Università degli studi di Padova

Facoltà di Ingegneria

Corso di Laurea Magistrale in Ingegneria Civile

Mixed time headway distributions: endogenous effects in motorways and rural two-lane two-way roads

Relatore: Dott. Ing. Riccardo Rossi
Correlatore: Prof. Dr. Ing. Bernhard Friedrich

Laureando: Federico Pascucci

Anno Accademico 2012-2013

PASCUCCI, F., *Mixed time headway distributions: endogenous effects in motorways and rural two-lane two-way roads*
Department of Structural and Transportation Engineering, University of Padua
KEYWORDS: time headway, trendless analysis, fundamental diagrams, gamma-GQM model.

Abstract

This thesis arises from the necessity to increase the research about time headway distribution, a topic study essential in many traffic engineering applications but characterized in Italy by a lack of information.

The interest in this topic is due to a research project of the Transportation Laboratory of the University of Padova in order to establish an information system of traffic flow phenomena in the Veneto area, with the aim to specify, calibrate and validate mathematical models of observed phenomena starting from traffic data recordings. Within the research project, this thesis seeks to identify TH probability density function for two-lane two-way road segments in different traffic situations. The interest in this field is due to its implications for driving simulator experiments with regard to vehicle generation.

In order to obtain the best results, it was necessary to enhance the methodology for analysis of headway data taking into consideration two contributions from scholarly literature. The first is the method of identifying sub-samples for use in statistical analysis. Here we refer to a trend test proposed by Luttinen in 1996 which aims to reach the stationarity of the headway variable. The second concerns the process of calibration of the probability density functions, through the implementation of the gamma-GQM model (especially suitable for representing headway distribution in all traffic context) and belonging to the mixed models.

Once the processing methodology was chosen, a time headway analysis was carried out as the endogenous traffic parameters changed - flow rate and flow composition (percentage of heavy vehicles in this case) - considered as affecting time headway distributions. Our attention focused on two-lane two-way rural roads in Northern Italy with data recorded both by inductive loops and radar sensors and with the aim to create a picture of the relationship between endogenous traffic conditions and headway distributions.

We completed a study of headways with data coming from A2 motorway Berlin-Dortmund in Germany and available from the *Institut für Verkehr und Stadbauwesen* of T.U. Braunschweig. This was done in order to compare the results with those obtained by other authors (Ha, Aron, Cohen 2011), who used a similar approach on the same kind of road, to test the affinity of the results and validate the procedure.

The originality of this thesis is in the application of a trendless analysis and a powerful mixed model - on the two-lane two-way roads, which are often neglected in headway literature in favor of motorways.

The robustness of the methodology combined with its capacity to reveal useful in-

formation about the traffic state permit us to assert that an adequate study of headway variable must follow this approach.

Acknowledgment

First and foremost I offer my sincerest gratitude to my supervisor, Dr. Ing. Riccardo Rossi, who has supported me throughout my thesis with his knowledge and his advice. His suggestions and inputs (all rigorously via Skype) allowed me to develop this work, in total freedom, under his watchful eye however; for his reason, I would like to thank him for his confidence in my work.

I would like to express my deep gratitude to Prof. Bernhard Friedrich who welcomed me at the IVS of the Technical University of Braunschweig. I am grateful for the great respect he has always demonstrated to me and for support with my research.

I would like to thank the guys of the IVS who even indirectly helped me, Steffen, Luciano, Elke and all the others. Thanks to Anke for the valuable conversations in german. Special thanks and a heartfelt hug to Jannis for his assistance and for showing me two essential personal qualities, which he combines masterfully: seriousness and professionalism (typically german) and expressivity (typically italian).

My sincere thanks also goes to Prof. Claudio Meneguzzer, who has never withheld to me his assistance, from my first Erasmus aspiration, to the much realer flowcharts.

A big “thank you” to Paul Loffler for his helpfulness and his English editorial work. I hope he will even edit this very sentence.

I would like to thank my mother and father for their continuous love, their supports in my decisions and their blind faith in me; my sister Giulia and my friend Serena for their feminine point of view which forced me to review various kind of personal prejudices over the years.

Last but not the least, my crony Federico, to whom I dedicate this thesis because it is similar to our friendship altogether: much time to develop it, and a high prize if I should sell it. I’m obviously joking: no one would ever buy this thesis.

Contents

Nomenclature

EDF Empirical Density Function

FR Flow Rate

GQM Generalized Queuing Model

IID Indipendent and Identically Distributed

IW Inverse Weibull

IW Inverse Weibull

K-S Kolmogorov-Smirnov

kdf Kernel density function

LOS Level Of Service

PDF Probability distribution function

pdf Probability density function

SL Significance level

SMS Space Mean Speed

TH Time Headway

TLTW Two-lane two-way

Chapter 1

Introduction

Time headway (TH) is a measure of the temporal space between two vehicles, and is defined as the elapsed time between the arrival of the leading vehicle and the following vehicle at a designated test point. Since it is measured from the front bumper to the front bumper, it is thereby the sum of the time used by a vehicle to pass the observation point (*occupancy time* or *detector clearance time*) and the time interval (*gap*) to the arrival of the next vehicle.

In the traffic flow theory TH is more important than gap because it holds a direct relationship with flow rate (FR), i.e. in a sample of n observations the reciprocal of mean TH equals the traffic volume[1] q counted during the same period of time:

$$\bar{h} = \frac{\sum_{i=1}^{n} t_i}{n} = \frac{1}{q} \tag{1.1}$$

The connection between TH and FR is similar to that between space headway (or *spacing*) and density, but "FR is more meaningful and more easily measured by practicing traffic engineers rather than density" [17]: for this reason working with TH rather than spacing is preferred. Another motivation is that TH is not as sensitive to vehicle speeds as the space headway [17]: the driver of a constrained vehicle adjusts his spacing depending on the speed of his predecessor basing on safety considerations, while headway is straight the result of the pair speed-velocity.

1.1 Importance of headway studies

Knowledge about headway is fundamental in traffic flow theory and for this reason it has been studied for many years. At first, thanks to his direct relation with traffic volume, headway was used to estimate the traffic capacity of a straight roadway section by processing short TH distributions. The latter have also been associated with the level of service (LOS): in the 1985 Highway Capacity Manual (HCM 1985) the LOS on two-lane

[1] In this work we mean the variable *traffic volume* as the measured number of vehicle in any time period, that always needs be specified. On the other hand, *flow rate* (FR) is the number of vehicle passing the cross section in a time period fixed at 1 hour.

rural highways is approximated by the percentage of vehicle in platoon, which correspond to the proportion of headways less than five seconds.

Nowadays accurate modeling and analysis of vehicle headway distribution helps traffic engineers to maximize roadway capacity and minimize vehicle delays. At unsignalized intersections and roundabouts, the TH distribution determines the opportunity for merging and crossing. As explained by Sullivan and Troutbeck [22]:

“Vehicle proceed through the intersection as determined by the vehicle priority rules. Here *gap acceptance* methods, which use the distribution of headway in the major traffic stream, determine the intersection performance and capacity. Gap acceptance theory states that vehicle can proceed into an intersection during any time-gap in the opposing traffic streaming greater than the critical gap”.

The knowledge of the arrival time of vehicle at signalized intersection could be used to minimize time delay and maximize the network capacity by adjusting or coordinating signal timing plans of the local site. Furthermore, as suggested by Luttinen, “in vehicle actuated traffic signals the control, the extension time in particular, is very sensitive to the arrival pattern” [17]. Therefore TH distribution is quite relevant in the study of optimal traffic signal control.

Knowledge about TH is also useful to generate vehicle in microscopic traffic simulation models, many of which have been developed to solve different traffic problems for interrupted or uninterrupted facilities. As Jang [16] explain:

“A key component determining the performance of such models is the generation of vehicle arrival times as input to the simulations. Hence, traffic simulation researchers have devoted considerable effort to developing theoretical models that adequately describe actual headway distributions”.

In recent years, other fields of interest have also been represented by safety analysis, with regard to the minimum TH that must be obeyed in case the leading vehicle suddenly stops, by ITS applications and by driving simulation experiments in reference to vehicles generation.

1.2 Historical background of interest

In the field of TH analysis, many authors gave their contribution on the basis of this scheme:

- development of a headway model on the basis of some assumption
- analysis of some statistical properties of TH
- calibration and validation of the model

Starting from Adam, who formulated the idea of arrival as a Poisson process and assumed TH negative exponentially distributed, several authors have proposed many more sophisticated models. In particular, the assumption of two different distributions for following and non-following vehicles paved the way to combined and, later, mixed models, which are supposed to be the best choice for TH modeling. We refer to the work of Buckley [6] who introduced the Semi-Poisson model, and those of Cowan [7] and Branston [3] who derived the GQM model starting from the queuing theory. After these works, headway science mainly focused on finding the proper simple model to apply to follower vehicles.

The aim has always been double: to find the proper TH model in relation to the traffic context, which could provide best performances in terms of fitting and computational effort for the calibration at first; then to analyze the variation of TH distribution at different traffic levels and varying endogenous conditions. Luttinen [17] made a major contribution to headway studies by summarizing all existing TH models and giving primary importance to the statistical analysis of data, which is the necessary step to the choice of the appropriate model.

We keep under primary attention three recent branches of work developed at:

- Delft University of Technology, by Hoogendorn, Bovy and Botma [14, 15], who applied the GQM model on two-lane rural roads in order to find differences in vehicle-classes and to develop a mixed-vehicle-type model. In addition, they analyzed the possibility of capacity estimation through TH models.
- University Paris-Est, by Ha, Aron and Cohen [10, 11, 12], who analyzed a large number of TH models on A6 motorway data in France with the aim to find the distribution which performs best in terms of goodness-of-fit and computational effort and paying close attention to the proper estimation method. Their attention focused also on the variation of the parameters of the gamma-GQM function, which appeared to be the best choice through all theoretical models.
- University of Padova, by Rossi and Gastaldi [21], who applied simple models on two-lane two-way roads with the aim to find the most suitable and to analyze the variation of the shape of the distribution as endogenous traffic conditions change.

1.3 Purpose of this study

The purpose of this study is to improve the methodology of analysis of the TH variable, in order to supply a valid instrument for analysis of TH distributions in different contexts and traffic conditions. The attention is focused on two main issues:

- the individuation of a method to select sub-samples from a the entire set of headway data following the Luttinen method for investigation of stationary intervals, since the TH variable can be inspected only with stationary data.
- the need for a good TH model; the studies mentioned in section 1.2 represent a valid contribution for the choice of an appropriate distribution.

Once the methodology is defined, we aim to find variation of TH distribution varying endogenous parameters which include the flow rate and the flow composition (percentage of heavy vehicles). This application is particularly aimed at the creation of an information system of traffic flow phenomena on rural roads, which include the knowledge of TH probability density function in relation to different traffic situation.

We will at first apply our methodology on data coming from the A2 motorway Berlin-Dortmund in Germany, in order to find some similarities with the work done by the research group of the University Paris-Est, which studied TH distributions on A6 motorway in France, by using roughly the same methodology. This step will allow us to test our methodology and to compare the results.

The final step of this work, and indeed the main purpose, is to discover if endogenous parameters are significant in our two-lane two-way rural roads and, if so, how they affect the shape of the distribution on a cartesian plane.

Finally, the results will show whether a further investigation is needed.

Chapter 2

Headway variable

The basic assumption on this research, and generally on all headway studies, is to consider TH as a random variable. This step is necessary to pave the way to the *theory of probability* and, in our case, to a trend analysis focused on the identification of stationary sub-samples.

2.1 Road traffic as a random process

Road traffic has been firstly considered as a random[1] process in 1936 thanks to the publication of Adam called indeed “Road Traffic Considered as a Random Series”. In this work, he compared traffic as a distribution of points along a line, in two different ways:

a) each point indicating the position of a vehicle at a given point, and the line representing a length of road;

b) each point indicating the moment at which vehicles passed a given place, and the line representing a period of time.

The variable under consideration, whether it is the position in space or the moment in time, appears to be very irregular. Distribution showing similar irregularities, explain Adam, are studied in the *theory of probability* under the name of *random distributions of points* (a) or *random series of events* (b). We remind that a series of events is defined as being *random* when:

1. each event is completely independent of any other event;
2. equal intervals of time are equally likely to contain equal numbers of events.

Similar conditions, expressed in terms of space instead of time, define a random distribution of points. In the statistical analysis of traffic flow, option (b) is chosen because the assumption of a random series leads to results of more practical utility.

Adam wrote that “freely-flowing traffic is found to conform so well to the distribution given by a random series” [1], and that departures from this process are produced by many factors, through which the increase in the traffic, difficulty in passing other vehicles and saturation.

[1]The term *random* is considered synonymous of *stochastic.*

Theory of Probability provides the basis for the mathematical description of random series. The essential element of this theory is the concept of a probability function which gives the mathematical probability that a random variable X will assume a specific value x:

$$f(x)dx = probability(x < X \leq x + dx) \tag{2.1}$$

where *f(x)* is the probability density function (pdf). Probability may be expressed even as a probability distribution function (PDF):

$$F(x) = probability(X \leq x) \tag{2.2}$$

that gives the probability that a random variable X is not greater than some given value x. The PDF may be defined in terms of the pdf by:

$$F(x) = \int_0^x f(k)dk \tag{2.3}$$

and it follows that:

$$\int_0^\infty f(x)dx = 1 \tag{2.4}$$

Continuous distribution function are used in traffic flow theory to represent random variables, which can take on a continuum of values, such as headways, speeds and quantities representing driver behavior. The objective is to derive a distribution which adequately represents the behavior of the random variable, and the choice of the pdf is often a matter of mathematical convenience [18].

2.2 Headway as a random variable

Probability theory applied to traffic flow usually concern the probability of vehicle passage, or arrival, in a specific time period (i.e. the random variable is the number of vehicle passing, or arriving, at a particular point). However, as suggested by McLean [18], "for some derivations, it is more convenient to think it in terms of the passage of the time space, or *gap*, between vehicles, and the random variable may be the number of gaps greater than a specific value". Similarly *headway* can be used as a random variable, since it is composed by the sum of *gap* and *occupancy time.*

This way we moved from the concept of *number of events* in an interval, to that of *interval of time* passed between two events, which is the headway. In substance, headway assumed as a random variable is used to study road traffic, considered as a random process.

In this case, calling headway T, expressions 2.1 and 2.2 may be re-written as follows:

$$f(t)dt = \Pr(t < T < t + dt) \qquad [pdf] \tag{2.5}$$

$$F(t) = \Pr(T \leq t) \qquad [PDF] \tag{2.6}$$

2.3 Shape of the TH distribution

Starting from a sample of headways in the arrival order (Fig. 2.1), the pdf is usually estimated by the *histogram method*, that consist in partitioning the range of data into class intervals, and counting the number of observation belonging to different bins.

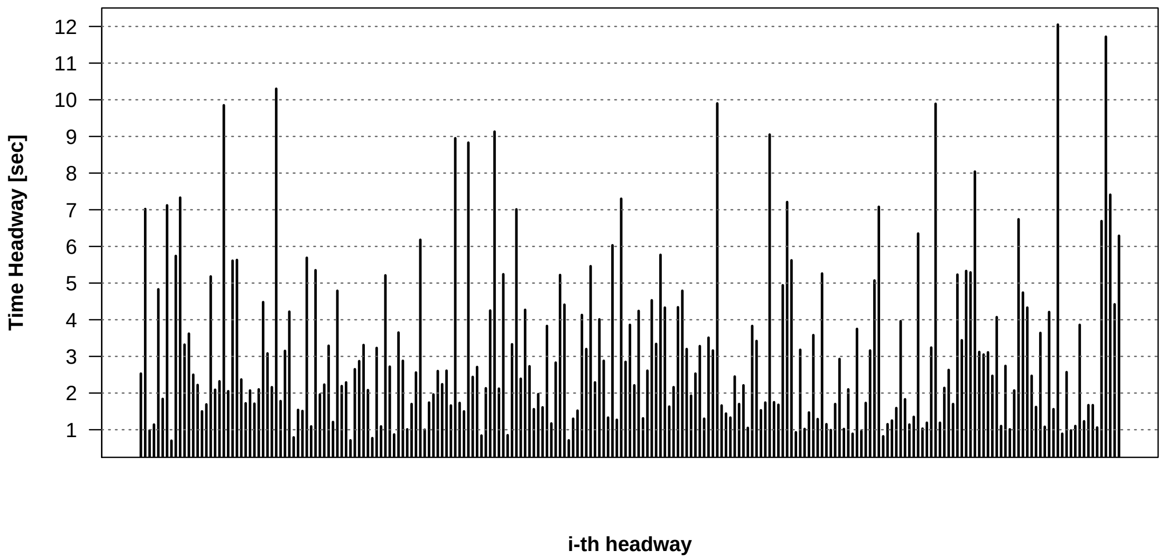

Figure 2.1: Headway values in arrival order, n=225

The number of elements in each bin is then divided by the number of total observation n and the bin width h, in order that the total area under the histogram is equal to one. The method is very straightforward but the value of h has a considerable effect on the shape of the histogram (Fig. 2.2), especially at low values of time headways. Another drawback is that the discontinuity of the histogram doesn't allow to lead a statistical analysis of headway variable - i.e. the mode or the peak heigh cannot be calculated precisely.

To overcome these problems, the analysis of empirical data in this thesis is managed by a Kernel density estimation that assures the properties of smoothness and continuity. The kernel method produces a continuos estimate of the pdf, avoiding the pitfalls of the histogram method. The kernel density estimator with kernel $K(x)$ at point t is defined as:

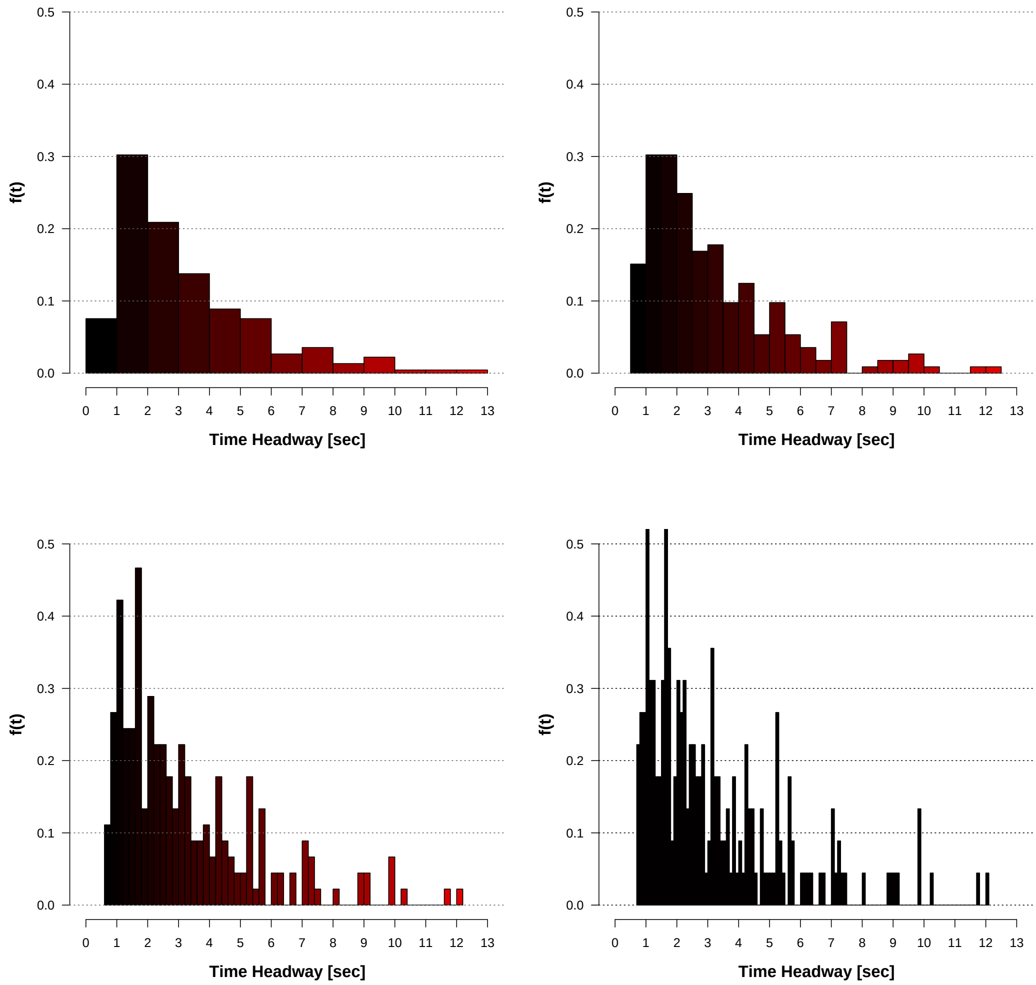

Figure 2.2: Histograms with different bin width (1, 0.5, 0.2, 0.1 [sec])

$$f_n(t) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{t-t_i}{h}) \tag{2.7}$$

where h is the smoothing parameter. The estimates are based on the *Epanechnikov kernel*, which is supposed to be the most efficient kernel function:

$$K(x) = \begin{cases} \frac{3}{4\sqrt{5}}(1 - 0.2\,x^2), & \text{if } \mid x \mid < \sqrt{5} \\ 0, & \text{otherwise.} \end{cases} \tag{2.8}$$

The benefits of the kernel density function (kdf) are quite evident in Fig. 2.3, where it is put in comparison with the histogram method. The shape of the density function now in very clear, the curve is “unimodal, bell-shaped and skewed to the right”[17].

While statistical values like mean and variance could be directly obtained by the sample of headway, only through the kdf two measures of particular interest in headway may be discovered precisely:

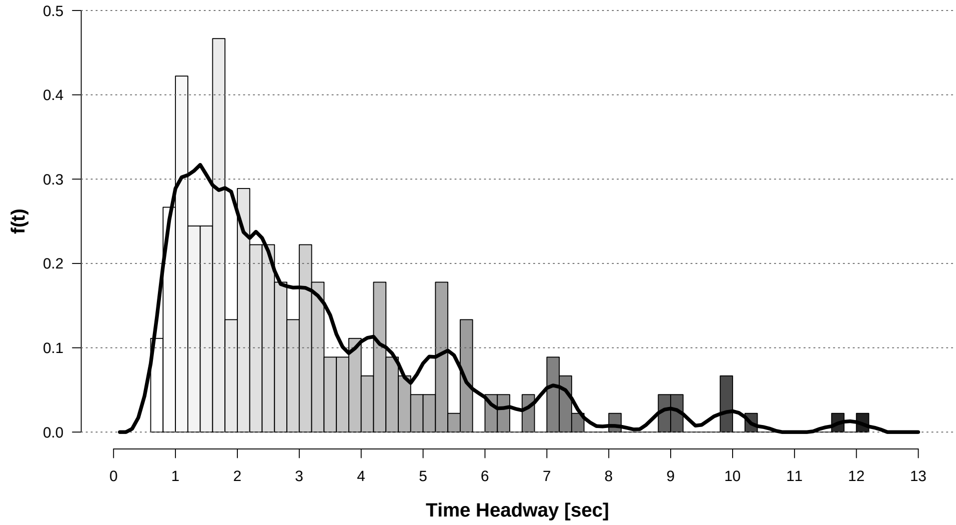

Figure 2.3: Kernel density function in comparison with histogram method

mode is the point where density reaches its maximum, so it is an approximation of the most frequent value in the distribution. It can be seen as an estimate of the headway that most drivers select when they are following the vehicle ahead, so it is a measure of typical platoon behavior (Summala and Vierimaa [23])

peak height is the measure of the frequency at mode.

The value of the mode may differ from the type of road, the type of vehicle, the type of drivers and weather conditions; his inverse value may be considered as a crude, and extremely theoretical, estimation of road capacity. On the other hand, peak height is extremely dependent from the traffic state, i.e. it increases as the flow rate arises.

Chapter 3

Trendless analysis

Real properties of headways can be inspected only with trendless data. The statistical moments of a sample of TH and the parameters of the distribution must be representative of a specific state; the trend analysis is the proper instrument to identify all the non-random variation which must be eliminated from the data set, otherwise the parameters change as a function of time. For this reason, we look for intervals where the headway variable is stationary[1].

Before giving the mathematical definition of *stationary random process*, we take a step back and consider headway as the interval between consecutive events, and the arrival of vehicle as the event itself. By doing so, we could define a *stationary series of event* through the definition given by Cox and Lewis [8], which consist of three points:

1. the distribution of the number of events in a fixed interval $[t_1', t_1'']$ is invariant under translation, i.e. is the same for $[t_1' + h, t_1'' + h]$ for all h;
2. the joint distribution of the number of events in fixed intervals $[t_1', t_1'']$,$[t_2', t_2'']$ is invariant under translation, i.e. is the same for $[t_1' + h, t_1'' + h]$,$[t_2' + h, t_2'' + h]$ for all h;
3. quite generally the same invariance property must hold for the joint distribution of the number of events in a set of k fixed intervals, for all k=1,2,... .

Notice that 1 and 2 are just special cases of 3, and 1 is a special case of 2.

In terms of random intervals $\{X_i\}$ between successive events, the *stationarity of a series of events* implies that all the X_i have the same marginal distribution function. The sequence $\{X_i\}$ is then a *stationary sequence of random variable* (or *stationary random process*) if the joint distribution of any k of the intervals between events, for all k=1,2,... , is invariant under a translation along the time axis.

It is observed that, since we evaluate variables along the time axis, we are referring to the concept of *stationary in time*.

The same concept could be expressed by the definition of McLean [18]:

“If the probability structure measured at any fixed point, or over any fixed length, is invariant with time, the system is said to be stationarity in time, and it can be represented by a stationarity random process. The terms uniform and traffic steady are sometimes

[1]The trendless is considered to be equivalent to stationarity

used in traffic flow literature to describe this characteristic. [...]

Analogous to *stationarity in time* concept, *stationarity in space* can be defined:

"A system is said to be *stationary in space* when the probability structure observed over time at any point is the same as that observed at any other point. If the process is also stationarity in time, it means that the probability structure observed over time at any fixed point is the same as the probability structure that would be observed by sampling a number of points at a fixed instant of time (ergodic hypothesis)" [18]

Saying that "distribution is invariant" - terms used by both authors - means that all the statistical moments are equal and independent from time, mean included. This definition is very restrictive, and gives to the statement the diction of strong stationary process. Since this condition "would be quite impracticable to verify with empirical data" [8], we can assume a *weak stationarity*, which requires that only that the first and the second-order properties of the process are invariant under a time shift. This is confirmed by McLean: "[...] the prediction should take the form of a probability distribution but, for practical purpose, mean and variance will often suffice" [18].

Furthermore we specify that, thanks to the relation:

$$q = \frac{1}{\bar{h}} \tag{3.1}$$

we can assert that in stationary condition flow rate is constant because it can be expressed as the opposite of the mean headway (which is a the first-order moment of the variable). As pointed out by Cox and Lewis [8], and assumed by Luttinen [17], it is possible that the flow rate might remain constant, even thought the detailed structure of the process changes: this possibility is however ignored by both the authors, and we can assume that constant flow rate condition is equal to stationarity.

The fact remains that the concept of *stationarity of a random variable* is different from *stationarity in traffic flow theory*, that states: "Traffic flow is said to be stationary when traffic volume q is invariant with the road section and density k is invariant with time". Since the definition given by Cox and Lewis about a *stationary series of event* concern only the time variable, the same definition cannot entail stationarity as intended by traffic flow theory.

3.1 Trend tests

Trend tests are statistical instruments for detecting trend of a measured variable, and may be used in our field with the aim to detect trendless samples. The use of tests for trend has not been a routine in the study of headway, in fact many authors have performed their statistical analysis on big sample of data, without taking the issue of stationarity into consideration. [15, 14, 25]. In other cases, the subdivision in time slices was due to other factors like weather conditions [13].

Sometimes the problem of nonrandom variation has been overcome by collecting samples at fixed time slices of length short enough to avoid any significant trend. Branston [3] applied a "short-term" sampling procedure on a two-lane section of the M4 motorway in West London, in which "data are sampled over time periods of a minute, or so". Zwahlen et al. [27] has subdivided a sample of 3-days data on Ohio freeways into 15-min time intervals to study the distribution of headways of free-flowing traffic ahead of work zones. Ha et al. [12] have applied a *short time sampling method* on RN118 national road in France to calculate aggregated variables during a fixed period of 6 minutes. Rossi and Gastaldi [21] assumed steady traffic condition by a 15 minutes time interval subdivision in two-lane two-ways roads in the province of Venice. However, as Luttinen has pointed out:

"[...] it is possible to have a significant trend in a 5-10 minute time period. One minute period, on the other hand, is too short, at least under light flow conditions". It follows that a fixed subdivision period in little intervals can eliminate long-term variations of the variable but may produce biased evaluation of headway properties if small samples with nearly equal means are combined together.

The proper approach is to obtain headways samples through a trend analysis. Luttinen [17] considered three different tests for trend - derived from statistical science - and evaluated their ability to detect trend through a Monte Carlo method: a linear trend was assumed for generating pseudo-random variates and the results of the different tests were compared. He concluded that the *exponential ordered scores test* - proposed by Cox and Lewis [8] in 1966 - was computationally more strenuous, but the more reliable method to detect trend. Therefore it was selected by the author as the prime method of testing trend in headway studies.

Other authors have cared about sampling data over time periods during which average flow was approximately constant. Breiman and Lawrence [5] have developed a method called *area test* which can detect sub-samples which own the property of constancy of flow rate; this method has been later applied by different authors, among which Branston [3] - who called it *"long term procedure"* - and recently by Ha et al. [12] - by the name of *"long time sampling method"*. Both the authors have used this test as an alternative to the fixed subdivision still used in their same works.

It is interesting to notice that Ha et al. [12] has obtained very long period samples using the *area test*; for this reason they deduced that two parameters of the procedure are not really rigorous, and finally applied all the three tests proposed by Luttinen [17] to detect more realistic stationary intervals.

3.2 Exponential ordered scores test

The test has been built by Cox and Lewis [8] to test the null hypothesis H_o that variables X_1,...X_n are *independent and identically distributed* random variables not necessarily exponentially distributed - so that the series is what is called a *renewal process*. This is quite useful in the study of headways considering that the arrival process is not completely a Poisson process. Luttinen [17] has excellently summarized the test revising that for headway studies.

The test attaches the score:

$$s_{r,n} = \frac{1}{n} + ... + \frac{1}{n-r+1}, \qquad\qquad r = 1, ..., n \tag{3.2}$$

to the rth order statistic in a sample of size n, where $t_{(r)}$ is the rth longest headway. Then is calculated the test statistic:

$$V = \sum_{i=1}^{n} s_n(i) \left(i - \frac{n+1}{2}\right) \tag{3.3}$$

The test statistic is asymptotically normally distributed with mean zero and variance:

$$\sigma^2(V) = \sum_{i=1}^{n} \left(i - \frac{n+1}{2}\right)^2 K_{2,n} \tag{3.4}$$

where:

$$K_{2,n} = 1 - \frac{1}{n-1} \left(\frac{1}{n} + \frac{1}{n-1} + ... + \frac{1}{2}\right). \tag{3.5}$$

The procedure is repeated incrementing the sub-sample size by 1 vehicle until the test reported trend at 5% level of significance, that is a common threshold for not rejecting the null-hypothesis. Once this condition is reached, the sub-sample size is decremented until the level of significance is over 70%. Fig. 3.1 depicts the boundary values referred to the standard normal distribution: the “*forward step*” ends when the point goes into the red zones ($z = \pm 1.96$), while *the* “*backward step*” - then the whole process - runs out once the oblique bars zone in reached ($z = \pm 0.39$).

The aim is that the sub-sample obtained satisfy three boundary conditions:

1. the sub-sample size should not be less than 100 elements, otherwise the sub-sample would be miserable.
2. the sampling period should not be less than 5 minutes: this condition is necessary to avoid an insufficient number of elements in sub-samples.
3. the sampling period should not be greater than 40 minutes: as Luttinen [17] explains, we aim to “keep the test sensitive to local trends”.

If one of these condition is not satisfied, the first observation in the sub-sample that seemed to cause the trend was removed, and the process was repeated.

The procedure has been implemented in R language[20] by an iterative process, that can detect the trendless sub-samples and evaluate the sub-sample’s features as time period and flow rate. The program takes approximately 5 minutes for a sample of 10^4 headways.

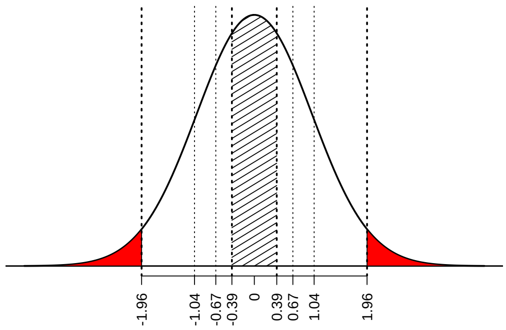

Figure 3.1: Standard normal distribution

In figure 3.2 is shown a graphical output of the procedure, where red and green line represent respectively the forward and backward step thresholds; black rings are drawn in the forward steps, and were filled by blue in the backward ones. In this case the procedure has detected a satisfactory sub-sample composed of 335 headways.

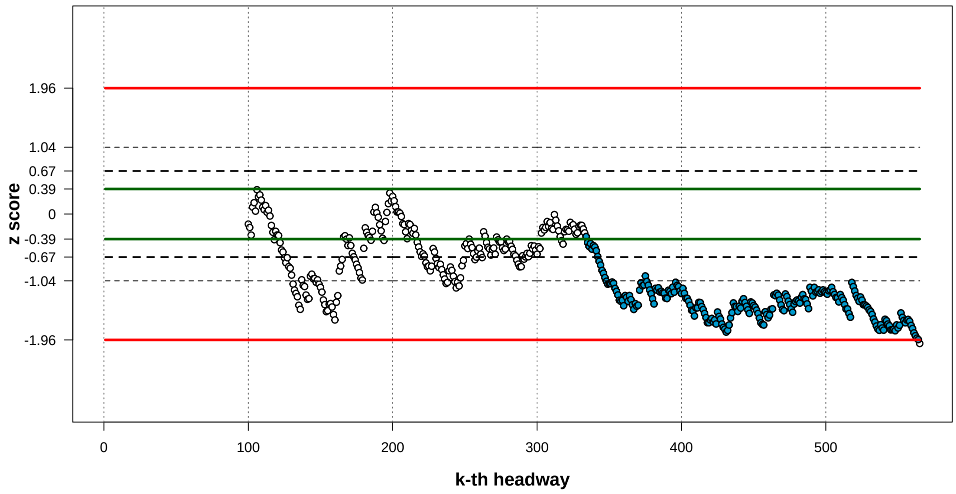

Figure 3.2: P-values, exponential ordered scores test

This test has been applied for all the sample data used in this work in order to obtained trendles sub-samples. The results of the test will be shown later case by case.

Chapter 4

Identification process

Traffic engineers are interested in examining the shape and the structure of TH distribution because it can reveal particular aspects of traffic flow, among which the traffic level, the relation between follower and free flowing vehicles and the correlation between consecutive headways. For this reason, we are interested in what Luttinen calls the *process of identification*, that is "the process of finding, from a set of model, the best theoretical model to describe the data"[17].

The basic assumption is that TH is a random variable with a specific distribution, still unknown.

4.1 Theoretical headway models

Since Adams [1], road traffic has been assumed to be a random series because this could lead to results of more practical utility. Consequently, traffic flow has been considered a *Poisson process* and headway - time between each pair of consecutive events - as a negative exponentially distributed variable. The proper model to describe traffic flow, under the assumption of a Poisson process, is called *negative exponential distribution*, and is expressed as:

$$F(t) = \begin{cases} 1 - e^{-\lambda t}, & \text{if } t \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

where λ is the *scale parameter* and correspond to the reciprocal of the mean headway:

$$\lambda = \frac{1}{\bar{t}} \tag{4.2}$$

With the aim of evaluating all headway distribution existing Luttinen [17] based on three considerations that are *reasonability*, *applicability* and *validity* of the model. The sequence is here proposed again in order summarize strengths and weaknesses of the distributions under analysis.

reasonability It has been demonstrated that under *weak* assumptions (no driver interaction and IID speeds) the headway process at some reasonably distance from the entry point can be assumed as a Poisson's process - that is what is called *Poisson tendency of low density traffic*. Even Branston [3] wrote that "on multi-lane unidirectional roads operating under light flow conditions, experiences of headway processes observed across all lanes has been reasonably consistent with the realization of a Poisson process". However, in a general traffic flow condition the assumption of a Poisson process is almost entirely unreliable, because some peculiar factors of traffic flow prevent traffic from behaving as a random phenomenon. These factors will be discussed later.

applicability The model has been widely used for his clarity; the parameter λ - the only one of the distribution - can be directly estimated by the sample thanks to the relation 4.2; furthermore it is connected to the flow rate, which is the most important factor that affects headway distribution.

validity As displayed in Fig. 4.1, the distribution overestimates the frequency of extremely short headways and underestimates it nearby the mode. Even the tail of the distribution is overestimated. Finally, the model has mode 0, that is improper through-and-through.

All these considerations allow to assert that the negative exponential distribution must be considered as the point of departure and the benchmark in headway studies. More accurate and detailed headway models would be useful for improving *reasonability* and *validity*, on the other hand the mathematical structure would always be more complicated - that means a low score in the field of applicability. Thanks to the classification made by Ha et al. [12], a fast overview of the main headway models used in literature is here made through the distinction simple-combined-mixed models.

4.1.1 Simple models

The most simple assumption in TH modeling is to assume the same behavior for all vehicles without any kind of distinction. Simple models are characterize by a simple mathematical formulation, combined with a low number of participating parameters, that makes the model easy and convenient. Common statistic distributions are used, like exponential (exposed above), gamma and log-normal.

reasonability Since each models is based on different assumptions, a single distribution is reasonable only in particular circumstances. Negative exponential distribution describes situations where vehicles without physical length can move freely [17], so it is reasonable under weak traffic conditions. Log-normal distribution has a great relation with *car-following*, therefore it is an attractive distribution only for follower headways.

applicability Simple models usually does not present problems in parameter estimation, and relatively easy and efficient subroutines can be used to generate model variates; therefore the applicability of simple models is absolutely good.

validity As stated before, these models have some limitations, foremost is they have difficulty modeling short headways. Therefore goodness-of-fit tests always show low levels of acceptability in terms of p-values, that make these models unreliable.

In order to improve the goodness-of-fit, it is used to add a location parameter τ that displace the whole distribution to the right by τ seconds.

Inverse Weibull distribution

Rossi and Gastaldi [21] applied different simple models to data recorded on two-lane two-way roads in the province of Venice, with the aim to evaluate which one could fit the data best. They concluded that for these kinds of roads, the Inverse Weibull (IW) distribution was the most suitable function and the most appropriate, through all the simple models, to represent data. Since traffic data used in this thesis come from the same roads analyzed in that research, the IW distribution is taken into account for a comparison with more elaborated models. The structure of the model is given by:

$$f(x) = \alpha\, \beta\, \left(\frac{1}{\beta\,(x - x_m)}\right)^{\alpha+1} exp\left(-\left(\frac{1}{\beta\,(x - x_m)}\right)^{\alpha}\right) \tag{4.3}$$

where α is the shape parameter, β is a mixture of shape and scale parameter, and x_m is the minimum headway of the sample. The visual fit as shown in Fig. 4.2 of the pdf is rather satisfactory.

4.1.2 Combined models

The inability of simple models to describe both the sharp peak and the long tails of the headway distribution suggest to assume two categories of vehicle - that is *follower* and *non-follower*[1] - each one generates headways that belongs to one category. In fact, since vehicle in different categories have different headway property, the division should be included in the model, which accordingly becomes a mixture of two distributions [17]. According to the queueing theory, the model consist in a system made of two servers, only one of which may serve at a time; each vehicle is served by server 1 with probability ϑ and by server 2 with probability (1-ϑ) - since the queue is continuous (Fig. 4.3). The result is that a headway H can be either the variate U (following headway) or the variate V (non-following headway), so that the two variables are independent.

The total distribution is a linear combination between the two components:

[1] We assume for convenience all non-followers to be leader, even when they have no followers. In this sense, we use the terms leader and non-follower as synonymous.

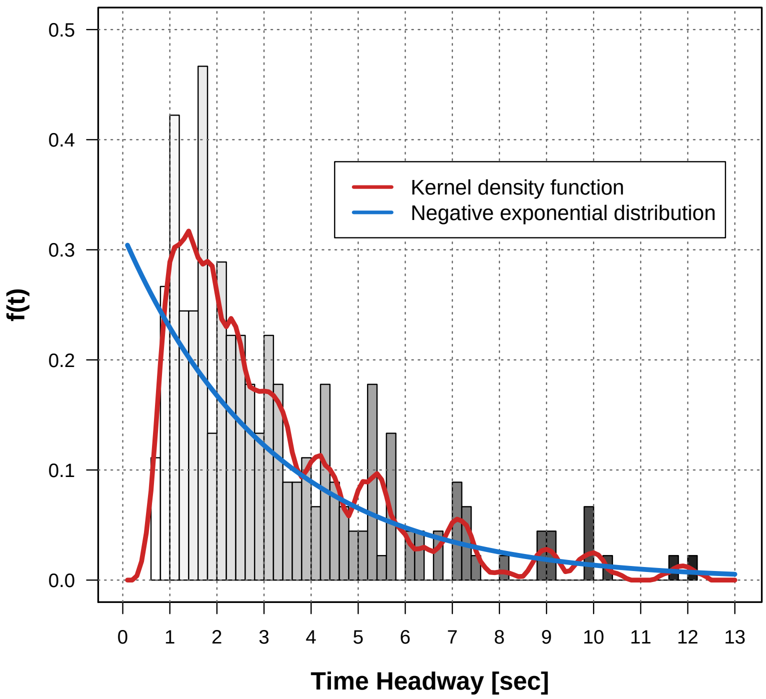

Figure 4.1: Negative exponential distribution, kdf and histogram of a sequence of headways

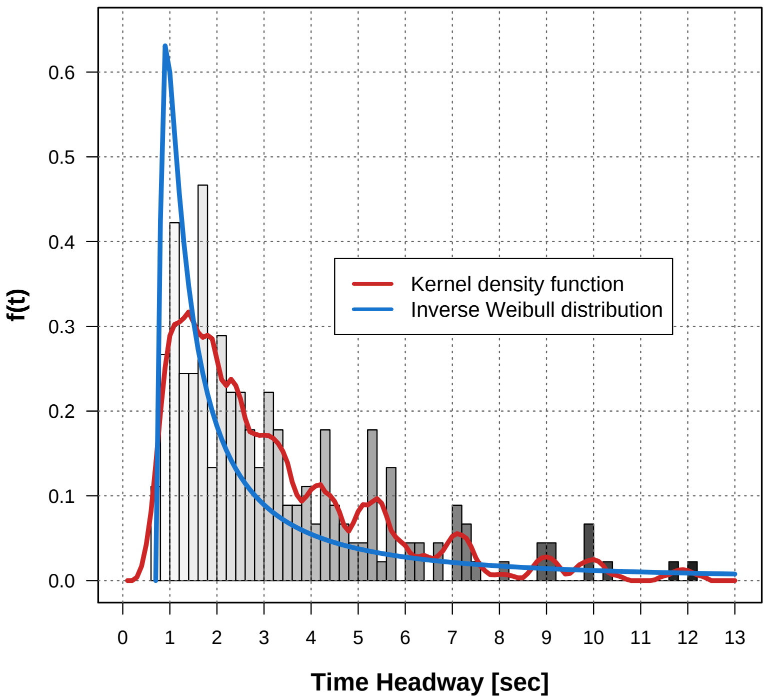

Figure 4.2: Inverse Weibull distribution, kdf and histogram of a sequence of headways

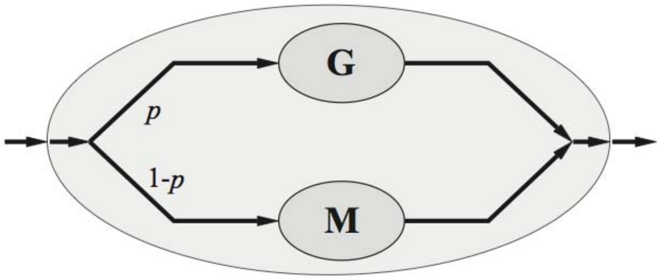

Figure 4.3: Queuing model for a combined headway distribution

$$f(t) = \theta g(t) + (1 - \theta)k(t) \tag{4.4}$$

where *g(t)* is the distribution of the variate U, *k(t)* that one of V. It is common to assume an exponential distribution for the function *k(t)*, while for *g(t)* different function have been proposed by the researcher. Note well that if traffic flow is composed only by non-following vehicles - that means $\theta = 0$ - the model reduces to an exponential distribution. To this category belong Cowan M3, Hyperlang and Hyperexponential distribution.

4.1.3 Mixed models

Mixed models are also based on the subdivision between leader and follower vehicles, but a precise threshold to divide headways doesn't exist. In fact the subdivision should take into consideration factors like the vehicle type (both of the leader and the follower) or the driver characteristic of the follower, who may assume a different security distance depending on a large number of factor (i.e. age, level of attention or driving aggressiveness). Therefore in these models follower headway range is not confined under a fixed value, but on the contrary it is extended through all the spectrum, without any limitation. Their probability density function is representative of the *zone of emptiness*[2] consequent to each vehicle and maintained for safety reason. In a high density traffic state it is assumed to be bell-shaped and right skewed, like lognormal or gamma models; for low volumes even an exponential distribution can be used without committing many errors.

Leading headway, which is not influenced by the vehicle ahead, owns two main characteristics:

- it is the result of a Poisson process;
- his range is infinitely extended to the right, but confined on the left by a fixed threshold in combined models, a variable value in mixed ones.

While the first point allows us to use an exponential distribution for leaders, the second one forces us to modify it in order to take into account that leading headways are generally absent in the zone 0-3 seconds. This question has been faced in two different ways in

[2]even called *empty zone* by other authors.

literature, through a diverse mathematical transposition: we refer to the *Semi-Poisson model*, proposed by Buckley [6] in 1968, and *Generalized Queuing Model*, proposed by Cowan [7] in 1975 (under the name of Cowan M4) and Branston [3] in 1976.

Semi-Poisson model

In the semi-Poisson model proposed by Buckley [6] the conjecture is that non-following headway V is always greater than following headway U. Given the equation 4.4 - where the distinction follower-leader is still valid for mixed models - *h(t)* is an exponential function with parameter λ modified so as to include only headways greater than a random variate sampled from *g(t)*.

Luttinen described the Semi-Poisson model as a queuing system (Fig. 4.4), similar to that of combined model, but with a modified exponential service time distribution M'.

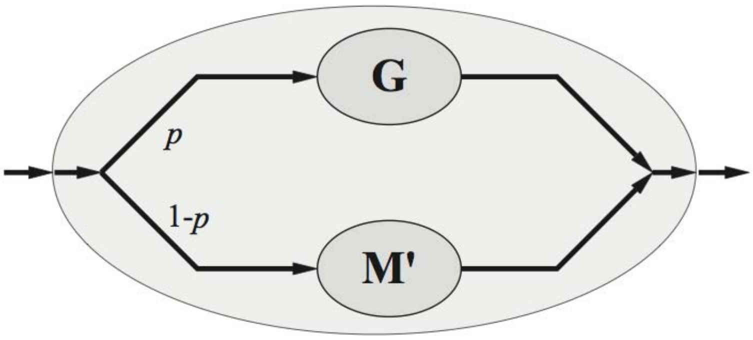

Figure 4.4: Queuing model for the semi-Poisson distribution

Function *g(t)* characterizes the zone of emptiness in front of each vehicle and is associated with the fluctuation in car-following [24]. Many authors have proposed different functions for *g(t)* like normal, truncated normal or gamma, with different results depending on many factors. As said before, even exponential distribution could be used in case of low traffic situation with restricted vehicle interaction, but it would be unsuitable in higher traffic conditions. Anyway, best results were obtained with gamma distribution, that appears to be the best model for following headways.

GQM model

Generalized Queuing Model (GQM) proposed by Cowan [7] and Branston [3] assumes that each headway is composed by two mutually independent random variable:

$$H = X + Y \tag{4.5}$$

where X is the *following component*, Y the *free-following* one. As Hoogendorn et al. [15] wrote: "the independence of these random variable is plausible since both result from different independent process: the choice of the driver to follow his predecessor at a certain time headway on the one hand and the arrival process of the free-flowing vehicles on the other hand".

In the GQM model proposed by Cowan [7] and Branston [3], the component Y is participating with a probability of (1-ϑ), while X is always present and represent the *empty zone*. Headway series is then a mixture of following vehicle, with variable U=X, and non-following ones, with variable V=X+Y.

While Y is assumed to be exponentially distributed, X can be described by any simple distribution; the more suitable to describe following component are however gamma and lognormal.

The pdf of GQM model is given by:

$$f(t) = \theta g(t) + (1-\theta)\lambda\, e^{-\lambda t} \int_0^t g(x)\, \exp(\lambda\, x)\, dx \tag{4.6}$$

where the second terms is the convolution of the function $f(t)$ and $g(t)$ (density of the sum of two variables) and represent the distribution of non-following headways. The analogy of the server is shown in Fig. 4.5: GQM model is assumed to be a queuing model with Poisson arrivals and a general service time distribution, that is a M/G/1 queuing system. As long as the server is busy it produces follower headways V, and with a probability of ϑ the server experiences some idleness during the interdeparture period, which is the sum of the service time X and the idle time Y (Gross and Harris [9]).

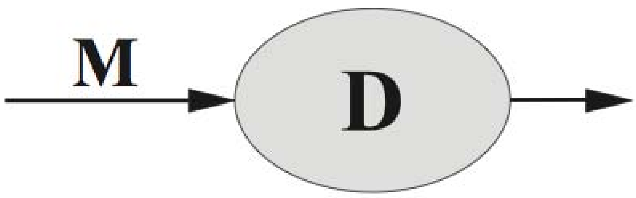

Figure 4.5: Queuing model for the GQM distribution

We want now to compare the arrival process given by GQM model with a simple Poisson process, mathematically expressed by the exponential distribution, in order to highlight the advantages step by step. Branston [3] identified two main physical factors that prevent traffic in a single lane from behaving as a random process, that is:

- *Stability and safety considerations*: following vehicle must assume larger headways than those allowed by a random process - this behavior is well explained by the car-following theory. Size of following headways are determined by factors like the length of the leader vehicle and following driver's perception of a safe separation from leader, affected by both speeds.
- *Lack of overtaking opportunities*: vehicle who wants to overtake are forced to catch up, in respect to the previous vehicle. This factor is an additional distinction to a free Poisson process.

Both combined and mixed models take into account the first consideration by the distinction between following and non-following vehicles, mathematically expressed by two different functions. In the matter of the lack of overtaking opportunities, Branston [3] explains that it is taken into account "by allowing the proportion of following vehicles - ϑ - and the parameter of the negative exponential distribution of gaps - λ - to vary [...]" from fixed values.

This model has been chosen by many authors in headway study, in particular Hoogendoorn et al. [15]"because of its clear theoretical foundation" and the excellent results obtained in previous studies, and Ha et al. [12] who proved the benefits of the application of this model in terms of goodness-of-fit.

Considerations

We report here three consideration about combined and mixed models, very useful to appreciate their distinction inside and out:

- both kinds of model are a linear combination of following and non-following headways (U and V). The structure of the model is always based on equation 4.4;
- U and V are independent in combined model, on the other hand they have a probabilistic relationship in mixed models;
- in the case of $\vartheta = 0$, mixed models doesn't reduce to an exponential distribution.

Both models represent the best choice through all headway distributions in terms of reasonability and validity, with the unique disadvantage in applicability, not simple but quite laborious. The choice between the models is at modelist discretion, since the goodness-of-fit is excellent in both cases, and both models could be criticized for their basic assumptions.

In fact Wasielewski [24] criticized to the GQM model for two reasons:

- it seems lacking in physical basis: each non-following headway is composed by a term drawn from the exponential distribution supplemented by a following headway, which has no physical reason to be added since leaders are not absolutely influenced by the vehicle ahead. This consideration induced him to use the SPM model in his work; but Luttinen [17] has showed that even in that case the free headway was a sum of two random variables, and in some way influenced by non-following headway distribution.
- the choice of the function for follower headways seems inherently arbitrary.

Furthermore, even Luttinen [17] criticized GQM model because his formulation is not based on the principles of the traffic flow dynamics; in fact Branston [3] derived GQM model from queuing theory.

From a theoretical point the basic mathematical assumption of both models can inherently be criticized or bashed.

Our engineering approach permit us to take some liberties and to focus on the performances of the model (the field *validity*), in order to look for a distribution that fits data at best and with the minimal effort. In this sense, a good research work has been done by Ha et al. [12], who looked for the best theoretical model combined with the most useful calibration method, found that:

- The goodness-of-fit time for the SPM is very long while the goodness-of-fit time of the GQM is very short. This time indicates the time necessary to compute the statistic tests (in their case: K-S, chi-squared and A-D tests) after obtaining the estimated parameters.
- GQM distribution provides a better fit than SPM using the gamma function as g(t) in both cases.

4.2 Estimation method

Once the model has been chosen, the parameters of the distribution must be calculated. Various method for estimation have been used in literature to calibrate the parameter values, among which the most used have been:

- method of moments;
- minimum chi-square method;
- maximum likelihood method.

Each method has his own peculiarity, here briefly summarized.

Method of moments

It requires the solution of a system that equalizes the first k population moments to the corresponding expressions for the k samples moments - where the first moment (mean headway) is simply obtained by the traffic volume. The method has been widely used for his simplicity but is, however, not efficient, mainly because the information in the data is reduced to the sample moments ([17]). Method of moments is quite efficient for normal distributions, but usually not appropriate for skew distribution like headway ones; anyway it can be used to give first approximations for more efficient methods.

Minimum chi-square method

The whole sample is subdivided in m mutually exclusive and exclusive classes, then the estimators are obtained by minimizing the chi-square statistic:

$$\chi^2 = \sum_{i=1}^{m} \frac{\left[n_i - n\, p\,(i \mid \overline{\theta}\,)\right]^2}{n\, p\,(i \mid \overline{\theta}\,)} \tag{4.7}$$

where n_i is the number of observation in the class i, n is the sample size, and $p\,(i \mid \overline{\theta}\,)$ is the probability of an observation falling into class i conditional upon parameter vector $\overline{\theta}$. The method is computationally simple and easy to implement, but has strong limitations because of the data subdivision in classes, that is troublesome especially in small samples. For this reason, if the sample is not copious, this method must be absolutely avoided.

Maximum likelihood method

This method aims to maximize the agreement of the selected model with the observed data, that is to say to find particular parameters values that make the observed results the most probable, given the model. For a random variable x, likelihood function of $\hat{\theta}$ is defined as:

$$L\,(\hat{\theta}) = \prod_{i=n}^{n} f\,(x_i \mid \hat{\theta}\,) \tag{4.8}$$

The parameters $(\theta = \theta_1, \ldots, \theta_k)$ of the distribution are obtained by maximizing the log-likelihood function, that is more easy to calculate:

$$ln\,L\,(\hat{\theta}) = \sum_{i=n}^{n} ln\,f\,(x_i \mid \hat{\theta}\,) \tag{4.9}$$

The maximization is reached by a set of equations where the partial derivates are equal to 0; this way the solution can be obtained analytically, and the resulting set of parameters $(\theta = \theta_1, \ldots, \theta_k)$ are called maximum likelihood estimators. The MLE is assumed to be the best method to calibrate the parameters for his properties of consistency, asymptotic normality and efficiency.

4.3 Goodness-of-fit test

The goodness-of-fit tests shows how well a statistical model fits a set ob observations. It tests the hypothesis that a particular distribution $F(x \mid \theta)$ with parameter $(\theta = \theta_1, \ldots, \theta_k)$ is the true distribution $G(x)$ which has generated the data:

$$H_0 : G(t) = F(x \mid \theta) \qquad\qquad (null\ hypothesis) \tag{4.10}$$

against the *alternative hypothesis* that the distributions differ. Assuming that H_0 is true, the probability to reject the null hypothesis - that is called *type one error* - is called *significance probability*, while the maximum probability of rejecting H_0 is called *significance level* of the test. The general procedure consists on defining a test statistic which measures the distance between the hypothesis and the data; knowing how the test statistics are distributed, it is possible to calculate the *significance probability*, even called *p-value*, that is a general measure of the goodness of the model. It is common to reject the null hypothesis if the p-value is smaller then 0.05.

Kolmogorov-Smirnov test

The Kolmogorov-Smirnov statistic (D) is based on the largest absolute vertical difference between $F(x \mid \theta)$ and $F_n(x)$ - that is the empirical density function. It is defined as:

$$D_n = \sup_x \mid F_n(x) - F(x \mid \theta) \mid \tag{4.11}$$

and the null hypothesis H_0 is that $F_n(x)$ is the empirical density function (EDF) of a random sample generated by $F(x \mid \theta)$. The significance probability of the test is:

$$p = 1 - F_D(D \mid n) \tag{4.12}$$

where $F_D(\cdot)$ is the PDF of the K-S distribution. The advantage of K-S test is quite relevant when the sample is small; in this case the chi-square test is troublesome because of the arbitrary discretization of the distribution function in bins, and the test would be less powerful.

4.4 Selection of the method

A general process for headway fitting consists basically on three choices:

1. selection of the model;
2. choice of the estimation method
3. evaluation of the calibrated model by a goodness-of-fit test.

The characteristic of each model, thanks to the three consideration proposed by Luttinen - *reasonability, applicability, validity* - should essentially lead to the selection of the model to use for fitting, depending on the aim of the research. If the characteristics of the road allow to use even simple distribution - that is to say the model is valid - the advantage is quite evident in the field of applicability. Otherwise, in the absence of the apt conditions, one should choose the composed or mixed models.

Regarding the estimation method, the most appropriate method has to be chosen by the criteria of performance and facility, well explained by Ha et al. [12]:

performance it concerns the results of the goodness-of-fit test obtained using the estimated parameter values, commonly measured by a statistic test, or better by a p-value;

facility is the rapidity of the method to carry out the estimation process, even called *execution time.*

As a consequence, the most appropriated estimation method may be chosen according to the type of data examined and the theoretical model adopted, in reference to the accuracy of the results aimed.

On the other hand, the selection of the goodness-of-fit test is independent from the previous choices done. The K-S test is absolutely better than for example a chi-squared test, beyond the model selected or the empirical data characteristic.

In headway literature various model and diverse estimation methods has been applied - an excellent summary about different authors choices has been done by Ha et al. [12].

The choices operated in this work are here explained and justified:

1. In this research gamma-GQM distribution - i.e. GQM model with *g(t)* corresponding to the *gamma distribution* - was chosen. Ha et al. [12] applied different models to samples of data from RN118 national roadway in France and deduced that the gamma-GQM without location parameter has to be considered as the best among all the time headway models, because it provides highest fits by high index of p-values in the goodness-of-fit test. This is consistent with the aim of this research, that is to calibrate the model that fits the data best. Furthermore, a comparison between gamma-GQM and Inverse Weibull (IW) distribution has been operated, for the reasons already explained.
2. In order to estimate the four parameters of the model, log-likelihood method was adopted for both distributions. The values of the IW has been calculated analytically by R code [20] through the minimization of the log-likelihood function. On the other hand, to obtain the set of parameters for the gamma-GQM, a numerical procedure has been applied, which calculates iteratively the log-likelihood function for a set of pre-determined parameters. The procedure is explained in Appendix A.
3. The evaluation of the goodness-of-fit is done by the K-S test because of his remarkable use in headway literature, then the possibility to compare test statistics of previous works with those coming from this research.

4.4.1 Gamma-GQM model

As exposed before, the gamma-GQM is a GQM that uses gamma distribution as the *g(t)* function, and is given by:

$$f(t|\alpha,\beta,\lambda,\theta) = \theta f_1(t) + (1-\theta) f_2(t) = \theta \frac{\beta^\alpha\, t^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta t} + (1-\theta)\frac{\beta^\alpha}{\Gamma(\alpha)} \lambda\, e^{-\lambda t} \int_0^t u^{\alpha-1}\, e^{-u(\beta-x)}\, dx \tag{4.13}$$

The knowledge of the role of each parameter is essential to understand their variation as the traffic conditions change.

Parameter θ is, above all, the most important factor that affect headway distribution. From the mathematical point of view, it establishes the weight between the distribution of *following* and *free* vehicles; if $\theta = 1$, the gamma-GQM reduces to a pure gamma distribution, while when $\theta = 0$ the first term of the equation disappears and the headways follow a mixed exponential-gamma distribution, consequent to the convolution theorem. From the traffic flow point of view, θ is the part of vehicles which owns the following-component only, that means following vehicles; a value of θ close to 0 means that almost all vehicles are free. Fig. 4.6 illustrates the contribution of following and non-following headways to the total pdf.

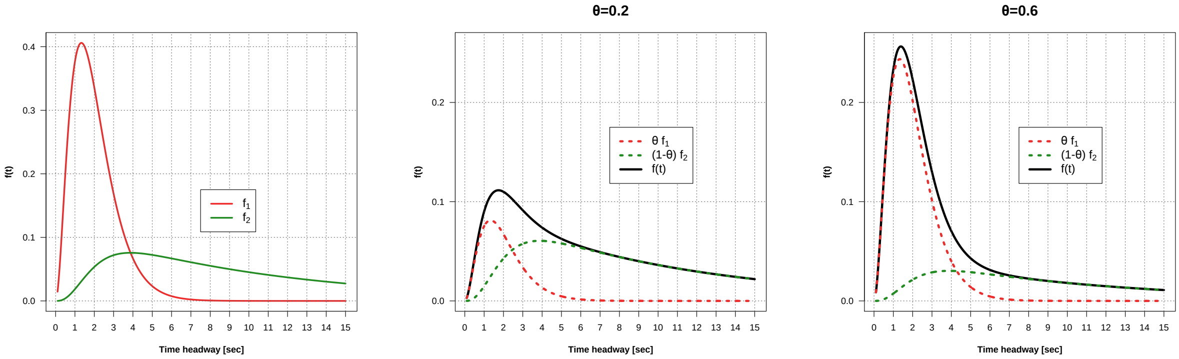

Figure 4.6: Density function of following and non-following headways by different values of θ

Parameter λ represent the intensity of the exponential tendency and has "an important role in decreasing the slope of its tail to fitting long headways"[13]. Since it derives from the exponential distribution, it is connected to the free headways: the value $1/\lambda$ denotes indeed the mean headway of the Y component. For a sensitivity analysis of parameter λ on gamma-GQM model, see Fig. 4.7.

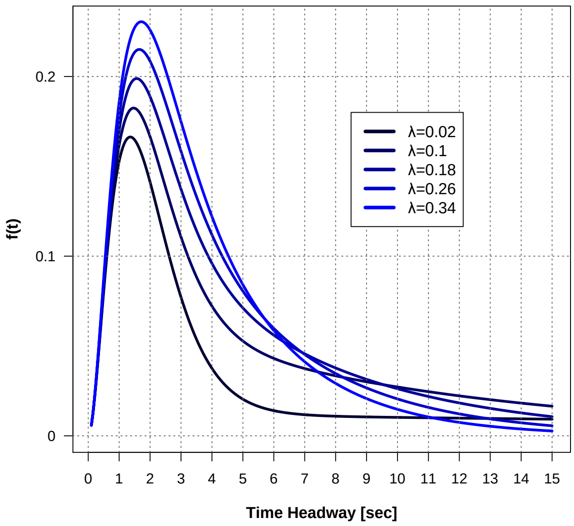

Figure 4.7: Sensitivity of parameter λ on gamma-GQM distribution

Specification of the empty zone

In gamma-GQM model the empty zone is described by the gamma distribution, called sometimes Pearson-III. Since Ha et al. [12] said that the best choice for headway modeling is to use the gamma-GQM without location parameter (associate to the LL method), no location parameter τ is used. Therefore, gamma distribution is defined only by α and

β, respectively *form* and *shape* parameter, with expectation E(X) and the variance given by:

$$E(X) = \alpha/\beta \tag{4.14}$$

$$var(X) = \alpha/\beta^2 \tag{4.15}$$

Hoogendoorn [15] says that "a straightforward interpretation of α and β from a traffic-flow theory point of view is not possible for the individual parameters". Despite that, the ratio α/β, is extremely relevant. Differences in the empty zone shape may derive from a series of factors, like the attention level of drivers, driver composition (with respect to the trip purpose, that is the distinction commuters-recreational), ambient conditions and vehicle composition (passengers car-trucks). Hoogendoorn et al. [15] studied how empty zone distribution change when varying two factors, namely the time of day and the type of vehicle, which affect TH distribution considerably.

Finally, "the coefficients of α and β are relatively large compared with other parameters" [15]; this implies that "small disturbances in the sample may cause relatively large differences in the estimate of α and β"[15] , and that a large number of samples is needed in order to detect a clear evidence.

Chapter 5

Results

In this chapter we want to present the results obtained by applying the same approach to the study of headways on motorways and two-lane two-way roads, including trendless analysis for the identification of stationary sub-samples and the calibration of the gamma-GQM function. Furthermore the achievement of the stationary condition allows to build the fundamental diagrams of traffic flow, to appreciate the real benefits of this procedure. The estimated parameters will be firstly analyzed as the flow rate increases; then we will take into account the type of lane (for motorway) and the percentage of heavy vehicles (for TLTW roads) . Finally, a graphical output will show the shape of the distribution, then the sensitivity of the model to endogenous traffic conditions.

5.1 Three-lanes motorway

Recently, the “Institut für Verkehr und Stadtbauwesen” at the TU Braunschweig has been carrying out a project about the effectiveness of fixed speed cameras on the three-lane A2 motorway Berlin-Dortmund. Double-loop detectors, installed at different road sections over the segment Hannover-Braunschweig, enable determination of headway (0.01 second resolution), speed, vehicle length and vehicle type for each lane of the motorway. Thanks to the high data supply, the set for the analysis has been collected on the strength of our requests. In particular, one road section has been chosen, located between the villages of Wunstorf and Bad Nenndorf (Fig. 5.1) in the province of Hannover, at km. 248.22 toward Dortmund, with a level terrain and lacking of on-ramps or off-ramps close to the observation point; moreover, time periods has been identified as whole days - three weekdays as Monday, Wednesday and Friday in March 2013 - with good weather conditions corresponding to the lack of precipitations.

Data are available for all lanes and both directions, so that each recording session - corresponding to one day in our case - contains six sets of data, mentioned as in figure 5.2 for convenience. Let us remark that under German legislation there is no speed limit in the motorway for passengers-car, while some categories of heavy vehicles are forced to respect speed limits (e.g. 100 km/h for busses). This condition will affect the shape of headway distribution, that would vary according to the type of lane - slow, middle or fast

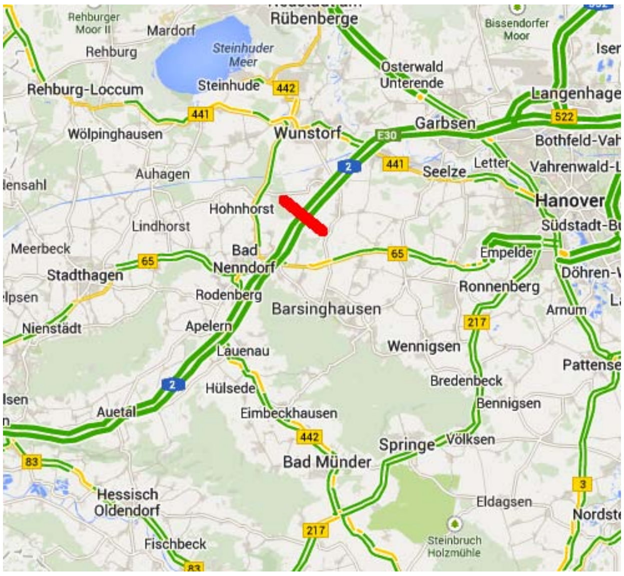

Figure 5.1: Location of the road section, Germany

- than to the vehicular traffic composition; indeed heavy vehicles are present in the slow lane with percentage closer than 70%, while the fast one is totally heavy vehicles-free.

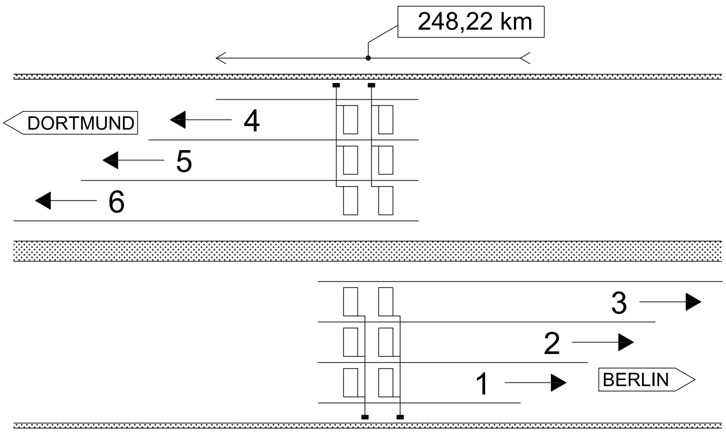

Figure 5.2: Sketch of the road section

5.1.1 Selection of trendless samples

As a first step, data belonging from the night period was discarded so that the analysis could focus on the day period 6 a.m.- 9 p.m.. This choice is justified by two main reason: firstly low traffic condition have always had a low appeal in traffic engineering applications, secondly trend test would have barely detected trendless samples passing the boundary conditions expressed in paragraph 3.2. Indeed, the union of conditions (1) and (3) leads to the restriction *FR >150*, that is rarely observed during the night period. All data sets that were part of the study are listed in Tab. 5.1, by the specification of the day of recording, the total number of car passages (observations) and the percentage of heavy vehicles (if the length is more than 7.5 metres, the passing vehicle is considered as heavy). Fig. 5.3 shows the mean values of the number of observations and % of heavy vehicles for each lane among all days.

Lane	Date	Obs.	% HV	Lane	Date	Obs.	% HV
1	Mon 4.3.13	10,356	61.1	4	Mon 4.3.13	11,684	71.7
1	Wed 27.2.13	11,629	78.7	4	Wed 27.2.13	11,220	74.4
1	Fri 1.3.13	12,105	73.5	4	Fri 1.3.13	11,054	54.1
2	Mon 4.3.13	15,125	1.4	5	Mon 4.3.13	15,984	6.8
2	Wed 27.2.13	14,674	1.7	5	Wed 27.2.13	14,850	8.1
2	Fri 1.3.13	18,422	1.6	5	Fri 1.3.13	17,090	3.2
3	Mon 4.3.13	9,412	0	6	Mon 4.3.13	11,329	0.1
3	Wed 27.2.13	8,886	0	6	Wed 27.2.13	10,737	0.2
3	Fri 1.3.13	13,019	0	6	Fri 1.3.13	12,955	0.1

Tab. 5.1: Days of traffic survey and main characteristics of traffic data

Afterwards all data-sets were submitted to trend analysis through *exponential ordered scores test* (par. 3.2). The number of trendless samples obtained is resumed in Tab. 5.2, according to the lane and the day of observation.

Lane	Mon	Wed	Fri	Tot
1	38	36	36	110
2	50	40	38	128
3	39	35	39	113
4	37	35	34	106
5	45	35	33	113
6	36	32	34	102

Tab. 5.2: Number of intervals detected, with lane and day specification

A clear overview is plotted in Fig. 5.4: each point in the graph represent a stationary sub-sample detected by trend test and is characterized by the value of FR and the number of records contained. Moreover, the grade of the connection line of any point with the origin measures the time length of the sub-sample, that must be comprised in the range 5-40 minutes (red dotted lines).

Finally, a visual acknowledgment of the ability of trend test to detect stationary intervals is plotted in Appendix A.

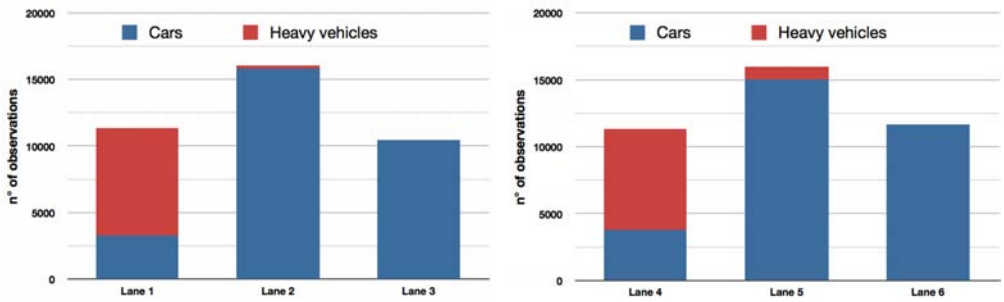

Figure 5.3: Mean values for all lanes

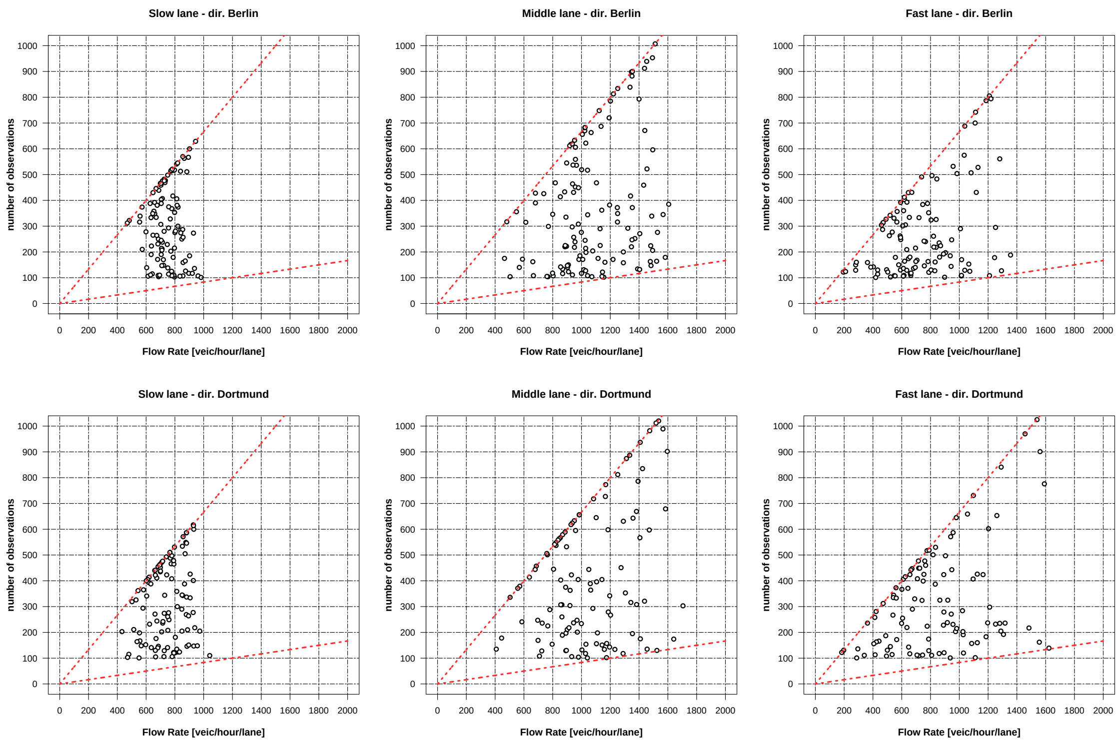

Figure 5.4: Flow rate and sample largeness for all lanes

5.1.2 Fundamental diagrams

Fundamental diagrams of traffic flow reveal the relations between the three macroscopic variable flow rate q, density k and space mean speed v_s, expressed respectively in $[veic/h]$, $[veic/km]$ and $[km/h]$ and connected together by the well-known relation:

$$q = k\,v \tag{5.1}$$

Usually fundamental diagrams are used to establish the capability of a road system; in this study only the relation $q - v_s$ will be inspected, with the aim to show how stationary samples identified by trend test behave in the $q - v_s$ diagram.

The flow rate value is directly calculated from the sub-sample, by applying a coefficient of 2 for heavy vehicles. In regard to v_s, the identification of trendless samples - then of steady flow conditions, although only in time - allows to estimate the space mean speed as the harmonic average of vehicle speeds recorded at the section during time intervals. The formula is here reminded, assuming v_i as the istantaneous speed of the i-th vehicle and n the number of passing vehicles in the sample:

$$v_s = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{v_i}} \tag{5.2}$$

Firstly, it is interesting to compare the points obtained by trend analysis to those obtained by fixed time subdivision, 5 minutes for instance. The comparison (Fig. 5.5) is focused on all the lanes in the direction of Dortmund on Wednesday in this case.

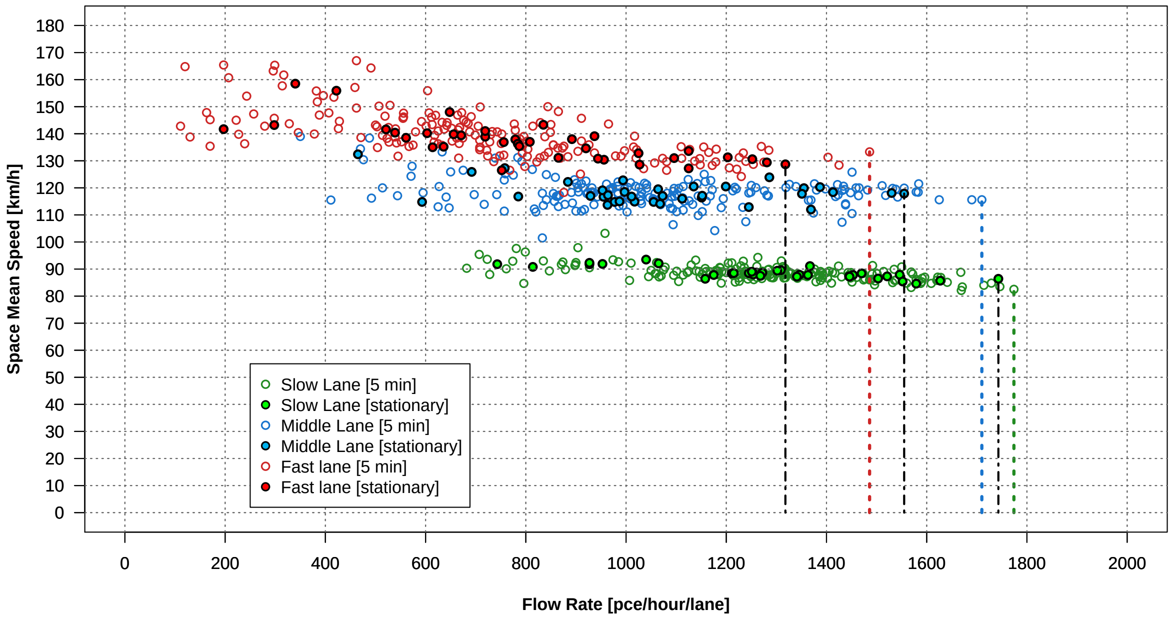

Figure 5.5: Space mean speed / Flow rate pairs estimated for each interval and different lanes (27.2.2012, dir. Dortmund)

The distribution of points by trend test, especially in slow and middle lane, is less scattered; this is quite normal, since stationary points belong to greater intervals in time (form 5 to 40 minutes), consequently they are more gathered up. However, the benefit of trend test can be appreciated on the right part of the diagram, where dotted coloured lines has been drawn: it is clear how a fixed time subdivision is not correct when estimating the maximum number of vehicles that can cross a section, because it would not be representative of stationary conditions. In this case, the 5 minutes partition leads to an overstate of the maximum traffic volume passed through the section: the colored line is always more on the right than the corresponding black line, that is for the stationary intervals. It is important to note that for the fast lane this difference is just shy of 200 pce/hour/lane. If the road had reached capacity, the point furthest to the right would have corresponded to the maximum FR that can cross the section, and a correct process for removing trend would be recommended not to commit heavy errors.

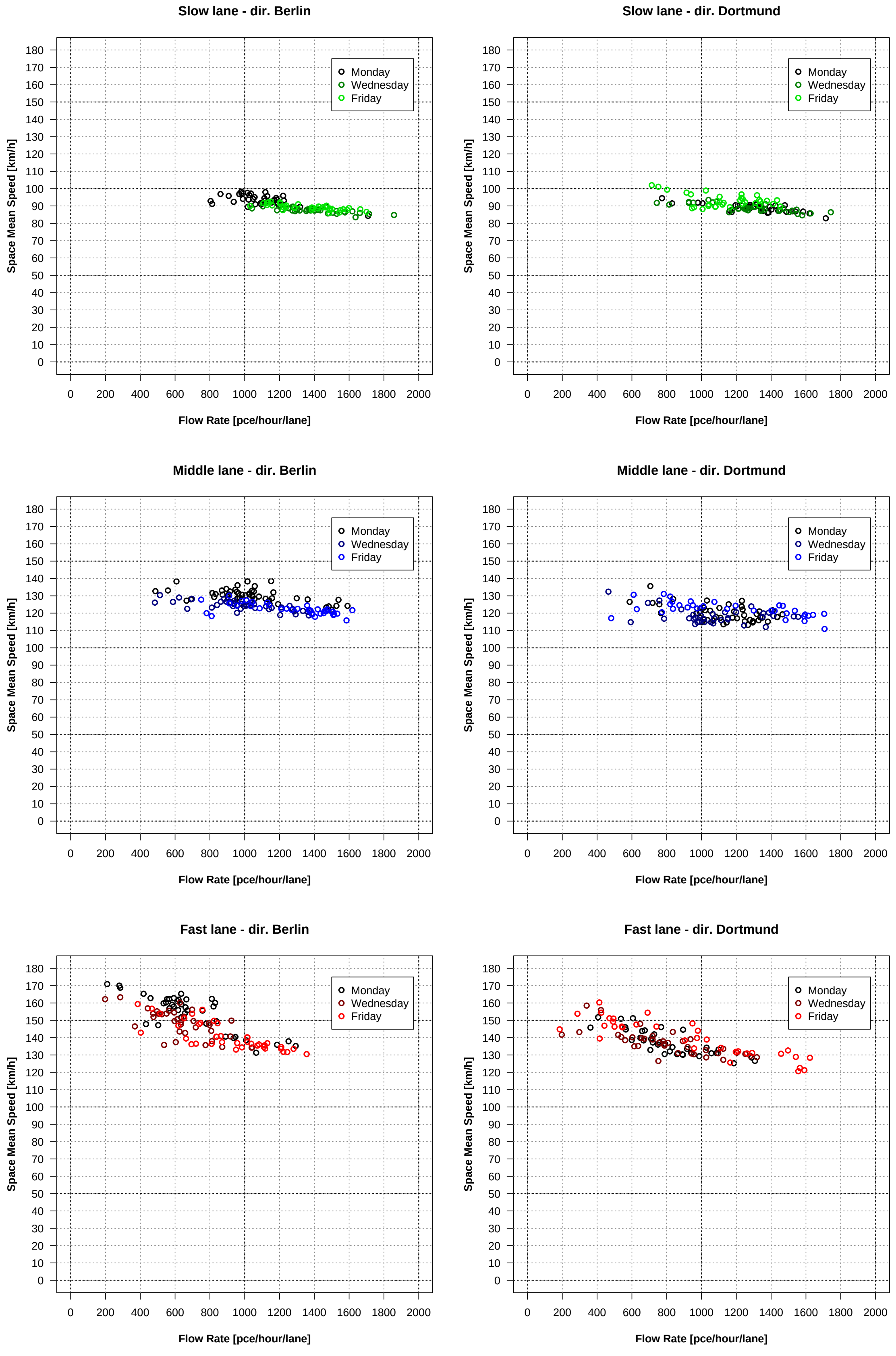

Figure 5.6: Space mean speed / Flow rate pairs for each lane and different weekday

Fundamental diagrams for all lanes are plotted in Figure 5.6, by the distinction of the day of the week (gradation of color).

The dispersion of the points arises with the speed of the lane. The highest flow rate is reached in the slow lane, because of the high concentration of HV; here traffic conditions are more stable and the arrival pattern is more uniform.

5.1.3 Mode and peak height

The benefits of the kdf could be particularly appreciated by the accurate estimate of the values of mode and peak height, with a precision of 10e-2. Fig. 5.7 reveals the trend of the mode of the estimated kdf with respect to FR. Looking at the results, we can assert that there no significant correlation with FR exist, in all lanes; d moreover, in middle and slow lane, mode is less scattered as FR increases; finally, mode is basically higher as the speed of the lane increases. Results are similar to the ones obtained by Ha et al. [10], who founded that, as the speed of the lane decreases, mode increases and become more scattered. It should be noted the mode value is dependent not only on speed, but also on vehicular composition: a high percentage of HV (longer than 7.5 m) increases the mode value because of the length of the vehicle.

In the matter of peak height, a clear tendency is evident in Fig. 5.8: peak height always increase as the flow rate arise. Peak height on fast lane is generally higher than that on slower lanes for same FR. Results are quite gathered around an hypothetical regression curve.

5.1.4 Estimation of time headway distribution

We assume headway to be distributed according to the gamma-GQM distribution. Thanks to the *maximum likelihood method* the pdf was estimated for each sub-sample, then and the K-S test was performed to evaluate the goodness-of-fit. Besides the lane and direction specification, sub-samples were classified according to the value of flow rate; classes of FR by a range of 200 [veh/hours/lane] were created in order to estimate the mean values of all the parameters of the pdf belonging to the interval, and to obtain a set of pdf for each FR range, lane and direction. The goodness-of-fit is valued by the p-value, that is the significance probability of the test.

In Table 5.3 are provided the mean values of the parameters within the respective FR class. In some cases the estimation process have led to extremely high values of α and β (more than 50 units): the process was then interrupted and the sample discarded, so that mean values wouldn't have been affected. Either way, the reference parameter of the gamma function is not the absolute value of α or β, but their ratio, which is pretty stable.

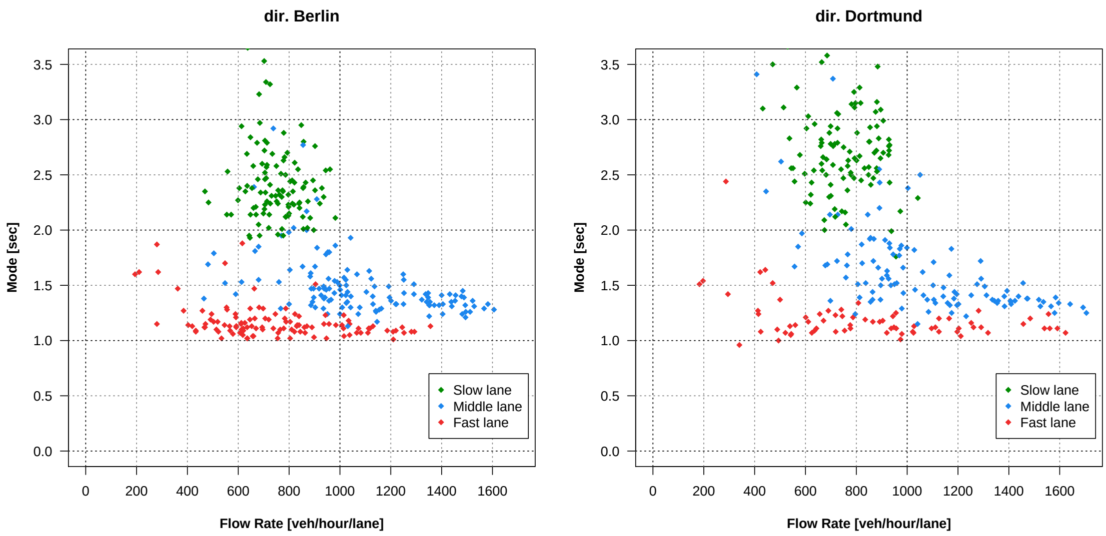

Figure 5.7: Mode of empirical headway distribution

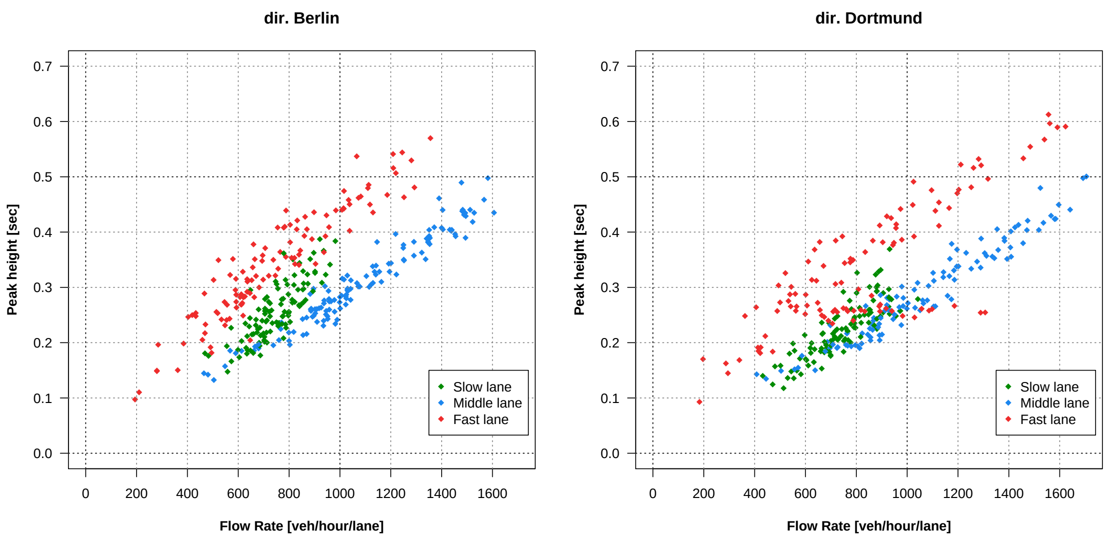

Figure 5.8: Peak height of empirical headway distribution

P-values

Firstly, it is important to point out how gamma-GQM has a great capacity to fit data and a high flexibility. P-values obtained by K-S test are really high, frequently close to 1, for sure over the 0.05 threshold.

In the analysis of a multi-sample data, it would be desirable to find a single measure to describe the overall acceptability of the null hypothesis: this is possible through the *combination of probabilities*, a statistical method well described by Luttinen [17]. In particular, we refer to the *moving probability*, where p-values are arranged in ascending order according to FR and the combined probability of k contiguous elements (with k here assumed equal to 9) is computed: the aim is to discover if the result of the goodness-of-fit is somehow dependent from FR. In Fig. 5.9 p-values are plotted in ascending order according to the FR, while the black line correspond to the moving probability.

Graphs reveal how no significant correlation exists between FR and the p-value of the K-S test for gamma-GQM.

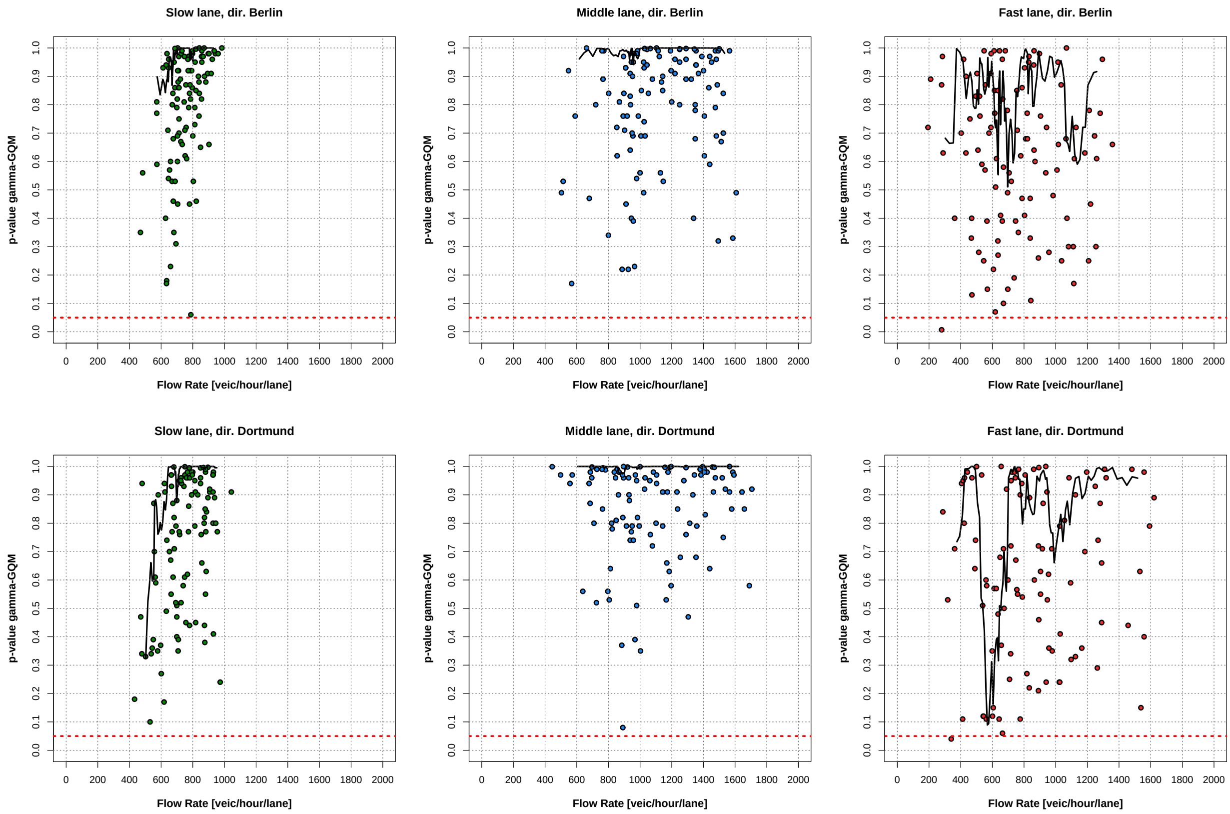

Figure 5.9: All lanes. Significance level for K-S test of gamma-GQM model and 9-point moving probabilities

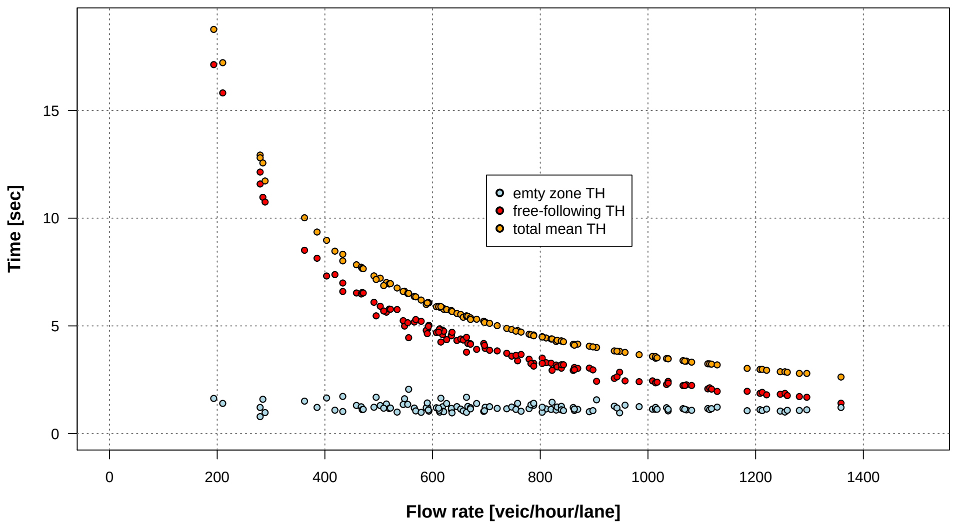

Figure 5.10: Role of the components of Eq. 5.3 at determining the mean headway

Mean of the distribution

Particular attention deserves the theoretical value of mean headway:

$$E(h) = E(U) + \frac{(1-\theta)}{\lambda} \tag{5.3}$$

where E(U) is the mean headway of the empty zone - derived from the gamma distribution - equal to α/β. Parameter λ represent the inverse ratio of the mean headway of the free-component V, present only for non-following headways. The role of both components at determining the mean headway could be easily understood through Fig. 5.10, in which both components are plotted in ascending order according to flow rate.

The mean of the empty zone is fairly stable and not influenced by FR. On the other hand, the second term decreases as the FR increases - this is due to the presence of λ and θ, which are sensitive to the traffic intensity.

Dependence from flow rate and lane type

Curves described in Tab. 5.3 are plotted in Fig. 5.11 with lane distinction. It is evident how FR affects considerably headway curve: peak height increases and the distribution arises, while the bell-shape becomes thinner. Moreover, differences through the lanes are observable in the curve width and height, as well as the mode value (around 2 seconds in the slow lane, lower in middle and fast lanes).

FR [vph]	**lane**	α	β	λ	θ	**p**
400-600	1	16.88	8.52	0.1906	0.135	0.62
600-800	1	13.24	6.4	0.2666	0.245	0.75
800-1000	1	13.44	6.34	0.3173	0.399	0.87
400-600	2	24.21	22.52	0.1757	0.046	0.59
600-800	2	7.2	5.51	0.2309	0.144	0.87
800-1000	2	15.2	15.32	0.3257	0.087	0.71
1000-1200	2	15.68	16.47	0.3836	0.136	0.84
1200-1400	2	21.6	22.5	0.4976	0.141	0.88
1400-1600	2	30.03	33.81	0.5888	0.146	0.79
200-400	3	15.76	16.13	0.0775	0.167	0.63
400-600	3	7.95	6.7	0.1206	0.319	0.65
600-800	3	9.11	7.92	0.1533	0.369	0.57
800-1000	3	9.48	8.36	0.1922	0.432	0.64
1000-1200	3	9.52	8.41	0.2254	0.497	0.55
1200-1400	3	11.09	10.25	0.2599	0.541	0.61

FR [vph]	**lane**	α	β	λ	θ	**p**
400-600	4	10.4	4.84	0.2094	0.108	0.49
600-800	4	9.41	4.75	0.2901	0.151	0.75
800-1000	4	8.21	4.15	0.3713	0.334	0.81
400-600	5	6.33	4.41	0.1627	0.153	0.97
600-800	5	7.26	6.4	0.2456	0.069	0.86
800-1000	5	7.58	6.63	0.3142	0.133	0.79
1000-1200	5	13.24	13.32	0.3908	0.249	0.83
1200-1400	5	9.04	8.24	0.5077	0.21	0.85
1400-1600	5	9.78	9.09	0.5718	0.26	0.92
1600-1800	5	10.49	9.53	0.6775	0.29	0.8
200-400	6	9.43	8.29	0.0813	0.209	0.53
400-600	6	7.49	6.36	0.1157	0.314	0.64
600-800	6	8.39	7.18	0.159	0.38	0.6
800-1000	6	8.67	7.44	0.1948	0.464	0.63
1000-1200	6	10.06	8.76	0.2362	0.501	0.59
1200-1400	6	9.71	8.37	0.2715	0.544	0.74
1400-1600	6	10.25	8.62	0.3141	0.633	0.63

Tab. 5.3: Values of gamma-GQM function

The scatter plot of gamma-GQM parameters against their corresponding flow rate is

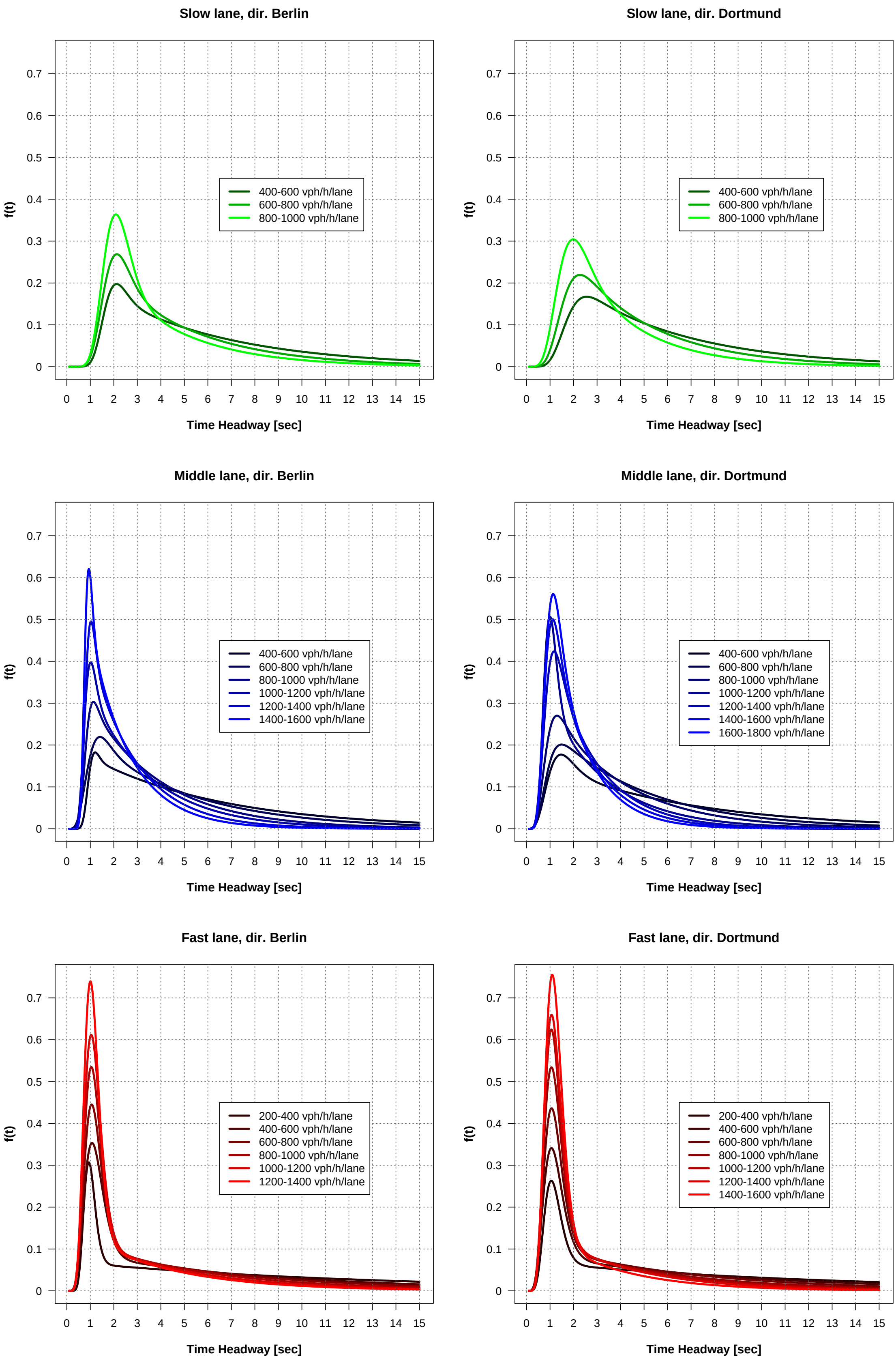

Figure 5.11: All lanes. Gamma-GQM models estimated by range of flow rate

depicted in Fig. 5.12, for Dortmund direction.

The parameter λ arises linearly as the FR increases, in fact λ is directly related to the flow rate in the exponential distribution. Independently of the lane, λ is gathered for low traffic volumes and more scattered for high volumes. On the other hand θ seems to be related to TV only in the fast lane, despite the points are scattered. On the other lanes the parameter varies widely, and the trend is not significative. This is consistent with the results of Ha et al. [10] in the French motorway, who founded that a slight tendency of θ can be observed only in fast lanes.

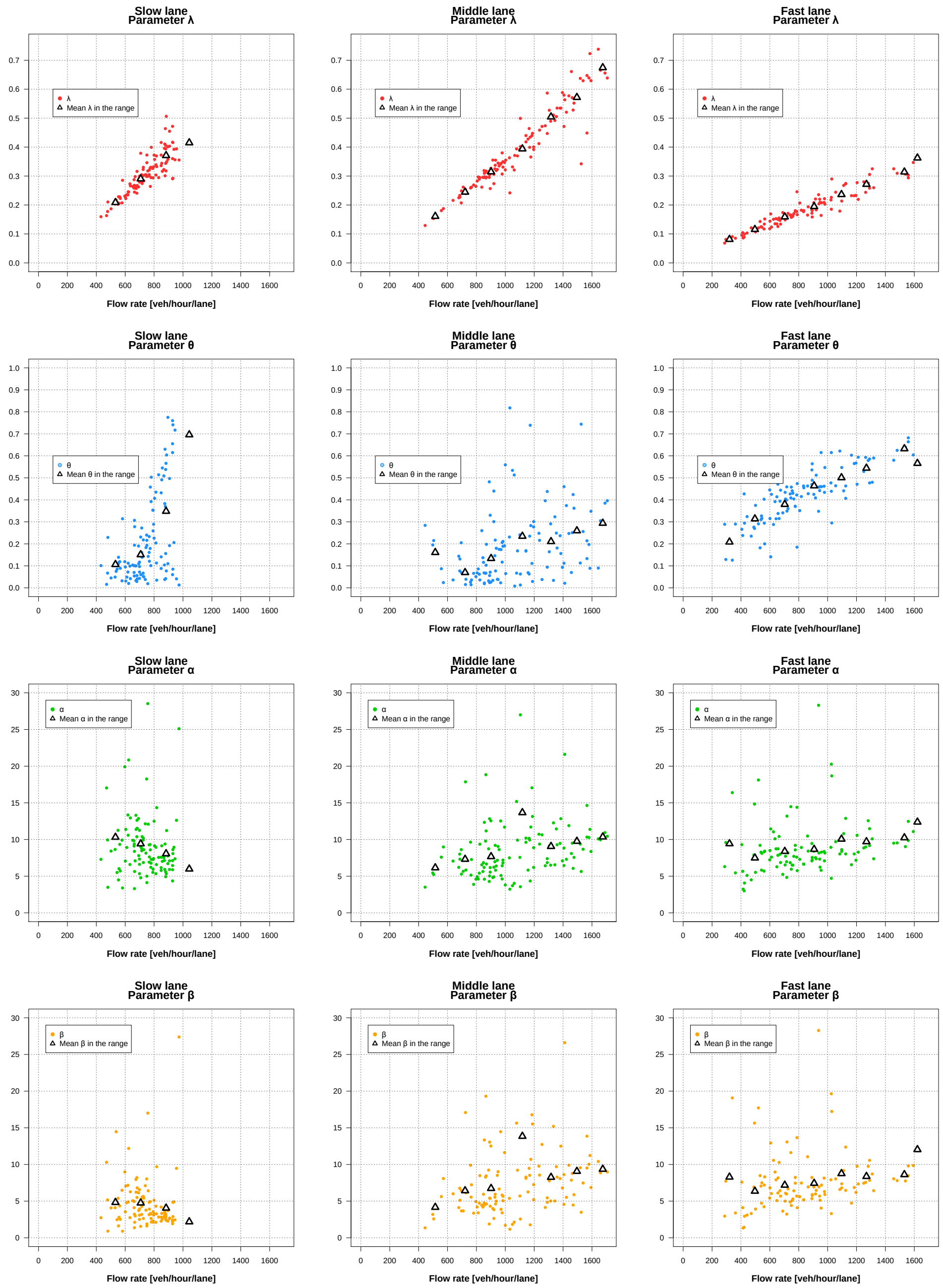

Figure 5.12: Plot of gamma-GQM parameters λ, θ,α and β against flow rate, dir. Dortmund

Even α and β reveal a tendency only in the fast lane, due to their high correlation with θ - this fact has been even noticed by [11]. In fact, Fig.5.13 compares different gamma-GQM distributions with the same value of FR: while λ is more stable for a similar value of FR, values of α, β and θ are more variable, and "small values of ϑ are associated with high values of α and β"[11].

The empty zone

The study of the parameters of the gamma function is useful for testing the model hypothesis of a fixed distribution *g(t)* for the empty zone; this assumption could be truthful, looking at the results in Tab. 5.4, since the mean and variance of the empty zone doesn't reveal any particular trend.

TV/LANE	1	2	3	4	5	6
200-400	-	-	0.98	-	-	1.14
400-600	1.98	1.08	1.19	2.15	1.43	1.18
600-800	2.07	1.31	1.15	1.98	1.13	1.17
800-1000	2.12	0.99	1.13	1.98	1.14	1.17
1000-1200	-	0.95	1.13	-	0.99	1.15
1200-1400	-	0.96	1.08	-	1.10	1.16
1400-1600	-	0.89	-	-	1.08	1.19
1600-1800	-	-	-	-	1.10	-

TV/LANE	1	2	3	4	5	6
200-400	-	-	0.06	-	-	0.14
400-600	0.23	0.05	0.18	0.44	0.33	0.18
600-800	0.32	0.24	0.15	0.42	0.18	0.16
800-1000	0.33	0.06	0.14	0.48	0.17	0.16
1000-1200	-	0.06	0.13	-	0.07	0.13
1200-1400	-	0.04	0.11	-	0.13	0.14
1400-1600	-	0.03	-	-	0.12	0.14
1600-1800	-	-	-	-	0.12	-

Tab. 5.4: Mean (left) and variance (right) of the empty zone

Beyond the independence from the FR, the ratio α/β seems approximately constant for a same lane (we will deal with this question later, with reference to TLTW roads). For the moment we only certify how this ratio reflects the type of road (or the lane type of the same road) under analysis; in Fig. 5.14, a linear regression is performed for all lanes in the direction of Dortmund, revealing a common attitude of the point in the diagram, then a close relationship.

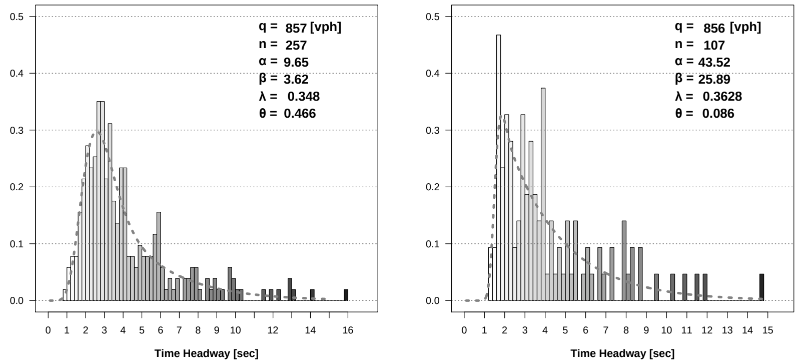

Figure 5.13: Estimated parameters for sub-samples with similar values of FR (slow line, dir.Berlin, Friday)

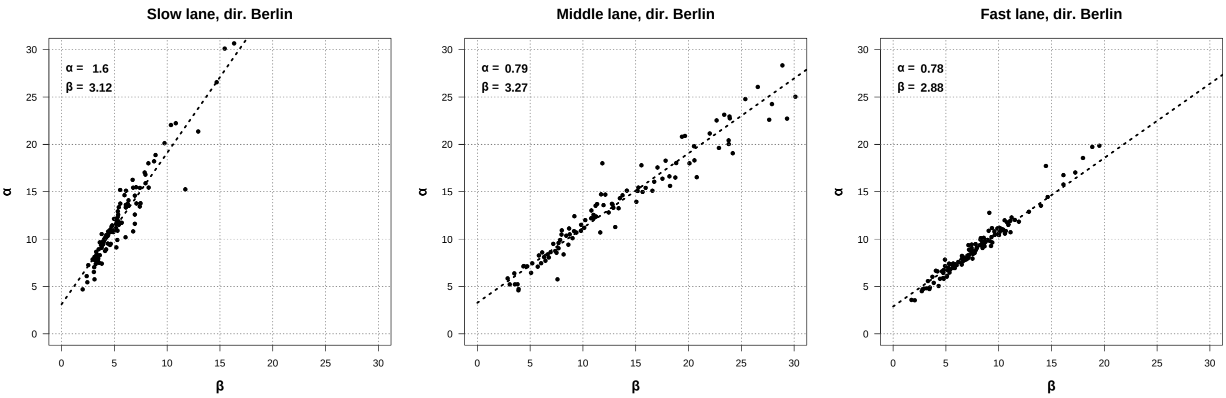

Figure 5.14: All lanes. Relation between α and β (slow, middle, fast)

5.2 Two-lane two-way roads

The Transportation Laboratory of the University of Padova is involved in a research project aimed at establishing an information system of traffic flow phenomena in the north-east Italy. The plan is to collect and manage traffic data on operations concerning interrupted and uninterrupted traffic flow conditions, in order to specify, calibrate and validate mathematical models of observed phenomena. Within this research project, this thesis seeks to identify TH probability density function for two-lane two-way road segments on different traffic situations. The data used for analysis of headway come from traffic surveys carried out on rural road network in Veneto, in particular:

- ATRs loop detector-based recordings in the province of Venice, coming from a continuous survey carried out by the Traffic Monitoring System for other research project;
- ATRs radar detector-based recordings in the province of Verona, coming from specific survey carried out by Transportation Laboratory of University of Padova.

The features obtained in the survey for each vehicle were road direction, arrival hours, arrival minutes, arrival seconds (to the tenth of a second), instantaneous speed and vehicle length; the aim is to obtain information not only about arrival time, but also concerning traffic composition and stability of traffic flow.

Traffic flow observations at a certain point (cross-section) of a road segment are useful in describing the traffic flow characteristics of the entire segment only if we accept the hypothesis that the segment is homogeneous in geometric and functional terms. Furthermore, cross-sections belong to similar road segments, having carriage width ranging from 6.80 to 7.40 meters; they are located in flat terrain and are perfectly straights. Traffic data were collected during good weather and dry surface conditions. In this way, exogenous conditions are assumed stable and don't affect TH distributions significantly[1].

[1] With the term *exogenous* we mean that the effect is not a component of the traffic flow and cannot be computed from traffic data [11].

Section	place	n° of days	Date	tecnology
A1	Bibione (VE)	1	Sat 29/08/2009	Inductive loops
A2	Trezze (VE)	1	Sat 29/08/2009	Inductive loops
A3	Calcroci (VE)	1	Fri 22/03/2013	Inductive loops
A4	Spinea (VE)	1	Thu 21/03/2013	Inductive loops
B	Verona	6	Thu 1/12/2011 - Tue 6/12/2011	Radar sensor
C	Lazise (VR)	6	Thu 5/4/2012 - Tue 10/4/2012	Radar sensor

Tab. 5.5: Duration and of traffic survey and technology used.

Section	n°of lanes	total vehicles	HV (%)	Directional Split (%)	Total samples
A1	2	9,551	1.4	51/49	2
A2	2	8,839	3.2	55/45	2
A3	2	4,517	6.3	48/52	2
A4	2	13,777	8.1	50/50	2
B	1	47,297	15.2	-	6
C	1	50,420	3.6	-	6

Tab. 5.6: Main characteristics of traffic data.

We hypothesize that the shape of TH distributions could depend on flow rate and on the flow composition (for the same flow rate), so we selected six cross-sections: sections A, B and F are characterized by a low percentage of HV, while sections C, D and E by a higher percentage. Overall, the traffic conditions represented cover a range of values from 100 to 1,100 veh/lane/hour. In some peak periods, meta-stable conditions were observed.

Tab. 5.5 show the time period and duration of the on-field surveys of traffic data collected for each cross-section examined, while and 5.6 illustrates the main characteristic of traffic flow and the total number of samples on hand.

Finally, we assume sections A1, A2, A3 and A4 as coming from a same hypothetical section named A, because of the homogeneity of geometrical and functional characteristic, as well as the technology of recording and the geographical context (province of Venice).

5.2.1 Selection of trendless samples

A total of 20 samples is submitted to the trend analysis. At first, data from inductive loops were selected for the period 7 a.m.-19 p.m. because headway longer than 10 minutes - only found in the night period - were not recorded by the sensor. Moreover, the statistical analysis is not interesting at very low traffic levels. On the other hand, no preliminary action has been taken for sections B and C, because data was always recorded by radar sensors: in this way we could test the trend analysis even in the night period, in presence of low traffic volumes. The graphical output of the trend analysis can be found in Appendix A by the same notations used for motorway; for the sections of Verona and Lazise, notice how the test rejects the night period (grey line), as expected. Characteristics of trendless intervals are resumed in Tab.5.7, 5.8 and Fig. 5.15.

Section	direction	n° of records	n° of intervals detected	mean time length [min:sec]
A1	1	4881	12	28:40
	2	4670	14	25:41
A2	1	4895	11	33:20
	2	3944	15	22:39
A3	1	2181	15	38:52
	2	2336	16	38:01
A4	1	6951	25	28:32
	2	6826	24	29:16

Tab. 5.7: Trendless sub-samples detected for A sections.

Section	Thu	Fri	Sat	Sun	Mon	Tue	total
B	34 [29:46]	34[31:31]	33[32:44]	27[34:30]	31[30:46]	32[31:10]	191
C	31[29:58]	31[31:14]	36[29:11]	31[31:45]	29[32:30]	27[32:59]	185

Tab. 5.8: N° of detected intervals and mean time length, sections B and C

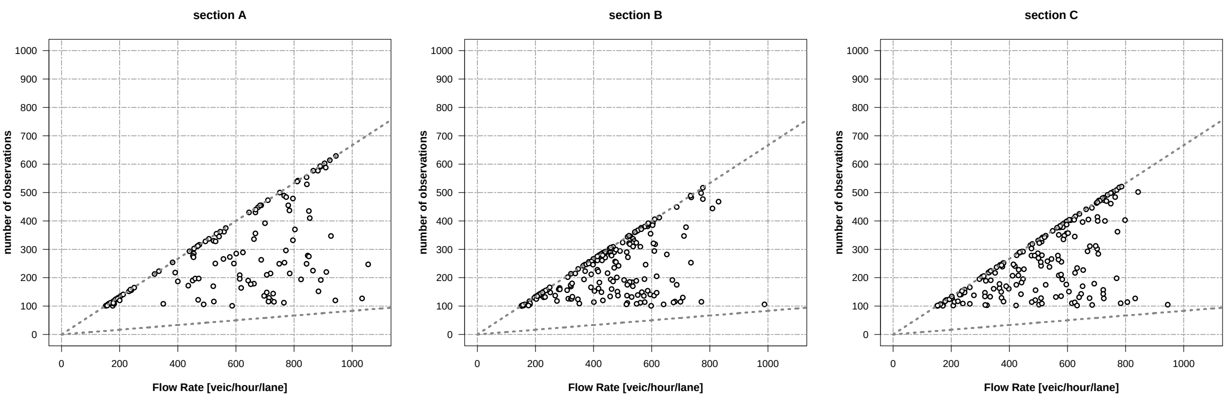

Figure 5.15: Flow rate and sample largeness for all the sections

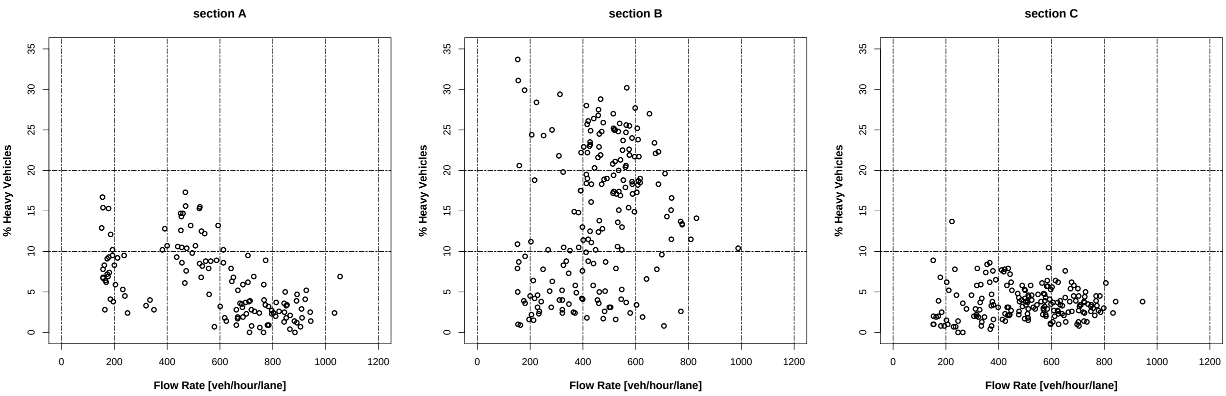

Figure 5.16: All sections. Numbers of trendless time intervals by range of FR and %HV

Each sub-interval was then classified according to its FR into six ranges (0- 200, 200-400, 400-600, 600-800, 800-1000, 1000-1200 veh/hour/lane) and to its %HV in three ranges (<=10%, >10% and <=20%, >20%). Fig. 5.16 shows the total number of trendless time intervals for all sections, classified by range of FR and %HV. Section B is the most complete one in both variables. On the other hand, section C present always low percentage of heavy vehicles (under 5%).

5.2.2 Fundamental diagrams

The traffic conditions observed at the cross-sections A are clearly described by the SMS/FR diagrams of Fig. 5.17: the set of pairs were estimated both with reference to the trendless time intervals and to 5-minute intervals.

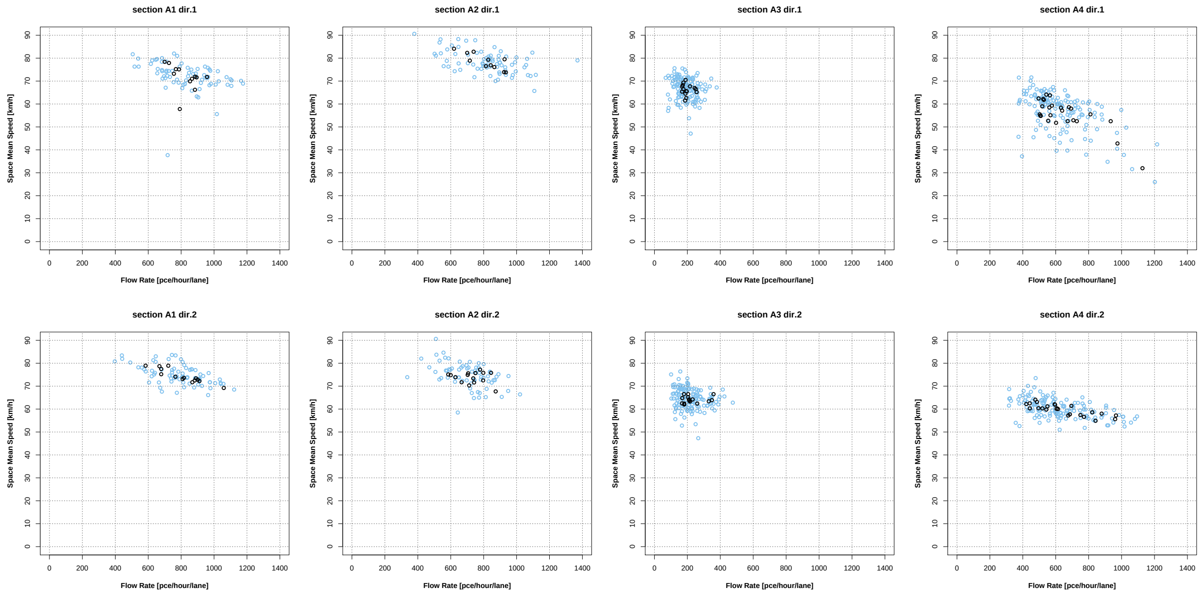

Figure 5.17: Sections A. SMS/FR pairs estimated for each trendless time interval (black rings) and for 5-minute time intervals (blue rings).

The distribution of points representing the state of the analyzed sections shows how the sampled time periods cover the whole domain of the flow rate, ranging from free-flow conditions to a value close to the capacity of the road segments (about 1,400 pce/hour/lane) at an average speed of 70-90 Km/h. The numerousness of the sub-samples for sections B and C permits to obtain a complete SMS-FR diagram, from free-flowing to metastable conditions, allowing a correct estimation of the real capacity of the roads (Fig. 5.18).

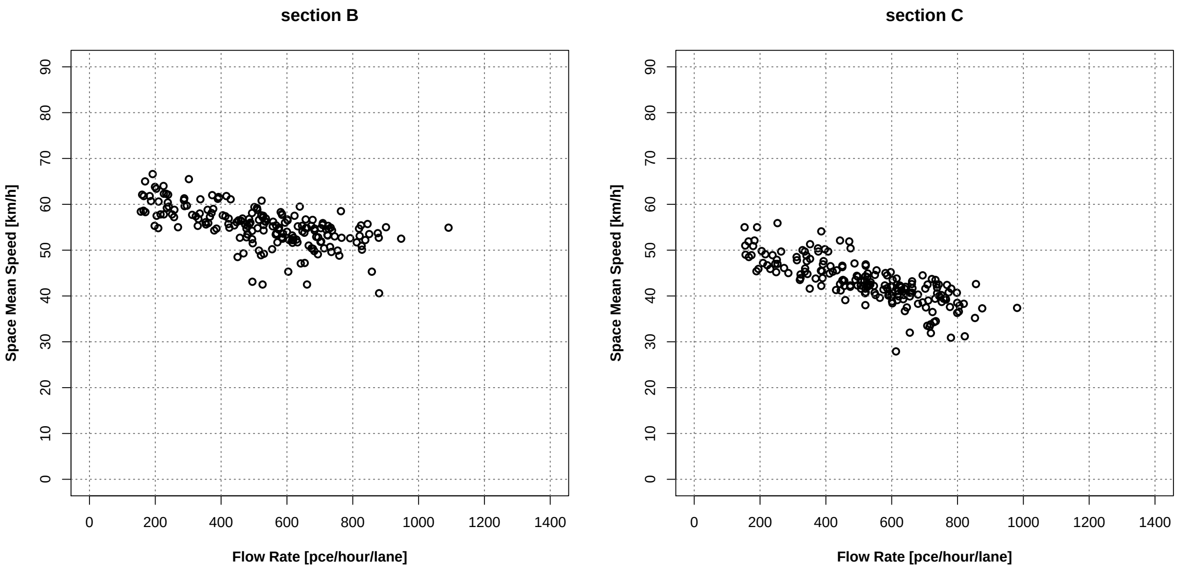

Figure 5.18: Sections B and C. SMS/FR pairs estimated for each trendless time interval

5.2.3 Mode and peak height

The mode and the peak heights were estimated using kernel estimation for density, with smoothing parameter equal to 0.2. There seems to be no significance correlation between mode values and flow rate (Fig . 5.19). On the other hand, Fig. 5.20 exhibit the tendency of peak height to increase as the flow rate arise. Both results are consistent with those obtained by Luttinen [17] that referred to two-lane two-way roads in Finland.

5.2.4 Estimation of time headway distribution

For each time intervals the parameters of the IW and the gamma-GQM distribution were estimated. Despite we know beforehand that gamma-GQM provides best fits thanks to his complex mathematical structure, the calibration of the IW has been carried out to recall the work done by the University of Padova about TH models in two-lane two-way roads [21], which came to the conclusion that IW was the most suitable model - among simple models - for this type of road to represent the real headway distribution. In order to evaluate the model performances and to make a comparison, models were calibrated by the same method (log-likelihood) and evaluated by the same goodness-of-fit test (K-S).

P-values

The results of goodness-of-fit test confirm that gamma-GQM fits data better than IW. We display the performances of the models through the associated p-value of the K-S test, which express the significance level (SL). To investigate factors which could cause a bad fitting for IW, in Fig. 5.22 we plotted the pair FR/number of observations of every

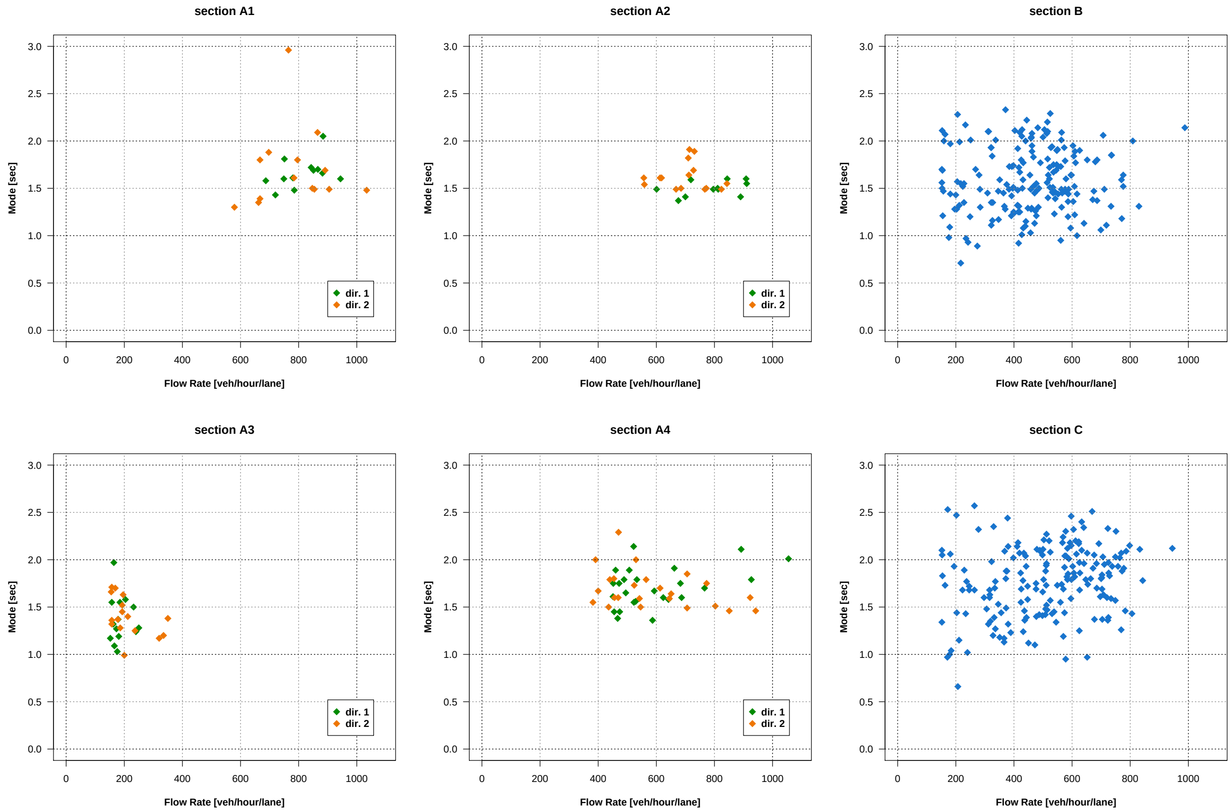

Figure 5.19: All sections: mode of empirical headway distributions

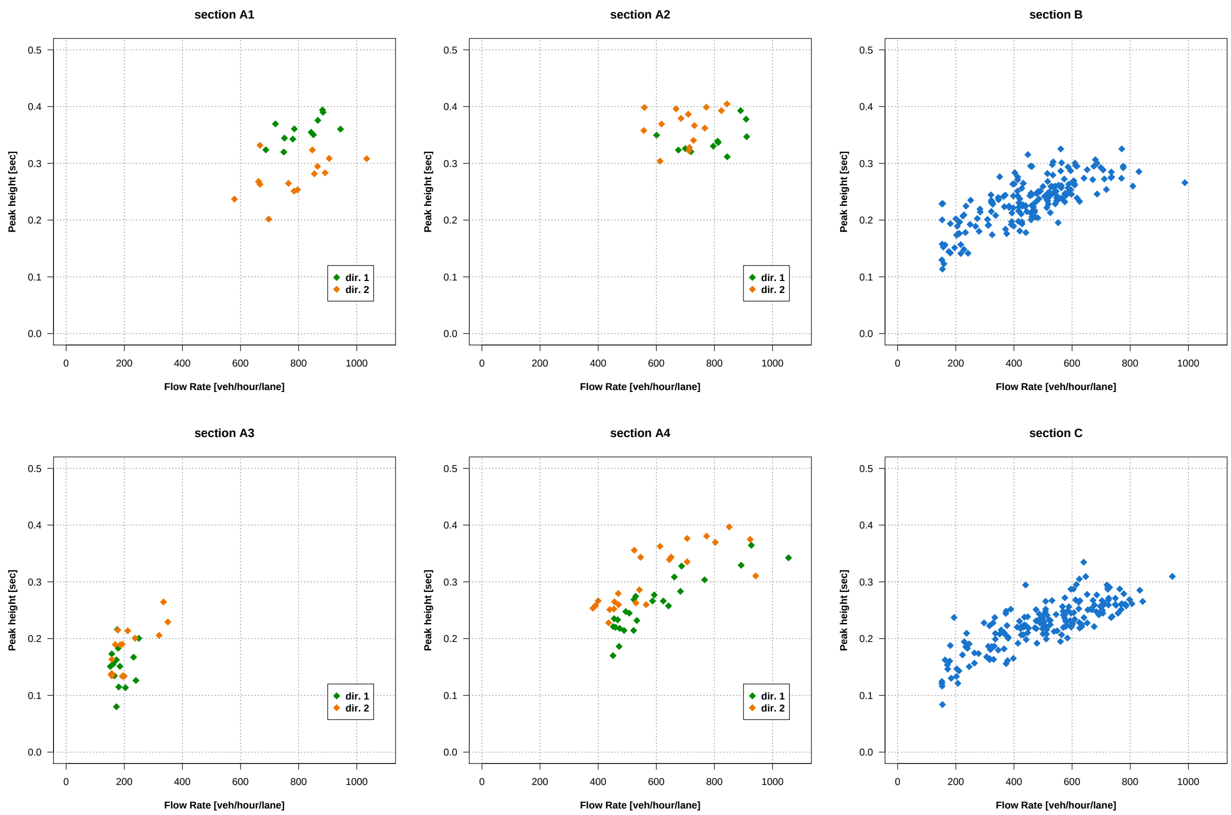

Figure 5.20: All sections: peak height of empirical headway distributions

sub-sample. It seems that IW has difficult calibrating the model for samples of more than 300-350 records; moreover, bad fitting were obtained in low traffic conditions.

On the other side, gamma-GQM fits data extremely well and is significant for all samples. A possible correlation of the SL with FR is investigated by the method of *moving probability* (Fig. 5.23): it seems that the SL lightly decreases as the FR increases but the fits, however, is always excellent.

This result is also strengthened from the qualitative point of view when we look at Fig. 5.24, which compares the gamma-GQM model, IW model and Kernel density estimation of the observed THs (we have chosen three intervals in the range 600-800 veh/hour/lane of high sizes)

All the estimated parameters obtained by calibration are then classified on the strength of the class of FR [vph] and %HV, like in the subdivision operated in Figure 5.16. For each box the average values for each parameter are calculated, in order to identify a representative set of values for each FR/%HV class. From a graphical point of view, this operation consists in finding the dotted black curve in Fig. 5.21 starting from the set of light distributions.

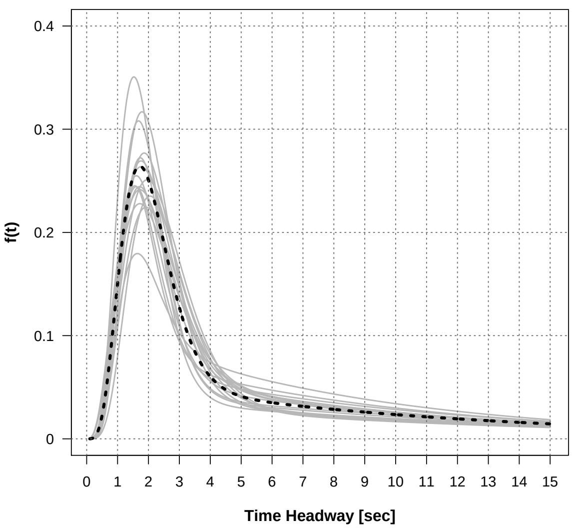

Figure 5.21: Average curve (black dotted line) from a set of pdf's (light solid lines)

Flow rate dependence

As written by Ha et al. [11], FR is the first and the most important factor affecting the scale values of TH; this fact is confirmed by the tendency of the parameter λ, computed in each FR class as the average value among all the sub-samples (Tab. 5.9, 5.10, 5.11). Notice how, for increasing FR:

- the mean decreases, in fact it is the inverse of the FR;

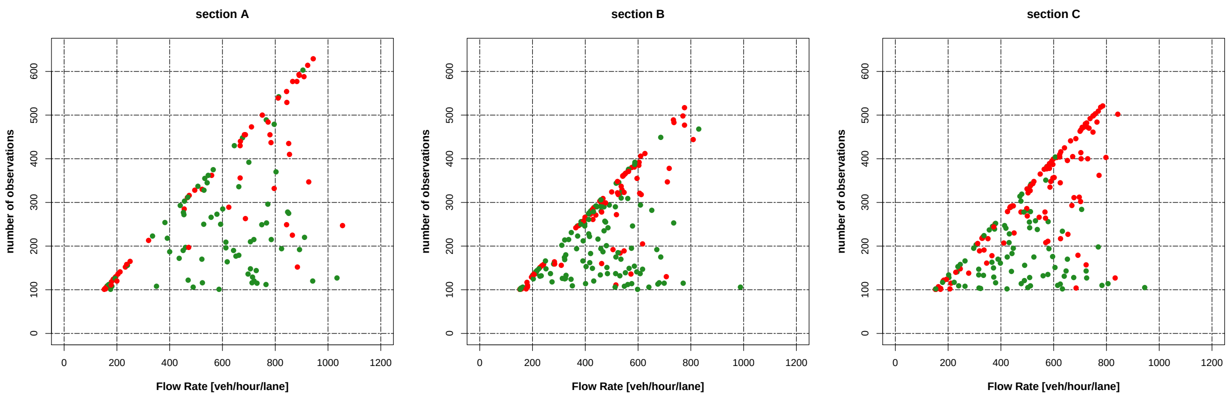

Figure 5.22: All sections. Acceptability of the IW fitting for each time interval, identified by the pair FR/number of observations: SL≥0.05 (green) and SL<0.05 (red)

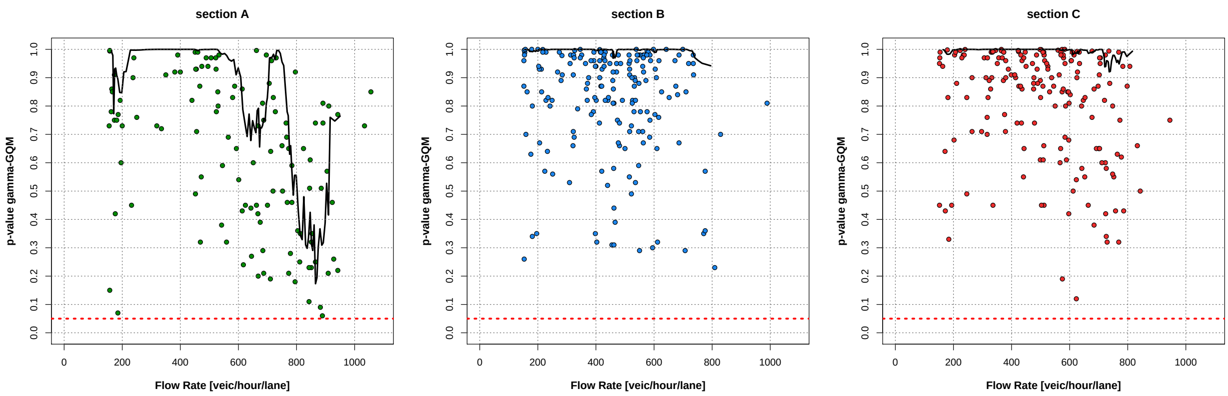

Figure 5.23: All sections. Significance level for K-S test of gamma-GQM model and 9-point moving probabilities

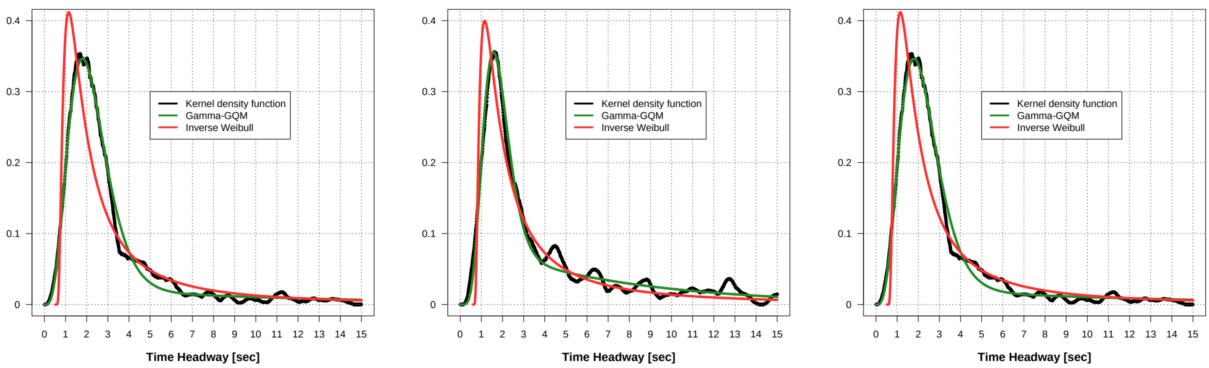

Figure 5.24: TH sample corresponding to trendless time interval with high size, 600≥FR>800. Kernel density estimation, gamma-GQM and Inverse Weibull pdf's.

- the variance decreases. The influence between vehicles arises, then headways becomes more gathered;
- the mode increases, at least in stable traffic conditions. Note that this fact was non evident in Fig. 5.19.

From a graphical point of view the curves raise and the bell-shape becomes more accentuated, and furthermore, after reaching the peak, the distribution slopes down more rapidly (see Fig. 5.25).

FR range	n° of sub-intervals	α	β	λ	ϑ	mean	variance	mode
0-200	22	6.37	4.16	0.0388	0.266	20.42	486.64	1.32
200-400	11	5.45	3.32	0.0587	0.357	12.6	187.19	1.38
400-600	31	5.76	2.99	0.0987	0.48	7.2	54.09	1.65
600-800	41	5.69	2.92	0.1182	0.628	5.1	27.3	1.64
800-1000	25	5.54	2.82	0.1464	0.671	4.21	16.06	1.65
1000-1200	2	5.5	2.45	0.2263	0.721	3.48	6.38	1.89

Tab. 5.9: Section A. Gamma-GQM models parameters for all flow rate ranges [veh/hour/lane].

FR range	n° of sub-intervals	α	β	λ	ϑ	mean	variance	mode
0-200	14	2.69	1.54	0.0324	0.352	21.74	618.08	1.14
200-400	49	2.61	1.42	0.0526	0.469	11.92	193.06	1.18
400-600	97	2.57	1.29	0.0771	0.597	7.21	69.29	1.26
600-800	28	2.46	1.24	0.0934	0.69	5.3	37.12	1.22
800-1000	3	2.32	1.15	0.1325	0.719	4.14	17.74	1.2

Tab. 5.10: Section B. Gamma-GQM models parameters for all flow rate ranges [veh/hour/lane].

FR range	n° of sub-intervals	α	β	λ	ϑ	mean	mode	variance
0-200	11	2.71	1.5	0.0366	0.291	21.21	1.21	532.07
200-400	43	2.53	1.29	0.0575	0.444	11.63	1.25	169.72
400-600	79	2.58	1.21	0.0857	0.583	7	1.37	58.63
600-800	58	2.88	1.24	0.1113	0.679	5.2	1.58	27.81
800-1000	4	2.2	1	0.1376	0.724	4.22	1.26	16.79

Tab. 5.11: Section C. Gamma-GQM models parameters for all flow rate ranges [veh/hour/lane].

As made for motorway, we report all the estimated value of the parameter with the aim to analyze their sensitivity and their variance with respect to the FR. The parameter λ, as founded for motorway, is strictly related with FR because it represents the arrival intensity of the free-following vehicles; it plays an important role in fitting the decreasing slope of long TH. It becomes more scattered as FR increase 5.26. Even ϑ is related to FR but not by a linear dependence like λ - it seems that it comes closer to the 1 value asymptotically . Moreover, both λ and ϑ appear to be more sprinkled in the A sections, probably because the geometrical and physical homogeneity among the sections is not absolute.

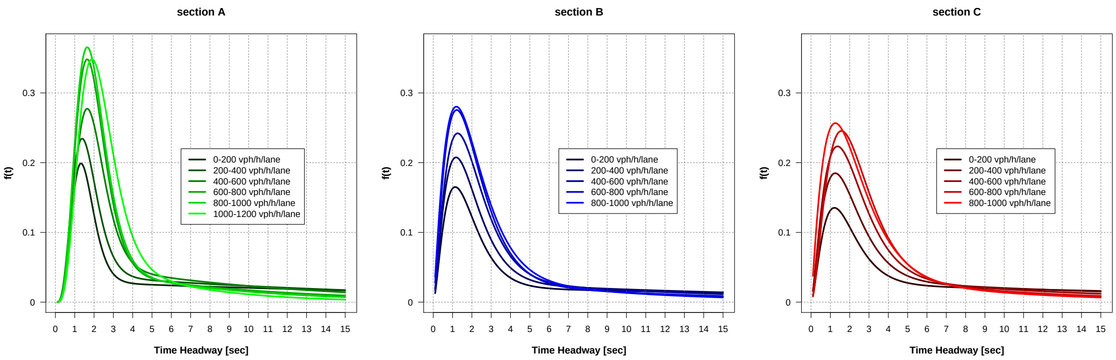

Figure 5.25: All sections. Gamma-GQM models estimated by range of flow rate

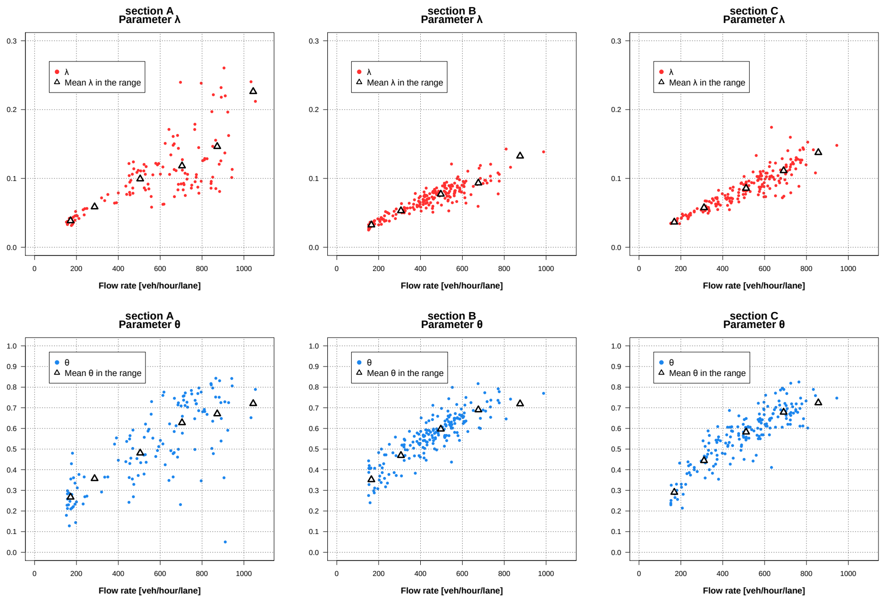

Figure 5.26: Variations of gamma-GQM parameter λ and ϑ according to FR

The study of the empty zone, then the *g(t)* function, deserves a special attention. At first, we show the punctual trend of α and β in Fig. 5.27: parameters appear scattered and without any tendency, substantially independent from FR. We report in Tab, 5.12 all the values of mean and variance of the empty zone, drawn in Fig. 5.28.

section A			*section B*			*section C*		
FR range	mean	variance	FR range	mean	variance	FR range	mean	variance
0-200	1.53	0.37	0-200	1.75	1.14	0-200	1.81	1.2
200-400	1.64	0.5	200-400	1.83	1.29	200-400	1.96	1.52
400-600	1.93	0.64	400-600	1.99	1.54	400-600	2.13	1.76
600-800	1.95	0.67	600-800	1.98	1.59	600-800	2.32	1.86
800-1000	1.96	0.7	800-1000	2.02	1.75	800-1000	2.21	2.22
1000-1200	2.24	0.92						

Tab. 5.12: All sections. Mean and variance of the empty zone for different FR ranges.

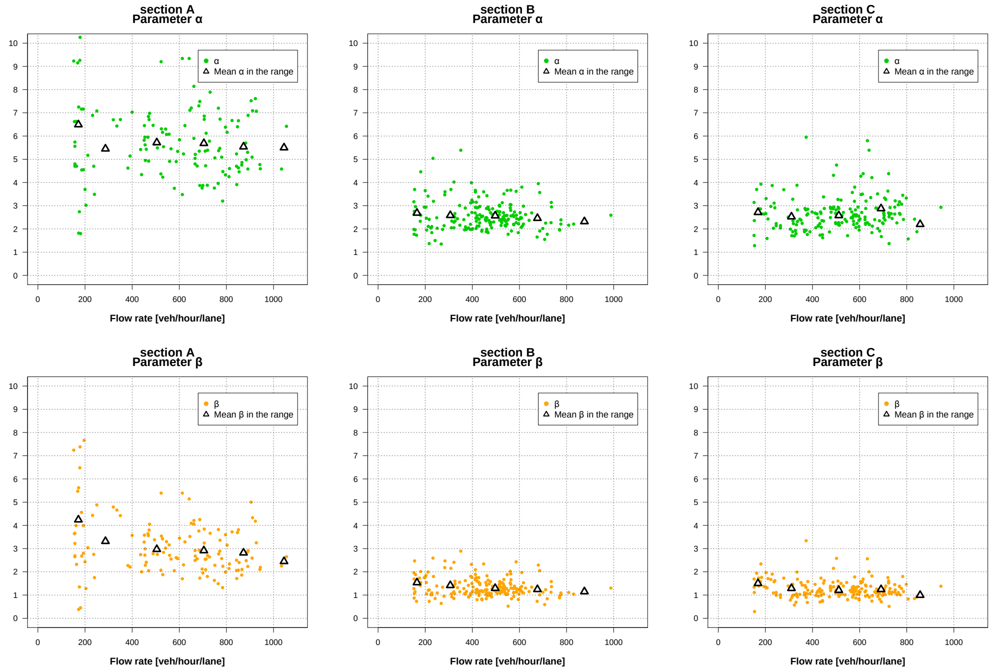

Figure 5.27: Variations of gamma-GQM parameter α and β according to FR

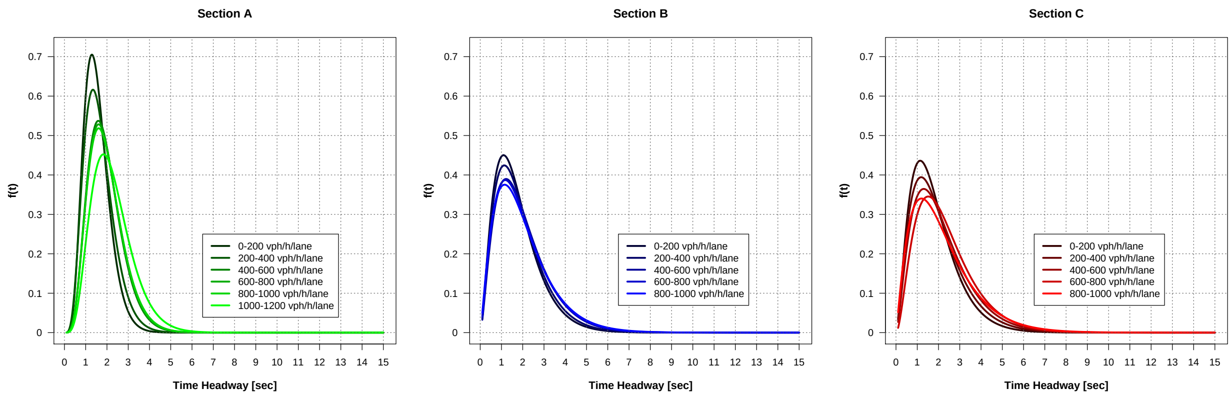

Figure 5.28: Empty zone distribution for different FR ranges.

Note that mean and variance arise whit FR; from a graphical point of view, we observe lowering in the peak height, while the curve moves to the right. This output originates two main problems:

1. it appears to contradict the model hypothesis of a fixed distribution for tracking headway, independent of the FR or traffic flow level;
2. if capacity could be estimated by the mean of the empty zone, we would have different capacity values according to the FR. In our case, it would decrease as FR arises.

About the last point, some elucidations must be given. The estimate of the capacity of a road through:

$$C = \frac{3600}{E(X)} \tag{5.4}$$

proposed by Minderhoud et al. [19], holds as reference value the mean of the empty zone of mixed models E(X), both for the Semi-Poisson and the GQM. This way capacity of a road would require only a sample of time headway at a given cross section to be estimated. Applying 5.4 to our case, we would obtain the results in Tab. 5.13.

section A		*section B*		*section C*	
FR range	capacity	FR range	capacity	FR range	capacity
0-200	2353	0-200	2061	0-200	1994
200-400	2191	200-400	1966	200-400	1837
400-600	1868	400-600	1813	400-600	1688
600-800	1846	600-800	1820	600-800	1555
800-1000	1833	800-1000	1784	800-1000	1628
1000-1200	1604				

Tab. 5.13: Capacity estimation of the road through the empty zone distribution.

Minderhoud et al. themselves [19] explain that those values of capacity estimated by 5.4 might be interpreted as theoretical values instead of practical capacity values. Estimation of capacity through GQM model “leads to a substantial overestimation of the observed road-capacity, probably caused by the implicit assumption that the distribution of the constrained headway at capacity level can be compared with the distribution at below capacity level”[19]. Values observed in Tab. 5.13 are certainly high for a two-lane two-way road, whose real capacity could be approximately 1200 veh/hour/lane.

Basically, an estimation of capacity when vehicle are not almost all followers couldn’t be a sincere, and this could happen only for $\theta > 0.9$, where congestion is already likely to occur.

Minderhoud et al. concluded that “capacity estimation based on headway models does not seem to be the best solution to derive a reliable design capacity value”.

Dependence from the percentage of heavy vehicles

In headway literature some authors have analyzed the differences in headway distribution among different vehicular categories, mainly discerning from passenger cars and trucks (heavy vehicles). Mostly, they confronted the problem by separating passenger headway from truck headway, and analyzing data separately.

Hoogendorn and Bovy [15] studied accurately headway distribution for different vehicle types using gamma-GQM model with location parameter d. They distinguished headway from passenger cars, articulated and non-articulated vehicles, and calibrated the model according to the type-specific headway. They found important results in the variation of the gamma-GQM parameter.

Ye et Zhang [25] did an accurate vehicle type-specific headway study on the basis of 24-h traffic data from I-35 in Austin (Texas), discerning headways according to the previous and consequent driver type. The first result was that “for all flow conditions, the truck-truck headway is the largest and the car-car headway is the smallest [...] because of the longer size of trucks and the lesser braking capabilities of trucks” [25]. According to the traffic flow condition - congested or uncongested situation - they found different suitable headway distribution. The paper demonstrates the necessity for a vehicle type-specific headway analysis for all traffic flow conditions with the exception of a very low traffic flow level, where differences are not clear. Variances should then be found at higher traffic flow levels.

Ai et al. [2] analyzed headway data in different traffic conditions discerning on previous and consequent driver type for each headway, finding differences in truck-car and truck-truck headway distributions when traffic shifts between congested and uncongested conditions.

Zala et al. [26] divided in car-follows-car and truck-follows-truck headway on the National Highway-8 (NH-8) in India, and analyzed data separately. They found that truck-follows-truck have higher headway range and headway mean than car-follows-car (4.8 against 2.78).

In substance, the method used by all authors in the approach to vehicle-type-specific headway distribution has always been:

1. Record all headways that cross a section, with the vehicle-type specification;
2. Subdivide headway according to the vehicle type (some authors took into account even the type of vehicle of the leader that the vehicle is following, creating different previous-successive combinations [25][26]);
3. analyze distributions separately and calibrate the models.

This way the separation between vehicle type would provide “additional insight into the plausibility of the headway distributions and parameter values and into the car-following behavior of distinct vehicle classes varying across different periods” [15].

In order to create a joint headway distribution comprehensive of all vehicles (in different percentages), the approach is to amount the probabilities through an *aggregate* model. Hoogendorn and Bovy [15] created an aggregate headway distribution using gamma-GQM model. They investigated whether vehicle-type segregation improves the descriptive accuracy of the model with respect to real-life sample, and they founded that the distribution estimates were virtually identical and the K-S distance practically the same. Therefore, this approach doesn’t improve headway modeling, but permit to distinguish and appreciate the differences between vehicle-type-specific headway distributions.

Despite this thesis doesn’t estimates headway distribution according to vehicle type, it is useful to remind some results obtained by other authors about type-specific pdf shapes.

Hoogendorn and Bovy [15] found different headway distributions for passenger cars, non-articulated trucks and articulated trucks working on two-lane one-way roads, working on 2-hours samples and separating headway belonging to different vehicle types subsequent the headway. Then they analyzed distributions separately, estimating gamma-GQM parameters. One of the results was that articulated trucks have a bigger mean and variance than cars (respectively: 2.90 vs 1.57 and 2.08 vs. 0.64), while articulated truck behavior is more similar to that of cars (respectively 1.63 and 0.74). From a graphical point of view, we report in Fig. 5.29 headway distributions for different vehicular categories:

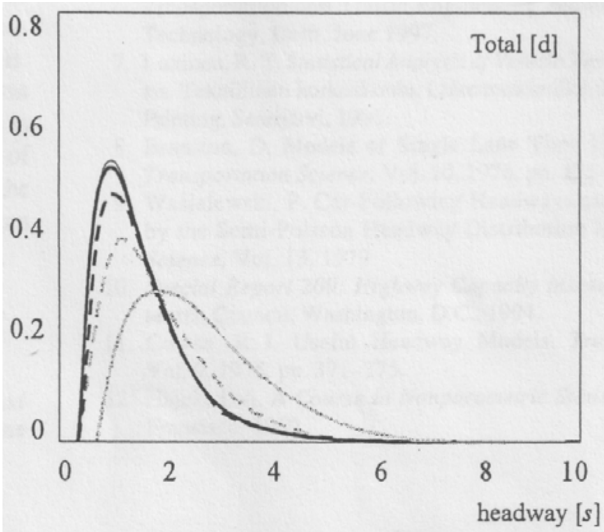

Figure 5.29: Gamma-GQM for different vehicular classes, Hoogendorn and Bovy [15]

Articulated trucks (light solid line) and non-articulated trucks (black dotted line) headway distributions are more shifted to the right in comparison with passenger car (black solid thin line). Their resulting distribution is represented by light dotted line (total trucks). The total headway distribution (black solid thick line) traces substantially the passenger car TH distribution, with only a little decrease of the peak height. Note that the truck percentage is 11%.

In this thesis we are not interested on analyzing variations of headway distribution for a different vehicular class, but we aim to test if the increasing of the %HV affects the shape of the distribution prominently, for a same value of FR.

The benefit will be found for example in microsimulation models, where vehicles are generated by the software: setting the parameter of FR and %HV, we could select the appropriate parameters of the gamma-GQM model as the proper probabilistic distribution.

For this reason, we kept fixed the FR class subdivision [200 vph], adding a classification based on the %HV. This was possible for section A (thresholds: 5% and 10%) and section B (thresholds: 10% and 20%); in fact section C owns a derisory %HV.

Tab. 5.14 and 5.15 resume all the results, while all gamma-GQM curves are drawn in Fig. 5.30 and 5.31 for each flow rate range.

FR range	HV%	n	α	β	λ	ϑ	mean	variance	mode	ph
0-200	≤5	3	7.77	4.85	0.0439	0.217	19.43	406.44	1.44	0.17
0-200	>5 et ≤10	13	5.91	3.88	0.0382	0.28	20.34	492.56	1.3	0.2
0-200	>10	6	6.66	4.42	0.0376	0.261	21.16	522.83	1.31	0.2
200-400	≤5	5	6.08	4.1	0.0657	0.312	11.96	159.85	1.28	0.24
200-400	>5 et ≤10	4	4.95	2.88	0.0471	0.322	16.11	305.9	1.42	0.19
200-400	>10	2	4.88	2.24	0.0644	0.539	9.34	112.3	1.77	0.25
400-600	≤5	2	4.33	2.38	0.0901	0.597	6.29	50.46	1.43	0.32
400-600	>5 et ≤10	12	6.01	3.26	0.1028	0.467	7.03	51.03	1.59	0.29
400-600	>10	17	5.76	2.87	0.0967	0.475	7.43	56.76	1.72	0.26
600-800	≤5	29	5.2	2.62	0.1178	0.64	5.04	26.67	1.64	0.34
600-800	>5 et ≤10	11	6.65	3.48	0.1194	0.604	5.22	28.31	1.66	0.37
600-800	>10	1	9.34	5.39	0.1166	0.525	5.81	35.26	1.58	0.41
800-1000	≤5	24	5.48	2.8	0.1457	0.669	4.23	16.3	1.64	0.36
800-1000	>5 et ≤10	1	7.07	3.25	0.1623	0.725	3.87	11.11	1.9	0.39
1000-1200	≤5	1	4.58	2.25	0.2405	0.652	3.48	6.92	1.67	0.33
1000-1200	>5 et ≤10	1	6.42	2.65	0.212	0.789	3.42	5.61	2.08	0.37

Tab. 5.14: Section A. All flow rate ranges (veh/hour/lane). Gamma-GQM models parameters and characteristics by range of percentage of heavy vehicles (thresholds: 5% and 10%).

FR range	HV%	n	α	β	λ	ϑ	mean	variance	mode	ph
0-200	≤10	9	2.67	1.42	0.0326	0.366	21.31	597.37	1.22	0.16
0-200	>10 et ≤20	1	1.98	1.28	0.0327	0.275	23.72	679.23	0.82	0.14
0-200	>20	4	2.91	1.86	0.0319	0.339	22.29	650.64	1.06	0.18
200-400	≤10	30	2.61	1.44	0.0522	0.458	12.21	200.64	1.17	0.21
200-400	>10 et ≤20	12	2.58	1.36	0.0551	0.505	10.87	164.10	1.21	0.22
200-400	>20	7	2.66	1.46	0.0502	0.457	12.64	216.53	1.18	0.21
400-600	≤10	20	2.51	1.22	0.0783	0.584	7.37	69.52	1.29	0.23
400-600	>10 et ≤20	34	2.68	1.4	0.0751	0.604	7.19	71.62	1.24	0.26
400-600	>20	43	2.51	1.24	0.0782	0.598	7.17	67.47	1.27	0.24
600-800	≤10	7	2.45	1.24	0.0941	0.689	5.28	36.74	1.21	0.28
600-800	>10 et ≤20	14	2.46	1.22	0.0984	0.687	5.19	33.98	1.24	0.27
600-800	>20	7	2.49	1.29	0.0827	0.698	5.58	45.7	1.19	0.29
800-1000	>10 et ≤20	3	2.32	1.15	0.1325	0.719	4.14	17.74	1.2	0.28

Tab. 5.15: Section B. All flow rate ranges (veh/hour/lane). Gamma-GQM models parameters and characteristics by range of percentage of heavy vehicles (thresholds: 10% and 20%).

The visual outcome is plain: there is no clear tendency doesn't exist for both study cases. The separation in FR classes has been made in order to nullify the effect of this parameter on headway distributions, as visible in Tab. 5.14 and 5.15, each %HV class has a different mean headway due to the scarcity of data in the sub-samples. This way the influence of this parameter cannot be completely nullified thereby concealing the %HV variable. In fact not only the standard deviation, but even the parameter λ (and θ to a

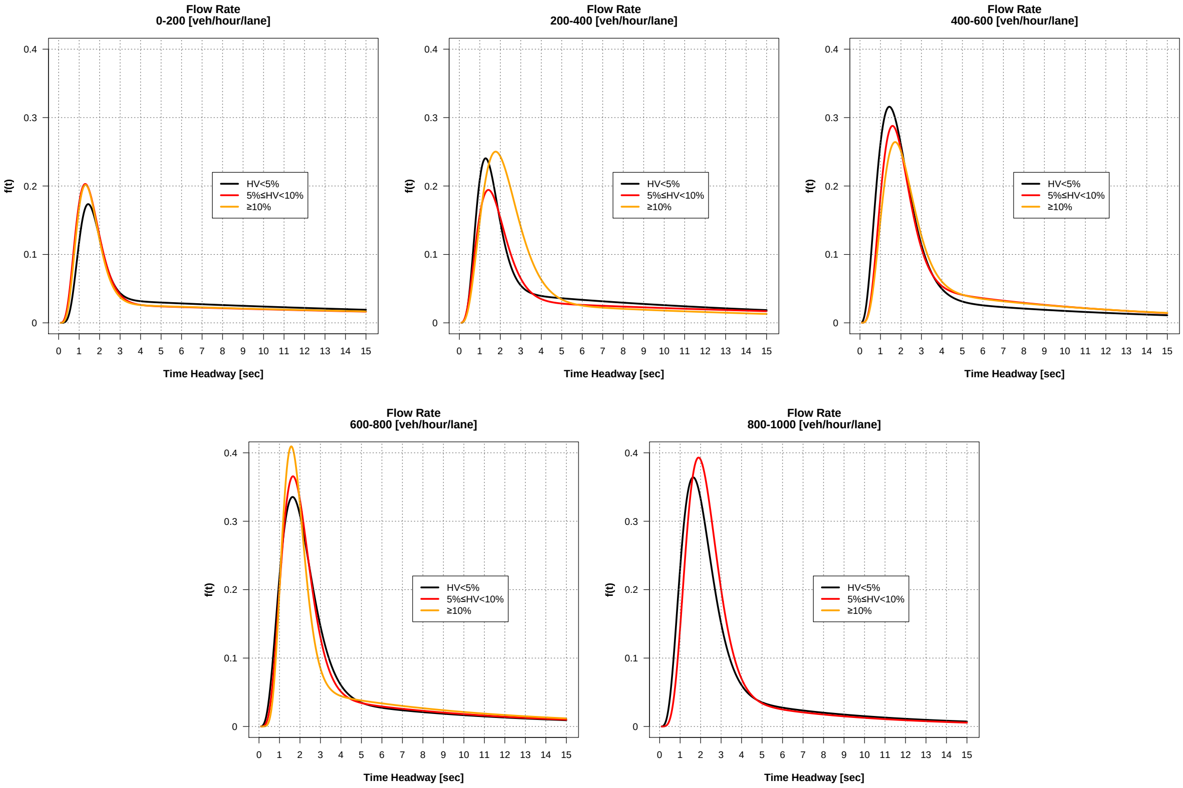

Figure 5.30: Section A. Gamma-GQM models estimated by range of percentage of heavy vehicles for alla FR ranges.

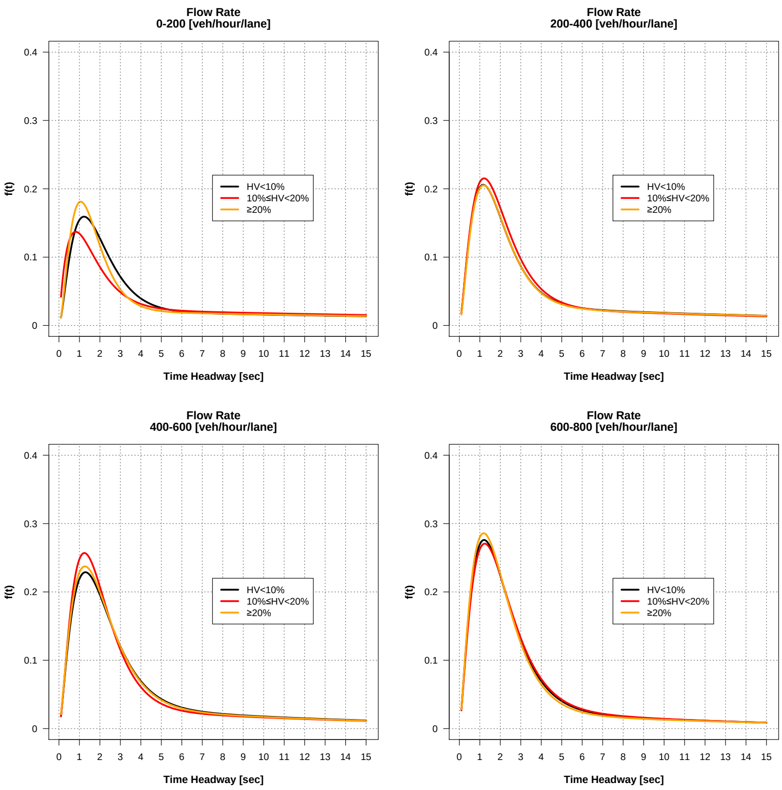

Figure 5.31: Section B. Gamma-GQM models estimated by range of percentage of heavy vehicles for alla FR ranges.

lesser extent) are extremely sensitive to FR.

Furthermore, all the parameters of the gamma-GQM are very highly correlated and very sensitive to other macroscopic variables like traffic occupancy and SMS, according to the study of Ha et al. [11]. In this sense, the percentage of heavy vehicle seems not to be relevant in this study.

5.3 Critics and considerations about gamma-GQM model

The application of the gamma-GQM distribution yielded excellent results in terms of goodness-of-fit, both on the A2 motorway and two-lane two-way roads; the model fits data well in any traffic condition and on different type of roads. Some question are however still unclear, like the traffic flow meaning of each parameter and the foundation of the mathematical assumption.

Firstly, we point out the high correlation between all parameters. Results obtained by calibration showed that "the low value of ϑ is usually associated with particularly high values of α and β"[11]. Therefore the interdependence between α, β, and ϑ is evident.

This is one reason which doesn't allow to interpret the different estimates of the parameters in a straightforward way from a traffic-flow theory point of view, because "interdependence of the parameters can cause disturbances when real-life TH data are estimated" [11]. For this reason, the study of the relationships between macroscopic variables and TH distribution - done for example by Ha et al. [11] - is very strenuous and complicated, and deserves more-detailed study.

To this uncertainty we add up the expression of model and its basic assumption, criticized by many authors [24, 14]. As explained in the paragraph 4.1.3, it is criticized the fact that a free-flowing driver takes tracking headway into consideration. In regard to this, it is still not clear if the *g(t)* distribution "describes the distribution of the tracking headway or [..] the tracking behavior of the various participants in the traffic process[14]". In addiction, Ha et al. [11] wrote that ϑ concerns the participating part of following vehicles, but interpreting it only in this way is not enough.

Still about *g(t)* function, the more or less arbitrary choice of the distribution "does not realistically model the behavior of tracking drivers at different flow rates" [14]. We notice that:

- On A2 motorway in Germany, we haven't found significant variation in mean and variance, and the empty zone seems stable;
- Ha et al. [11] founded on A6 motorway in France that the mean of the empty zone was fairly stable, despite the variance decreases according to the traffic volume;
- Hoogendorn and Botma [14] founded that the capacity of the road through eq. 5.4 decreases (than the mean headway of the empty zone increases) as FR arises on

two-lane one-way rural road in Nederland;
- We found a significant increase in mean and variance on two-lane two-way roads in Italy.

It seems that flow rate has a considerable effect on the distribution of the empty zone when speed of the road is lower and overtaking is often unfeasible; the increasing of FR on two-lane two-way roads causes an increasing attention to driving bigger than in motorway, because the lane is tighter and there is an influence of the vehicles in the opposite directions.

However, despite the weakness of the model from a strict mathematical point of view and a consequent incomplete knowledge of the physical meaning of the parameters, the model fits data at best among all distributions in terms of goodness-of-fit, and it is assumed to be the more complete and reliable among all.

Chapter 6

Conclusions

The aim of this work was to enhance the methodology of analysis of headway, by executing a proper study of this variable on two-lane two-way roads. This work has shown that:

- The research of the stationary condition is an advantage not only in headway studies but also in the study of capacity through fundamental diagrams. Some modification to the Luttinen's procedure was possible thanks to higher computational capacity of modern technology, thereby allowing a more detailed inspection of trendless conditions. The R programming language was especially useful in its simplicity and power.
- The gamma-GQM model, due to its deficiency of a solid mathematical base and a strict correlation through the variables, doesn't allow a direct comparison of its parameters through a traffic engineering interpretation. The fit is, however, extremely good, for confirming previous results in literature about gamma-GQM to be the best model for headway data.
- This study has confirmed the tendency of the distribution to rise as the flow rate increases thereby allowing a detailed evaluation of statistical values like mean, variance and mode. In motorways a significant difference of the distribution among lanes was observed, due to the differences in speed and composition of vehicles. In two-lane two-way roads the study of the parameter HV for a common value of flow rate did not show any clear tendency. Although the vehicle type specific distribution of heavy vehicles has higher values of mean and variance with respect to passenger car distribution, the effect of HV here was not visible. In fact, beyond the high correlation through parameters and the strong sensitivity to other macroscopic variables, the same flow rate has a masking effect on the HV variable and could not be completely nullified.

Using this approach, the study of headways can be reused in the future for different types of roads in order to expand the scope of study. A direct benefit of this research could be the implementation of probability density distributions in microsimulation software or driving simulators with regard to vehicle generation, as well as a bench test.

Appendix A

Gamma-GQM parameters estimation

The procedure to estimate the parameters α, β, λ and ϑ of the gamma-GQM distribution for a sample of headway h is based on the log-likelihood method. Since it was not possible to compute the maximum of the log-likelihood function analytically because of the complex mathematical structure, a numerical procedure in R language has been implemented (Fig. A.1). The pattern is to calculate the value of the log-likelihood function (L) on a set of pre-determined values: at iteration *iter* the program select the vectors:

$$\vec{\alpha}^{iter} = \begin{pmatrix} \alpha_1^{iter} \\ \alpha_2^{iter} \\ \alpha_3^{iter} \\ \alpha_4^{iter} \\ \alpha_5^{iter} \end{pmatrix} \quad \vec{\beta}^{iter} = \begin{pmatrix} \beta_1^{iter} \\ \beta_2^{iter} \\ \beta_3^{iter} \\ \beta_4^{iter} \\ \beta_5^{iter} \end{pmatrix} \quad \vec{\lambda}^{iter} = \begin{pmatrix} \lambda_1^{iter} \\ \lambda_2^{iter} \\ \lambda_3^{iter} \\ \lambda_4^{iter} \\ \lambda_5^{iter} \end{pmatrix} \quad \vec{\theta}^{iter} = \begin{pmatrix} \theta_1^{iter} \\ \theta_2^{iter} \\ \theta_3^{iter} \\ \theta_4^{iter} \\ \theta_5^{iter} \end{pmatrix}$$

and computes numerically the log-likelihood function for all the possible combination of the parameters. All vectors contain 5 elements (ni=nj=nk=nm=5), for a total of 5^4 combinations; they are arranged in order to create a 4-dimensional wire-frame, that is what we call *investigation zone* (the distance between consecutive elements of a vector is a fixed step, whose length is halved decreases at every iteration).

For each “A block” (see Fig. A.2) the program detects the set of values:

$$\vec{\xi}_{ott} = (\alpha_{ott}; \beta_{ott}; \lambda_{ott}; \theta_{ott})$$

which maximize the log-likelihood function (L_{max}). If the solution belongs to the boundary of the wire-frame, the investigation zone is shifted and the process is repeated; if it is inside the wire-frame, the investigation zone is thickened around the solution and a new iteration begins.

First iteration assumes:

$$\vec{\alpha}^1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \quad \vec{\beta}^1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \quad \vec{\lambda}^1 = \begin{pmatrix} 0.1 \\ 0.125 \\ 0.15 \\ 0.175 \\ 0.2 \end{pmatrix} \quad \vec{\theta}^1 = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.3 \\ 0.4 \\ 0.5 \end{pmatrix}$$

We assume 6 iteration as sufficient to reach an adequate parameter approximation (10e-2 precision for α and β, 10e-4 for λ and 10e-3 for ϑ approximately).

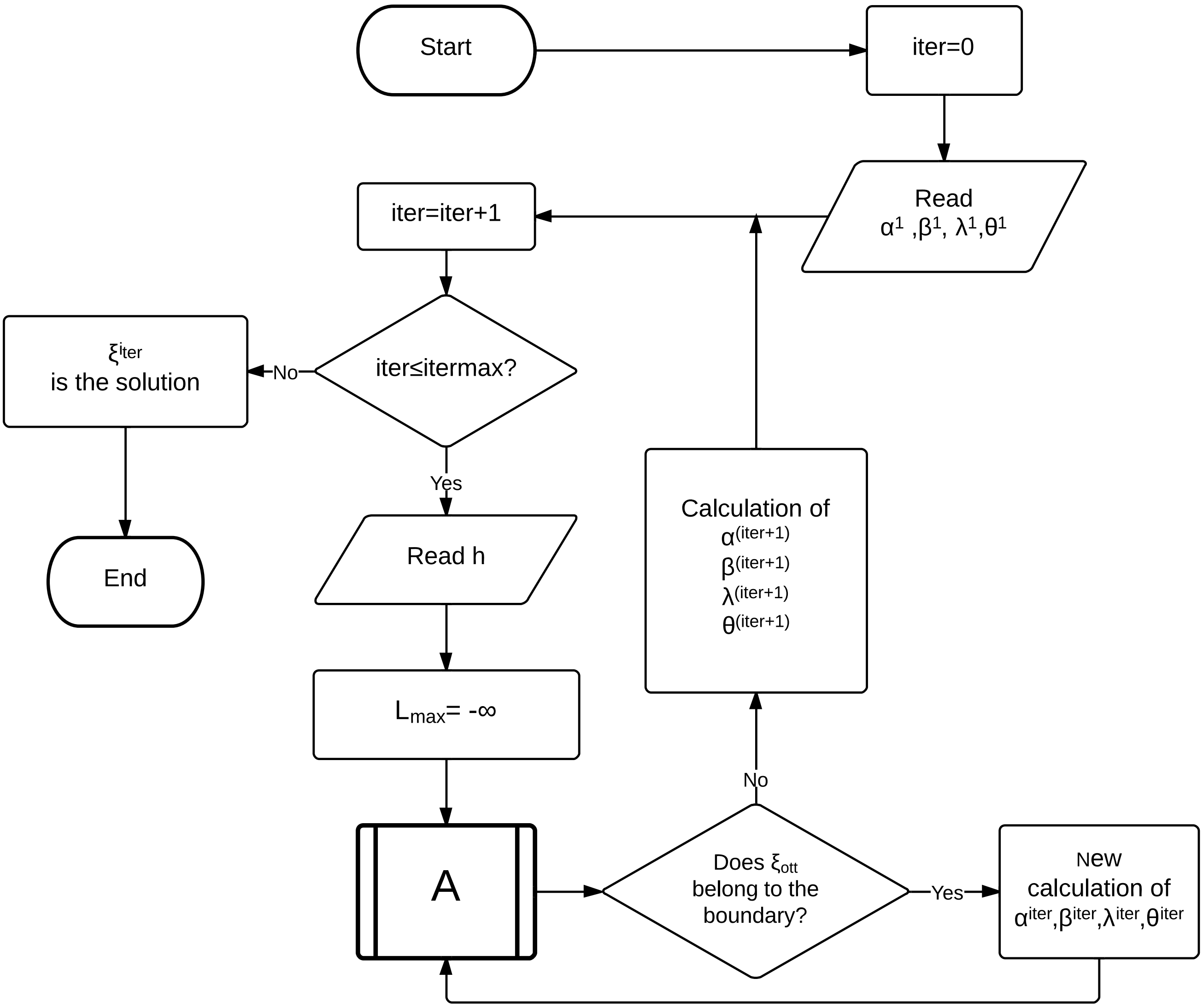

Figure A.1: Flowchart of the process of parameter estimation for gamma-GQM

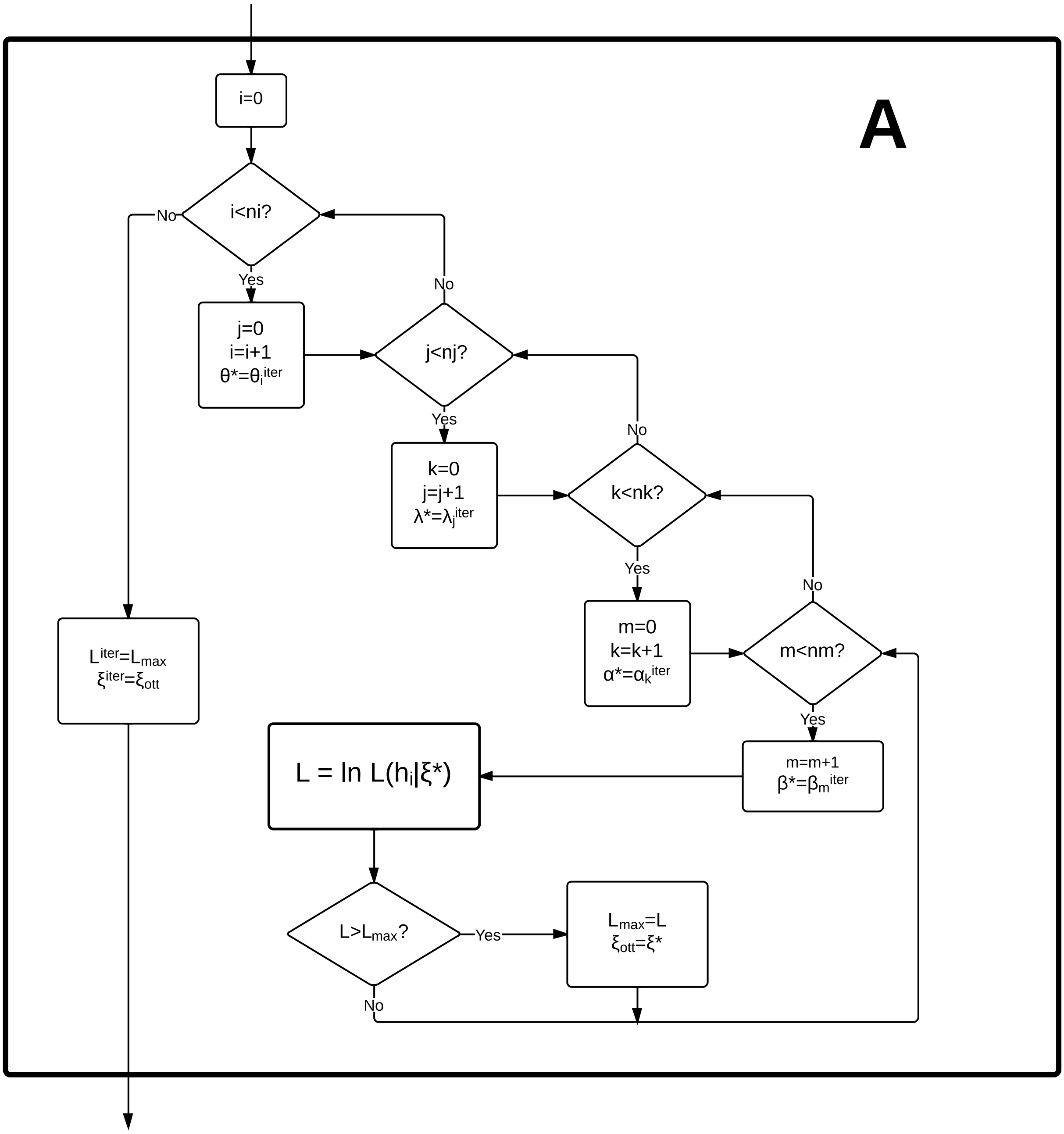

Figure A.2: Flowchart of A block of the process

Appendix B

Trendless analysis

In Chapter 5 we have applied the exponential ordered score test to different samples in order to identify the periods of stationarity. Since headway is the opposite of flow rate and the concept of stationarity involves a restrained variation of first order moment, we wanted to check the ability of the test representing flow rate (veh/hour/lane) computed as 10-minute moving means by using two different colours to distinguish trendless intervals.

We report here:

- all the lanes of motorway in Berlin direction, on Wednesday (Fig. B.1);
- all the A sections (Bibione, Trezze, Calcroci, Spinea) for two-lane two-way roads in both directions (Fig. B.2,B.3);
- section B (Verona) and C (Lazise) on Thursday and Saturday week-day (Fig. B.4).

Figure B.1: Identification of trendless time intervals and corresponding flow rate (veh/hour/lane). Motorway A2, dir. Berlin, Wednesday 1.3.13

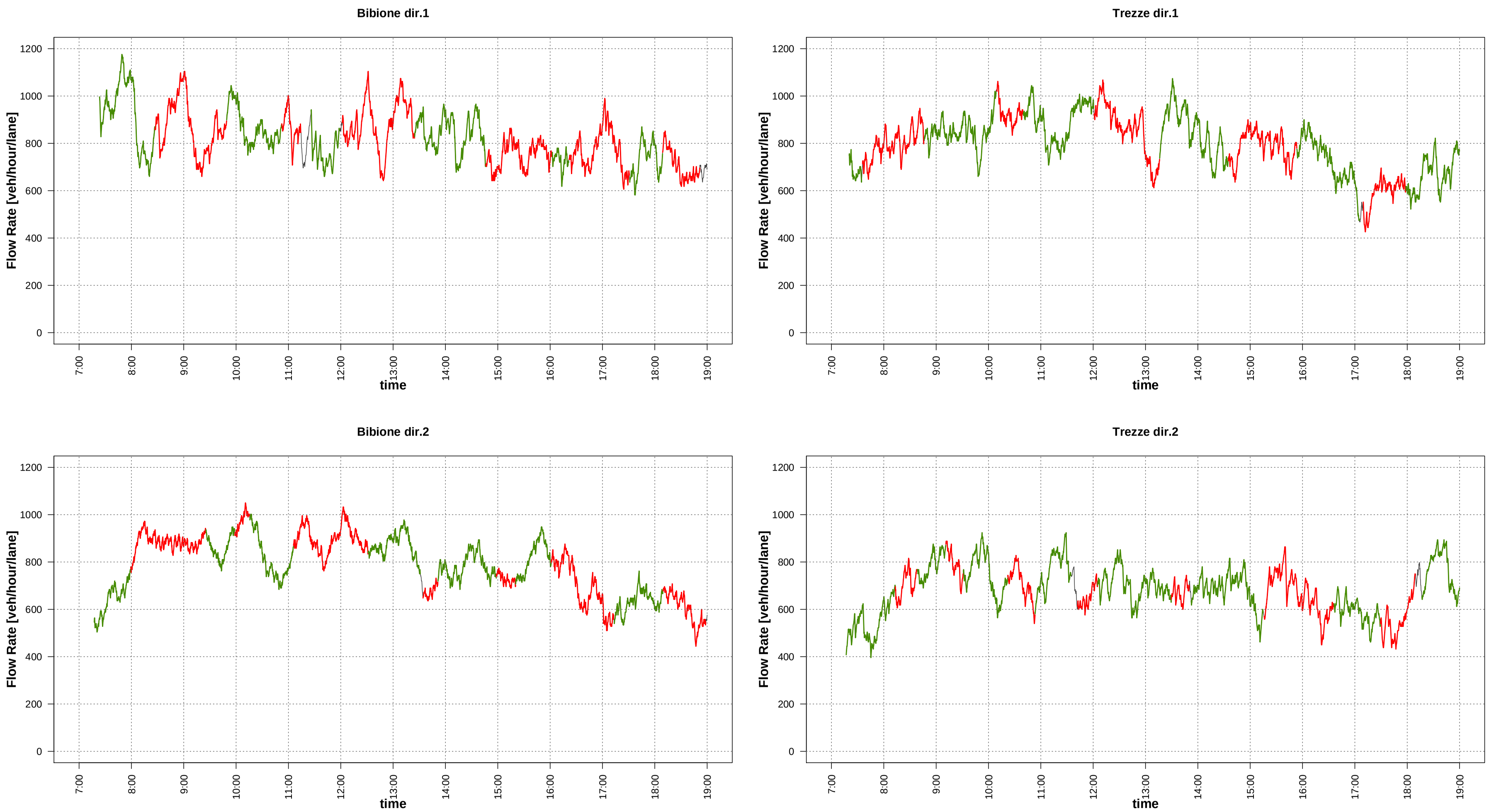

Figure B.2: Identification of trendless time intervals and corresponding flow rate (veh/hour/lane). Bibione and Trezze, both directions.

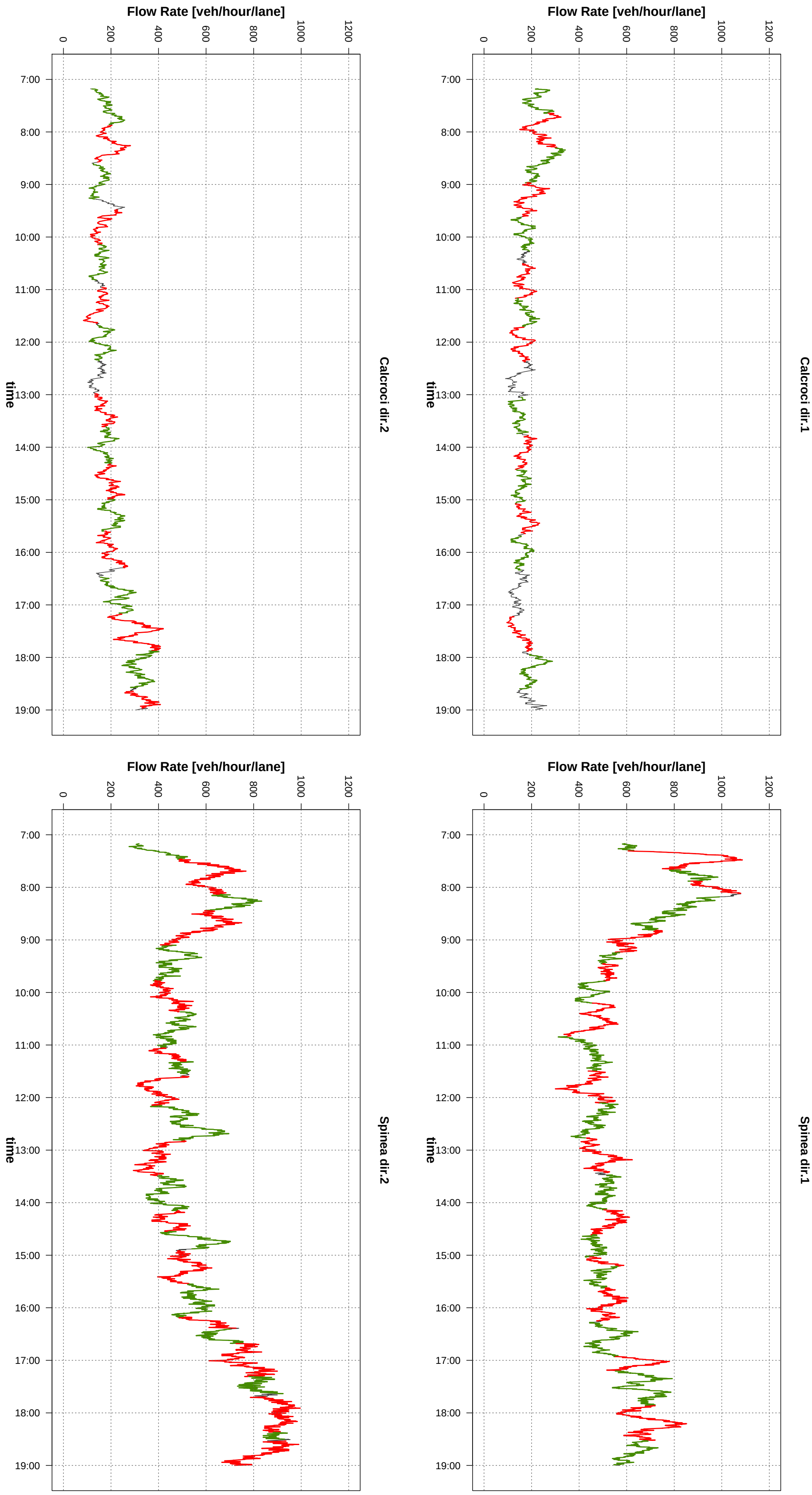

Figure B.3: Identification of trendless time intervals and corresponding flow rate (veh/hour/lane). Calcroci and Spinea, both directions.

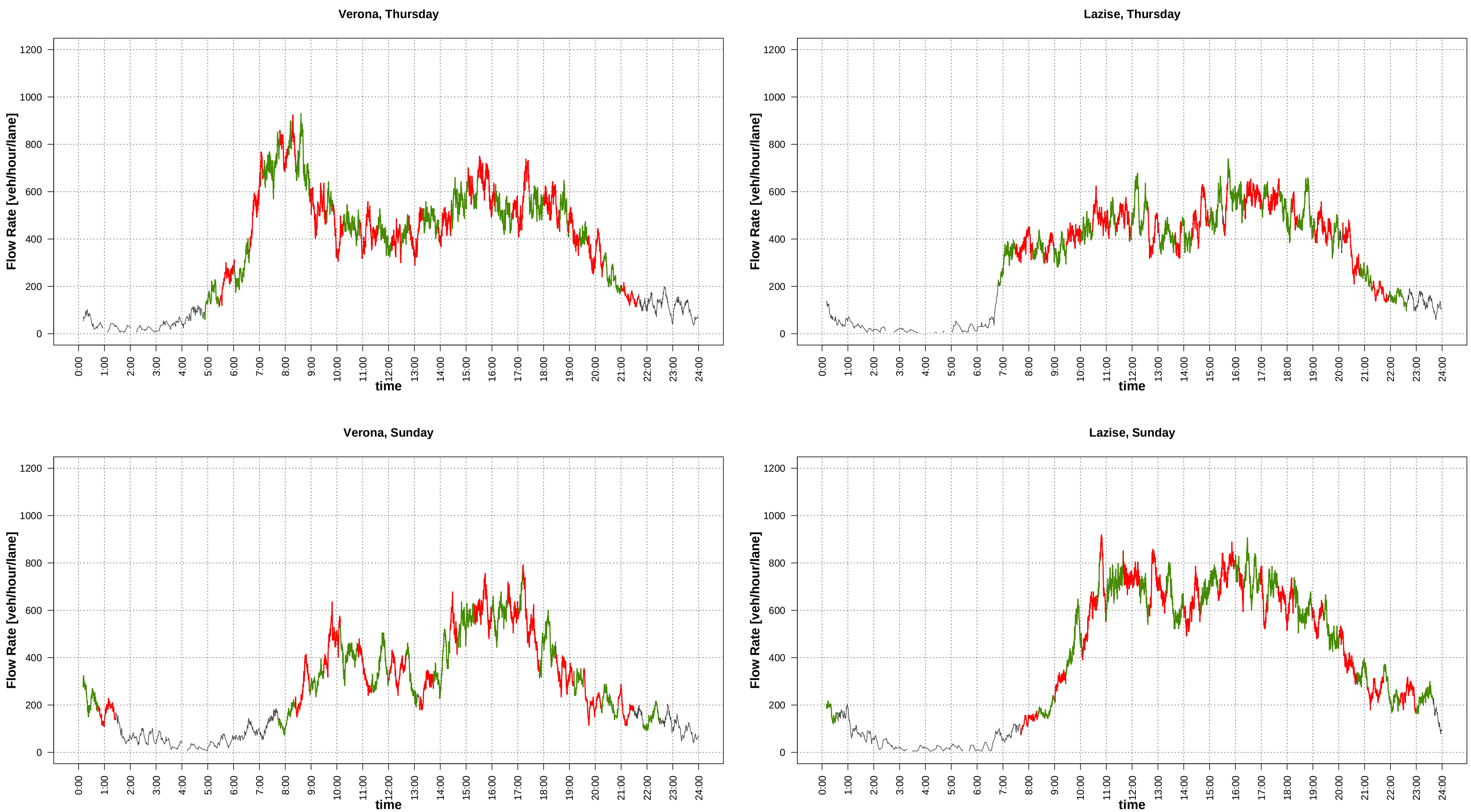

Figure B.4: Identification of trendless time intervals and corresponding flow rate (veh/hour/lane). Verona and Lazise, same direction, Thursday and Saturday week-day.

Bibliography

[1] Adams, W. F., 1936. Road traffic considered as a random series. Journal of the Insflow rate (pce/hour/lane) computed as 5-minute moving meanstitution of Civil Engineers 4 (1), 121–130.

[2] Ai, Q., Yao, Z., Wei, H., and Li, Z., 2010. Identifying Characteristics of Freeway Traffic Headway by Vehicle Types Using Video Trajectory Data. ICCTP 2010, pp. 1984-1992.

[3] Branston, D., 1976. Models of Single Lane Time Headway Distributions. Transportation Science 10 (2), 125-148.

[4] Breiman, L., Gafarian, A. V., Lichtenstein, R. & Murthy, V. K., 1968. An experimental analysis of single-lane time headways in freely flowing traffic. 4th Int. Sympos. Traffic, Karlsruhe.

[5] Breiman, L., Lawrence, F., 1973. Time scales, fluctuations and constant flow periods in uni-directional traffic. Transportation Research 7, 77–105.

[6] Buckley, D. J., 1968. A semi-Poisson model of traffic flow. Transportation Science 2(2), 107–133.

[7] Cowan, R. J., 1975. Useful headway models. Transportation Research 9(6), 371–375.

[8] Cox, D. R., Lewis, P. A. W., 1966. The Statistical Analysis of Series of Events. Methuen & Co Ltd., London.

[9] Gross, D., Harris, C. M., 1985. Fundamentals of Queueing Theory. 2th ed., John Wiley & Sons, New York.

[10] Ha, D.-H., Aron M., Cohen, S., 2010. Comparison of Time Headway Distribution in Different Traffic Contexts. Proceeding of the 12th World Conference of Transportation Research, Vol. 1261 (2010).

[11] Ha, D.H., Aron, M., Cohen, S., 2011.Variations in Parameters of Time Headway Models According to Macroscopic Variables, Fundamental Diagram, and Exogenous Effects. Transportation Research Record 2260 (2011), 102-112.

[12] Ha, D.H., Aron, M., Cohen, S., 2012. Time Headway Variable and Probabilistic Modeling. Transportation Research Part C 25 (2012), 181-201.

[13] Habtemichael, F.G., Santos, L. de P., El Faouzi, N-E., 2010. Time headway distribution parameters as performance indicators of motorway traffic and driver behavior: a comparison between good and adverse weather conditions. 91th TRB 2012 Annual Meeting.

[14] Hoogendoorn, S.P., Botma, H., 1997. Modeling and Estimation of Headway Distri-

butions. Transportation Research Record 1591 (1997), 14-22.

[15] Hoogendoorn, S.P.,Bovy, P.H.L., 1998. New Estimation Technique for Vehicle-Type-Specific Headway Distributions. Transportation Research Record 1646 (1998), 18-28.

[16] Jang, J., 2011. Analysis of Time Headway Distribution on Suburban Arterial. KSCE Journal of Civil Engineering 16(4), 644-649.

[17] Luttinen, R.T., 1996. Statistical Analysis of Vehicle Time Headway. PhD Thesis. University of Technology Lathi Center, Neopoli, Lathi, Finland.

[18] McLean, J. R., 1989. Two-Lane Highway Traffic Operations: Theory and Practice. Gordon and Breach Science Publishers, New York.

[19] Minderhoud, M., Botma, H., Bovy, P.H.L., 1997. Assessment of Roadway Capacity Estimation Methods. Transportation Research Record 1572 (1997), 59-67.

[20] R Development Core Team, 2006. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

[21] Rossi, R., Gastaldi, M., 2012. An Empirical Analysis of Vehicle Time Headways on Rural Two-lane Two-way Roads. 15th meeting of the EURO Working Group on Transportation, September 2012.

[22] Sullivan, D. P., Troutbeck, R. J., 1994. The use of Cowan's M3 headway distribution for modeling urban traffic flow. Traffic Engineering and Control 35(7/8), 445–450.

[23] Summala, H., Vierimaa, J., 1980. Vehicle speeds, speed differences and headways at 10 locations in autumn 1978: Analyses of driver behavior. Finnish National Board of Public Roads and Waterways, Helsinki.

[24] Wasielewski, P., 1979. Car-following headways on freeways interpreted by the semi-Poisson headway distribution model. Transportation Science 13 (1), 36–55.

[25] Ye, F., Zhang, Y., 2009. Vehicle type-specific headway analysis using freeway traffic data. Transportation Research Record 2124 (2009), 220-230.

[26] Zala, L.B., Modi, K.B., Desai, T.A., Roghelia, A.N., 2011. Headway distribution for NH-8 traffic at Vaghasi Village location. National Conference on Recent Trends in Engineering & Technology, 13-14 May 2011.

[27] Zwahlen, H.T., Oner, E., Suravaram, K.R., 2007. Approximated Headway Distributions of Free-Flowing Traffic on Ohio Freeways for Work Zone Traffic Simulations. Transportation Research Record 1999, 131-140.