

## 中文摘要

**摘要：**高效的交通是现代化城市发展不可缺少的部分。自上世纪后半叶，全球许多大城市在发展进程中逐渐遇到拥堵严重、效率低下等交通问题以及由此衍生出的一系列环境和社会问题，这在一定程度上制约了城市的发展。在这一背景下，人们提出了智能交通系统（ITS）的概念。交通检测技术作为智能交通系统的基础部分，对整个系统的运行效率起到重要作用。现代检测技术主要以断面检测器、车牌识别和浮动车为主，其中浮动车数据的获取成本相对较低，测量范围大，目前在采集行程时间数据的方法已获得广泛应用；

本文首先对现有的交通检测技术，尤其是行程时间采集方法进行了综述，在此基础上重点阐述了现有的行程时间的断面数据和浮动车数据估计方法。针对现有技术的不足，提出了一种基于小样本浮动车数据的信号灯路段行程时间估计方法，该方法能够较好的克服在浮动车小样本数据前提下，信号灯延误造成的数据波动带来的估计误差。最后，通过在实际道路上进行交通调查验证本方法的有效性，分别采用断面检测和浮动车检测方式对多个路段交通状态进行调查并估计路段行程时间，结果显示浮动车估计行程时间与断面估计行程时间相似，且采用本算法后行程估计时间误差较采用直接平均法减小。说明本文提出的方法能够有效的提高行程时间估计精度，在工程实践中具有应用价值。

**关键词：**浮动车；小样本；行程时间估计；信号灯路口；交通检测技术

## ABSTRACT

**ABSTRACT:** An Effective transportation system is important for development of modern city. From 70's of last century, transportation problems such as severe congestion and low efficiency, which are followed by a series of environment and social problems, restrict the development speed of modern cities to some extent. In view of this point, the concept of Intelligent Transportation System (ITS) is introduced to the whole world. Traffic detection technique, as a fundamental component of many applications in ITS, has important impact on the operation efficiency of ITS applications.

This thesis introduces traffic detection technique, especially on travel time collection methods existed nowadays firstly. Then different travel time estimation methods including estimated from traffic detectors and from float car data are overviewed in detail. By pointing out problems in current methods, a travel estimation method based on floating car data of small sample size is introduced in this thesis. This method can reduce errors caused by data fluctuation induced by signal period in short arterial road in center area of big cities. Finally, based on real project, empirical analyses are provided to prove the efficiency of this method. Analysis result shows the consistency of travel time estimated from detectors and float car data. Moreover, comparison between errors in estimated travel time from this method and common method gives another proof for method efficiency. Conclusions are drawn that the method provided in this thesis can improve travel time estimation accuracy.

**KEYWORDS:** Floating car; small sample size; travel time estimation; signalized intersection; traffic detection technique

## 图目录

图 2-1 检测器安装位置示意图.....	3
图 2-2 GPS 定位系统示意图.....	4
图 2-3 检测器安装位置示意图.....	7
图 3-1 有交叉口检测路段示意图.....	14
图 4-1 鲁班路立交至漕溪北路匝道高架和地面主干道地图.....	21
图 4-2 318 国道（沪青平公路）从外环至赵巷镇道路地图.....	22
图 4-3 中山南二路检测器设备安装示意图.....	28
图 4-4 三一八国道检测器设备安装示意图.....	28
图 4-5 路段 1（龙华西路-天钥桥路和）行程时间原始点（ $\Delta$ :遇到红灯的点， ○:未遇到红灯点）、修正点（●）.....	30
图 4-6 路段 1（龙华西路-天钥桥路和）样本点简单平均曲线（--）、修正后的 样本点平均曲线（:）.....	30
图 4-7 路段 1（龙华西路-天钥桥路和）检测器估计行程速度与浮动车估计行 程速度对比.....	34
图 4-8 路段 32（航东路-环西一大道）检测器估计行程速度与浮动车估计行程 速度对比.....	34
图 5-1 主干道断面 D1 雷达测速仪数据与 MC 数据比较.....	53
图 5-2 主干道断面 D2 雷达测速仪数据与 MC 数据比较.....	53
图 5-3 主干道断面 D3 雷达测速仪数据与 MC 数据比较.....	53
图 5-4 318 国道断面 D1 雷达测速仪数据与 MC 数据比较.....	54
图 5-5 318 国道断面 D2 雷达测速仪数据与 MC 数据比较.....	54
图 5-6 318 国道断面 D3 雷达测速仪数据与 MC 数据比较.....	54
图 5-7 主干道断面 D1 雷达测速仪与 MC 样本组数据比较.....	55
图 5-8 主干道断面 D1 MC 数据相对误差.....	55
图 5-9 主干道断面 D1 相对误差分布曲线.....	55
图 5-10 主干道断面 D2 雷达测速仪与 MC 样本组数据比较.....	56
图 5-11 主干道断面 D2 MC 数据相对误差.....	56
图 5-12 主干道断面 D2 相对误差分布曲线.....	56
图 5-13 主干道断面 D3 雷达测速仪与 MC 样本组数据比较.....	57
图 5-14 主干道断面 D3 MC 数据相对误差.....	57
图 5-15 主干道断面 D3 相对误差分布曲线.....	57
图 5-16 国道断面 D1 雷达测速仪与 MC 样本组数据比较.....	58
图 5-17 国道断面 D1 MC 数据相对误差.....	58

图 5-18 318 国道断面 D1 相对误差分布曲线 .....	58
图 5-19 国道断面 D2 雷达测速仪与 MC 样本组数据比较 .....	59
图 5-20 国道断面 D2 MC 数据相对误差 .....	59
图 5-21 318 国道断面 D2 相对误差分布曲线 .....	59
图 5-22 国道断面 D3 雷达测速仪与 MC 样本组数据比较 .....	60
图 5-23 国道断面 D3MC 数据相对误差 .....	60
图 5-24 318 国道断面 D3 相对误差分布曲线 .....	60
图 5-25 主干道平均法、调和平均法、自适应加权平均法路段平均行驶时间 .....	61
图 5-26 318 国道平均法、调和平均法、自适应加权平均法路段平均行驶时间 .....	61
图 5-27 中山南二路相位图 .....	64
图 5-28 三一八国道相位图 .....	65

## 学位论文版权使用授权书

本学位论文作者完全了解北京交通大学有关保留、使用学位论文的规定。特授权北京交通大学可以将学位论文的全部或部分内容编入有关数据库进行检索，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。

(保密的学位论文在解密后适用本授权说明)

学位论文作者签名：段子昂

导师签名：姜伟

签字日期：2008年6月5日

签字日期：2008年6月5日

## 独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京交通大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：段子勇      签字日期：2008年6月5日

## 致谢

本论文的工作是在我的导师关伟教授的悉心指导下完成的，关伟教授严谨的治学态度和科学的工作方法给了我极大的帮助和影响，使我更加明白了如何发现和解决问题，如何做人做事。在此衷心感谢多年来关伟老师对我的关心和指导。

王喜富教授对于我的科研工作和论文都提出了许多的宝贵意见，在此表示衷心的感谢。

在实验室工作及撰写论文期间，何蜀燕、刘惠玲、王均、贾洪旭、郑鑫、张学炜等同学对我论文中的研究工作给予了热情帮助，在此向他们表达我的感激之情。

另外也感谢我的家人，他们的理解和支持使我能够在学校专心完成我的学业。

# 1 引言

## 1.1 研究背景

高效的交通是现代化城市发展不可缺少的部分。在城市发展进程中，人口和经济的不断增长产生了大量的交通需求，然而，交通基础设施由于受到自然和经济资源的限制，发展缓慢，不能满足高速增长的交通需求。自上世纪后半叶，全球许多大城市在发展进程中逐渐遇到拥堵严重、效率低下等交通问题以及由此衍生出的一系列环境和社会问题，这在一定程度上制约了城市的发展。在这一背景下，人们提出了智能交通系统（ITS）的概念，它采用信息技术、计算机技术、控制技术等手段对传统交通运输系统进行改造，从而达到增强系统运行效率、提高系统的可靠性和安全性的目的。30年来，为了解决日益严重的交通拥堵问题，智能交通在世界各地的应用日趋成熟，其中美国、日本和欧洲已成为世界智能交通系统的三大基地。90年代初，我国学者开始关注国际上智能交通系统的发展，1995年以后，我国关于智能交通系统的研究、试验和国际交流日益频繁。交通部已将智能交通系统列入中长期规划。

为了使智能交通系统领域达到最高的运行效率，新思想和新技术不断融合到新的应用中，目前已经涉及的领域包括先进交通管理系统（ATMS）、先进出行者信息系统（ATIS）、先进公共运输系统（APTS）、商用车辆运营系统（CVO）、先进车辆控制和安全系统（AVCSS）、先进巡航辅助高速公路系统（AHS）和先进乡村运输系统（ARTS）等。城市集成交通系统作为先进交通管理系统的子系统，主要在大城市等交通密集地区得到广泛应用，通过先进的测量、监控、信息和控制方法，为出行者和其他道路使用者以及交通管理人员提供实时、适当的交通信息和最优路径诱导，试图使交通流保持最佳状态。这一系统的基础是实时可靠的交通流数据，只有基于高质量的实测交通流数据，控制系统才能够做出正确合理的决策，诱导交通流达到合理分布。目前获取实时交通数据的方法主要有通过在路边埋设的检测器，如线圈、视频设备、微波检测器等，以及移动式检测方案，如浮动车（floating car, or probe vehicle）<sup>[1-7]</sup>和手机获取行程时间等方法。其中路边埋设检测器成本高，维护复杂，不易大规模铺设；浮动车数据往往通过安装在出租车/公交车上的GPS收发装置获得，成本相对较低，测量范围大，但对于城市郊区的覆盖程度较差，目前浮动车采集行程时间数据的方法已获得广泛应用<sup>[8-15]</sup>；手机获取行程时间数据为近年来正在兴起的交通信息采集手段，该方法几乎不需要

成本，即使在郊区道路也有较高的手机持有率，但由于这一方法的行程时间估计精度受到路网分布及基站布设影响较大，以及安全、隐私等问题影响，还没有得到广泛的应用。基于道路行程时间的交通流诱导控制手段包括为车载 GPS 设备推荐行驶路线、在路边的可变信息牌（variable message sign, VMS）上显示交通状态和估计行程时间、以及将实时交通状态显示在网页上供用户随时参考等。这些发布的交通诱导信息直接影响交通参与者的出行方式和路径，影响城市道路交通流分布和交通运行畅通程度。因此，如何获得准确可靠的交通数据，并且据此估计路段行程时间成为交通诱导控制系统研究的关键问题之一。

## 1.2 论文研究目标

尽管随着 GPS 定位精度的提高，浮动车技术用于交通数据检测已经得到了越来越广泛的应用，但在路网密集的大城市市区，这一方法仍存在一些待解决的问题，如带信号灯路段行程时间估计精度问题。由密集路网中路段长度较短，车辆的信号灯延时对车辆行程时间造成较大影响，则通过样本平均方法得到的车辆平均行程时间会与真实值偏差增大，降低数据的可靠性，尤其在样本点数量不足的情况下，样本均值具有偶然性，不能反映真实道路状况。针对这一问题，本文提出了一种估计带信号灯路段行程时间的方法，这一方法结合了相对易获得的信号灯相位信息，有效降低了估计误差，使路段行程时间估计值更加接近真实值，能够提高交通诱导与控制系统的运行效率。

## 1.3 论文结构和主要内容

根据研究内容，本文一共分为五章：

第一章 引言，主要介绍研究背景和论文研究目标。

第二章 相关研究综述，对当前浮动车技术和小样本估计方法进行简要介绍。

第三章 小样本路段行程时间估计方法，对本文的核心算法进行介绍，阐述了算法的主要原理，计算过程并进行误差分析，从理论角度证明该算法能够提高估计精确度。

第四章 实证分析，建立测试系统，通过实地采集浮动车数据和基准数据，比较路段行程时间估计值，从实证角度证实算法的有效性。

第五章 结论，对算法进行总结，提出不足及改进方向。

## 2 相关研究综述

### 2.1 行程时间采集方法

行程时间和旅行速度是交通诱导和控制的重要信息，目前采集行程时间的主要方法有浮动车法、车牌识别方法、断面检测器法等。

#### 2.1.1 断面检测器法

检测器通常安装在路边，检测附近断面的流量、车速、占有率等数据。断面检测器采集的信息不能直接提供行程时间数据，需要经过转换。通常情况下，一个路段的行程时间需要通过安装在该路段上的多个断面检测器数据推算获得。在假定路段车速均匀的情况下，路段行程时间可由路段长度和路段平均速度计算出。检测器安装示意图如图 2-1 所示。

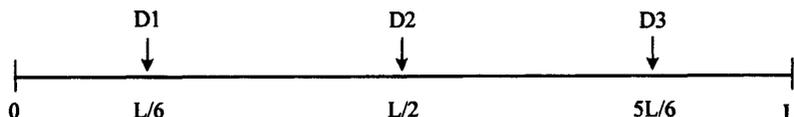


图 2-1 检测器安装位置示意图

假定路段总长度为  $L$ ，平均每 300 米至 500 米布设一个检测器，检测器编号分别为  $D_1, D_2, \dots, D_n$ ，则为了最大程度的利用每个检测器的数据，宜将检测器均匀铺设，即检测器  $D_i$  的安装位置为  $\frac{(2i-1)L}{2n}$ 。图 2-1 为  $n=3$  时的检测器安装位置示意图。在实际情况中由于受具体路况限制，检测器安装在接近理论值的地方即可。

常用的断面交通检测器包括线圈检测器、微波检测器、视频检测器、以及各种便携式检测器，如气压管检测器、地磁检测器等。

#### 2.1.2 浮动车法

浮动车即为在道路上跟随车流行驶的车，在车辆行驶时由人工或车载设备记录下一系列信息，如定位信息、速度信息、车辆 ID、时间等，利用这些信息可以估计路段行程时间、交通状态等重要交通数据。浮动车数据收集方法包括人工记录车辆经过地点和时间、信标定位、三角定位和 GPS 定位方法等<sup>[16]</sup>。

人工记录方法通过车辆在预先设定的道路上行驶，直接采集车辆行程时间。

其他车载装置通过通过电磁技术、雷达技术、GPS 定位以及电信蜂窝网络等实现位置、时间、速度信息等采集工作，据此间接获得车辆行程时间。其中 GPS 定位方法在大规模浮动车数据采集中得到广泛应用，GPS 定位法利用全球定位系统获取定位、方向和速度信息，结构示意图见图 2-2。

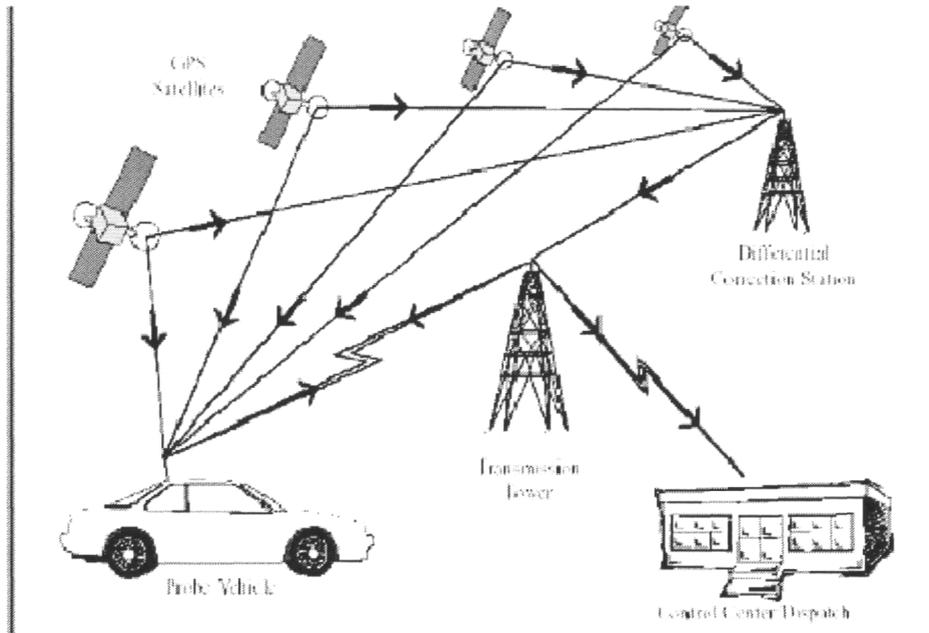


图 2-2 GPS 定位系统示意图

利用 GPS 定位法的浮动车检测系统的优点在于：

- 1) 数据采集的成本低廉并且基础设施的建设投入将非常少。
- 2) GPS 定位数据精度高。
- 3) GPS 产品已日趋成熟，GPS 数据格式有统一标准，各类 GIS 软件都提供 GPS 数据识别接口。

使用 GPS 采集的数据，应用地图匹配、路径推测等相关的计算模型和算法进行处理，将浮动车位置数据和城市道路在时间和空间上关联起来，最终得到浮动车所经过道路的车辆行驶速度以及道路的行车行程时间等交通拥堵信息。如果在城市中部署足够数量的浮动车，并将这些浮动车的位置数据通过无线通讯系统定期、实时地传输到一个信息处理中心，由信息中心综合处理，就可以获得整个城市动态、实时的交通拥堵信息，这些信息进一步将通过各种形式进行发布，如控制中心显示，VMS 显示，网页形式发布等。

### 2.1.3 车牌识别方法<sup>[17]</sup>

车牌识别方法包括人工车牌识别和自动车牌识别两种方法，其核心思想是识别完整行经指定路段的车辆，通过记录同一车辆的驶入时间和驶出时间，进而计算出车辆在该路段的行程时间。

人工牌照法通常在调查路段的起点和终点，每个运行方向设置一个由观察员和记录员组成的二人小组。在数据采集开始前，先根据标准时间校准记录用表，数据采集开始后，观察员报出在指定方向上通过的交通流中汽车牌照号码的最后三位数字、车型以及秒表的读数。目前，也可通过人工交通调查仪记录车牌信息，由于调查仪有自动记录时间功能，消除了秒表记录容易产生的误差，可以获得更持久和准确的数据。人工牌照法能够较准确的测得不同路段的行程时间、各种车型的平均行程速度及车头时距。但此方法的局限在于①中能测得位于起点和终点之间的行程时间；②多车道路段上由于视野受到阻碍，容易错记或漏记车牌号码；③劳动强度大，对人员要求高。

自动牌照识别法通过在道路重要监测点安装车牌照识别设备（LPR），利用车牌识别系统的摄像设备，识别出通过监测点的车辆信息及车牌照，据此可获得车辆密度、队长、排队规模等交通信息，防范和观察交通事故。这种方法除了可以对行程时间进行识别外，还可以统计多种交通参数，为交通诱导系统提供必要的交通流信息。

## 2.2 行程时间估计方法

### 2.2.1 由断面检测数据估计行程时间

断面检测器不能直接检测行程时间，只能测量流量、速度和占有率数据。目前由断面检测器数据估计行程时间的方法主要有两种：第一种方法是先估计点速度，再将其转换成行程时间，这一方法的前提是路段上行驶的车辆为等长并且匀速<sup>[18-23]</sup>，在此基础上，一些研究人员提出了不需要等速度假设的行程时间估计方法<sup>[23,425]</sup>，但是，这种方法在速度估计和行程时间转换方面都易产生误差。另一种方法基于宏观或微观的随机交通流模型，直接根据交通流特点将检测器时间转换成行程时间<sup>[24-27]</sup>。若目标路段上安装有  $n$  个检测器，则该路段行程时间可根据路段上所有检测器的点速度得到，算法可采用平均法、调和平均法和自适应加权平均法<sup>[28]</sup>。对于有信号灯的路段，行程时间中还需加上延误时间，延误时间由道路容量、车辆最大流出率、道路饱和度和相位等因素决定<sup>[29-31]</sup>。

#### 1) 断面点速度估计法

早期交通检测方法中以线圈检测为主，单线圈检测器仅能提供车流量  $N(t)$  和占有率  $O(t)$ ，因此研究人员提出了点速度的估计方法  $\hat{s} = \hat{n} / g\hat{o}$ ，其中  $\hat{s}, \hat{n}, \hat{o}$  分别为速度、流量和占有率， $g$  为一常数修正系数。

若考虑到车辆速度方差，则速度、占有率和车流量间的关系为

$$\frac{N_i}{O_i} \left( \frac{\bar{l}}{T} \right) \left[ \frac{\sigma_s^2}{\bar{s}^2} + 1 \right] = \bar{s} \quad (2-1)$$

进一步若考虑到测量误差，则有

$$\frac{N_i}{O_i} = \frac{\bar{s}T}{\bar{l}} \left[ \frac{\bar{s}^2}{\sigma_s^2 + \bar{s}^2} \right] - \frac{\Delta O_i}{O_i} \frac{\bar{s}T}{\bar{l}} \left[ \frac{\bar{s}^2}{\sigma_s^2 + \bar{s}^2} \right] \quad (2-2)$$

其中  $\sigma_s^2$  为速度方差， $\Delta O_i$  为占有率测量误差， $\bar{l}$  为平均车长<sup>[20]</sup>。

估计出点速度后，即可根据路段长度计算行程时间。

## 2) 结合交通流特征直接计算行程时间

结合交通流特征的方法如文献<sup>[31]</sup>中所述，对于一段有一个入口匝道和一个出口匝道的路段，使用断面检测数据估计路段行程时间公式如下：

$$tt(t) = \frac{\Delta x}{v(t)} = \frac{\Delta x}{q(t)} \cdot k(t) \quad (2-3)$$

其中

$$q(t) = \alpha[q_u(t) + q_{on}(t)] + (1 - \alpha) \cdot [q_d(t) + q_{off}(t)] \quad (2-4)$$

$$k(t) = k(t-1) + \frac{1}{L \cdot \Delta x} \cdot (q_u(t) + q_{on}(t) - (q_d(t) + q_{off}(t))) \quad (2-5)$$

$tt(t), v(t), q(t), k(t)$  分别为  $t$  时刻的估计行程时间、速度、流量和密度； $q_u(t), q_{on}(t), q_d(t), q_{off}(t)$  分别为  $t$  时刻上游流量、入口匝道流量、下游流量和出口匝道流量， $L$  为车道数量， $\Delta x$  为路段长度。

## 3) 由多个点速度估计行程时间

假定测量路段总长度为  $L$ ，在其上布设  $n$  个检测器，检测器编号分别为  $D_1, D_2, \dots, D_n$ ，若要最大程度的利用每个检测器的数据，则检测器  $D_i$  的安装位置应为  $\frac{(2i-1)L}{2n}$ 。但在实际情况中由于受具体路况限制，检测器只能安装在接近理论值的地方。设检测器  $D_i$  的实际安装位置为  $L_i$ ，每个检测器的有效间距为  $l_i$ ，则有

$$l_i = \frac{L_{i+1} - L_{i-1}}{2} \quad (i = 0, 1, \dots, n+1) \quad (2-6)$$

其中  $L_0$  和  $L_{n+1}$  为辅助变量：

$$\begin{cases} L_0 = -L_1 \\ L_{n+1} = L + L_n \end{cases} \quad (2-7)$$

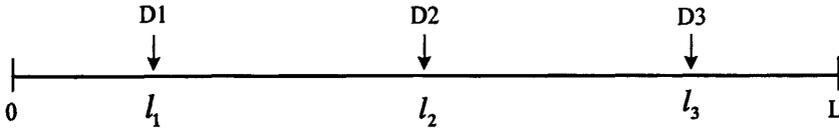


图 2-3 检测器安装位置示意图

采用三种算法将断面速度转化为车辆平均行驶速度：算术平均法、调和平均法和自适应加权平均法。

设在采样时间段 \$t\$ 内经过检测器 \$D\_i\$ 的车辆数为 \$N^{(i)}\$，记第 \$k\$ 辆车的断面行驶速度为 \$u\_k^{(i)}\$，估计的车辆平均行驶速度用 \$\bar{u}\_i\$ 表示。则：

1) 算术平均法：

$$\bar{u}_s = \frac{1}{\sum_t N^{(i)}} \sum_t \sum_k u_k^{(i)} \quad (2-8)$$

2) 调和平均法

$$\bar{u}_s = \frac{\sum_t N^{(i)}}{n \cdot \sum_t \left( \frac{l_i}{L} \sum_k \frac{1}{u_k^{(i)}} \right)} \quad (2-9)$$

3) 自适应加权平均法<sup>[28]</sup>

首先计算检测器 \$D\_i\$ 处的时间平均速度 \$\bar{u}\_i^{(i)}\$ 和标准差 \$\sigma^{(i)}\$

$$\bar{u}_i^{(i)} = \frac{1}{N^{(i)}} \sum_k u_k^{(i)} \quad (2-10)$$

$$\sigma^{(i)} = \sqrt{\frac{1}{N^{(i)} - 1} \sum_k (u_k^{(i)} - \bar{u}_i^{(i)})^2} \quad (2-11)$$

引入加权因子 \$\omega^{(i)}\$，有

$$\bar{u}_s = \sum_t \omega^{(i)} \bar{u}_i^{(i)} \quad (2-12)$$

其中

$$\omega^{(i)} = \frac{1}{(\sigma^{(i)})^2 \sum_r \frac{1}{(\sigma^{(i)})^2}} \quad (2-13)$$

#### 4) 信号灯路口延误计算

交叉口平均延误在高峰和平峰时段分别测量，路段行程时间可由路段行驶时间加上相应时段的平均延误估算得到。

文献[29]提出以下公式计算路口平均延时：

$$d = \frac{T(1-T_g/T)^2}{2(1-(T_g/T)x)} + \frac{x^2}{2q(1-x)} - 0.65 \left( \frac{T}{q^2} \right)^{\frac{1}{3}} x^{2+s(T_g/T)} \quad (2-14)$$

其中  $T_r$  为红灯时间， $T_g$  为绿灯时间， $T$  为信号灯周期， $q$  为上游流量， $s$  为车辆疏散流量， $x$  为道路饱和度。

道路饱和度定义为

$$x = q/S/(T_g/T) \quad (2-15)$$

$$\bar{u}_s \leftarrow \frac{L}{\frac{L}{\bar{u}_s} + d} \quad (2-16)$$

由断面检测器估计的路段行程时间加上信号灯延时即为实际路段平均行程时间，可使用这一数值与浮动车估计的行程时间做比较。

### 2.2.2 由浮动车数据估计行程时间

由于断面检测数据是截面数据，而路段平均行程时间是反映路段区间的参数，所以对于动态时变交通系统，利用线圈数据估计平均行程时间很难满足实时性和准确性约束条件<sup>[33]</sup>。由于浮动车在道路上行驶，直接测量动态行程时间，因此被越来越广泛用作行程时间估计方法。但路段行程时间是指一段时间内经过该路段的所有车辆的平均行驶速度，若浮动车样本量过小，则其均值将距真实值有较大的偏差<sup>[34,35]</sup>，因此浮动车样本量问题一直是该领域研究的热点问题。如研究浮动车样本数量与行程时间估计误差的关系，这方面的研究包括估计给定精度要求下路段上所需的最小浮动车辆数<sup>[36-41]</sup>、估计给定浮动车占总车流比例情况下的行程时间误差<sup>[42-47]</sup>等。但即使在系统中投入大量浮动车，由于车辆在路网中分布的不均匀性，仍会存在部分路段的样本量过小，这样就需要选择合适的小样本估计方

法,从小样本数据中估计出高精度的行程时间。

统计学中的小样本估计方法主要包括 Bayes 方法、Bootstrap 和 Bayes-Bootstrap 方法、蒙特卡罗 (MC) 仿真方法、最小二乘法—卡尔曼平滑—(LKL)、最小绝对残差和法(MSAR)、小子样的相容性检验方法、分位点方法、修正极大似然法(MMLE)、序贯压缩法(SR)、综合的修正极大似然法和序贯压缩法 CMSR)等,分别对不同类型的系统有一定的适用性。

在实际行程时间估计方面的相关研究包括样本中位数估计行程时间<sup>[7]</sup>、指数平滑方法<sup>[33]</sup>,以及与断面检测数据进行数据融合估计行程时间的方法<sup>[32,48-52]</sup>,如 Kalman 滤波和 BP 神经网络等方法。

下面将对这些方法进行简要介绍。

统计学中对小样本的定义一般以样本容量  $n=50$  作为分界,但 50 并不是严格的规定,如果母体的分布接近正态分布,则所需的子样还可以小些(工程实践中,一般认为样本容量  $n<50$  为小样本)。

### 1) Bayes 方法<sup>[53]</sup>

Bayes 可靠性估计是一个半经验方法,该方法假设可靠性有一个“验前分布”即先验分布,根据这个验前分布和样本信息可导出“验后分布”,从而根据这个验后分布做出可靠的 Bayes 点估计与区间估计。

设  $\theta$  是总体分别  $p(x|\theta)$  中的参数,从总体随机抽取  $X = (x_1, x_2, \dots, x_n)$ ,根据  $\theta$  的先验信息取一先验分布  $\pi(\theta)$ ,用贝叶斯公司算得后验分布  $\pi(\theta|x)$ 。作为  $\theta$  的估计可选用后沿分布  $\pi(\theta|x)$  的某个位置特征量,如最大值,中位数,期望值等。

使后验分布  $\pi(\theta|x)$  达到最大的值  $\theta_{md}$  称为  $\theta$  的最大后验估计;后验分布  $\pi(\theta|x)$  的中位数  $\theta_{mc}$  称为  $\theta$  的后验中位数估计,后验分布  $\pi(\theta|x)$  的期望值  $\theta_e$  称为  $\theta$  的期望值估计,这三个估计都称为贝叶斯估计。

Bayes 方法的特点就在于利用验前分布,由于有了这个验前分布,不需要很大的子样也可以得到较好的概率估计值,这是 Bayes 估计的优点。然而验前分布确定的是否合理,将直接影响系统可靠性的评定结果,因此,在 Bayes 统计推断方法中,验前信息的运用是一个关键问题。不同形式的验前分布,将引起不同的统计分析后果,对于 Bayes 估计或检验,将具有不同的风险。

### 2) Bootstrap 和 Bayes-Bootstrap 方法<sup>[54]</sup>

Bootstrap 方法实质是一个再抽样过程,即用现有的资料去模仿未知的分布。运用这种方法可以对参数进行区间估计或统计假设检验。

设  $x_1, x_2, \dots, x_n$  为  $F(x)$  的样本,  $\theta = \theta(F)$  为总体分布的未知参数,  $F$  为抽样分布函数,  $\hat{\theta} = \hat{\theta}(F)$  为  $\theta$  的估计,记  $T_n = \hat{\theta}(F_n) - \theta(F)$ ,它表示了估计误差,

记  $x^* = (x^*_1, x^*_2, \dots, x^*_n)$  为从  $F_n$  中抽样获得的再生样本, 称为 Bootstrap 子样;  $F^*_n$  是由  $x^*$  所获得的抽样分布。记  $R_n = \hat{\theta}(F^*_n) - \hat{\theta}(F)$ , 称  $R_n$  为  $T_n$  的自助统计量, 利用  $R_n$  分布 (在给定  $F_n$  之下) 去模仿  $T_n$  的分布, 这就是 Bootstrap 方法的中心思想。

Bayes-Bootstrap 方法 (也称随机加权法), 它将自助统计量  $R_n$  中的  $\hat{\theta}(F^*_n)$  换为  $\hat{\theta}_v = \theta(\sum_{i=1}^n V_i f_i(x))$ , 其中  $x = (x_1, x_2, \dots, x_n)$ ,  $f_i(x)$  是某个 Borel 函数,  $V_i$  为具有 Dirichlet 分布  $D(1, 1, \dots, 1)$  的随机变量。记  $D_n = \hat{\theta}_v - \hat{\theta}(F_n)$  称为随机加权统计量。Bayes-Bootstrap 方法是一种估计误差的统计处理方法, 实质是用随机加权的统计变量  $D_n$  去模仿估计误差  $T_n$  的分布。

下面以指标  $P$  为例说明如何运用 Bayes-Bootstrap 方法。要对均匀性指标  $P$  进行评定, 首先必须找到指标  $P$  的统计量并确定其分布。设有指标  $P$  的  $L$  个样本  $p_1, p_2, \dots, p_n$ , 指标  $P$  的统计量:

$$\bar{P} = \frac{1}{L} \sum_{i=1}^n P_i \quad (2-17)$$

运用 Bayes-Bootstrap 方法确定  $\bar{P}$  的分布。

统计量  $\bar{P}$  有估计偏差

$$T_n = \bar{P} - P \quad (2-18)$$

由此构造 Bayes-Bootstrap 统计量

$$D_n = \bar{P}_v - \bar{P} \quad (2-19)$$

其中

$$\bar{P}_v = \sum_{i=1}^n V_i P_i \quad (V_1, V_2, \dots, V_n) \sim \text{Dirichlet} D_n(1, 1, \dots, 1) \quad (2-20)$$

由  $D_n$  的分布去估计  $T_n$  从而获得统计量  $\bar{P}$  的分布, 具体步骤如下。

(1) 产生  $N$  组 ( $N$  足够大) Dirichlet 分布  $D_n(1, 1, \dots, 1)$  的随机矢量序列  $V(1), V(2), \dots, V(N)$ , 其中  $V(i) = (V_{i1}, V_{i2}, \dots, V_{in})$ ,  $i=1, 2, \dots, N$ ;

(2) 对每个  $V(i)$  计算相应的  $\bar{P}_v(i) = \sum_{j=1}^n V_{ij} P_j$  从而计算  $D_n(i) = \bar{P}_v(i) - \bar{P}$   $i=1, 2, \dots, N$ ;

(3) 以  $D_n(i)$   $i=1, 2, \dots, N$ ; 作为  $T_n$  的估计, 作直方图;

(4) 对直方图, 用最小二乘曲线拟合方法, 获得  $T_n$  的分布密度函数  $f(t)$ ;

(5) 对分布密度函数进行一次积分, 可获得  $T_n$  的分布函数  $F(t)$

(6) 由  $T_n = \bar{P} - P$ , 对给定的指标  $P$ , 可以获得统计量  $\bar{P}$  的条件分布函数  $F(t/p)$  及条件分布密度函数  $f(t)$ 。

### 3) 蒙特卡罗 (MC) 仿真方法

蒙特卡罗 (MC) 仿真方法是一种数字方法, 包括以概率统计理论为其主要理论, 以随机抽样为其主要手段的两个核心问题。可概括如下, 为求解数学、物理、工程技术以及生产管理等方面的问题, 首先建立一个概率模型或随机过程, 使它的参数等于问题的解; 然后通过对模型或过程的观察或抽样试验来计算所求参数的统计特征, 最后给出所求解的近似值。

该方法的基本特点是: 方法及程序结构简单, 大量简单重复抽样; 概率意义的收敛, 收敛速度慢; 方法适用性强。可以对复杂的随机性问题进行求解, 不要对数学模型的简化和假设; 进行计算机的大量试验, 不会有不可克服的困难; 做出的解答能否反映实际, 与原来所建立的数学模型是否正确及输入信息的质量有关, 而与本方法无关。

简言之, 蒙特卡罗法是模拟随机变量  $h_1, h_2, \dots, h_n$  的函数  $\eta = (h_1, h_2, \dots, h_n)$  得到抽样值  $\eta_1, \eta_2, \dots, \eta_N$ , 经过统计处理后, 得到  $\eta$  的概率分布或各阶矩的统计估计, 最后得到问题的近似解, 因此, 用蒙特卡罗法求解的问题, 其解答往往是某事件出现的概率或某个随机变量的数字特征。

参数的概率分布函数的提取:

1. 当样本较多时, 可以采用频率求得经验分布函数。
2. 当数据个数较少, 但知道该参数的分布类型。
3. 当已知参数只有两个时, 则认为该参数是服从均匀分布的随机变量。
4. 当已知数据为 3 个时, 则采用三角分布。

用蒙特卡罗法模拟资源量需要用大量的随机数, 产生  $\eta = (h_1(t), h_2(t), \dots, h_n(t))$  样本, 用来统计  $\eta$  的近似分布。

#### 4) 中位数法

若第  $k$  时段进入路段的车辆数为  $N_k$ , 当  $n_k < 15\%N_k$  时, 其中  $n_k$  为在第  $k$  时段内有有效的 GPS 数据样本数, 按小样本计算方法估计路段平均参数。由于实际 GPS 数据样本数量小, 样本分布出现非对称的概率较大, 直接用 GPS 数据样本估计总体路段平均行程时间均值误差也较大。此时采用非参数方法, 用 GPS 数据的顺序统计量及中位数估计总体路段平均行程时间。顺序统计量是充分统计量, 即 GPS 样本包含车流总体分布中未知参数信息, 总体的路段平均行程时间可以从 GPS 数据的行程时间样本顺序统计量推断。

设在第  $k$  时段有  $n_k$  个有效行程数据样本  $t_k(1), t_k(2) \dots t_k(n_k)$ , 将行程时间样本从小到大排序后, 得到行程时间样本顺序统计量  $t_{(1)}, t_{(2)} \dots t_{(n_k)}$ , 从行程时间样本顺序统计量可计算行程时间样本中位数  $M$  为

$$M = \begin{cases} (t_{n_k/2} + t_{(1+n_k)/2})/2, n \text{ 为偶数} \\ t_{(1+n_k)/2}, n \text{ 为奇数} \end{cases} \quad (2-21)$$

假设行程时间分布为正态分布, 则用中位数  $M$  表示的路段行程时间的置信区间为  $1-0.5^{n_k-1}$ 。

### 5) 指数平滑法

设第  $k$  时段交通流平均速度为  $v_k$ , 有  $n_k$  个有效速度样本值  $v_k^1, v_k^2 \dots v_k^{n_k}$ , 则  $v_k^m (m \in 1, 2 \dots n_k)$  可写成如下形式:

$$v_k^m = v_k + Z_k^m \quad (2-22)$$

式中  $Z_k^m$  为零均值白噪声, 代表第  $m$  辆车在  $k$  时段的瞬时速度对该时刻其所在交通流平均速度的偏差。

假设同一路段相邻时间间隔的速度估计方差受限, 即有

$$v_k = v_{k-1} + \eta_k \quad (2-23)$$

式中  $\eta_k$  为独立同分布的随机变量。若已知  $\theta_k$  是第  $k$  个抽样间隔内路段的平均行程速度估计值, 且前一间隔的估计值为  $\theta_{k-1}$  则路段平均速度为

$$\theta_k = (1 - f_k)\theta_{k-1} + f_k \frac{1}{n_k} \sum_{m=1}^{n_k} v_k^m \quad (2-24)$$

其中  $0 < f_k < 1$  为时间段  $k$  的权值。当  $f_k = \frac{n_k}{1+n_k} \cdot |C_k|$  时可使估计方差最小。 $C_k$  为误差跟踪信号, 其定义为

$$C_k = \frac{E_k}{A_k} \quad (2-25)$$

且有

$$E_k = r \cdot e_k + (1-r) \cdot E_{k-1}, A_k = r \cdot |e_k| + (1-r) \cdot A_{k-1} \quad (2-26)$$

$$\text{其中 } r = \frac{n_k}{n_k + 1}, e_k = \left( \frac{1}{n_k} \sum_{m=1}^{n_k} v_k^m \right) - \theta_{k-1}$$

根据文献[33], 采用 VISSIM 对 3 个路段进行仿真分析, 得到指数平滑行程速度比平均行程速度的误差分别减小了 2.07%、7.26%和 13.56%, 误差方差分别减小 38.30%、59.3%和 38.64%; 采用指数平滑方法后, 各路段速度估计精度分别由原来的 94%、89%、86.4% 上升至 95%、90.8%和 86.8%。

### 6) 结合断面检测数据进行平均行程时间评估

文献[32]提出基于 Kalman 滤波的断面检测数据和浮动车数据融合的算法。该

算法思想是首先得到断面检测数据行程时间估计和浮动车行程时间估计，然后将二者结合起来，应用 Kalman 滤波算法估计路段密度，从而计算行程速度。

浮动车行程时间估计如下：

$$tt_p = \frac{\sum_{n=1}^N (t_d^n - t_u^n)}{N} \quad (2-27)$$

然后估计断面行程时间，与浮动车行程时间结合，使用 Kalman 滤波估计路段密度和行程时间：

$$\text{状态方程为 } k(t) = k(t-1) + u(t) + w(t) \quad (2-28)$$

$$\text{测量方程为 } tt(t) = H(t) \cdot k(t) + v(t) \quad (2-29)$$

$w(t)$  和  $v(t)$  分别为系统误差和测量误差，均假设为高斯分布白噪声，其中  $u(t)$  和  $H(t)$  由下式计算：

$$u(t) = \frac{1}{L \cdot \Delta x} \cdot (q_u(t) + q_{on}(t) - (q_d(t) + q_{off}(t))) \quad (2-30)$$

$$H(t) = \frac{\Delta x}{\alpha[q_u(t) + q_{on}(t)] + (1-\alpha) \cdot [q_d(t) + q_{off}(t)]} \quad (2-31)$$

使用该算法对一段长度为 6 英里的高速路段进行仿真分析，给定 5% 的浮动车占有率，日常拥堵情况下使用断面检测、浮动车检测与 Kalman 滤波算法的行程时间误差分别为 10.6%，10.8% 和 7.6%，有突发事件情况下三者分别为 16%，14.3% 和 9.8%。

### 3 小样本路段行程时间估计方法

#### 3.1 行程时间估计问题的提出

目前主要的行程时间估计方法包括断面检测方法和浮动车检测方法，由于断面检测器不能直接测量行程速度，需要使用流量数据和占有率数据进行估算，容易产生误差。浮动车方法能够直接测量行程速度，但可能由于样本量过少导致估计偏差增大，因此需要采用合适的小样本估计方法。但现有的浮动车的行程时间小样本估计方法多针对没有信号控制的高速路段，这种情况下样本分布集中，且为单峰分布，使用现有方法可有效降低估计误差。但实际情况中城市路网大部分由带信号灯的交叉路口组成，如要对这些路段进行交通状态分析，必须考虑信号灯延时，得到可反映真实情况的行程时间。如果样本量足够，则其均值能够反映路段交通状态，若样本量不足，仅按照平均值求解将造成比无信号灯路段更大的误差。因此，本文提出一种考虑信号灯延时的小样本行程时间估计方法，能够较好的解决这一问题。

本文研究对象为城市主干路，其中包括若干带信号灯的路段，如图 3-1 所示。

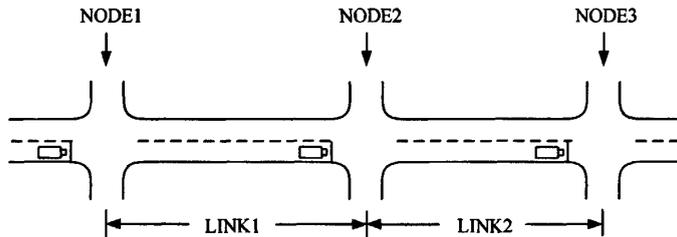


图 3-1 有交叉口检测路段示意图

已知浮动车在每个路段的包含信号灯延时的行程时间和信号灯相位情况，则可通过以下算法降低估计误差：首先是通过数据预处理移除边界值和错误点，然后根据相位情况对样本值进行修正，用修正后的样本计算样本均值，最后进行误差分析，证明该算法能够有效降低估计误差。

#### 3.2 数据预处理方法

由于受到交通信号控制的影响，浮动车测量的路段行程时间较为分散，同时样本数量的限制使得小概率事件能够对平均行程时间的计算造成很大影响，如偶然发生的在拥挤时段车辆高速通过交叉口的现象，或者在非拥挤时段车辆偶然遇

到的较大延误。因此为了获得更合理的结果，应首先对原始数据进行预处理。

设车辆在路段  $i$  的行程时间分布函数为  $F_i(x)$ ，则当样本  $x_0$  满足  $\frac{\alpha}{2} \leq F_i(x_0) \leq 1 - \frac{\alpha}{2}$  时接受该值，即对样本值过小和过大的点进行过滤。本项目中  $\alpha$  取 0.1，即样本数据的置信系数取 90%，过滤时间最短的 5% 的样本点和时间最长的 5% 样本点。

### 3.3 行程时间估计算法

在有信号灯控制的路段，车辆行程时间在遇到红灯和未遇到红灯的条件下具有不同的分布，即

$$f(x) = f(x|R)P(R) + f(x|\bar{R})P(\bar{R}) \quad (3-1)$$

假设在这两种情况下行程时间均服从正态分布，即

$$f(x|R) \sim N(\mu + c_r, \sigma_r^2) \quad (3-2)$$

$$f(x|\bar{R}) \sim N(\mu - c_g, \sigma_g^2) \quad (3-3)$$

其中  $R$  为事件“遇到红灯并停车”，其概率  $P(R) = \omega_r$ ，同理，事件“未遇到红灯，不停车”的概率为  $P(\bar{R}) = \omega_g = 1 - \omega_r$ ， $\mu$  为平均行程时间， $c_r$  和  $c_g$  为两种情况下对平均行程时间的调整参数。

由上述单路段的行程时间分布函数，可知路段的行程时间期望和方差分别为

$$\mu = \omega_r \mu_r + \omega_g \mu_g \quad (3-4)$$

$$\sigma^2 = \omega_r \sigma_r^2 + \omega_g \sigma_g^2 + \omega_r \omega_g (\mu_r - \mu_g)^2 \quad (3-5)$$

将条件  $R$  下所有样本点数值减小  $c_r$ ，条件  $\bar{R}$  下所有样本点数值增加  $c_g$ ，记处理后的样本值为  $x_1$ ，则经过修正后的样本分布为

$$f(x_1) = f(x_1|R)P(R) + f(x_1|\bar{R})P(\bar{R}) \quad (3-6)$$

其中

$$f(x_1|R) \sim N(\mu, \sigma_r^2) \quad (3-7)$$

$$f(x_1|\bar{R}) \sim N(\mu, \sigma_g^2) \quad (3-8)$$

修正后的样本点的期望和方差分别为

$$\hat{\mu} = \omega_r \mu_r + \omega_g \mu_g \quad (3-9)$$

$$\hat{\sigma}^2 = \omega_r \sigma_r^2 + \omega_g \sigma_g^2 \quad (3-10)$$

若假设  $\sigma_r = \sigma_g = \sigma_0$ ，则有  $\sigma_2^2 = \sigma_0^2$ 。

则可知修正过的均值  $\bar{x}_1$  为行程时间的无偏估计，且比原始均值估计更为有效。

算法步骤如下：

- step 1 按时间顺序寻找相邻  $n$  个样本点（如 5 个点），这些样本点内行程时间最长的车辆包含了红灯延误，行程时间最短的车辆没有经历红灯。实际上，满足这一条件的概率为  $P_m = 1 - \omega_r^n - \omega_g^n$ ，样本点个数越多，可能性越大。例如，当  $\omega_r = 0.5$  时，5 个样本点中既存在遇到红灯停车的点，又存在未遇到红灯的点的概率为 93.5%。
- step 2 由样本点中的最长行程时间  $x_{\max}$ 、最短行程时间  $x_{\min}$  和相位信息可估计出该时间段内的行程时间阈值  $\theta = x_{\max} \cdot \omega_r + x_{\min} \cdot \omega_g$ ，其中  $\omega_r = T_r / T$ ， $\omega_g = 1 - \omega_r$ ，认为行程时间大于  $\theta$  的样本点包含了红灯延误，行程时间小于  $\theta$  的样本点未经历红灯。
- step 3 计算所有包含延误的样本点的平均行程时间  $\bar{x}_r$  和未包含延误的样本点平均行程时间  $\bar{x}_g$ ，由此估计参数  $c_r$  和  $c_g$ ，即  $c_r = \omega_g (\bar{x}_r - \bar{x}_g)$ ， $c_g = \omega_r (\bar{x}_r - \bar{x}_g)$ 。
- step 4 对原始样本点进行处理，即对于所有遇到红灯的样本点，将其值减小  $c_r$ ，对于所有未遇到红灯的样本点，将其值增加  $c_g$ 。
- step 5 最后对处理后的数据进行 5 分钟平均，得到平均行程时间变化曲线。

当估计扩展路段（由多个小路段组成）的行程时间时，设某车辆依次行驶过路段  $i=1, 2, \dots, n$ ，则车辆从路段 1 出发到完成路段  $n$  的总行程时间  $X$  为

$$X_n = \sum_{i=1}^n x(i) \quad (3-11)$$

设每个路段上车辆行驶时间互相独立，则当  $n$  足够大时，由中心极限定理可得

$X_n$  服从均值为  $\sum_{i=1}^n \mu(i)$ ，方差为  $\sum_{i=1}^n \sigma^2(i)$  的正态分布。

对扩展路段中的每个样本点都按照其所在路段  $i$  的相位情况进行修正，得到修正值  $x_1(i)$ ，那么修正后的扩展路段行程时间为

$$\hat{X}_n = \sum_{i=1}^n x_1(i) \quad (3-12)$$

其均值为  $\sum_{i=1}^n \hat{\mu}(i)$ ，方差为  $\sum_{i=1}^n \hat{\sigma}^2(i)$  的正态分布。

进行扩展路段行程速度估计的算法步骤如下：

- step 1 对每个路段中的样本点进行修正。
- step 2 按照车辆行驶路径找到每个路段中该车辆所对应的行驶时间，相加得到多路段行程时间。
- step 3 最后对处理后的数据进行 5 分钟平均，得到多路段平均行程时间变化曲线。

### 3.4 算法误差分析

由于期望  $\mu = E(x) = \omega_r \cdot E(x|R) + \omega_g \cdot E(x|\bar{R})$ ，可得到  $c_r$  和  $c_g$  的关系：

$$\omega_r \cdot c_r = \omega_g \cdot c_g \quad (3-13)$$

平均行程时间  $\mu$  的估计值  $\hat{\mu}$  可由  $\mu_r = \mu + c_r$  和  $\mu_g = \mu + c_g$  的估计值  $\hat{\mu}_r$  和  $\hat{\mu}_g$  得到，即

$$\hat{\mu} = \omega_r \cdot \hat{\mu}_r + \omega_g \cdot \hat{\mu}_g \quad (3-14)$$

由均值的区间估计可得到  $\mu_r$  和  $\mu_g$  的置信系数为  $1-\alpha$  的区间  $I_r$  和  $I_g$ ：

$$I_r = [\bar{x}_r - \frac{S^*}{\sqrt{n_r}} \cdot t_{1-\frac{\alpha}{2}}(n_r - 1), \bar{x}_r + \frac{S^*}{\sqrt{n_r}} \cdot t_{1-\frac{\alpha}{2}}(n_r - 1)] \quad (3-15)$$

$$I_g = [\bar{x}_g - \frac{S^*}{\sqrt{n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_g - 1), \bar{x}_g + \frac{S^*}{\sqrt{n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_g - 1)] \quad (3-16)$$

其中  $\bar{x}_r$  为条件  $R$  下样本点均值， $\bar{x}_g$  为条件  $\bar{R}$  下样本点均值。则  $\mu$  的置信系数为  $1-\alpha$  的区间为，

$$I = [\delta - \lambda, \delta + \lambda] \quad (3-17)$$

其中

$$\delta = \omega_r \cdot \bar{x}_r + \omega_g \cdot \bar{x}_g \quad (3-18)$$

$$\lambda = \omega_r \cdot \frac{S_r^*}{\sqrt{n_r}} \cdot t_{1-\frac{\alpha}{2}}(n_r - 1) + \omega_g \cdot \frac{S_g^*}{\sqrt{n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_g - 1) \quad (3-19)$$

这时可认为  $\hat{\mu}$  在  $1-\alpha$  的估计误差  $e = \lambda / \delta$ 。

则其期望  $\mu_1 = \mu$ ，方差  $\sigma_1^2 = \omega_r \cdot \sigma_r^2 + \omega_g \cdot \sigma_g^2$

在  $\sigma_r^2$  接近  $\sigma_g^2$  时可近似认为  $x_1 \sim N(\mu, \sigma_0^2)$ ，其中  $\sigma_0^2 \approx \sigma_r^2 \approx \sigma_g^2$ 。则  $\mu$  在  $1-\alpha$  下的置信区间为  $[\bar{x}_1 - \frac{S^*}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1), \bar{x}_1 + \frac{S^*}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1)]$ ，估计误差为

$$e_1 = \frac{S^*}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\bar{x}_1} \quad (3-20)$$

若认为样本标准差和均值都是路段行程时间分布的无偏估计，可有  $S^* \approx S_r^* \approx S_g^* \rightarrow \sigma_0$ ， $\bar{x}_1 \approx \omega_r \cdot \bar{x}_r + \omega_g \cdot \bar{x}_g \rightarrow \mu$ ，且  $\omega_r / n_r \approx \omega_g / n_g \approx (\omega_r + \omega_g) / (n_r + n_g) \rightarrow r$ ，则原始数据估计误差：

$$\begin{aligned} e &= \left[ \omega_r \cdot \frac{S_r^*}{\sqrt{n_r}} \cdot t_{1-\frac{\alpha}{2}}(n_r - 1) + \omega_g \cdot \frac{S_g^*}{\sqrt{n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_g - 1) \right] \cdot \frac{1}{\omega_r \cdot \bar{x}_r + \omega_g \cdot \bar{x}_g} \\ &\approx \left[ \omega_r \cdot \frac{\sigma_0}{\sqrt{n_r}} \cdot t_{1-\frac{\alpha}{2}}(n_r - 1) + \omega_g \cdot \frac{\sigma_0}{\sqrt{n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_g - 1) \right] \cdot \frac{1}{\mu} \\ &> \left[ \omega_r \cdot \frac{\sigma_0}{\sqrt{n_r}} + \omega_g \cdot \frac{\sigma_0}{\sqrt{n_g}} \right] \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\mu} \\ &= \left[ r \cdot \sqrt{n_r} \cdot \sigma_0 + r \cdot \sqrt{n_g} \cdot \sigma_0 \right] \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\mu} \\ &> \left[ r \cdot \sqrt{n_r + n_g} \right] \cdot \sigma_0 \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\mu} \\ &= \frac{\sigma_0}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\mu} \\ &\approx \frac{S^*}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\bar{x}_1} = e_1 \end{aligned} \quad (3-21)$$

即由原始数据均值  $\bar{x}$  估计的行程时间误差  $e$  大于由处理后的  $\bar{x}_1$  估计的误差  $e_1$ 。

对于扩展路段行程时间  $X_n$  在置信系数  $1-\alpha$  下的区间估计为

$$I_x = [\bar{X}_n - \frac{S_x^*}{\sqrt{n_r + n_g}} t_{1-\frac{\alpha}{2}}(n_r + n_g - 1), \bar{X}_n + \frac{S_x^*}{\sqrt{n_r + n_g}} t_{1-\frac{\alpha}{2}}(n_r + n_g - 1)] \quad (3-22)$$

则这一估计的误差为

$$e_2 = \frac{S_x^*}{\sqrt{n_r + n_g}} t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\bar{X}_n} \quad (3-23)$$

以下证明多路段的行程时间估计误差小于组成多路段的各路段误差的均值，即

$$e_2 < \bar{e}_1 < \frac{1}{n} \cdot \sum_{i=1}^n \frac{S^*(i)}{\sqrt{n_r(i) + n_g(i)}} \cdot t_{1-\frac{\alpha}{2}}(n_r(i) + n_g(i) - 1) \cdot \frac{1}{\bar{x}(i)} \quad (3-24)$$

若认为样本标准差和均值都是多路段行程时间分布的无偏估计，则有  $\bar{x}(i) = \mu(i)$ ， $\bar{X}_n = \sum_{i=1}^n \bar{x}(i) = \sum_{i=1}^n \mu(i)$ ， $S_x^* = \sum_{i=1}^n \sigma_0^2(i)$ ， $S^*(i) = \sigma_0(i)$ ，且对于任意  $i$  均都有

$$\omega_r(i)/n_r(i) \approx \omega_g(i)/n_g(i) \approx (\omega_r(i) + \omega_g(i))/(n_r(i) + n_g(i)) = r \quad (3-25)$$

又由在路段中车辆守恒，因此对任意  $i$  均有

$$n_r(i) + n_g(i) = n_r + n_g \quad (3-26)$$

若假设  $S_r^*(i)/\bar{x}(i) \approx S_g^*(i)/\bar{x}(i) \approx \sigma_0(i)/\bar{x}(i) = \beta$ ，即每个路段的方差均值比接近，则有：

$$\bar{e}_1 = \frac{1}{n} \cdot \sum_{i=1}^n \frac{\sigma_0(i)}{\sqrt{n_r + n_g}} \cdot \frac{1}{\bar{x}(i)} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) = \frac{\beta}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \quad (3-27)$$

而

$$\begin{aligned}
e_2 &= \frac{S_x}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\bar{X}_n} \\
&= \frac{\sqrt{\sum_{i=1}^n \sigma_0^2(i)}}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\sum_{i=1}^n \bar{x}(i)} \\
&< \frac{\sum_{i=1}^n \sigma_0(i)}{\sum_{i=1}^n \bar{x}(i)} \cdot \frac{1}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \quad (3-28) \\
&= \frac{\beta}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \\
&= \bar{e}_1
\end{aligned}$$

即多路段的行程时间估计误差比从单路段估计要低。若进一步假设在特殊情况下，如果路段划分均匀，则每个路段行程时间的均值和方差都接近，即对于任意路段*i, j*均有 $\bar{x}(i) \approx \bar{x}(j) = \mu$ ,  $\sigma_0(i) \approx \sigma_0(j) = \sigma_0$ ，则

$$\begin{aligned}
e_2 &= \frac{\sqrt{\sum_{i=1}^n \sigma_0^2(i)}}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \cdot \frac{1}{\sum_{i=1}^n \bar{x}(i)} \\
&= \frac{\sigma_0}{\sqrt{n} \cdot \mu} \cdot \frac{1}{\sqrt{n_r + n_g}} \cdot t_{1-\frac{\alpha}{2}}(n_r + n_g - 1) \quad (3-29) \\
&= \frac{1}{\sqrt{n}} \cdot \bar{e}_1
\end{aligned}$$

在路段划分理想的情况下（路段划分大致均匀），多路段的行程时间的估计误差仅为单路段的 $1/\sqrt{n}$ ，这说明考虑的路段数量越多，行程时间的估计误差越小，样本均值越接近实际平均行程时间。由于在实际情况中路段不太可能分割绝对均匀，实际上多路段行程时间的估计误差将下降至单路段误差均值的 $1/\sqrt{n} \sim 1$ 倍之间。

## 4 实证分析

### 4.1 测试系统概述

为了对上文提出的浮动车数据测试方法进行实证分析，验证该方法的可靠性，使用浮动车数据与多种断面交通流检测数据共同对相同路段进行测量，通过对二者数据比较对算法的可行性进行验证。

实际数据采集选择上海市内的快速路、地面主干道和一条国道进行。快速路选取内环高架从鲁班路立交至漕溪北路匝道的双向道路（单向长度约 4.7km）；地面主干道选取中山南二路从鲁班路至漕溪北路的双向道路（单向约 4.7km），如图 4-1 所示；国道选取 318 国道（沪青平公路）从外环至赵巷镇的双向道路（单向约 18km），如图 4-2 所示。

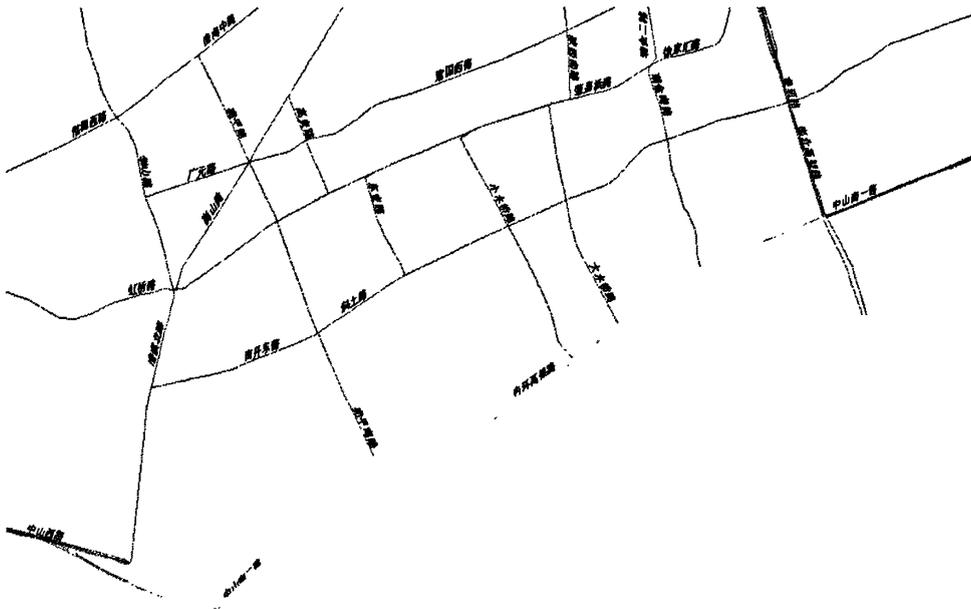


图 4-1 鲁班路立交至漕溪北路匝道高架和地面主干道地图

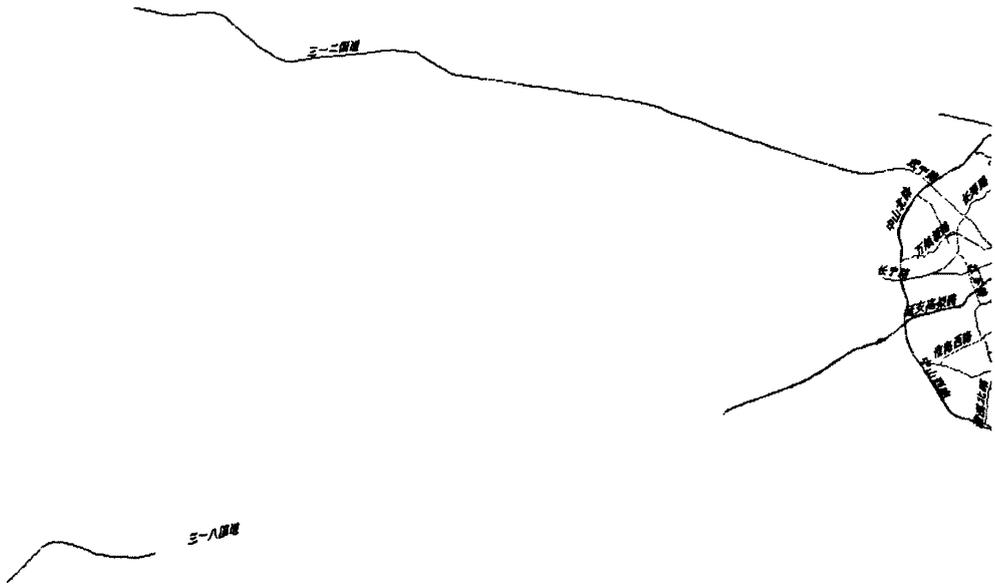


图 4-2 318 国道（沪青平公路）从外环至赵镇道路地图

三条检测道路按照交叉口或匝道分成若干检测路段(LINK)，检测设备将给出每个 LINK 的平均行程时间和平均速度。其中，

- 鲁班路立交至漕溪北路匝道的地面主干道以交叉口的中点为节点(NODE)，每相邻两个 NODE 确定双向两个 LINK，双向道路编号及每段道路详细位置见下表：

路段编号	长度(m)	起点路名	终点路名
1	393.2	龙华西路	天钥桥路
2	441.6	天钥桥路	双峰路
3	326.6	双峰路	宛平南路
4	573.2	宛平南路	东安路
5	251.3	东安路	船厂路
6	452.1	船厂路	小木桥路
7	306.2	小木桥路	大木桥路
8	264.9	大木桥路	兆丰路
9	293.3	兆丰路	瑞金南路
10	388.7	瑞金南路	打浦路
11	388.7	打浦路	瑞金南路
12	293.3	瑞金南路	兆丰路
13	264.9	兆丰路	大木桥路
14	306.2	大木桥路	小木桥路
15	452.1	小木桥路	船厂路

16	251.3	船厂路	东安路
17	573.2	东安路	宛平南路
18	326.6	宛平南路	双峰路
19	441.6	双峰路	天钥桥路
20	393.2	天钥桥路	龙华西路

- 三一八国道以交叉口的中点为节点(NODE)，每相邻两个 NODE 确定双向两个 LINK，双向道路编号及每段道路详细位置见下表：

路段编号	长度 (m)	起点路名	终点路名
21	538.8	环西一大道	航东路
22	284.3	航东路	空港路
23	685.3	空港路	航新路
24	551.2	航新路	吴宝路
25	240.5	吴宝路	七莘路
26	1032.7	七莘路	华翔路
27	1032.7	华翔路	七莘路
28	240.5	七莘路	吴宝路
29	551.2	吴宝路	航新路
30	685.3	航新路	空港路
31	284.3	空港路	航东路
32	2048.1	航东路	环西一大道

- 鲁班路立交至漕溪北路匝道的高架道路主要以上下匝道口为 NODE 来划分高架 LINK，当匝道口间距离较远时，会在之间人为增加一个 NODE，双向道路编号及每段道路详细位置见下表：

路段编号	长度 (m)	起点路名	终点路名
33	476.6	宛平南路匝道	船厂路
34	860.9	船厂路	瑞金南路匝道
35	456	瑞金南路匝道	瑞金南路
36	353.1	瑞金南路	打浦路匝道
37	403.7	打浦路匝道	南北高架鲁班路
38	477.7	瑞金南路匝道	小木桥路
39	599.98	小木桥路	宛平南路匝道

所得到的数据主要包括 2007 年 6 月 4 日至 6 月 8 日间项目组在上海市三条道路上采集到的车辆断面行驶速度、路段行程时间和交叉口延时数据。这一组数据通过浮动车调查和设备调查两种方式获得，其中断面行驶速度主要由地面检测器 MetroCount5600（以下简称 MC）和 NC97（以下简称 NC）获得，少量由雷达测

速仪得到, 路段行程时间通过浮动车调查和 DV 拍摄获得, 交叉口延时由 TDC12 统计得到。另外还有辅助数据如红绿灯相位调查。主要调查数据及获得方式汇总在下表中:

表格 4-1 主要数据来源

数据种类	获得方式	地点	有效检测时间	
断面速度	MC	主干道 龙华西路至天钥桥路	6月4日	
		318国道 航东路至环西一大道	6月6日	
	NC97	主干道 龙华西路至天钥桥路	6月4日	
		雷达测速仪	主干道 龙华西路至天钥桥路	6月4日 7:20~8:45
	行程时间	浮动车	主干道	6月4日 7:00~12:00
			318国道	6月6日 7:40~12:00
高架路	6月8日 14:00~17:00			
交叉口延误	TDC12	主干道 天钥桥路口	6月4日 7:30~12:00	
		318国道 环西一大道路口	6月6日 7:30~12:00	

## 4.2 浮动车数据采集

使用两周时间进行数据采集工作。浮动车用于采集每天固定时间段内所有检测路段的平均速度数据。具体方案为在工作日的 7:00~9:00 时段进行浮动车测试, 这是因为在这两小时内包括了高峰时间段 (7:00~8:00) 和非高峰时间段 (8:00~9:00)。由于周末的流量高峰比平时来得晚一些, 因此周六、日在 9:00~11:00 时间段采集数据。

采用图 4-1 和图 4-2 所示路段进行检测, 检测周期为 5 分钟, 平均速度 40km/h, 道路覆盖率达到 65% 的情况下计算各路段所需最小浮动车数量及误差情况。

- 单向 4.7km 的往返路段总长度为 9.4km, 代入式(1)得到  $n = 4$ 。置信水平为

$$95\% \text{时, } Z_{\alpha/2} = 1.645, \frac{\sigma}{d} = 1.22$$

► 单向 18km 的往返路段总长度为 36km, 代入式(1)得到  $n = 8$ 。置信水平为

$$95\% \text{时, } Z_{\alpha/2} = 1.645, \frac{\sigma}{d} = 1.72$$

由以上可知, 鲁班路立交至漕溪北路匝道的高架路及地面主干道分别应有 4 辆检测车往返行驶, 318 国道应有 8 辆检测车往返行驶, 即三条路段同时检测时则应至少有 16 辆检测车。当交通处于高峰时段时, 由于车辆平均速度下降, 若要维持同样的覆盖率则需要增加检测车数量。

每辆检测车内配备数据调查人员 2 名, 检测设备包括数据记录表格、电子秒表等。在每个数据采集阶段内, 检测车在检测道路上来回采集数据。执行方法是测试车从同一起点间隔 30~60 秒钟发车, 并相互在很小的时间差里通过每个路段。在行使过程中, 每辆车以浮动跟车法 (Floating Car Method) 随车流前进, 并在最大程度上避免以下驾驶行为: 故意加速超车或抢时通过交叉口信号灯, 以及故意怠慢而以最慢车速行驶。每辆车独立的在起点和终点之间连续测点采集数据。在每辆车里, 两位工作人员的分工是: 一人观察行车位置, 并确定该车在每个 NODE 的准确时间, 另一人将此时间作仔细记录。

## 4.3 基准数据采集

为了验证算法的有效性, 项目中采用断面检测方法获得基准数据, 通过安装在同一路段的多个检测器共同采集交通流量和速度数据, 能够获得准确度较高的行程时间。

### 4.3.1 断面检测器选择

针对城市道路的特点, 我们将根据具体断面情况, 采用多种设备联合调查的方式, 以充分发挥各种检测技术的优势, 获得最为准确的交通数据。目前选用的交通调查设备包括两大类: 固定型和移动型。

固定型设备包括 MetroCount5600 气压管式交通调查仪、NC97 地磁式交通调查仪、Vantage 视频检测器、TDC12 手持式多功能交通调查仪, 以及 STALKER 手持式雷达测速仪。其中 MetroCount5600 和 NC97 安装在路面上, 能够测量交通流量、速度和车型多种数据, Vantage 视频检测器可通过拍摄的交通流视频对速度、流量进行分析, TDC12 和雷达测速仪用于人工采取小样本数据, 以校对仪器测量

结果。

移动型设备包括车载 GPS 定位仪和手持 GPS 定位仪，用于浮动车检测方式，通过获取小样本的行程时间数据对平均行程时间进行估计。

各设备性能参数及现场安装要求如下

#### 1) MetroCount 5600

- 性能:

MetroCount 5600 为全球公认的检测精度最好的便携式交通数据采集设备，如果安装合理，车流量精度可达到 100%，车速、车型精度可达到 99% 以上。它采用气压管检测技术，利用车轴距进行分车型。其缺陷是安装和拆卸时需要阻断交通（每车道不超过 2 分钟，可以在凌晨或深夜进行设备安装），需要在路面钉钉子（射钉）。

- 安装要求:

- 检测断面现场能够允许安装人员夜间 1—4 点内进入，以方便安装和拆卸设备。所检测断面要有安装人员避车的区域，如紧急停车带。
- 每台设备要在车道的两个边缘各钉 4 或 6 个水泥钉（32mm 或 52mm 依据路面压实情况）。
- 设备每时每刻都需要有人看守，防止被盗或异常情况的发生。
- 现场安装人员 4 人以上，根据断面的多少而定。

#### 2) NC97

- 性能:

按以往调查经验，NC97 的数据精度仅次于 MC5600，如果安装合理，车流量精度可达到 98%，车速精度可达到 95% 以上。尤其在车速较慢时，其数据仍可达到很高的精度。它采用地磁检测技术，利用车长分型，分型精度大于 90%。其缺陷是安装和拆卸时需要阻断交通（每车道不超过 1 分钟，可以在凌晨或深夜进行设备安装），安装的方向精度会影响其检测精度。

- 安装要求:

- 检测断面现场能够允许安装人员夜间 1—4 点内进入，以方便安装和拆卸设备。
- 所检测断面要有安装人员避车的区域，如紧急停车带。
- 由于安装采用沥青胶带，粘在路面上，所以断面需要整洁、干燥。如果有水需使用水泥钉固定。
- 设备每时每刻都需要有人看守，防止被盗或异常情况的发生。

- 现场安装人员 3 人以上，根据断面的多少而定。

### 3) Vantage 视频检测器

- 性能：  
视频检测可提供全面的交通数据信息和事故检测信息，如道路上车辆的车流量、平均速度、平均车头时距、车道占用率，车型分类等，并可为调查人员提供可视图像，使数据更直观。
- 安装要求：
  - 采用便携摄像机录像，回到办公室再进行数据处理。
  - 摄像机架设在 6—10 米的高度，可以在过街天桥上，临街楼上或使用升降装置（放在路边即可）。

### 4) TDC12 多功能交通调查仪

- 性能：  
TDC12 是手持式人工调查交通流量的电子设备，在交通调查中往往以短时间的人工调查方式为精度基准，来验证自动观测设备的数据精度。在过街桥或路边即可进行数据记录，使用较为方便，但是，其不能采集车辆平均速度数据，只能验证流量精度。
- 安装要求：
  - 设备无需现场安装，便携方式采集，采集人员站在路边或过街天桥上即可。
  - 调查人员需 2 人以上。

### 5) Basic 雷达测速仪

- 性能：  
雷达测速仪是手持式用于采集车辆速度的设备，用来验证检测速度精度。

#### 安装要求：

- 设备无需现场安装，便携方式采集，采集人员站在路边或过街天桥上即可。
- 调查人员需 2 人以上。

## 4.3.2 断面检测器部署

便携式检测设备采集固定路段的全天交通流数据。便携式检测设备第一周和第二周测量不同的地点，测量路段长度约 1 公里左右（3 车道）或 1.5 公里左右（2

车道), 保证获得一周所有时间内的路段平均速度。

中山南二路的断面检测路段为龙华西路至天钥桥路, 检测器安装布局情况如图 4-3 所示。测量 D1、D2、D3 三处的交通数据时, 最外车道采用 MetroCount5600 进行测量, 靠内二车道采用 NC97 测量。

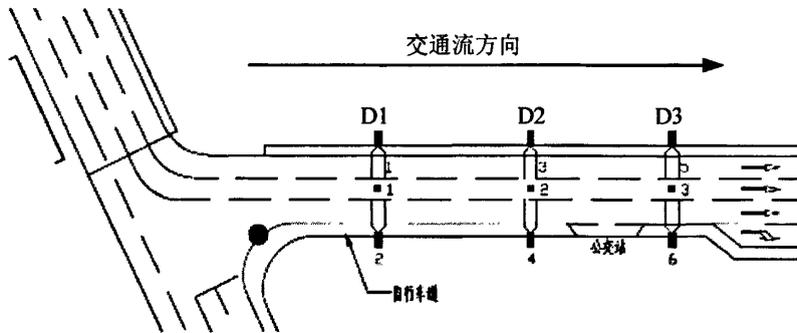


图 4-3 中山南二路检测器设备安装示意图

三一八国道断面检测路段为环西一大道至航东路, 检测器安装布局情况如图 4-4 所示, 此处 D1~D6 均采用 MetroCount 进行检测。

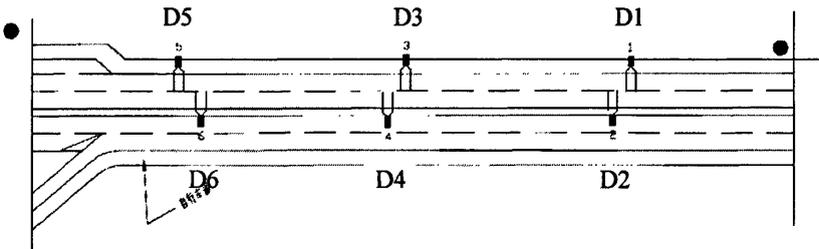


图 4-4 三一八国道检测器设备安装示意图

MetroCount 可直接测得每车道的单车数据, 包括车辆的断面速度、经过时间、车辆累计数量等, NC97 可测量采样时间间隔内的流量和速度平均值。同时安排人员在路旁进行人工流量 (使用 TDC12) 和速度 (使用 Basic 雷达测速仪) 记录工作, 由于人工方法采集的小样本数据用于原始数据的校验, 不需要长时间检测, 因此采取流动检测方法, 在某个检测点检测半小时后可换到下一检测点。

#### 4.3.3 断面检测器数据精度检验

通常情况下, 由 MC 检测得到的交通流数据具有较高的精度, 但在非常拥堵时测量误差会增大。为了确认检测数据的有效性, 我们将路边雷达测速仪测量的速度与 MC 的检测速度进行比较, 估计设备检测数据的精度。

用  $u_r(t)$  和  $u_m(t)$  分别表示  $t$  时刻雷达测速仪和 MC 检测到的单车车速。由于在检测时间段内，雷达测速仪样本数量与 MC 样本数量不一致，采用下述方法选取相同数量的样本点。

1. 寻找与  $u_{ri}(t)$  时间间隔 5 秒内的速度最接近  $u_{ri}(t)$  的  $u_{mi}(t), i=1, 2, \dots, m$ ;
2. 寻找与  $u_{mi}(t)$  时间间隔 5 秒内的速度最接近  $u_{mi}(t)$  的  $u_{rj}(t), j=1, 2, \dots, m$ ;
3. 若  $i=j$ ，则找到一组可比较样本。
4. 以此类推找到每一检测点的所有可比较样本组，设为： $u_{ri}(t), u_{mi}(t), i=1, 2, \dots, n$ ;

以雷达测速仪速度为准，则 MC 检测到的单车车速的相对误差

$$\varepsilon_i = \frac{u_{ri}(t) - u_{mi}(t)}{u_{ri}(t)} * 100, i=1, 2, \dots, n;$$

$$\bar{\varepsilon} = \sum_{i=1}^n \varepsilon_i / n;$$

平均相对误差

$$MAPE = \frac{\sum_{i=1}^n |\varepsilon_i|}{n}$$

由数据分析可知（参见附录 C：断面检测器精度检验结果），主干道 D2、D3 的 MC 数据由于受公交车道和红绿灯影响较大，相对误差具有较大的方差，其余检测点相对误差的均值基本在零附近，方差也比较一致。雷达测速仪速度数据与 MC 速度数据趋势基本一致，且彼此间存在正态分布的误差，因此在实际计算过程中这些误差可作为白噪声处理。

表格 4-2 断面检测器误差分析

路段	检测点	$\bar{\varepsilon}$ (%)	MAPE (%)	$\mu$ (%)	$\sigma$ (%)
主干道	D <sub>1</sub>	-4.1404	11.4226	-3.3085	8.1501
	D <sub>2</sub>	-6.1694	11.7019	0.8363	17.8315
	D <sub>3</sub>	-4.4328	10.4602	-4.5015	12.6955
国道	D <sub>1</sub>	0.6383	11.2938	2.0959	8.6378
	D <sub>2</sub>	0.135	11.6454	1.4467	8.8406
	D <sub>3</sub>	1.7175	11.1354	1.7035	8.4624

## 4.4 数据对比及结果分析

### 4.4.1 浮动车行程时间误差分析

对每个路段进行数据修正及均值处理,得到的修正点及 15 分钟平均行程时间曲线如图 4-5 和图 4-6 所示,由图 4-5 可看出,遇到红灯的样本点数值降低,未遇到红灯的样本点数值增加,修正后的数据点较方差减小。图 4-6 说明了修正后的行程时间曲线比不加修正的情况更加光滑,也更加接近实际情况。例如,在 10:00 至 10:30 之间,原始样本行程时间突然增加,而人工实际观测中并未发生拥堵情况,说明使用均值估计出现了偏差,而修正均值估计有效的改善了这一情况。

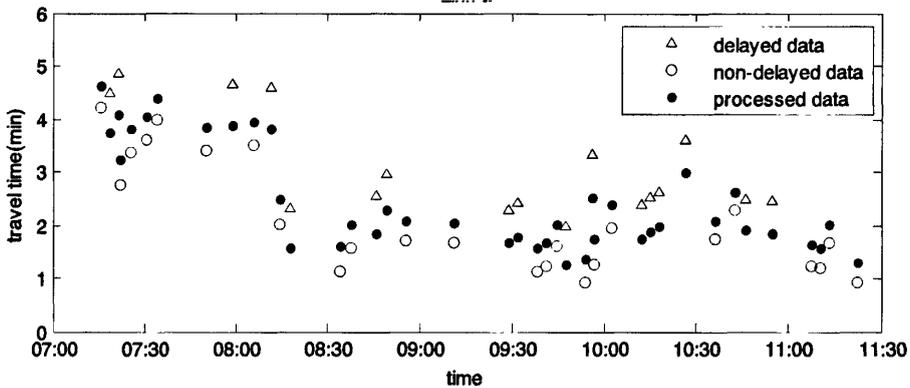


图 4-5 路段 1 (龙华西路-天钥桥路和) 行程时间原始点 (△:遇到红灯的点, ○:未遇到红灯点)、修正点 (●)

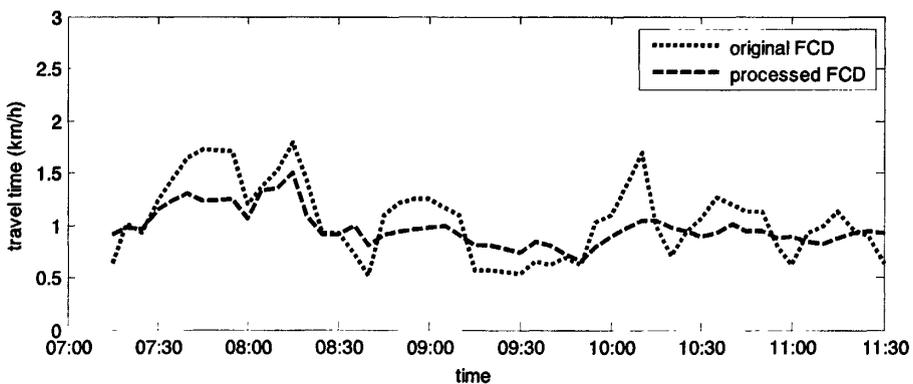


图 4-6 路段 1 (龙华西路-天钥桥路和) 样本点简单平均曲线 (---)、修正后的样本点平均曲线 (·)

将分析时段分为高峰期和非高峰期两种对计算结果进行分析,所有路段的均

值与误差对比如下表所示,为了表达简便,下述各表中使用方法 I 表示简单平均法,方法 II 表示本文的修正均值算法。另外,由于路段长度不一样,路段间行程时间相差较大,不能反映路段实际交通情况。为了对能够更直观的理解路段交通状态,表中统一使用行程速度作为估计结果。

所有路段统计结果如表格 4-3 和表格 4-4 所示,可看出对于某一路段,其均值在修正前后可能具有较大差异,如高峰时间段中的路段 7,路段 16 和路段 25,其修正后均值与原始均值分别为 20.69/15.74, 17.13/12.80 和 22.66/17.44,相对差达到 31%,34%和 30%,这样的差距可以使交通诱导系统对当前交通情况作出错误判断。但对所有路段的速度均值再次平均,则修正前后数值相差不大。这也说明在样本量足够的情况下,修正样本均值与原始均值一致,是实际情况的无偏估计。修正后的样本标准差和在 95%置信水平下的估计误差较原始样本明显下降,说明本算法能够有效降低估计误差。

对比非高峰时段与高峰时段,路段速度均值有所增长,高峰时段速度相差大的路段 7,路段 16 和路段 25 中,其修正前后非高峰速度分别为 15.85/14.51, 14.99/11.90 和 14.86/14.04,相对差为 9.23%,26%和 5.84%,原始样本均值数据质量要优于比高峰时段。由所有非高峰数据,计算得单路段均值平均误差为 9.0%,高峰数据的单路段均值平均误差为 10.6%,说明在非高峰时期,使用简单均值法得到的结果比高峰时间段的质量高。但从样本标准差和 95%置信水平下的误差方面比较,修正样本方法在非高峰时期并没有比在高峰时期的精度高。

表格 4-3 高峰时间段(7:00~9:00)统计结果比较

路段编号	方法 I (km/h)		方法 II (km/h)		方法 I 估计误差% (95%置信水平)	方法 II 估计误差% (95%置信水平)
	均值	样本标准差	均值	样本标准差		
1	8.76	2.76	8.44	1.08	20.42	14.38
2	21.42	8.69	24.93	7.57	23.76	15.76
3	15.46	10.58	11.34	1.85	27.28	16.29
4	11.84	3.20	11.87	2.05	22.11	15.50
5	28.20	3.02	28.41	1.48	17.82	10.18
6	25.00	11.22	24.63	5.70	32.21	18.08
7	20.69	7.05	15.74	3.45	30.83	16.00
8	19.19	8.20	19.91	6.02	28.44	16.04
9	16.88	8.52	14.01	4.22	33.85	19.73
10	8.24	2.65	7.56	0.99	31.25	17.08
11	13.89	6.81	15.91	8.40	16.67	9.71
12	21.13	5.91	20.96	3.20	29.30	13.49
13	15.89	9.22	13.45	2.65	31.50	15.14
14	36.09	5.67	37.35	3.24	13.56	8.59

15	28.04	12.96	34.43	14.57	30.48	19.27
16	17.13	9.40	12.80	3.91	35.05	22.01
17	24.23	14.29	26.06	11.83	28.92	17.28
18	34.03	8.55	34.51	4.71	25.04	11.87
19	10.90	1.55	10.29	0.64	25.20	13.25
20	23.04	6.42	23.48	2.91	25.00	14.67
21	30.72	6.11	30.60	3.08	18.96	10.31
22	19.24	5.53	21.01	2.80	28.21	13.98
23	26.31	4.19	26.96	2.48	19.14	13.25
24	27.34	10.49	24.40	4.32	26.08	14.90
25	22.66	9.37	17.44	5.70	36.55	23.85
26	19.00	6.30	17.58	2.23	27.59	17.28
27	28.15	5.92	26.23	2.51	20.35	11.95
28	15.73	7.94	14.66	3.44	28.53	19.39
29	29.91	8.89	28.70	4.03	18.80	11.27
30	27.58	5.24	28.21	2.63	15.15	11.41
31	19.33	11.42	15.69	3.29	29.59	13.95
32	12.66	3.50	12.67	1.94	23.25	12.32
平均	21.21	7.24	20.64	4.25	25.65	14.94

表格 4-4 非高峰时间段(9:00~12:00)统计结果比较

路段 编号	方法 I (km/h)		方法 II (km/h)		方法 I 估计误差% (95%置信水平)	方法 II 估计误差% (95%置信水平)
	均值	标准差	均值	标准差		
1	12.33	3.83	12.21	1.78	29.58	18.46
2	26.85	6.52	28.86	3.72	29.74	21.60
3	19.50	10.51	16.28	2.54	22.83	15.06
4	17.73	8.01	15.89	2.51	24.66	17.59
5	35.63	3.75	35.72	1.99	15.62	7.00
6	33.18	12.23	26.58	6.38	34.57	16.61
7	15.85	7.31	14.51	3.52	34.25	24.48
8	28.20	8.38	29.19	6.79	36.83	31.40
9	19.17	8.42	18.07	4.45	30.69	24.65
10	19.87	11.03	17.81	5.32	16.55	12.55
11	32.21	4.02	34.43	2.76	29.01	22.62
12	31.00	8.86	27.60	4.59	27.50	20.72
13	16.79	7.30	13.77	2.71	30.51	24.89
14	44.10	3.82	45.13	1.64	20.13	15.95
15	38.79	11.49	38.23	9.90	39.31	31.42
16	14.99	9.85	11.90	3.68	30.99	25.84
17	31.67	12.07	31.78	13.26	33.09	26.49
18	37.38	7.76	38.16	5.62	22.64	11.96

19	13.90	6.93	11.09	1.45	23.74	13.30
20	22.40	4.81	21.98	2.46	30.65	16.15
21	32.39	5.83	33.00	3.36	26.62	19.79
22	26.90	9.37	24.39	5.34	30.83	23.25
23	23.83	4.37	24.58	2.49	19.13	11.38
24	26.06	10.33	25.09	4.52	29.85	15.62
25	14.86	7.77	14.04	3.44	34.83	21.66
26	21.51	7.08	19.19	2.86	25.47	16.57
27	27.69	5.45	26.11	1.86	20.58	14.23
28	18.82	8.93	16.11	3.28	31.33	24.81
29	28.00	5.98	26.70	2.49	23.64	12.84
30	27.54	5.54	27.62	3.03	20.63	11.29
31	20.20	9.58	15.71	3.30	34.65	25.40
32	17.85	4.70	20.25	2.21	19.77	13.72
平均	24.91	7.56	23.73	3.91	27.51	19.04

对于扩展路段行程时间情况，统计其误差与相应的单路段平均误差，如表格 4-5 所示。由表格 4-5 可知大部分情况下在扩展路段使用修正样本会大幅降低估计误差，其平均误差降低率为 41%。但其误差降低状况不稳定，个别情况下，如路段 11~20 的高峰期误差，扩展路段误差反而大于单路段平均情况。说明扩展情况下的行程时间估计受路段划分情况和交通状况影响较大。

表格 4-5 扩展路段误差与单路段误差均值对比（95%置信水平）

路段编号	高峰期		非高峰期	
	扩展路段估计误差 (%)	单路段平均估计误差 (%)	扩展路段估计误差 (%)	单路段平均估计误差 (%)
路段 1~10	11.96	15.90	5.05	18.94
路段 11~20	16.47	14.53	7.77	20.93
路段 21~26	9.24	15.60	14.17	18.05
路段 27~32	6.55	13.38	5.16	17.05

#### 4.4.2 浮动车与断面估计行程时间对比

浮动车估计路段行驶速度与固定设备估计路段行驶速度进行比较，得到曲线和误差分析如表格 4-6、表格 4-7、图 4-7、图 4-8 所示，由此可知断面检测器估计行程时间与浮动车估计行程时间趋势相似，路段 1 由于速度值本身较小，尽管速度误差仅为 1~2 km/h，相对误差也达到 10%以上，但这种差距不会造成交通状态的误判。路段 32 在非高峰时段两种估计值较吻合，相差仅 5%，但在高峰时段

趋势相近但数值相差较大。两个路段上都具有检测器行程速度估计值在高峰时段偏高的特性，分析其原因可能是高峰期时信号灯对车辆的延误可能达到 1 个周期以上，这样由于本文中的断面检测数据估计方法仅估计 1 个信号灯周期内的误差，造成行程速度估计值偏高。

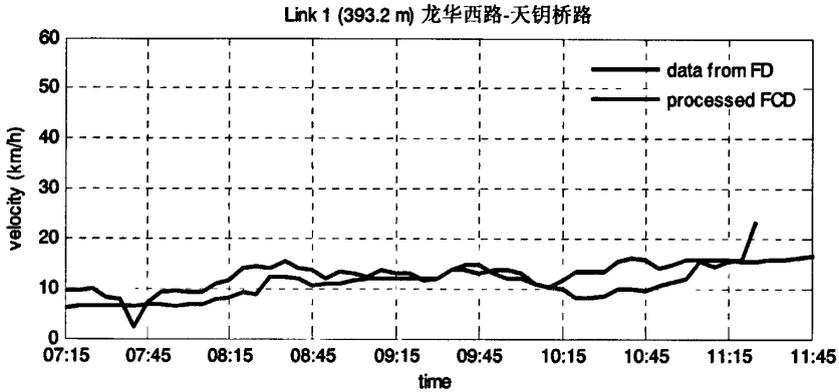


图 4-7 路段 1（龙华西路-天钥桥路和）检测器估计行程速度与浮动车估计行程速度对比

表格 4-6 路段 1（龙华西路-天钥桥路和）固定检测器与浮动车估计行程速度比较

时段	固定检测器速度		浮动车速度		误差分析		
	均值	时段均值标准差	均值	时段均值标准差	误差均值	相对平均误差 (%)	误差标准差
高峰期 (7:00~9:00)	10.83	3.12	8.44	2.32	-2.4	-17.88	1.71
非高峰期 (9:00~12:00)	13.75	1.6	12.2	3.02	-1.67	-11.69	2.58

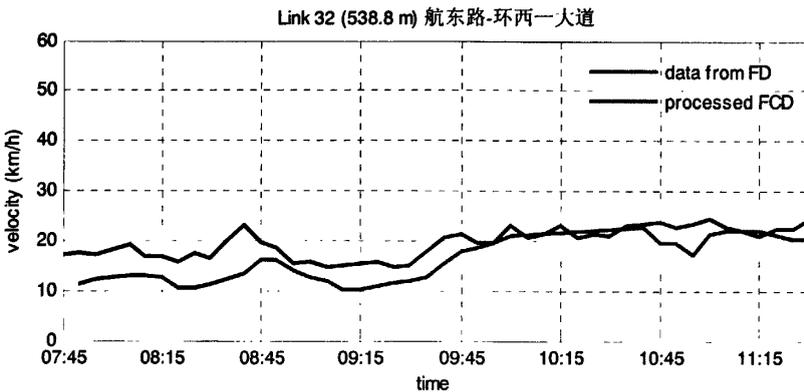


图 4-8 路段 32（航东路-环西一大道）检测器估计行程速度与浮动车估计行程速度对比

表格 4-7 路段 32 (航东路-环西一大道) 固定检测器与浮动车估计行程速度比较

时段	固定检测器速度		浮动车速度		误差分析		
	均值	时段均值标准差	均值	时段均值标准差	误差均值	相对平均误差 (%)	误差标准差
高峰期 (7:00~9:00)	17.47	2.24	12.67	1.93	-4.81	-27.33	1.64
非高峰期 (9:00~12:00)	22.13	1.41	20.25	1.69	-1.24	-5.33	1.8

## 5 结论与展望

本文提出了一种由小样本浮动车数据估计带信号灯路段行程时间的方法，并对实际数据进行实证分析，得到结论如下：

### 1) 相比传统方法，本文提出的方法能够减小估计误差

在 10 分钟内只有约 8 个行程时间数据的情况下，本算法提供了准确度更高的行程时间估计值，表现在样本方差和误差的降低，降低幅度约为 40%，这一特性在高峰期和非高峰时段没有明显差异。

### 2) 由本文提出的方法估计得到的平均行程时间能够反应实际交通情况

根据与断面检测数据对比结果，发现本方法估计得到的平均行程速度与断面检测数据结果趋势一致，在非高峰时段差异较小，二者平均速度差大约在 1~2km/h 之间；高峰时段二者间可能具有由多个信号灯延误导致的系统误差，这使得断面数据估计行程速度偏高。这说明估平均行程速度能够反应实际交通情况。

综上，本方法能够较好的对带信号灯路口交通状态进行评估，具有实际应用价值。

对信号灯路段的小样本行程时间估计问题还可在以下方面展开进一步研究：

### 1) 结合 GPS 定位数据进一步完善算法

本文中算法直接使用单量浮动车的行程时间样本数据对平均行程时间进行估计，但在大多数浮动车应用中系统采集到的是固定间隔的 GPS 定位数据，需要经过处理才能得到路段行程时间，进而应用本算法。因此将 GPS 数据处理和小样本估计算法有机结合起来是下一步研究方向之一。

### 2) 提高算法的通用性

应用本文提出的算法估计行程时间的必要前提之一是已经事先得知欲评估路段的信号灯相位，而得到这一确切信息具有一定的难度，这在一定程度上限制了算法的通用性。通过对数据的实际观察发现，相位信息与样本数据分布有一定联系，因此先根据历史数据推测出信号灯相位，再按本文的算法计算平均行程时间是提高算法通用性的一个途径。

## 参考文献

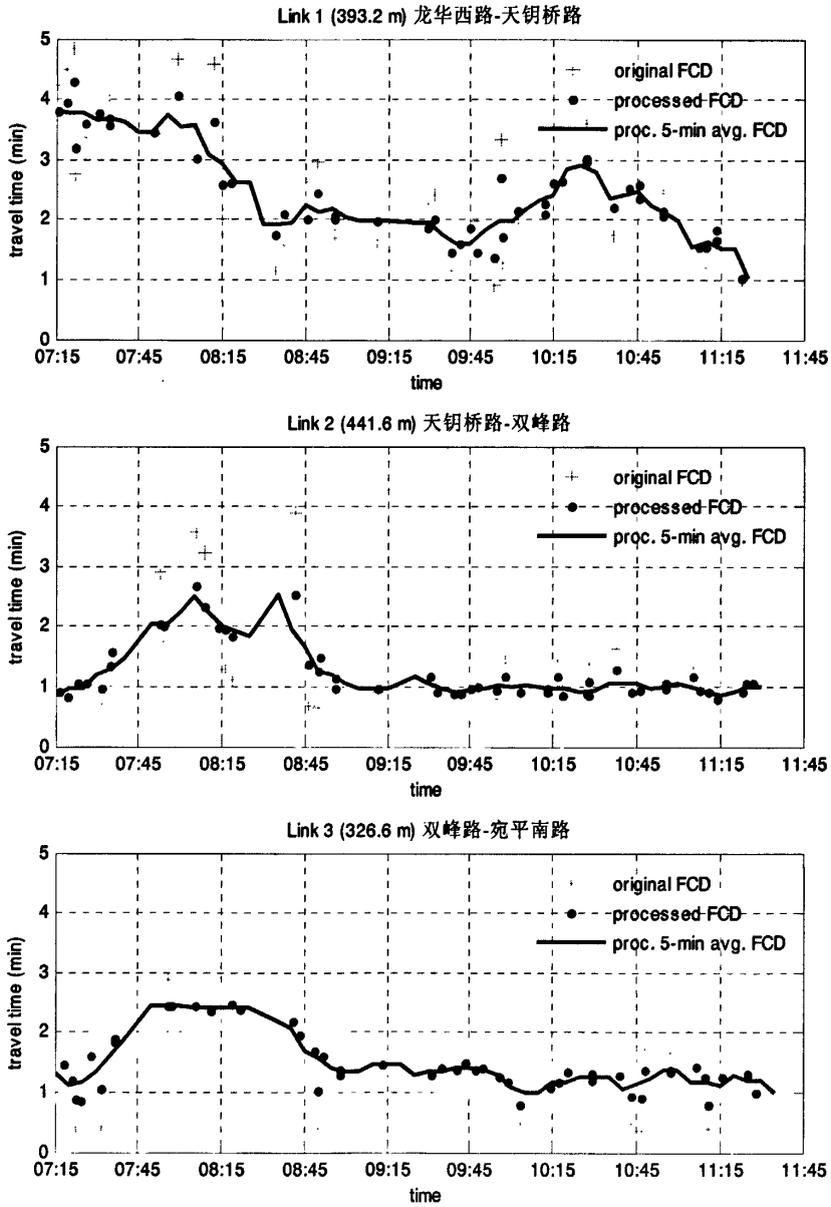
- [1] Boker, G. (2000) "Traffic state analysis on the basis of floating car data". *Automatisierungstechnik*, Vol. 48, No. 8, pp. 365-371.
- [2] Huber, W., Lädke M. and Ogger R. (1997) "Extended floating-car data for the acquisition of traffic information". In *Proceedings of the 4th World Congress on Intelligent Transport Systems*. Berlin, Germany.
- [3] Kerner, B.S., et al. (2005) "Traffic state detection with floating car data in road networks". In *Proceedings of the 8th Intelligent Transportation Systems*. Vienna, Austria.
- [4] Nakata, T. and Takeuchi J. (2004) "Mining traffic data from probe-car system for travel time prediction". In *Proceedings of International Conference on Knowledge Discovery and Data Mining: ACM Press New York, NY, USA*.
- [5] Schäfer, R.P., et al. (2002) "Analysis of Travel Times and Routes on Urban Roads by Means of Floating-car Data". In *Proceedings of the 9th World Congress on Intelligent Transport Systems*.
- [6] 孙晓峰, 吴建平, 基于浮动车数据采集技术的城市交通网络功能评价方法研究. *现代交通技术*, 2005. 2(6): p. 55-58.
- [7] 王力, 张海, 范耀祖, 基于探测车技术的路段平均速度估计模型. *交通运输系统工程与信息*, 2006. 6(4): p. 29-33.
- [8] 朱丽云, 温慧敏, 孙建平, 北京市浮动车交通状况信息实时计算系统. *城市交通*, 2008. 6(1): p. 77-80.
- [9] Boyce, D., Kirson A. and Schofer. J (1993) "Design and Implementation of ADVANCE". In *Proceedings of 3rd International Conference on Vehicle Navigation and Information Systems: Ottawa*, pp. 415-426.
- [10] Kersken, U. and Baierl W. (1998) "Method and telematics equipment for the collection and transferring the relevant motoring information", European patent publication EP 0892 379 A2, holder: Robert Bosch. GmbH, day of notification: 25.06.1998.
- [11] Bishop, R. (2004) *Floating Car Data Projects Worldwide: a selective review* [EB/OL]. <http://ivs.ollrce.net/archivep/2004/apr/>, 2004.
- [12] Kroes, E., Hagemeyer F. and Linssen J. (1999) "A new, probe vehicle-based floating car data system: Concept, implementation and pilot study". *Traffic engineering & control*, Vol. 40, No. 4, pp. 200-204.
- [13] 路加 (2003). *交通拥挤的度量方法与基于浮动车的交通拥挤检测*. 自动化系. 北京, 清华大学. 硕士学位论文.

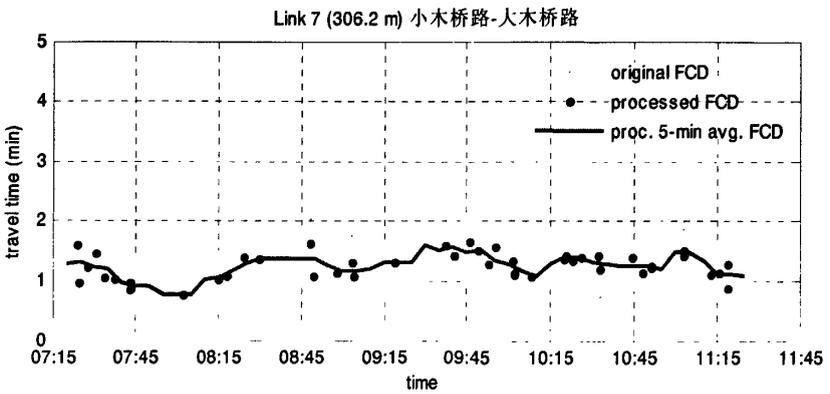
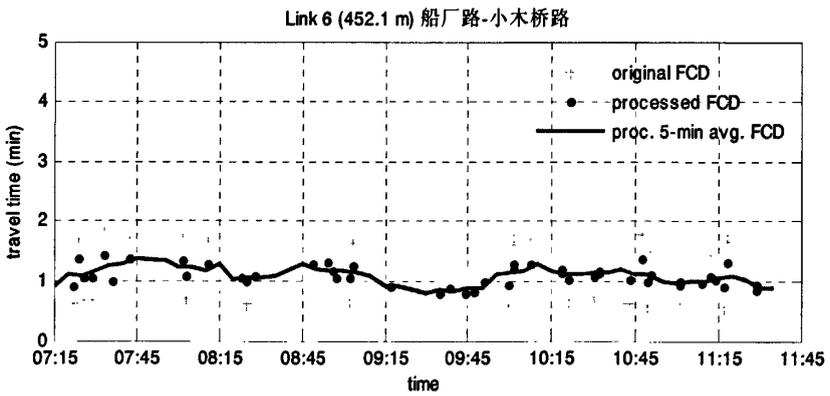
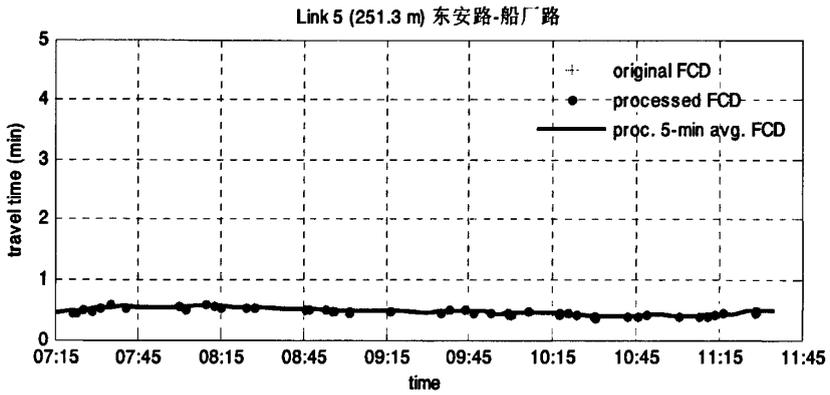
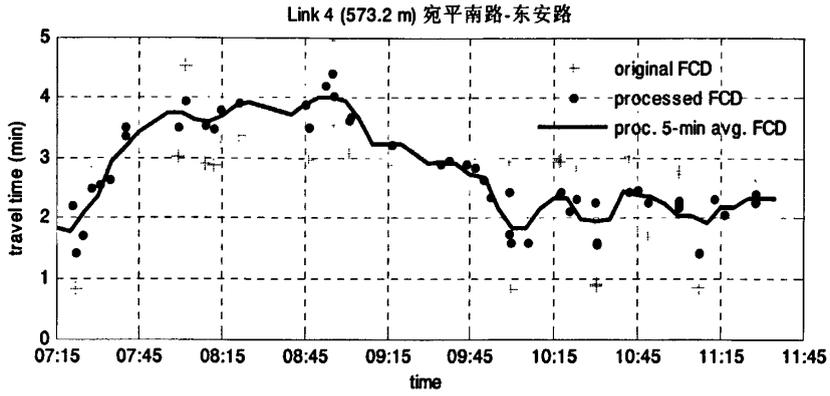
- [14] 翁剑成, 荣建, et al. (2007). "基于浮动车数据的行程速度估计算法及优化." 北京工业大学学报 33(5): 459-464.
- [15] 章威, 徐建闽, et al. (2007). "基于浮动车技术的城市路况计算方法." 交通运输系统工程与信息 7(1): 43-47.
- [16] Turner, S. M., W. L. Eisele, et al. (1998). "Travel Time Data Collection Handbook." Federal Highway Administration, Report FHWA-PL-98-035, Mar.
- [17] Huisken, G. and E. C. Van Berkum (2003). "A Comparative Analysis of Short-Range Travel Time Prediction Methods." Proc. of 82nd Annual TRB meeting.
- [18] Hall, F. and B. Persaud. Evaluation of Speed Estimates Made with Single-Detector Data from Freeway Traffic Management Systems. In Transportation Research Record 1232, 1989, pp. 9-16.
- [19] Pushkar, A., F.L. Hall, and J.A. Acha-Daza. Estimation of Speeds from Single-loop Freeway Flow and Occupancy Data Using Cusp Catastrophe Theory Model. In Transportation Research Record 1457, pp. 149-157, 1994.
- [20] Dailey, D. J. A Statistical Algorithm for Estimating Speed from Single Loop Volume and Occupancy Measurements. Transportation Research B, Vol. 33, pp. 313-322, 1999.
- [21] Hellinga, B. R. (2002) Improving Freeway Speed Estimates from Single-Loop Detectors. Journal of Transportation Engineering, vol. 128, No.1, pp58-67.
- [22] Luk, J. On-line Journey Time Estimation. Modeling and Monitoring the Performance of Urban Traffic Control Systems, Australian Road Research Board. ARRB SR 43, 1989.
- [23] Dailey, D. J. Travel-time Estimation Using Cross-correlation Techniques. Transportation Research B, Vol. 27, pp. 97-107, 1993.
- [24] Nam, D. H. and D. R. Drew. Traffic Dynamics: Method for Estimating Freeway travel Times in Real Time From Flow Measurements. Journal of Transportation Engineering, Vol. 122, 1989, pp. 185-191.
- [25] Petty, K.F., Bickel P., Ostland, M., Rice, J., Schoenberg, F., Jiang, J., and Ritov, Y. (1998) Accurate Estimation of Travel Times from Single Loop Detectors, Transportation Research A, Vol. 32, No. 1, pp.1-17.
- [26] Ran B, Roupail N M, Tarko A, et al., Towards a class of link travel time functions for dynamic assignment models on signalized network. Transportation Research Part B, 1997, 31(4): 277-290.
- [27] Box, G.E.P. and G.C. Tiao, Bayesian Inference in Statistical Analysis. 1973, DTIC Research Report AD0760786.
- [28] 邹亮, 徐建闽, 朱玲湘, 温惠英, 基于行程时间估计的快速路检测器布设密度优化方法研

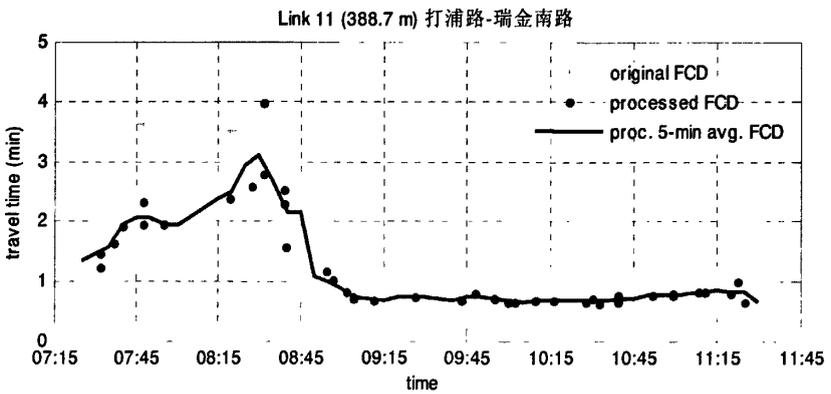
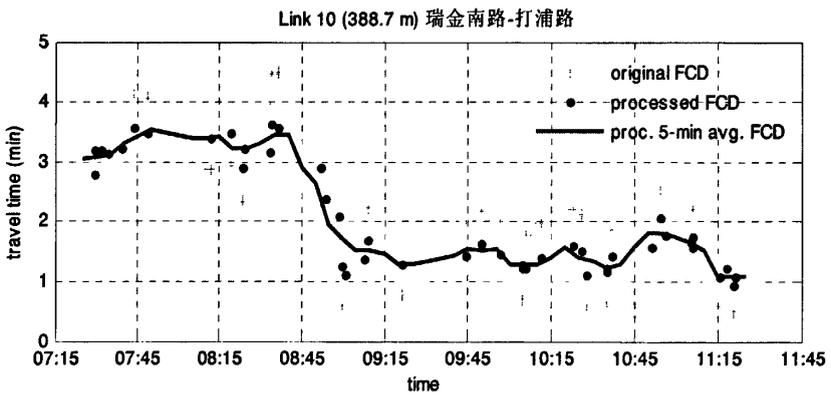
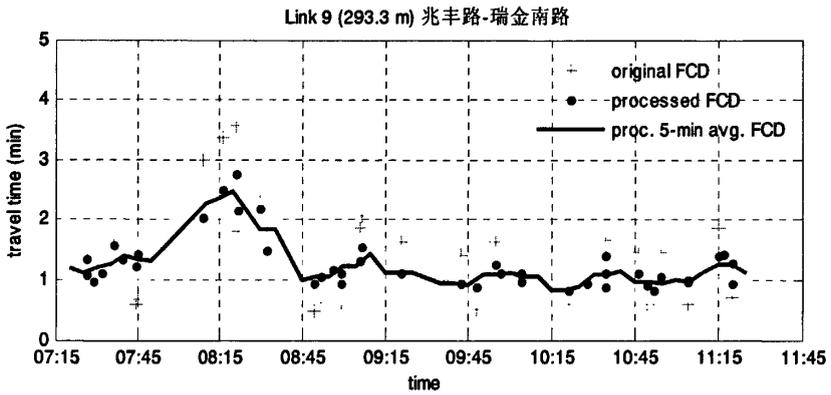
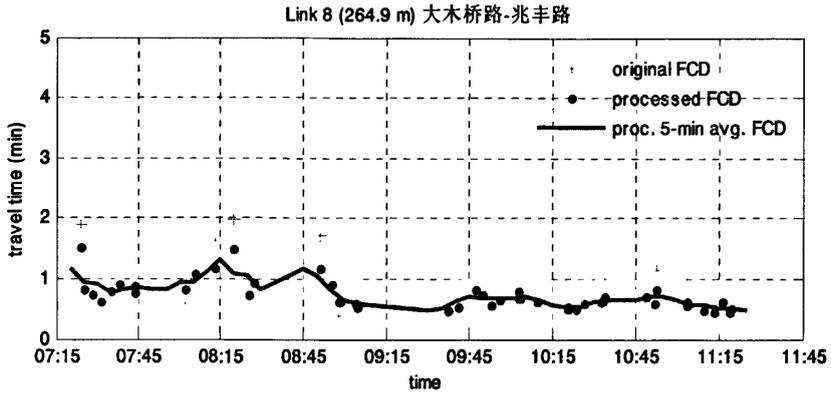
- 究. 公路交通科技, 2007. 24(6): p. 114-117.
- [29] Webster, F. V. (1958) Traffic Signal Settings. Road Research Laboratory Technical Paper No. 39, HMSO, London.
- [30] Catling, I. (1977) "A Time-Dependent Approach to Junction Delays". Traffic Engineering and Control, pp. 520-523, 526.
- [31] Lin, F. B. and F. Mazdeyasa (1983). "Delay Models of Traffic Actuated Signal Controls." Transportation Research Record 905: 33-38.
- [32] Lianyu Chu et, al. Adaptive Kalman Filter Based Freeway Travel time Estimation. Trans Res Board. 2005.
- [33] 张和生, 张毅, 温慧敏, 胡东成, 利用 GPS 数据估计路段的平均行程时间. 2007, 37(3): 533-537.
- [34] Wang, Z. (2004) "Using Floating Cars to Measure Travel Time Delay: How Accurate Is the Method?" In Transportation Research Record: Journal of the Transportation Research Board, No. 1853, TRB, National Research Council, Washington, D.C., pp. 84-93.
- [35] Turksma, S. (2000) "The Various Uses of Floating Car Data". In Proceedings of Road Transport Information and Control.
- [36] 张存保, 杨晓光, 严新平, 移动交通检测系统中探测车的样本数量. 中国公路学报, 2007. 20(1): p. 96-101.
- [37] 姚琛, 基于路段覆盖率的浮动车样本数量研究. 山东理工大学学报, 2006. 20(2): p. 96-98.
- [38] Janc, M., Probe Vehicle Population and Sample Size for Arterial Speed Estimation. Computer-Aided Civil and Infrastructure Engineering, 2002. 17(11): p. 53-60.
- [39] Chen, M. and S.I.J. Chien, Determining the Number of Probe Vehicles for Freeway Travel Time Estimation by Microscopic Simulation. Transportation Research Record, 2000. 1719: p. 61-68.
- [40] Cohen, S., et al. Probe vehicle sample sizes for travel time estimation on equipped motorways. in Road Transport Information and Control, 2002. Eleventh International Conference on (Conf. Publ. No. 486). 2002.
- [41] 涂智, 李昊, et al. (2006). "基于最小浮动车样本数量的道路覆盖率与交通信息更新周期研究." 中国铁道科学 27(5): 127-131.
- [42] 陈小鸿, 冯均佳, et al. (2007). "基于浮动车数据的行程时间可靠度特征研究." 城市交通 5(5): 42-45.
- [43] Nanthawichit, C., T. Nakatsuji, et al. (2003). "Application of Probe-Vehicle Data for Real-Time Traffic-State Estimation and Short-Term Travel-Time Prediction on a Freeway." Transportation Research Record 1855: 49-59.
- [44] Liu, K., T. Yamamoto, et al. (2006). "An Analysis of the Cost Efficiency of Probe Vehicle Data

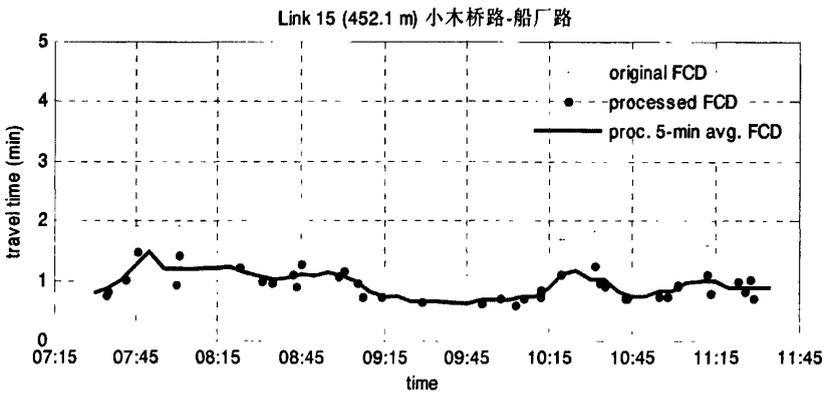
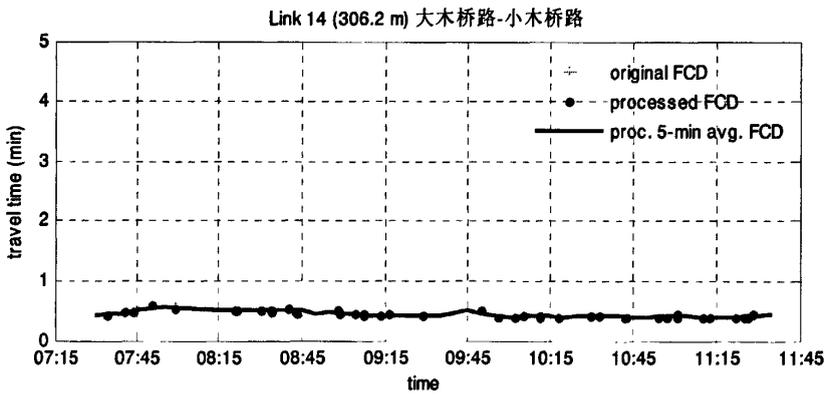
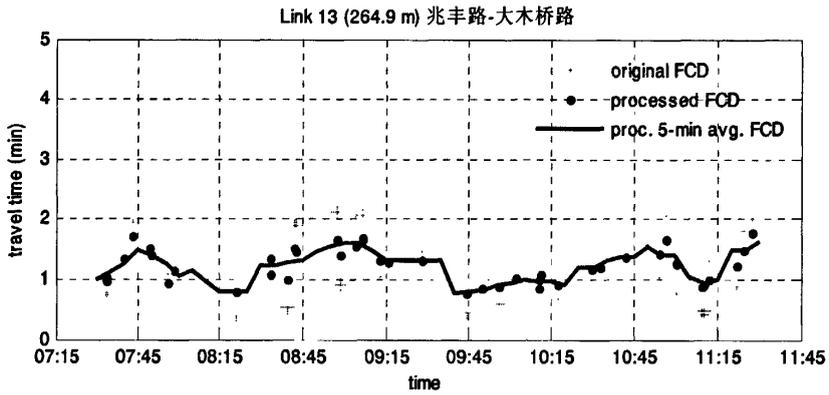
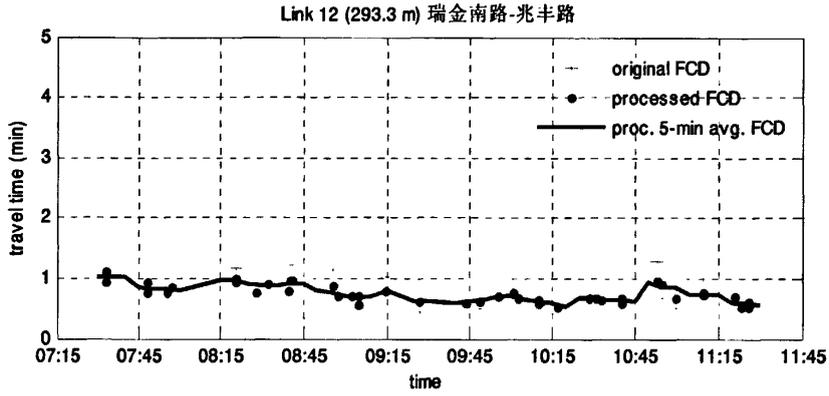
- at Different Transmission Frequencies." *International Journal of ITS Research* 4(1): 21-28.
- [45] 贾元华, 吴贵福, et al. (2006). "道路网络中移动型检测器最小样本量优化方法研究." *交通与运输(B07)*: 32-35.
- [46] 王力, 张海, et al. (2007). "移动交通流检测技术探测车样本比例模型研究." *交通运输系统工程与信息* 5.
- [47] Hellinga, B. R. and L. Fu (2002). "Reducing bias in probe-based arterial link travel time estimates." *Transportation Research Part C* 10(4): 257-273.
- [48] 张存保, 严新平, 固定检测器和移动检测器的交通信息融合方法. *交通与计算机*, 2007. 25(3): p. 14-17, 22.
- [49] 邹亮, 徐建闽, et al. (2007). "基于浮动车移动检测与感应线圈融合技术的行程时间估计模型." *公路交通科技* 24(6): 114-117.
- [50] 钱寒峰, 林航飞, et al. (2007). "浮动车车速处理分析系统中的数据融合技术." *计算机工程与应用* 43(31): 230-232.
- [51] Ivan, J. N. and V. Sethi (1998). "Data Fusion of Fixed Detector and Probe Vehicle Data for Incident Detection." *Computer-Aided Civil and Infrastructure Engineering* 13(5): 329-337.
- [52] 温惠英, 徐建闽, et al. (2006). "基于灰色关联分析的路段行程时间卡尔曼滤波预测算法." *华南理工大学学报: 自然科学版* 34(9): 66-69.
- [53] Box, G. E. P. and G. C. Tiao (1973). *Bayesian Inference in Statistical Analysis*, DTIC Research Report AD0760786.
- [54] Efron, B. and R. Tibshirani, *An Introduction to the Bootstrap*. 1993: Chapman & Hall/CRC.

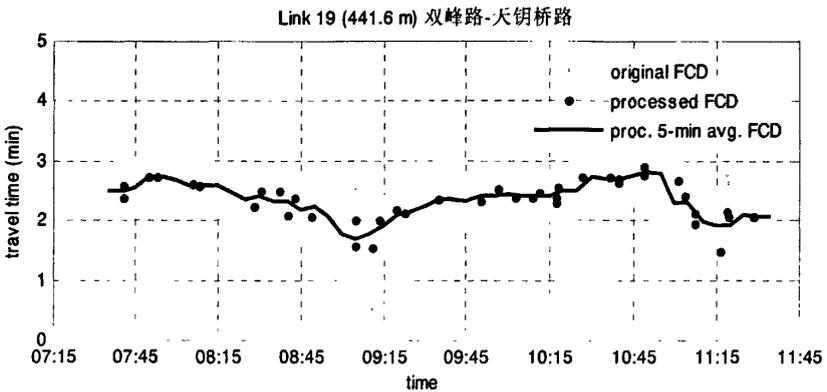
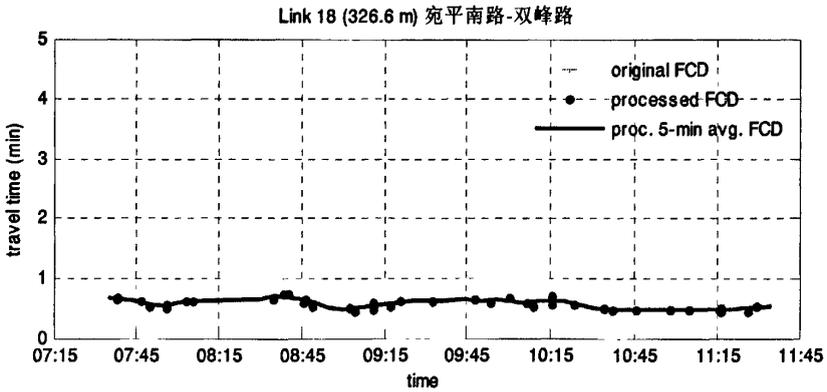
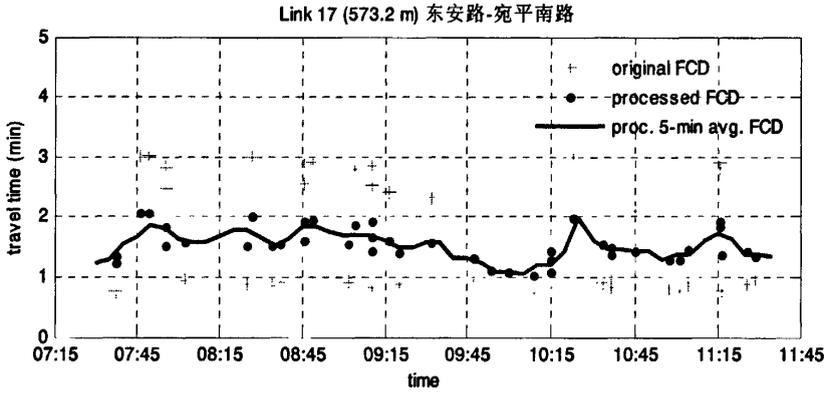
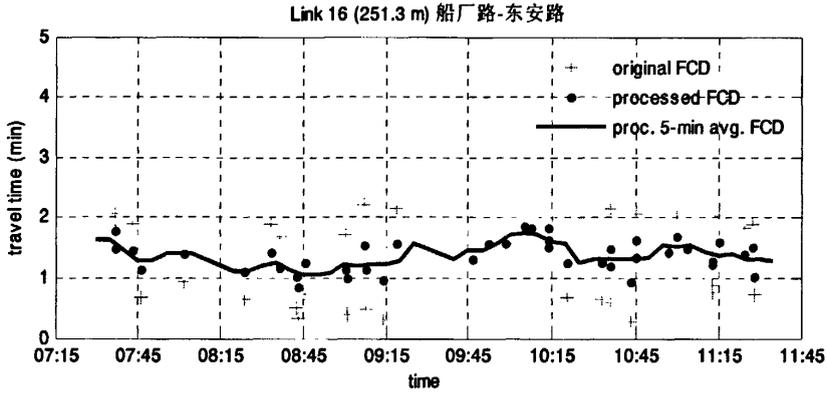
### 附录 A: 浮动车全部路段行程时间计算结果

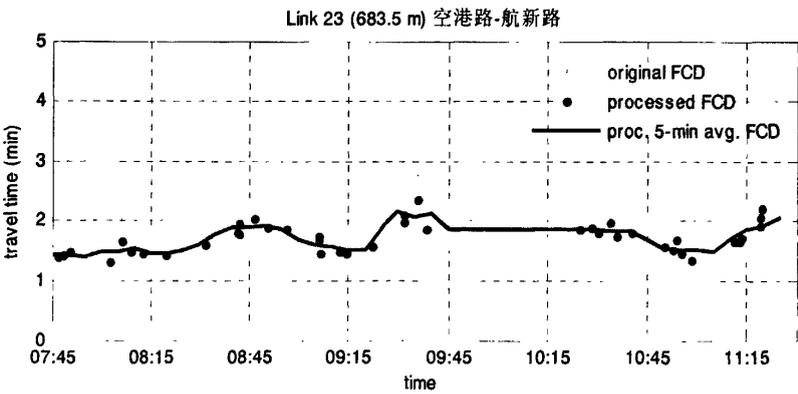
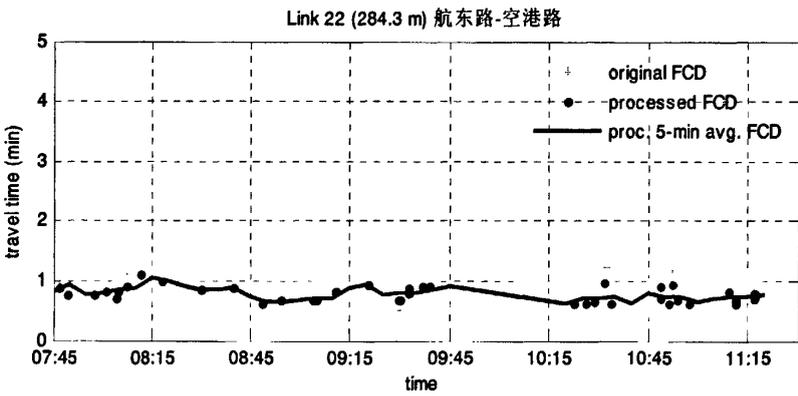
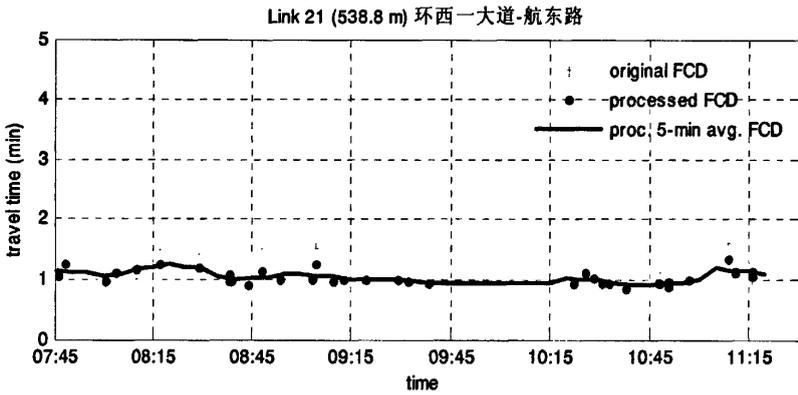
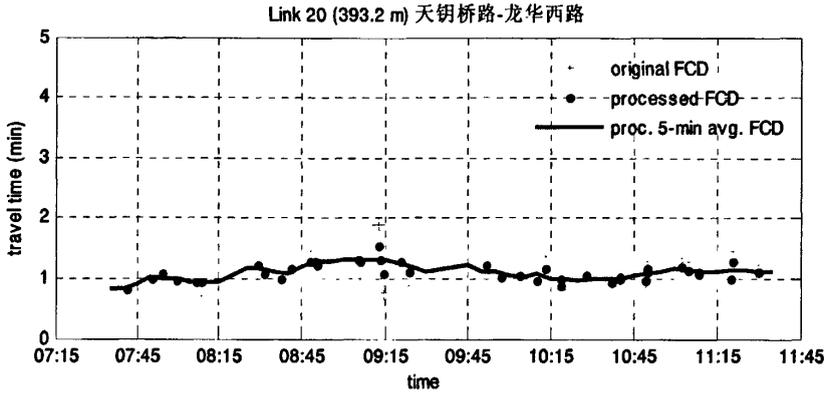


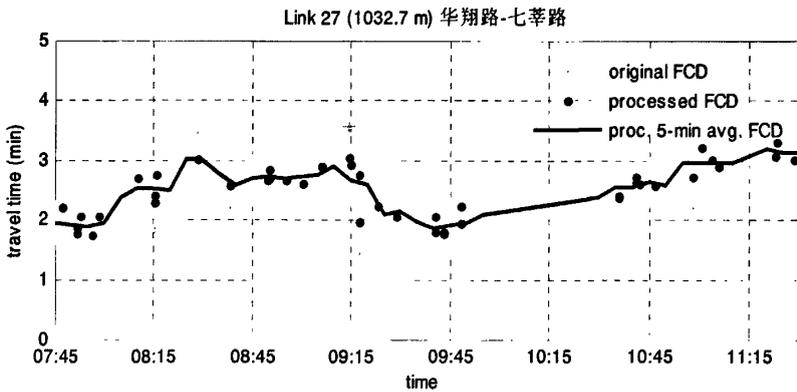
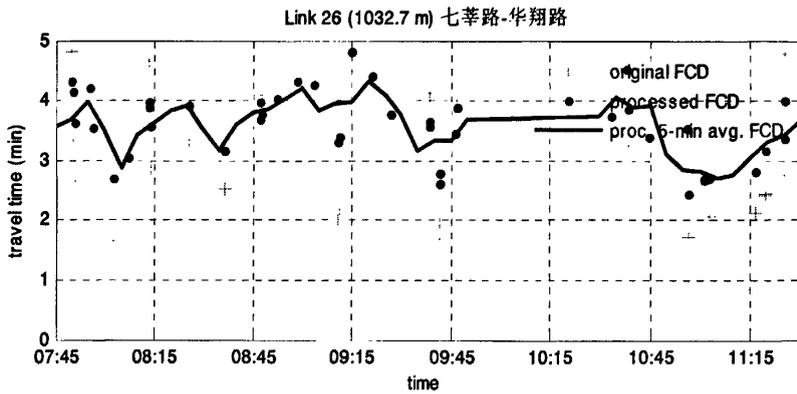
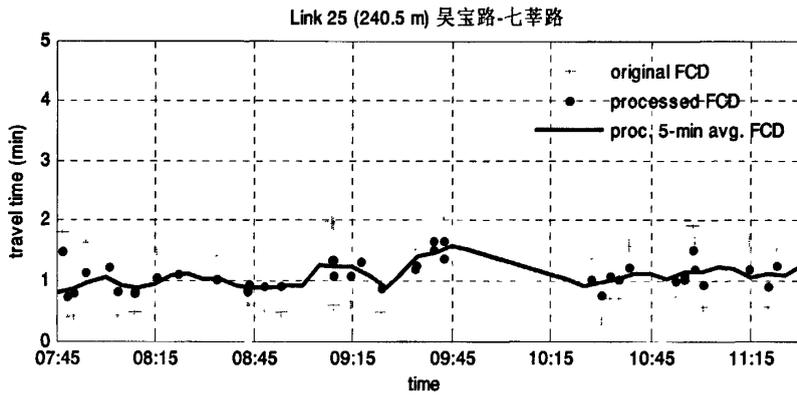
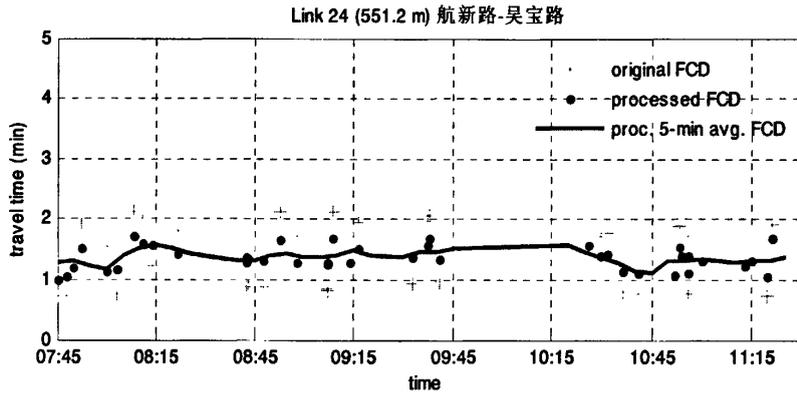


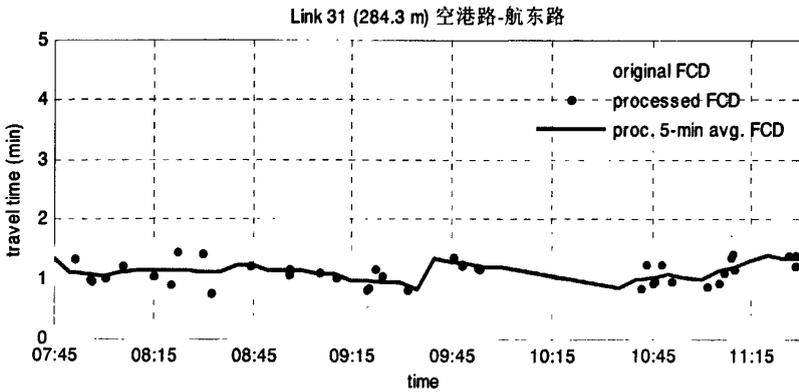
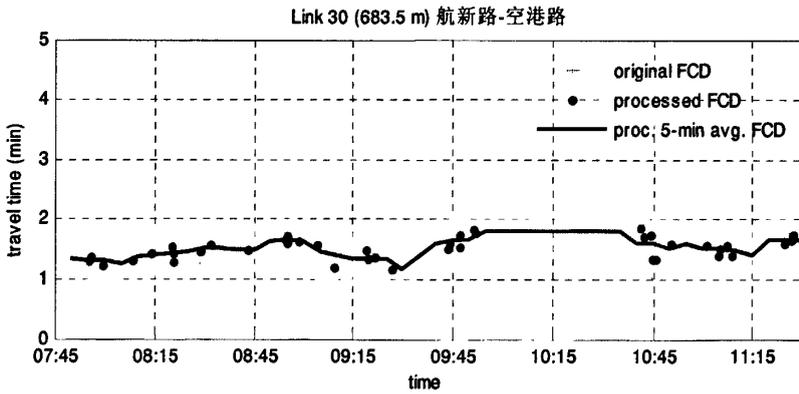
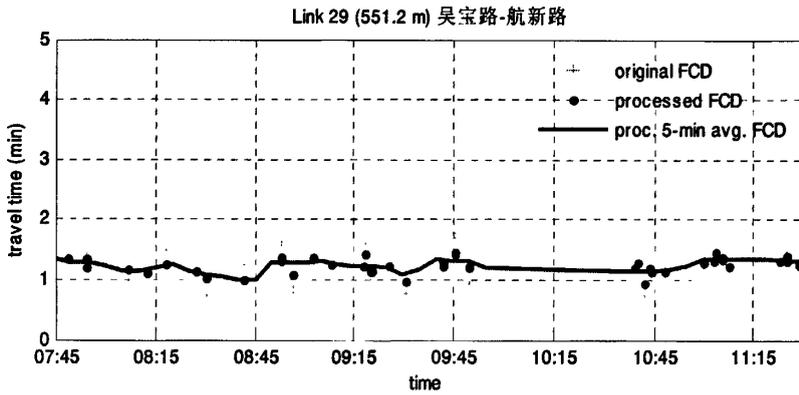
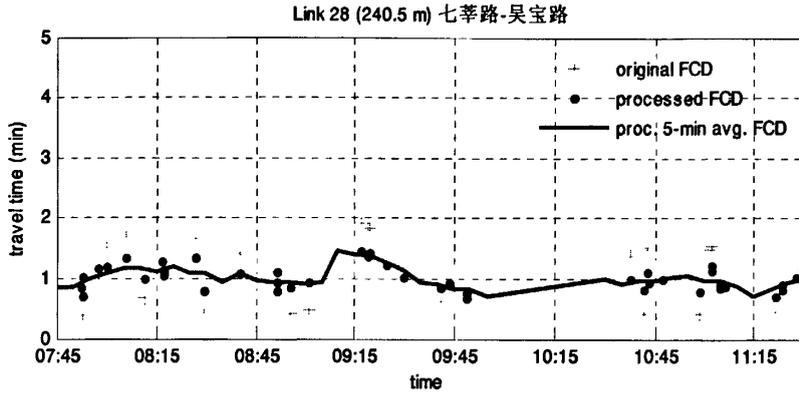


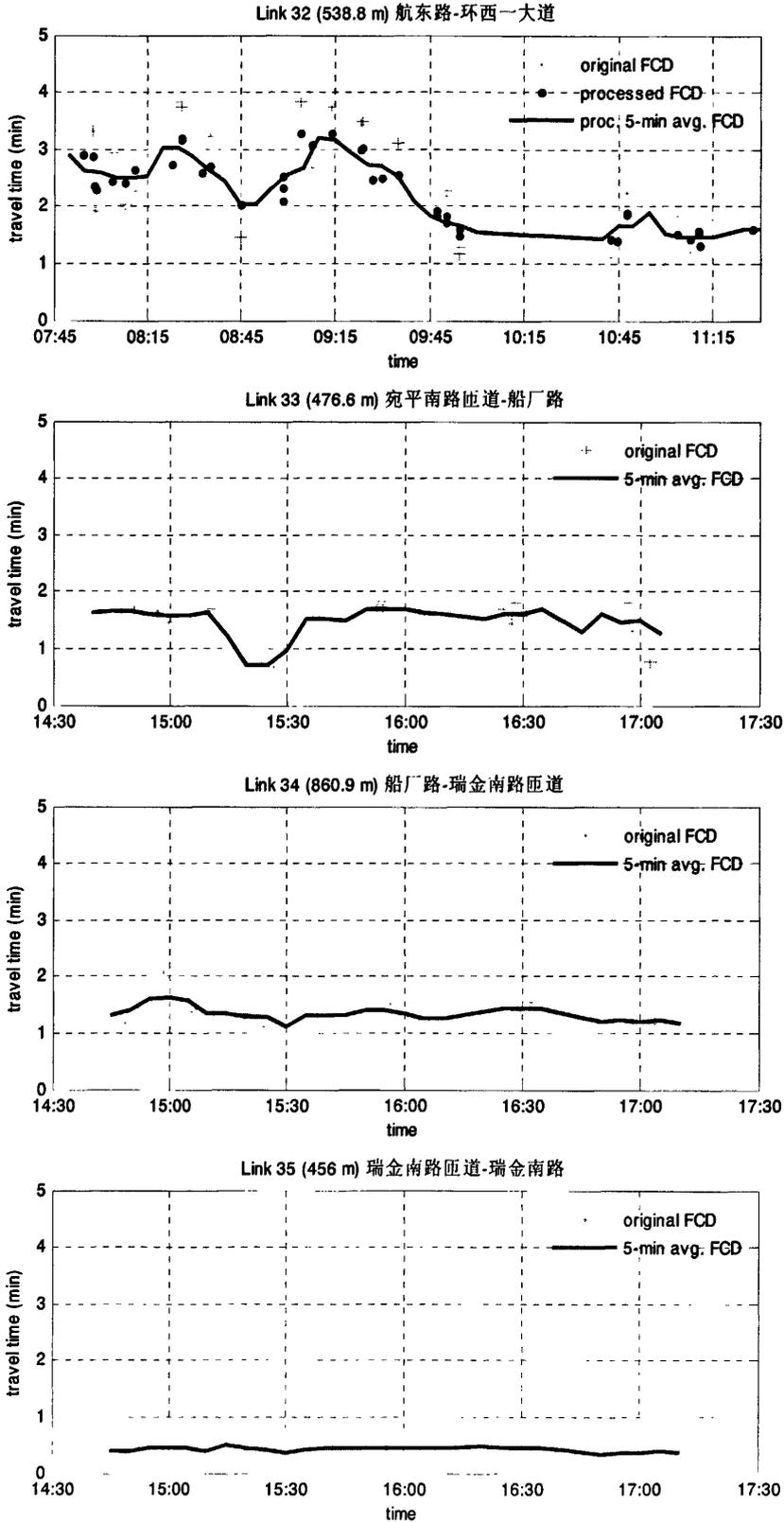


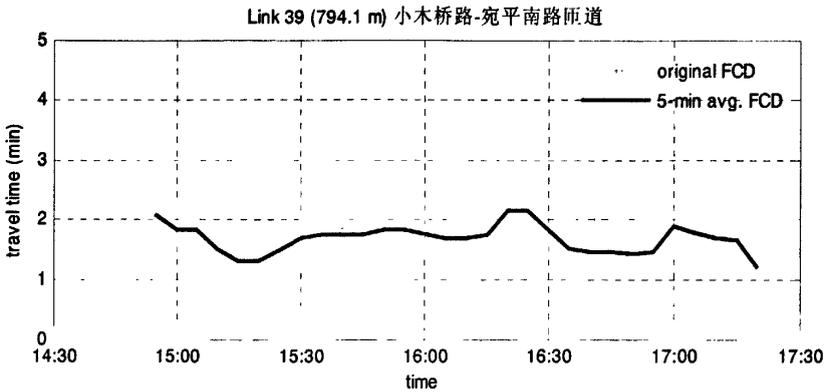
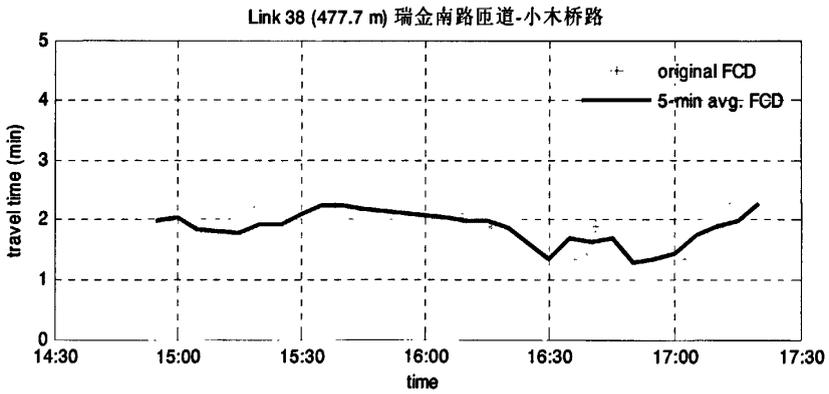
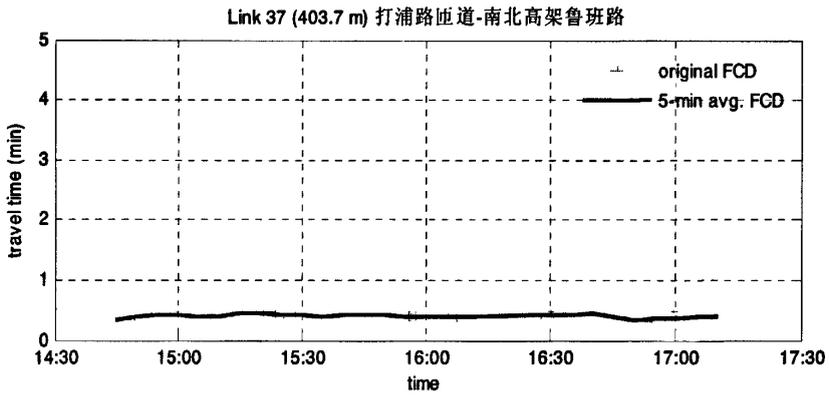
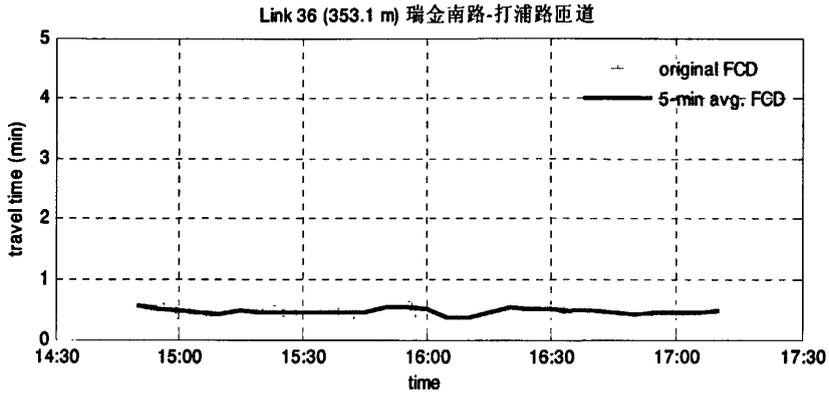






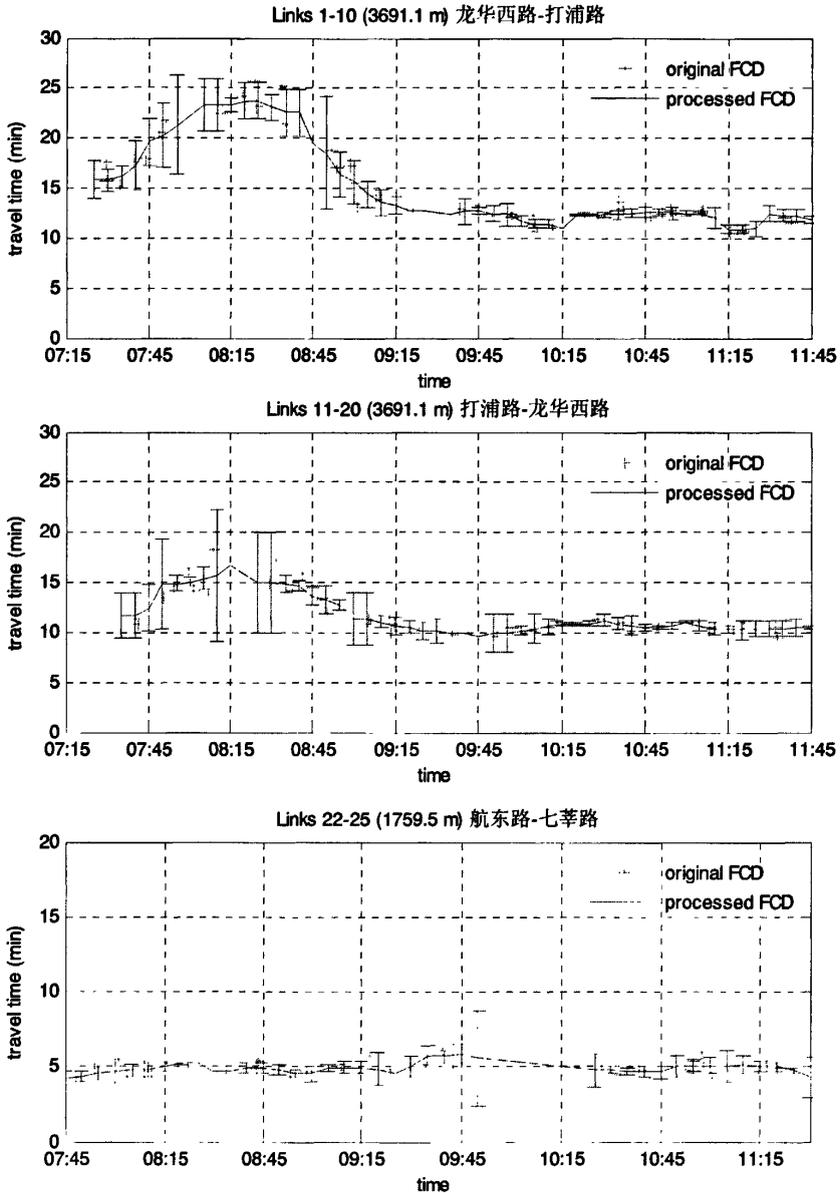


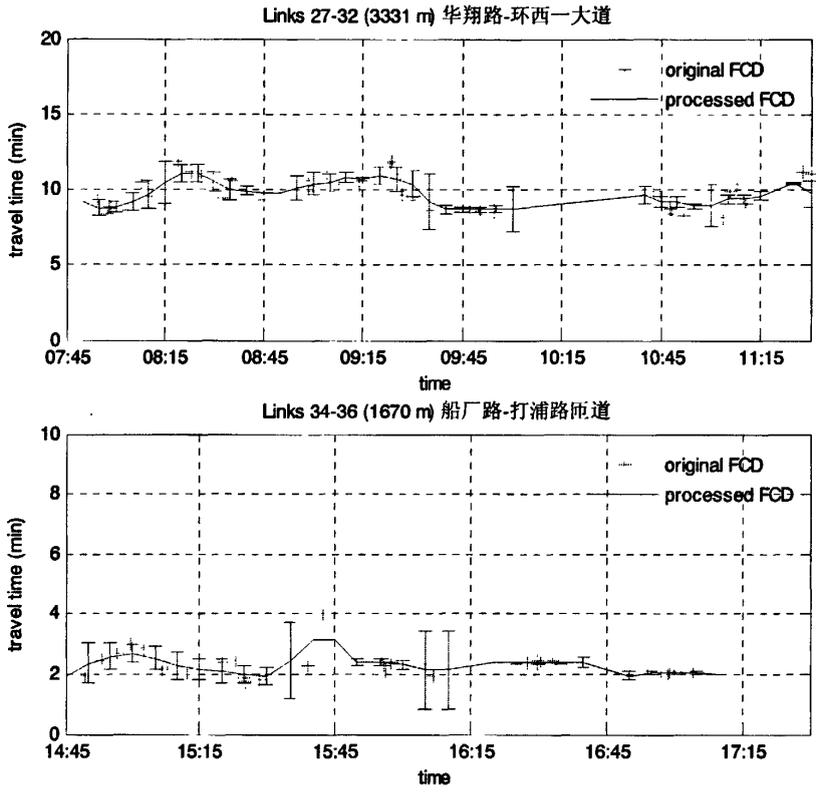




## 附录 B: 浮动车扩展路段行程时间

以 Links 1-10、Links 11-20、Links 22-26、Links 27-32、Links 34-36 为多路段计算行程时间置信区间:





### 附录 C: 断面检测器精度检验结果

如表格 4-1 所示, 雷达测速仪和 MC 在 2007 年 6 月 4 日 7:20~8:45 同时对主干道的龙华西路至天钥桥路路段进行检测, 因此使用该时间段内最外道 MC 资料和雷达测速仪数据进行比较。3 个断面雷达测速仪和 MC 实测数据比较图如下:

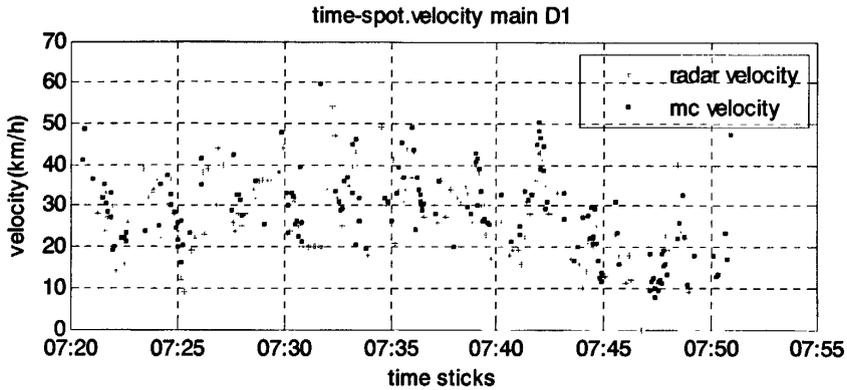


图 5-1 主干道断面 D1 雷达测速仪数据与 MC 数据比较

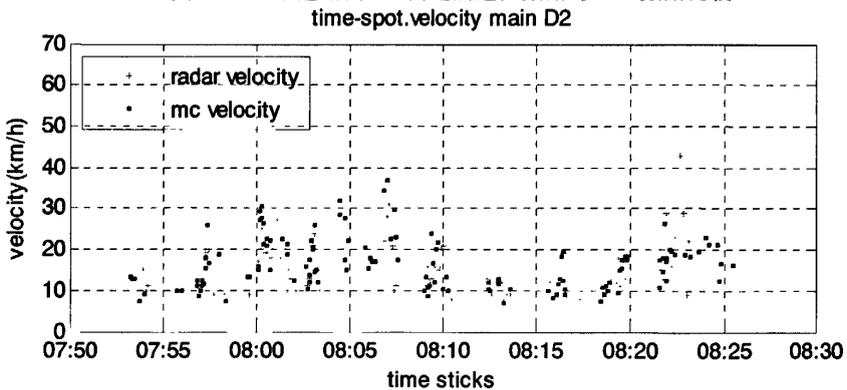


图 5-2 主干道断面 D2 雷达测速仪数据与 MC 数据比较

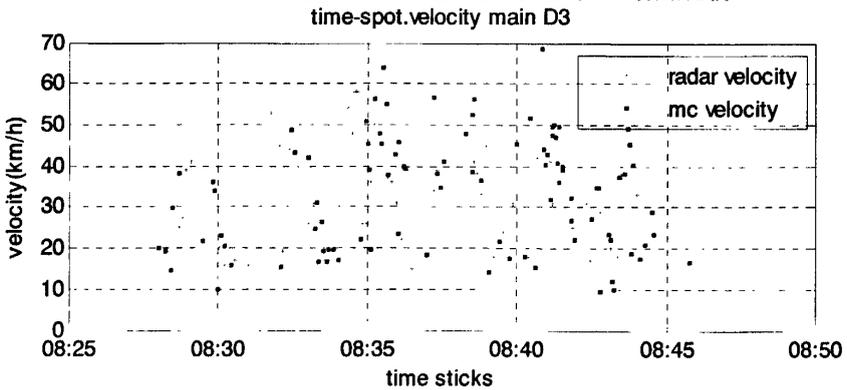


图 5-3 主干道断面 D3 雷达测速仪数据与 MC 数据比较

同样,雷达测速仪和 MC 在 6 月 6 日 7:30~9:40 同时对 318 国道的航东路至环西一大道路段进行检测, 因此使用该时间段最外侧车道 MC 数据和雷达测速仪数据进行比较。3 个断面雷达测速仪和 MC 实测数据比较图如下:

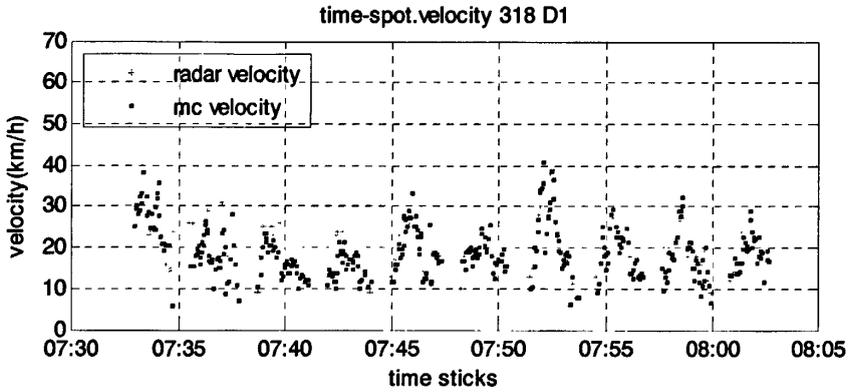


图 5-4 318 国道断面 D1 雷达测速仪数据与 MC 数据比较

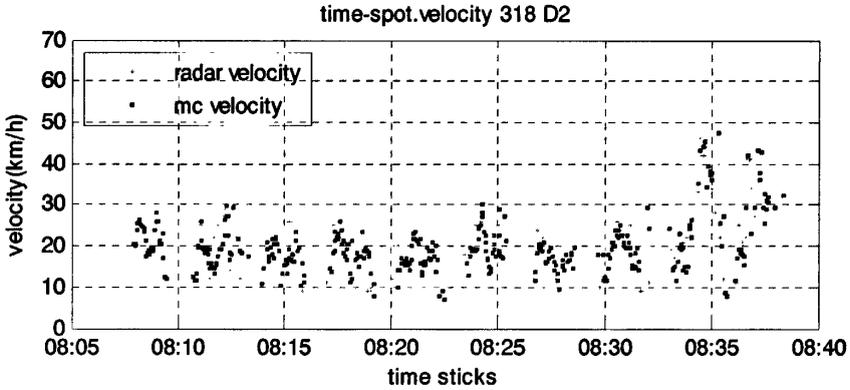


图 5-5 318 国道断面 D2 雷达测速仪数据与 MC 数据比较

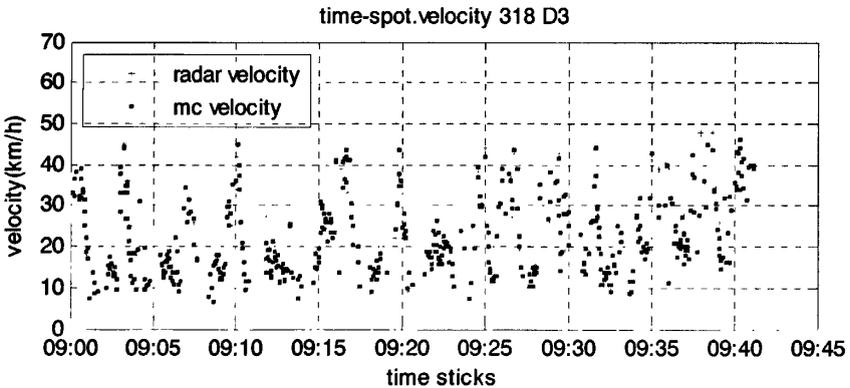


图 5-6 318 国道断面 D3 雷达测速仪数据与 MC 数据比较

根据检验算法分别找出主干道和国道各检测点样本数量相同的雷达测速仪数据和 MC 数据。对这两组数据进行比较, 并以雷达测速仪数据为准, 计算 MC 数据的相对误差。

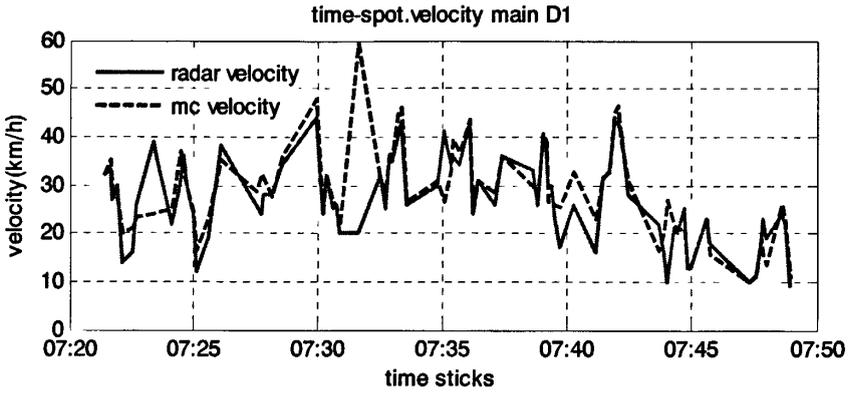


图 5-7 主干道断面 D1 雷达测速仪与 MC 样本组数据比较

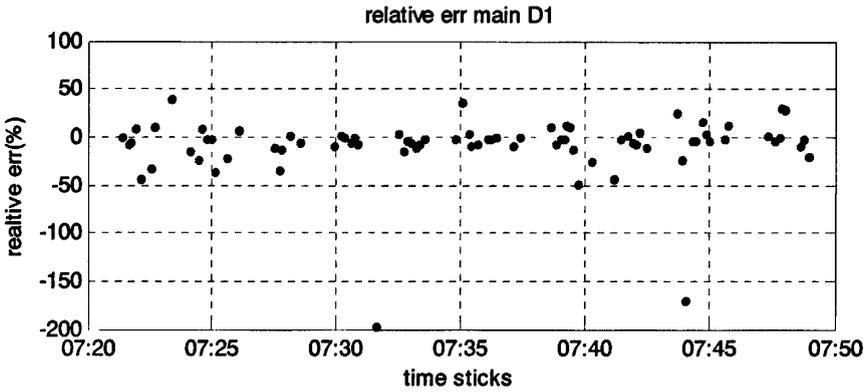


图 5-8 主干道断面 D1 MC 数据相对误差

$$\bar{\varepsilon} = -4.1404\%$$

$$MAPE_{main}^{D_1} = 11.4226\%$$

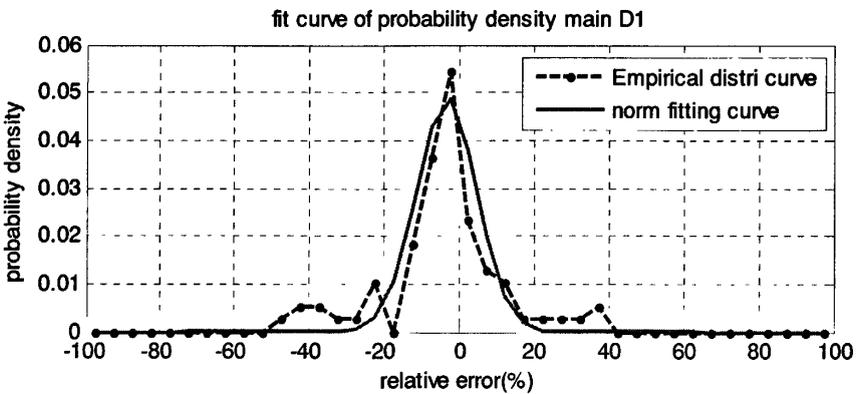


图 5-9 主干道断面 D1 相对误差分布曲线

$$\mu = -3.3085, \sigma = 8.1501;$$

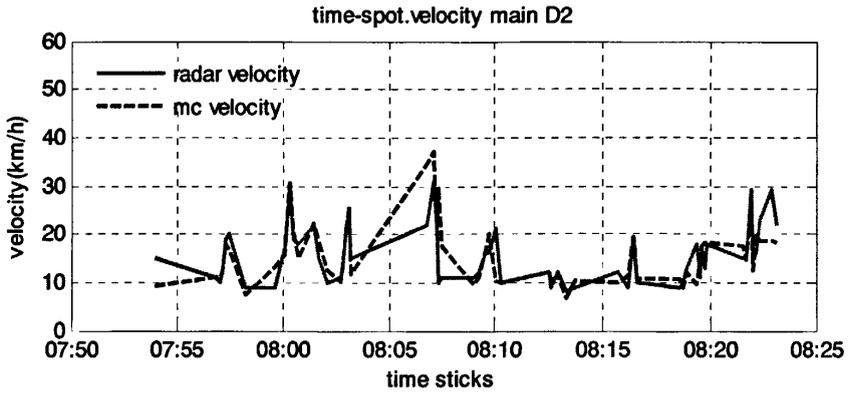


图 5-10 主干道断面 D2 雷达测速仪与 MC 样本组数据比较

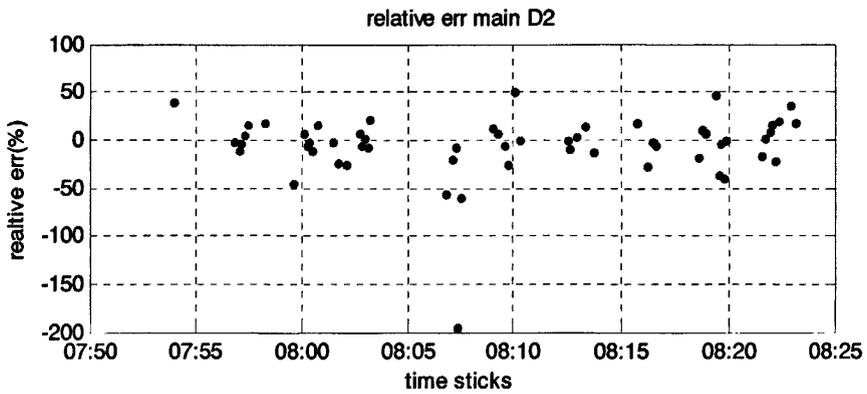


图 5-11 主干道断面 D2 MC 数据相对误差

$$\bar{\varepsilon} = -6.1694\%$$

$$MAPE_{main}^{D_2} = 11.7019\%$$

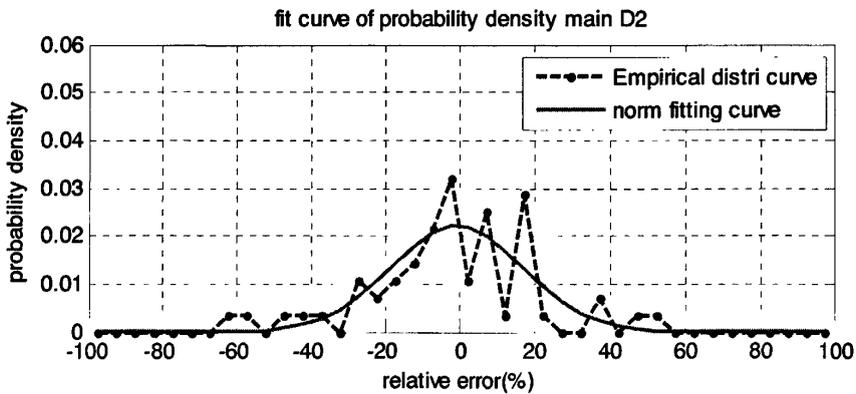


图 5-12 主干道断面 D2 相对误差分布曲线

$$\mu = -0.8363, \sigma = 17.8315;$$

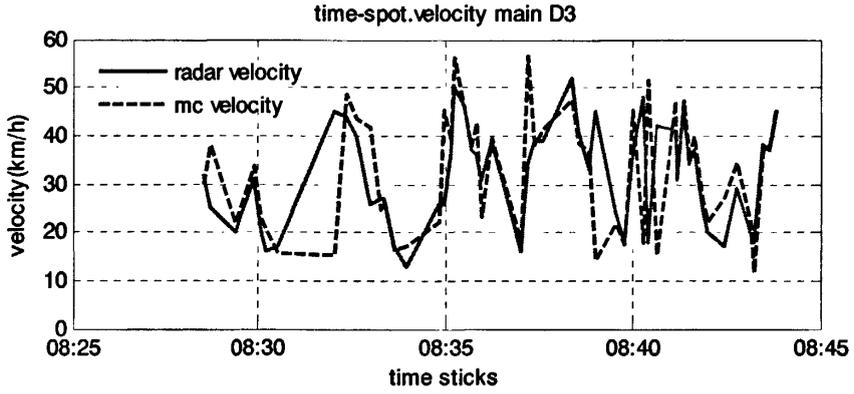


图 5-13 主干道断面 D3 雷达测速仪与 MC 样本组数据比较

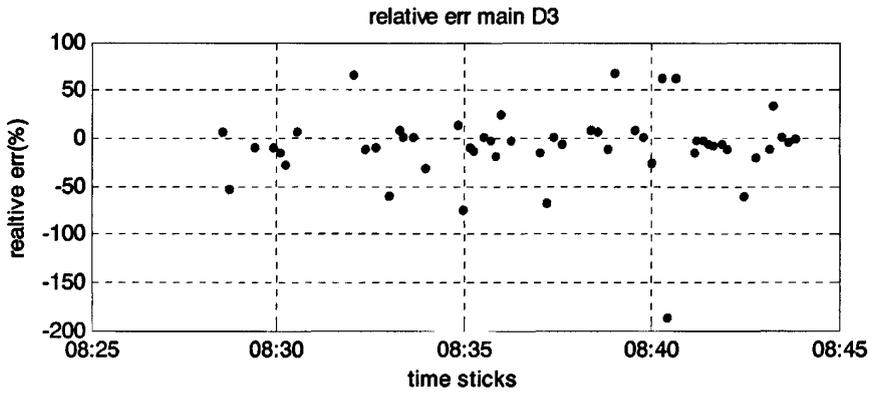


图 5-14 主干道断面 D3 MC 数据相对误差

$$\bar{\varepsilon} = -4.4328\%$$

$$MAPE_{main}^{D_3} = 10.4602\%$$

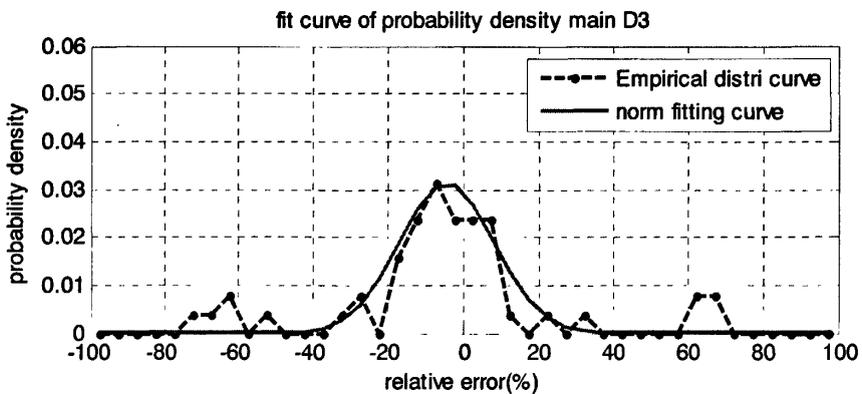


图 5-15 主干道断面 D3 相对误差分布曲线

$$\mu = -4.5015, \sigma = 12.6955;$$

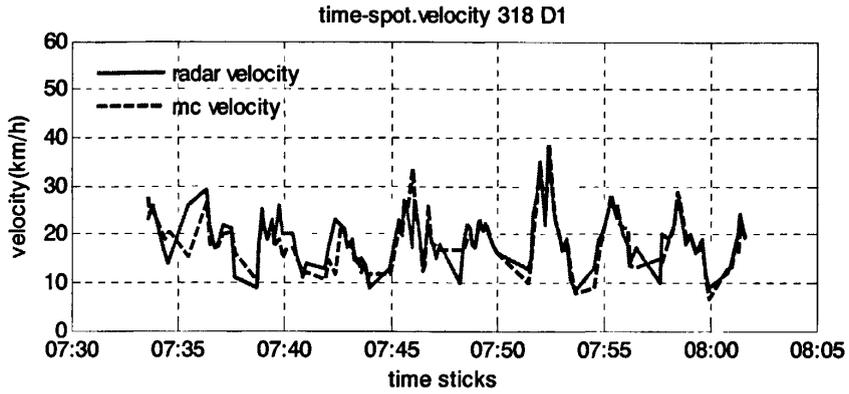


图 5-16 国道断面 D1 雷达测速仪与 MC 样本组数据比较

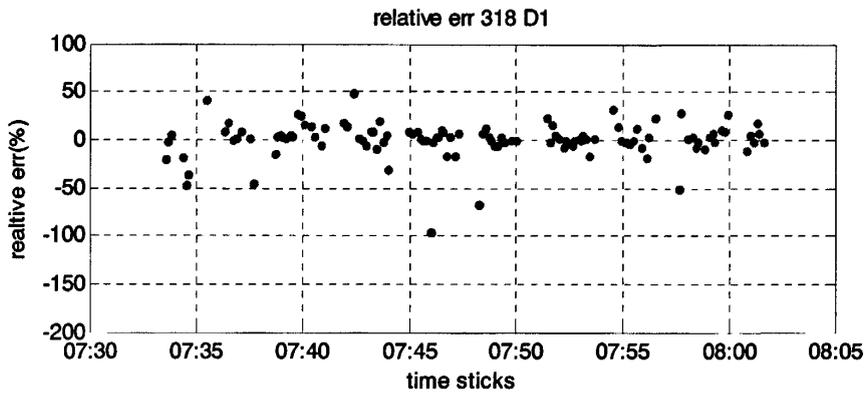


图 5-17 国道断面 D1 MC 数据相对误差

$$\bar{\varepsilon} = 0.6383\%$$

$$MAPE_{318}^{D1} = 11.2938\%$$

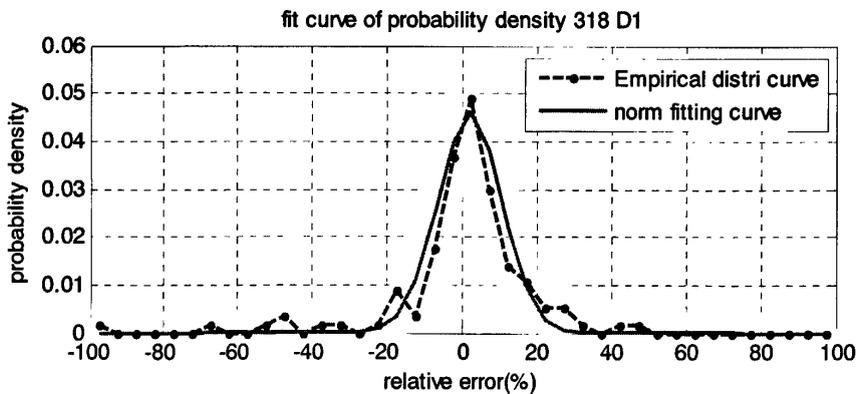


图 5-18 318 国道断面 D1 相对误差分布曲线

$$\mu = 2.0959, \sigma = 8.6378;$$

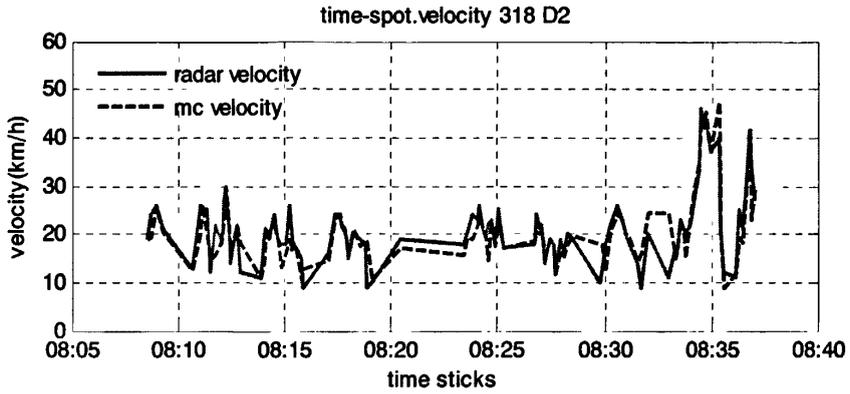


图 5-19 国道断面 D2 雷达测速仪与 MC 样本组数据比较

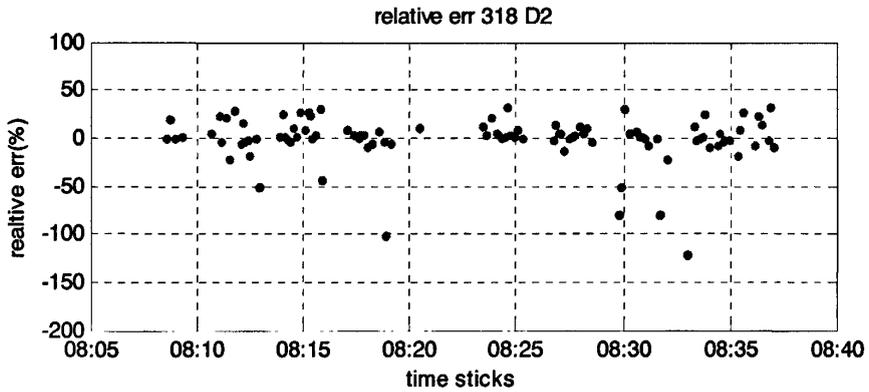


图 5-20 国道断面 D2 MC 数据相对误差

$$\bar{\varepsilon} = 0.135\%$$

$$MAPE_{318}^{D_2} = 11.6454\%$$

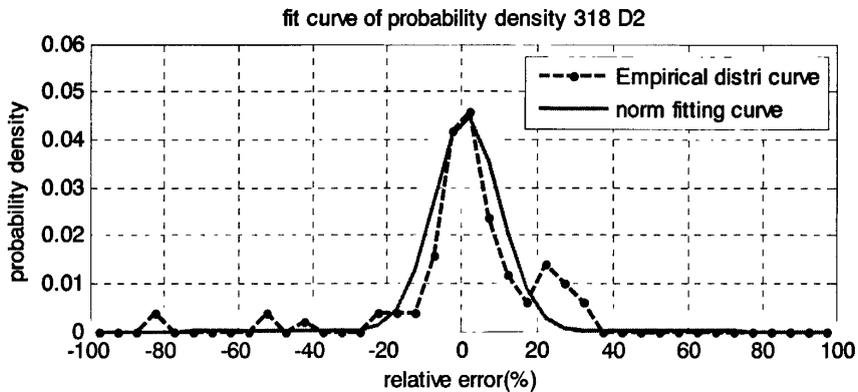


图 5-21 318 国道断面 D2 相对误差分布曲线

$$\mu = 1.4467, \sigma = 8.8406;$$

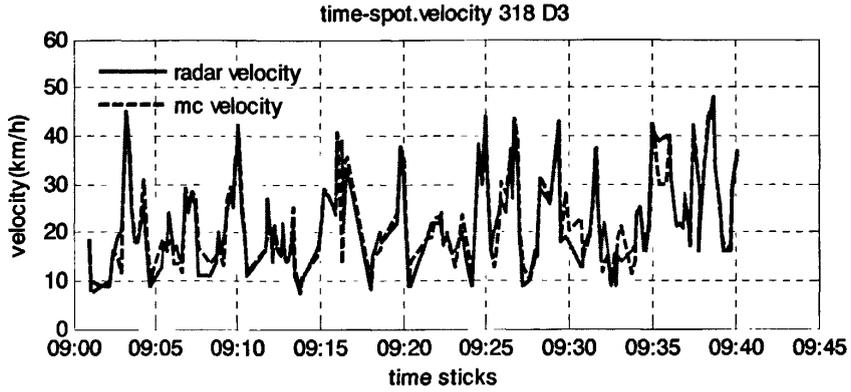


图 5-22 国道断面 D3 雷达测速仪与 MC 样本组数据比较

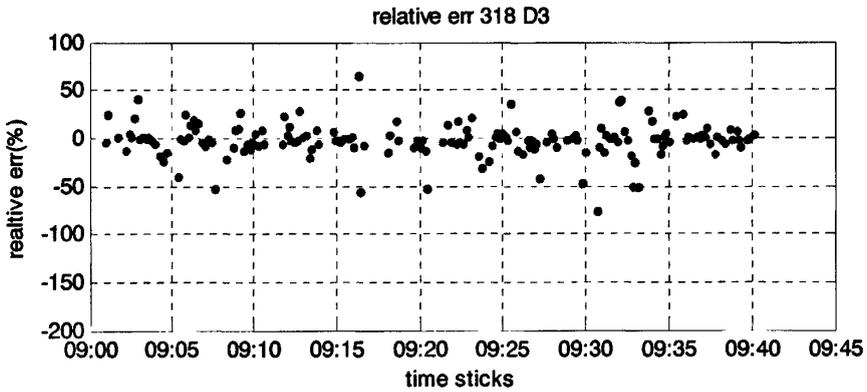


图 5-23 国道断面 D3MC 数据相对误差

$$\bar{\varepsilon} = 1.7175\%$$

$$MAPE_{318}^{D_3} = 11.1354\%$$

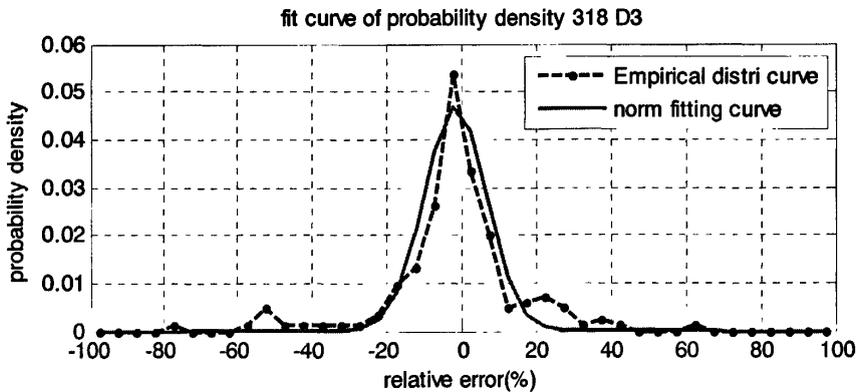


图 5-24 318 国道断面 D3 相对误差分布曲线

$$\mu = -1.7035, \sigma = 8.4624;$$

## 附录 D: 断面检测器数据估计路段行程时间结果

使用 3 种路段速度估计算法（平均法、调和平均法、自适应加权平均法）得到的路段平均行驶时间（不包含延误）如下图所示：

### 1) 主干道：

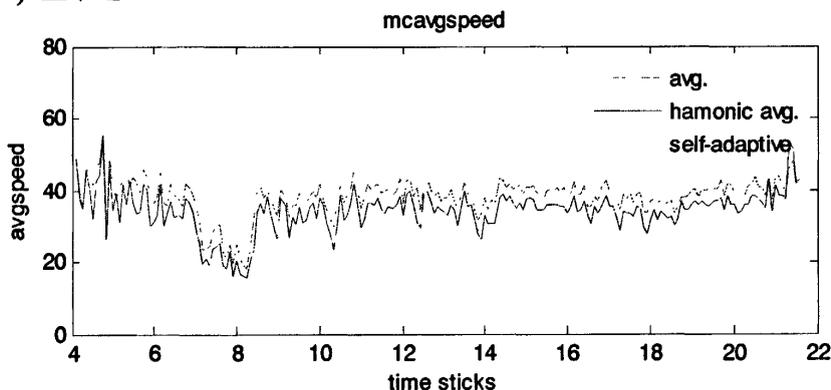


图 5-25 主干道平均法、调和平均法、自适应加权平均法路段平均行驶时间

由此可见三种方法所得到的路段速度比较近似，主干道速度曲线中，使用算法 1 得到的路段行程速度较高，调和平均算法在早高峰期的行驶速度与自适应算法基本相同，而在平峰时期的速度低于自适应算法。自适应算法除了在凌晨 5 点左右有一异常点，其他时段均显示出合理的交通流特性。

### 2) 318 国道

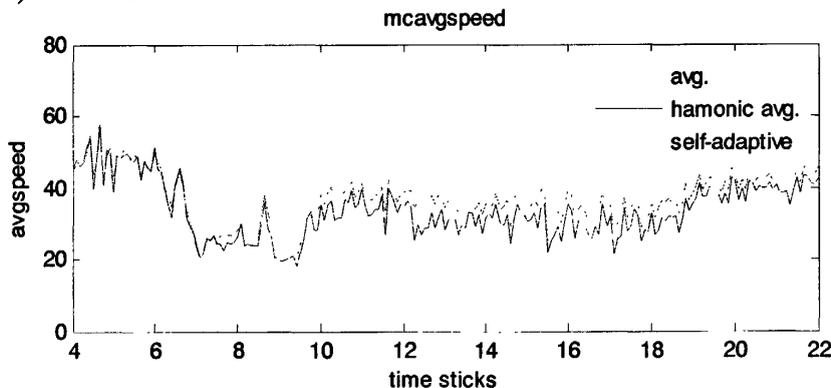
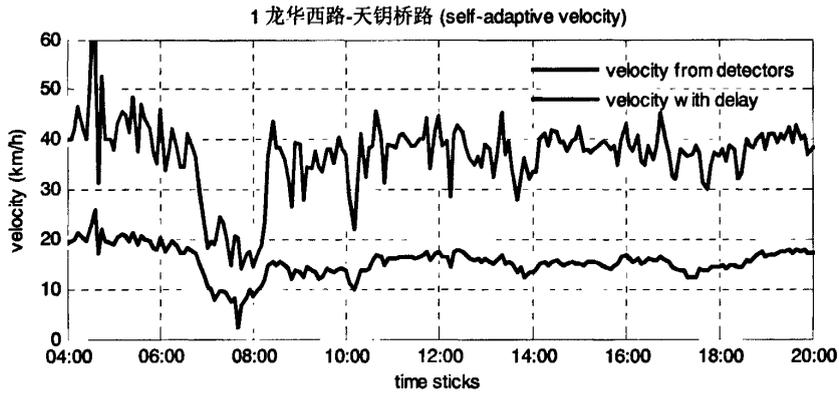
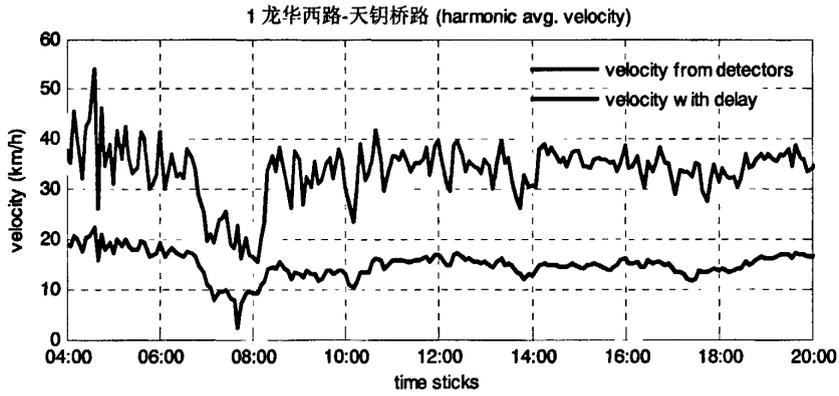
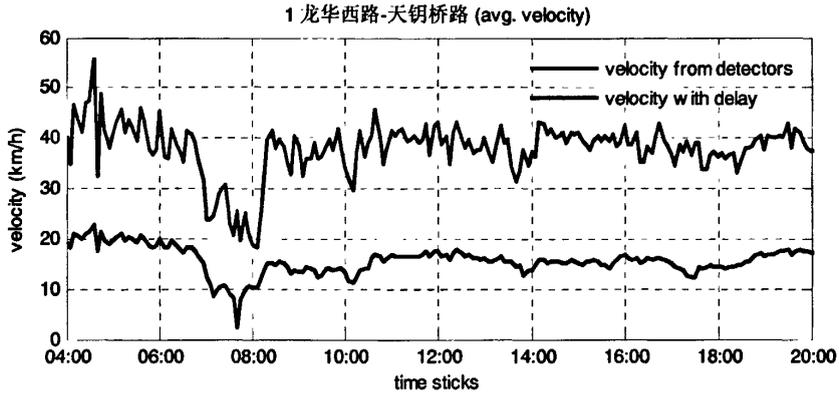


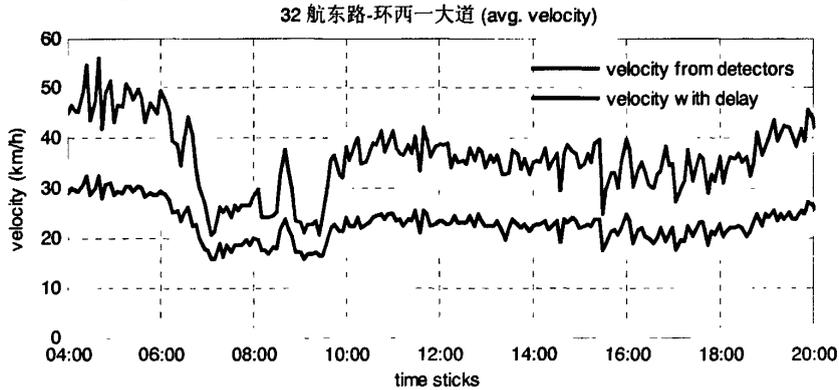
图 5-26 318 国道平均法、调和平均法、自适应加权平均法路段平均行驶时间

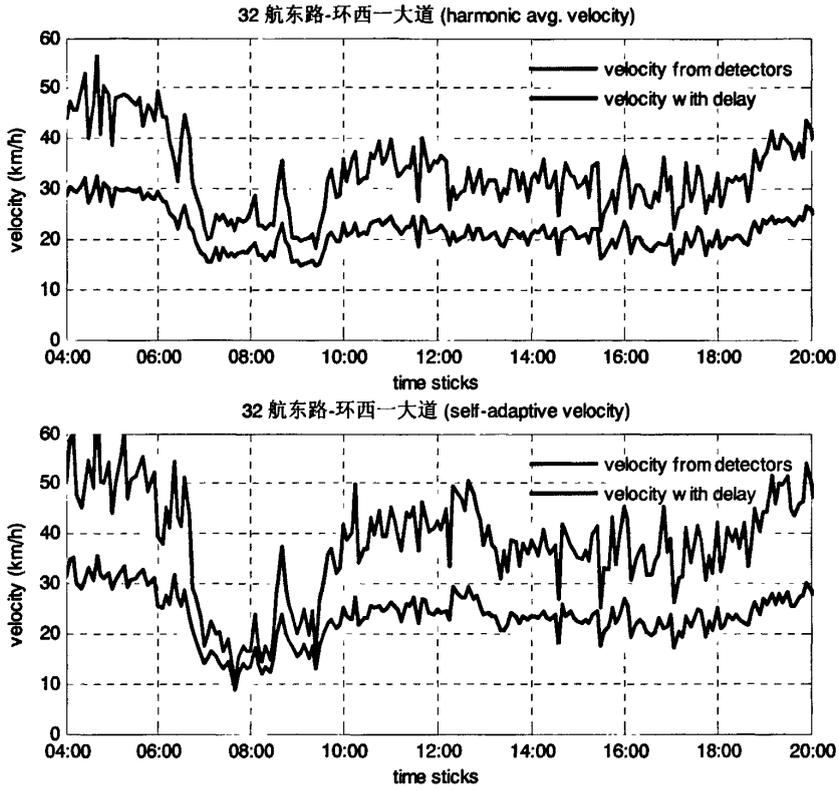
包含延误的路段平均行程时间计算：以检测断面 D2 检测到的流量为上游进入流量，路段疏散流量取 1350veh/km，得到包含延误的路段平均速度如下：

### 1) 主干路：



2) 318 国道:





由上各图可见，在相同的平均速度下，三种不同的延迟算法给出类似的结果，本文选择自适应加权平均法作为延迟计算方法。

附录 E: 信号灯路口相位调查图

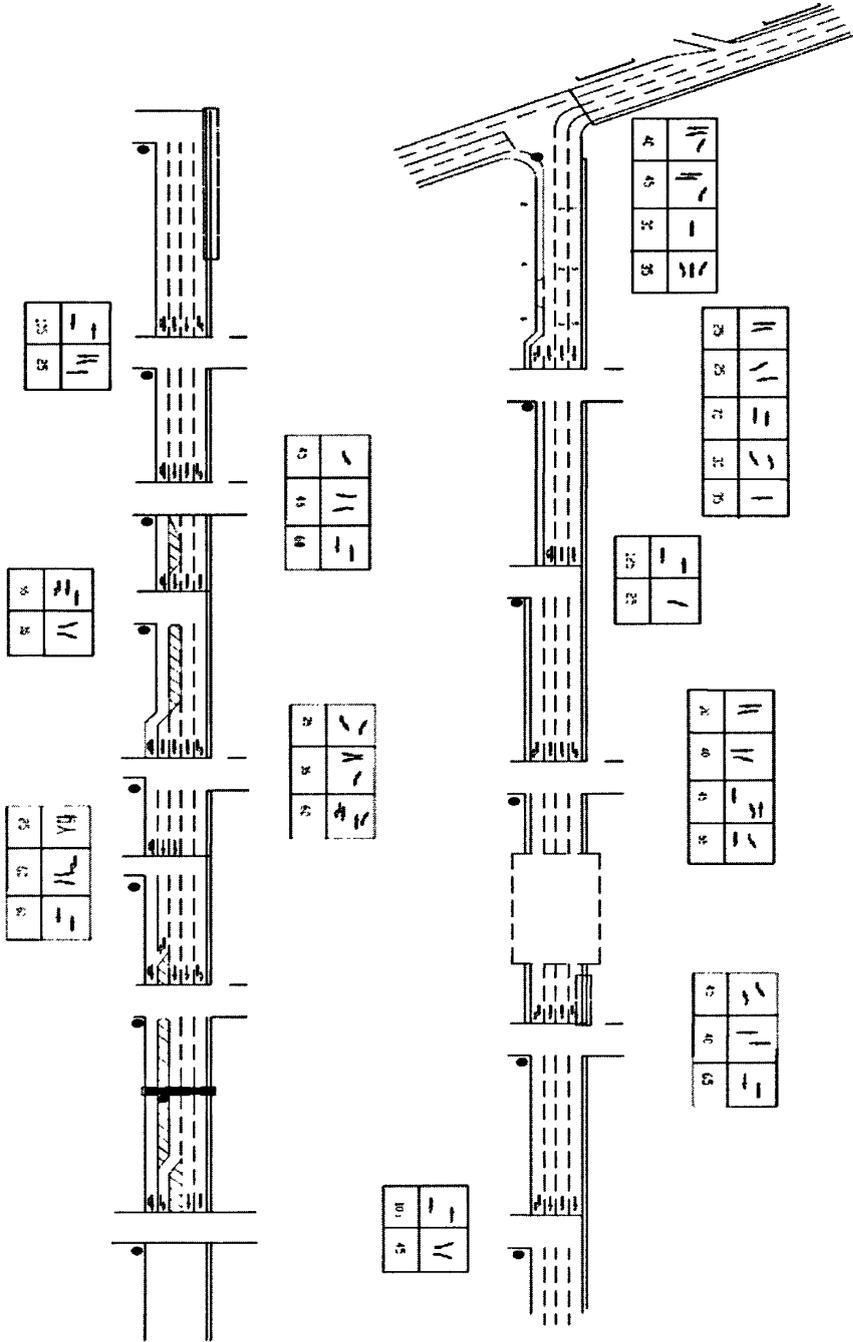


图 5-27 中山南二路相位图

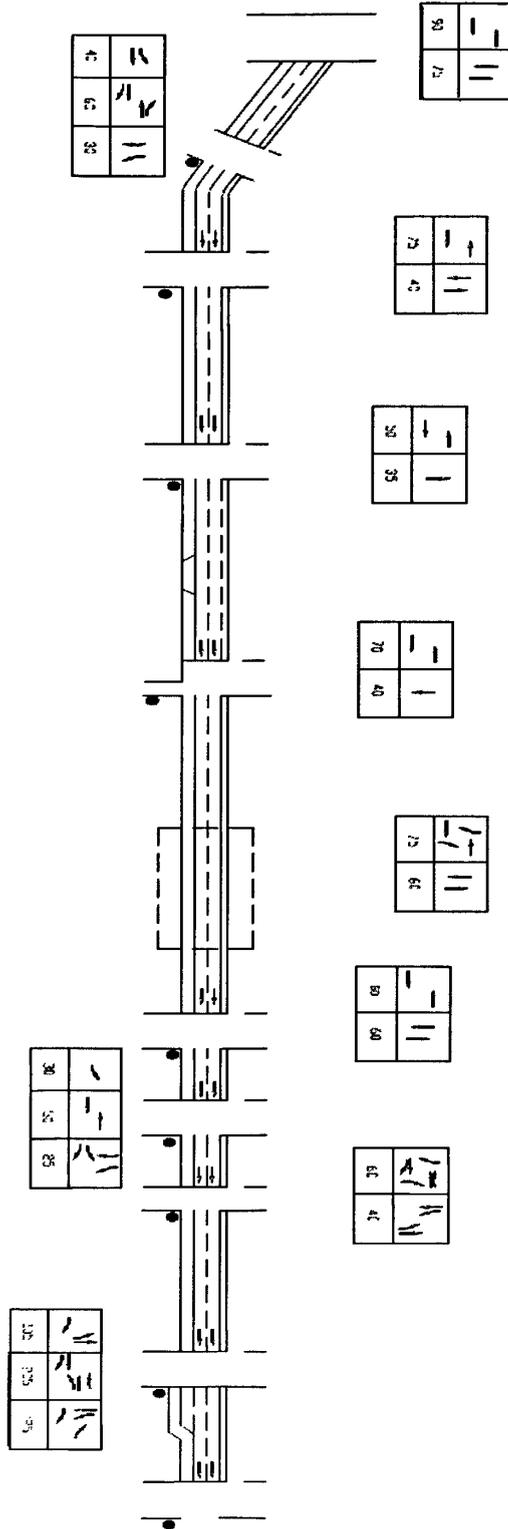


图 5-28 三一八国道相位图

## 作者简历

段学勇，男，生于1974年10月。1993年9月至1997年7月在西安电子科技大学攻读学士学位，2006年9月至今在北京交通大学系统工程与控制研究所攻读硕士学位。