

# 快速路网实时 OD 预测的时间颗粒度选择

杜豫川, 孙轶凡, 陈赣浙

(同济大学 道路与交通工程教育部重点实验室, 上海 200092)

**摘要:** 常见的基于实测数据的 origin-destination(OD)预测方法分为两类:一类基于历史信息,即根据上一天或上一周同一日相同时段的数据进行预测,简称同比预测法;另一类则是根据同一天相邻时间段的数据预测本时段的 OD,简称环比预测法. 预测所用基础数据的时段长度称为时间颗粒度. 时间颗粒度的大小对 OD 预测结果的稳定性、准确性具有重要影响. 针对上海快速路网,采用 ADF 单位根检验和  $K$ -Means 聚类分析方法,研究时间颗粒度对预测结果的影响,提出了时间颗粒度选择的建议,同比预测方法相比环比预测法更容易得出稳定、合理的预测结果,30~60 min 的时间颗粒度预测效果较好.

**关键词:** 时间颗粒度; 快速路 OD 预测; 波动性; 相似性  
**中图分类号:** U495 **文献标志码:** A

## Time Granularity Selection for Expressway OD Real-time Prediction

DU Yuchuan, SUN Yifan, CHEN Ganzhe

(Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 200092, China)

**Abstract:** Two methods, year-on-year prediction method and circularity-ratio prediction method, are often resorted to in the real-time origin-destination(OD) prediction. The former depends on historical information, such as the data about the previous day or week. The latter depends on the data of adjacent period to predict the real-time OD. The size of time granularity, the time period used in the real-time prediction of expressway network has a significant effect on the stability and accuracy of OD prediction results. The Augmented Dickey-fuller Test (ADF) unit root test and  $K$ -means algorithm are employed to study the effect of prediction results, and the size of time granularity is recommended. Compared with circularity-ratio prediction method, year-on-year prediction method can lead to the more stable and accurate result and a more satisfied time granularity is

between 30 minutes and one hour.

**Key words:** time granularity; expressway origin-destination (OD) prediction; volatility; similarity

对城市快速路网而言,每个上下口匝道都可以看成一个起点(O点)或终点(D点). 它的 OD 信息是进行快速路交通运营和建设绩效分析的基础. 随着实测数据的累积,利用这些数据进行 OD 预测已经成为人们获取 OD 的重要途径. 通常根据 OD 矩阵的预测周期可以分为静态预测方法和动态预测方法两大类. 由于实时交通管理的需求,动态 OD 预测逐渐成为当今研究的热点. 对于这类预测方法,国内外学者进行了很多的研究,包括时间序列法、神经网络法<sup>[1-2]</sup>、支持向量机方法、卡尔曼滤波法<sup>[3-5]</sup>、非参数回归方法<sup>[6]</sup>、贝叶斯推论法<sup>[7-9]</sup>等. 目前,学者们对预测方法的研究已经非常深入,但在预测基础数据的时间颗粒度差异对预测结果的影响方面讨论很少. 事实上,快速路流量在不同的时间颗粒度下,对预测结果有很大影响,预测时选择的时间颗粒度大小可能直接影响预测的效果. 颗粒度过小可能造成预测结果的波动,过大则可能造成预测结果的不准确,需要根据实际数据进行科学分析.

动态预测方法根据基础数据的不同可分为两类,第一类方法需要一个先验的 OD 矩阵、路径和交通分配模型,它是基于历史数据的预测方法,以历史的 OD 预测当前时段的 OD,这里将其定义为同比预测法. 另一类则采用时变的路段交通量进行预测,以上一时段的 OD 预测当前时段的 OD,这里将其定义为环比预测法. 无论是同比预测或是环比预测,都需要采用一个时段长度(时间颗粒度)作为预测的基础. 本文将分别采用环比预测法和同比预测法,研究时间颗粒度的大小对快速路 OD 预测结果的影响,

收稿日期: 2015-07-30

基金项目: “十二五”国家科技支撑计划(2014BAG03B05)

第一作者: 杜豫川(1976—),男,教授,博士生导师,工学博士,主要研究方向为智能交通系统、运输经济分析. E-mail: ycdu@tongji.edu.cn

通讯作者: 孙轶凡(1990—),男,硕士生,主要研究方向为智能交通系统. E-mail: sunyf1201@126.com

探讨实时 OD 的预测效果.

### 1 快速路 OD 数据的环比波动性分析

设定一个交通流量统计时段(时间颗粒度),按此时段将快速路流量数据处理成时间序列数据. 相邻时段 OD 预测结果波动性越大,说明该时间颗粒度下的 OD 预测越不可靠. 环比预测需要建立在波动性小、平稳性大的数据基础上,因此在对快速路实时 OD 预测前,需要对不同时间颗粒度下的快速路 OD 数据进行波动性分析. 时间序列的平稳性检验就是一种进行波动性分析的重要手段和方法.

#### 1.1 时间序列的平稳性检验

平稳性是一个很重要的特性,平稳时间序列的基本特性在未来一段时间里将维持不变,而不平稳的时间序列会产生结构变动,而这种结构变动将削弱预测结果的可靠性. 分析表明,上海快速路网上的 OD 交通量存在平稳性,但可能由于时间颗粒度选择不当影响预测结果的平稳性,导致预测的不可靠.

常见的时间序列的平稳性检验方法有以下几种:图形分析法、样本自相关函数检验法、特征根检验法、游程检验法和单位根检验法<sup>[10]</sup>. 其中单位根检验是现代计量经济分析中检验一个时间序列是否平稳的正式方法,包括 DF (Dickey-Fuller Test) 检验, ADF (Augmented Dickey-Fuller Test) 检验, PP 检验, KPSS 检验以及 ERS 检验等多种. ADF 检验是对 DF 检验的扩展,是目前实证研究使用最多的工具. 所以,选择 ADF 单位根检验对快速路 OD 波动性进行分析.

#### 1.2 基于 ADF 检验的快速路 OD 波动性分析方法

在进行波动性分析时,按照不同时间颗粒度可以得到 OD 结果的不同时间序列,再应用 ADF 单位根时间序列平稳性检验方法,得出检验结果.

ADF 单位根检验的回归模型包含没有常数项和时间趋势、仅含有常数项、含有常数项和时间趋势三种情况,检验时,有各自的临界值.

模型 I :含有常数项和时间趋势项

$$\Delta Y_t = \alpha_1 + \alpha_2 t + \rho Y_{t-1} + \beta_i \sum_{i=1}^m \Delta Y_{t-i} + u_t \quad (1)$$

模型 II :仅含有常数项

$$\Delta Y_t = \alpha_1 + \rho Y_{t-1} + \beta_i \sum_{i=1}^m \Delta Y_{t-i} + u_t \quad (2)$$

模型 III :不含常数项和时间趋势项

$$\Delta Y_t = \rho Y_{t-1} + \beta_i \sum_{i=1}^m \Delta Y_{t-i} + u_t \quad (3)$$

式中: $t$  是时间变量; $\alpha$  是常数项; $\alpha_2 t$  是时间趋势项; $u_t$  是残差.

采用  $t$  分布检验  $\rho=0$ , 所得的统计量即为 ADF 值. 在给定的显著性水平下,如果 ADF 统计量小于临界值,则参数  $\rho$  显著地不等于 0,则序列  $Y_t$  不存在单位根,说明序列是平稳的<sup>[11]</sup>.

为了保证不同 ADF 检验平稳性评价结果的准确性,采用 ADF 值的对应概值  $p$  作为评价指标,在一定的显著性水平下( $\alpha=0.05$ ):如果  $\alpha=0.05$ ,则拒绝原假设,序列不存在单位根,序列是平稳的,而且  $p$  值越小,序列平稳性越显著;如果  $p>\alpha$ ,则接受原假设,序列存在单位根,序列是不平稳的,而且  $p$  值越大,序列的不平稳性越显著.

通过平稳性检验结果评价,定量分析了快速路 OD 数据在不同的时间颗粒度下的时间序列的平稳性,而时间序列的平稳性即是快速路 OD 数据的波动性;所以同样的,快速路 OD 数据的波动性也采用 ADF 值对应概值  $p$  作为评价指标进行分析, $p$  值越大,快速路 OD 的波动性越显著, $p$  值越小,则 OD 的平稳性越好,该时间颗粒度下的环比预测效果越好.

基于 ADF 检验的快速路 OD 波动性分析流程设计如下(图 1):

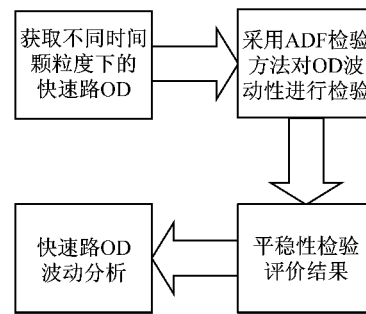


图 1 基于 ADF 检验的快速路 OD 波动性分析流程  
Fig. 1 Analysis process of expressway Origin-Destination (OD) volatility by ADF unit root test

## 2 快速路 OD 数据的同比相似性分析

不同日期的快速路 OD 流量作为两列时间序列,具有一定程度的相似性. 相似性越大的 OD 数据,同比预测的效果越好. 聚类分析就是一种进行相似性分析的重要方法. 本文选择 K-means 聚类分析进行 OD 相似性分析.

### 2.1 K-means 聚类分析

K-means 算法采用距离作为相似性的评价指标,即认为两个对象的距离越近,其相似度越大. 算

法认为簇是由距离靠近的对象组成的,因此把得到紧凑且独立的簇作为最终目标<sup>[12]</sup>.

在 K-means 方法中,  $k$  个初始类聚类中心点的选取对聚类结果具有较大的影响,因为在该算法第一步中是随机的选取任意  $k$  个对象作为初始聚类的中心,初始地代表一个簇.在每次迭代中,对数据集中剩余的每个对象,根据其于各个簇中心的距离将每个对象重新赋给最近的簇.当考察完所有数据对象后,一次迭代完成,新的聚类中心被计算出来.继续迭代,直到结果收敛<sup>[13]</sup>.

### 2.2 基于 K-means 算法的快速路 OD 相似性分析方法

基于 K-means 算法的快速路 OD 相似性分析是将某一时间颗粒度下每一天的数据序列作为样本进行聚类,OD 数据划分成若干的类别,对于不同的分类数,反复应用 K-means 算法聚类,会得出不同的聚类结果.对于不同的聚类结果,其合理性、代表性可采用  $F$  检验方法进行分析评价.

设样本空间由  $m$  个样本组成,每组样本  $r_i$  包含  $n$  个特征观测值,对于快速路 OD 流量而言,观测样本的特征值为一天之内按每个时间颗粒度计算的 OD 流量值( $x$ ). 所得数据表示如下:

$$r_i = (x_1, x_2, \dots, x_n) \quad (4)$$

式中:  $i = (1, 2, \dots, m)$ . 总体样本的中心向量为

$$\bar{r} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) \quad (5)$$

式中:

$$\begin{aligned} \bar{x}_1 &= \frac{1}{m} \sum_{i=1}^m x_{1i}, & \bar{x}_i &= \frac{1}{m} \sum_{i=1}^m x_{ii} \\ \bar{x}_n &= \frac{1}{m} \sum_{i=1}^m x_{ni} \end{aligned} \quad (6)$$

设分类数为  $k$ ,第  $j$  类的样本数为  $m_j$ ,其样本记为  $r_1^{(j)}, r_2^{(j)}, \dots, r_{m_j}^{(j)}$ ,第  $j$  类的聚类中心向量为

$$\bar{r}^{(j)} = (\bar{x}_1^{(j)}, \bar{x}_2^{(j)}, \dots, \bar{x}_n^{(j)}) \quad (7)$$

式中:

$$\begin{aligned} \bar{x}_1^{(j)} &= \frac{1}{m_j} \sum_{i=1}^{m_j} x_{1i}^{(j)}, & \bar{x}_i^{(j)} &= \frac{1}{m_j} \sum_{i=1}^{m_j} x_{ii}^{(j)} \\ \bar{x}_n^{(j)} &= \frac{1}{m_j} \sum_{i=1}^{m_j} x_{ni}^{(j)} \end{aligned} \quad (8)$$

$F$  统计量为

$$F = \frac{\sum_{j=1}^k n_j \|\bar{r}^{(j)} - \bar{r}\|^2 / (c-1)}{\sum_{j=1}^k \sum_{i=1}^{m_j} \|r_i^{(j)} - \bar{r}^{(j)}\|^2 / (m-k)} \quad (9)$$

式中:  $\|r_i^{(j)} - \bar{r}\| = \sqrt{\sum_{x=1}^m (r_x^{(j)} - \bar{r})^2}$  为第  $j$  类样本的样本中心  $\bar{r}^{(j)}$  与总体样本中心  $\bar{r}$  的欧几里德距离.  $\|r_i^{(j)} - \bar{r}^{(j)}\|$  为第  $j$  类样本  $r_i^{(j)}$  总体样本

中心的欧几里德距离. 统计量  $F$  服从自由度为  $(c-1, m-k)$  的  $F$  分布.

在同一  $F$  分布下,  $F$  值越大,表明类间的差别更显著. 在一定的显著性水平下 ( $\alpha = 0.05$ ), 如  $F$  值大于临界阈值  $F(c-1, m-k)$ , 则说明在该显著性水平下,类间的差别显著,类内差别不显著即相似性显著,分类结果是合理的.

为了对不同的分类结果进行准确评价,采用  $F$  值的对应概值  $p$  作为评价指标. 在一定的显著性水平下 ( $\alpha = 0.05$ ): ①多种分类情况下,  $p \leq \alpha$ , 那么  $p$  值最小的那个分类结果是类间的差别与类内相似性最显著的,即其分类结果是最合理,最可接受的; ②多种分类情况下,  $p \leq \alpha$ , 但  $p$  值并不随分类数变化而发生明显变化. 此时可以认为分类数最少的分类结果是最合理,最可接受的.

基于 K-means 聚类算法的快速路 OD 相似性分析流程设计如图 2 所示:

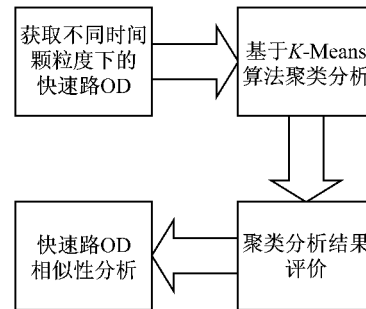


图 2 基于 K-means 算法的快速路 OD 相似性分析流程  
Fig.2 Analysis process of expressway similarity by K-means algorithm

## 3 实例分析与方法比较

### 3.1 研究范围确定

对上海市快速路进行分区,选择其中 6 个交通小区作为研究对象,它们分别是延安中区北侧、延安中区南侧、南北北区西侧、南北北区东侧、内环西北区内圈、内环西北区外圈,包含 30 个 OD 对. 具体小区区域位置及编号信息如图 3 所示.

### 3.2 快速路 OD 波动性分析

OD 数据采用 2008 年 12 月 8 日—2008 年 12 月 21 日的实际 OD 数据. 通过 Eviews6.0 中的 ADF 检验模块对研究范围内确定的 6 个交通小区间的 OD 流量(共计 30 个 OD 对)进行不同时间颗粒度下的波动性分析. 以 1 号小区至 2 号小区的 OD 流量作为分析样例.

对各个时间颗粒度下 O1(延安中区南侧)D2(南

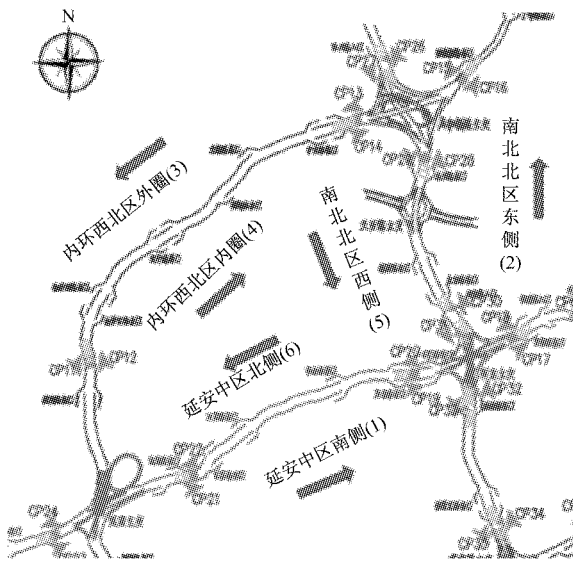


图 3 研究区域位置及编号  
Fig. 3 The location and numbers in study

北北区东侧)流量,取 12 月 8 日数据进行 ADF 检验. 首先取时间颗粒度为 5 min, 检验结果显示, ADF 检验值 = -1.442 574, 对应概值  $p=0.561 3$ , 大于 5% 显著性水平的临界值, 所以接受原假设, 即时间颗粒度为 5 min 时, O1D2 交通量序列是非平稳的.

分析中采用的时间颗粒度为: 5 min, 10 min, 15 min, 30 min, 1 h 和 2 h.

将各个时间颗粒度下 O1D2 流量的 ADF 检验结果进行汇总, 如图 4 所示.

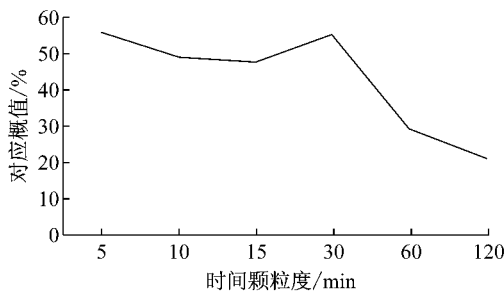


图 4 O1D2 流量波动性变化趋势  
Fig. 4 The trend of O1D2 flow volatility

由图 4 可以发现: 在 5% 的显著性水平下, 5 min, 10 min, 15 min, 30 min, 1 h, 2 h 这几个颗粒度下的 O1D2 流量的波动性都较显著; 随着时间颗粒度的增加, 序列的不平稳程度呈现下降的趋势, 即 O1D2 流量的波动性呈现逐步减弱的趋势.

对剩下的 29 个 OD 对分别进行分析(取 12 月 8 日的数据进行分析), 将对应概值汇总于图 5.

由图 5 可知: 在显著性水平下, 5 min, 10 min, 15 min, 30 min, 1 h 和 2 h 这几个时间颗粒度下的

OD 流量波动性都比较显著; 大部分 OD 对流量在这些时间颗粒度下, 随着时间颗粒度的增大, 波动性逐渐减弱.

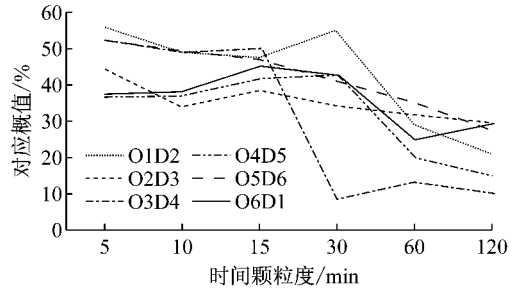


图 5 OD 对流量波动性变化趋势  
Fig. 5 The trend of all OD flow volatility

对于一个时间序列而言, 平稳性是预测的重要基础. 综合以上 6 个交通小区 30 个 OD 对的流量在不同时间颗粒度下的波动性分析, 可以认为: 研究范围内快速路 OD 流量在 5 min, 10 min, 15 min, 30 min, 1 h 和 2 h 这几个时间颗粒度下, 波动性都比较显著, 所以环比预测法进行快速路 OD 进行预测的话, 所得结果的可靠性不高.

### 3.3 快速路 OD 相似性分析

K-Means 算法聚类分析在 SPSS 20.0 环境下进行, 研究对象仍为所确定的 6 个小区(30 个 OD 对), 时间跨度为 12 月 8 日—12 月 21 日; 研究的时间颗粒度为 5 min, 10 min, 15 min, 30 min, 1 h, 2 h, 6 h, 12 h 和 1 d; 聚类数为 2~7 类. 图 6 是 O1D2 聚类对应概值变化的示例.

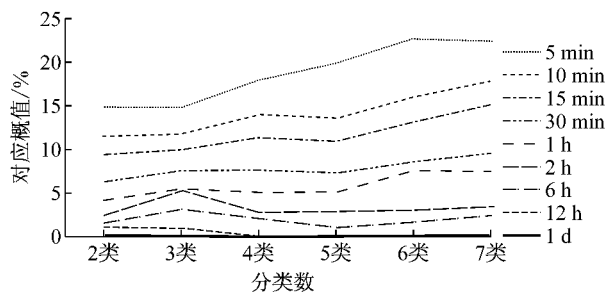


图 6 O1D2 聚类结果  
Fig. 6 The clustering results of O<sub>1</sub>D<sub>2</sub>

通过图 6 以及其他 OD 对不同聚类数的分析, 可以发现: 随着时间颗粒度的增大,  $F$  统计量所对应的概值逐渐变小; 在 5 min, 10 min, 15 min, 30 min, 1 h, 2 h 这几个时间颗粒度下, 随着分类数的增加, 对应的概值呈现上升的趋势; 当时间颗粒度下大于 2 h 时, 对应概值随着分类数增加并不发生明显的变化, 且一直维持在较低的水平.

因此可以认为:时间颗粒度越大,不同日期同一时间段的快速路 OD 流量相似性越显著;对于不同日期同一时间段的快速路 OD 而言,分为两类是可接受的;在一定的显著性水平下( $\alpha=0.05$ ),O1D2 流

量相似性显著的最小时间颗粒度及对应分类数为 1 h(2 类).

对剩余的 29 对 OD 进行分析并将汇总于表 1.

表 1 各 OD 对流量相似性显著最小时间颗粒度及对应分类数

Tab.1 Minimum time granularity and number of classification at the situation of OD flow with significant similarity

起点 \ 讫点	D1	D2	D3	D4	D5	D6
O1	—	1 h(2 类)	10 min(2 类)	10 min(2 类)	15 min(2 类)	5 min(2 类)
O2	10 min(2 类)	—	15 min(2 类)	15 min(2 类)	5 min(2 类)	5 min(2 类)
O3	10 min(2 类)	30 min(2 类)	—	5 min(2 类)	30 min(2 类)	30 min(2 类)
O4	15 min(2 类)	30 min(2 类)	5 min(2 类)	—	15 min(2 类)	10 min(2 类)
O5	1 h(5 类)	30 min(7 类)	30 min(2 类)	1 h(2 类)	—	30 min(2 类)
O6	5 min(2 类)	10 min(2 类)	10 min(2 类)	15 min(2 类)	5 min(2 类)	—

通过表 1 可以发现,对于所研究范围内的 30 个 OD 对而言,相似性显著的最小时间颗粒度及对应分类数有多种情况. 由于对于快速路规划及区域诱导而言,很多时候需要从整个路网的角度进行考虑,通过表 1 可以发现,在研究范围内,30 min 和 1 h 的时间颗粒度可以使 90%以上 OD 达到相似性条件,30 min 以下的时间颗粒度相似性显著的比例较低.

因此可以认为:30 min(2 类)是同比预测中的快速路 OD 相似性显著的最小时间颗粒度及对应分类数;30 min 以上的时间颗粒度下的快速路 OD 相似性都是显著的.

表 2 是研究范围内 O1D2 流量在 30 min(2 类)情况下的聚类成员具体情况.

表 2 O1D2 聚类成员

Tab.2 O1D2 clustering members

数据日期	聚类	距离	数据日期	聚类	距离
12月8日(周一)	1	253.367	12月15日(周一)	1	271.716
12月9日(周二)	1	270.158	12月16日(周二)	1	193.414
12月10日(周三)	1	175.431	12月17日(周三)	1	198.663
12月11日(周四)	1	167.457	12月18日(周四)	1	143.724
12月12日(周五)	1	170.790	12月19日(周五)	1	253.237
12月13日(周六)	2	156.218	12月20日(周六)	2	159.648
12月14日(周日)	2	151.072	12月21日(周日)	2	181.613

通过表 2 不难发现在 30 min(2 类)情况下,所有工作日的 OD 数据都被分到了同一类中,而周末的数据分在另一组. 结合研究范围内的所有 OD 对的聚类成员具体情况,可以发现:29 个 OD 对的工作日被分入了同一类内,占总数的 96.7%. 因此在 30 min 的时间颗粒度下,可以用工作日的快速路 OD 同比预测工作日,用双休日的快速路 OD 同比预测双休日的快速路 OD.

### 3.4 快速路 OD 同比环比预测对比

从波动性分析中,可以得到:

(1) 环比预测时,在小于 2 h 的颗粒度下,快速路 OD 流量的波动性都较显著;随着时间颗粒度的增加,快速路 OD 的波动性呈现逐步减弱的趋势.

(2) 同比预测时,时间颗粒度越大,快速路 OD 流量相似性越显著;30 min 以上的时间颗粒度下的快速路 OD 相似性都是显著的.

(3) 用工作日 OD 同比预测工作日 OD 的效果、双休日 OD 同比预测双休日 OD 的效果较好.

以上分析的实际交通状况不包含交通事故,分析结果表明,在一天中快速路的波动性较大,引起这种波动性的原因可能是人们在一天里的通勤时间不一,这种波动性既可能是时间颗粒度的影响,也可能反应了不同时段内 OD 确实不同的事实,这也是环比预测法的波动性比较大的重要原因. 在相似性分析中,不同工作日同一时段的快速路 OD 具有较好的相似性,说明在不同工作日,不同的出行者趋向于选择各自稳定的出行时间. 而周末由于没有工作影响,人们出行的相似性并没有工作日明显. 结合波动性和相似性的对应概值,如图 7 所示.

从图 7 可以发现:环比预测需要较大的时间颗粒度才能让 OD 波动性下降到合适的波动范围内. 而同比预测则在一个较小的时间颗粒度下,OD 就能

达到合适的波动范围内。

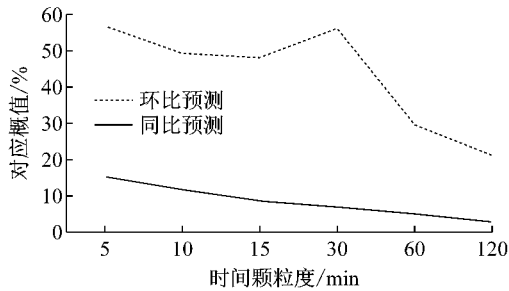


图 7 O1D2 流量波动性变化趋势

Fig.7 The trend of Origin1-Destination2 flow volatility

在快速路交通管理中,常常需要一个时段内的 OD 流量预测.因此可以认为:在没有突发交通事故的情况下,比起利用一天内实时数据的环比预测,利用历史数据中同一时段的快速路 OD 同比预测更容易得到较合理的结果.

## 4 结论

在研究时间颗粒度对预测结果影响的基础上,采用上海快速路 OD 数据对环比预测方法和同比预测方法进行了比较,提出了预测时采用的时间颗粒度建议.得出如下结论:

(1)时间颗粒度对于 OD 预测结果的稳定性、准确性具有显著的影响.

(2)采用环比预测所得结果的波动较大,所需要的时间较长;利用同一时段的历史数据进行同比预测,选择 30~60 min 的时间颗粒度的 OD 预测效果可达到较好精度.

(3)与环比预测方法相比,同比预测方法更容易得出稳定、合理的预测结果.

由于 OD 数据受限于时间跨度,在分类中无法对较多分类进行分析.同时,因为缺乏突发事件下的时间颗粒度选择的分析,该研究还存在不足,有待进一步完善.

## 参考文献:

[1] Musson Lorenzo, Grant-Mullersusan, 陈海波. 基于线圈数据的高速公路 OD 矩阵预测神经网络法[J]. 交通运输系统工程与信息, 2010, 10(1): 88.  
Musson Lorenzo, Grant-Muller Susan, Chen Haibo. A neural network approach to motorway OD matrix estimation from loop counts[J]. Journal of Transportation Systems Engineering and Information Technology, 2010, 10(1): 88.

[2] 羊文琦. 基于遗传算法和 BP 神经网络的区域性公路交通量

预测研究[D]. 成都:西南交通大学, 2015.

YANG Wenqi. Prediction of highway traffic volume based on genetic algorithms and BP neural network [J]. Chengdu: Southwest Jiaotong University, 2015.

- [3] Ashok K, Ben-Akiva M. Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems[C]// International Symposium on Transportation and Traffic Theory. Amsterdam: Elsevier, 1993: 465-484.
- [4] Ashok K. Dynamic trip table estimation for real time traffic management systems[D]. Cambridge: MIT, 1992.
- [5] 吴鼎新. 实时 OD 预测中卡尔曼滤波算法的应用[C]// Proceedings of International Conference on Engineering and Business Management (EBM2010). 武汉: 武汉大学, 2010: 3756-3758.  
WU Dingxin. Application of Kalman Filtering method to real-time OD estimation [C] // Proceedings of International Conference on Engineering and Business Management (EBM2010). Wuhan: Wuhan University, 2010: 3756-3758.
- [6] 董春娇. 多状态下城市快速路网交通流短时预测理论与方法研究[D]. 北京: 北京交通大学, 2011.  
DONG Chunjiao. Theoretical research for short-term traffic flow prediction in multi traffic states on urban expressway network[D]. Beijing: Beijing Jiaotong University, 2011.
- [7] 王金梅. 道路交通动态 OD 矩阵预测方法的研究[J]. 宁夏工程技术, 2002, 1(4): 362.  
WANG Jinmei. The study and analysis of model algorithmic for dynamic origin-destination matrix estimation and prediction [J]. Ningxia Engineering Technology, 2002, 1(4): 362.
- [8] 邵昀泓, 程琳, 王炜. 最大熵模型在交通分布预测中的应用[J]. 交通运输系统工程与信息, 2005, 5(1): 83.  
SHAO Yunhong, CHEN Lin, WANG Wei. Application of entropy-maximizing(EM) model to traffic distribution forecast [J]. Journal of Transportation Systems Engineering and Information Technology, 2005, 5(1): 83.
- [9] 于秋红. 基于贝叶斯方法的分层优化 OD 估计模型[J]. 公路, 2014(11): 123.  
YU QiuHong. A hierarchical optimization model to OD estimation based on Bayesian method[J]. Highway, 2014(11): 123.
- [10] 王立平, 万伦来, 等. 计量经济学理论与应用[M]. 合肥: 合肥工业大学出版社, 2008.  
WANG Liping, WAN Lunlai. Theory and application of econometrics [M]. Hefei: Hefei University of Technology Publisher, 2008.
- [11] Dickey D A, Fuller W A. Distribution of the estimators for autoregressive time series with a unit root[J]. Journal of the American Statistical Association, 1979, 74: 427.
- [12] 张建辉. K-means 聚类算法研究及应用[D]. 武汉: 武汉理工大学, 2007.  
ZHANG Jianhui. Research and application of K-means clustering algorithm [D]. Wuhan: Wuhan University of Technology, 2007.
- [13] 张建萍, 刘希玉. 基于聚类分析的 K-means 算法研究及应用[J]. 计算机应用研究, 2007, 24(5): 166.  
ZHANG Jianping, LIU Xiyu. Cluster analysis-based K-means algorithm and its application [J]. Application Research of Computers, 2007, 24(5): 166.