测值与实际值之间拟合度越好,实际值与预测值之间有着更为相似的演化趋势。

各模型预测效果曲线如图 7 所示,绝对误差曲线如图 8 所示,预测性能的评价指标值如表 1 所示。由图 7 可知,SSA-CKF-LSSVM 模型预测曲线与实测曲线的拟合效果最佳;模型的 EC 值为 0.96(均值),最接近于 1,同样说明拟合效果最佳;而 ARMA 模型 EC 值仅为 0.87(均值),这是因为该预测模型具有一定的滞后性。由图 8 可知,SSA-CKF-LSSVM 模型的绝对误差波动范围最小,表明该模型预测稳定性较好,特别是在交通流变化相对剧烈时(如 7:00~9:00 时段)。由表 1 可知,SSA-CKF-LSSVM 模型的三项误差(MAE、MAPE、MSE)均低于其他模型。综上,基于奇异谱分析和组合核函数 LSSVM 的预测模型预测性能良好,能提高短时交通流量的预测精度。

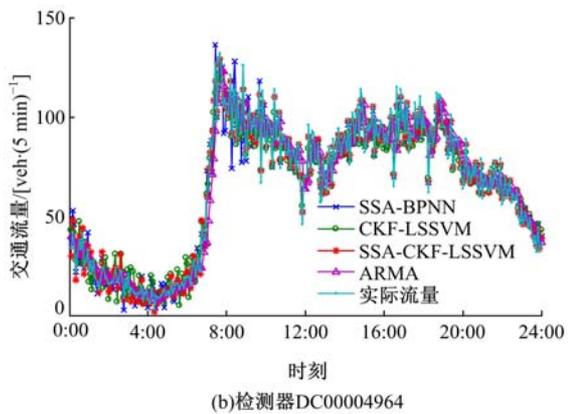
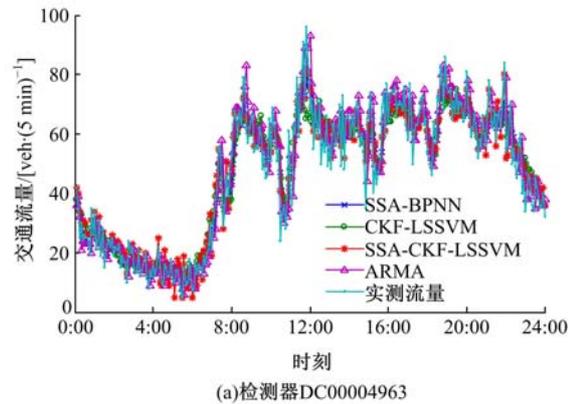


图 7 各模型交通流量的预测效果

Fig. 7 Effect of traffic flow prediction models

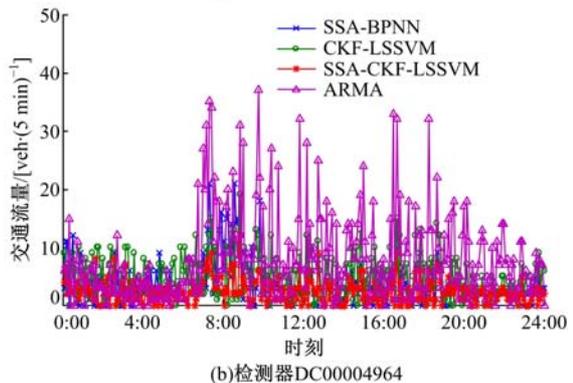
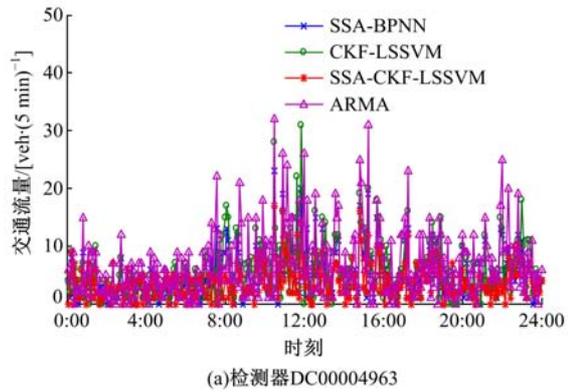


图 8 各模型预测的绝对误差

Fig. 8 Absolute error of traffic flow prediction models

表 1 各预测模型的评价指标值  
Table 1 Performance comparison of prediction models

模型	检测器 DC00004963				检测器 DC00004964			
	MAE/ [veh · (5 min) <sup>-1</sup> ]	MAPE/ %	MSE/ [veh · (5 min) <sup>-1</sup> ] <sup>2</sup>	EC	MAE/ [veh · (5 min) <sup>-1</sup> ]	MAPE/ %	MSE/ [veh · (5 min) <sup>-1</sup> ] <sup>2</sup>	EC
SSA-BPNN	4.78	4.98	38.33	0.94	4.52	7.54	25.92	0.95
CKF-LSSVM	5.66	5.90	55.14	0.93	5.89	8.61	64.95	0.91
SSA-CKF-LSSVM	3.75	3.94	23.42	0.96	3.61	5.92	16.30	0.96
ARMA	7.90	8.19	92.11	0.86	8.08	10.3	102.17	0.88

## 5 结束语

针对交通流数据含有噪声成分,短时交通流预测精度不高的问题,提出了基于奇异谱分析和组合核函数最小二乘支持向量机的短时交通预测模型(SSA-CKF-LSSVM)。运用奇异谱分析方法,将交通流时间序列分解和重构,消除了原序列的噪声成分。使用降噪后的交通流数据对组合核函数最小二乘支持向量机进行训练,并通过 PSO 算法确定模型参数。以厦门市莲前西路的实测交通流数据为例,验证了新建模型的预测效果并与其他 3 种模型的预测结果进行对比分析,结果表明新建模型预测性能良好,能够有效提高交通流预测精度。

### 参考文献:

[ 1 ] 邴其春,杨兆升,周熙阳,等. 基于向量误差修正模型的短时交通参数预测[J]. 吉林大学学报:工学版,2015,45(4):1076-1081.  
Bing Qi-chun, Yang Zhao-sheng, Zhou Xi-yang, et al. Short-term traffic parameters prediction method based on vector error correction model[J]. Journal of Jilin University(Engineering and Technology Edition), 2015,45(4):1076-1081.

[ 2 ] van Lint J W C, Hoogendoorn S P, van Zuylen H J. Accurate freeway travel time prediction with state-space neural networks under missing data [J]. Transportation Research Part C: Emerging Technologies,2005,13(5):347-369.

[ 3 ] Stathopoulos A, Karlaftis M G. A multivariate state space approach for urban traffic flow modeling and prediction[J]. Transportation Research Part C: Emerging Technologies,2003,11(2):121-135.

[ 4 ] Sun H, Liu H, Xiao H, et al. Use of local linear regression model for short-term traffic forecasting[J]. Transportation Research Record: Journal of the Transportation Research Board, 2003, 1836: 143-

150.

[ 5 ] Chandra S R, Al-Deek H. Predictions of freeway traffic speeds and volumes using vector autoregressive models[J]. Journal of Intelligent Transportation Systems,2009,13(2):53-72.

[ 6 ] Wang Y, Papageorgiou M. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach[J]. Transportation Research Part B: Methodological,2005,39(2):141-167.

[ 7 ] 龚勃文,林赐云,李静,等. 基于核自组织映射-前馈神经网络的交通流短时预测[J]. 吉林大学学报:工学版,2011,41(4):938-942.  
Gong Bo-wen, Lin Ci-yun, Li Jing, et al. Short-time traffic flow prediction based on KSOM-BP neural network[J]. Journal of Jilin University(Engineering and Technology Edition), 2011,41(4):938-942.

[ 8 ] Hong W C, Dong Y C, Zheng F F, et al. Forecasting urban traffic flow by SVR with continuous ACO [J]. Applied Mathematical Modelling, 2011,35(3): 1282-1291.

[ 9 ] Suykens J A K, van Gestel T, Brabanter D E. Least Squares Support Vector Machines[M]. Singapore: World Scientific, 2002.

[10] Xiao Y, Liu J J, Hu Y, et al. A neuro-fuzzy combination model based on singular spectrum analysis for air transport demand forecasting[J]. Journal of Air Transport Management,2014,39:1-11.

[11] 罗勇,匡翠林,卢辰龙,等. 基于 SSA 的 GPS 坐标序列去噪及季节信号提取[J]. 大地测量与地球动力学,2015,35(3):391-395.  
Luo Yong, Kuang Cui-lin, Lu Chen-long, et al. GPS coordinate series denoising and seasonal signal extraction based on SSA[J]. Journal of Geodesy and Geodynamics,2015,35(3):391-395.

[12] Smits G F, Jordan E M. Improved SVM regression using mixtures of kernels[C]// Proceedings of the 2002 International Joint Conference, Honolulu, NJ, 2002:2785-2790.