

# 短时交通流复杂动力学特性分析及预测\*

张洪宾<sup>1)2)†</sup> 孙小端<sup>1)</sup> 贺玉龙<sup>1)</sup>

1)(北京工业大学, 交通工程北京市重点实验室, 北京 100124)

2)(德州学院汽车工程学院, 德州 253023)

(2013年7月29日收到; 2013年11月13日收到修改稿)

为揭示短时交通流的内在动态特性, 利用非线性方法对交通流混沌特性进行识别, 为短时交通流的预测提供基础. 基于混沌理论对交通流时间序列进行相空间重构, 利用C-C算法计算时间延迟和嵌入维数, 采用Grassberger-Procaccia算法计算吸引子关联维数, 通过改进小数据量法计算最大Lyapunov指数来判别交通流时间序列的混沌特性. 针对局域自适应预测方法在交通流多步预测中预测器系数无法调节的问题, 提出了交通流多步自适应预测方法. 通过实测数据计算, 结果表明: 2, 4和5 min三种统计尺度的交通流时间序列均具有混沌特性; 改进的小数据量法能够准确地计算出最大Lyapunov指数; 构建的交通流多步自适应预测模型能够有效地预测交通流量的变化. 为智能交通系统诱导和控制提供了依据.

**关键词:** 交通流量, 混沌, 最大Lyapunov指数, 多步自适应预测

**PACS:** 05.45.Tp, 05.45.-a, 01.75.+m

**DOI:** 10.7498/aps.63.040505

## 1 引言

实时准确的短时交通流预测是实现智能交通诱导和控制的重要前提与关键. 短时交通流是指观测尺度不超过15 min的交通流, 从交通流实时管理的角度而言, 希望为城市交通控制系统以及诱导系统等提供短期交通流信息. 交通控制的最大周期一般是3 min, 交通诱导的周期一般为5 min, 因此, 如何在5 min内准确地预测交通流是实现智能交通的关键. 短期交通流预测一般以路段某一断面观测数据所构成的时间序列为基础, 随着观测数据统计时间尺度的缩短, 实时性逐步提高. 交通流时间序列信号由于受到各种复杂因素的制约和影响, 表现出高度的非线性. 其是否存在混沌特性以及精确地识别交通流的混沌特性, 是保证短时交通流预测效果的前提与基础.

由于交通系统受到各种复杂因素的制约和影响, 表现出高度的非线性和复杂性, 难以根据系统

结构建立动力学方程, 通常采用时间序列分析的方法来研究, 通过计算该序列的最大Lyapunov指数来判定是否具有混沌特性. 目前计算最大Lyapunov指数的方法主要有: Wolf法<sup>[1]</sup>, Jacobian方法<sup>[2]</sup>、P范数算法<sup>[2]</sup>、奇异值分解法<sup>[3]</sup>、小数据量算法<sup>[4]</sup>、基于小波变换的最大Lyapunov指数算法<sup>[5]</sup>. 小数据量算法因计算量小, 容易操作, 对小数据组比较可靠, 已得到广泛应用. 文献<sup>[6]</sup>将小数据量法应用于交通流混沌特性的识别; 文献<sup>[7]</sup>通过用虚假临界点法计算嵌入维数完善小数据量法, 研究了微观仿真交通流和实测车头间距时间序列, 但没有将系统相空间中时间延迟和嵌入维数两个参数联系起来.

基于相空间重构思想, 目前已提出了多种预测混沌时间序列的方法, 主要包括: 全局预测法、局域预测法<sup>[8]</sup>、自适应预测法<sup>[9]</sup>、局域自适应预测法<sup>[10,11]</sup>和局域支持向量机预测<sup>[12]</sup>. 自适应预测法只需很少的训练样本就能对混沌序列做出很好的预测, 能自适应地跟踪混沌的运动轨迹, 预测精度较高, 适合

\* 国家重点基础研究发展计划 (批准号: 2012CB723303) 和国家自然科学基金青年科学基金 (批准号: 51308058) 资助的课题.

† 通讯作者. E-mail: hongbin\_1979@163.com

小数据量的情况, 便于实际应用. 已有成果对该方法的多步预测性能讨论很少. 局域自适应预测法结合了局域法和自适应法的优点, 用自适应的方法调节权值, 预测性能大大提高, 文献 [6] 和文献 [13] 基于 Volterra 滤波器的自适应预测方法, 构建了交通流量 Volterra 自适应预测模型, 但该方法在交通流量多步预测时, 因真实值未知而无法调节预测器参数.

本文首先基于混沌理论对交通流时间序列进行相空间重构, 利用 C-C 方法计算嵌入维数和延迟时间, 采用 Grassberger-Procaccia (G-P) 算法计算吸引子关联维数, 通过改进小数据量法计算交通流时间序列的最大 Lyapunov 指数来判断交通流的混沌特性; 针对局域自适应预测方法在交通流量多步预测中无法调节预测器参数, 根据交通流时间序列短期可预测性及自适应算法自适应跟踪混沌运动轨迹的特点, 基于 Volterra 自适应预测模型, 提出交通流多步自适应预测方法; 最后通过实测交通流数据验证算法的有效性.

## 2 交通流混沌特性分析

### 2.1 交通流时间序列相空间重构

相空间重构是用动力系统方法分析非线性时间序列的基础, 也是一种非线性建模方法. 设单变量时间序列为  $\{x(t) | t = 1, 2, \dots, N\}$ , 由延迟坐标状态相空间重构法可得延迟矢量和轨迹矩阵为

$$\mathbf{X} = [\mathbf{X}_1 \quad \mathbf{X}_2 \quad \dots \quad \mathbf{X}_M]^T$$

$$= \begin{bmatrix} x_1 & x_2 & \dots & x_M \\ x_{1+\tau} & x_{2+\tau} & \dots & x_{M+\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & \dots & x_{M+(m-1)\tau} \end{bmatrix}, \quad (1)$$

式中,  $M = N - (m - 1)\tau$  为相点数,  $m$  为嵌入维数,  $\tau$  为时间延迟. 只要  $m$  和  $\tau$  选取恰当, 根据 Takens<sup>[14]</sup> 定理就可以在拓扑等价的意义上恢复原来系统的动力学形态, 相点间的连线是描述系统在  $m$  维相空间中的演化轨迹, 重构后的样本数为 Takens 确定了该相空间嵌入维数的界限值,  $m$  应大于  $2d$ ,  $d$  为该状态空间奇异吸引子的维数.

在重构相空间过程中, 时间迟延和嵌入维数的选取直接影响相空间重构的质量. 嵌入维数太低会出现吸引子的自交性, 太高则使点与点之间的距离

太远. 时间延迟太小, 重构吸引子相邻点的相关性太强, 吸引子的分析很容易被噪声干扰, 时间延迟太大, 本来较近的向量也会拉远, 而导致不确定的系统状态<sup>[15]</sup>.

### 2.2 最佳嵌入维数 $m$ 和延迟时间 $\tau$ 的确定

传统的观点认为嵌入维数和延迟时间是相互独立的, 需要单独进行或依次进行计算. 近来的研究表明, 嵌入维数和时间延迟是紧密相关的, 时间延迟的选取不应独立于嵌入维数, 而应结合时间窗  $\tau_w = (m - 1)\tau$  来确定所求参数值.

C-C<sup>[16]</sup> 算法通过序列的关联积分来构成统计量, 统计量代表了非线性时间序列的相关性, 通过统计量和延迟时间的关系同时得出  $\tau$  和  $\tau_w$ , 从而确定出嵌入维数  $m$ , 关联积分定义为

$$C(m, r, t) = \frac{2}{M(M-1)} \sum_{1 \leq i < j \leq M} H(r - d_{ij}), \quad (2)$$

式中,  $M$  为相点数,  $r$  为邻域半径的大小,  $d_{ij} = \|X_i - X_j\|$  为相空间中两点之间的欧氏距离;  $H(z)$  为 Heaviside 阶跃函数定义为

$$H(z) = \begin{cases} 1 & (z > 0) \\ 0 & (z \leq 0) \end{cases}. \quad (3)$$

将时间序列  $\{x(t), t = 1, 2, \dots, N\}$  划分为  $t$  个不相交的时间序列, 对  $t$  个不相交的子时间序列, 有

$$S(m, N, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, N/t, r, t) - C_s^m(1, N/t, r, t)]. \quad (4)$$

对所有  $S(m, N, r, t)$  取平均值

$$\bar{S} = \frac{1}{M \cdot J} \sum_{m=1}^M \sum_{j=1}^J S(m, r_j, t), \quad (5)$$

式中,  $M, J$  分别为嵌入维数和  $r$  的数目. 定义  $\Delta S(m, t)$  为关于  $r$  的最大偏差, 有

$$\Delta S(m, t) = \max\{S(m, r_j, t)\} - \min\{S(m, r_j, t)\}. \quad (6)$$

根据 Brock 等<sup>[17]</sup> 对几种重要渐进分布的数学统计结论, 在实际中常取  $m = 2, 3, 4, 5$ ,  $r_i = i\sigma/2$ ,  $i = 1, 2, 3, 4$ , 其中  $\sigma$  指时间序列的标准差. 分别计算下面 3 个变量:

$$\bar{S}(t) = \frac{1}{16} \sum_{m=2}^5 \sum_{j=1}^4 S(m, r_j, t), \quad (7)$$

$$\Delta S(t) = \frac{1}{4} \sum_{m=2}^5 \Delta S(m, t), \quad (8)$$

$$S_{\text{cor}}(t) = \Delta \bar{S}(t) + |\bar{S}(t)|, \quad (9)$$

$\Delta \bar{S}(t)$  的第一个极小值对应的时间或  $\bar{S}(t)$  的第一个零点对应的时间为时间延迟  $\tau$ ;  $S_{\text{cor}}(t)$  取得最小值时, 对应的时间为时间窗口  $\tau_W$ . C-C 方法利用统计结果得到的, 容易操作, 计算量小, 具有较强的抗噪声能力.

### 2.3 吸引子关联维数的确定

关联维数是相空间中吸引子复杂度的度量. 计算关联维数比较适用的方法是由 Grassberger 和 Procaccia 提出的 G-P 算法<sup>[18]</sup>, 主要计算过程如下:

1) 利用时间序列  $\{x(t_i), i = 1, 2, \dots, N\}$ , 先给定一个较小的嵌入维数  $m_0$ , 对应一个重构空间

$$Y_{t_i} = [x(t_i), x(t_i + \tau), x(t_i + 2\tau), \dots, x(t_i + (m_0 - 1)\tau)];$$

2) 计算关联函数

$$C(r, m) = \frac{1}{N(N-1)} \sum_{I, J} H(r - \|Y(t_i) - Y(t_j)\|); \quad (10)$$

3) 对于  $r$  的某个适当范围, 吸引子的维数  $d$  与累积分布函数  $C(r, m)$  应满足对数线形关系, 即  $d(m) = \left\lfloor \frac{\ln C(r, m)}{\ln r} \right\rfloor$ , 从而由拟合求出对应于  $m_0$  的关联维数估计值  $d(m_0)$ ; 一般是根据所取得不同  $r$  值和与其对应的  $C(r, m)$  值做出  $\ln C(r, m) - \ln r$  曲线, 其直线部分的斜率就是关联维数值  $d(m_0)$ ;

4) 增加嵌入维数  $m$ , 重复以上步骤, 直到关联维数  $d$  在一定误差范围内不再随着  $m$  的增加而增加, 此时得到的  $d_s = d(m_c)$  就是时间序列的吸引子关联维数.

如果时间序列存在分数维吸引子, 则随着嵌入维数  $m$  的增加, 关联维数  $d$  也增加, 但增加率逐渐减少. 当  $m$  增加到一定时,  $d$  趋于饱和值  $d_s$ ,  $d_s$  就是该时间序列的吸引子关联维数, 通常是非整数,  $d_s$  所对应的最小  $m$  就是描述该时间序列吸引子所需的最小嵌入维数  $m_c$ . 如果  $m_c$  不存在, 那么关联维的估计值  $d(m)$  将随  $m$  的增长无穷大, 从而可知被识别的时间序列是一个随机系统.

### 2.4 交通流量时间序列最大 Lyapunov 指数的计算

Lyapunov 指数  $\lambda$  是定量描述混沌特性的量.  $\lambda < 0$ , 对应于稳定的不动点和周期运动;  $\lambda > 0$ , 对应于混沌吸引子. 为减少计算量, 提高操作的可行性, 采用文献<sup>[19]</sup>给出的小数据量改进算法. 计算步骤如下:

1) 对时间序列  $\{x(t), t = 1, 2, \dots, N\}$  进行快速傅里叶变换, 计算平均周期  $P$ ;

2) 用 C-C 法求时间延迟  $\tau$  和嵌入维数  $m$ , 重构相空间  $\{Y_j | j = 1, 2, \dots, M\}$ ;

3) 找相空间中每个点  $Y_j$  的最紧邻点  $Y_{\hat{j}}$ , 并限制短暂分离, 即

$$d_j(0) = \min_{x_{\hat{j}}} \|Y_j - Y_{\hat{j}}\| \quad (|j - \hat{j}| > p); \quad (11)$$

4) 对相空间中每个点  $Y_j$ , 计算出该邻点对的  $i$  个离散时间步后的距离  $d_j(i)$ ,

$$d_j(i) = |Y_{j+i} - Y_{\hat{j}+i}| \quad (i = 1, 2, \dots, \min(M - j, M - \hat{j})); \quad (12)$$

5) 对每个  $i$ , 求出所有的  $\ln(d_j(i))$  平均  $y(j)$ , 即

$$y(j) = \frac{1}{q\Delta t} \sum_{j=1}^q \ln d_j(i); \quad (13)$$

式中,  $q$  为非零  $d_j(i)$  的数目, 并用最小二乘法作回归直线, 该直线的斜率就是最大 Lyapunov 指数  $\lambda$ . 该方法充分利用了时间序列的空间演变信息, 对小数据组比较可靠, 计算量小, 容易操作.

## 3 交通流多步自适应预测方法

用局域自适应预测模型对已知序列进行单步预测时, 首先需要知道延迟矢量  $\mathbf{X}_n$  的邻域内的最近几个点. 根据重构轨迹. 计算延迟矢量  $\mathbf{X}_n$  与前面的  $n - 1$  个延迟矢量  $\mathbf{X}_i (i = 1, 2, 3, \dots, n - 1)$  的距离

$$d(i) = \|\mathbf{X}_i - \mathbf{X}_n\|_2, \quad (14)$$

找出  $m$  个最近点的  $T$  次迭代后的像  $x_j^*(n + (m - 1)\tau)$  ( $j = 1, 2, 3, \dots, m$ ) 组成重排矢量  $\mathbf{X}_n^*$ :

$$\mathbf{X}_n^* = \begin{bmatrix} x_1^*(n + (m - 1)\tau) \\ x_2^*(n + (m - 1)\tau) \\ \vdots \\ x_m^*(n + (m - 1)\tau) \end{bmatrix}. \quad (15)$$

然后根据  $x_j^*(n + (m - 1)\tau)$  ( $j = 1, 2, 3, \dots, m$ ) 进行线形自适应预测, 得

$$\hat{x}_T(n + (m - 1)\tau + T) = \sum_{j=1}^m w_j(n + (m - 1)\tau) x_j^*(n + (m - 1)\tau). \quad (16)$$

根据最陡下降原理可得递推公式为

$$\begin{aligned} e(n + (m - 1)\tau + T) &= x(n + (m - 1)\tau + T) \\ &\quad - \hat{x}_T(n + (m - 1)\tau + T), \end{aligned} \quad (17)$$

$$\begin{aligned} w_j(n + (m - 1)\tau + 1) &= w_j(n + (m - 1)\tau) \\ &\quad + 2\mu e(n + (m - 1)\tau + T), \end{aligned} \quad (18)$$

其中  $\mu$  为控制算法收敛的系数.

交通流量多步自适应预测方法的计算过程为: 首先根据 (1) 式重构已知  $N$  点序列的轨迹并令  $i = 1$ , 令预测步长  $T = i$ ; 然后根据 (14)—(18) 式用局域自适应预测模型对已知序列进行  $i$  步预测, 由延迟矢量  $\mathbf{X}_M$  得到对  $i$  个将来值的预

测值  $\hat{x}_{T=1}(N + i)$ , 并把  $\hat{x}_{T=i}(N + i)$  加入原始序列, 即令  $x(N + i) = \hat{x}_{T=i}(N + i)$ . 当  $i = i + 1$  时重复计算, 最后输出对将来值的多步预测值  $\hat{x}_{T=1}(N + i)$  ( $i = 1, 2, 3, \dots, k$ ). 该算法根据已知样本得到对将来值的预测值并根据上一步预测值  $\hat{x}_{T=i}(N + i)$  与下一步预测值  $\hat{x}_{T=i+1}(N + i)$  之差自适应调节滤波器系数, 明显好于局域自适应预测算法的多步预测性能, 并且该算法适合小数据量的情况 [20].

### 4 实例分析

实验中的交通流量数据来自北京市主干路天坛东路由南向北方向某路段断面, 采样时间为 2013 年 7 月 15 日—2013 年 7 月 19 日连续 5 个工作日, 观测时间为 7:00—17:00 连续 10 h. 为全面分析短时交通流的复杂动力学特性, 采用 2, 4 和 5 min 三种不同统计尺度记录交通流量数据, 分别得到 1500, 750 和 600 个数据, 流量曲线分别如图 1, 2, 3 所示.

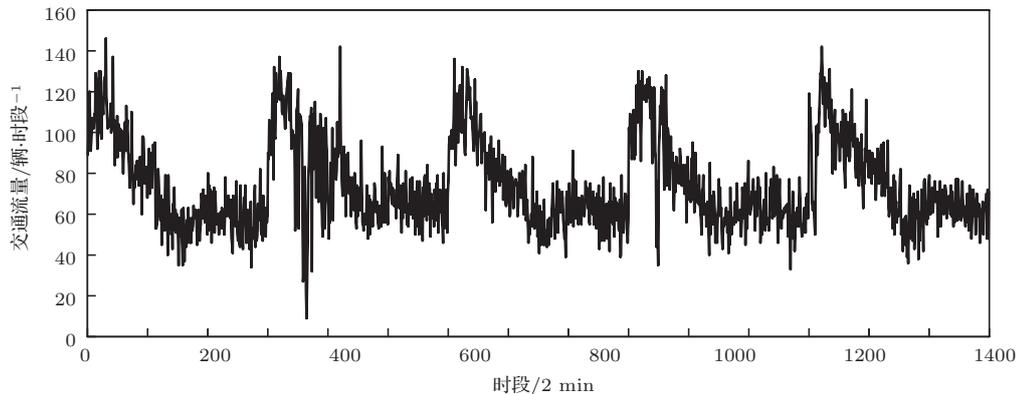


图1 连续5个工作日 2 min 观测尺度交通流量时间序列图

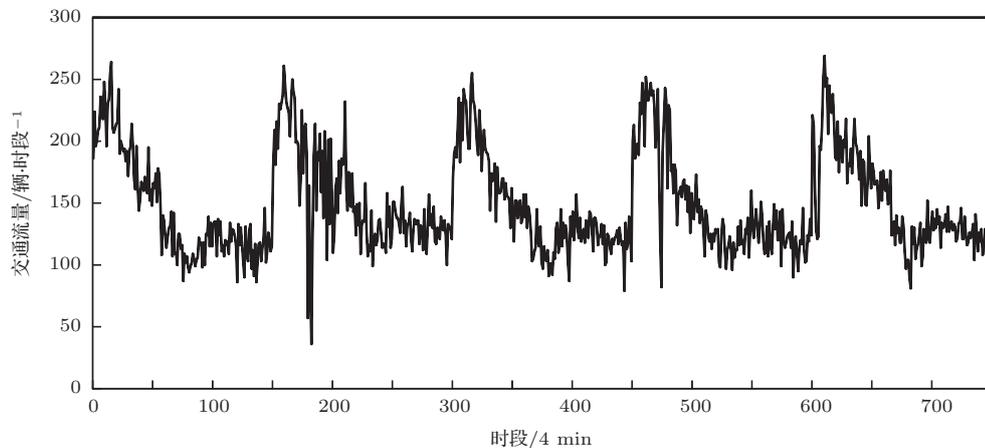


图2 连续5个工作日 4 min 观测尺度交通流量时间序列图

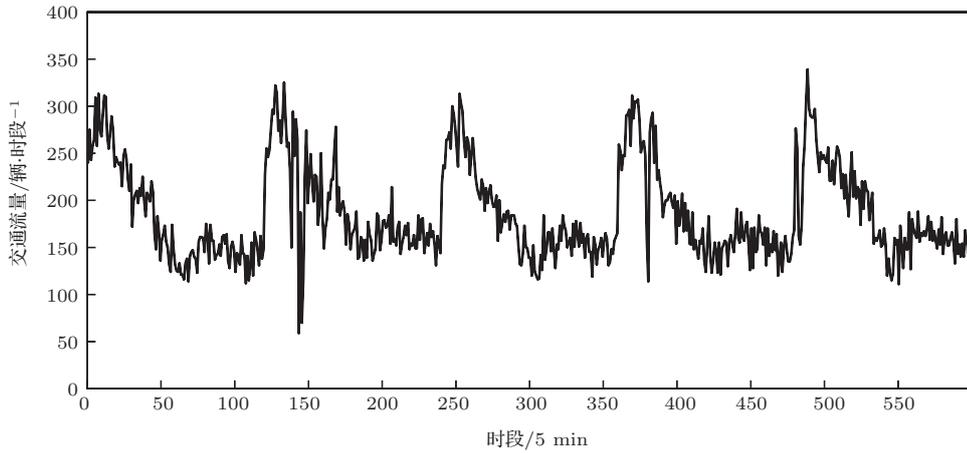


图3 连续5个工作日5 min 观测尺度交通流量时间序列图

由图1, 2, 3连续5个工作日交通流量时间序列曲线可以看出, 三种不同观测尺度的交通流量变化趋势基本相同, 具有一定的日相似性, 偶尔出现特殊情况, 这些特殊情况给短时交通流预测提出了新的挑战.

#### 4.1 交通流混沌特性的识别

如图1交通流量2 min 观测尺度时间序列所示, 由C-C方法计算延迟时间  $\tau = 11$ , 嵌入窗宽  $\tau_w = 25$ , 得出嵌入维数  $m = 3$ ; 根据G-P算法求出饱和关联维  $d_s$  和相应最小嵌入维数  $m_c$ , 图4和图5分别为交通流量2 min 观测尺度的  $\ln C(r) - \ln r$  关系曲线和关联维数  $D(m) - m$  变化曲线; 最后根据小数据量改进法, 计算得到最大Lyapunov指数  $\lambda_1 = 0.3025$ , 为确保结果的可靠性, 同时采用Wolf法计算最大Lyapunov指数为0.4137.

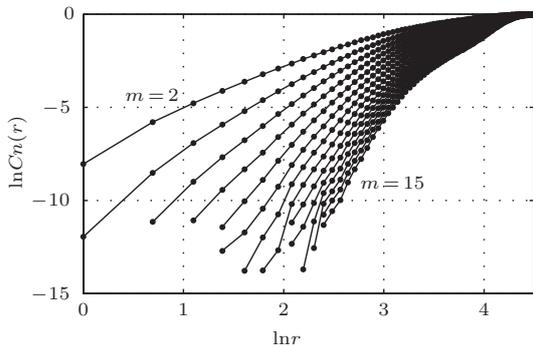


图4 2 min 交通流量  $\ln C(r) - \ln r$  关系

由图4可以看出, 随着嵌入维数  $m$  的增加, 曲线趋于饱和. 由图5可以确定, 饱和关联维数  $d_s = 1.4643$  和最小嵌入维数  $m_c = 3$ , 这说明交通流时间序列有分数维吸引子存在. 通过对最大

Lyapunov指数  $\lambda_1$  的计算, 得出  $\lambda_1 > 0$  的结论, 说明2 min 统计尺度的交通流量时间序列具有混沌特性.

如图2交通量4 min 观测尺度时间序列所示, 由C-C方法计算延迟时间  $\tau = 10$ , 嵌入窗宽  $\tau_w = 26$ , 得出嵌入维数  $m = 4$ ; 根据G-P算法求出饱和关联维  $d_s$  和相应最小嵌入维数  $m_c$ , 图6和图7分别为交通流量4 min 观测尺度的  $\ln C(r) - \ln r$  关系曲线和关联维数  $D(m) - m$  变化曲线; 最后根据小数据量改进算法, 计算得到最大Lyapunov指数  $\lambda_2 = 0.1106$ , 采用Wolf法计算最大Lyapunov指数为0.1388.

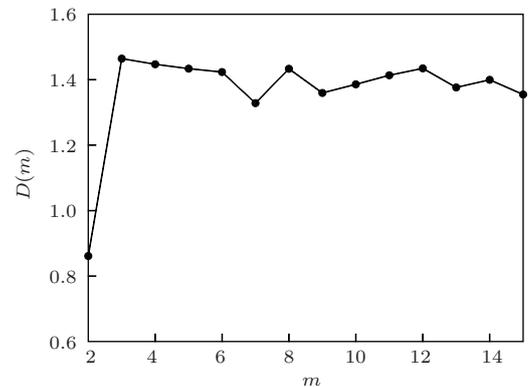


图5 2 min 交通流量  $D(m) - m$  关系

由图6可以看出, 随着嵌入维数  $m$  的增加, 曲线趋于饱和. 由图7可以确定, 饱和关联维数  $d_s = 1.8894$  和最小嵌入维数  $m_c = 4$ , 这说明交通流时间序列有分数维吸引子存在. 通过对最大Lyapunov指数  $\lambda_2$  的计算, 得出  $\lambda_2 > 0$  的结论, 说明4 min 统计尺度的交通流量时间序列具有混沌特性.

如图3交通流量5 min观测尺度时间序列所示,由C-C方法计算延迟时间嵌入窗宽 $\tau_w = 21$ ,得出嵌入维数 $m = 4$ ;根据G-P算法求出饱和关联维 $d_s$ 和相应最小嵌入维数 $m_c$ ,图8和图9分别为交通流量5 min观测尺度的 $\ln C(r)$ - $\ln r$ 关系曲线和关联维数 $D(m)$ - $m$ 变化曲线;根据小数据量改进算法,计算最大Lyapunov指数 $\lambda_3 = 0.1087$ ,采用Wolf法计算最大Lyapunov指数为0.1371.

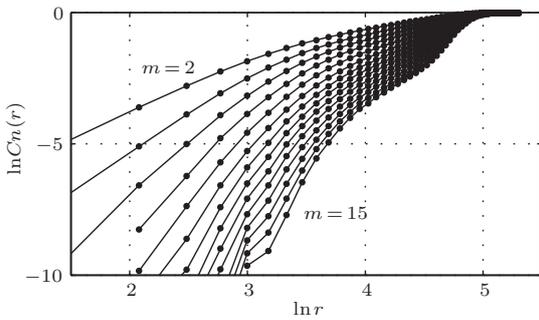


图6 4 min 交通流量  $\ln C(r)$ - $\ln r$  关系

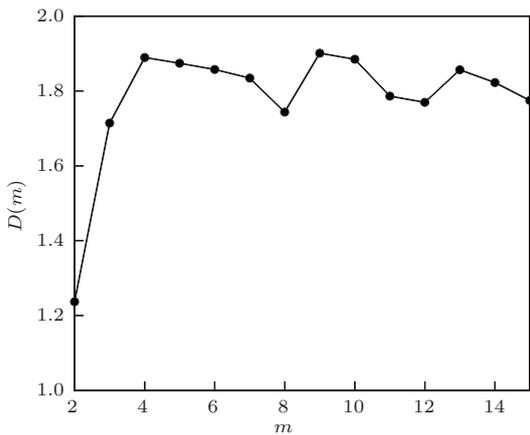


图7 4 min 交通流量  $D(m)$ - $m$  关系

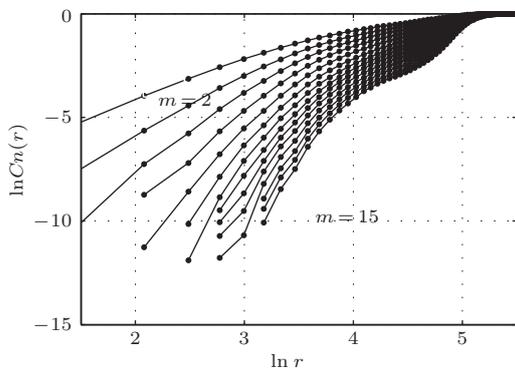


图8 5 min 交通流量  $\ln C(r)$ - $\ln r$  关系

由图8可以看出,随着嵌入维数 $m$ 的增加,曲

线趋于饱和.由图9可以确定,饱和关联维数 $d_s = 1.9293$ 和最小嵌入维数 $m_c = 3$ ,这说明交通流时间序列有分数维吸引子存在.通过对最大Lyapunov指数 $\lambda_3$ 计算,得出 $\lambda_3 > 0$ 的结论,说明5 min统计尺度的交通流时间序列具有混沌特性.

根据以上分析,观测尺度为2,4和5 min的交通流量时间序列都具有混沌特性,也就是说,5 min以内的不同观测尺度对交通流混沌特性没有明显影响,为采用与混沌理论相适应的短时交通流预测方法提供了依据.

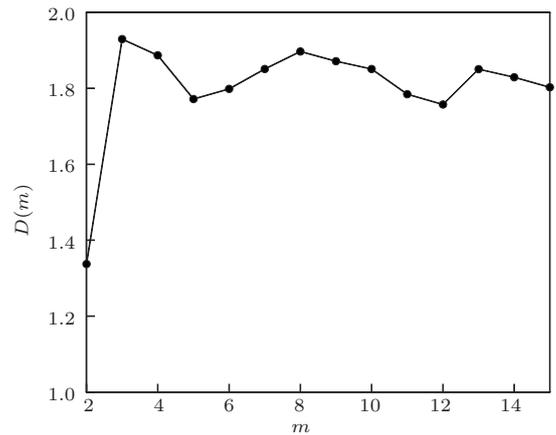


图9 5 min 交通流量  $D(m)$ - $m$  关系

#### 4.2 交通流量多步自适应预测

选用2013年7月15日—2013年7月19日连续5个工作日5 min统计尺度的600个样本中前480个训练预测模型,然后对后120个进行多步预测,预测效果的评价指标采用平均绝对误差MAE (mean absolute error)和平均绝对百分比误差MAPE (mean absolute percentage error).采用Volterra自适应预测模型和改进后的多步自适应预测模型进行预测,预测结果如图10所示,预测误差分析见表1.

由图10可以看出,两种预测方法得到交通流量的预测值与实际值的变化趋势基本一致,但改进的多步自适应预测算法预测值更接近实际值,结果更能反映交通流量变化的趋势和规律;通过表1两种预测方法的误差指标对比,可以看出改进的多步自适应预测算法的预测效果好,可以满足智能交通控制和诱导的要求.另外,交通流量多步自适应预测方法需要很少的训练样本,预测模型简单,适合小数据量交通流量的预测要求.

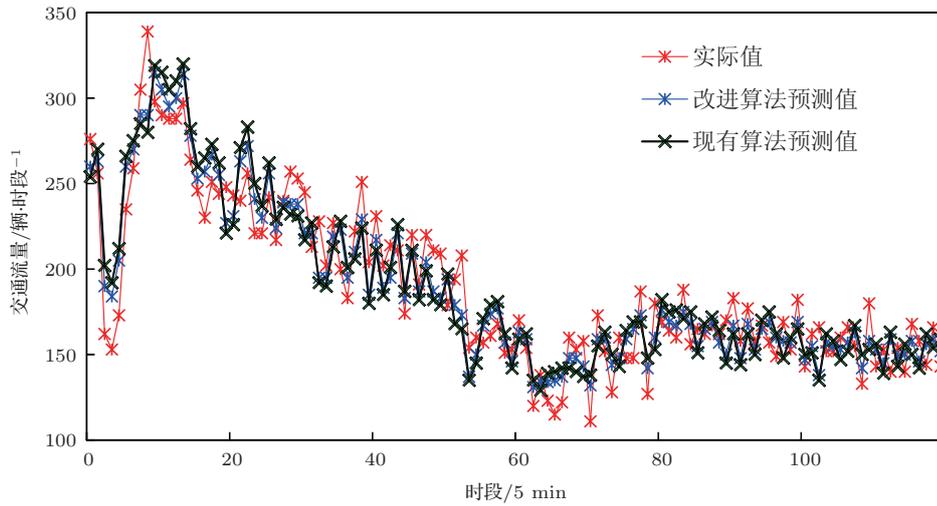


图10 交通流量实际值与预测值对比

表1 两种预测方法误差指标对比表

模型	MAE/辆·5 min <sup>-1</sup>	MAPE/%
现有算法	19	10.04
改进算法	13	7.16

## 5 结 论

1) 通过C-C算法将时间延迟和嵌入维数建立联系,应用改进小数据量法对实测交通流量时间序列进行混沌判定.结果表明,连续5个工作日的2,4和5 min三种统计尺度的交通流量时间序列均具有混沌特性,该方法可以准确地判断交通流时间序列中是否含有混沌特性,对准确掌握不同统计尺度的短时交通流内部演变规律具有重要意义,为采用与混沌理论相适应的预测方法提供了依据.

2) 针对局域自适应预测方法在多步预测中无法调节预测器参数,根据交通流时间序列短期可预测性及自适应算法自适应跟踪混沌运动轨迹的特点,基于Volterra自适应预测模型,提出交通流多步自适应预测方法.实验结果表明,改进的多步自适应预测算法预测性能好于自适应预测算法的多步预测性能,预测精度较高,能够很好反映地交通流量时间变化的趋势和规律.

## 参考文献

[1] Wolf A, Swift J B, Swinney H L, Vastano J A 1985 *Physica D* **16** 285

[2] Barana G, Tsuda I 1993 *Phys. Lett. A* **175** 421  
 [3] Briggs K 1990 *Phys. Lett. A* **151** 27  
 [4] Rosenstein M T, Collins J, Deluca C J 1993 *Physica D* **65** 117  
 [5] Chen Z, Liang P 2000 *J. Guizhou Normal Univ. (Natural Science)* **18** 58(in Chinese) [陈琢, 梁蓓 2000 贵阳师范大学学报(自然科学版) **18** 58]  
 [6] Zhang Y M, Qu S R, Wen K G 2009 *China Civil Engineer. J.* **42** 119 (in Chinese) [张玉梅, 曲仕茹, 温凯歌 2009 土木工程学报 **42** 119]  
 [7] Lu Y, Chen Y H, He G G 2007 *Systems Engineer. Theor. Pract.* **27** 85 (in Chinese)[卢宇, 陈宇红, 贺国光 2007 系统工程理论与实践 **27** 85]  
 [8] Ding T, Zhou H C 2004 *Sys. Engrn. Electron* **26** 338 (in Chinese) [丁涛, 周惠成 2004 系统工程与电子技术 **26** 338]  
 [9] Zhang J S, Xiao X C 2000 *Acta Phys. Sin.* **49** 403 (in Chinese)[张家树, 肖先赐 2000 物理学报 **49** 403]  
 [10] Gan J C, Xiao X C 2003 *Acta Phys. Sin.* **52** 1096 (in Chinese)[甘建超, 肖先赐 2003 物理学报 **52** 1096]  
 [11] Gan J C, Xiao X C 2003 *Acta Phys. Sin.* **52** 2996 (in Chinese)[甘建超, 肖先赐 2003 物理学报 **52** 2996]  
 [12] Zhang J S, Dang J L, Li H C 2007 *Acta Phys. Sin.* **56** 67 (in Chinese)[张家树, 党建亮, 李恒超 2007 物理学报 **56** 67]  
 [13] Zhang Y M, Qu S R 2010 *Appl. Res. Comput.* **27** 4486 in Chinese) [张玉梅, 曲仕茹 2010 计算机应用研究 **27** 4486]  
 [14] Takens F 1981 *Dynamical System and Turbulence, Lecture Notes in Mathematics* (Vol. 898) (Berlin: Springer-Verlag) p230  
 [15] Dong L, Gao S, Liao X Z 2007 *Acta Energiæ Solaris Sin.* **28** 1290 (in Chinese) [冬雷, 高爽, 廖晓钟 2007 太阳能学报 **28** 1290]  
 [16] Kim H S, Eykholt R, Salas J D 1999 *Physica D* **127** 48  
 [17] Brock W A, Hsieh D, Lebaron A B 1991 *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence* (Cambridge: MTT Press) p217  
 [18] Grassberger P, Procaccia I 1983 *Physica D* **9** 1898  
 [19] Lü J H, Lu J A, Chen S H 2002 *Chaotic Time Series Analysis and Applications* (Wuhan: Wuhan University

Press) p116 (in Chinese) [吕金虎, 陆君安, 陈士华 2002 混沌时间序列分析及其应用 (武汉: 武汉大学出版社) 第 116 页]

[20] Meng Q F, Zhang Q, Mu W Y 2006 *Acta Phys. Sin.* **55** 1666 (in Chinese)[孟庆芳, 张强, 牟文英 2006 物理学报 **55** 1666]

## Analysis and prediction of complex dynamical characteristics of short-term traffic flow\*

Zhang Hong-Bin<sup>1)2)†</sup> Sun Xiao-Duan<sup>1)</sup> He Yu-Long<sup>1)</sup>

1) (*Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China*)

2) (*School of Automobile Engineering, Dezhou University, Dezhou 253023, China*)

( Received 29 July 2013; revised manuscript received 13 November 2013 )

### Abstract

In order to reveal the internal dynamic property of short-term traffic flow, the nonlinear analysis method is used to identify the chaotic property of traffic flow which is the basis for the prediction of the traffic flow time series. Traffic flow time series is reconstructed in phase-space based on chaos theory. The embedding dimension and delay time are first calculated via the C-C method. The correlative dimension of attractor is then calculated with the Grassberger-Procaccia method. The largest Lyapunov exponent of traffic flow set is calculated on the basis of the improved small data set method to verify the presence of the chaos in traffic flow time series. A novel multi-step adaptive prediction method is proposed to solve the problem of adjusting the filter parameters of the chaos local adaptive prediction method during traffic flow multi-step prediction. The traffic flow time series are found to have chaotic properties in different statistical scales of 2, 4, and 5 min and show that the improved small data set method can accurately evaluate the chaotic property for traffic flow time series, and that the multi-step adaptive prediction method is capable of effectively predicting its fluctuation, which provides a useful reference for traffic guidance and control.

**Keywords:** traffic flow, chaos, maximum Lyapunov exponent, multi-step adaptive prediction

**PACS:** 05.45.Tp, 05.45.-a, 01.75.+m

**DOI:** 10.7498/aps.63.040505

\* Project supported by the National Basic Research Program of China (Grant No. 2012CB723303) and the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 51308058).

† Corresponding author. E-mail: [hongbin\\_1979@163.com](mailto:hongbin_1979@163.com)