Number of Lane Changes Determined by Splashover Effects in Loop Detector Counts

Victor L. Knoop PhD (corresponding author)

Delft University of Technology	Université de Lyon
	INRETS-ENTPE
Transport & Planning	Laboratoire d'Ingénierie
	Circulation Transports (LICIT)
Stevinweg 1	Rue Maurice Audin
2628 CN Delft, the Netherlands	69518 Vaulx-en-Velin, France
Phone: +31 15 27 81723	+33 4 72 04 77 15
v.l.knoop@tudelft.nl	

R. Eddie Wilson
Christine Buisson PhD
Bart van Arem

INTERNAL REPORT TU DELFT

Please cite this work as: Knoop, V.L., Buisson, C., Wilson, E. and Van Arem, B. (2011) Number of Lane Changes Determined by Splashover Effects in Loop Detector Counts , in: Proceedings of the 90th Annual Meeting of the Transportation Research Board, 23-27 January 2011, Washington D.C.

ABSTRACT

Lane changes are important in quantifying traffic, both for operational purposes as well as for planning purposes. Traditional in-lane loop detectors do not provide this information, hence traffic engineers have estimated lane changes using other data sources. This paper provides a method to estimate the number of lane changes based on the number of vehicles which are measured by two loops in adjacent lanes. It is argued that two adjacent loops only record a vehicle simultaneously if the vehicle is changing lanes. A stochastic process determines whether a lane changing vehicle will "straddle", i.e. be detected by two loops. The resulting probability distribution of straddles is then determined by the number of lane changes. This stochastic process is analyzed and used to develop a method to get from a straddle count to a distribution of lane changes. In the evaluation considered here, the number of lane changes per aggregation period varies by a factor 10. The estimation procedure is very well capable of determining this varying number of lane changes. Using the number of straddles as input (measured at the road) the number of lane changes can be predicted with an accuracy of 10%.

1 INTRODUCTION

Despite the fact that they play a crucial role in traffic management and operations, lane changes are not detected by traditionally used detectors. This is due to the nature of the measurement device. Loop detectors are point detectors able only to measure what happens locally and therefore no information on lane changing is available. As a consequence, lane changing models are calibrated on other types of data, for example the distribution of traffic counts over the lanes (*1*) or video data.

Lane changes are important in two aspects of traffic: first it appears that they induce stop and go waves (e.g., Ahn and Cassidy (*2*)) and explain a part of traffic congestion (e.g., Cassidy et al. (*3*)), which both are operational aspects. Second, a better estimate of the number of lane changes improves estimating the OD matrix. They are known for a city as a whole, in a static way and from long and costly surveys. We will show that for operational applications like freeway management, our method permit to obtain a dynamic local OD matrix in case where classical methods are infective, for instance in the situation of a couple of successive merge and diverges of freeways.

The purpose of the present paper is to establish an operational method to estimate the number of lane changes by an innovative approach to loop detector data. The principle exploits the non-zero width of the car and of the physical detector. The basis of the methodology is that it is studied what fraction of the cars is detected by the loops *in two adjacent lanes* (this is called a splashover of the signal). A vehicle which activates two detectors will be called a “straddling” vehicle. The essence of the proposed methodology is that the number of “straddling” vehicles for a detector gives an estimation for the number of lane changing vehicles at the detector and thus in the road section.

It is a stochastic process whether a vehicle will induce a splashover signal during a lane change. Therefore, there is not a one-to-one relationship between the number of lane changes in a section and the number of measured vehicles straddling at the detector. However, this paper will derive the stochastic relationship between the number of measured straddles and the number of lane changes.

The set-up of the remainder of the paper is as follows. First, section 2 gives an overview of the state of the art in local measurements as well as the estimation of the number of lane changes. Then, section 3 explains which data is typically available from loop detectors and how we use them to find the lane changes and the straddles. Section 4 shows the microscopic properties of a lane change for a single vehicle and some properties of straddling which could be derived from these data. This is a foundation for the theory of a tool to predict the lane change distribution. In section 5 the mathematical basis of the method is explained, and the method is calibrated. Section 6 shows how the method works in practice and shows the results of the use of the tool on a different data set. Section 7 shows the influence of the speed and the distance between different detector sites. This section also contains an example case showing the use of the number of lane changes for an OD matrix estimation. Finally, the conclusions are presented in section 8.

2 LITERATURE REVIEW

Loop detectors are the most traditional way to measure traffic. If the detectors are constructed in such a way that the lateral distance between two detectors in two adjacent lanes is smaller than a vehicle width, a vehicle should always be detected by at least one detector, but possibly by two. Such a double count is usually considered an erroneous count (*4*). Using a good loop layout, i.e., loops which are wide enough related to the lane width, it can be avoided that cars can pass in between two detectors. Coifman (*5*) show differences in detection rates and the errors of detectors and how these vary among different detectors. Magnetometers work basically the same as loop detectors, but also have the same errors, and even more vehicles pass unnoticed (*6*).

Several projects have start up in the past 5-10 years to collect traffic data on video, for instance the NGSIM project (*7*) or video data recorded with a helicopter (e.g., (*8*)). The resulting trajectory data can give the number of lane changes, since it provides spatio-temporal information. However, the treatment of video-data is not yet as automated that even if one could collect large quantities of data, or collect it at a regular basis. It would take too much time to extract the trajectories from all these data (*9, 10, 11*). Therefore, this needs to be done off-line and for small quantities of data, hence studies to a statistical effect, requiring large amounts of data, cannot be performed using this type of data.

Lane changes have been studied for several decades. For instance, Wardrop (*12*) already discusses the amount of lane changes that should be measured as function of density. In a microscopic view, the work of Gipps (*13*) shows situations in which drivers want to change lanes and can do so. Also more recently, there have been theories of lane changes (e.g., Laval and Daganzo (*14*)). McDonald et al. (*15*) were among the first to publish on the systematic empirical observations on changes. This was done by analyzing video images on a three-lane freeway. They found that "it is extremely difficult to the model (of lane changes) for a wide condition of flows and conditions". A very large empirical collection of data has been carried out by Olsen et al. (*16*) showing much details of individual lane changes, such as the reasons and the duration. Also, Duret and Buisson (*17*) has shown the effects of lane changes.

Knoop et al. (*18*) took on the challenge left by McDonald et al. (*15*) again and showed a stochastic relationship between the number of lane changes and the density. However, the relationship is a statistical one, which might be site-dependent. It is therefore still useful to have a measurement of the number of lane changes independent of the density. Other measures for the number of lane changes are difficult to find.

3 DESCRIPTION OF THE DATA USED IN THIS PAPER

The aim of this paper is to show how the frequency of lane changing on highways can be estimated using inductance loop detector data from a single detector site. The key requirement is that at this site, each lane of the highway is equipped with a (double) inductance loop detector and that there is some facility for identifying when a single vehicle overlaps (or *straddles*) the detectors in adjacent lanes. These conditions are realized in a recent study on the M42 motorway in the United Kingdom, and moreover, this study also includes the reconstruction of individual vehicle trajectories. Thus lane changes can be observed independently of straddles and the correlation between lane-changing and straddles can be analyzed. Firstly section 3.1 gives general background details of the M42 study and then section 3.2 gives further details of the inductance loop hardware.

3.1 Details of the instrumented section

The data used for this paper come from the M42 motorway near Birmingham in the United Kingdom which has an unprecendented resolution of inductance loop detectors: in a 16km section of the North-bound carriageway there are 183 detector sites, at each of which is located a double inductance loop detector in each lane of the highway. In normal operation, these detectors record 1-minute traffic statistics, but during September 2008 an approximately 1 mile section (see 1) incorporating 16 detector sites was enhanced so that amongst other improvements, full individual vehicle data was recorded at each site. At each site, the time and lane number of each passing vehicle are recorded in addition to estimates of the vehicle's length and speed.

Since the detector sites are separated by only 100m nominally (of course, exact separations are known), vehicles do not often decelerate or accelerate too markedly between sites, so it becomes possible to re-identify vehicles from site to site, and thus in effect follow them down the highway (an idea pioneered by Coifman and Cassidy (*19*)). For further details, see (*20, 21*)). Figure 2 illustrates how the re-identification process works. In particular, lane changes may be identified and they have been analyzed by Knoop et al. (*18*).

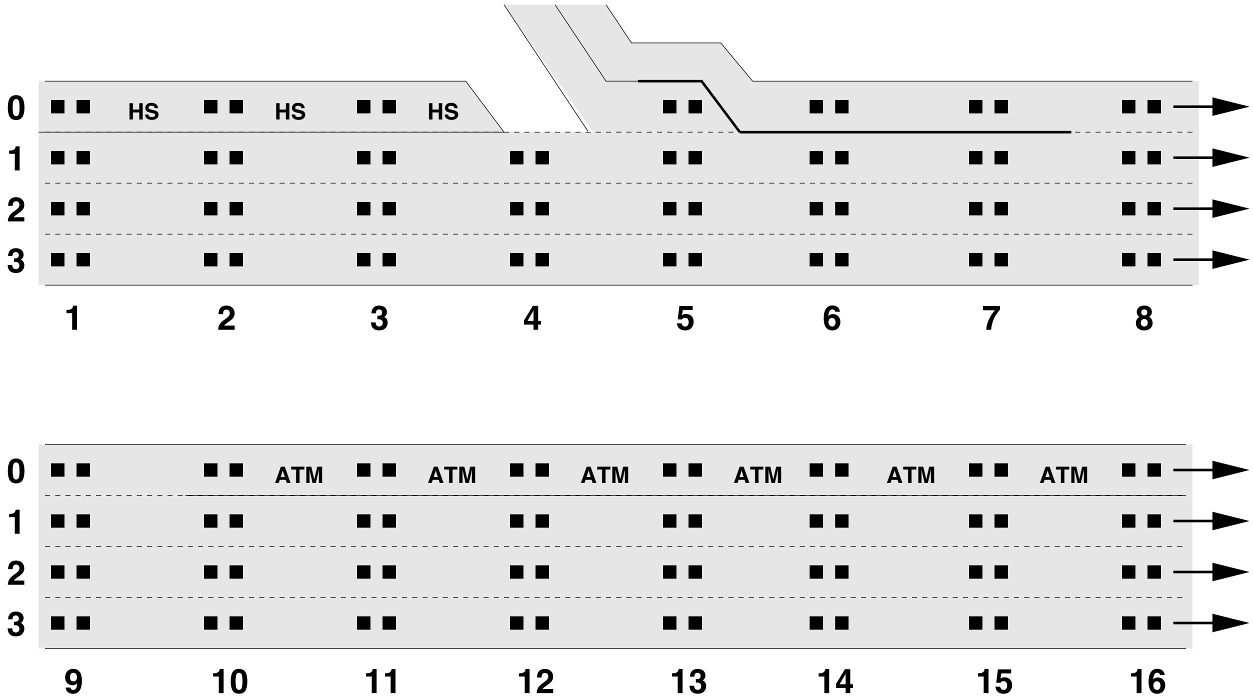


FIGURE 1 Geometry of the instrumented section showing the location of double loop detector sites. Note that in the UK, slower vehicles drive on the left. HS denotes the hard shoulder (emergency lane) and ATM (Active Traffic Management) denotes hard shoulder which can be activated as an ordinary running lane in peak periods. The on-ramp has both a right-hand *forced merge* lane and a slower left-hand lane typically used by heavy goods vehicles (HGVs). When ATM is turned off, HGVs are required to join the main carriageway (lanes 1–3) between sites 8 and 10 inclusive. In this study we try to remove the effect of the on-ramp as far as possible and focus on the section between sites 14 and 16.

One drawback of this approach is that the error rate of the re-identification algorithms is not known exactly. However, as part of a commercial trial, a small sample was manually cross-checked with video data and this process indicated an error rate in free flow conditions of order 1 in 1000 (personal communication, IDRIS Diamond Consulting Services, (*22*)). In contrast, in very congested or strongly dynamic traffic conditions, the re-identification process fails entirely. This study is therefore restricted to fluid traffic conditions (speeds of over 20 m/s = 44 mph).

For the purposes of this study, we have constrained ourselves to a particularly clean stretch of data (with few equipment outages etc.) taken from the period 1 October to 30 November 2008 inclusive; data from periods with failures is discarded completely.

3.2 Details of the inductance loop hardware

The loop detectors considered here are based on the principle of electromagnetic induction. The inductance of coils of wire buried in the road surface is altered by the presence of any nearby objects that conduct electricity, and thus the proximity of (metallic) vehicles close to the loop may be detected by signal processing electronics located in the road-side outstations.

In standard inductance loop installations, the auxiliary electronics detects the times at which the inductance in the loop rises and falls below given threshold levels, corresponding roughly to the passage of a vehicle's front and rear end. For single loop installations, this permits the direct counting of vehicles

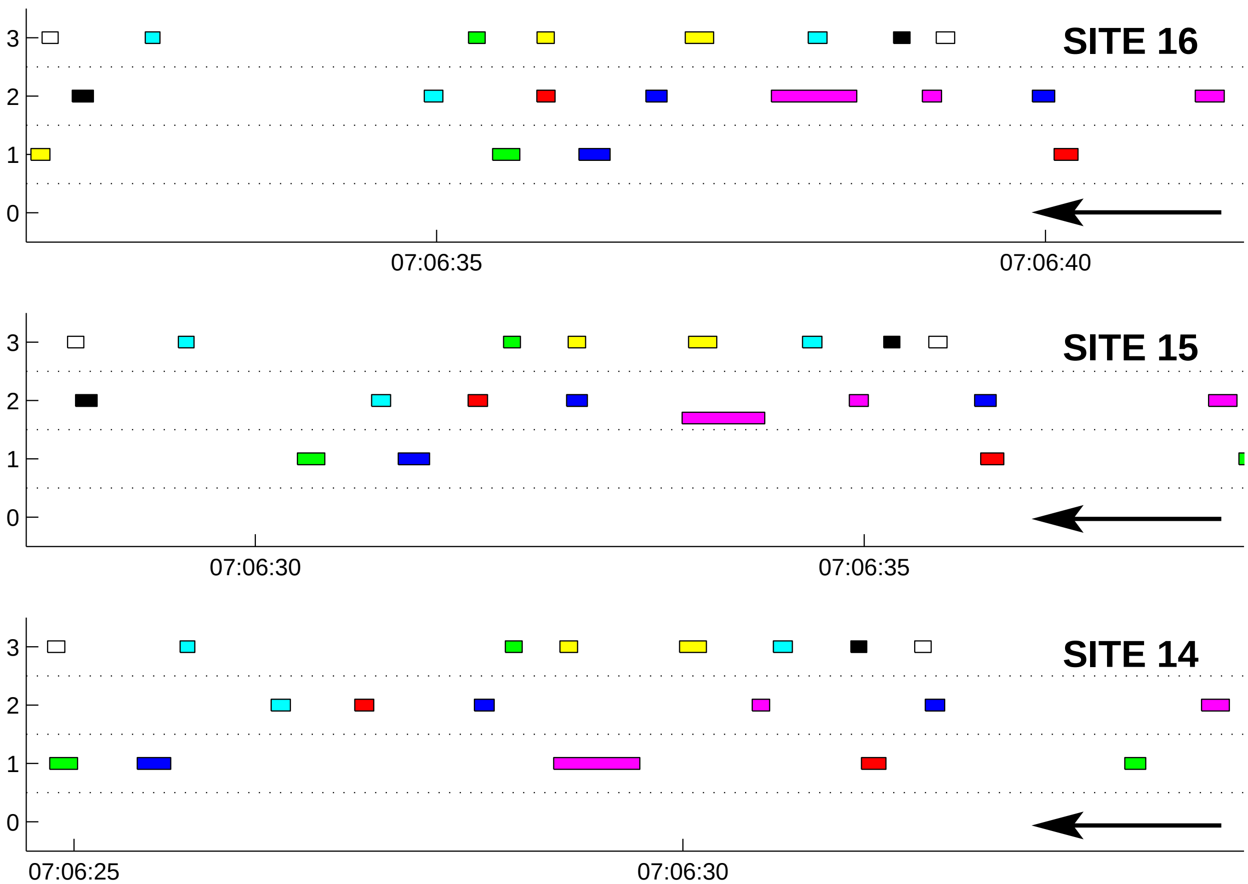


FIGURE 2 Visualization of a short section of Individual Vehicle Data from three inductance loop sites labeled according to figure 1. Since the time axis runs to the right, vehicles apparently drive to the left in each pane of the picture, thus in effect providing a 'helicopter view' of the traffic at each of the sites. The coloring of each vehicle helps indicate the re-identification of vehicles between sites as obtained by our automated algorithms. In this example, we observe a Heavy Goods Vehicle pulling out from lane 1 to lane 2 between sites 14 and 15, which has been *straddling* at site 15.

and the measurement of occupancy. However, in double inductance loop installations as we have here, two loops are installed which are separated longitudinally by several meters. Consequently each passing vehicle gives an "on" and "off" time at each of the two loops and these four times may be processed via divided differences to estimate the vehicle's length and speed.

However, the loops in the instrumented section are equipped with the advanced IDRIS (*22*) signal processing system which also takes into account the inductance signature of each passing vehicle: that is the inductance profile of each passing vehicle is captured as a fine-resolution function of time. In particular, this inductance signature can be used to help re-identification and also to classify vehicle type (since axles show up very clearly in the signature).

When there is a signal at the detectors in two adjacent lanes, the system compares the two signatures in order to determine whether they are due to two separate vehicles, or to one single vehicle that *straddles* the two detectors. The latter is sometimes known as a "splashover" effect, see Lee and Coifman (*4*), and its analysis can be used to correct for errors such as double counting. However, we will use this information to estimate the lateral position of each vehicle on the road. For the case at hand, the lanes are 3.5 meters wide. The detector at lane 2 is 2.4 meters wide, and the one on lane 3 is 1.8 meter wide. The distance between the two detectors is approximately 1.5 meter.

In data thus obtained from the Idris system, each vehicle record has in addition to its lane data a straddling flag which takes one of three values:

- *no straddle*: the vehicle is in the centre of its lane.
- *straddle left*: the vehicle is overlapping partially with the lane to its left.
- *straddle right*: the vehicle is overlapping partially with the lane to its right.

Note that a vehicle changing lane to the right, from lane i to lane $i + 1$, should pass through the following sequence of lateral positions: (1) no straddle in lane i, (2) straddle right in lane i, (3) straddle left in lane $i + 1$, (4) no straddle in lane $i + 1$. However, because the duration of a lane change is small compared with the typical spacing between detector sites, it may be that no straddles are recorded for a given lane change at any detector site.

Finally, note that there are various tolerances which can be set within the Idris system so as to affect the detection of straddles, although experiments upon these have not been performed in the present study. But in particular, it may be possible to lower the detection threshold parameters so that straddles (and hence lane changes) are detected more frequently, although this will involve a trade-off with an increasing number of false positives when no lane changes occur.

4 LANE CHANGES AND STRADDLING – A MICROSCOPIC VIEW

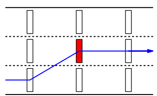
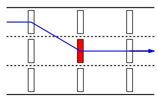
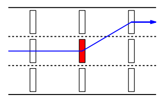
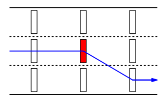
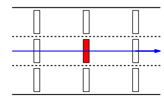
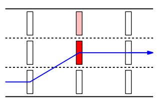
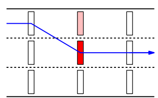
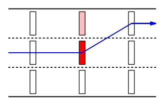
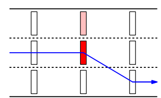
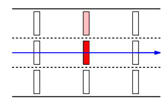
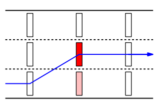
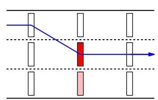
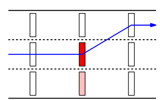
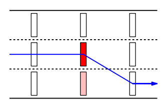
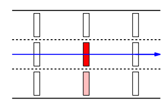
This section now considers the different ways in which the straddle flag at a single detector can correspond to lane changes either up- or downstream of the detector at site 15 (table 1), and we count the various combinations as they occur in a single day of sample data. As we have discussed already, we focus on the straddle flag recorded at detector site 15 and use lane data from sites 14, 15 and 16 (a 200m section) to detect whether a lane change has occurred. Furthermore, we focus our attention on vehicles where are recorded in the middle lane at site 15, as this is the most interesting and general since it invokes changes both to the right and the left in and out of the lane.

Table 1 shows the possible combinations of straddling and lane changing upstream and downstream of the detector site. We focus on one day, the 31 October 2008. 28,101 vehicles pass the middle lane at the detector site. Due to re-identification issues, we exclude the vehicles which pass at less than 20 m/s. The remaining 22,502 vehicles (80%) are analyzed, and grouped according to the possibilities of straddling and lane changes as shown in figure 1.

The lower half of table 1 shows the core ratios which are discussed in the next two paragraphs. Let us focus on the cases with straddling. No vehicles straddle to the ”wrong” side (e.g., first straddling right and than changing lanes to the left downstream or upstream changing from the right and straddling to the left at site 15). This analysis also shows that is is very unlikely that a vehicle that does not change lane, straddles. In fact, only 3 out of the 21,323 (21,320+1+2) cars that go straight on do so. Therefore, we shall simplify matters and assume that none of the vehicles that do not change lanes are detected straddling.

It also shows that there is no considerable difference indicating a asymmetrical lane change manoeuvre. The differences in lane changes left and right within for instance the right straddling (40 cases with a lane change left downstream versus 56 cases with a lane change right upstream) are in line with the number of lane changes without straddle (190 lane changes left downstream versus 306 lane changes right downstream). Since there is not a considerable difference, in the sequel of this paper, all lane changes and all straddles are analyzed together. This means that we will focus on three cases: not changing lanes and no straddle, changing lanes and no straddle and changing lanes with a straddle.

TABLE 1 The different possibilities of straddling and lane changing (upstream, “upstr”, and downstream, “downstr”) at site 15. The figure shows three lanes at sites 14, 15 and 16. The vehicles which are studied are the vehicles which are detected at site 15 at the middle lane (principally - sometimes some splashover effects), hence this detector is colored red. The movement of the vehicle is indicated with an arrow. The splashover effects in another lane, caused by a straddling vehicle, are indicated by pink shading.

LC / Straddle	Left Upstr	Right Upstr	Left Downstr	Right Downstr	None
None	 190	 248	 243	 306	 21320
Left	 0	 48	 48	 0	 3
Right	 40	 0	 0	 56	 3

	Number of obs.	Percentage	Comment
Total number of vehicles observed on lane 2 detector 15	22505		
Number of vehicles observed also on lane 2 on detector 14 and 16 (did not changes lane neither between 14 and 15 nor between 15 and 16)	21326	21326/22505 =94.8%	Inverse of lane change rate: 95% does not change lane; a little over 5% changes lane in 200m
Number of vehicles observed to straddle, but are also on lane 2 on detector 14 and 16 (did not change lanes either between 14 and 15 nor between 15 and 16)	6	6/21326 =0.28%	false positives: less than 3 per 1000
Number of vehicles observed on lane 2 on detector 15 and 16 and either on lane 1 or on lane 3 on detector 14 (change lane upstream of detector 15)	526	526/22505 =2.3%	More than 2% of the total number of all vehicles changed lanes in the 100 meters between detector 14 and detector 15
Number of vehicles observed on lane 2 on detector 15 and 16 and either on lane 1 or on lane 3 on detector 14 (change lane before detector 15) and made a straddle on detector 15	88	88/526 =0.167	Straddle rate: more than 15% of the vehicles that did change lane upstream of detector 15 generated a stradle on detector 15
Number of vehicles observed on lane 2 on detector 14 and 15 and either on lane 1 or on lane 3 on detector 16 (change lane downstream of detector 14)	653	653/22505 =2.9%	More than 2% of the total number of vehicles changed lanes in the 100 meters between detector 15 and detector 16
Number of vehicles observed on lane 2 on detector 14 and 15 and either on lane 1 or on lane 3 on detector 16 (change lane after detector 15) and made a straddle on detector 15	104	104/653 =15.9%	Straddle rate: more than 15% of the vehicles that changed lanes downstream of detector 15 generated a stradle on detector 15

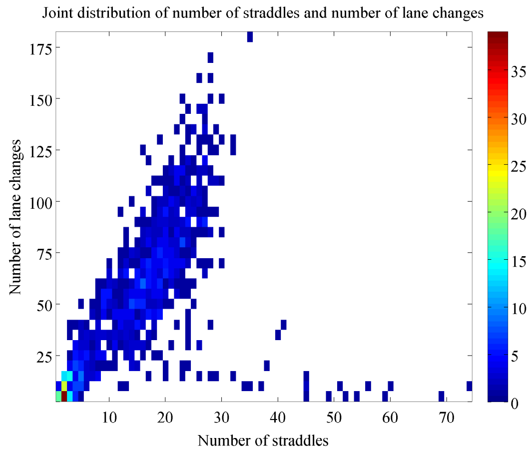


FIGURE 3 The joint distribution of the number of straddles and the number of lane changes in the calibration set. Note that there is a distribution for each bin with a fixed number of straddles.

5 CALIBRATING THE STRADDLING PROCESS

In this section we show the relationship between straddling and lane changing, and provide the numerics for our case study. To perform this calibration, we use a subset of the available data. 2/3 of the two-month data set (see section 3.1) is used for calibration purposes and 1/3 is used to check the predictive value of the straddling later (section 6); the split is random, on a day-by-day basis.

The data is aggregated over intervals of 60 minutes, for which both the number of stradlles, N_S, and the number of lane changes, N_{LC}, are known, because they are both measured. We choose an aggregation interval of 60 minutes since it gives a reasonable number of straddles to work with. A shorter time would be possible, resulting in more aggregation intervals with fewer recorded straddles and lane changes in each of the intervals. Intervals where 20% of the vehicles had a speed of less then 20 m/s were considered (possibly partly) congested and therefore, due to re-identification issues, discarded. However, after this selection, more than 80% of the data is still remaining.

A naive method would be to simply relate the number of lane changes to the number of straddles. Figure 3 shows a figure which shows, in colors, how scattered the data in fact is. A known value for the number of straddles will not provide a crisp value for the number of lane changes. This is due to the stochastic process which lies underneath the detection of straddles from lane changes.

As described in section 4, the straddling process is assumed only to take place once there is a lane change. Furthermore, the stochastic part is that not all vehicles will produce a straddle while changing lanes. It is possible for a vehicle to perform the complete lane change between upstream and downstream detectors without resulting in a straddle flag at either detector. If we assume that all vehicles are the same, the number of straddles recorded is a a binomial function of the number of lane changes. This is to say that for each lane change, there is an independent draw whether there is a straddle recorded or not. The relative number of times that a straddle is recorded, depends on the detector configuration, the distance between the detectors and the threshold setting for a straddling flag.

The probability of getting s straddles in lc trials is thus given by a binomial distribution function

$$P(N_S = s) = \binom{s}{lc} p^s (1-p)^{lc-s} \tag{1}$$

In this equation, p is the rate of "success" for a lane changer to trigger a straddle. Thus, $(1-p)$ is the "failure" rate, that is the rate at which the lane changer does not trigger a straddle (but does change lanes).

N_S is the stochastic variable indicating the number of straddling vehicles measured in an aggregation period. Likewise, N_{LC} will be used to indicate the stochastic variable of the number of lane changes. Note that the number of straddles is a stochastic variable, and only related with a distribution function to the number of lane changes. The distribution function in equation 1 implies that on average

$$N_S = pN_{LC} \tag{2}$$

The only goal of the calibration procedure, is *to find* p. To this end, a maximum likelihood estimator is used, which was used on all data of the calibration set at once. We therefore took together all lane changes and all straddles for the non-congested aggregation intervals. The maximum likelihood estimator gave for our data set to an p of 16.4%, being the percentage of cars that is recorded straddling while performing a lane change. The 95% confidence bounds for p are 16.1% and 16.7%, which shows that p is estimated quite accurately due to the large quantities of data. If one would use one day of data – which we did for testing purposes – the 95% confidence intervals of p are typically 4% (compared to 0.6% in the case with 40 days of data). Two days will lead to a confidence interval of p of around 3%.

6 RECOVERING THE NUMBER OF LANE CHANGES FROM THE STRADDLES

There are two goals of this section:

1. Find the distribution of lane changes for aggregation periods with a given number of straddles in another data set than used to determine p, but recorded at the same site.

2. Show the accuracy of this estimate.

The first goal basically is applying Bayes' rule

$$P(N_{LC}|N_S = S_0) = P(N_S|N_{LC})\frac{P(N_{LC})}{P(N_S)} \tag{3}$$

For the prior $P(N_{LC})$ we use $P(N_{LC} = n) = (1/p)P(N_S = (p * n))$, where $P(N_S)$ is the experimental probability density function. $P(N_S|N_{LC})$ is a binomial function with success rate p.The remainder of the section shows how this rule is applied in practice.

The distribution of the number of lane changes (goal 1) can be read from a matrix showing the joint distribution of straddling and lane changing (see figure 4a). This matrix shows how often combinations of straddling and lane changing occur. Once this matrix is estimated for a validation set, one row of the matrix will give the distribution of the number of lane changes given the number of measured straddles in an aggregation period.

The different parts of this section will explain the different steps to estimate this matrix. The first step is to estimate the total distribution of lane changes (see section 6.1). The second step is to divide this distribution over different values for the number of straddles (section 6.2). Section 6.3 shows the normalization which is required to estimate the number of aggregation intervals in which a specific number of lane changes is given. This last point is only relevant to check the accuracy of the estimation (goal 2 of the section). The results in terms of accuracy for a validation set are given in section 6.4.

6.1 From the measurement of the number of straddles to the distribution of the number of lane changes

The starting point for the method is a set of data for which the number of straddles is measured – here, we will use a validation set which is also used to test the accuracy of the method (goal 2). We will assume here the number of aggregation intervals is large enough to have a good distribution function for the number of straddles, which is required to test the accuracy of the estimate. We applied the method to 1/3 of the data set which was kept separate from the total data set for this purpose; we apply the same restriction as on the calibration set in order to have only uncongested data.

The first step to get to the distribution of the number of lane changes for an interval with a certain number of straddles, is to have an overall distribution for the number of lane changes (see figure 4a). This is based on the observed distribution of straddles per aggregation time (60 minutes) in the validation set, P_S (see figure 5a). The idea is to stretch this distribution to come to the distribution of lane changes, since the average number of straddles per lane change is known from the calibration (p, 16.4%).

Since the bin of one straddle count corresponds to several bins of lane change counts, the distribution function has to be interpolated to give the lane change distribution. That means that for each integer value lc for the number of lane changes, the corresponding average number of straddles is calculated: $s = p * lc$. The value s is generally non-integer, and we call its nearest-neighbors integer numbers s_+ and s_- with $s_- <= s <= s_+$. The estimated probability density function of the lane change $P_{LC}(lc)$ is weighted combination of the measured probability density function of the straddling in s_+ and s_-, where the weights depend on the distance between s and $s-$ and $s+$ respectively.

$$P_{LC}(lc) = (s_+ - s)P_S(s_-) + (s - s_-)P_S(s_+) \quad (4)$$

In case s happens to be an integer number, $s_- = s_+ = s$ and $P_{LC}(lc) = P_S(s)$.

The function P_{LC} finally needs to be normalized such that:

$$\sum_{lc=0}^{\infty} P_{LC}(lc) = 1 \quad (5)$$

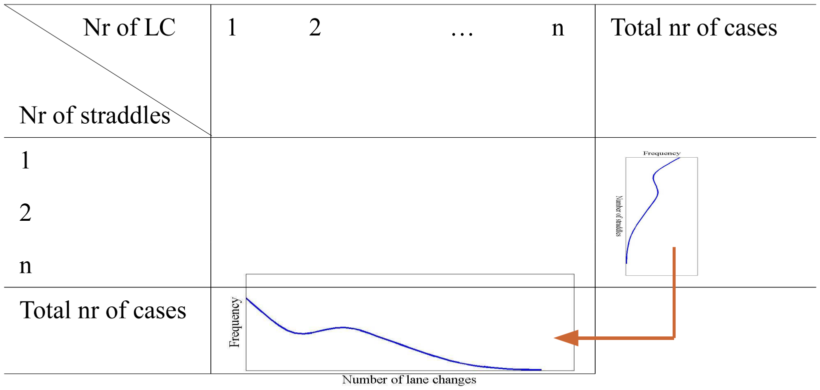
The accuracy of the result can be tested since for the calibration set the number of lane changes is also measured. The estimated probability distribution function and the measured probability distribution function are shown in figure 5b, which shows that the estimation works quite well.

6.2 From the distribution of lane changes to the joint distribution of the number of lane changes and straddles

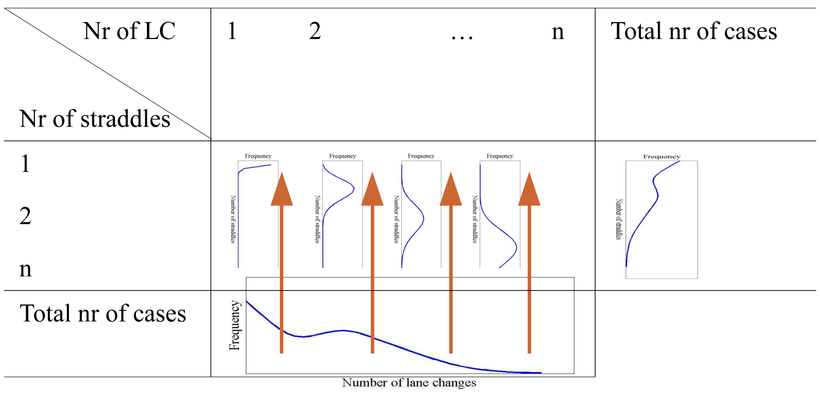
The distribution function for the number of lane changes need to be split up for different number of straddles, Only then, we can find the distribution for the number of lane changes given the number of straddles in an aggregation interval (goal 1). We do this as follows. For each of the bins of the total distribution function of the number of lane changes lc, the distribution of the number of straddles is given by the theory, equation 1. Once the number of lane changes is known (i.e., the column number) as well as the parameter of the binomial distribution function (p from the calibration), this matrix can be filled, column by column. This spreads the distribution of the number of lane changes to a bi-dimensional, joint, distribution showing the coincident probability between the number of lane changes and the number of straddles. Figure 4b shows this graphically.

6.3 Normalization of the distribution function for the number of lane changes for a given number of straddles

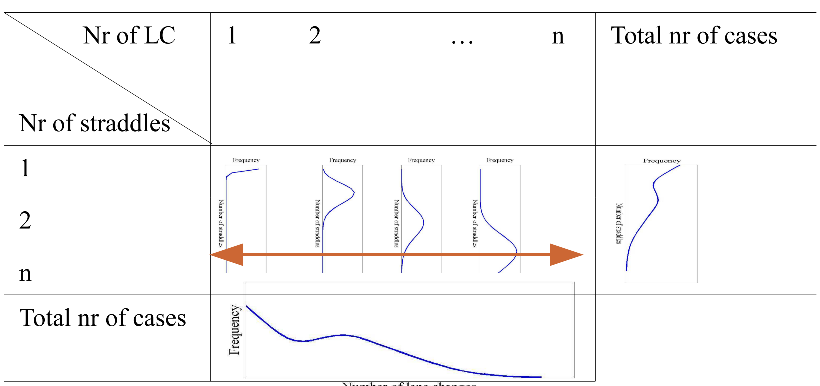
One row indicates the probability of lane changes for a certain value of straddling. However, to check the accuracy of this estimated probability density function with the measured distribution (goal 2 of this section), we need to have a probability distribution function *within one bin of number of straddles* which sums to one. Therefore, to get to a number of observations for each number of lane changes lc within a bin



(a) Step 1 in the process of finding the distribution of lane changes: stretch the measured distribution of straddling is stretched (using $1/p$) to obtain an estimate for the total lane change distribution

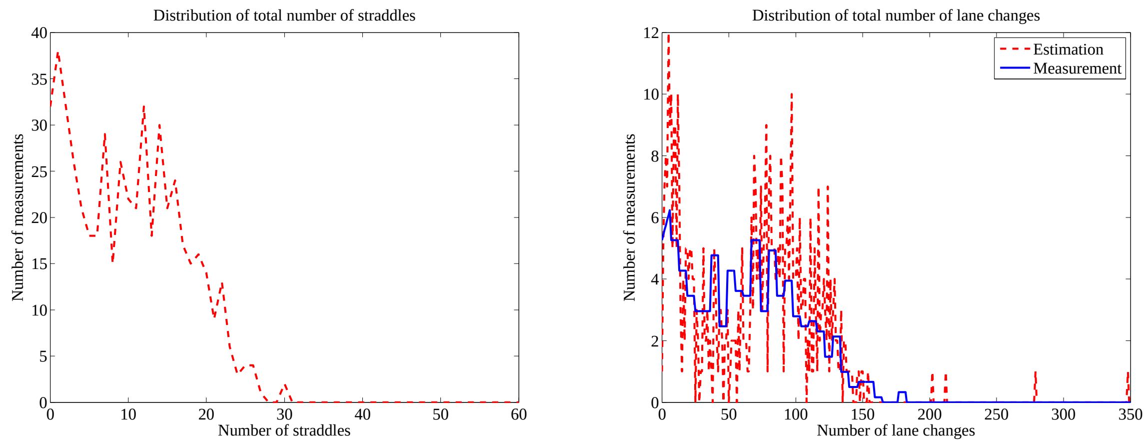


(b) Step 2: The aggregation periods in each of the bins of equal number of lane changing are distributed over different straddling numbers. To this end, a binomial distribution is used.



(c) Step 3 in the process of finding the distribution of lane changes: multiply the distribution by a scale factor such that sum of the obtained numbers in each of the bins of equal lane changes (for the same number of cars straddling) add up to the measured number of straddling.

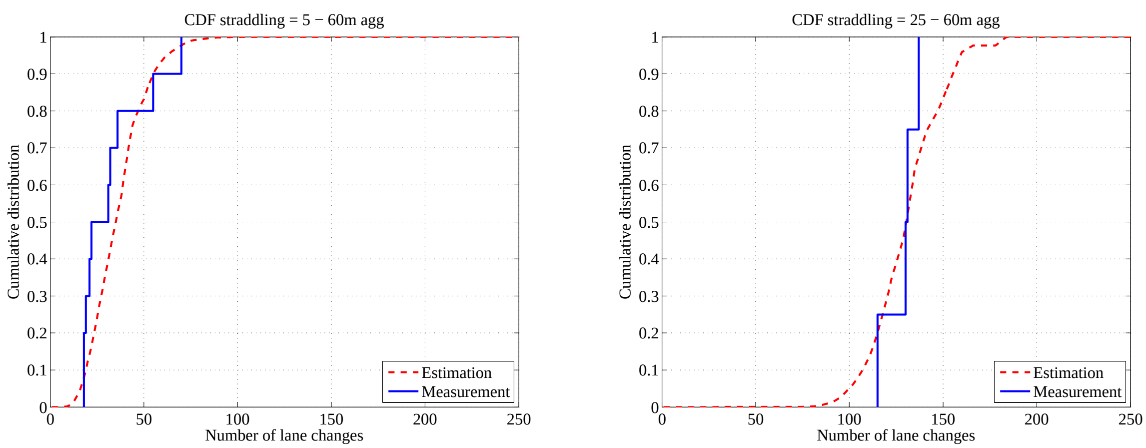
FIGURE 4 The process of finding the distribution of lane changes. The basic idea is to construct a matrix which shows how often a combination of lane changing and straddling is measured.



(a) Measured distribution of the number of straddles

(b) Estimated and measured distribution of the number of lane changes

FIGURE 5 Estimation of the distribution of the total number of lane changes based on the distribution of the total number of straddles



(a) 5 cars straddling

(b) 25 cars straddling

FIGURE 6 The estimated distribution of lane changes and the measured distribution for the validation set in a cumulative curve.

of straddling s, we scale the distribution in a row in such a way that the sum of the straddles in the matrix is in agreement with the observed number of straddles in the data set. Now, the resulting numbers after the scaling are the expected frequencies of lane changing in the set of aggregation intervals with that number of straddling. Figure 4c shows this graphically.

6.4 Results

Sections 6.1 to 6.3 have described how the method works operationally: now we present results to evaluate the accuracy of the method (goal 2 of this section). Figure 6 shows that the estimation of the distribution of lane changes is quite accurate. The relevant number for estimating OD matrices is the mean number of lane changes for a given number of straddles: $\langle N_{LC}(N_S = S_0)\rangle$. We compare this with the mean expected number of lane changes for a given number of straddles, $\left\langle \tilde{N_{LC}}(N_S = S_0)\right\rangle$. The relative error between the

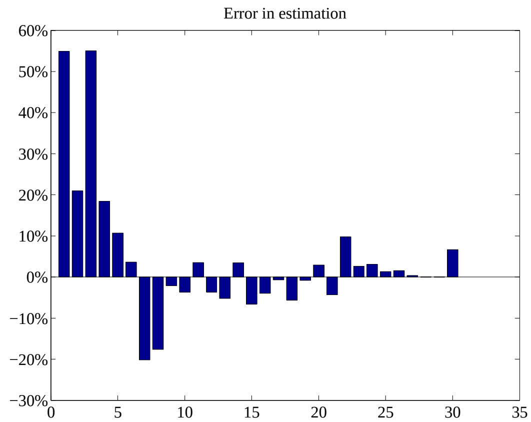


FIGURE 7 The relative error of the estimation of the validation set, calculated by the difference of the mean of the estimation and the mean of the calculation divided by the measured number of lane changes. A negative number means an underestimation of the number of lane changes.

two, $\frac{\langle \tilde{N_{LC}} \rangle - \langle N_{LC} \rangle}{\left\langle \tilde{N_{LC}} \right\rangle}$ is shown in figure 7 for all values of straddling. Even for the cases with very low numbers of straddle observations (see figure 5), the estimation is quite good (never more than 60%). The error is mostly below 10%. This is very good, since the variation in the number of lane changes between the high and the low number of straddles is more than a factor 10.

7 GENERALISABILITY AND INTERPRETATION

Section 5 showed the calibration of p which is the quotient between the number of straddles recorded and the number of lane changes in 200 meters (namely, from site 14 to 15 and from site 15 to 16 (see figure 1) on our site. This quotient depends directly on the length of the considered interval, D_{loop}, and the length over which a vehicle is straddling, D_S:

$$p = \frac{D_S}{D_{\text{loop}}} \tag{6}$$

The length of the straddle will depend on all site specific elements, such as

- the width of a lane;
- the width of a detector;
- the sensitivity setting of a detector;
- the width of the car;
- the length of the lane change manoeuvre, possibly influenced by its speed.

The length of the distance over which is interested in lane changes, will generally depend on the road layout and the detector spacing (one arguably will use road cell units which with a detector in the middle). This section will discuss the influences the speed of vehicles and of the detector spacing. The possible variations of other influencing factors cannot be determined from this site. If one likes to apply the method to another site where one of the other influencing factors are different, a separate calibration will be needed, or an assumption of the effect of the differences.

Section 7.3 shows an example of an OD matrix estimation where the method is useful.

7.1 Speed of the vehicles

A first order approximation is to assume a lane change to have a constant duration in time. Olsen et al. (*16*) shows some properties, including a average lane change time of 6.28 seconds. Because the detectors are not capturing vehicles passing at the edges, there is a shorter time during which a vehicle will activate both detectors. This time is called the straddling time, T_{strad}, which we expect to be constant. We can than rewrite equation 6 into

$$p = \frac{D_S}{D_S} = \frac{vT_{\text{strad}}}{D_{\text{loop}}} = \beta v, \tag{7}$$

in which $\beta := \dfrac{T_{\text{strad}}}{D_{\text{loop}}}$. We will now analyse the effects of the speed on p, thereby find β, and relate that to a straddling time.

Due to re-identification problems, cars driving under 20 m/s were discarded, but the speed dependency from 20 m/s and up to the speed limit (70 miles per hour, 31 m/s) will be analysed to check the assumption of a constant straddling time. To this end, we grouped vehicles in bins with similar speed (bin width=1 m/s), and calculated the straddling quotient p for the lane changing vehicles using the same method as in section 5 (a maximum likelihood estimator). The results, including the 95% confidence intervals are plotted in figure 8a. The figure shows that the assumption of a constant straddling time is quite accurate, at 1.12s. This gives confidence that this assumption can also be extrapolated to other speed intervals.

7.2 Length of considered distance

To find the influence of larger distances between the detectors, we reduce the number of considered detectors (similar to a situation where less detectors are available), and repeat the above analysis for the selection of detectors. Due to the effects of the on-ramp, we prefer to stay as far upstream as possible, and the data from lanes further upstream that 10 will not be used. Furthermore, we prefer to have a detector in the middle of the analysed area, which resulted in the detector combinations as shown in table 2. Note that with detector

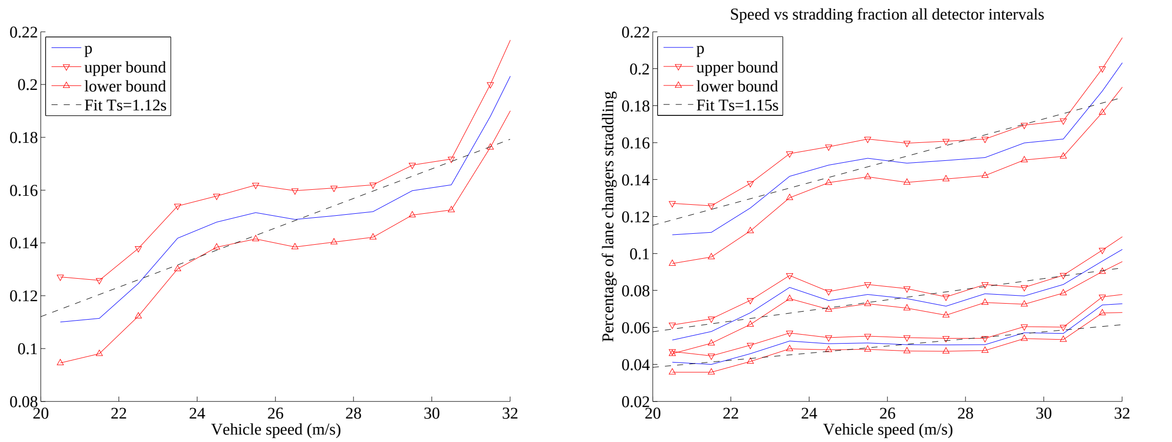
TABLE 2 The combination of detectors used

start detector	end detector	straddle detector	length of stretch	T_{strad} (fit)
14	15	16	200m	1.12s
12	14	16	400m	1.14s
10	13	16	600m	1.19s

intervals of 400 meters, one would be interested in the lane changes over an interval of 400 meters. This corresponds to our test of detector 12, 14 and 16; in this case, however, we needed detectors at the ends of the section for the calibration process, whereas one would normally choose all detectors to be in the middle of a section (and non at the edges). The fitted values for T_{strad} are almost identical. Figure 8 shows the different lines for $p(v)$. It also shows the lines assuming a constant T_{strad} *for all detector intervals*, 1.12s. It is likely that this also holds for longer detector intervals.

7.3 Relevance for OD matrix - a practical example

One of the applications of the straddling method presented here is the case of freeway OD matrix estimation. It is very common to see the type of situation described in figure 9. For this configuration, a dynamic



(a) p as function of speed (b) p for different intervals for which lane changes are recorded

FIGURE 8 The values of p as a function of the speed, and the fit with a constant straddling time per lane change

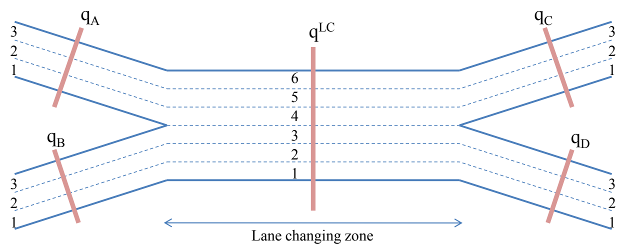


FIGURE 9 Example of a weaving zone with detectors

estimation of the OD matrix is very useful for traffic management purposes, especially if one of the exits leads to a congested area. Usually the flows of each entry and exit are equipped with loop detectors. Using q as symbol for the flow, and Q as symbol for the demand, we have the following determining equations:

$$
\begin{aligned}
q_a &= Q_{AC} + Q_{AD} && (8)\\
q_b &= Q_{BC} + Q_{BD} && (9)\\
q_c &= Q_{AC} + Q_{BC} && (10)\\
q^{LC} &= Q_{AD} + Q_{BC} && (11)
\end{aligned}
$$

Q_d can de directly derived from $Q_d = Q_a + Q_b - Q_c$, and does not add any information.

With the method presented in the paper, we can calculate the lane change flow

$$q^{LC} = \frac{N_{LC}}{\Delta T} \tag{12}$$

In case there is one lane per direction (in stead of 3 shown in the figure 9, the crossing flows ($Q_{AD} + Q_{BC}$) equals the lane changing flows. In case there are more lanes, there might be lane changing for vehicles still heading to the same destination (for instance, a vehicle heading from A to C and changing from lane 2 to lane 1). One could overcome this problem, and find the crossing flows, q^{LC} by focussing on the straddles to the right on the detector on lane 3 and straddles to the left for on the detector on lane 4. These lane change maneuvers indicate a switch of destination. After calibrating this process, the crossing flows will be available by the procedure described in this paper, however, still not showing whether vehicles move from lane 3 to lane 4 or from lane 4 to lane 3.

To find these, equations 10 and 11 are combined:

$$Q_{AC} - Q_{AD} = q_c - q^{LC} \tag{13}$$

Subtracting equation 13 from equation 8, we find:

$$2Q_{AC} = q_c - q^{LC} - q_a \tag{14}$$

Similarly, the other flows can be determined and we have access to the full OD matrix.

8 DISCUSSION AND CONCLUSION

This paper proposed a method to find the number of lane changes in a traffic stream by reading out loop detector data. From the number of simultaneous activations of loops in adjacent lanes an estimate of the number of lane changes can be made.

The best way to use the proposed methodology is to include a site-specific calibration. This could be done by data from spaced detectors, as done in this study, but there are alternatives, for instance counting the number of lane changes (really manual, or using a video) and compare this with the straddling measured in the same time. This gives a initial set to calibrate the relationship between straddling on a detector and the number of lane changes, and it will provide the interval bounds as well. Afterwards, the straddling of the detectors can be used as (calibrated and validated) estimation for the number of lane changes. The more data is available, the more accurate the estimation will be.

The paper calibrated this method for one detector site and checked the validity of the estimation for the same site. The results of the method were very good: the number of estimated lane changes was on average within 10% off the measured value. A sensitivity analysis showed that the method can also be applied for a longer inter-detector distances. It showed that the straddling time remains equal at approximately 1.1 second per lane change. This in turn means that the straddling distance, which determines the straddling rate, depends on the speed. In case the speed in the validation set and the calibration set differs, one has to take this variability into account.

However, in future it needs to be studied to what extent this measure is site-specific. The straddling duration will depend on the width between the detectors, the width of the car, the trajectory a driver chooses and the detector characteristics. This can be country specific. The first two elements can be found in the literature or design guidelines (per country); for the third one, one has to make an additional assumption.

ACKNOWLEDGEMENT

Research sponsored by the Netherlands Organisation for Scientific Research (NWO - project "How traffic jams start") and HTAS SPITS, and the collaboration between the universities in Bristol, Delft, and Lyon is supported by COST Action MULTITUDE. R.E. Wilson is supported by the UK Engineering and Physical Services Research Council grant number EP/E055567/1. Thanks to IDRIS Diamond Consulting Services (*22*) and the Highways Agency for providing the data.

REFERENCES

[1] Choudhury, C., M. E. B. Akiva, T. Toledo, A. Rao, and G. Lee, State dependence in lane changing models. In *Proceedings of the 17th International Symposium on Transportation adn Traffic Theory*, Elsevier, London, 2007.

[2] Ahn, S. and M. Cassidy, Freeway Traffic Oscillations and Vehicle Lane-Change Maneuvers. In *Proceedigs of the International Symposium of Transportation and Traffic Theory* (R. E. Allsop, M. G. H. Bell, and B. G. Heydecker, eds.), Elsevier, Amsterdam, 2007, pp. 691–710.

[3] Cassidy, M. J., C. F. Daganzo, K. Jang, and K. Chung, *Transportation and Traffic Theory 2009: Golden Jubilee*, Springer, chap. Spatiotemporal Effects of Segregating Different Vehicle Classes on Different Lanes, pp. 57–74, 2009.

[4] Lee, H. and B. Coifman, Algorithm to Identify Splashover Errors at Freeway Loop Detectors. In *Proceedings of the 89th Annual Meeting of the Transportation Research Board*, 2010.

[5] Coifman, B., Vehicle Level Evaluation of Loop Detectors and the Remote Traffic Microwave Sensor. *Journal of Transportation Engineering*, Vol. 132, No. 3, 2006, pp. 213–226.

[6] Medina, J. C., A. Hajbabaie, and R. F. Benekohal, Wireless Magnetometers for Traffic Data Collection at Signalized Intersections. In *Proceedings of the Summer Meeting of the Traffic Flow Committee of the Transportation Research Board*, Annecy, France, 2010.

[7] NGSIM, *Home of the Next Generation Simulation cummunity*. www.ngsim.fhwa.dot.gov/, 2008.

[8] Hoogendoorn, S., H. J. Van Zuylen, M. Schreuder, B. G. H. Gorte, and M. Vosselman, Microscopic traffic data collection by remote sensing. *Transportation Research Record: Journal of the Transportation Research Board No.1885*, 2003, pp. 121 – 128.

[9] Angel, A., M. Hickman, and P. Mirchandani, Methods of Analyzing Traffic Imagery Collected From Aerial Platforms. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 4, No. 2, 2003, pp. 99 – 107.

[10] Zhang, L., V. Kovvali, N. Clark, D. Sallman, and V. Alexiadis, *NGSIM-VIDEO User's Manual*. US Department of Transportation, Federal Highway Administration, 2007.

[11] Knoop, V. L., S. P. Hoogendoorn, and H. J. Van Zuylen, Processing Traffic Data collected by Remote Sensing. *Transportation Research Records, Journal of the Transportation Research Board*, Vol. 2129, 2009, pp. 55–63.

[12] Wardrop, J., Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, Part II*, Vol. I, No. 2, 1952, pp. 325–378.

[13] Gipps, G. P., A Model for the Structure of Lane-Changing Descisions. *Tranportation Research Part B: Methodological*, Vol. 20, 1986, pp. 403–414.

[14] Laval, J. A. and C. F. Daganzo, Lane-changing in traffic streams. *Transportation Research Part B: Methodological*, Vol. 40, No. 3, 2006, pp. 251–264, doi: DOI: 10.1016/j.trb.2005.04.003.

[15] McDonald, M. M., Brackstone, and D. Jeffery, Simulation of lane usage characteristics on 3 lane motorways. In *Proceedings of the 27th ISATA Conference*, Germany, 1994, pp. 365–72.

[16] Olsen, E. C. B., S. E. Lee, and M. J. .; Wierwille, Walter W.; Goodman, Analysis of Distribution, Frequency, and Duration of Naturalistic Lane Changes. *Human Factors and Ergonomics Society Annual Meeting Proceedings,*, Vol. 5, 2002, pp. 1789–1793.

[17] Duret, S., A.and Ahn and C. Buisson, Passing rates to measure relaxation and impact of lane-changing in queue. *Computer-Aided Civil and Infrastructure Engineering*, Vol. Vol.4, No. Special issue on Advanced Traffic Computational Modeling and Data Management, 2011.

[18] Knoop, V. L., J. Colé, R. E. Wilson, C. Buisson, and B. Van Arem, The Number of Lane Changes in Traffic. In *Proceedings of the Traffic Flow Theory and Characteristics Committe Summer Meeting of the Transportation Research Board*, 2010.

[19] Coifman, B. and M. Cassidy, Vehicle reidentification and travel time measurement on congested freeways. *Transportation Research Part A*, Vol. 36, 2002, p. 899917.

[20] Wilson, R. E., From Inductance Loops to Vehicle Trajectories. In *Proceedings of Symposium on the Fundamental Diagram: 75 Years*, 2008.

[21] Wilson, R. E. http://www.enm.bris.ac.uk/trafficdata/, http://www.enm.bris.ac.uk/trafficdata/.

[22] IDRIS Consulting. http://www.idris-technology.com/, 2010.