

DOI: 10.11918/j.issn.0367-6234.201703102

交叉口短时流量 CEEMDAN-PE-OSELM 预测模型

田秀娟¹, 于德新^{1,2}, 邢雪³, 商强¹, 王树兴⁴(1. 吉林大学 交通学院, 长春 130022; 2. 吉林省道路交通重点实验室, 长春 130022;
3. 吉林化工学院 信息与控制工程学院, 吉林 吉林 132022; 4. 山东高速股份有限公司, 济南 250000)

摘要: 为提高交叉口短时交通流预测精度, 以历史交通流量数据为基础, 提出一种基于自适应噪声完整集成经验模态分解 (complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN) - 排列熵 (permutation entropy, PE) - 在线序贯极限学习机 (online sequential extreme learning machine, OSELM) 组合预测模型 (CEEMDAN-PE-OSELM)。首先对交通流历史时间序列数据进行 CEEMDAN 分解, 得到多个本征模态函数 (intrinsic mode function, IMF) 分量; 通过 PE 算法对 IMF 分量进行重组, 形成具有复杂度差异的重组子序列。然后, 分别构建重组子序列 OSELM 预测模型, 将预测结果相加得到最终预测流量。最后选取一实际交叉口, 进行模型验证分析。结果表明: CEEMDAN-PE-OSELM 模型的 MAE、MAPE 和 MSE 的值均低于其他模型, 预测误差最小; EC 值为 0.963, 高于 ARIMA 模型的 EC 值 (0.898), 最接近于 1, 预测精度最高, 稳定性最好。就同一预测模型而言, 经过 CEEMDAN-PE 处理的模型的各项误差明显降低, 预测精度有所提高。

关键词: 时间序列; 短时流量预测; 组合预测; 经验模态分解; 极限学习机

中图分类号: U491

文献标志码: A

文章编号: 0367-6234(2018)03-0083-07

Prediction model of CEEMDAN-PE-OSELM for intersections short-term traffic flow

TIAN Xiujuan¹, YU Dexin^{1,2}, XING Xue³, SHANG Qiang¹, WANG Shuxing⁴

(1. College of Transportation, Jilin University, Changchun 130022, China; 2. Jilin Province Key Laboratory of Road Traffic, Changchun 130022, China; 3. College of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022, Jilin, China; 4. Shandong High-Speed Company Limited, Jinan 250000, China)

Abstract: To improve the prediction accuracy of intersection short-term traffic flow, a new CEEMDAN-PE-OSELM model is developed based on the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), permutation entropy (PE) and online sequential extreme learning machine (OSELM). Firstly, traffic flow historical time series are decomposed by CEEMDAN algorithm. Secondly, PE algorithm is used to recombine the IMF components obtained by CEEMDAN, and a series of restructured subsequences can be obtained, which have a significant difference in terms of complexity. Then, the OSELM prediction models are proposed for each restructured subsequence respectively, and the final results are got by adding the prediction results. Finally, a typical intersection is verified the effect and performance of the hybrid prediction model. Results show that the values of MAE, MAPE and MSE of CEEMDAN-PE-OSELM prediction model are lower than other models, and get a minimal error. The EC value of the improved model is 0.963, which is higher than that of ARIMA model (0.898) and the most close to 1. The CEEMDAN-PE-OSELM prediction model has the highest precision and best stability, and the errors decrease obviously.

Keywords: time series; short-term traffic prediction; hybrid prediction; empirical mode decomposition; extreme learning machine

作为智能交通系统 (intelligent transportation systems, ITS) 重要子系统, 交通控制和诱导系统皆以交通状态预测为前提, 而短时交通流量预测是交通状态预测的基础。通常, 信号控制周期小于 3 min, 诱导周期为 5 min。虽然采集设备日趋完善, 但对于

高效的控制和诱导系统而言, 检测设备采集的信息仍具有一定的滞后性。此外, ITS 快速发展对交通流预测提出了更高的要求。因此, 研究高效准确的短时 (通常为 5 min) 交通流预测方法具有重要意义。针对短时交通流预测, 国内外进行了大量研究, 建立了多种模型, 大体分为 4 类: 1) 统计模型, 如时间序列^[1]、Klaman 模型^[2]等。2) 非线性模型, 如混沌理论^[3]、小波理论模型^[4]等。3) 智能模型, 如神经网络^[5]、支持向量机模型^[6]等。4) 组合预测模型, 如 AFSA-LSSVM^[7]、PSR-RELM 模型^[8]等。混沌理论、

收稿日期: 2017-03-21

基金项目: 国家科技支撑计划项目 (2014BAG03B03);

山东省省管企业科技创新项目 (20122150251-1)

作者简介: 田秀娟 (1990—), 女, 博士研究生;

于德新 (1972—), 男, 教授, 博士生导师

通信作者: 于德新, yudx@jlu.edu.cn

神经网络及支持向量机在分析复杂非线性不确定系统方面具有明显优势。但传统神经网络多采用梯度下降法,易陷入局部最优,且学习速度慢。极限学习机(extreme learning machine, ELM)^[9]学习速度快,泛化性能好,广泛应用于工程问题。OSELM 算法^[10]是对 ELM 的改进,分批训练历史数据,支持可变样本数。交通流量数据在实际采集和传输中,缺失与错误不可避免,应根据情况分批训练,因此,OSELM 算法能满足交通流量预测要求。此外,交叉口交通流数据具有非平稳性,进行先验分析有助于提高预测精度。CEEMDAN^[11]是对 EEMD^[12]的改进,提高分解效率同时重构误差几乎为零。基于现有研究,本文提出一种基于 CEEMDAN-PE-OSELM 的组合预测模型,以提高交叉口短时流量预测精度。

1 CEEMDAN-PE 数据分解与重组

对交叉口流量时间序列进行 CEEMDAN 分解;然后利用 PE 算法计算 IMF 分量的排列熵值,对序列重组。

1.1 CEEMDAN 方法

EMD 方法是按照不同波动尺度对原始序列进行分解,得到不同幅值的 IMF 分量^[13]。CEEMDAN 方法则通过在每一分解阶段添加自适应白噪声,计算唯一余量以获取 IMF 分量,分解过程完整^[11]。

记原始交通流量时间序列为 $s(n)$, 第 $i(1, 2, \dots, I)$ 次添加的标准正态分布白噪声序列为 $v^i(n)$, 则第 i 次分解交通流量序列可以表示为 $s^i(n) = s(n) + v^i(n)$ 。定义 $E_k(\cdot)$ 和 $F_{IM,k}$ 分别为第 k 个由 EMD 和 CEEMDAN 产生的模态分量。CEEMDAN 步骤如下。

步骤 1 与 EEMD 方法^[12]相同, CEEMDAN 也是针对交通流时间序列 $s(n) + \varepsilon_0 v^i(n)$ 进行 I 次分解。因此,可以通过 EMD 方法获取第 1 个模态分量为

$$F_{IM,1}(n) = \frac{1}{I} \sum_{i=1}^I F_{IM,1}^i(n) = \overline{F_{IM,1}(n)}. \quad (1)$$

步骤 2 在第 1 阶段($k=1$), 计算第 1 个唯一的余量流量序列为

$$r_1(n) = s(n) - F_{IM,1}(n). \quad (2)$$

步骤 3 对流量序列 $r_1(n) + \varepsilon_1 E_1(v^1(n))$ 进行 $i(1, 2, \dots, I)$ 次分解,直到得到第 1 个 EMD 流量模态分量为止。同理,可得到第 2 个模态分量为

$$F_{IM,2}(n) = \frac{1}{I} \sum_{i=1}^I E_1(r_1(n)) + \varepsilon_1 E_1(v^1(n)). \quad (3)$$

步骤 4 类似地, 计算其他阶段的第 $k(2, 3, \dots, K)$ 个余量序列;然后依据步骤 3, 得到第 $k+1$ 个模态分量,即

$$r_k(n) = r_{k-1}(n) - F_{IM,k}(n), \quad (4)$$

$$F_{IM,k+1}(n) = \frac{1}{I} \sum_{i=1}^I E_1(r_k(n)) + \varepsilon_k E_k(v^i(n)). \quad (5)$$

步骤 5 当余量序列不能继续分解时,即余量信号的极值点个数至多不超过两个,算法结束。此时,共得到 K 个模态分量,余量序列的最终结果为

$$R(n) = s(n) - \sum_{k=1}^K F_{IM,k}. \quad (6)$$

则原始交通流量时间序列 $s(n)$ 最终被分解为

$$\sum_{k=1}^K F_{IM,k} \text{ 和 } R(n), \text{ 即 } s(n) = \sum_{k=1}^K F_{IM,k} + R(n).$$

综上, CEEMDAN 分解过程完整,能够对原始交通流量时间序列精确分解。

1.2 PE 算法

PE 算法^[14]可以衡量系统复杂度,表征时间序列的随机性和动力学突变特性。概念简单,计算速度快,抗干扰能力强,特别适用于非线性数据,对时间序列变化的敏感性很高,并且具有很好的鲁棒性^[15],已广泛应用于各种时间序列。

交叉口流量数据具有一定的随机性和非平稳性,使得 CEEMDAN 分解得到的交通流 IMF 分量较多。因此,为了减小计算规模,本文采用 PE 算法进行 IMF 重组,步骤如下。

步骤 1 首先对 CEEMDAN 得到的各 IMF 流量序列 $\{X(i), i=1, 2, \dots, N\}$ 进行相空间重构^[16],得到相空间矩阵

$$Y = \begin{bmatrix} x(1) & x(1+\tau) & \cdots & x(1+(m-1)\tau) \\ \vdots & \vdots & \vdots & \vdots \\ x(j) & x(j+\tau) & \cdots & x(j+(m-1)\tau) \\ \vdots & \vdots & \vdots & \vdots \\ x(K) & x(K+\tau) & \cdots & x(K+(m-1)\tau) \end{bmatrix}. \quad (7)$$

式中 m 和 τ 分别表示嵌入维数和延迟时间; $j=1, 2, \dots, K$ 。

步骤 2 将上述矩阵的每行作为一个分量,则共有 K 个分量,存在关系 $K = n - (m-1)\tau$ 。以第 j 个分量 $(x(j), x(j+\tau), \dots, x(j+(m-1)\tau))$ 为例,按照各元素数值大小,进行升序排列,即

$$x(i+(j_1-1)\tau) \leq x(i+(j_2-1)\tau) \leq \cdots \leq x(i+(j_m-1)\tau). \quad (8)$$

式中 j_1, j_2, \dots, j_m 表示各分量元素所在列序号。如存在 $x(i+(j_1-1)\tau) = x(i+(j_2-1)\tau)$, 则按照 j_1 和 j_2 值大小排序。综上,对于任意流量时间序列 $X(i)$ 的相空间重构矩阵的每一行向量都可以得到一组序列 $S(l)$ 为

$$S(l) = \{j_1, j_2, \dots, j_m\}. \quad (9)$$

式中 $l=1,2,\dots,k$, 且存在 $k \leq m!$, m 维相空间共有 $m!$ 种不同的映射符号序列.

步骤3 计算每种符号序列 $S(l)$ 出现的概率 P_1, P_2, \dots, P_k . 按照 Shannon 熵的形式, 可以将流量时间序列 $X(i)$ 的第 k 种不同符号序列的排列熵 $H_p(m)$ 定义为

$$H_p(m) = - \sum_{j=1}^k P_j \ln P_j. \quad (10)$$

由上式可知, 当 $P_j = 1/m!$ 时, $H_p(m)$ 达到最大值 $\ln(m!)$. 为了方便, 将 $H_p(m)$ 进行标准化处理

$$H_{PE}(m) = H_p(m) / \ln(m!), \quad (11)$$

式中 $0 \leq H_{PE}(m) \leq 1$, 大小表示时间序列的随机程度. 值越大, 随机性越强; 反之, 随机性越弱. 嵌入维数 m 和延迟时间 τ 是 PE 算法两个重要参数. 研究表明, 当样本量较小时, 嵌入维数 m 通常较小; 对于样本量大于 1 024 的时间序列来说, 当 $m = 6$ 时, 即可获得稳定的排列熵值; 而延迟时间 τ 对于排列熵的影响较小^[17].

通过以上步骤, 计算各 IMF 排列熵. 依据排列熵值接近程度, 对 IMF 分量合并, 得到重组子序列.

2 CEEMDAN-PE-OSELM 组合预测模型构建

将重组子序列作为输入数据构建 OSELM 预测模型, 将多个预测结果相加得到最终预测流量.

2.1 OSELM 算法

在 OSELM 算法训练中, 仅需输入当前批数据更新网络权值, 无需重复扫描历史数据, 泛化能力强. 分为初始化和序贯学习两个阶段, 步骤如下.

步骤1 从给定的交叉口交通流量训练集 $D = \{(x_i, t_i) \mid x_i \in R^n, t_i \in R^m, i = 1, L, N\}$ 中选取少量的流量数据 $D_0 = \{(x_i, t_i) \mid i = 1, N_0\}$, 初始化, 其中 L 为隐含层节点数, N_0 满足 $N_0 \geq L$.

步骤2 随机生成隐含层节点参数 $(a_i, b_i), i = 1, 2, \dots, L$; 计算得到初始隐含层输出矩阵为

$$H_0 = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \cdots & \vdots \\ G(a_1, b_1, x_{N_0}) & \cdots & G(a_L, b_L, x_{N_0}) \end{bmatrix}_{N_0 \times L}. \quad (12)$$

步骤3 计算初始输出权重 $\beta^{(0)} = P_0 H_0^T T_0$, 其中 $P_0 = (H_0^T T_0)^{-1}, T_0 = [t_1, \dots, t_{N_0}]^T$.

步骤4 初始化加入网络的样本批次 $k = 0$.

步骤5 计算新加入的第 $k + 1$ 批训练数据

$$D_{k+1} = \{(x_i, t_i) \mid i = (\sum_{j=0}^k N_j) + 1, \dots, (\sum_{j=0}^{k+1} N_j)\}$$

的隐含层输出矩阵 H_{k+1} , 其中 N_{k+1} 为第 $k + 1$ 批新样本的个数.

$$H_{k+1} = \begin{bmatrix} G(a_1, b_1, x_{(\sum_{j=0}^k N_j)+1}) & \cdots & G(a_L, b_L, x_{(\sum_{j=0}^k N_j)+1}) \\ \vdots & \cdots & \vdots \\ G(a_1, b_1, x_{\sum_{j=0}^k N_j}) & \cdots & G(a_L, b_L, x_{\sum_{j=0}^k N_j}) \end{bmatrix}_{N_{k+1} \times L}. \quad (13)$$

步骤6 计算更新输出权重 $\beta^{(k+1)}$ 和 P_{k+1} .

$$P_{k+1} = P_k - P_k H_{k+1}^T (I + H_{k+1} P_k H_{k+1}^T)^{-1} H_{k+1} P_k, \quad (14)$$

$$\beta^{(k+1)} = \beta^{(k)} + P_{k+1} H_{k+1}^T (T_{k+1} - H_{k+1} \beta^{(k)}). \quad (15)$$

其中 $T_{k+1} = [t_{(\sum_{j=0}^k N_j)+1}, \dots, t_{\sum_{j=0}^{k+1} N_j}]^T$. 当逐个获得训练数据时, 有 $N_{k+1} = 1$ 存在, 式(14)、(15)可以分别简化为

$$P_{k+1} = P_k - \frac{P_k h(x_{k+1}) h^T(x_{k+1}) P_k}{1 + h^T(x_{k+1}) P_k h(x_{k+1})}, \quad (16)$$

$$\beta^{(k+1)} = \beta^{(k)} + P_{k+1} h(x_{k+1}) (t_{k+1}^T - h^T(x_{k+1}) \beta^{(k)}). \quad (17)$$

式中 $h(x_{k+1}) = [G(a_1, b_1, x_{k+1}) \cdots G(a_L, b_L, x_{k+1})]$.

步骤7 令 $k = k + 1$, 返回步骤5继续执行, 直至算法结束.

其中, 步骤1至步骤4为初始化阶段, 步骤5至步骤7为序贯学习阶段. 当 $\text{rank}(H_0) = L$ 时, OSELM 和 ELM 算法学习能力相同.

2.2 组合预测模型构建

构建交叉口短时流量 CEEMDAN-PE-OSELM 组合预测模型, 流程如图1所示.

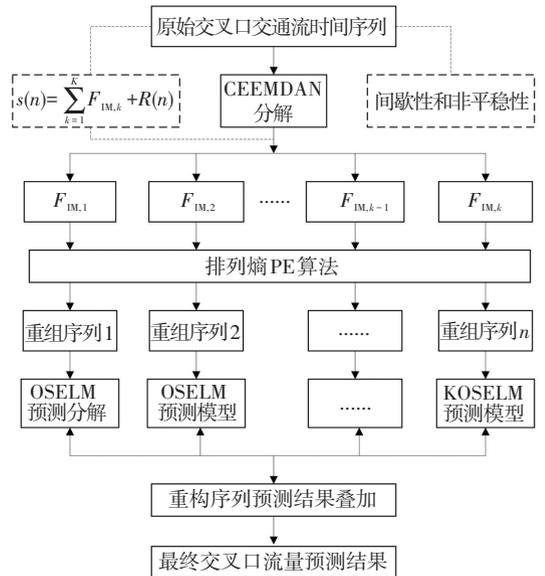


图1 CEEMDAN-PE-OSELM 交叉口流量预测流程

Fig.1 Flowchart of intersection traffic flow prediction model based on CEEMDAN-PE-OSELM

为提高预测精度, 在建立预测模型之前将流量重组序列进行归一化处理, 即

$$X'(i) = \frac{(X(i) - \min\{X(i)\})}{(\max\{X(i)\} - \min\{X(i)\})} \quad (18)$$

反之,可以反归一化处理,即

$$X(i) = X'(i) \cdot (\max\{X(i)\} - \min\{X(i)\}) + \min\{X(i)\} \quad (19)$$

式中 $X'(i)$ 为归一化流量数据, $X(i)$ 为重组流量序列数据.

3 实例验证

选取一典型十字交叉口,进行短时流量预测.该交叉口几何特性如图 2 所示.

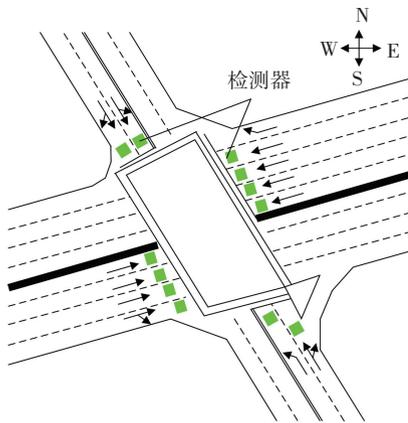


图 2 交叉口几何特性

Fig.2 Geometry characteristics of the intersection

采集东进口连续 5 个工作日的 5 min 流量数据,如图 3 所示. 经过预处理,得到 1 440 个可用数据点. 前 4 个工作日的 1 152 个数据点用于预测模型建立,第 5 个工作日的的数据用于模型验证.

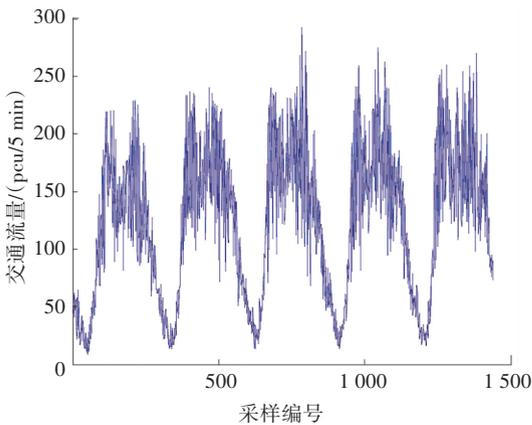


图 3 交叉口东进口 5 min 流量分布

Fig.3 Traffic flow distribution of the eastern approach at the interval of 5 min

3.1 交通流量序列 CEEMDAN-PE 分解重组

通过 MATLAB 软件编程,对原始交叉口流量序

列进行 CEEMDAN 分解,分解时添加 $I = 500$ 组标准差为 0.2 的白噪声信号,结果如图 4 所示. 图 4 中交叉口流量原始时间序列被分解为 10 个随机性不同的 IMF 分量,相邻 IMF 分量波动程度相似.

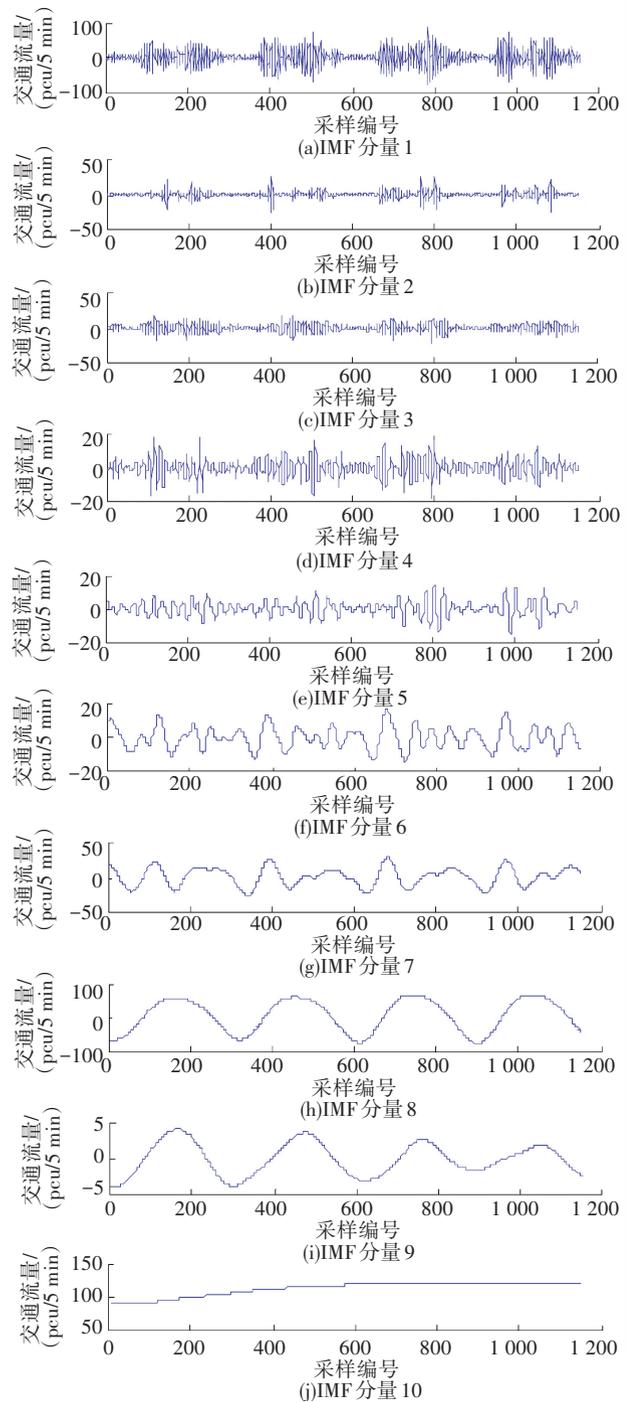


图 4 交叉口流量数据 CEEMDAN 分解结果

Fig. 4 Intersection traffic data decomposition results by CEEMDAN algorithm

图 5 为流量序列 CEEMDAN 分解百分比误差图示. 由图 5 可知,该流量时间序列的分解误差非常小,达到了 10^{-14} 级;说明了交叉口流量序列得到了完全分解,证明了该方法的有效性.

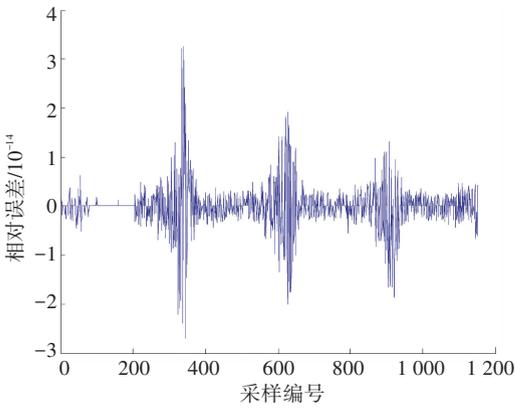


图5 CEEMDAN 分解相对误差

Fig.5 Relative percentage error of CEEMDAN

根据前文,取嵌入维数 $m = 6$,延迟时间 $\tau = 3$,计算 $F_{IM,k}$ 排列熵值,结果见表1. $H_p(m)$ 为 $F_{IM,k}$ 排列熵值, $H_{PE}(m)$ 为标准化排列熵值, $\Delta H_{PE}(m)$ 为相邻分量标准排列熵值之差. 各 IMF 分量标准化排列熵值分布如图6所示.

表1 各 IMF 分量的排列熵值

Tab.1 Permutation entropy values of each IMF component

$F_{IM,k}$	$H_p(m)$	$H_{PE}(m)$	$\Delta H_{PE}(m)$
$F_{IM,1}$	5.509	0.837	0.015
$F_{IM,2}$	5.410	0.822	0.043
$F_{IM,3}$	5.123	0.779	0.128
$F_{IM,4}$	4.280	0.651	0.154
$F_{IM,5}$	3.271	0.497	0.101
$F_{IM,6}$	2.605	0.396	0.077
$F_{IM,7}$	2.099	0.319	0.065
$F_{IM,8}$	1.670	0.254	0.018
$F_{IM,9}$	1.551	0.236	0.131
$F_{IM,10}$	0.693	0.105	—

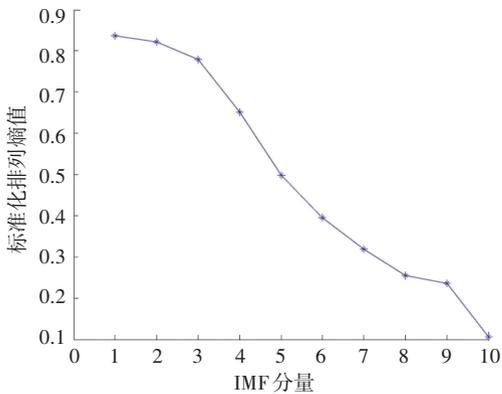


图6 各 IMF 分量标准化排列熵值

Fig.6 Standardized permutation entropy values of each IMF component

将 $\Delta H_{PE}(m)$ 值小于 0.1 的相邻 $F_{IM,k}$ 合并,且综合比较同一 $F_{IM,k}$ 与前后 $F_{IM,k}$ 之间的差值,保证运

行效率和结果精确度. 由表1和图6可知,随机性最强的3个分量为 $F_{IM,1}$ 、 $F_{IM,2}$ 和 $F_{IM,3}$,对预测结果影响最大,且标准化 PE 值相差甚小,因此将 $F_{IM,1}$ 、 $F_{IM,2}$ 和 $F_{IM,3}$ 合并; $F_{IM,4}$ 和 $F_{IM,5}$ 虽然相邻,但标准化 PE 值与相邻的其他分量相差较大,因此作为单独分量; $F_{IM,6}$ 和 $F_{IM,7}$ 相邻且标准化 PE 值相差较小,可合并; $F_{IM,8}$ 和 $F_{IM,9}$ 的标准化 PE 值差异为 0.018,将其合并. 由于 $F_{IM,9}$ 与 $F_{IM,10}$ 的标准化 PE 值之差和 $F_{IM,8}$ 与 $F_{IM,9}$ 的标准化 PE 值之差相差较大,为了保证精确性,不进行合并. 综上,可以得到6组重组序列: $F_{IM,1} + F_{IM,2} + F_{IM,3}$, $F_{IM,4}$, $F_{IM,5}$, $F_{IM,6} + F_{IM,7}$, $F_{IM,8} + F_{IM,9}$, $F_{IM,10}$. 重组序列分布如图7所示.

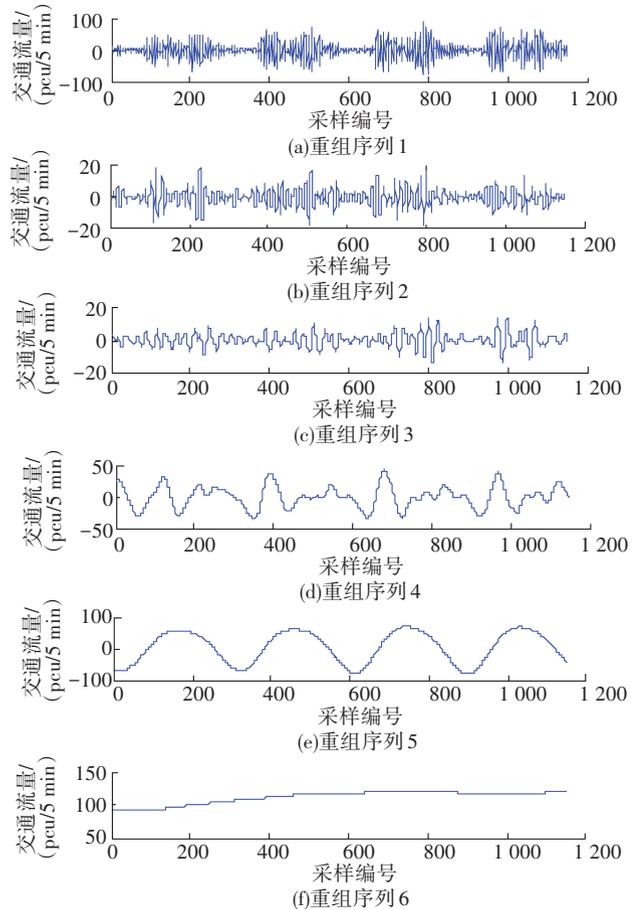


图7 经 CEEMDAN-PE 处理的交叉口流量重组序列

Fig.7 Restructured intersection traffic flow subsequences obtained by CEEMDAN-PE

3.2 预测结果分析

采用滚动单步预测,模型输入为预测点前两个小时流量序列,输出为预测值序列. 则 1 152 个数据可构建 1 128 组输入-输出数据,组成训练集. 分别构建 ARIMA、BP、ELM、OSELM、CEEMDAN-PE-BP、CEEMDAN-PE-ELM 和 CEEMDAN-PE-OSELM,对比预测效果.

OSELM 与 ELM 需要提前确定隐含层节点数 L ; 选取 $N_0 = L$ 个样本数据作为初始化阶段训练样

本,用于增量学习的样本数据量为 $1\ 128 - N_0$. 经过交叉验证试验,得到最优 L 为 30.

预测评价指标选取平均绝对误差 E_{MA} 、平均相对误差 P_{MAE} 、均方误差 S_{ME} 和均等系数 C_E 分别为

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

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

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

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

其中 \hat{y}_i, y_i 分别为第 i 时刻的预测值和真实值, n 为

样本量. E_{MA} 、 P_{MAE} 、 S_{ME} 越小,误差越小; C_E 越接近 1,预测效果越好.

各模型的预测效果曲线如图 8 所示,预测性能评价指标见表 2.

表 2 各预测模型性能指标对比

Tab.2 Comparisons of prediction performance indexes for each prediction model

模型	E_{MA}	$P_{MAE}/\%$	S_{ME}	C_E
ARIMA	22.09	17.13	872.59	0.898
BP	20.39	15.50	723.37	0.907
ELM	17.75	13.88	576.71	0.917
OSELM	14.68	11.73	386.34	0.932
CEEMDAN-PE-BP	13.52	9.97	311.91	0.940
CEEMDAN-PE-ELM	11.25	8.44	223.65	0.949
CEEMDAN-PE-OSELM	8.65	6.51	114.33	0.963

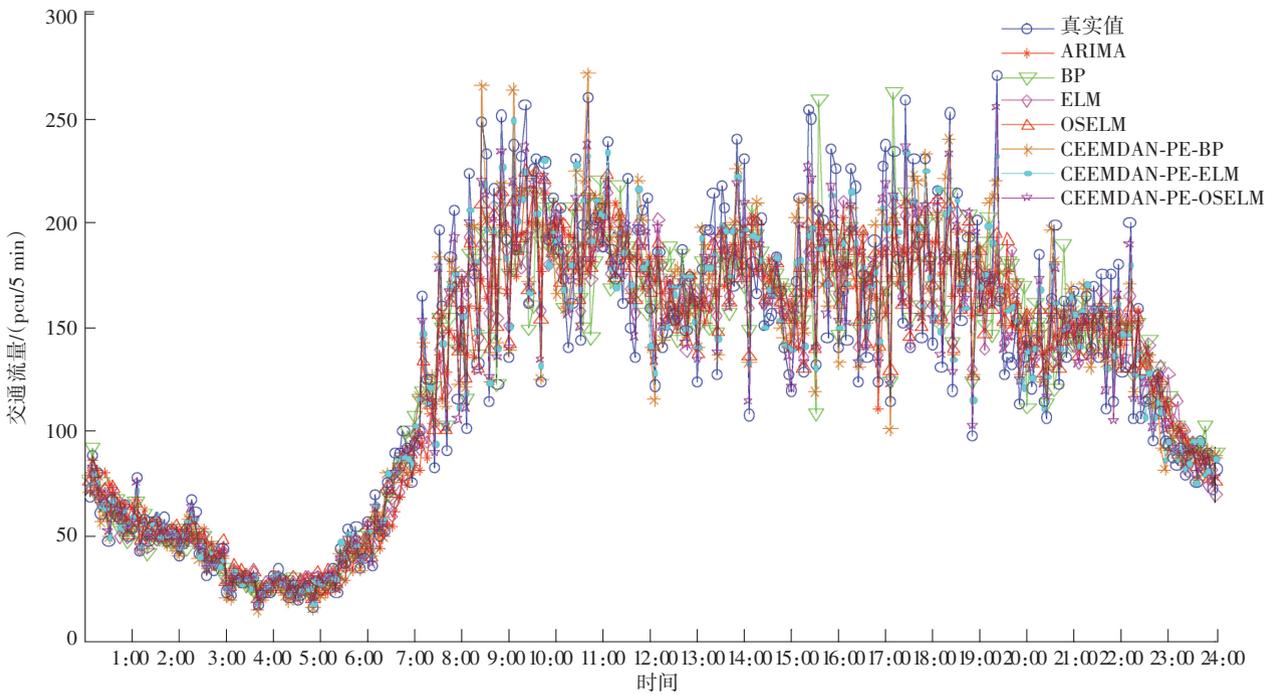


图 8 各模型预测结果

Fig.8 Prediction results of each model

由图 8 可知,与其他预测模型相比,本文 CEEMDAN-PE-OSELM 模型流量预测值与真实值最为接近. ARIMA 模型与真实值的拟合程度最低,与真实值的偏离程度最大;BP 模型、ELM 模型和 OSELM 模型的预测值与真实值的拟合程度逐渐提高,但是相比于经过序列分解与重组的 CEEMDAN-PE-BP 模型、CEEMDAN-PE-ELM 模型及 CEEMDAN-PE-OSELM 模型,与真实值的偏离程度仍然很大.

由表 2 可知,CEEMDAN-PE-OSELM 模型的

E_{MA} 、 P_{MAE} 和 S_{ME} 的值均低于其他模型,说明其预测误差最小,预测精度最高;且 C_E 值为 0.963,高于 ARIMA 模型的 C_E 值(0.898),相比于其他方法, C_E 值最接近于 1,说明其预测值与真实值的偏离程度最小,预测效果最好,模型的稳定性最好. 就同一预测模型而言,经过 CEEMDAN-PE 处理的模型的各项误差明显降低,预测精度有所提高.

综上,基于 CEEMDAN-PE-OSELM 的交叉口短时交通流预测模型性能较好,能够满足预测要求.

4 结 论

1) 基于交叉口历史流量时间序列,提出了一种基于 CEEMDAN-PE-OSELM 的交叉口短时交通流量预测方法,该方法与其他经典预测模型相比,模型流量预测值与真实值最为接近,平均绝对误差、平均相对误差、均方误差均最小。

2) 预测模型的均等系数为 0.963,高于 ARIMA 模型的均等系数 0.898;且与其他模型相比,均等系数最接近 1,与真实值的偏离程度最小,预测效果最好。

3) 经过 CEEMDAN-PE 分解与重组的模型的各项误差明显降低,预测精度有所提高。

参考文献

- [1] LINGRAS P, SHARMA S C, OSBORNE P, et al. Traffic volume time-series analysis according to the type of road use [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2000, 15(5): 365-373.
- [2] GAO Junwei, LENG Ziwen, ZHANG Bin, et al. The application of adaptive Kalman filter in traffic flow forecasting [J]. *Advanced Materials Research*, 2013, 680: 495-500.
- [3] 张玉梅,曲仕茹,温凯歌. 交通流量的混沌特性分析及预测模型研究[J]. *土木工程学报*, 2009, 42(1): 119-123.
ZHANG Yumei, QU Shiru, WEN Kaige. Chaotic property analysis and prediction model study for traffic flow time series [J]. *China Civil Engineering Journal*, 2009, 42(1): 119-123.
- [4] 丁恒,郑小燕,刘燕,等. 基于小波分析的随机交通流组合预测方法研究[J]. *系统仿真学报*, 2012, 24(2): 377-381.
DING Heng, ZHENG Xiaoyan, LIU Yan, et al. Stochastic traffic series time combination forecasting based on wavelet analysis [J]. *Journal of System Simulation*, 2012, 24(2): 377-381.
- [5] JIANG X, ADELI H. Dynamic wavelet neural network model for traffic flow forecasting [J]. *Journal of Transportation Engineering*, 2005, 131(10): 771-779.
- [6] 杨兆升,王媛,管青. 基于支持向量机方法的短时交通流量预测方法[J]. *吉林大学学报(工学版)*, 2006, 36(6): 881-884.
YANG Zhaosheng, WANG Yuan, GUAN Qing. Short-time traffic prediction method based on SVM [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2006, 36(6): 881-884.
- [7] 刘静. 基于 AFSA-LSSVM 的短时交通流量预测[J]. *计算机工程与应用*, 2013, 49(17): 226-229.

- LIU Jing. Short-time traffic prediction model based on LSSVM optimized by artificial fish swarm algorithm [J]. *Computer Engineering and Applications*, 2013, 49(17): 226-229.
- [8] 商强,杨兆升,李志林,等. 基于相空间重构和 RELM 的短时交通流量预测[J]. *华南理工大学学报(自然科学版)*, 2016, 44(4): 109-114.
SHANG Qiang, YANG Zhaosheng, LI Zhilin, et al. Short-term traffic flow prediction based on phase space reconstruction and RELM [J]. *Journal of South China University of Technology (Natural Science Edition)*, 2016, 44(4): 109-114.
- [9] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications [J]. *Neuro Computing*, 2006, 70(1/2/3): 489-501.
- [10] LIANG N Y, HUANG G B, SARATCHANDRAN P, et al. A fast and accurate online sequential learning algorithm for feedforward networks [J]. *IEEE Transactions on Neural Networks*, 2006, 17(6): 1411-1423.
- [11] TORRES M E, COLOMINAS M A, SCHLOTTHAUER G, et al. A complete ensemble empirical mode decomposition with adaptive noise [C]//2011 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE, 2011: 4144-4147.
- [12] WU Z H, HUANG N E. Ensemble empirical mode decomposition: A noise-assisted data analysis method [J]. *Advances in Adaptive Data Analysis*, 2009, 1(1): 1-41.
- [13] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. *Proceedings of the Royal Society A—Mathematical Physical and Engineering Sciences*, 1998, 454(1971): 903-995.
- [14] BANDT C, POMPE B. Permutation entropy: a natural complexity measure for time series [J]. *Physical Review Letters, The American Physiological Society*, 2002, 88(17): 174102.
- [15] ZUNINO L, ROSSO O A, SORIANO M C. Characterizing the hyperchaotic dynamics of a semiconductor laser subject to optical feedback via permutation entropy [J]. *IEEE Journal of Selected Topics in Quantum Electronics*, 2011, 17(5): 1250-1257.
- [16] 吕金虎,陆君安,陈士华. 混沌时间序列分析及其应用[M]. 武汉: 武汉大学出版社, 2002.
LÜ Jinhu, LU Jun'an, CHEN Shihua. *Chaotic time series analysis and application* [M]. Wuhan: Wuhan University Press, 2002.
- [17] YAN R, LIU Y, GAO R X. Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines [J]. *Mechanical Systems and Signal Processing*, 2012, 29(1): 474-484.

(编辑 魏希柱)