

基于多维时空的NPCA-PSR-IGM(1,1)组合模型的短时交通流预测

殷礼胜 高贺* 魏帅康 孙双晨 何怡刚

(合肥工业大学电气与自动化工程学院 合肥 230009)

摘要: 针对城市短时交通流序列非线性和混沌性的特点, 为提高短时交通流的预测精度, 该文提出一种基于多维时空的非线性主成分分析(NPCA)和相空间重构(PSR)的改进灰色(IGM(1,1))组合预测模型。首先, 使用数据相关性的非线性主成分分析算法对多维交通流量序列进行时空降维, 同时保留影响预测点的主要交通流量数据, 从而提高建模的精确度; 其次, 利用多维时空交通流量序列相空间重构放大交通流量内部的细微特征, 以使其内在规律得以充分展现, 进一步提升预测精度; 最后, 结合背景值改进的灰色模型适应于线性、非线性以及所需数据少的特点, 进行短时交通流预测。实验结果表明, NPCA-PSR-IGM(1,1)组合预测模型的平均相对误差相比NPCA-PSR-GM(1,1)组合预测模型减小3.12%, 其标准偏差相对PCA-PSR-IGM(1,1)组合预测模型从15.7091下降到2.0589。同时与最新的预测模型相比, 该组合预测模型也提高了预测精度, 达到了较好的预测效果。

关键词: 短时交通流预测; 多维时空; 非线性主成分分析; 相空间重构; 改进灰色模型

中图分类号: TN911.7; U491.1

文献标识码: A

文章编号: 1009-5896(2021)04-1035-07

DOI: [10.11999/JEIT200026](https://doi.org/10.11999/JEIT200026)

Short-term Traffic Flow Prediction Based on NPCA-PSR-IGM (1,1) Combined Model of Multi-dimensional Space-time

YIN Lisheng GAO He WEI Shuaikang SUN Shuangchen HE Yigang

(School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

Abstract: In view of the nonlinear and chaos of urban short-term traffic flow sequence, this article proposes a combined prediction model based on multi-dimensional spatio-temporal Nonlinear Principal Component Analysis (NPCA) and Phase Space Reconstructed (PSR) Improved Gray Model (IGM(1,1)) in order to improve its forecast accuracy. First, the data correlation NPCA algorithm is used to reduce the spatial and temporal dimensions of multi-dimensional traffic flow sequences, while preserving the main traffic flow data that affects the predicted points, so as to improve the accuracy of the modeling. Phase space reconstruction amplifies the subtle features inside the traffic flow, so that its internal laws can be fully displayed, and improve further the prediction accuracy. Finally, the gray model combined with the improved background value is adapted to the characteristics of linearity, non-linearity and less required data. Short-term traffic flow is predicted. The experimental results show that the average relative error of the NPCA-PSR-IGM (1,1) combination prediction model is 3.12% smaller than that of the NPCA-PSR-GM (1,1) combination prediction model, and its standard deviation is relative to the PCA-PSR-IGM (1,1) combination prediction model has dropped from 15.7091 to 2.0589. At the same time, compared with the latest prediction model, the combined prediction model also improves the prediction accuracy and achieves a better prediction effect.

Key words: Short-term traffic flow prediction; Multidimensional space-time; Nonlinear Principal Component Analysis (NPCA); Phase Space Reconstruction (PSR); Improved Gray Model (IGM(1,1))

收稿日期: 2020-01-05; 改回日期: 2020-08-22; 网络出版: 2020-09-17

*通信作者: 高贺 gaohe1104@163.com

基金项目: 国家自然科学基金(51577046, 61673153), 教育部科学技术研究重大项目(313018), 安徽省科技计划重点项目(1301022036)

Foundation Items: The National Natural Science Foundation of China (51577046, 61673153), The Key Grant Project of Chinese Ministry of Education (313018), Anhui Provincial Science and Technology Foundation of China (1301022036)

1 引言

智能交通的发展与社会经济的发展和联系愈加紧密,智能交通系统的研究已然成为当今热点研究课题^[1,2]。短时交通流预测是实现交通诱导和控制的关键,同时也是智能交通系统管理的基础^[3,4]。短时交通流具有较强的混沌性和非线性等,各专家针对其特点提出很多智能预测模型和算法。梅朵等人^[5]采用主成分分析(Principal Component Analysis, PCA)与遗传粒子群支持向量机结合的方法来预测交通流量,该方法有效地简化了预测路段时空相关性极强的多个交通流预测变量,但其忽略了交通流量的非线性以及序列的混沌性。王科伟等人^[6]采用多维相空间重构(Phase Space Reconstructed, PSR)对交通流数据进行混沌时间序列分析,但该模型没有考虑城市交通的空间特性以及交通流量的非线性。钱伟等人^[7]采用组合模型的短时交通流量预测,该模型具有训练时间短、抗干扰性强的特点,但该方法所需大量历史交通流量数据,对于采集数据较少的情况,结果未知。佟健颖等人^[8]采用基于时空关联性的深度残差网络预测算法,该算法考虑到交通流量数据的时空相关性并利用深度残差网络进行预测,从而提高预测精度,但其忽略了交通流数据的非线性和混沌特性。

从多维时空交通流量序列的非线性、混沌的内在规律角度来看,以上研究有的忽略了序列的非线性或者序列的混沌性甚至两者均没考虑。综上所述,本文针对这些研究的不足在构建模型时,考虑交通流量序列的非线性的同时结合序列的混沌性,尝试提出一种基于多维时空的非线性主成分分析和相空间重构的改进灰色(NPCA-PSR-IGM(1,1))组合预测模型。并用该模型对合肥市芜湖路与徽州大道交叉口部分交通流量进行预测和对比分析,通过NPCA-PSR-GM(1,1)组合预测模型、PCA-PSR-IGM(1,1)组合预测模型以及文献^[7]和文献^[8]所提出的预测模型的预测值的比较,以验证本文模型在精度提升上的优越性。

2 多维时空NPCA-PSR-IGM(1,1)组合预测模型相关理论

2.1 基于多维时空交通流量相关性的非线性主成分分析原理

对 p 个路段的采样点进行交通流量数据采集,每个路段等时间连续采样 l 次,即 p 为路网中空间的数量, l 为各空间中时间采样次数,可以表示为 Y_1, Y_2, \dots, Y_p ,其中 $Y_i = (y_{i,1}, y_{i,2}, \dots, y_{i,l}), i = 1, 2, \dots, p$ 。计算各观测点与预测点交通流量序列的相关系数 ρ_j ,定义为

$$\rho_j = \frac{E(Y_1 Y_j) - E(Y_1)E(Y_j)}{\sqrt{E(Y_1^2) - E^2(Y_1)}\sqrt{E(Y_j^2) - E^2(Y_j)}} \quad (1)$$

其中, $j = 2, 3, \dots, p$; Y_1 为预测点的交通流量序列; Y_j 为各观测点的交通流量序列; $E(Y)$ 为序列 Y 的数学期望。

记空间降维后的交通流量序列为 Y'_1, Y'_2, \dots, Y'_m ,其中 $Y'_i = (y'_{i,1}, y'_{i,2}, \dots, y'_{i,l}), i = 1, 2, \dots, m$,一般 $0 < m \leq p$ 。也可用 $m \times l$ 矩阵表示,即 $Y'_{m \times l} = [Y'_i]_{m \times l}$ 。首先进行中心化对数比变换处理成标准矩阵 Y^* ,其矩阵元素定义为

$$y_{i,j}^* = \lg(y_{i,j}) - \frac{1}{l} \sum_{j=1}^l \lg(y_{i,j}) \quad (2)$$

建立协方差矩阵 R 并计算特征值和特征向量。

$$R = (R_{i,j})_{l \times l} \quad (3)$$

其中, $R_{i,j} = \frac{1}{m} \sum_{t=1}^m (y_{t,i}^* - \bar{y}_i^*)(y_{t,j}^* - \bar{y}_j^*)$; $\bar{y}_i^* = \frac{1}{m} \sum_{i=1}^m y_{t,i}^*$; $\bar{y}_j^* = \frac{1}{m} \sum_{j=1}^m y_{t,j}^*$ 。求得协方差矩阵 R 的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l$ 和对应的特征向量 u_1, u_2, \dots, u_l 。

方差贡献率和累计方差贡献率分别定义为 $\eta_j = \lambda_j / \sum_{j=1}^l \lambda_j \times 100\%$, $\eta(l) = \sum_{j=1}^l \eta_j$ 。主成分的个数取决于累计方差贡献率,通常累计方差贡献率在85%~95%时对应的前 n 个主成分便包含 l 个原始序列所提供的绝大多数的信息^[9]。 n 个主成分对应的特征向量为 $U_{l \times n} = [u_1, u_2, \dots, u_n]$,则 m 个采样点的 n 个主成分构成的矩阵为 $Z_{m \times n} = Z_{m \times l}^* U_{l \times n}$ 。以此得到主成分降维后的多维交通流量序列,不妨记为 X_1, X_2, \dots, X_m ,其中 $X_i = (x_{i,1}, x_{i,2}, \dots, x_{i,n}), i = 1, 2, \dots, m$ 。

2.2 多维时空交通流量序列的混沌相空间重构原理

已知多维时空交通流量序列经过数据相关性的非线性主成分分析后,将原始序列在空间和时间上降维,得到交通流量序列 X_1, X_2, \dots, X_m ,其中 $X_i = (x_{i,1}, x_{i,2}, \dots, x_{i,n}), i = 1, 2, \dots, m$ 。根据Takens嵌入定理^[10],研究复杂的多维时间序列数据首先需要进行相空间重构。其主要目的是构造出表征原序列的高位空间特性。对上述多维序列重构相空间如式(4)。

$$V_i = [v_{1,i}, v_{2,i}, \dots, v_{m,i}] \quad (4)$$

其中, $v_{j,i} = [x_{j,i}, x_{j,i+\tau_j}, \dots, x_{j,i+(d_j-1)\tau_j}]$; τ_i, d_i 分别为第 i 个交通流量序列的延迟时间和嵌入维数。经过上式的相空间重构,可生成 $M = n - \max_i [(d_i - 1)\tau_i]$ 个相空间点。

将多维时空交通流量序列进行相空间重构后，每个相点代表交通路网的一个状态。假设相点 V_k 的最邻近相点 V_a ，则 V_k 和 V_a 之间的距离为 Z_k

$$Z_k = \min_{i=1,2,\dots,M-1} (\|V_k - V_i\|) = [(v_{1,k} - v_{1,i})^2 + (v_{2,k} - v_{2,i})^2 + \dots + (v_{m,k} - v_{m,i})^2]^{1/2} \quad (5)$$

其中， $v_{j,i} = [x_{j,i}, x_{j,i+\tau_j}, \dots, x_{j,i+(d_j-1)\tau_j}]$; $x_{j,i}$ 表示路段 j 的第 i 个交通流量。

2.3 多维时空相空间重构的改进灰色模型IGM(1,1)

将重构相空间的最邻近相点距演变序列 Z 累加，可使其规律性变得更强^[11]，得到1阶累加序列(1-AGO(Accumulated Generating Operation sequences)) $Z^{(1)}$ ，即

$$Z^{(1)} = [Z_1, (Z_1 + Z_2), \dots, (Z_1 + Z_2 + \dots + Z_M)] \quad (6)$$

$Y^{(1)}$ 是 $Z^{(1)}$ 的紧邻均值生成序列： $Y^{(1)} = (y_2^{(1)}, y_3^{(1)}, \dots, y_M^{(1)})$ ，其中， $y_t^{(1)} = 0.5(Z_{t-1}^{(1)} + Z_t^{(1)})$, $t = 2, 3, \dots, M$ 。

这里采用背景值改进的灰色模型IGM(1,1)预测交通流量在相点空间中相点演变规律。

常规灰色模型GM(1,1)的1阶线性微分方程为

$$\frac{dZ_t^{(1)}}{dt} + aZ_t^{(1)} = b \quad (7)$$

解方程得

$$\hat{Z}_t^{(1)} = \left(Z_1 - \frac{b}{a} \right) e^{-a(t-1)} + \frac{b}{a} \quad (8)$$

其中，参数 a, b ，可按最小二乘法求得。

由式(7)可知，预测精度取决于参数 a, b ，而参数 a, b 的值依赖于背景值 $Y^{(1)}$ 的构造形式。因此，背景值 $Y^{(1)}$ 的值是直接影响灰色模型精度的关键因素。而背景值 $Y^{(1)}$ 一般可以看作 $Z_{t-1}^{(1)}$ 和 $Z_t^{(1)}$ 与横坐标围成的梯形的面积^[12]，通过分析发现常规灰色模型对于序列数据变化平缓时，实际面积值接近梯形面积，因此偏差较小。但当序列出现指数增长趋势时，实际面积与梯形面积值有较大的差值，模型偏差较大。因此，当背景值用面积表示时，其背景值越接近实际面积，模型误差越小。

一般的指数形式为 $y = ae^{b(x-c)} + d$ ，其中 a, b, c, d 均为待定参数，而为了确定其形式至少需要4点，为便于计算这里简化指数一般形式为 $y = ae^x + b$ ，此时有

$$\begin{cases} Z_{t-1}^{(1)} = ae^{t-1} + b \\ Z_t^{(1)} = ae^t + b \end{cases} \quad (9)$$

因此改正后的背景值为

$$\begin{aligned} \bar{y}_t^{(1)} &= \int_{t-1}^t Z_t^{(1)} dt = \int_{t-1}^t (ae^t + b) dt \\ &= \frac{(e-2) \times Z_t^{(1)} + Z_{t-1}^{(1)}}{e-1} \end{aligned} \quad (10)$$

通过上式背景值的改进，建立改进的灰色模型IGM(1,1)。由于改进的背景值比常规的背景值更加接近于实际，所以误差更小。

3 基于多维时空NPCA-PSR-IGM(1,1)组合预测模型构建与算法

本文选定合肥市芜湖路与徽州大道交叉口出口处作为预测点，取样时间为2018年10月15日上午7点到8点45分，时间间隔5 min记录1次。本模型不失一般性，为节省篇幅不再考虑晚高峰。如图1可知，预测点主要受到附近6个观测点的影响，共得6组数据，每组22个时间点交通量作为实验数据，取每组前16个时间点交通流量作为拟合数据，记为 Y_1, Y_2, \dots, Y_6 ，其中， $Y_i = (y_{i,1}, y_{i,2}, \dots, y_{i,16})$, $i = 1, 2, \dots, 6$ 。预测点 Y_1 的后6个时间点的交通流量作为预测对比真实值。

3.1 数据相关性的主成分分析降维

取路网中的6个采样点，分别为观测点1—观测点6，其中观测点1也是预测点。各观测点与预测点按式(1)进行数据相关性分析，计算得到各点相关系数 ρ ，结果见表1。

一般认为相关系数取0.7以上为强相关^[13]。由表1可知，影响预测点交通流量的主要因素有观测点1、观测点2以及观测点3的交通流量，见图1中虚线圈内所示。

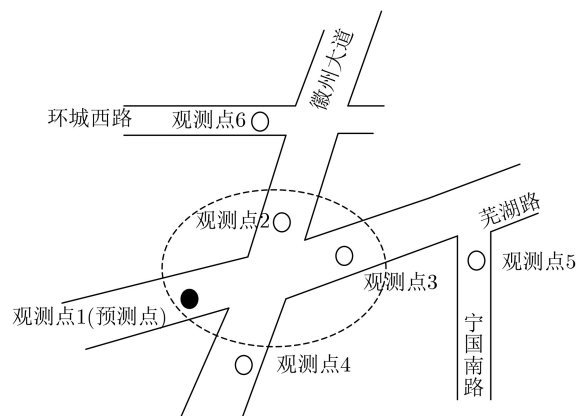


图1 芜湖路与徽州大道交叉口示意图

表1 各观测点与预测点交通流量数据相关系数

观测点序号	1	2	3	4	5	6
ρ	1.000	0.800	0.919	0.601	0.454	0.312

空间降维后的交通流量序列，可用矩阵表示为 $Y'_{3 \times 16} = [y'_{i,j}]_{3 \times 16}$ ，通过计算可知前13个主成分的累计方差贡献率已经达到87.38%，说明只需用这13个主成分就能较好地概括原始交通流量。

3.2 多维时间序列的相空间重构参数的选取

结合多维时空交通流量序列的情况，使用自相关法计算各序列的延迟时间，得 $\tau_i = 1, i = 1, 2, 3$ 。使用GP算法^[14]，对观测点1的交通流量序列算出在不同维数下 $\ln r$ 和 $\ln C(r)$ 的值，将嵌入维数 $d_1 = 1$ 至 $d_1 = 6$ 情况下的曲线绘制如图2所示。

若去除曲线斜率为0和斜率为 ∞ 的直线段，每条曲线所包含的直线斜率，可作为各嵌入维数对应的关联维数，结果如表2所示。

由表2可知饱和关联维数 D 趋于0.5，因 $d_1 \geq 2D + 1$ ，则可确定嵌入维数2。路网中各路段采样点的交通流量序列的嵌入维数按照上述方法计算得到，同时由 $M = n - (d - 1)\tau$ 计算对应的嵌入相点数。由于附近各路段交通环境的相似性和复杂性，一般取各维交通流量序列的嵌入相点数中最小作为统一的相点数^[15]，从而得知，多维时空交通流量序列的嵌入相点为 $M = n - \max_{1 \leq i \leq 3} [(d_i - 1)\tau_i] = 12$ 。

3.3 多维时空的NPCA-PSR-IGM(1,1)组合预测模型

多维时空交通流量序列经相空间重构后得到的相空间相点 $V_i (i = 1, 2, \dots, 12)$ ，计算其最邻近相点距，可以得到一个包含12个最邻近相点距离的序列 Z ，对 Z 作1阶累加得1-AGO，记为 $Z^{(1)}$ ，由最小二乘法计算得到1阶微分方程的参数 a, b ，进而求得序列 Z 的估计值，由式(15)解得预测点的估计值 Y_1^* 。组合预测模型的具体结构图如图3所示。

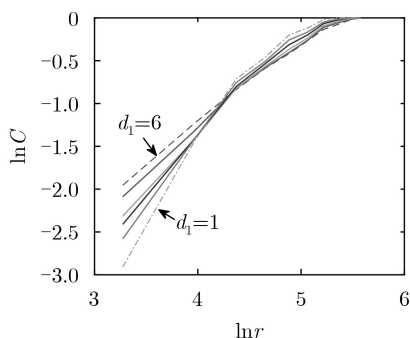


图2 观测点1的交通流量序列 $\ln r - \ln C$ 关系图

表2 观测点1的嵌入维数与关联维数

嵌入维数	关联维数	嵌入维数	关联维数
1	0.06	4	0.36
2	0.09	5	0.47
3	0.28	6	0.49

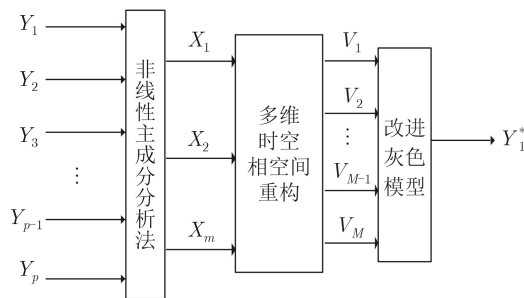


图3 多维时空NPCA-PSR-IGM(1,1)组合预测模型

3.4 基于多维时空的NPCA-PSR-IGM(1,1)组合预测算法

短时交通流的非线性和混沌性特征，很难采用单一的预测方式得到较好的预测结果。本文采用数据相关性的非线性主成分分析和相空间重构的改进灰色组合预测算法对短时交通流进行预测，具体预测步骤如下：

步骤1 对多维时空交通流量序列进行空间降维。首先，对 p 个路段等时间间隔连续采样 l 次，得到多维时空交通流量数据，然后由相关系数求解公式计算各观测点与预测点的相关系数 $\rho_j (j = 2, 3, \dots, p)$ ，从而剔除与目标路段相关性较低的观测点，此时路段数为 $m (m=3)$ 。

步骤2 利用步骤1得到空间降维后的交通流量序列，进行时间维度上的非线性主成分分析。由累计方差贡献率的计算公式确定主成分的个数 $n (n = 13)$ ，相应的累计方差贡献率取85% ~ 95%时，也就说明前 n 个的采样序列已经包含了对预测点影响较大的主要信息。

步骤3 对降维后的交通流量序列进行相空间重构。首先，计算各交通流量序列的延迟时间 τ_i 和嵌入维数 $d_i (i = 1, 2, \dots, m)$ ，其次，由 $M = n - \max_i [(D_i - 1)T_i]$ 从而得到重构后的多维相点数 $M (M = 12)$ ，所以重构后的相空间向量可表示为 $V_i (i = 1, 2, \dots, M)$ 。

步骤4 找到每个相点的最邻近相点并计算出其值，得到相点的最邻近相点距演变序列 Z ： $Z = (Z_1, Z_2, \dots, Z_M)$ 。

步骤5 利用步骤4得到的 Z 进行1阶累加，得到 $Z^{(1)}$ ，如式(8)。以序列 $Z^{(1)}$ 为基础，构造背景值改进的序列，即序列 $\bar{Y}^{(1)} = (y_2^{(1)}, y_3^{(1)}, \dots, y_M^{(1)})$ ，其中， $\bar{Y}^{(1)}$ 的计算采用式(10)得到，即 $y_t^{(1)} = \frac{(e - 2) \times Z_t^{(1)} + Z_{t-1}^{(1)}}{e - 1}$ 。

步骤6 建立

$$\left. \begin{aligned}
 B &= \begin{bmatrix} -y_2^{(1)} & -y_3^{(1)} & \dots & -y_M^{(1)} \\ 1 & 1 & \dots & 1 \end{bmatrix}^T \\
 Y &= [Z_2, Z_3, \dots, Z_M]^T
 \end{aligned} \right\} \quad (11)$$

按照最小二乘法求出 $\hat{a} = [a, b]^T = (B^T B)^{-1} B^T Y$, 解得参数 a, b ($a = -0.0699, b = 226.6869$)。采用灰色模型的1阶微分方程对 $Z^{(1)}$ 序列进行预测得到灰色预测累加还原值为 \hat{Z} 。

步骤7 求得相点距离后, 反演计算相点的位置, 即得交通流量的预测值。设时刻 k 的相点 V_k 的最邻近相点为 V_a, V_k 与 V_a 距离为

$$\hat{Z}_k = \min_{i=1,2,\dots,M-1} (\|V_k - V_i\|) = [(v_{1,k} - v_{1,i})^2 + (v_{2,k} - v_{2,i})^2 + \dots + (v_{m,k} - v_{m,i})^2]^{1/2} \quad (12)$$

其中, $v_{j,i} = [x_{j,i}, x_{j,i+\tau_j}, \dots, x_{j,i+(d_j-1)\tau_j}]$; \hat{Z}_k 可由灰色预测得出, 仅 $x_{k,i}$ 为未知的预测值, 通过解式(12)方程可以得出 $x_{k,i}$ 。

步骤8 输出预测点的预测值。

4 实验结果与分析

4.1 评价标准

本文以平均相对误差和标准偏差为评价模型预测精度的好坏的标准, 平均相对误差和标准偏差公式分别为

$$\bar{\Delta} = \frac{1}{n} \sum_{k=1}^n |\Delta_k| \quad (13)$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{k=1}^n (x(k) - \bar{x}(k))^2} \quad (14)$$

其中, $\bar{\Delta}$ 为平均相对误差; $\Delta_k = \frac{x(k) - \bar{x}(k)}{x(k)}$, Δ_k 为相对误差; σ 为标准偏差; n 为交通流量的个数; $x(k)$ 为第 k 个实际交通流量, $\bar{x}(k)$ 为第 k 个预测交通流量。

4.2 实验对比与分析

本文分别使用基于多维时空的非线性主成分分析和相空间重构的常规灰色(NPCA-PSR-GM(1,1))组合预测模型(如图4)、基于多维时空的主成分分析和相空间重构的改进灰色(PCA-PSR-IGM(1,1))组合预测模型以及基于多维时空的非线性主成分分析和相空间重构的改进灰色(NPCA-PSR-IGM(1,1))组合预测模型进行拟合仿真, 以此分别对预测点后6个时间点(序列号17~22)的交通流量

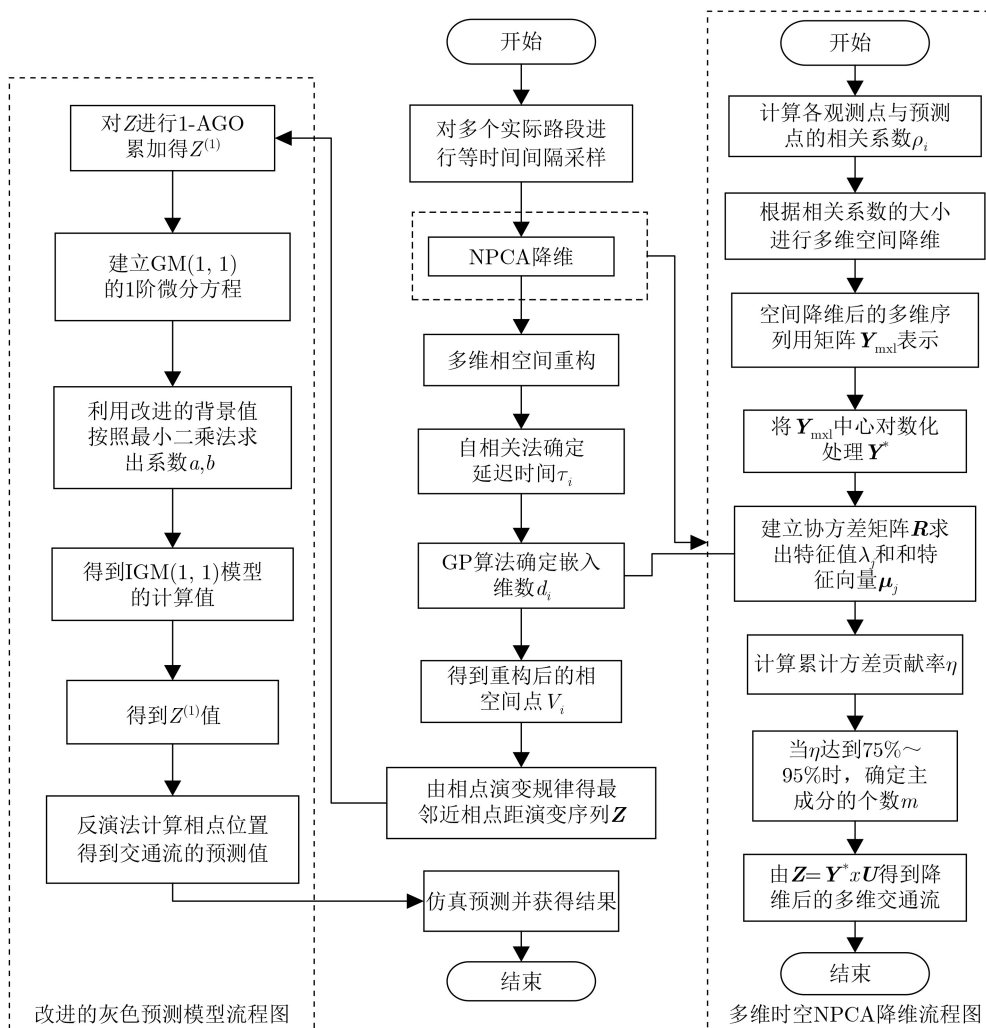


图4 基于多维混沌时空NPCA-PSR-IGM(1,1)组合预测算法流程图

进行预测, 预测点的真实值与3种组合预测模型的预测结果趋势如图5所示。由图5进一步分析可以得到3种组合预测模型预测的平均相对误差和标准偏差如表3所示。

由图5可知, 基于多维时空的NPCA-PSR-IGM(1,1)组合预测模型预测效果更好, 基于多维时空的PCA-PSR-IGM(1,1)组合预测模型其效果最差。由表3可知, 该组合预测模型相对于NPCA-PSR-GM(1,1)组合预测模型和PCA-PSR-IGM(1,1)组合预测模型其预测结果的平均相对误差分别下降了3.12%和7.53%, 同时其标准偏差分别从6.4336下降到2.0589和15.7091下降到2.0589。为了证明本文所提组合模型的优劣, 将组合预测结果与文献[7]以及文献[8]所提出的预测方法进行对比, 预测点的真实值与最新预测模型的预测结果趋势如图6所示。同时, 得出性能指标如表3所示。

由图6的预测结果, 结合表3预测性能指标可以看出, 在实际获取的交通流量数据较少的情况下,

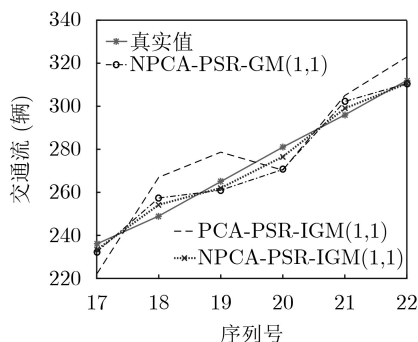


图5 3种组合预测模型预测结果图

表3 3种组合预测模型预测误差统计

组合预测模型	平均相对误差(%)	标准偏差
NPCA-PSR-GM(1,1)	3.33	6.4336
PCA-PSR-IGM(1,1)	7.65	15.7091
文献[7]模型	4.34	9.5833
文献[8]模型	0.69	4.0010
NPCA-PSR-IGM(1,1)	0.12	2.0589

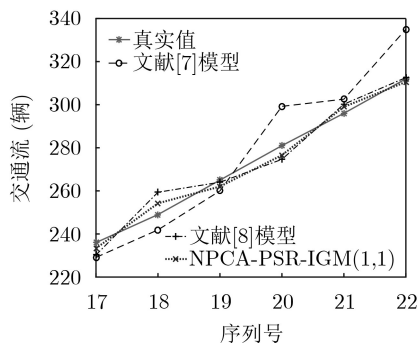


图6 本文组合预测模型与最新模型预测结果

文献[7]所提出的组合预测模型其优势没有得以体现并不能达到较好的预测效果。同时, 文献[8]所提出的基于时空关联性的深度残差网络预测模型, 虽然考虑到短时交通流量数据的时空相关性, 利用大数据进行深度学习构建深度残差模型, 达到不错的预测效果, 但本文所提出的组合预测模型不仅涉及交通流量的时空相关性, 而且考虑到交通流量数据的非线性以及混沌性, 从而进行非线性主成分分析剔除无关数据和利用多维时空相空间重构凸显交通流量数据的内部规律, 进一步提高预测精度。

5 结束语

本文首先运用数据相关性的非线性主成分分析对多维时空交通流量序列进行时空降维, 其次应用多维时空相空间重构放大交通流量序列的内部细微特征, 最后结合改进灰色模型在保留传统模型易于计算的优点的同时在预测精度上也有所提升。通过仿真结果可知, 该组合预测模型相对于多维时空NPCA-PSR-GM(1,1)组合预测模型、多维时空PCA-PSR-IGM(1,1)组合预测模型以及一些最新的预测模型有着更高的预测精度。这种组合模型为短时交通流预测提供了一种新的想法, 但本文没有就天气和不同节假日因素对路段的影响, 因此这将是本文研究的下一步内容。

参考文献

- [1] DARAGHMI Y A, YI C W, and CHIANG T C. Negative binomial additive models for short-term traffic flow forecasting in urban areas[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(2): 784–793. doi: 10.1109/TITS.2013.2287512.
- [2] DAI Guowen, MA Changxi, and XU Xuecai. Short-term traffic flow prediction method for urban road sections based on space-time analysis and GRU[J]. *IEEE Access*, 2019, 7: 143025–143035. doi: 10.1109/ACCESS.2019.2941280.
- [3] 殷礼胜, 唐圣期, 李胜, 等. 基于整合移动平均自回归和遗传粒子群优化小波神经网络组合模型的交通流预测[J]. *电子与信息学报*, 2019, 41(9): 2273–2279. doi: 10.11999/JEIT181073. YIN Lisheng, TANG Shengqi, LI Sheng, et al. Traffic flow prediction based on hybrid model of auto-regressive integrated moving average and genetic particle swarm optimization wavelet neural network[J]. *Journal of Electronics & Information Technology*, 2019, 41(9): 2273–2279. doi: 10.11999/JEIT181073.
- [4] MACKENZIE J, RODDICK J F, and ZITO R. An evaluation of HTM and LSTM for short-term arterial traffic flow prediction[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 20(5): 1847–1857. doi:

- 10.1109/TITS.2018.2843349.
- [5] 梅朵, 郑黎黎, 冷强奎, 等. 基于时空GPSO-SVM的短时交通流预测[J]. 交通信息与安全, 2017, 35(2): 68–74, 120. doi: 10.3963/j.issn.1674-4861.2017.02.010.
MEI Duo, ZHENG Lili, LENG Qiangkui, *et al.* A prediction model for short-term traffic flow based on space-time GPSO-SVM[J]. *Journal of Transport Information and Safety*, 2017, 35(2): 68–74, 120. doi: 10.3963/j.issn.1674-4861.2017.02.010.
- [6] 王科伟, 徐志红. 基于混沌时间序列的道路断面短时交通预测模型[J]. 交通运输工程与信息学报, 2010, 8(1): 70–74. doi: 10.3969/j.issn.1672-4747.2010.01.014.
WANG Kewei and XU Zhihong. Chaotic-time-series-based Short-term traffic flow forecast model of road cross-section[J]. *Journal of Transportation Engineering and Information*, 2010, 8(1): 70–74. doi: 10.3969/j.issn.1672-4747.2010.01.014.
- [7] 钱伟, 车凯, 李冰锋. 基于组合模型的短时交通流量预测[J]. 控制工程, 2019, 26(1): 125–130.
QIAN Wei, CHE Kai, and LI Bingfeng. Short-term traffic flow prediction based on combined models[J]. *Control Engineering of China*, 2019, 26(1): 125–130.
- [8] 佟健颀, 黎英, 王一旋. 基于深度残差网络的短时交通流量预测[J]. 电子测量技术, 2019, 42(18): 85–89. doi: 10.19651/j.cnki.emt.1902768.
TONG Jianjie, LI Ying, and WANG Yixuan. Deep residual network for short-term traffic flow prediction[J]. *Electronic Measurement Technology*, 2019, 42(18): 85–89. doi: 10.19651/j.cnki.emt.1902768.
- [9] 王肖锋, 张明路, 刘军. 基于增量式双向主成分分析的机器人感知学习方法研究[J]. 电子与信息学报, 2018, 40(3): 618–625. doi: 10.11999/JEIT170561.
WANG Xiaofeng, ZHANG Minglu, and LIU Jun. Robot perceptual learning method based on incremental bidirectional principal component analysis[J]. *Journal of Electronics & Information Technology*, 2018, 40(3): 618–625. doi: 10.11999/JEIT170561.
- [10] 睦萍, 郭英, 李红光, 等. 基于混沌吸引子重构和Low-rank聚类的跳频信号电台分选[J]. 电子与信息学报, 2019, 41(12): 2965–2971. doi: 10.11999/JEIT180947.
SUI Ping, GUO Ying, LI Hongguang, *et al.* Frequency-hopping transmitter classification based on chaotic attractor reconstruction and low-rank clustering[J]. *Journal of Electronics & Information Technology*, 2019, 41(12): 2965–2971. doi: 10.11999/JEIT180947.
- [11] MA Ziji, DONG Yanru, LIU Hongli, *et al.* Forecast of non-equal interval track irregularity based on improved grey model and PSO-SVM[J]. *IEEE Access*, 2018, 6: 34812–34818. doi: 10.1109/ACCESS.2018.2841411.
- [12] YIN Kedong, GENG Yan, and LI Xuemei. Improved grey prediction model based on exponential grey action quantity[J]. *Journal of Systems Engineering and Electronics*, 2018, 29(3): 560–570. doi: 10.21629/JSEE.2018.03.13.
- [13] ROOPA H and ASHA T. A linear model based on principal component analysis for disease prediction[J]. *IEEE Access*, 2019, 7: 105314–105318. doi: 10.1109/ACCESS.2019.2931956.
- [14] ZHANG Shuqing, HU Yongtao, JIANG Wanlu, *et al.* Chaos phase space reconstruction based on symbolic analysis and multi-component conditional entropy[C]. 2015 International Conference on Computational Intelligence and Communication Networks (CICN), Jabalpur, India, 2015: 612–616. doi: 10.1109/CICN.2015.125.
- [15] FAN Yuhang, CHANG Dingge, WANG Yanbo, *et al.* Research on partial discharge identification of power transformer based on chaotic characteristics extracted by G-P algorithm[C]. The 2nd International Conference on Electrical Materials and Power Equipment (ICEMPE), Guangzhou, China, 2019: 577–581. doi: 10.1109/icempe.2019.8727289.
- 殷礼胜: 男, 1974年生, 博士, 副教授, 研究方向为复杂系统建模、非线性时间序列预测、交通流预测等。
高 贺: 男, 1993年生, 硕士生, 研究方向为交通流预测、智能控制系统。
魏帅康: 男, 1995年生, 硕士生, 研究方向为交通流预测、复杂系统建模。
孙双晨: 男, 1995年生, 硕士生, 研究方向为交通流预测、智能控制系统。
何怡刚: 男, 1966年生, 博士, 教授, 研究方向为通信信道建模与检测、复杂电磁分析与建模等。

责任编辑: 马秀强