The relationship between road safety and congestion on motorways

Paula Marchesini & Wendy Weijermars

R-2010-12

The relationship between road safety and congestion on motorways

A literature review of potential effects

R-2010-12
Paula Marchesini & Wendy Weijermars
Leidschendam, 2010
SWOV Institute for Road Safety Research, The Netherlands

Report documentation

Number:	R-2010-12
Title:	The relationship between road safety and congestion on motorways
Subtitle:	A literature review of potential effects
Author(s):	Paula Marchesini & Wendy Weijermars
Project leader:	Wendy Weijermars
Project number SWOV:	01.6.1.2
Keywords:	Congestion (traffic), motorway, rear end collision, collision, accident, severity (accid, injury), speed, traffic concentration, mathematical model.
Contents of the project:	This literature review investigates the relationship between congestion and safety at road sections of the main road network (mainly motorways) and specifically looks at unstable and congested traffic conditions.
Number of pages:	28
Price:	€ 8,75
Published by:	SWOV, Leidschendam, 2010

SWOV Institute for Road Safety Research
P.O. Box 1090
2260 BB Leidschendam
The Netherlands
Telephone +31 70 317 33 33
Telefax +31 70 320 12 61
E-mail info@swov.nl
Internet www.swov.nl

Summary

Mobility has been increasing significantly in the last few decades and will continue to increase. On road stretches which have insufficient capacity, traffic becomes congested. Traffic congestion has a negative impact on the economy and on the quality of people's lives. Road users experience delay and stress, and environmental pollution increases. The effects of traffic congestion on traffic safety, however, are less obvious.

This literature review investigates the relationship between congestion and safety at road sections of the main road network (mainly motorways) and specifically looks at unstable and congested traffic conditions. The hypotheses that are explored are the following:

1. Traffic congestion levels have been increasing significantly in the last few decades, and this trend seems to continue. The general perception is that crash frequency increases with increasing congestion levels. Nevertheless, it is expected that severe crashes will not increase under these traffic conditions.
2. Crash frequency, severity and type are expected to be affected by the changing flow conditions that occur when traffic starts to become congested.
3. Once traffic is congested, fewer (serious) crashes are expected to occur within the queue. In contrast, at the tail of the queue more severe rear-end crashes are to be expected.
4. It is expected that the effects of congestion on safety depend on the extent to which drivers are surprised by the congestion. This may, in its turn, depend on the type of congestion, the location of the queue and the use of variable message signs.

Two types of studies are discussed; 1) studies that use aggregated data and compare various locations and/or different time periods (e.g. peak and non-peak) and 2) studies that are based on short observation periods and analyse which traffic conditions result in crashes.

With regard to the first hypothesis, results of different studies are not consistent. Some studies find that high volume to capacity (V/C) ratios result in higher crash rates but less severe crashes. Another study finds that crash rates decrease at high density levels, and one of the studies does not find any relationship between congestion and crash frequency nor between congestion and crash severity. None of these studies explicitly provide information about the influence of congestion on crash rate.

The results regarding the second hypothesis seem to be more consistent. The studies suggest that crash likelihood increases as speed variability increases (a typical indicator for unstable traffic conditions). Also large speed differences between lanes and density variability seem to increase crash likelihood.
Crash severity seems to decrease with increasing volumes (or V/C ratios). Golob, Recker & Pavlis (2008) report that crash severity does not seem to change during the transition from free flow to congested conditions, yet

decreases once traffic is congested. Furthermore, the studies seem to show consistent results with regard to crash type. Rear-end crashes are more likely to occur during unstable conditions.

The third hypothesis could not be fully proved, due to the scarcity of the literature on this topic. Nevertheless, the studies which rely on disaggregated data provide some insight in this respect. The study by Golob, Recker & Pavlis (2008) (also based on disaggregated data) reports some findings that can be related to the situation within the queue. It is observed that once traffic is congested, crash severity is greatly reduced when all lanes present similar flow conditions. The authors find that under these conditions of uniformity, crashes with (fixed) objects are more likely to occur. When only the left and interior lanes are congested, rear-end and side-impact crashes are more likely.

The fourth hypothesis about the effects on safety of structural and incidental congestion could not be investigated on the basis of the available literature.

In summary, the literature discussed provides insight into the relation between unstable traffic conditions and crash frequency, severity and type. Moreover, it offers a better understanding of rear-end crash occurrence at the tail of the queue. Although the available literature does not provide a complete picture of the relationship between congestion and safety, it does present an overview of the current status of research on this subject. However, several questions are still to be answered. As increasing congestion seems to pose a significant problem to traffic safety, there is need for further research. Therefore, we recommend investigating the relationship between congestion and traffic safety on motorways in more detail, by linking data on traffic volumes with congestion and crash data.

Contents

1. Introduction

Mobility has been increasing significantly in the last few decades and will continue to increase (Janssen, Okker & Schuur, 2006). In the Netherlands, the number of kilometres driven increased by 5 percent between 2000 and 2008 (KiM, 2009) and car use increased by 10 percent between 2000 and 2008 (KiM, 2008). The number of kilometres driven on Dutch motorways increased from 55,6 billion in 2000 to 63,3 in 2007 (DVS, 2008). On road stretches which have insufficient capacity, traffic becomes congested.

The delay due to congestion[1] on motorways increased by 58 percent between 2000 and 2008, from 30,8 million hours to 48,7 million hours (DVS, 2009). *Figure 1*, below, shows that delay mainly increased during peak periods (KiM, 2008).

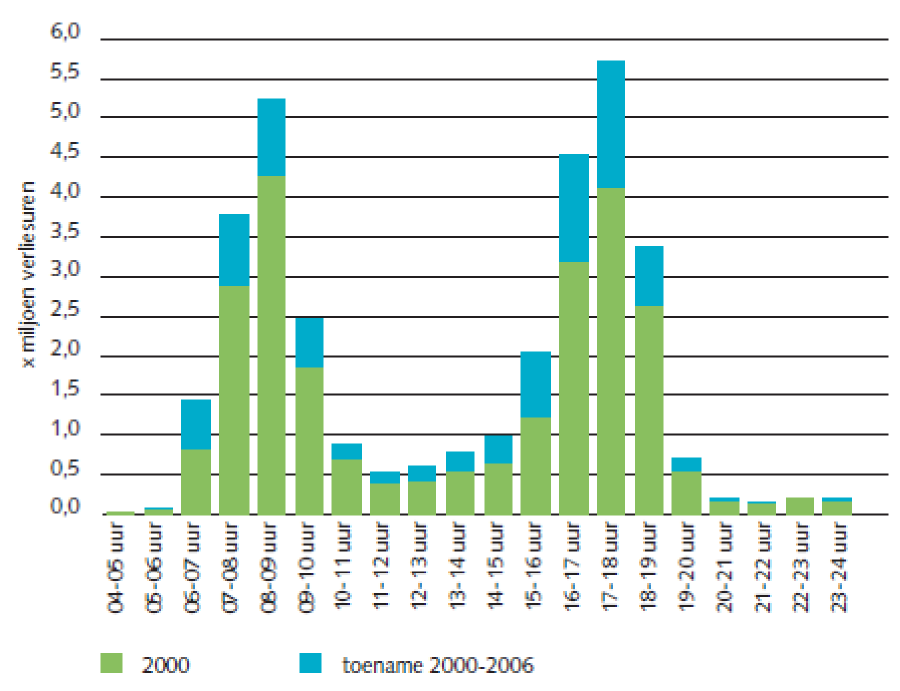

Figure 1. *The amount of traffic delay (y-axis) in relation to the time of the day (x-axis) and in blue the increase of traffic delay between 2000 and 2006 (Source: KiM, 2008).*

Traffic congestion has a negative impact on the economy (by decreasing productivity) and on the quality of people's lives. Road users experience delay and stress and environmental pollution increases (more noise, more emissions and more fuel consumption).

---

[1] This is defined as the delay for vehicles that travel with a speed lower than 50km/h compared to a reference speed of 100 km/h.

The effects of traffic congestion on traffic safety, however, are less obvious. Rietveld & Shefer (1994) suggest that congestion might have a positive effect on safety by decreasing the number of fatalities as speeds decrease. Although this statement seems logical, when looking at the traffic conditions in more detail the effects of congestion on safety are less apparent. As traffic flow increases and density approaches its critical value, traffic flow is said to be unstable. Under these conditions, any small disturbance may lead to crashes. Once the traffic jam is formed, rear-end crashes may occur at the tail of the queue due to large differences in speed. Furthermore, motorways giving evidence of structural congestion (i.e. there where demand is almost always higher than capacity) cause road users to seek alternative routes - often perceived as faster routes - which are almost inevitably roads with a higher accident risk.

It is important to gain a clear understanding of how traffic flow processes affect safety, in order to better understand developments in traffic safety and in order to identify possibilities for improvement. This report focuses on the effects of unstable and congested traffic situations on traffic safety.

1.1. Problem statement

The literature review in this report focuses on the relationship between traffic volume and safety at road sections of the main road network (mainly motorways) and specifically looks at unstable and congested traffic conditions. A literature review on the relationship between congestion and safety on secondary roads can be found in Duivenvoorden (2010).

The literature review is intended to provide insight into the following hypotheses:

1. Traffic congestion levels have been increasing significantly in the last few decades and this trend seems to continue. The general perception is that crash frequency increases with increasing congestion levels. Nevertheless, it is expected that severe crashes will not increase under these traffic conditions.

2. As traffic **starts to become congested**, the average speed decreases but the number of interactions increases and flow conditions are unstable. It is expected that fewer *severe* crashes will occur given the lower speeds. Unstable flow conditions are expected to result in an increase in the frequency of rear-end crashes. Also, run-off-road and side-impact crashes may occur, given the manoeuvre drivers will make in order to avoid hitting the tail of a queue. Frequency, severity and type of crash are therefore expected to be affected by the changing flow conditions.

3. Once **traffic is congested**, vehicles decelerate and accelerate within short distances. Speed in the queue is low and as a result fewer crashes are expected to occur within the queue. On the other hand, traffic that approaches the queue might be surprised by the queue and therefore, more severe rear-end crashes are to be expected at the tail of the queue.

4. There are two types of congestion: structural or incidental (also referred to as recurrent and non-recurrent congestion). Structural congestion occurs when traffic demand is higher than capacity, while incidental

congestion is the result of occasional conditions such as a crash, bad weather or road works which alter the traffic flow[2]. Drivers are able to anticipate the occurrence of structural congestion more easily because it takes place at about the same locations at about the same time of the day. Incidental congestion, on the other hand, might surprise drivers and lead to more unexpected behaviour. The extent to which congestion surprises drivers also depends on other factors, such as the distance from which the queue is visible (e.g. a queue located just after a curve will result in greater surprise) and the use of technology to warn drivers (e.g. variable message signs or in-car technology). It is expected that the effects of congestion on safety will vary according to the type of congestion and the extent to which the congestion surprises the driver.

5. There is a relationship between congestion levels and route choice. Drivers opt for alternative routes in an attempt to reach their destination faster than by simply joining or entering a queue. This route choice may be based on the driver's familiarity with the network or on information provided by either variable message signs or in-vehicle systems. When traffic diverts from congested motorways to secondary roads, a safety problem might arise as different types of roads show different levels of safety. Motorways, for instance, are safer than secondary roads (SWOV, 2007). Consequently, route choice management has to be addressed from a mobility perspective as well as a safety perspective. This is a very broad topic which is dealt with in a different research project (Dijkstra & Drolenga, 2008).

1.2. Structure of the report

Chapter 2 of this report provides a brief overview of traffic flow theory while *Chapter 3* provides insights based on a literature review on congestion and safety. Finally, *Chapter 4* discusses the conclusions that can be drawn on the basis of the literature and gives recommendations regarding further research.

---

[2] Congestion in the Netherlands is mainly (about 80%) due to high traffic demand, 12 – 15% is due to traffic accidents and 4% to road works (DVS, 2008).

2. Introduction to traffic flow theory

Traffic flow can be described at a macroscopic level by three main variables; volume, density and mean speed. These variables can be defined as follows (Hoogendoorn, 2007):

“The volume of traffic (*traffic volume*) is the number of vehicles passing a cross-section of a road in a unit of time.”

“The *density* of a traffic flow is the number of vehicles present on a unit of road length at a given moment. Just like the volume the density can refer to a total road, a roadway, or a lane.”

“The *mean speed* is the arithmetic mean of the speeds of the vehicles passing a point on a roadway. The mean speed can be measured at a point (cross section) or at a given moment. The space-mean speed is the arithmetic mean of the speeds of the vehicles on a road section at a given moment.”

Under stationary and homogeneous flow conditions, the relationship between the variables traffic volume (q), density (k) and speed (u) can be described by the following equation:

$$q = k \cdot u$$

where u is the space-mean speed.

The relations between pairs of these variables are described by the so called *fundamental diagrams*. *Figure 2* below shows the theoretical fundamental diagrams.

From these graphs various interesting points emerge that distinguish free flow from congested conditions:

u_o is the free flow speed, that is, the speed if $k=0$ and $q=0$
u_c is the capacity speed
q_c is the critical traffic volume
k_c is the critical density
k_j is the jam density

The region of the graphs that takes values of density smaller than the *jam density* is called the free flow branch, whereas the region that takes density values larger than the critical is called the congested branch.

These diagrams provide information about the relationship between the variables but not about the causality. The shape of these diagrams has been studied extensively and various mathematical models and empirical relationships can be found in the literature (Hall, 2001). In this report, this topic will not be discussed further.

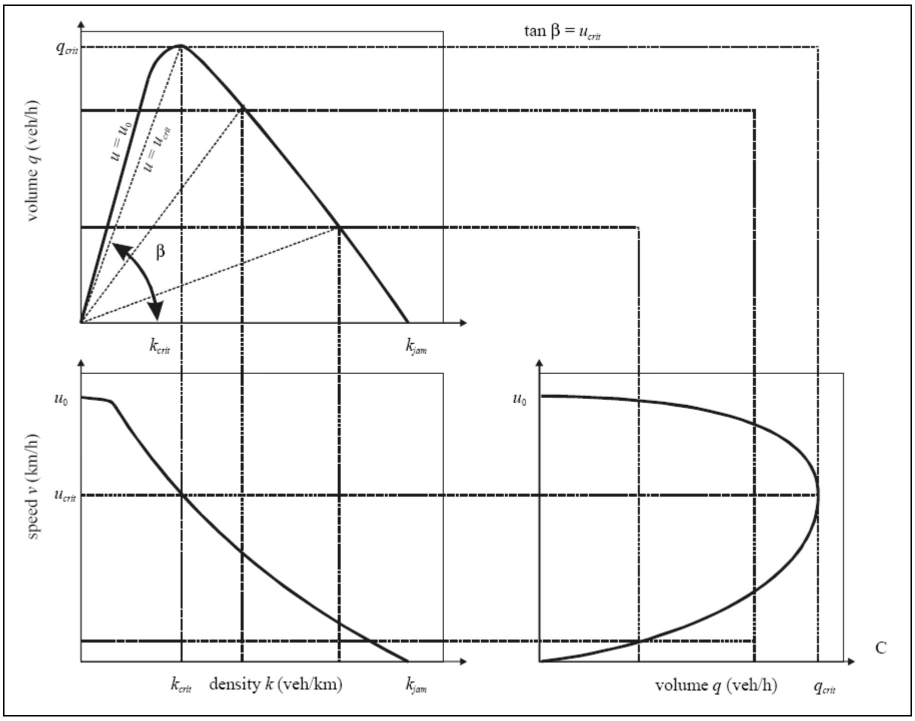

Figure 2. *Three interrelated forms of the fundamental diagrams (Hoogendoorn, 2007).*

2.1. What happens when the traffic flow becomes congested

Congested traffic flow is characterized by a situation in which road users cannot drive at their desired speed because they are constrained by the presence of other road users. In contrast, a free flow situation occurs when road users do not influence each other's behaviour.

Smooth traffic flow can suddenly change due to, for instance, a decelerating or slow moving vehicle or a partial blocking of the road. This sudden change generates discontinuities in the traffic stream in the form of kinematic shock waves. A shock wave is a rapid change in the traffic conditions (flow and density) that propagates in time and space (Van Driel, 2007; Hoogendoorn, 2007). The shock wave is characterized by the sudden change in the individual vehicle speeds downstream of the disturbance.

Shockwaves can be classified in various ways. For instance, they can be classified according to the direction in which they propagate. A *backward* moving shockwave propagates upstream of the disturbance while a *forward* moving shockwave propagates in the same direction as the traffic. Shockwaves can also be differentiated with regard to the result of their propagation. As a result of a *forming* shock wave the congested section increases (e.g. vehicles joining the queue) and as a result of a *recovery* shock wave the congested section decreases; e.g. vehicles leaving the queue (May, 1990).

Assume, for instance, a reduction of capacity on a section of a motorway (that is to say, a bottleneck where the road section changes e.g. from three to two lanes). If the flow approaching the bottleneck (q_i) is larger than the capacity of the bottleneck (q_o), a queue will start growing (backward forming shockwave). As soon as q_i becomes smaller than q_o the queue will start to dissipate (forward recovery shockwave). See also *Figure 3*. In a bottleneck there is also a frontal stationary shockwave located in the place where demand is larger than capacity.

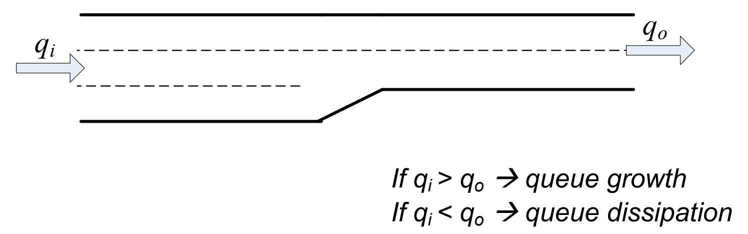

Figure 3. *Bottleneck.*

As the traffic flow becomes congested, there are some particular conditions that can be distinguished when vehicles approach, enter, and leave the jam. When approaching a jam, drivers have to decelerate in order to adapt to the speed of the queue. This is not instantaneous as the driver's reaction time and the reaction time of the car entail a delay. Drivers, nevertheless, can anticipate and start decelerating before they queue. If the driver's reaction time and the reaction time of the vehicle equal the driver's anticipation, a stable continuation of the jam will take place, taking in a constant inflow. Within a jam, drivers keep a larger headway when driving in stop & go conditions than when driving in heavily congested or free flow conditions. The quickness of acceleration responses as drivers leave the queue has an impact on how fast the queue will dissolve (Van Driel, 2007). The input/output flow, coupled to the cause of the traffic disturbance (e.g. bottleneck, slow moving vehicle), also determines the pace of growth and continuation of the jam.

2.2. How to measure congestion

There are several ways of measuring congestion and each of them involves the use of different parameters such as volume, density, occupancy, queue length, travel time, delay, speed, volume to capacity ratio (V/C) and level of service (LOS). Nevertheless, there is still ongoing research into what a good measure for congestion is (for instance Jun & Lim, 2008). V/C and LOS have long been used as defined in the Highway Capacity Manual. LOS is a quality measure describing operational conditions within a traffic stream. LOS A implies free flow conditions whereas LOS F indicates congested flow. V/C ratio is defined as the volume divided by capacity. A V/C ratio of 0.8 or 0.85 is a commonly accepted threshold for congested flow conditions.

The definition of congestion on motorways that is used by Rijkswaterstaat (the Dutch Ministry of Waterways and Public Works) is an average speed of 50km/h (or lower). Other calculation methods use a variety of the listed parameters and will be discussed in the next chapter.

3. Literature review

The literature reviewed in this report focuses on the hypotheses presented in *Chapter 1*. The literature search was performed using SWOV and Delft University of Technology libraries and Google's search engine. The key words used were the following:
Accident, crash, safety, flow, volume, capacity, congestion, V/C, freeways, motorways, density, occupancy, speed, rear-end crash/accident, multi-vehicle crash/accident, queue, traffic jam.

There are various studies that examine the relation between congestion and safety, and their outcomes are quite diverse. The main differences are to be found in the research methodology and in the ways in which congestion is measured.

Section 4.1 deals with the effects of congested traffic on the frequency, severity and type of crash. *Section 4.2* discusses literature that deals with crash type at the tail of and within the queue.

3.1. Relationship between flow conditions and crash severity, frequency and type

The studies that deal with the relationship between flow conditions and crash severity, frequency and type can be split into two categories. The first uses data established during long observation periods (e.g. hours) and compares congestion and safety levels at different locations or different periods in time. The data is then aggregated over a time span, e.g. a year. The second uses short observation periods (e.g. 5 minutes) and analyses which traffic conditions lead to crashes.

3.1.1. *Studies that use aggregated data*

Several studies investigate the relationship between congestion and safety by considering average hourly traffic flow. Chang & Xiang (2003) studied the relationship between congestion and crashes on arterial roads and motorways. Data on five arterial road sections and five motorway sections was employed for the study. The motorway sections studied comprised 179 links and 5402 crashes. The *volume per lane* was used as a surrogate measure for congestion. Crash frequency (AF) and accident rate (AR) were calculated according to the following formulas:

$$AF = \frac{\text{number of accidents on a link}}{\text{link length}}$$

$$AR = \frac{\text{number of accidents on a link}}{\text{link length} \cdot \text{AADT}}$$

The statistical analysis shows that on links with higher AADT (annual average daily traffic) more crashes occur. Moreover, the authors find that during peak-congestion periods road links with higher volumes per lane show higher accident rates. For off-peak periods, no relationship between volume and accident rate is found. The research also looks into the relation between crash severity and congestion and the results show that the crashes that take place at motorways with higher ADDT are likely to be less severe.

Zhou & Sisiopiku (1997) assessed the relationship between V/C ratios and accident rate on a 26km (16 miles) stretch of an urban motorway in Michigan. This study uses V/C ratios as a measure of the traffic flow conditions. They argue that this measure is more representative of the accident exposure as it includes geometric and road operational characteristics. Three sections of the motorway with rather similar traffic volume and geometry were selected. Data on traffic flow and crashes on these sections were collected and aggregated for the analysis. In order to obtain a wide range of V/C ratios, they calculated hourly V/C ratios and accident rates. For each hour of the day, the accident rate was obtained from the following equation:

$$\text{hourly acc rate} = \frac{\text{avg.hourly number of accidents per year} \cdot 10^8}{\text{avg.hourly traffic volume} \cdot 365 \cdot 16\text{miles}}$$

They report a U-shaped relationship between total accident rates and V/C ratios (*Figure 4*). This function implies that hours with low V/C ratio (e.g. at night time) show high accident rates. As the V/C increases the accident rate decreases, reaches a minimum and increases again for higher V/C ratios.

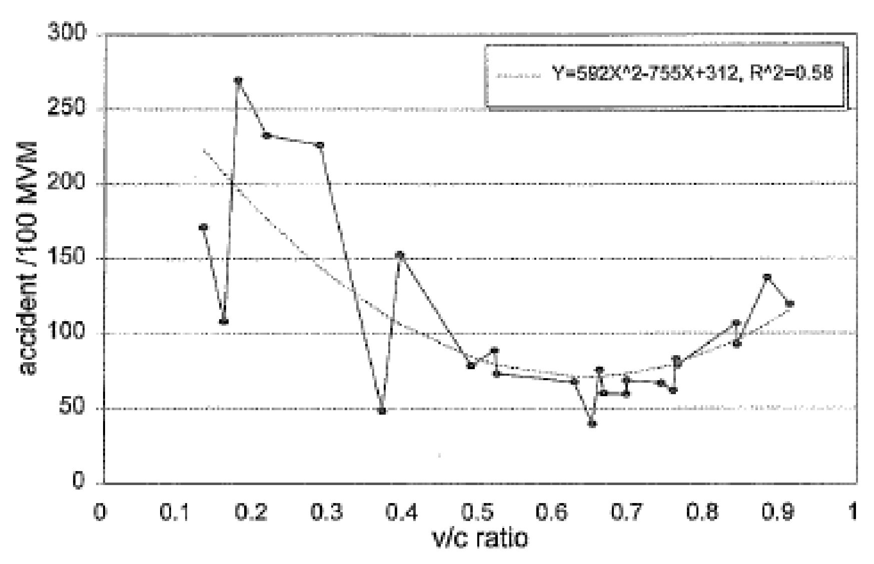

Figure 4. *Accident rate vs. V/C ratio (Zhou & Sisiopiku, 1997).*

A U-shaped function is also found between V/C ratios and multiple-vehicle accident rates as well as between V/C ratios and property-damage-only accident rates. However, for injury and fatal crashes increasing V/C ratios result in decreasing accident rates. The relations found in this study might also have been influenced by factors other than V/C ratios, e.g. light conditions and driver conditions (alcohol, fatigue).

Lord, Manar & Vizioli (2005) studied the influence of volume, density and V/C ratios on the occurrence of crashes on rural and urban motorways. They employed hourly flow data from loop detectors located at various sections of a 40.5 km rural motorway and a 5 km urban motorway. The crashes used for the analysis cover a five-year period. The vehicle densities as well as the V/C ratios were calculated on the basis of the method proposed by the Highway Capacity Manual 2000 (Transportation Research Board, 2000). They find similar results to those obtained by Zhou and Sisiopiku (1997); as the V/C ratio increases, so does the likelihood of multi-vehicle crashes. With regards to density, the study concludes that as density increases, crash frequency increases, reaches a maximum and then decreases again. The relationship found between volume and crash risk suggests that the latter decreases as the volume increases. The authors do not make a distinction between congested and non-congested flow regimes when fitting the models to the data, and they suggest that such a distinction could better capture the nature of the relationship between crashes and traffic flow. In this study, the relations might again have been influenced by other factors such as light conditions.

In contrast to the above mentioned studies, Wang, Quddus & Ison (2009), who use a congestion index as a measure of congestion, report that there is no relationship between congestion and crash frequency or crash severity. The congestion index (CI) was generated by the difference between the actual travel time (T) and the free flow travel time (T_0) divided by the free flow travel time, as it had been defined by Tailor et al. (2000) (In Wang, Quddus & Ison, 2009).

$$CI = \frac{T - T_0}{T_0}$$

The congestion index (CI) was calculated for each road segment and for each hour and subsequently an average CI per road segment was determined. A higher value of CI represents a higher level of congestion. The authors compare 70 road segments along the M25 motorway in England. Besides the CI, other explanatory variables are used in their statistical models such as AADT, segment length, horizontal radius, gradient and number of lanes. They also analyze peak and off-peak periods separately, but no relationship between crashes and the congestion index CI is found. This study, however, employs hourly flow, delay and speed (averaged over a year) and therefore, the actual impact of varying flow conditions over short periods of time might not have been captured.

None of the studies mentioned above explicitly provide information about the influence of unstable and congested traffic flow on crash rate. The studies use data aggregated over time and therefore do not capture varying flow conditions. The next section discusses research that attempts to capture those varying flow conditions. These studies investigate which traffic flow conditions lead to crashes and therefore also provide information about the processes that lead to crashes.

3.1.2. *Studies on crash precursors*

Rather recently, a new approach has come into use which involves disaggregated data. The reason for using disaggregated data is the possibility of better capturing short term variations in traffic flow conditions.

Several studies have been carried out that deal with ‘crash precursors’, which are the traffic conditions that exist before a crash occurs.

Oh et al. (2001) studied the crash likelihood on a motorway in California by using real-time data, that is, online data from loop detectors. Several traffic flow parameters (5-minute averages and standard deviations of speeds, flows and occupancies) are tested in order to identify which one(s) relate(s) best to crash likelihood. They find that the standard deviation of speed in these 5 minutes is a useful measure to characterize two different traffic conditions; normal and disruptive. They define a disruptive traffic condition as the condition that leads to the occurrence of a crash. Their study suggests that crashes are more likely to occur as speed variation increases.

The same result was obtained by Lee, Saccomanno & Hellinga (2002) in their study on the effect of crash precursors on crash potential. Apart from speed variation in time (given by the coefficient of variation in speed, calculated as the standard deviation of speed divided by the average speed), they also consider density and speed variation across lanes as crash precursors. The data used to calculate the crash precursors was obtained from loop detectors on a stretch of an urban motorway in Toronto. Speed variation was calculated over a period of 5 minutes before the crash occurrence, whilst density was calculated for the time of the crash. Their findings suggest that an increase in speed variation, not only in time but also across lanes, is linked to an increased likelihood of crash occurrence. The same relationship is found for increasing density. Lastly, they conclude that crash likelihood during a peak period is higher than during an off-peak period.

In a more recent study by the same authors (Lee, Hellinga & Saccomanno, 2003) a model was developed that relates crash precursors to crash potential by dealing with some of the shortcomings from the earlier study. The study area and period considered are the same, but one of the crash precursors is replaced; speed variation across lanes was found to have no influence on the crash potential. It is therefore replaced by the average speed difference between upstream and downstream ends of a road section (Q). The latter parameter is considered to account for the influence on crash occurrence during queue formation and dissipation. They then study the influence of speed variation (CVS), density (D) and average speed difference (Q) on crash occurrence. In order to correct for the probability of crash occurrence due to traffic volume, a measure of the exposure is included (given by the product of the AADT and the length of the road segment). One of the shortcomings of the earlier study by Lee, Saccomanno & Hellinga (2002) was the subjective determination of the observation time period. A rather arbitrary 5-minute observation period was employed in the first study. In the follow-up study the observation time period is determined in such a way that the difference between the values of the crash precursors (Q, CVS and D) for the crash and non-crash cases is maximized. They find that 8 minutes before a crash the difference in CVS between a crash and a non-crash case is maximum. Similarly, observation periods of 3 and 2 minutes provide the maximum differences in D and Q, respectively, for crash and non-crash cases. They conclude that crash potential increases with higher levels of speed variation, density and exposure.

A slightly different approach was taken in a study by Hourdos et al. (2006) as they only used a high-crash motorway section for the analysis. They investigate the relationship between traffic flow parameters and crashes (including near-misses, i.e. an evasive manoeuvre performed in order to avoid a crash) on a mile-long motorway section in Minneapolis. The data used for the study was obtained from video camera observations and traffic measurements at several locations. They point out that the benefit of using video camera is that it overcomes the inaccuracy of the time of the crash recorded by the police. The study includes 30 crashes and 122 near misses. They fit a model to the data that includes several traffic flow parameters (average speed, coefficient of variation of speed, traffic pressure, kinetic energy, etc.) as well as environmental conditions (pavement, visibility and sun position) as explanatory variables. The traffic pressure was obtained by multiplying density and speed variance.
It is concluded that large speed differences between lanes and density variability increase crash likelihood. Moreover, the sun facing the driver has a strong effect on crash likelihood, while visibility and pavement conditions only have a minor effect.

Golob, Recker & Pavlis (2008) presented a method that assesses the relationship between traffic flow parameters and the type of crash, severity, location and number of vehicles involved. Data on traffic volume and lane occupancy from six motorways was obtained from loop detectors during a six-month period. The motorways studied have three or more lanes but only three (the leftmost, one interior and the rightmost lane) were taken into account for the analysis. They find that data from loop detectors 20 minutes before a crash is sufficient to describe unsafe traffic conditions. An analysis of various traffic flow parameters shows that eight ‘Traffic Flow Factors’ can explain the relationship between traffic flow conditions and safety. The eight traffic flow factors are the following:

- Factor 1: Congestion in left and interior lanes
- Factor 2: Volume level
- Factor 3: Synchronized lane conditions (traffic conditions changing in the same way on all lanes)
- Factor 4: Right lane perturbation
- Factor 5: Volume variation across lanes
- Factor 6: Conforming curb lanes (degree to which right lane volumes are related to left and interior lanes volumes)
- Factor 7: Systematic volume change
- Factor 8: Synchronized flow in left and interior lanes (degree to which volumes and densities in the outer lanes are synchronized).

Their study finds that when the left and interior lanes are congested the likelihood of severe crashes decreases. If all lanes have similar traffic volumes, then congestion reduces crash severity by more than half.
It is found that the effect of congestion on severity (severity decreasing due to low speeds) is offset by unstable flow conditions such as the change from free flow to congestion. When there is a rather uniform volume level across lanes, the volume variation in time indicates a higher likelihood of injury crashes.

With regards to the type of collision, the authors find that when left and interior lanes are congested, the likelihood of rear-end crashes as well as of side-impact crashes is higher, although the latter to a lesser degree. Crashes with (fixed) objects are more likely to occur when all lanes present similar congestion levels. They suggest that these crashes might be caused by drivers trying to avoid rear-end and side-impact crashes. The study also suggests that rear-end crashes are more likely to occur under unstable flow conditions. This study provides interesting insights into the effects of varying flow conditions. Nevertheless, it only takes into consideration flow conditions prior to a crash and across lanes.

A rather different approach was adopted by Abdel-Aty & Pande (2005) who analyze consecutive locations. Unlike Golob, Recker & Pavlis (2008), they also include traffic flow conditions when there are no crashes as well as the longitudinal variation of the flow conditions. This approach allows for more insight into the differences between traffic flow conditions in crash and non-crash situations.
Using a spatial-temporal approach they develop a method for identifying locations that are crash prone, based on dynamic flow conditions. The method provides a spatial approach as it takes into account several consecutive locations to measure flow conditions during a certain period of time.
The raw data consisted of 30 seconds average speed, vehicle counts and lane occupancy (percentage of time the detector is occupied) measured at 28 loop detectors located along a 21km road stretch. This data was then aggregated into 5min averages and standard deviations, over the three lanes. The 30-minute time period preceding the crash was divided into 5-minute time slices, where time slice 1 corresponds to the 5-minute period before the crash.
Several locations along a road segment were considered for measuring flow conditions. Each crash was associated with the closest loop detector, called detector F (see Figure 5).
A coefficient of variation in speed (CVS) and an average occupancy (AO) were used as traffic flow indicators for the safety level of the road section. The coefficient of variation was obtained by the ratio between the standard deviation of speed in five-minute intervals and the average speed over the same interval.

$$CVS = \frac{\mathrm{SDS}_{5\mathrm{min}}}{\mathrm{AS}_{5\mathrm{min}}}$$

Data on time and day of the crashes was used to retrieve data on the pre-crash flow conditions (during an 8-month study period) from one loop detector downstream where the crash took place and five upstream (see *Figure 5*).

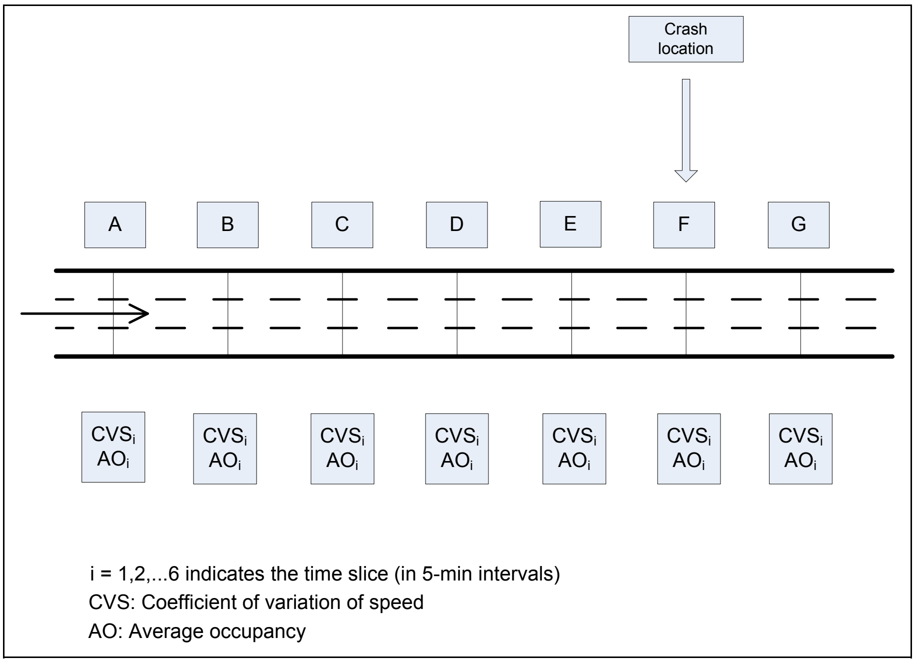

Figure 5. *Loop detector stations on a road section labelled A to G for a given crash (Adapted from Abdel-Aty & Pande, 2005).*

For instance, if a crash took place on a Wednesday at 5 p.m., then data (from loop detectors A to G) on speed, volume and occupancy for all Wednesdays of the eight-month study period, during 30 minutes before the time of the crash (i.e. from 4:30 till 5:00 p.m.) was obtained. The data was retrieved in this way in an attempt to distinguish traffic flow conditions for crash and non-crash cases and also to be able to adjust for external factors such as driver population or geometric features. The study shows that a high value of CVS represents frequent queue forming and dissipation and that under these conditions crashes are more likely to occur.

Furthermore, it is concluded that by observing the coefficient of variation pattern from sets of three detectors a crash prone *location* can be identified. This finding implies that loop detector data can be utilized to identify crash prone conditions.

An improvement to the latter approach was carried through in a follow-up study by Abdel-Aty, Uddin & Pande (2005) that also investigates crash prone conditions, but does so by developing separate models for low-speed and high-speed operating conditions. The database was expanded for both the number of road sections and the period of time for the crashes observed. The distinction between a low and high-speed regime was made on the basis of a histogram of the average speed observed from the loop detector located nearest to the crash and the corresponding percentage of crashes. The two regimes were separated at a speed of about 60km/h (37.5mph). It is observed that there are more crashes under the low-speed regime than under the high-speed regime.

In the study four explanatory variables are tested: 1) the coefficient of variation in speed (CVS), 2) average lane occupancy (AO), 3) standard deviation of volume (SV), and 4) average volume (AV). The resulting model for a low-speed regime suggests that high variation of speeds around the crash location is a crash precursor. Moreover, it is shown that the shock wave generated by the downstream propagation of the congested flow

regime travels in less than 5 minutes from two stations downstream the crash location to one station upstream.
Around the crash location, small variations in volume across lanes coupled to a high variation in speeds result in a higher crash likelihood.

A study by Lee, Abdel-Aty & Hsia (2006) investigated the crash precursors of side-impact crashes. They analyzed data from a 58 km motorway corridor and included crashes that took place over a period of four years. The data was plotted in a flow-occupancy graph which shows that side-impact crashes occur under uncongested flow conditions, while rear-end crashes occur under both congested and uncongested conditions. The authors state that the result is in line with the literature which suggests that side-impact crashes are related to lane change frequency and, therefore, they are also related to the level of congestion. The statistical analysis includes variables such as average speed, flow and occupancy as well as variation in speed and flow. Geometric and peak/off-peak factors were also considered. An overall average flow ratio (OAFR) measure was developed that accounts for lane change frequency. It was calculated on the basis of the average flows on each lane at a given time interval.
The observation time period was 5 to 10 minutes before the time of the crash. The authors find that overall average flow ratio, the coefficient of variation in flow and the peak/off peak period are good indicators of the likelihood of side-impact crashes. The study shows that when the flow variation across lanes is higher, side-impact crashes are more likely than rear-end crashes. During off-peak periods higher values of OAFR and of the coefficient of variation in flow are linked to increased odds of side-impact crashes. This implies that during congested periods (peak) the likelihood of rear-end crashes is higher when compared to side-impact crashes.

3.2. What happens at the tail of the queue and ‘within’ the queue

The literature on the impact of the traffic conditions in the queue on safety is not very extensive. Nevertheless, a few studies have dealt with this topic and they are discussed below.

It was mentioned in the previous section that Lee, Hellinga & Saccomanno (2003) studied the effect of crash precursors on crash potential. One of the crash precursors they took into account was Q, the speed difference between the upstream and downstream end of the road section. High speed differences within a short distance of a road section are characteristic traffic flow conditions of queue formation or dissipation. It was determined that a 2-minute observation period before the crash occurrence yielded the best crash precursor data. They found that high values in parameter Q increase the likelihood of crash occurrence. This finding suggests that crashes are likely to be more frequent during the formation and dissipation of the queue.

Abdel-Aty & Pande (2006) studied the characteristics of rear-end crashes on a motorway using a spatial-temporal approach.
Crash data from the period 1999 to 2003 yielded 1620 rear-end crashes. The crash location was linked to the nearest loop detector, which is called station F. The time period preceding the crash was divided into 5-minute time slices. Average speeds (over all lanes) 5 and 10 minutes before the crash occurrence were calculated (AS_1 and AS_2, respectively) for two

stations upstream (station D) and two stations downstream (station H) of the crash location, see *Figure 6*.

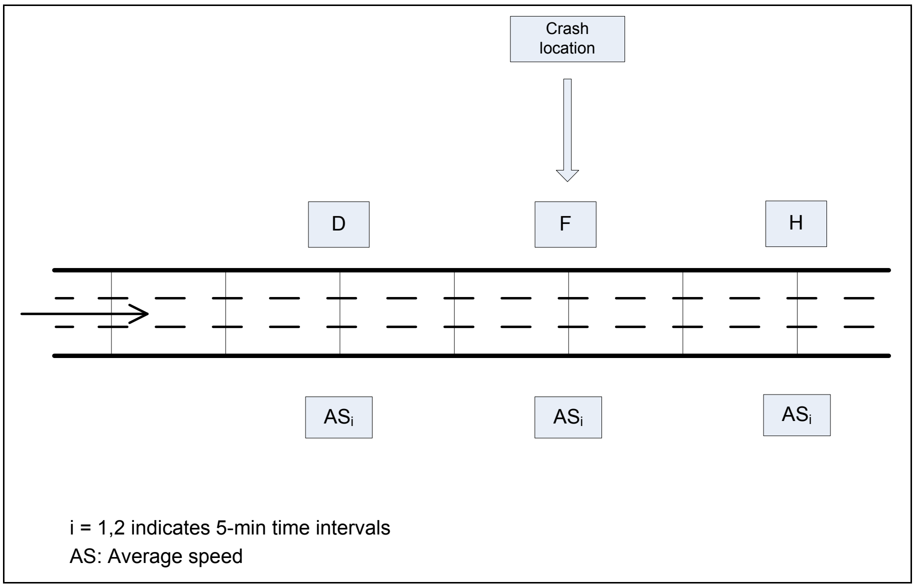

Figure 6. *Loop detector stations D, F and H for a given crash (Adapted from Abdel-Aty & Pande, 2006).*

They used speed conditions 5 to 10 minutes before the crash occurrence for further analysis (time slice 2). On the basis of the speed conditions at stations D, F and H, rear-end crashes are then grouped into two speed regimes (Note that the way in which the speed regime is classified differs from the study by Abdel-Aty, Uddin & Pande (2005) presented in *Section 3.1.2*). Crashes that took place under regime 1 present rather low speeds at all stations, 5 to 10 minutes before the crash took place. Crashes under regime 2, however, show rather high speeds at all stations. The severity of the crashes under these regimes is analyzed. It is concluded that both property damage-only and fatal crashes are more frequent under regime 2 (higher speeds) than under regime 1. The authors suggest that crashes classified under regime 2 occur as the queue starts to build up, since the regime is characterized by higher speeds in the 10 minutes before the crash. In contrast, the authors imply that crashes under regime 1 are related to the dissipation of the queue. The authors imply therefore, that rear-end crashes at the tail of the queue (when the queue is building up) are more likely to occur than at the head of the queue (when the queue is dissipating).

4. Conclusions and further research

This literature review was meant to investigate the relationship between congestion and safety on motorways. The hypotheses that were explored are the following:

1. Traffic congestion levels have been increasing significantly in the last few decades, and this trend seems to continue. The general perception is that crash frequency increases with increasing congestion levels. Nevertheless, it is expected that severe crashes will not increase under these traffic conditions.
2. Crash frequency, severity and type are expected to be affected by the changing flow conditions that occur when traffic starts to become congested.
3. Once traffic is congested, fewer (serious) crashes are expected to occur within the queue. In contrast, at the tail of the queue more severe rear-end crashes are to be expected.
4. It is expected that the effects of congestion on safety depend on the extent to which the drivers are surprised by the congestion. This may, in its turn, depend on the type of congestion, the location of the queue and the use of variable message signs.

4.1. Conclusions

Two types of studies were discussed that deal with the relationship between flow conditions and crash severity, frequency and type. The first uses aggregated data based on long observation periods (e.g. hourly periods), which are then aggregated over say a year. These studies compare various locations and/or different time periods (e.g. peak and non-peak). The second type of study uses short observation periods (e.g. 5 minutes) and analyses which traffic conditions result in crashes. These studies based on disaggregated data allow for the observation of flow changes in short intervals.
Studies with disaggregated data are more appropriate for investigating hypotheses 2 and 3. These studies are better able to capture short-term variations in traffic volumes that occur during the building up and breaking down of congestion.

Based on the literature discussed in the previous sections the following conclusions may be drawn. With regard to the first hypothesis, results of different studies are not consistent. Some studies (Chang & Xiang, 2003; Zhou & Sisiopiku, 1997) find that high V/C ratios result in higher crash rates but less severe crashes. According to Lord et al (2005), however, crash rates decrease at high density levels. Wang, Quddus & Ison (2009) find no relationship between congestion and crash frequency nor between congestion and crash severity. None of these studies explicitly provide information about the influence of congestion on crash rate.

When looking at the traffic processes in more detail, the results regarding the second hypothesis seem to be rather consistent. The studies suggest that crash likelihood increases as the speed variability increases (a typical indicator for unstable traffic conditions). Also large speed differences between lanes and density variability seem to increase the crash likelihood.

As mentioned before, crash severity decreases with increasing volumes (or V/C ratios) according to the studies based on aggregated data. Golob, Recker & Pavlis (2008), who based their study on disaggregated data, report somewhat different results: in the transition from free flow to congested conditions, crash severity does not seem to vary due to the traffic situation. However, once traffic is congested crash severity decreases.
Furthermore, the studies seem to show consistent results with regard to crash type. Golob, Recker & Pavlis (2008) indicate that under unstable conditions, rear-end crashes are more likely to occur. Even though Lee, Abdel-Aty & Hsia (2006) only differentiate between peak and non-peak periods, they arrive at a similar conclusion; rear-end crashes are more likely to occur during unstable flow conditions. This is supported by studies with aggregated data, which suggest that rear-end crashes are more frequent than side-impact crashes as volumes increase. In general, as traffic becomes congested, multiple vehicle crashes are more frequent than single vehicle crashes.

The third hypothesis could not be fully proved due to the scarcity of the literature on the topic. Nevertheless, the studies which rely on disaggregated data provide some insight in this respect. The study by Lee, Hellinga & Saccomanno (2003) suggests that crashes are more likely to occur during the formation and dissipation of the queue. The crash likelihood within the queue was not addressed in their study, nor in the study by Abdel-Aty & Pande (2006). They suggest that more rear-end crashes occur as the queue builds up than as the queue dissipates. The distinction between these two situations was made on the basis of the average speed prevailing before the crash, and therefore crashes in non-congested conditions could also be included in the sample. Thus, the study does not provide a complete picture about the crash frequency and severity at the tail of the queue. The study by Golob, Recker & Pavlis (2008) (also based on disaggregated data) reports some findings that could be related to the situation within the queue. It is observed that once the traffic is congested, crash severity is greatly reduced when all lanes present similar flow conditions. Under these conditions of uniformity, crashes with (fixed) objects are more likely to occur. When only the left and interior lanes are congested, rear-end and side-impact crashes are more likely.

The fourth hypothesis about the effects on safety of structural and incidental congestion could not be investigated on the basis of the available literature.

In summary, the literature discussed provides insight into the relations between unstable traffic conditions and crash frequency, severity and type. Moreover, it offers a better understanding of rear-end crash occurrence at the tail of the queue.

Although the available literature does not provide a complete picture of the relationship between congestion and safety, it does present an overview of the current status of research on this subject. However, several questions are still to be answered. As increasing congestion seems to pose a significant problem to traffic safety, there is need for further research.

4.2. Further research

There are a number of hypotheses that could not be fully investigated and therefore, further research is needed. The remaining hypotheses are the following:

1. Congestion increases crash rates and reduces crash severity.
2. Rear-end crashes at the tail of the queue are more severe than crashes within the queue.
3. Crashes are less frequent within the queue than at the tail of the queue.
4. It is expected that the effects of congestion on safety depend on the extent to which drivers are surprised by the congestion. Factors such as the location of queues, variable message signs or automatic incident detection might play an important role in the relationship between traffic safety and congestion.

There are three potential research methods that could be adopted for a follow-up study: modelling, observation or data analysis.
Modelling refers to the use of a traffic model (e.g. MIXIC) to analyse processes that occur in cases of unstable traffic flow or congestion.
Observation refers to the analysis of video images of the traffic on motorways. This makes it possible to observe the prevailing traffic conditions and the processes that take place as the flow becomes congested.
Data analysis refers to the study of data that is collected after-the-fact (in this case, crashes). Crash records can be linked to the NWB (Nationaal WegenBestand or translated “National Road Database”) and data on traffic volumes as well as congestion can be obtained from Rijkswaterstaat (the Dutch Ministry of Waterways and Public Works). This data would allow for the analysis of the circumstances under which crashes occurred.

The advantage of modelling is that different traffic situations can easily be simulated by the traffic model. The main drawback is that the extent to which the method provides answers to the hypotheses depends on how well the software represents reality. Observation provides more insight into processes that might result in crashes. However, in order to study the effects of congestion on safety, an important number of crash situations should be observed to allow for a statistically sound analysis. Such a requirement has the disadvantage of being extremely time-consuming. Modelling as well as observation is mainly useful to provide more insight into processes in unstable or congested traffic conditions that may result in crashes. The remaining research questions however, do not concern processes that result in crashes, but concern severity of crashes, the risk of crashes within a queue and factors that influence the relation between congestion and traffic safety. *Data analysis* is a more appropriate research method to answer these questions. Therefore, this method is recommended for further research. By linking crash data to data on traffic volume and flow conditions, crash rates during congestion and non-congestion conditions can be compared. In that way, insight can be obtained into the effect of congestion on crash rates and crash severity in general. Moreover, crash rates can be compared for different types of congestion (e.g. recurrent vs. non-recurrent). Finally, by analysing the time and location of crashes in more detail, it can be investigated in which part of a queue most crashes occur.

Interviews with experts in the field will help select the best approach to carry out the follow-up study. Cooperation with the Delft University of Technology, would offer more understanding of the theory on traffic processes and on the type of data required for investigating the hypotheses.
Rijkswaterstaat – Dienst Verkeer en Scheepvaart (the Dutch Ministry of Waterways and Public Works – Department of Traffic and Navigation) could be asked to provide information with regard to data availability on congestion and traffic volumes and to the type of data analysis they employ.

References

Abdel-Aty, M. & Pande, A. (2005). *Identifying crash propensity using specific traffic speed conditions.* In: Journal of Safety Research*,* vol. 36, nr. 1, p. 97-108.

Abdel-Aty, M. & Pande, A. (2006). *Properties of freeway rear-end crash clusters.* In: Proceedings of the 11th International Conference of the Hong Kong Society for Transportation Studies. Hong Kong.

Abdel-Aty, M., Uddin, N. & Pande, A. (2005). *Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways.* In: Transportation Research Record: Journal of the Transportation Research Board, nr. 1908, p. 51-58.

Chang, G. & Xiang, H. (2003). *The relationship between congestion levels and accidents*. Maryland State Highway Administration, Baltimore, United States.

Dijkstra, A. & Drolenga, J. (2008). *Safety effects of route choice in a road network: Simulation of changing route choice*. R-2008-10. SWOV Institute for Road Safety Research, Leidschendam.

Driel, C. van (2007). *Driver support in congestion. An assessment of user needs and impacts on driver and traffic flow.* Proefschrift University of Twente, Enschede, the Netherlands.

Duivenvoorden, K. (2010). *The relationship between traffic volume and road safety on the secondary road network; A literature review*. D-2010-2. SWOV Institute for Road Safety Research, Leidschendam.

DVS (2008). *Bereikbaarheidsmonitor Hoofdwegennet 2007*. Dienst Verkeer en Scheepvaart, Delft, The Netherlands.

DVS (2009). *Nationale Mobiliteitsmonitor 2009*. Dienst Verkeer en Scheepvaart, in opdracht van de Stuurgroep Nationale Mobiliteitsmonitor, Delft, The Netherlands.

Golob, T.F., Recker, W. & Pavlis, Y. (2008). *Probabilistic models of freeway safety performance using traffic flow data as predictors.* In: Safety Science*,* vol. 46, nr. 9, p. 1306-1333.

Hall, F. (2001). *Traffic stream characteristics*. In: Gartner, N., Messer, C.J. & Rathi, A.K. (red.), Traffic flow theory. A state-of-the-art report. Committee on Traffic Flow Theory and Characteristics, United States, p. 2.1-2.33.

Hoogendoorn, S.P. (2007). *Traffic flow theory and simulation.* Lecture notes course vk4821. Transportation and Traffic Engineering Section. Delft University of Technology, Delft.

Hourdos, J., et al. (2006). *Real-time detection of crash-prone conditions at freeway high-crash locations.* In: Transportation Research Record: Journal of the Transportation Research Board, nr. 1968, p. 83-91.

Janssen, L.H.J.M., Okker, V.R. & Schuur, J. (2006). *Welvaart en Leefomgeving*. Centraal Planbureau, Natuur en Milieu Planbureau, Ruimtelijk Planbureau, Den Haag, the Netherlands.

Jun, J. & Lim, I.-K. (2008). *A potential congestion severity measure from the percent of congestion and the duration of continuous congestion.* In: Proceedings of the 87th TRB Annual Meeting 2008. Washington DC, Transportation Research Board TRB.

KiM (2008). *Mobiliteitsbalans 2008*. Kennisinstituut voor Mobiliteitsbeleid KiM, Den Haag, The Netherlands.

KiM (2009). *Mobiliteitsbalans 2009*. Kennisinstituut voor Mobiliteitsbeleid KiM, Den Haag, The Netherlands.

Lee, C., Abdel-Aty, M. & Hsia, L. (2006). *Potential real-time indicators of sideswipe crashes on freeways.* In: Transportation Research Record: Journal of the Transportation Research Board, nr. 1953, p. 41-49.

Lee, C., Hellinga, B. & Saccomanno, F. (2003). *Real-time crash prediction model for application to crash prevention in freeway traffic.* In: Transportation Research Record: Journal of Transportation Research Board, nr. 1840, p. 67-77.

Lee, C., Saccomanno, F. & Hellinga, B. (2002). *Analysis of crash precursors on instrumented freeways.* In: Transportation Research Record: Journal of Transportation Research Board, nr. 1784, p. 1-8.

Lord, D., Manar, A. & Vizioli, A. (2005). *Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments.* In: Accident Analysis & Prevention*,* vol. 37, nr. 1, p. 185-199.

May, A.D. (1990). *Traffic flow fundamentals*. Prentice-Hall, Englewood Cliffs, NJ.

Oh, C., et al. (2001). *Real-time estimation of freeway accident likelihood.* In: Proceedings of the 80th TRB Annual Meeting 2001. Washington, D.C., Transporation Research Board TRB.

Rietveld, P. & Shefer, D. (1994). *Congestion and safety on highways : towards an analytical model.* In: Proceedings of the Third International Conference on Safety and the Environment in the 21st Century : lessons from the past, shaping the future. Tel Aviv, Israel, p. 56-78.

SWOV (2007). *Network Management and Sustainable Safety*. Fact sheet, November 2009. SWOV Institute for Road Safety Research, Leidschendam.

Transportation Research Board (2000). *Highway Capacity Manual 2000; metric version*. National Research Council, Transportation Research Board TRB / National Academy Press, Washington, D.C.

Wang, C., Quddus, M.A. & Ison, S.G. (2009). *Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England.* In: Accident Analysis & Prevention, vol. 41, nr. 4, p. 798-808.

Zhou, M. & Sisiopiku, V. (1997). *Relationship between volume-to-capacity ratios and accident rates.* In: Transportation Research Record: Journal of the Transportation Research Board, nr. 1581, p. 47-52.