

# 基于混沌理论的交通流短时预测模型

李洪萍, 裴玉龙

(哈尔滨工业大学 交通科学与工程学院, 黑龙江 哈尔滨 150090)

**摘要:** 交通流预测是交通系统可行性分析、交通设计和交通管控的基础, 短时预测是交通流预测的难点. 论文在分析现有交通流预测方法的基础上, 提出了一种基于混沌理论的交通流短时预测方法, 利用基于小数据量的 Wolf 改进算法计算了流率序列的最大 Lyapunov 指数. 将基于 Lyapunov 指数的一维预测模式具体化, 建立了交通流短时预测模型, 并对模型进行了改进, 改进后的预测结果具有较高的精度. 该模型在智能交通系统 (ITS) 的交通控制与诱导方面具有广阔的应用前景.

**关键词:** 交通流时间序列; 混沌; Lyapunov 指数; 短时预测模型

**中图分类号:** U491 **文献标识码:** A **文章编号:** 1007 - 855X (2006) 05 - 0095 - 05

## A Chaos Theory - Based Model of Short - Time Forecasting of Traffic Flow

LI Hong-ping, PEI Yu-bing

(School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China)

**Abstract:** Traffic flow forecasting is the basis of feasibility analysis of traffic system, traffic design, traffic management and control, in which the short - time forecasting is more difficult. The existing traffic forecasting models are analyzed, and a new short - time forecasting model based on chaos theory is developed. The Largest Lyapunov Exponent of flow rate series is estimated with the improved Wolf arithmetic method based on small data sets. A short - time forecasting model is developed from the one - dimensional forecasting mode based on the Largest Lyapunov Exponent. The forecasting precision of the improved model is relatively high and the model has a wide potential in traffic control and inducement in ITS field.

**Key words:** traffic flow series; chaos; Lyapunov Exponent; short - time forecasting model

### 0 引言

交通运输系统的目标确定、可行性分析、方案设计、方案评价与决策都需要了解系统整体或要素的未来发展变化规律, 这就需要借助于科学的预测技术和预测方法<sup>[1]</sup>. 在公路网和城市道路网络规划、道路规模确定及交通管理与控制中, 预测均具有举足轻重的作用.

交通流参数预测属于时间序列预测范畴, 即将观测值按照时间先后排列, 进行分析并对未来趋势进行预测. 交通流参数预测是进行交通管理与控制的前提, 预测的精度直接影响到交通管控措施的效果. 国内外学者对交通预测问题做了大量的研究和探讨, 针对预测目的的不同, 建立了多种模型与方法.

随着交通数据采集技术的发展, 利用视频数据采集设备或车辆传感器已可以直接观测得到交通量、车速的时间序列, 这些数据反映了交通系统在驾驶员、车辆、道路和环境的综合作用下发生的动态变化过程, 这种单一变量时间序列包含了十分丰富的混沌信息<sup>[2]</sup>, 蕴藏着与交通流状态相关联的变量系统的变化痕迹. 本文利用视频数据采集系统获取的高速公路的交通流率数据, 不直接考虑各种随机因素的影响, 而是

收稿日期: 2005 - 06 - 28 基金项目: 教育部博士点基金项目 (项目编号: 20030213030); 哈尔滨工业大学校基金资助项目 (项目编号: HIE 2002 76)

第一作者简介: 李洪萍 (1977 ~), 女, 讲师, 博士研究生. 主要研究方向: 交通流理论、道路通行能力.

E - mail: rockpm@sina.com

直接从流率序列入手,分析数据的混沌特性,计算最大 Lyapunov 指数并建立交通流短时预测模型. 预测结果表明这种方法在最大可预测时间范围内 ( $1/\lambda$ ), 具有较高的精度. 且该法计算简单, 具有很强的实用性, 在未来 ITS 应用中的交通实时控制与交通诱导方面具有广阔的应用前景.

## 1 交通流参数预测方法

常用的交通流预测方法有趋势外推法、移动平均法、指数平滑法、时间序列的分解方法、回归分析法等<sup>[1]</sup>. 趋势外推法依据事物发展的惯性原理, 用某种函数(直线、多项式、指数曲线、生长曲线、包络线等)拟合实测数据点, 对预测对象的某一参数的发展趋势进行外推. 移动平均法把原来时间序列的时距扩大, 采用逐项推移的方法计算时序平均数, 以消除原来序列时距较短受偶然因素影响而引起的波动, 分为一次  $N$  元移动平均法、二次移动平均法和加权移动平均法. 指数平滑法的原理与移动平均法的基本相同, 利用对历史数据进行平滑来消除随机因素的影响, 分为一次指数平滑法、二次指数平滑法和多次指数平滑法. 时间序列的分解方法将观测值序列分解为季节性、趋势性、周期性、随机性几部分并修匀数据, 除随机性外, 其余分别预测, 再组成一复合模型. 这些方法适用于长期预测, 如路段年平均日交通量(AADT)的预测.

近年来, 对交通流短时预测的需求日益增大, 许多学者应用了其它领域中各种成熟的预测模型, 取得了一定的效果, 包括 ARMA 法<sup>[3]</sup>、Kalman 滤波法<sup>[4]</sup>、神经网络方法<sup>[5]</sup>、小波方法<sup>[6]</sup>等. 文献[3]基于采用 ARMA( $p, d, 0$ )模型结构的时间序列分析方法, 提出一种短时交通流实时自适应预测算法, 算法中采用带遗忘因子的递推最小二乘方法进行参数估计, 采用基于线性最小方差预报原理的 Astrom 预报算法进行预报. 文献[4]假设路段当前交通流量是相邻上、下游断面流量的线性函数, 用状态方程和观测方程组成的状态空间模型描述滤波器, 建立了基于 Kalman 滤波理论的交通流预测模型. 文献[5]针对人工神经网络良好的预测能力和传统神经网络的寻优速度慢、实时性差、推广能力低等缺点之间矛盾, 提出一种基于线性独立函数的智能神经元模型, 提高神经网络的信息处理能力, 并将这种智能神经元组成的广义神经网络应用于交通流预测. 文献[6]尝试将小波分析用于交通流短时预测, 探讨了其可行性, 提出了一种基于多分辨率小波分解与重构的交通流短时预测方法. 这些方法理论基础比较成熟, 都认为交通流系统外部的随机因素引起了系统的不确定性, 在预测的过程中, 要对这些随机因素进行统计假设, 假设的合理性对模型精度有很大的影响.

基于混沌理论和相空间重构的预测方法, 可以不直接考虑各种随机因素的影响, 而是直接从实测时间序列入手, 获取系统的混沌特性参数, 进而进行预测. 文献[7]从非线性时间序列预测角度入手, 计算了延迟时间、嵌入维等交通流系统相空间重构参数, 并给出了一种基于相空间重构理论的局部预测方法.

## 2 时间序列的重构相空间

交通流状态变量的观测值  $Y(Q, V$  或  $K$  等) 可以看成是某一时间变量的函数, 即存在一维动力系统方程  $Y = f(x)$ . 由于交通流系统的复杂性, 特别是在同步流状态, 系统具有明显的混沌特性, 很难得到  $f(x)$  的简单、准确的表达式. 传统的研究方法通过引入大量的假设条件将系统进行简化, 得到  $f(x)$  的表达式, 在实际应用中, 参数标定的工作量较大且预测精度和预测有效时间难以掌握, 特别是当预测时段较短的时候.

在时间序列的分析中, 决定序列的可观测因素很多, 而且动力学方程往往是非线性的, 甚至是混沌的. 同时由于测量精度的实际限制、计算的复杂性, 以及可能存在的本质上的非确定性因素等多方面的困难, 严重制约着人们对时间序列内在机制的理解. 20 世纪 80 年代以来, 由于 Takens 对 Whitney 早期在拓扑学方面工作的发展, 使得深入分析时间序列的背景和动力学机制成为可能. 在确定性的基础上, 对序列动力学因素进行分析, 目前广泛采用的是延迟坐标状态空间重构法.

根据混沌理论, 具有  $n$  个状态变量  $x_1, x_2, \dots, x_n$  的系统随时间变化的非线性动力方程为:

$$dx_i/dt = f(x_1, x_2, \dots, x_n) \quad i = 1, 2, \dots, n \quad (1)$$

对于时间序列  $x_1, x_2, x_3, \dots, x_i, \dots$ , 适当选定嵌入维数  $m$  和时间延滞  $\tau$ , 重构相空间:

$$Y(t_i) = [x(t_i), x(t_i + \tau), x(t_i + 2\tau), \dots, x(t_i + (m - 1)\tau)] \quad i = 1, 2, \dots \quad (2)$$

$Y(t_i)$  即为动力系统轨迹. 按照 Takens 定理, 即可以找到一个合适的嵌入维, 即如果延迟坐标的维数  $m > 2d + 1$ ,  $d$  是动力系统的维数, 就可以在拓扑等价的意义下恢复吸引子的动力学特性<sup>[8]</sup>.

### 3 由一维时间序列计算最大 Lyapunov 指数的方法

#### 3.1 Lyapunov 指数

混沌运动的基本特点是运动对初值条件极为敏感. 两个很靠近的初值所产生的轨道, 随时间推移按指数方式分离, Lyapunov 指数就是表示系统在多次迭代中平均每次迭代所引起的指数分离中的指数, 与初始值的选取没有关系, 文献 [8] 对此有清晰的解释.

对于一维动力系统  $x_{n+1} = F(x_n)$  中, 原动力系统的 Lyapunov 指数为:

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| \frac{dF(x)}{dx} \right|_{x=x_i} \quad (3)$$

若  $\lambda < 0$ , 则意味着相邻点最终要靠拢合并成一点, 这对应于稳定的不动点和周期运动; 若  $\lambda > 0$ , 则意味着相邻点最终要分离, 这对应于轨道的局部不稳定, 如果轨道还有整体的稳定因素 (如整体有界、耗散、存在捕捉区等), 则在此作用下反复折叠并形成混沌吸引子. 故  $\lambda > 0$  可作为系统混沌行为的一个判据.

对于一般  $n$  维动力系统, 定义 Lyapunov 指数如下:

设  $F$  是  $R^n \rightarrow R^n$  上的  $n$  维映射, 决定一个  $n$  维离散动力系统  $x_{n+1} = F(x_n)$ , 将系统的初始条件取为一个无穷小的  $n$  维球, 由于演变过程中的自然变形, 球将变为椭球. 将椭球的所有主轴按其长度顺序排列, 那么第  $i$  个 Lyapunov 指数根据第  $i$  个主轴的长度  $P_i(n)$  的增加速率定义为

$$\lambda_i = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left| \frac{P_i(n)}{P_i(0)} \right|, \quad (i = 1, 2, \dots, n) \quad (4)$$

将全部的 Lyapunov 指数谱按大小排列为:  $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ , 其中最小的 Lyapunov 指数决定轨道收缩的快慢, 最大的 Lyapunov 指数则决定轨道发散即覆盖整个吸引子的快慢, 所有指数之和可以认为大体上表征轨线总的平均发散快慢.

#### 3.2 最大 Lyapunov 指数算法

Lyapunov 指数的数值计算方法有许多种, 包括定义法、Wolf 方法、Jacobian 方法、p - 范数方法等. 其中, Wolf 方法和 Jacobian 方法是由单变量时间序列提取 Lyapunov 指数的较实用的方法, 二者都是基于相空间重构的. Rosenstein 等人提出的在 Wolf 方法基础上改进的由小数据量提取 Lyapunov 指数的算法对小数据组比较可靠, 且计算量小、易于应用<sup>[9]</sup>, 计算流程见图 1. 笔者采用文献 [9] 中的算法, 相空间中任意两点的欧氏距离用公式 (5) 计算, 利用 Visual Basic 编制程序.

$$D_{ij} = d(y_i, y_j) = \left[ (x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{im} - x_{jm})^2 \right]^{\frac{1}{2}} \quad (5)$$

#### 3.3 流率时间序列的最大 Lyapunov 指数

论文采用的数据是利用 AutoScope 视频数据采集系统获取的北京市西二环路流率数据, 调查时段为 7.5 h, 双向 6 车道的平均流率为 6 703 veh/h, 平均速度为 44.5 km/h. 绘制密度—流率散点图, 将其与三相交通流理论的基本假设相比较, 发现当时交通流系统处于同步流状态, 通过功率谱和波形分析等方法判断系统具有混沌与分形特性, 利用 G - P 算法得到关联维  $d = 4.80$ , 对应的

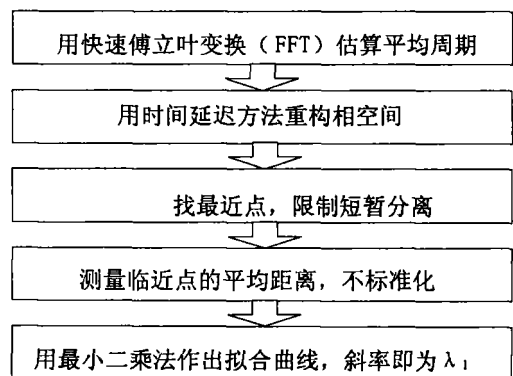


图1 由小数据量计算最大 Lyapunov 指数方法流程  
Fig.1 Flowchart calculating the largest Lyapunov Exponent from small data sets

最小嵌入相空间维数为 7.

将调查得到的单车数据按 2 分间隔进行统计,应用上述基于小数据量的 Wolf 改进算法计算 2 分流率序列的最大 Lyapunov 指数 (λ<sub>1</sub>),根据 Takens 定理, m ≥ 2d + 1,确定 m = 11;计算不同 P 值下的 λ<sub>1</sub> 值,结果表明 P = 4 为临界值.为进一步考虑 m 值对计算结果的影响,取定 P = 4 时,计算不同 m 值下的 λ<sub>1</sub> 值.计算结果见表 1,最后确定 λ<sub>1</sub> = 0.021. λ<sub>1</sub> > 0,进一步表明系统是混沌的,并可以利用基于 λ<sub>1</sub> 的一维预测模式进行交通流特性参数的预测.而最小嵌入相空间维数为 7,表明至少需要 7 个独立变量才能充分描述系统的演化特征,建立这样的模型是极其困难的,即使建成,模型的复杂性也不便于其应用.而基于 λ<sub>1</sub> 的预测模型则克服了这个困难.

### 4 基于最大 Lyapunov 指数的预测模型

#### 4.1 基于最大 Lyapunov 指数的一维预测模式

设参考态为 Y<sub>M</sub>, 它的最近邻态为 Y<sub>N</sub>, 经过 k 次演化后,有 [10]:

$$2^{1 \cdot k} = \frac{Y_{M+k} - Y_{N+k}}{Y_M - Y_N} \quad (6)$$

其中只有点 Y<sub>M+k</sub> 的最后一个分量 x(t<sub>n+1</sub>) 未知,故 x(t<sub>n+1</sub>) 是可预报的.

#### 4.2 预测模型

相应于由小数据量计算最大 Lyapunov 指数的方法,建立交通流短时预测模型:

在 Y<sub>j</sub> (j = 1, 2, ..., n - m + 1) 中寻找最大的下标值 k, Y<sub>k</sub> 的最近邻点 Y<sub>k̂</sub> 的下标 k̂ < k - P, 且 Y<sub>k</sub> 与 Y<sub>k̂</sub> 之间的距离 (利用公式 (5) 计算) 不等于零,即:

$$d_k(0) = \min_{\hat{k}} |Y_k - Y_{\hat{k}}| > 0, \hat{k} < k - P \quad (7)$$

构造向量 y<sub>n-m+2</sub>: {x<sub>n-m+2</sub>, x<sub>n-m+3</sub>, ..., x<sub>n</sub>, x<sub>n+1</sub>}, 以 Y<sub>k</sub> 为基点,向量 Y<sub>k</sub> - Y<sub>k̂</sub> 向前经过 n - m + 2 - k 次演化,到达 Y<sub>n-m+2</sub> - Y<sub>k̂+n-m+2-k</sub>, 此时有:

$$Y_{n-m+2} - Y_{k̂+n-m+2-k} = Y_k - Y_{k̂} \cdot e^{(n-m+2-k) \cdot \lambda_1} = d_k(0) \cdot e^{(n-m+2-k) \cdot \lambda_1} \quad (8)$$

可以求出式中唯一的未知数 x<sub>n+1</sub>.

如果所求 Y<sub>k</sub> 使等式左端向量距离过大,以致不能找到满足条件的 x<sub>n+1</sub> 时,需将此 Y<sub>k</sub> 过滤掉,重新计算寻找最大的下标值 k

将 x<sub>n+1</sub> 视为已知,重复 (1)、(2), 计算得到 x<sub>n+2</sub>, 依次类推以求出 x<sub>n+3</sub>, x<sub>n+4</sub>, ... 当误差会超过允许范围,预测结束.

### 5 应用 Lyapunov 指数进行交通流率预测

将 2 分流率序列的前 100 个值作为已知量,对后面的 125 个点进行预测.分别改变 m、λ<sub>1</sub> 和 P 的值,经过大量数值实验,发现当 m = 11, λ<sub>1</sub> = 0.022, P = 6 时,预测结果最好.

#### 5.1 一次性预测

读入前 100 个实测数据,运行一次程序,预测出后面的 125 个值,预测值与实测值的对比结果见图 2.通过误差计算与分析,发现预测初期的精度较高,随着预测次数的增加,误差稳步增大,然后开始出现不规则的起伏,但总体上是逐渐增大的.误差增大是长时间用预测值作为实测值进行后续预测,误差累计造成的.而后期误差的正负波动是算法中,对不满足条件的 Y<sub>k</sub> 进行过滤造成的.

#### 5.2 逐次预测

为消除一次性预测中,长时间利用预测值进行后续预测造成的误差,定时用实测值取代参与运算的已

表 1 流率序列的最大 Lyapunov 指数计算  
Tab 1 The Largest Lyapunov Exponents of flow rate series

P	m = 11		P = 4	
	λ <sub>1</sub>	m	λ <sub>1</sub>	m
2	0.019	7	0.028	
3	0.020	8	0.025	
4	0.021	9	0.023	
5	0.021	10	0.022	
6	0.021	11	0.021	
7	0.021	12	0.020	

经预测出的值来进行后续运算. 这样在取代的同时, 累计误差回到零点. 不妨选用 6 min 进行一次替换, 即一次运算只预测 3 个流率值, 然后用实测的流率值取代这 3 个值, 预测随后的 3 个值, 依次类推. 预测结果见图 3 可以看出, 同一次性预测相比, 逐次预测的结果与实测值吻合得很好. 进一步进行绝对误差分析, 发现预测值总体上偏低, 需要对上述预测模型进行改进.

### 5.3 模型的改进

根据误差分析结果, 在原预测值的基础上增加一常数项 (26.67), 对模型进行改进. 分析模型改进后的预测绝对误差, 发现改进后的模型预测值的绝对误差的波形与噪声的波形相似, 毋庸置疑, 其中很大一部分是原始数据自身的噪声引起的, 因为此前对实测数据没有进行任何去噪处理. 为此, 利用 Matlab 软件中的滤波函数对误差数据进行处理. 分析滤波后的相对误差 (图 4), 预测前期相对误差基本控制在  $\pm 8\%$  的范围内, 在预测的后期 (约 100 次以后), 连续出现较大的相对误差, 说明已达到最大可预测的时间尺度了.

### 5.4 最大可预测的时间尺度

通常认为轨线分离到原间距的数倍或十几倍, 轨道就不确定了. 由此可以确定混沌运动可预测的最大时间尺度为  $1/\lambda_1^{[18]}$ .  $\lambda_1$  越大, 蝴蝶效应越强, 运动的可预测性越差.

利用前面计算得到的  $\lambda_1$  值, 得到最大可预测时间为 45.45, 即最多可预测 45 个点 (90 min). 改进后的模型的计算结果表明最大可预测时间远远超过了该值, 进一步说明方法的可行性和模型的可靠性.

## 6 结论

采用基于小数据量的 Wolf 改进算法计算了同步流状态下流率序列的最大 Lyapunov 指数, 通过理论与试算找到参数值:  $m = 11, P = 4$ , 得到最大 Lyapunov 指数  $\lambda_1 = 0.021$ .  $\lambda_1 > 0$  进一步验证了交通流系统是混沌系统. 将基于最大 Lyapunov 指数的一维预测模式具体化, 建立交通流短时预测模型. 该法基于观测数据的时间序列客观存在的内在混沌特性进行的, 避免了以往预测方法中主观确定因素的影响, 文中采用及时更新数据进行后续预测、增加常数修正项等方法提高模型的预测精度和可预报时间, 最后通过对误差进行滤波处理以消除原始数据中的高频成分. 模型的预测精度基本在  $\pm 8\%$  范围内, 可预测时间也较长.

论文是将混沌预测模型应用于交通流短时预测的尝试, 模型就短时段、混沌的交通流特性参数预测取得较好的预测结果, 但就下述问题还有待于深入研究:

(下转第 105 页)

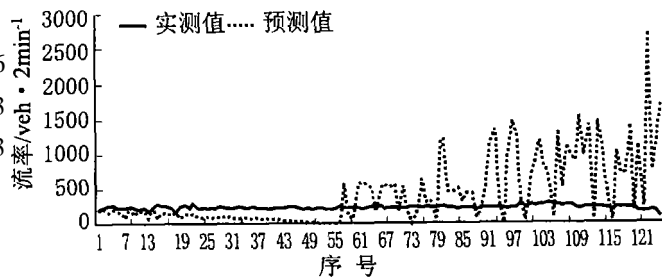


图2 流率序列的一次性预测结果  
Fig.2 Forecasting results of the flow rate series

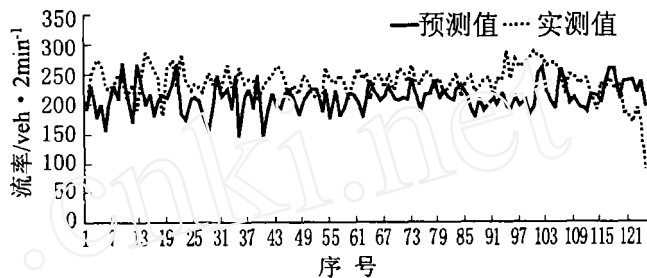


图3 流率序列的逐次预测结果  
Fig.3 Forecasting results of the flow rate series through stepwise operation

- 2) 本文推出的有效功率修正公式简单实用,适合我国国情,可对试验柴油机即时修正.并可对任意工况及整个外特性进行修正;
- 3) 修正公式可以用于高海拔地区,但必需修正标定油量;
- 4) 油耗线法作为一种方法,可用来核算标定功率,但修正系数  $K$  推荐使用式 (2) 或 (15) 计算;
- 5) ISO1585 修正公式不适宜高海拔的有效功率修正.

#### 参考文献:

- [1] Engine Rating Code - Diesel Engine SAEJ - 270[R]. SAE Hand Book, SAE1974.
- [2] 朱倩,冯良华,许忠厚,等.柴油机标定功率的核算方法[J].内燃机工程,1980,1(1):1-10.
- [3] 吴光夏.试论柴油机标定功率的换算方法[J].内燃机工程,1982,3(1):64-72.
- [4] 于瑞涛.汽车柴油机标定功率修正方法的若干问题[J].内燃机工程,1982,(3)1:73-82.
- [5] 王陈生.非增压柴油机功率标定与修正方法[J].内燃机工程,1982,3(1):83-87.

(上接第 99 页)

- 1)  $\rho$  的计算参数取值与模型取得最佳预测效果的取值略有差异,其原因需要通过更多的数值实验来分析确定.
- 2) 模型还有待于进一步改进,以提高预测精度.

#### 参考文献:

- [1] 姚祖康.道路与交通工程系统分析[M].北京:人民交通出版社,1996:125-130.
- [2] 汪树玉,刘国华,杜玉盖,等.大坝观测数据序列中的混沌现象[J].水利学报,1999,(7):22-26.
- [3] 韩超,宋苏,王成红.基于 ARMA 模型的短时交通流实时自适应预测[J].系统仿真学报,2004,16(7):1530-1535.
- [4] Iwao Okotani, Yorgos J Stephanedes. Dynamic Prediction of Traffic Volume Through Kalman Filtering Theory[J]. Transportation Research (Part B), 1994, 1(8B):1-11.
- [5] 谭国真,丁浩.广义神经网络的研究及其在交通流预测中的应用[J].控制与决策,2002,17(11):777-784.
- [6] 贺国光,马寿峰,李宇.基于小波分解与重构的交通流短时预测法[J].系统工程理论与实践,2002,9:101-106.
- [7] 宗春光,宋靖雁,任江涛,等.基于相空间重构的短时交通流预测研究[J].公路交通科技,2003,20(4):71-75.
- [8] 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.
- [9] Michael T. Rosenstein, James J. et al. A Practical Method for Calculating Largest Lyapunov Exponents from Small Data Sets [EB/OL]. Internet Resources: <http://www.physionet.org/physiotools/lyapunov/RosensteinM93.pdf> 1992.
- [10] 黄润生.混沌及其应用[M].武汉:武汉大学出版社,2000:288.