

基于SVM与数据降噪方法结合的交通流预测

吴础良, 秦玉涛

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2023年6月6日; 录用日期: 2023年8月6日; 发布日期: 2023年8月14日

摘要

准确预测道路交通流是智能交通系统的重要组成部分。由于受到一些外部因素的干扰, 含有噪声的原始交通流数据可能会导致预测性能下降, 为此, 本研究提出了一种称为经验独立模态分解(Empirical Independent Mode Decomposition, EIMD)的降噪方法。模拟实验结果表明EIMD方法优于其他降噪方法。在此基础上, 本文提出了一种结合EIMD与支持向量机(Support Vector Machine, SVM)的组合预测模型, 以提高预测精度。通过使用贵阳市路段的交通数据, 评估了不同模型(SVM、小波神经网络(Wavelet Neural Network, WNN)和卷积神经网络(Convolutional Neural Networks, CNN))结合降噪方法的预测性能。预测结果表明, 结合EIMD降噪方法的模型预测结果均优于未采用降噪方法的模型, 其中, EIMD结合SVM的组合模型预测性能高于其组合模型。研究结果表明, 本研究提出的方法能够降低噪声对预测性能的影响, 准确地预测道路交通流。这对于智能交通系统具有重要意义。

关键词

交通流预测, 降噪算法, 经验模态分解, 独立分量分析, 支持向量机

Traffic Flow Prediction Based on Combination of SVM and Data Noise Reduction Method

Chuliang Wu, Yutao Qin

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Received: Jun. 6th, 2023; accepted: Aug. 6th, 2023; published: Aug. 14th, 2023

Abstract

Ensuring precise forecasting of road traffic flux holds utmost importance in the advancement of

intelligent transportation systems. Nonetheless, the presence of disturbances in the original traffic flux data may impede the accuracy of predictions. In this investigation, we propose an Empirical Independent Mode Decomposition (EIMD) technique to effectively diminish disturbances. Our simulated outcomes showcase the superiority of the EIMD approach when compared to alternative techniques for reducing disturbances. Building upon this, we present a combined forecasting model that integrates EIMD and Support Vector Machine (SVM) to further enhance prediction accuracy. Through the utilization of traffic data from Guiyang City, we assess the performance of diverse combined models, which include Wavelet Neural Networks (SVM), Convolutional Neural Networks (WNN), and Convolutional Neural Networks (CNN), while employing various methods for reducing disturbances. The outcomes unveil that the model incorporating EIMD outperforms the model without disturbance reduction, and the combined model with SVM yields the most favorable prediction performance. These discoveries underline the efficacy of our proposed approach in mitigating the impact of disturbances and accurately forecasting road traffic flux, thereby making a significant contribution to the advancement of intelligent transportation systems.

Keywords

Traffic Flow Prediction, Noise Reduction Method, Empirical Mode Decomposition, Independent Component Analysis, Support Vector Machine

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着汽车数量的大幅攀升, 缓解交通拥堵越来越成为城市发展所要面对的重大课题。构建完善的智能交通预测系统, 是一项重要的应对措施。准确的交通流量预测能够辅助有关部门进行决策, 提前采取相关预警措施, 促进管理体系和管理能力现代化。因此, 研究短时交通流预测具有很强的实践意义[1]。

目前, 本文提出了大量的预测模型, 以通过关注不同的模型结构和计算过程来增强预测性能, 例如, 参数模型[2] [3] [4] [5], 反向传播神经网络(Back Propagation Neural Network, BP) [6], 支持向量机(Support Vector Machine, SVM) [7]和人工神经网络(Artificial Neural Network, ANN) [8]等。这些算法在一些特定的应用中表现出很高的性能。然而, 由于交通系统的时空变化具有不确定性和复杂性, 短时交通流数据在采集过程中常常受到强烈的噪声干扰, 导致交通流预测的可靠性和准确性下降。为此, 研究人员已经致力于设计数据降噪过程, 在应用预测方法之前对数据进行处理, 以提高交通流预测的准确性和可靠性[9] [10] [11] [12] [13]。

其中, 小波(Wavelet, WL)方法是一种被广泛采用的降噪模型。Xie 等[9]采用小波卡尔曼滤波模型消除短时交通量预测中的噪声影响, 表现出比单独的小波模型和卡尔曼模型更好的性能。Jiang 等[10]提出了一种改进的离散小波包变换, 利用统计自相关函数来选择小波方法中的分解层数, 来处理原始数据源中隐含的噪声。此外, Lu 等[11]利用小波变换将交通流数据分解为多尺度分量。而在研究[12]和研究[13]的工作中应用小波降噪方法提高交通流预测精度, 并提出了基于自组织神经网络和神经小波的几种改进的预测模型。综上所述, 降噪过程的加入是提高交通流预测精度的有效手段。

但是上述的研究中, 小波变换的性能受到所选择的小波基的影响。不同的小波基具有不同的时频特性, 选择不合适的小波基可能导致去噪效果不佳。同时, 小波降噪中的阈值选择是一个重要的问题。合

适的阈值选择可以实现较好的去噪效果, 但如何选择适当的阈值并没有一个通用的准则, 需要根据具体问题和应用进行调整。所以在这项研究中, 本文侧重于提出一种自适应的降噪方法, 不需要去选择基函数和阈值函数, 该方法结合了经验模态分解(Empirical Mode Decomposition, EMD)和独立分量分析(Independent Component Analysis, ICA), 用于去除噪声干扰。并通过模拟仿真实验证明了该降噪方法的有效性, 在此基础上, 通过对真实交通流的预测进行评估, 结果表明该降噪方法与预测模型的组合可以有效地提高预测精度, 其中 EIMD 与 SVM 的组合预测模型有最高的预测精度。

2. 方法原理

2.1. 小波(WL)

对于消除交通流数据中的噪声干扰, 可以采用 WL 方法[14]。降噪模型的目标是减少噪声等干扰因素对模型预测精度的影响, 尤其是在复杂的非线性交通数据集中。一维噪声模型可以用以下形式表示:

$$R(t) = Q(t) + \xi(t), \quad t = 1, 2, \dots, n \quad (1)$$

其中, $R(t)$ 是含有噪声的原始信号; $Q(t)$ 是没有噪声的纯净信号; $\xi(t)$ 为噪声; i 为时刻; t 为时刻; n 是信号长度。小波降噪的原理是抑制 $R(t)$ 的 $Q(t)$ 部分, 恢复 $R(t)$ 的 $Q(t)$ 部分。小波降噪模型能够保留降噪后的原始信号特征, 小波降噪方法的流程通常分为三个步骤, 如下所示:

- Step 1. 选择一个小波, 确定小波分解的层数 N , 然后计算信号的 N 层分解。
- Step 2. 对于从层 1 到层 N 的每一层的高频因子, 选取阈值用于阈值化。
- Step 3. 根据近似系数和修正后的小波系数, 利用小波逆变换重构降噪后的交通数据。

2.2. 经验模态分解(EMD)

EMD 在图像分析、齿轮箱故障诊断等领域广泛应用, 并取得了显著的成功。EMD 的主要特点是数据自适应性, 这使得它能够适应不同类型的信号, 并获得良好的分解效果[15] [16]。EMD 通过采用三次样本插值方法, 在信号中自适应地分解出多个具有物理意义的本征模态函数(Intrinsic Mode Function, IMF)。IMF 具有以下两个特征条件: 1) 在整个时间序列中, 信号的极值点和过零点的数目相等或相差 1; 2) 由极大值点和极小值点确定的包络线均值为零。对 EMD 模型的更详细解释如下所述:

Step 1. 从数据序列 $R(t)$ 中找出潜在的极值点, 通过采用三次样条插值方法连接极值点和最大值点, 形成上包络 $U(t)$ 和下包络 $L(t)$ 。

Step 2. 计算上包络 $U(t)$ 和下包络 $L(t)$ 的平均值, 以获得平均包络 $M(t)$ (参见公式(2))。

Step 3. 计算 $R(t)$ 和 $M(t)$ 之间的差值以获得 $I(t)$ (参见公式(3))。如果 $I(t)$ 满足 IMF 的条件, 则 $I(t)$ 是 IMF 之一, 否则需要重复上述三个子步骤, 直到 $I(t)$ 满足 IMF 的条件。

Step 4. 通过计算 $R(t)$ 与 $I(t)$ 的差值得到残差部分 $r(t)$, 重复上述三个子步骤得到新的 IMF, 直到残差数据分布单调或有一个极值点为止。

最后, 故原始信号 $R(t)$ 可表示为被分解为一系列 IMF 的 $I_i(t)$ 和残差部分 $r(t)$, 如以下公式所示:

$$M(t) = \frac{U(t) + L(t)}{2} \quad (2)$$

$$I(t) = R(t) - M(t) \quad (3)$$

$$R(t) = \sum_{i=1}^m I_i(t) + r(t) \quad (4)$$

2.3. 独立分量分析(ICA)

ICA 是一种统计方法, 主要用于信号处理和数据分析。其目标是将观测信号分解为相互独立的子信号, 这些子信号被称为独立分量。ICA 假设观测信号是由一组独立分量线性混合而成, 通过寻找最大可能独立的分离矩阵来实现分解[17]。

$$X(t) = AS(t) \quad (5)$$

其中, t 为离散时间变量; $X(t)$ 为 M 维观测信号集合; A 为 $M \times N$ 阶未知混合矩阵; $S(t)$ 为互相独立的子信号。为了在 A 和 $S(t)$ 均未知的情况下, 仅利用观测到的信号 $X(t)$, 尽可能真实地分离出源信号 $S(t)$, 可构建一个分离矩阵(解混矩阵) W , 使 $X(t)$ 经过分离矩阵 W 变换后, 得到 n 维输出列向量。这样, ICA 问题的求解可表示为:

$$Y(t) = WX(t) = WAS(t) = S(t) \quad (6)$$

其中, $Y(t)$ 为 $S(t)$ 的估计信号。由于解混时除了各 $S_i(t)$ 互相独立外, 对 $S(t)$ 与 A 无其他先验知识, 故称为“盲信源分解”。又由于分解结果要求 $Y(t)$ 中各分量互相独立, 故称“独立分量分析”。

通常把解混过程 W 分解成两步, 第一步先把观察矢量 $X(t)$ 经过线性系统 B 变换成中间输出 $Z(t)$, 称为球化; 第二步再把 $Z(t)$ 变换成 $Y(t)$ 。要求 $Z(t)$ 中各分量正交归一(即互相正交而且方差都等于 1)。

Fast ICA (Fast Independent Component Analysis) 是一种常用的 ICA 方法[18], 该方法基于非高斯性最大化原理, 采用牛顿迭代方法寻找非高斯性最大值, 对观测变量的大量采样点进行批处理, 每次从观测信号中分离出一个独立分量。对 FastICA 模型的更详细解释如下所述:

Step 1. 将 $X(t)$ 去均值, 然后加以球化得 $Z(t)$ 。

Step 2. 取任意初始矢量 $u_i(0)$, 要求它满足模长为 1。

Step 3. 求 $u_i(k+1) = E\left(Zf\left(u_i^T(k)Z\right)\right) - E\left(f'\left(u_i^T(k)Z\right)\right)u_i(k)$, 其中 k 为迭代序号, 不是时间序号, $E(\cdot)$ 可以通过对 Z 的各采样时刻求均值来估计, $f(\cdot)$ 是函数 $f(y) = y \cdot \exp(-y^2/2)$ 。

Step 4. 将 $u_i(k+1)$ 归一化。

Step 5. 如果 $|u_i(k+1)u_i^T(k)|$ 不接近 1, 则令 k 加 1, 回到 Step 2。否则迭代结束。输出最终的 $u_i(k+1)$ 作为 u_i 。

Step 6. 提取独立分量: $g_i(k) = u_i^T z(k)$ 。

2.4. 经验独立模态分解(EIMD)

EMD 方法分解出的各个 IMF 包含了原始信号在不同时间尺度上的局部特征, 该方法可用于提取交通流种高频信息。相比之下, ICA 方法通过提取独立源并利用随机噪声的弱非高斯性质来识别随机噪声。因此, 本研究提出了经验独立模态分解(Empirical Independent Mode Decomposition, EIMD)降噪方法, 结合了 EMD 和 ICA 的优势, 主要用于压制高频随机噪声。以下是具体方法的描述:

Step 1. 使用 EMD 方法将原始信号 $R(t)$ 分解为多个 IMFs, 记为 $D = \{I_1, I_2, \dots, I_k, \dots, I_m\}$ 。

Step 2. 计算每个 IMF 分量与原始信号 $R(t)$ 之间的相关系数。记相关系数中具有第一个局部最小值的下一个 IMF 为 I_k 。可以观察到, 从 I_k 开始, 后续 IMFs 与原始信号 $R(t)$ 的相关性开始增加, 可认为 I_k 为含噪 IMFs 与纯净 IMFs 的分离边界。计算相关系数的公式如为公式(7)所示。

$$r(R(t), I_i(t)) = \frac{\text{Cov}(R(t), I_i(t))}{\sqrt{\text{Var}[R(t)]\text{Var}[I_i(t)]}} \quad (7)$$

Step 3. 对包括 I_k 和 I_k 之前的 IMFs, 进行 ICA 方法的盲源分离, 用于进一步分离出高频的随机噪声。

Step 4. 选择频率最高的独立分量作为高频随机噪声, 并将其去除。之后, 将剩余的独立分量重构回IMFs。

Step 5. 将重构后的IMFs, 记为 $I'_i(t)$ 和没参与重构的剩余IMFs累加, 得到降噪后的信号 $Q(t)$ 。

$$Q(t) = \sum_{i=1}^k I'_i(t) + \sum_{i=m}^m I_i(t) \quad (8)$$

2.5. 小波神经网络(WNN)

WNN是在传统反向传播神经网络(Back Propagation Neural Network, BP) [19] [20]结构的基础上, 用母小波函数的伸缩平移代替BP隐节点的Sigmoid函数。小波函数具有多样性, 鉴于Morlet小波拥有较好的光滑性和时域局部性, 本文采用Morlet小波基函数替代隐含层的Sigmoid函数, 表达式为:

$$\psi(x) = \cos(1.75x) \exp(-x^2/2) \quad (9)$$

在小波神经网络训练过程中, 采用适应时刻估计方法(Adaptive Moment Estimation, Adam)梯度下降法对WNN的网络权值和小波基函数伸缩平移因子进行修正, 通过对参数的调整, 实现输出值逐步逼近期望值。

2.6. 卷积神经网络(CNN)

卷积神经网络CNN已经在图像分类, 语音识别, 自然语音处理等几大领域获得显著成功[21]。在CNN结构中, 包含两个特殊的神经网络层: 一是卷积层, 每个卷积层都包含多个特征映射, 每个特征映射是一个由多个神经元构成的“平面”, 通过一种卷积滤波器提取输入的一种特征。二是池化层, 其作用是基于局部相关性原理进行亚采样, 从而在减少数据量的同时保留有用信息。特别地, 在一维卷积中, 滑动窗口是在时间序列数据的宽度方向上进行滑动, 并通过相乘求和的方式进行卷积计算。通过一维卷积操作, 可以捕捉到时间序列数据中的局部模式和趋势特征, 从而进行有效的预测和分析。

2.7. 支持向量机(SVM)

SVM回归的目标是找到一个最优的回归函数(或称为超平面), 以尽可能地拟合训练数据, 并在保持较小的预测误差的同时, 最大化边界(即支持向量)与训练样本之间的间隔。SVM回归与传统的线性回归不同, 它通过使用核函数将数据映射到高维特征空间, 从而能够处理非线性关系。在高维特征空间中, SVM回归通过寻找一个最优的超平面来实现回归任务。该方法已被广泛应用于数据分析, 模式识别, 分类, 回归等领域[22] [23]。

3. 模拟实验与降噪结果讨论

3.1. 降噪性能评价

在仿真信号实验中, 若真实信号已知, 可采用均方根差(RMSE)和信噪比(SNR)作为降噪评价指标, 计算公式如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Q(t) - R(t))^2} \quad (10)$$

$$SNR = 10 * \lg \left(\frac{\sum_{t=1}^n Q(t)^2}{\sum_{t=1}^n (Q(t) - R(t))^2} \right) \quad (11)$$

其中, t 为时刻; n 是信号长度; $Q(t)$ 为降噪后信号; $R(t)$ 为真实信号。SNR 越大、RMSE 越小说明降噪效果越好, 反之, 说明降噪效果越差[24]。

3.2. 仿真交通流 I 的降噪分析

为了评估所提出的 EIMD 方法的有效性, 有必要模拟类似于现实场景的交通流信号, 因为没有实际的方法可以获得无噪声的交通流数据来定量评估降噪方法, 所以通过对交通流数据特征的分析, 可以看出, 交通流数据在高峰时段具有很强的周期性和波动性。因此, 模拟交通流是由公式(12)中不同频率的子信号组合而成, 模拟时间为 5 天, 每 5 分钟记录一次数据, 共 1440 个数据点。在模拟数据中加入不同信噪比的高斯白噪声, 并在高峰时段加入实际交通流波动。具体而言, 使用 EMD 对真实交通流数据进行分解, 得到 IMF1 和 IMF2, 然后将其添加到模拟数据中, 以用于模拟高峰时段波动。本研究使用的实际交通流数据来自贵阳市, 详情见 4.1 节。

$$\begin{cases} y_1 = a_1 * \sin(b_1 * x + c_1) \\ y_2 = a_2 * \sin(b_2 * x + c_2) \\ y_3 = a_3 * \sin(b_3 * x + c_3) \\ y_4 = a_4 * \sin(b_4 * x + c_4) \\ y_5 = a_5 * \sin(b_5 * x + c_5) \\ y_6 = a_6 * \sin(b_6 * x + c_6) \\ y_7 = a_7 * \sin(b_7 * x + c_7) \\ \xi_1 = \text{noise} \\ \xi_2 = \text{real noise} \\ y' = y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 \\ y = y' + \xi_1 + \xi_2 \end{cases} \quad (12)$$

其中, $a_1 = 378.5$, $b_1 = 0.002183$, $c_1 = -0.008533$, $a_2 = 140.3$, $b_2 = 0.02183$, $c_2 = -2.222$, $a_3 = 384.1$, $b_3 = 0.003959$, $c_3 = 1.87$, $a_4 = 49.37$, $b_4 = 0.04362$, $c_4 = 3.166$, $a_5 = 47.8$, $b_5 = 0.06548$, $c_5 = 0.6803$, $a_6 = 30.07$, $b_6 = 0.1093$, $c_6 = 2.926$, $a_7 = 191.7$, $b_7 = 0.004817$, $c_7 = 4.413$ 以及 ξ_1 表示添加的白噪声, ξ_2 表示添加的真实交通流高峰期噪声, y' 表示不含噪声的仿真交通流信号, y 表示含有噪声的仿真交通流信号。首先, 先加入 30 dB 白噪声, 具体仿真交通流信号如图 1 所示。

Table 1. Calculation results of correlation coefficients between each IMF and $R(t)$ (30 db white noise added)

表 1. 各 IMF 与 $R(t)$ (加 30 db 白噪声)相关系数计算结果

IMFs	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	Residual
相关系数	0.1294	0.1055	0.1175	0.3454	0.8339	0.8266	0.3835

接下来, 为验证 EIMD 降噪方法的有效性。在对比其他降噪方法时, WL 方法中基函数选取了不同小波类型, 包括 db4, coif2, haar, sym2, 其中分解层数设计为 3。在阈值处理方式上, 采用软阈值, 这可以使得重建信号比较光滑。在计算阈值的方法上, 则采用无偏似然估计原则(Rigrsure)。EIMD 中 IMF 分解次数设置为 6, 各 IMF 和原始序列 $R(t)$ 相关系数计算结果如表 1 所示, 由 2.4 内容可知, 应选取前 3 个 IMFs 作为 2.4 中 Step 3 的输入。

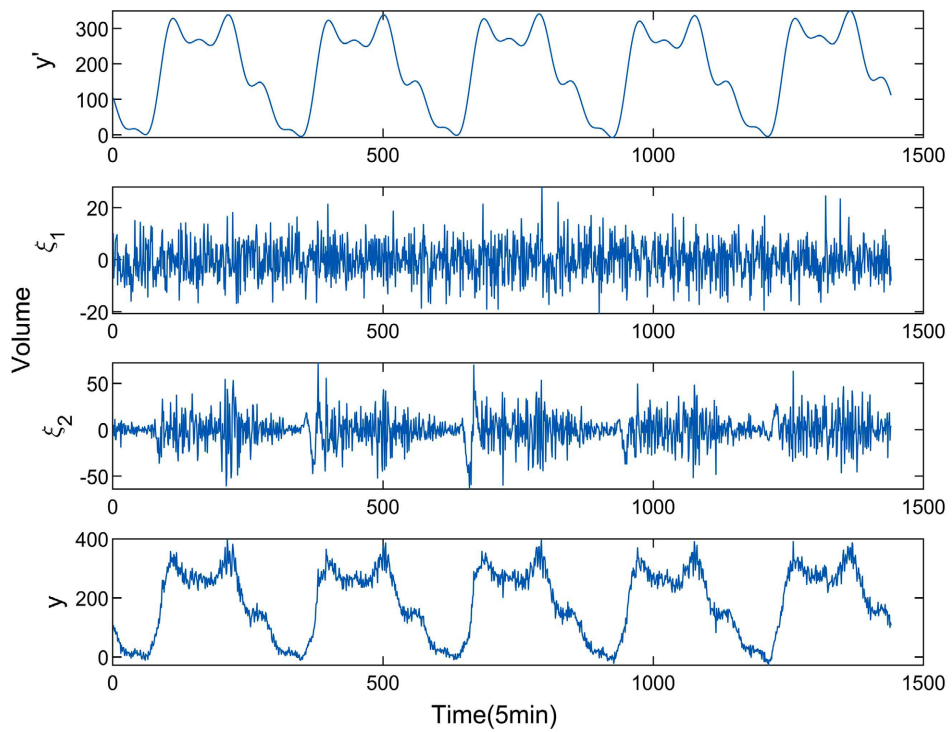


Figure 1. Simulation signal I and its components

图 1. 仿真信号 I 及其各组成部分

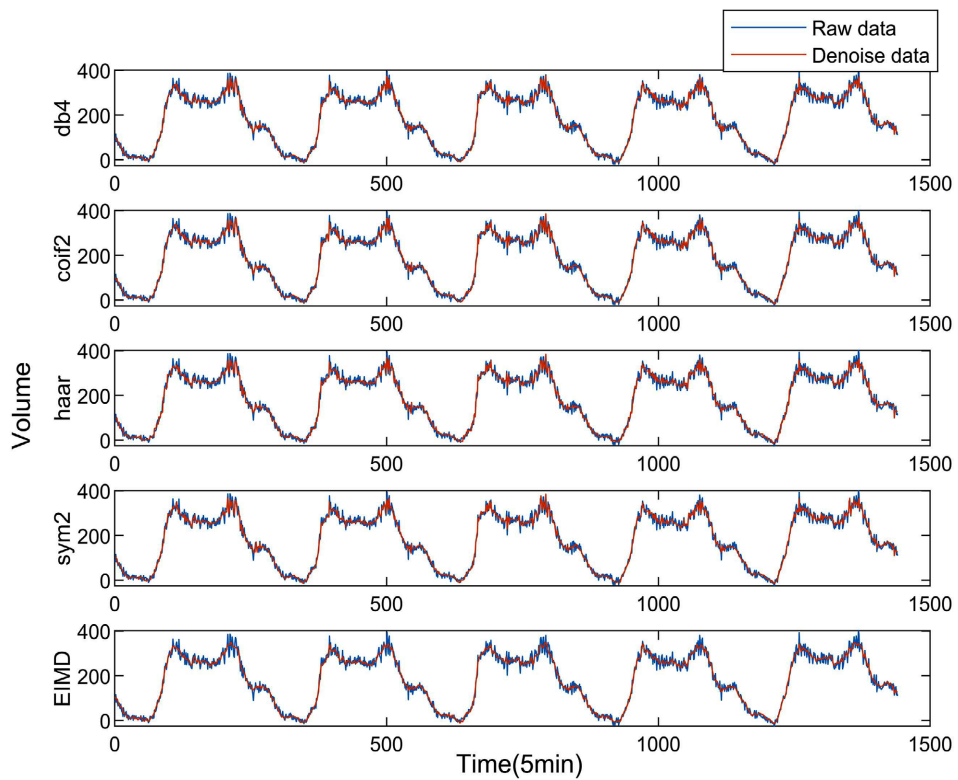


Figure 2. The denoising results of simulation signal I with noise by different denoising methods

图 2. 不同降噪方法对含噪的仿真信号 I 的降噪结果

Table 2. Denoising results (30 db white noise added)
表 2. 降噪结果(加入 30 db 白噪声)

	RMSE	SNR
原始信号	17.47	50.35
WL (db4)	10.52	60.46
WL (coif2)	10.75	60.03
WL (haar)	11.99	57.83
WL (sym2)	10.85	59.83
EIMD	8.97	63.62

Table 3. Denoising results (35 db white noise added)
表 3. 降噪结果(加入 35 db 白噪声)

	RMSE	SNR
原始信号	17.07	50.81
WL (db4)	11.84	58.08
WL (coif2)	11.00	59.55
WL (haar)	12.40	57.16
WL (sym2)	11.16	59.29
EIMD	7.89	66.14

Table 4. Denoising results (40 db white noise added)
表 4. 降噪结果(加入 40 db 白噪声)

	RMSE	SNR
原始信号	16.49	51.49
WL (db4)	11.54	58.60
WL (coif2)	11.50	58.66
WL (haar)	12.62	56.80
WL (sym2)	12.00	57.82
EIMD	8.35	65.08

Table 5. Denoising results (45 db white noise added)
表 5. 降噪结果(加入 45 db 白噪声)

	RMSE	SNR
原始信号	16.53	51.44
WL (db4)	11.52	58.63
WL (coif2)	11.47	58.73
WL (haar)	12.00	57.81
WL (sym2)	12.16	57.55
EIMD	8.21	65.41

利用 WL 和 EIMD 对仿真信号 I 进行降噪, 降噪结果如图 2 所示。图 2 清楚地表明, 本文提出的 EIMD 降噪方法与其他降噪方法相比, 产生的波形更清晰平滑, 有效地消除了高频部分的噪声。均方根误差和信噪比分析如表 2 所示。EIMD 方法的均方根误差较小, 信噪比较大, 降噪性能优于其他方法。为了进一步验证结果, 在原交通流信号中加入信噪比分别为 35 dB、40 dB 和 45 dB 的高斯白噪声, 计算其均方根误差和信噪比, 如表 3~5 所示。这些表中的数据一致地证明了本文提出的方法与其他降噪方法相比具有更强的降噪能力。

4. 预测结果比较与讨论

4.1. 数据来源

为了评估基于降噪方法与 SVM 的组合预测模型的有效性, 对贵州省贵阳市长岭南路与阳关大道交叉口环路检测器采集的真实交通流数据进行了分析。数据收集时间为 2021 年 3 月 1 日至 2021 年 3 月 5 日, 每 5 分钟测量一次, 每天总计 288 个数据点。前 4 天的数据(占总数据的 80%)作为训练集, 用于模型的训练, 第五天的数据(占总数据的 20%)作为测试集, 用于评估最终模型的预测性能。

4.2. 模型预测性能评价

实验结果评价用到 3 个误差指标, 分别为均方根误差 RMSE、平均绝对误差 MAE 和绝对百分比误差 MAPE, 计算公式为:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_{real}(t) - Y_{pred}(t))^2} \quad (13)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_{real}(t) - Y_{pred}(t)| \quad (14)$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{Y_{pred}(t) - Y_{real}(t)}{Y_{real}(t)} \right| \quad (15)$$

其中, $Y_{real}(t)$ 表示在 t 时刻的交通流的实际值; $Y_{pred}(t)$ 表示预测值, n 表示样本个数。

4.3. 预测结果

本研究采用了 WL 和 EIMD 方法对原始数据进行降噪处理。这些方法的参数在 3.2 节中有详细的相关设置, 特别地, EIMD 中各 IMF 和 $R(t)$ 相关系数计算结果如表 6 所示。

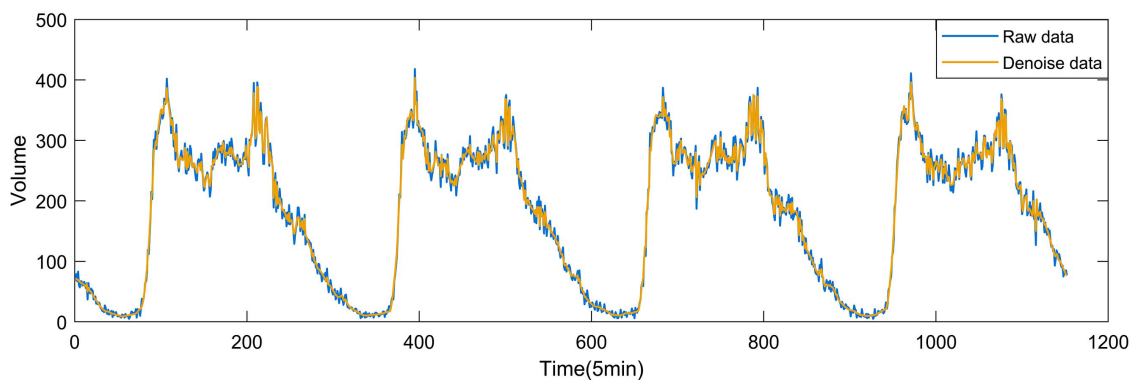
Table 6. Calculation results of correlation coefficients between each IMF and raw traffic flow

表 6. 各 IMF 与原始交通流的相关系数计算结果

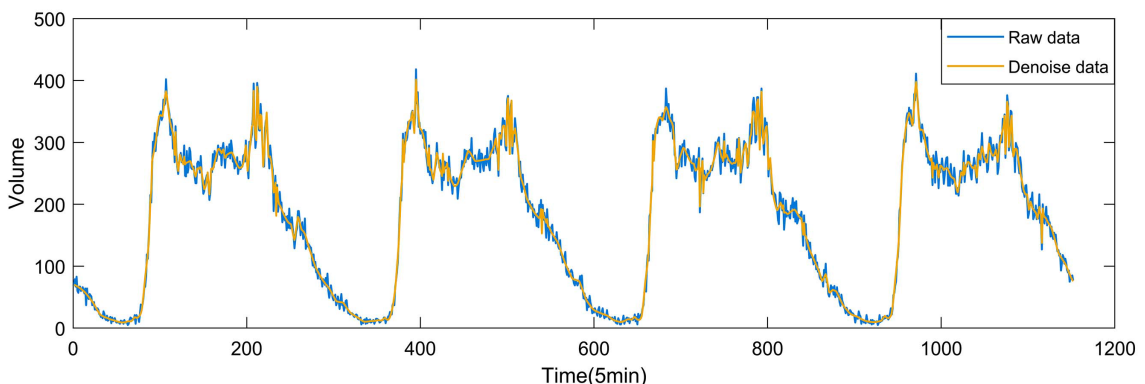
IMFs	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	Residual
相关系数	0.1182	0.1101	0.1614	0.5509	0.6900	0.3390	-0.0074

当使用 WL 和 EIMD 方法对原始数据进行降噪处理后, 图 3 显示了前 1152 个数据的降噪结果。黄线表示降噪后的数据, 蓝线表示原始数据。数据按周期划分, 每天包含 288 个数据点, 呈现出明显的高峰和低谷。在高峰时段, 数据的波动性明显高于非高峰时段。这是因为道路占用率增加以及高峰时段司机行为的影响所导致的。因此, 在高峰时段交通噪声比非高峰时段更为明显。值得注意的是, 在非高峰时段, 噪声干扰较小, 降噪结果与原始数据几乎一致。表 7 提供了采用不同降噪方法后的交通流与原始交通流的详细对比。根据表中的结果, 可以观察到 EIMD 方法显示出与原始数据最显著的偏差, 表明它有

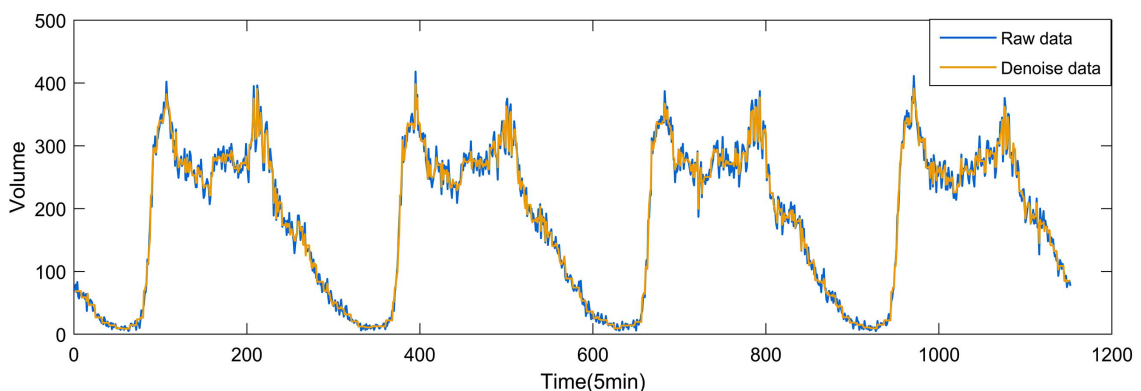
效地滤除了大量的噪声。将降噪后得到的数据, 使用 SVM 模型进行交通流预测。交通流预测方案包括 SVM、SVM + WL(db4)、SVM + WL(coif2)、SVM + WL(haar)、SVM + WL(sym2)和 SVM + EIMD。SVM 方法中, 采用步长为 6, 惩罚因子为 4.0, 径向基函数参数为 0.8。预测结果如图 4 所示。黑色圆圈表示真实数据, 蓝色符号表示预测数据。表 8 中列出了预测结果的 RMSE、MAE、MAPE。直接使用 SVM 模型进行预测时, RMSE 为 20.43, MAE 为 14.44, MAPE 为 11.99%。在进行不同降噪处理后, 预测的 RMSE、MAE、MAPE 都得到了不同程度上的减少。此外, 与 EIMD 的组合预测模型有着最好的预测精度, 预测精度, 预测的 RMSE、MAE、MAPE 分别减少了 33%、30%和 31.5%。综上所述, 降噪后的数据预测能够更好地捕捉交通流的变化趋势, 从而提高预测精度。与其他降噪方法与 SVM 的组合模型相比, EIMD 和 SVM 组合预测模型表现出更优越的性能和优势。



(a) WL(db4)



(b) WL(coif2)



(c) WL(haar)

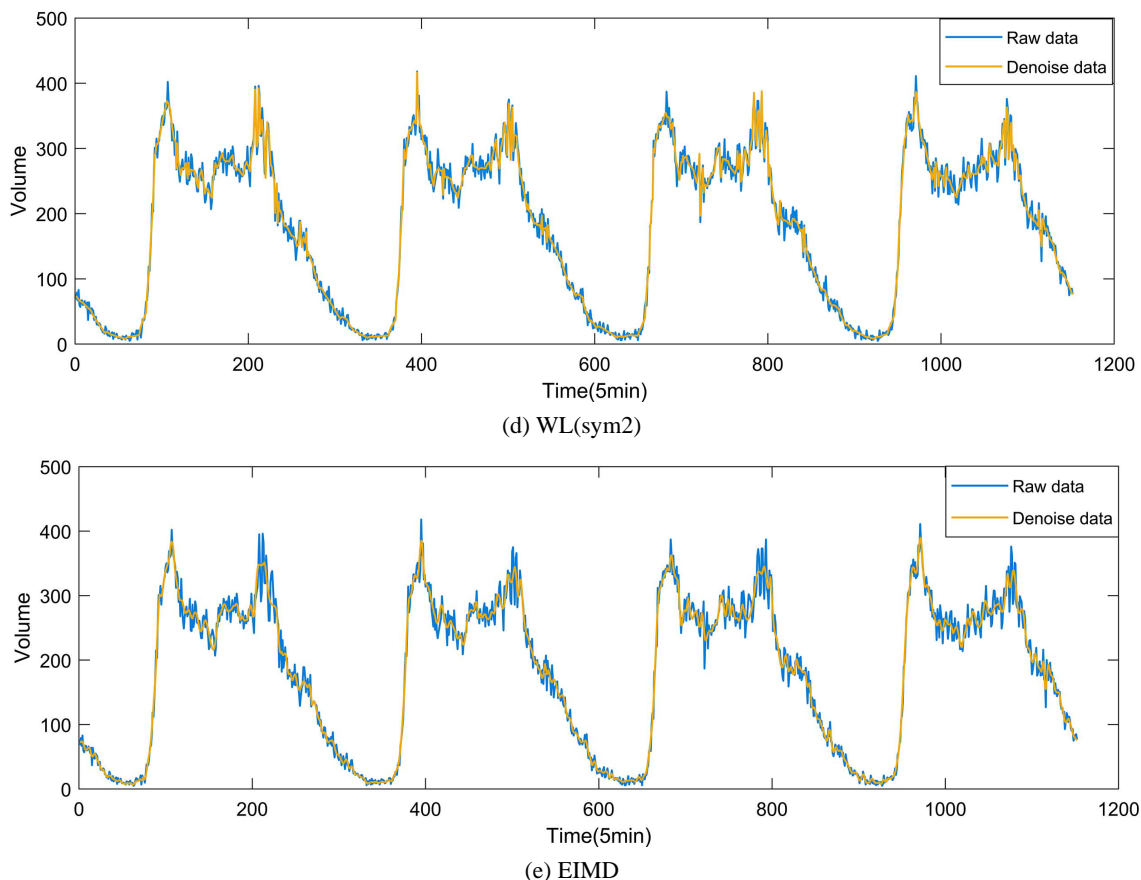


Figure 3. Denoising effects of different methods on traffic flow
图 3. 不同方法降噪方法对交通流的降噪效果

Table 7. Comparison of traffic flow and original traffic flow with different denoising algorithms

表 7. 采用不同去噪算法的交通流与原始交通流的比较

	RMSE	MAE	MAPE
WL (db4)	7.02	5.92	7.48
WL (coif2)	8.00	7.88	10.24
WL (haar)	7.85	6.81	8.62
WL (sym2)	8.03	7.78	10.04
EIMD	8.87	10.25	13.82

为了进一步验证 EIMD 方法的有效性, 我们引入了两种额外的预测方法, 并比较了不同的降噪方法与这些预测方法的组合预测模型。WNN 方法中, 输入层由 4 个节点组成, 采用 1 个输入层、1 个隐藏层、1 个输出层的标准结构。CNN 方法中, 采用步长为 8, 三个隐藏层, 每层使用 31 个卷积核, 它分别生成 16、32 和 32 个特征映射, 选取激活层为 RELU, 采用 Adam 梯度下降法。设置最大训练次数为 300 次, 初始学习率为 0.01, 学习率下降因子为 0.1。学习率降低 0.010×1 , 采用最大池化方法。表 9 和表 10 显示了不同降噪方法和预测模型的组合模型的预测性能比较结果。从表中结果可知, 在进行不同降噪处理后, 预测的 RMSE、MAE、MAPE 都得到了不同程度上的减少。此外, 与 EIMD 的组合预测模型有着最好的预测精度, 这与 SVM 的情况保持一致。

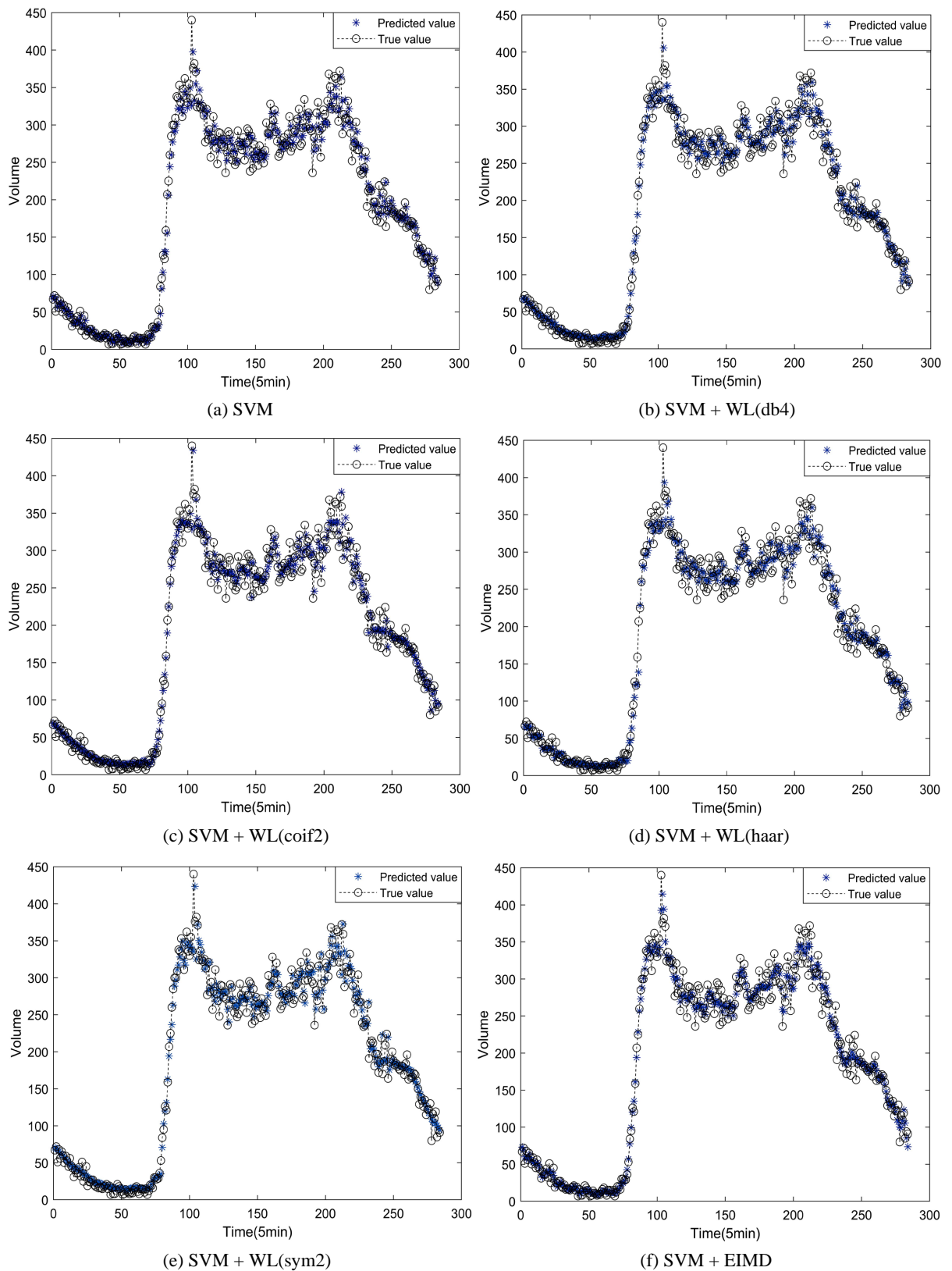


Figure 4. Prediction results of different denoising methods combined with SVM

图 4. 不同降噪方法结合 SVM 的预测结果

Table 8. Prediction results of different denoising algorithms combined with SVM
表 8. 不同降噪算法结合 SVM 的预测结果

	RMSE	MAE	MAPE
SVM	20.43	14.44	11.99
SVM + WL (db4)	18.61	12.99	10.20
SVM + WL (coif2)	17.85	12.01	9.86
SVM + WL (haar)	19.28	13.23	10.86
SVM + WL (sym2)	18.21	12.28	9.87
SVM + EIMD	13.61	10.05	8.21

Table 9. Prediction results of different denoising algorithms combined with WNN
表 9. 不同降噪算法结合 WNN 的预测结果

	RMSE	MAE	MAPE
WNN	22.66	17.15	16.22
WNN + WL (db4)	19.36	14.27	13.59
WNN + WL (coif2)	21.10	15.70	14.10
WNN + WL (haar)	20.41	14.83	13.06
WNN + WL (sym2)	21.35	14.53	13.89
WNN + EIMD	14.74	11.69	12.25

Table 10. Prediction results of different denoising algorithms combined with CNN
表 10. 不同降噪算法结合 CNN 的预测结果

	RMSE	MAE	MAPE
CNN	20.76	14.73	12.91
CNN + WL (db4)	18.22	12.73	10.42
CNN + WL (coif2)	17.94	12.21	10.63
CNN + WL (haar)	18.68	12.94	10.48
CNN + WL (sym2)	18.27	12.34	9.64
CNN + EIMD	13.66	10.01	8.34

从表 8, 表 9 和表 10 的预测结果中, 我们可以总结出以下几个有趣的发现:

1) 从表和图中的结果观察可以看出, 结合降噪方法的模型的预测结果都更好于不采用降噪方法的模型。这主要是因为降噪过程能够去除原始数据中的急剧变化部分, 使数据的变化趋势更加明显。通过对数据进行降噪处理, 模型在训练过程中变得更加稳定, 并能够获得更高的预测精度。

2) 在所有降噪方法中, EIMD 对预测结果的改善效果最好, 而各种 WL 的预测精度相似, 但都比 EIMD 的预测精度稍差。EIMD 预测精度高的原因在于它将原始数据分解为多个 IMFs 后, 通过识别到的多个含噪 IMFs 进行盲源分离, 压制了高频随机噪声。在 EIMD 与不同的预测方法(SVM, WNN, CNN)组合时, 在 RMSE、MAE 和 MAPE 上均表现出较好的结果。这说明 EIMD 方法对不同的预测模型都具有一定的通用性和有效性。此外, EIMD 和 SVM 的组合在预测精度方面表现最好, 优于其他组合模型。

3) WL 降噪模型的性能。在四种小波类型中, sym 相对产生最高的预测性能, coif 和 db 类型的预测精度相近, 低于 sym。haar 小波的预测性能不如其他三种小波。其原因可以归结为 sym 波形对交通流数据的变化模式表现出更好的拟合性能。coif 和 db 型小波适合处理具有对称特征的数据。由于交通量隐含的周期性, 它表示近似对称的模式。因此, coif 和 db 也可以产生相对较高的预测精度。

综上, 这些研究结果强调了 EIMD 方法在降低噪声和提高交通流预测精度方面的有效性和优越性。将 EIMD 与不同的预测方法相结合, 提高了预测性能, 为交通流路段提供了可靠的预测。这证明了使用 EIMD 作为交通流量预测任务中有价值的工具的潜力。

5. 结论

本文提出了一种结合 EMD 和 ICA 的 EIMD 方法, 该方法可以有效地去除交通流数据中的噪声。为了验证 EIMD 方法的有效性, 进行了仿真实验。在此基础上, 将数据降噪方法与机器学习模型相结合, 提出了一种新的交通流预测框架。该框架采用了 WL (haar, db4, sym2, coif2) 和 EIMD 降噪方法, 分别结合 SVM、WNN、CNN 三种预测模型对交通流进行预测, 通过分析 RMSE、MAE 和 MAPE 性能指标, 确定出准确、可靠的交通流预测模型。通过观察这项研究的结果, 我们可以得出几个有趣的结论。首先, 仿真结果表明, 在各种噪声环境下, EIMD 方法在降噪方面优于不同小波基的 WL 方法。其次, 结合降噪方法的模型预测性能均优于未采用降噪方法的模型。第三, 在考虑降噪算法与不同预测模型(SVM, WNN, CNN)的组合时, EIMD 与 SVM 的组合模型预测性能高于其他组合模型。第四, 在 WL 方法中, sym 小波类型的预测精度高于其他三种类型。

此外, 在未来的工作中我们考虑以下两个方面来进一步提高预测的准确性: 1) 引入更多特征, 除了基本的交通流数据, 可以考虑引入如天气数据、道路等级、交通事件等可能影响预测精度的因素。2) 探索先进的机器学习技术, 除了传统的机器学习方法, 可以尝试引入更先进的技术, 如深度学习模型、集成方法或混合模型。

参考文献

- [1] Habtemichael, F.G. and Cetin, M. (2016) Short-Term Traffic Flow Rate Forecasting Based on Identifying Similar Traffic Patterns. *Transportation Research Part C: Emerging Technologies*, **66**, 61-78. <https://doi.org/10.1016/j.trc.2015.08.017>
- [2] Dong, H., Jia, L., Sun, X., Li, C. and Qin, Y. (2009) Road Traffic Flow Prediction with a Time-Oriented ARIMA Model. 2009 5th International Joint Conference on INC, IMS and IDC, Seoul, 25-27 August 2009, 1649-1652. <https://doi.org/10.1109/NCM.2009.224>
- [3] Kumar, S.V. and Vanajakshi, L. (2015) Short-Term Traffic Flow Prediction Using Seasonal ARIMA Model with Limited Input Data. *European Transport Research Review*, **7**, Article No. 21. <https://doi.org/10.1007/s12544-015-0170-8>
- [4] Duan, P., Mao, G., Liang, W. and Zhang, D. (2018) A Unified Spatio-Temporal Model for Short-Term Traffic Flow Prediction. *IEEE Transactions on Intelligent Transportation Systems*, **20**, 3212-3223. <https://doi.org/10.1109/TITS.2018.2873137>
- [5] Wang, H., Zhang, R., Cheng, X. and Yang, L. (2022) Hierarchical Traffic Flow Prediction Based on Spatial-Temporal Graph Convolutional Network. *IEEE Transactions on Intelligent Transportation Systems*, **23**, 16137-16147. <https://doi.org/10.1109/TITS.2022.3148105>
- [6] Guo, X. and Zhu, Q. (2009) A Traffic Flow Forecasting Model Based on BP Neural Network. 2009 2nd International Conference on Power Electronics and Intelligent Transportation System (PEITS), Shenzhen, 19-20 December 2009, 311-314.
- [7] Yan, H., Qi, Y. and Yu, D.J. (2022) Short-Term Traffic Flow Prediction Based on a Hybrid Optimization Algorithm. *Applied Mathematical Modelling*, **102**, 385-404. <https://doi.org/10.1016/j.apm.2021.09.040>
- [8] Kumar, K., Parida, M. and Katiyar, V.K. (2013) Short Term Traffic Flow Prediction for a Non Urban Highway Using Artificial Neural Network. *Procedia-Social and Behavioral Sciences*, **104**, 755-764. <https://doi.org/10.1016/j.sbspro.2013.11.170>

- [9] Xie, Y., Zhang, Y. and Ye, Z. (2007) Short-Term Traffic Volume Forecasting Using Kalman Filter with Discrete Wavelet Decomposition. *Computer-Aided Civil and Infrastructure Engineering*, **22**, 326-334. <https://doi.org/10.1111/j.1467-8667.2007.00489.x>
- [10] Jiang, X. and Adeli, H. (2004) Wavelet Packet-Autocorrelation Function Method for Traffic Flow Pattern Analysis. *Computer-Aided Civil and Infrastructure Engineering*, **19**, 324-337. <https://doi.org/10.1111/j.1467-8667.2004.00360.x>
- [11] Lu, B. and Huang, M. (2009) Traffic Flow Prediction Based on Wavelet Analysis, Genetic Algorithm and Artificial Neural Network. 2009 *International Conference on Information Engineering and Computer Science*, Wuhan, 19-20 December 2009, 1-4. <https://doi.org/10.1109/ICIECS.2009.5362817>
- [12] Boto-Giralda, D., Díaz-Pernas, F.J., González-Ortega, D., Díez-Higuera, J.F., Antón-Rodríguez, M., Martínez-Zarzuela, M. and Torre-Díez, I. (2010) Wavelet-Based Denoising for Traffic Volume Time Series Forecasting with Self-Organizing Neural Networks. *Computer-Aided Civil and Infrastructure Engineering*, **25**, 530-545. <https://doi.org/10.1111/j.1467-8667.2010.00668.x>
- [13] Dunne, S. and Ghosh, B. (2013) Weather Adaptive Traffic Prediction Using Neurowavelet Models. *IEEE Transactions on Intelligent Transportation Systems*, **14**, 370-379. <https://doi.org/10.1109/TITS.2012.2225049>
- [14] Huang, D.W. (2003) Wavelet Analysis in a Traffic Model. *Physica A: Statistical Mechanics and Its Applications*, **329**, 298-308. [https://doi.org/10.1016/S0378-4371\(03\)00623-X](https://doi.org/10.1016/S0378-4371(03)00623-X)
- [15] Nunes, J.C., Bouaoune, Y., Delechelle, E., Niang, O. and Bunel, P. (2003) Image Analysis by Bidimensional Empirical Mode Decomposition. *Image and Vision Computing*, **21**, 1019-1026. [https://doi.org/10.1016/S0262-8856\(03\)00094-5](https://doi.org/10.1016/S0262-8856(03)00094-5)
- [16] Liu, B., Riemenschneider, S. and Xu, Y. (2006) Gearbox Fault Diagnosis Using Empirical Mode Decomposition and Hilbert Spectrum. *Mechanical Systems and Signal Processing*, **20**, 718-734. <https://doi.org/10.1016/j.ymsp.2005.02.003>
- [17] Draper, B.A., Baek, K., Bartlett, M.S. and Beveridge, J.R. (2003) Recognizing Faces with PCA and ICA. *Computer Vision and Image Understanding*, **91**, 115-137. [https://doi.org/10.1016/S1077-3142\(03\)00077-8](https://doi.org/10.1016/S1077-3142(03)00077-8)
- [18] Oja, E. and Yuan, Z. (2006) The FastICA Algorithm Revisited: Convergence Analysis. *IEEE Transactions on Neural Networks*, **17**, 1370-1381. <https://doi.org/10.1109/TNN.2006.880980>
- [19] 翟久刚, 田延飞, 严新平. 基于 BP 神经网络与残差分析的船舶交通流量预测[J]. 上海海事大学学报, 2013, 34(1): 19-22.
- [20] 马秋芳. 改进 PSO 优化的 BP 神经网络短时交通流预测[J]. 计算机仿真, 2019, 36(4): 94-98+323.
- [21] Narmadha, S. and Vijayakumar, V. (2023) Spatio-Temporal Vehicle Traffic Flow Prediction Using Multivariate CNN and LSTM Model. *Materials Today: Proceedings*, **81**, 826-833. <https://doi.org/10.1016/j.matpr.2021.04.249>
- [22] Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. *Machine Learning*, **20**, 273-297. <https://doi.org/10.1007/BF00994018>
- [23] 罗文慧, 董宝田, 王泽胜. 基于 CNN-SVR 混合深度学习模型的短时交通流预测[J]. 交通运输系统工程与信息, 2017, 17(5): 68-74.
- [24] 甄龙信, 王云龙, 邓小艳, 张伟锟. 基于集合经验模态的小波半软阈值降噪[J]. 探测与控制学报, 2018, 40(5): 53-57.