# Quantifying the Number of Lane Changes in Traffic

Victor L. Knoop PhD (corresponding author, v.l.knoop@tudelft.nl)

| | |
|---|---|
| Delft University of Technology | ENTPE-INRETS |
| 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 |

Jonathan Colé
ENTPE-INRETS, Laboratoire d'Ingénierie Circulation Transports (LICIT)
Rue Maurice Audin, 69518 Vaulx-en-Velin, France

R. Eddie Wilson
University of Bristol, Department of Engineering Mathematics
University Walk, Bristol, UK

Christine Buisson PhD
ENTPE-INRETS, Laboratoire d'Ingénierie, Circulation Transports (LICIT)
Rue Maurice Audin, 69518 Vaulx-en-Velin, France

Prof. Bart van Arem
Delft University of Technology
Stevinweg 1, 2628 CN Delft, The Netherlands

July 21, 2010

**ABSTRACT**
Lane changes are an important factor in traffic operations on motorways, but there frequency has not been intensively studied before. Here several measures of lane-changing are proposed – the absolute number of lane changes and the number of lane changes relative to other traffic characteristics, like density, distance travelled and total time spent. The paper then tests the different methods on data collected from the United Kingdom's M42 motorway under a range of free flow traffic conditions. This analysis shows that the number of lane changes per kilometer is rather flat, varying between 1 lane change per vehicle kilometer in very light traffic conditions to 0.5 lane change per vehicle kilometer at densities near capacity.

## 1 INTRODUCTION

Lane-changing influences traffic characteristics to a large extent. For instance, lane changes are necessary for overtaking, and lane-changing vehicles create voids in traffic streams causing under-utilisation of the road. However, little is known about the number of lane changes. One of the main reasons is that there is no good measure to quantify them, and therefore no empirical relationship between the traffic conditions and the number of lane changes has been established so far.

Lane-changing can also be used to to control traffic, directly or indirectly. For example, there is a relationship between the on-ramp flow, possibly controlled by ramp metering, and the number of lane changes. First of all, all vehicles entering the motorway need to change lanes. But moreover, the inflow also changes the number of courtesy gaps that have to be created, and hence increases the number of lane changes. If the number of lane changes influences the capacity, the ramp metering therefore influences the capacity as well. A speed limit and its enforcement can reduce the differences in speed between vehicles and therefore reduce the number of overtakings and hence the number of lane changes.

Thus, in order to describe the effects of the traffic control systems accurately, a measurement for the number of lane changes is required. The main contribution of this paper is that this measurement is defined, and is tested with data. This is a first paper to show the number of lane changes in a statistical way. It is beyond the scope of this paper to give a detailed analysis on of the traffic processes and vehicle-interactions which cause these lane changes.

There are several papers (e.g., (*1*)) which describe, model and predict lane changes. Most papers express the number of overtakings per kilometer per hour as introduced by Wardrop (*2*). From a theoretical point of view, he derives

$$N_O = \frac{k^2 \gamma(u)}{2} \tag{1}$$

in which $N_O$ is the number of overtakings per space and time, $k$ is the density, $\gamma(u)$ is the mean difference in speed.

To check this theory with data, or for further classification of the number of lane changes in data, a good measurement of the number of lane changes, preferably independent of the traffic situation, is useful. This measurement could be used as a parameter in describing the traffic characteristics (e.g., capacity), showing the influence of the number of lane changes. This paper proposes several lane change measurements in section 2. These will be tested against data from the UK, and this empirical analysis is the second contribution. The data description can be found in section 3, and the results in section 4. The paper ends with a discussion and conclusions, in section 5.

## 2 METHODS FOR QUANTIFYING THE LANE-CHANGING RATE

The number of lane changes can best defined in an area in space-time. The symbols used are listed in table 1.

We first discuss Edie's generalised variables (*3*). Let $A$ be the area in space time which is considered, with a surface $|A|$ [m s]. In this area, all vehicles together travel a distance $D(A)$; the total time spent by all vehicles is $T(A)$. Edie now defines the following equations for flow, density and average speed:

$$q(A) = \frac{D}{|A|} \text{(2a)}; \qquad k(A) = \frac{T}{|A|} \text{(2b)}; \qquad u = \frac{D}{T} = \frac{q(A)}{k(A)} \text{(2c)} \tag{2}$$

**TABLE 1 : The symbols used and their meaning**

| Symbol | meaning |
|---|---|
| $N$ | Number of lane changes [] |
| $q_{X1}$ | Flow at location $X1$ $[veh/h]$ |
| $k_\tau$ | Density at time $\tau$ $[veh/km]$ |
| $A$ | The area in space and time considered $[ms]$ |
| $D$ | The total distance travelled $[m]$ |
| $T$ | The total time travelled $[s]$ |
| $u$ | Average speed according to Edie's definitions $[m/s]$ |
| $n_A$ | Number of lane changes per time and space $[1/(ms)]$ |
| $n_D$ | Number of lane changes per time travelled $[1/s]$ |
| $n_T$ | Number of lane changes per distance travelled $[1/m]$ |
| $n_{TT}$ | Number of lane changes per time travelled squared $[1/s^2]$ |
| $n_{DD}$ | Number of lane changes per distance travelled squared $[1/m^2]$ |
| $n_{DT}$ | Number of lane changes per distance travelled per time travelled $[1/(ms)]$ |
| $n_{Dk}$ | Number of lane changes per time travelled per density$[1/s]$ |
| $n_{Tk}$ | Number of lane changes per distance travelled density$[1/m]$ |

Based on the expected relationships between the traffic flow variables and the number of lane changes, several method can be defined.

### 2.1 Method 1: Relative to area

The first method of expressing the number of lane changes is a relative measure per time and space.

$$n_A = \frac{N}{|A|} \tag{3}$$

This method simply expresses the number of lane changes, similar to the number of overtakings in equation 1. However, a relative number might be better for purposes which relate to traffic streams.

### 2.2 Methods 2 and 3: relative to the number of cars

One of the methods is to consider the number of lane changes relative to the flow. Of course, the number of lane changes grows proportionally with the area size in case the traffic conditions do not change:

$$n_D = \frac{N}{q\left(A\right)\left|A\right|} \tag{4}$$

Using relationship 2a, this is equal to $N/D$:

$$n_D = \frac{N}{D\left(A\right)} \tag{5}$$

This means that we have a lane change measurement which is proportional to the distance travelled. If the number of lane changes is proportional to the distance travelled, this is the best measurement, because it will remain constant independent of the traffic conditions. In fact, this measure then shows the behavioral aspect of travellers, rather than the density.

Alternatively, the number of lane changes can be considered relative to the density of vehicles:

$$n_T = \frac{N}{k(A)} \tag{6}$$

Using relationship 2b, this equals

$$n_T = \frac{N}{T(A)} \tag{7}$$

$n_T$ is constant if a driver changes lanes with an average time interval. $n_D$ and $n_T$ are connected by the average speed. Combining equations 5 and 7 with 2c gives:

$$n_T = n_D u \tag{8}$$

### 2.3 Further methods: relative to time and number of other cars

Methods 2 and 3 relate the number of lane changes to the number of cars. However, it can be expected that a driver will not have a constant number of lane changes every minute or kilometer, but that this number also depends on the number of other cars which cause him to change lanes. The length of the extended abstract is too short to give all computations. In short: using a similar reasoning as in methods 2 and 3, dividing by the flow or density is equal to dividing by the distance travelled or total time travelled. Therefore, the relative number of lane changes in equation 5 and 7 are divided once more by either the total distance travelled, the total time spent or the density:

$$n_{DD} = \frac{N}{D(A)} / D(A) = \frac{N}{(D(A))^2} \tag{9}$$

$$n_{TT} = \frac{N}{T(A)} / T(A) = \frac{N}{(T(A))^2} \tag{10}$$

$$n_{DT} = \frac{N}{D(A)} / T(A) = \frac{N}{T(A)} / D(A) = \frac{N}{D(A)\,T(A)} \tag{11}$$

$$n_{Tk} = \frac{N}{T(A)} / k(A) = \frac{N}{T(A)\,k(A)} \tag{12}$$

$$n_{Dk} = \frac{N}{D(A)} / k(A) = \frac{N}{D(A)\,k(A)} \tag{13}$$

## 3 DATA COLLECTION ENVIRONMENT AND AGGREGATION METHOD

The data used for this paper come from the *Active Traffic Management* section of the M42 motorway near Birmingham in the United Kingdom see (*4*). This highway is equipped with dynamic speed control systems and features hard shoulder running in peak times, expanding the width of the carriageway from 3 lanes to 4 lanes in each direction. In addition, the traffic detection systems have unprecedented resolution, with a nominal spacing of 100 metres between inductance loop sites. In their usual operation, each site captures the flow, (time-average) speed and occupancy (as a percentage) in 1-minute aggregate units for each of the lanes.

During 2008, 16 consecutive sites on the North-bound carriageway were enhanced as part of a commercial technical evaluation project so that, amongst other improvements, the full Individual Vehicle Data of all vehicles driving through a one-mile section is recorded, see Wilson (*5*).

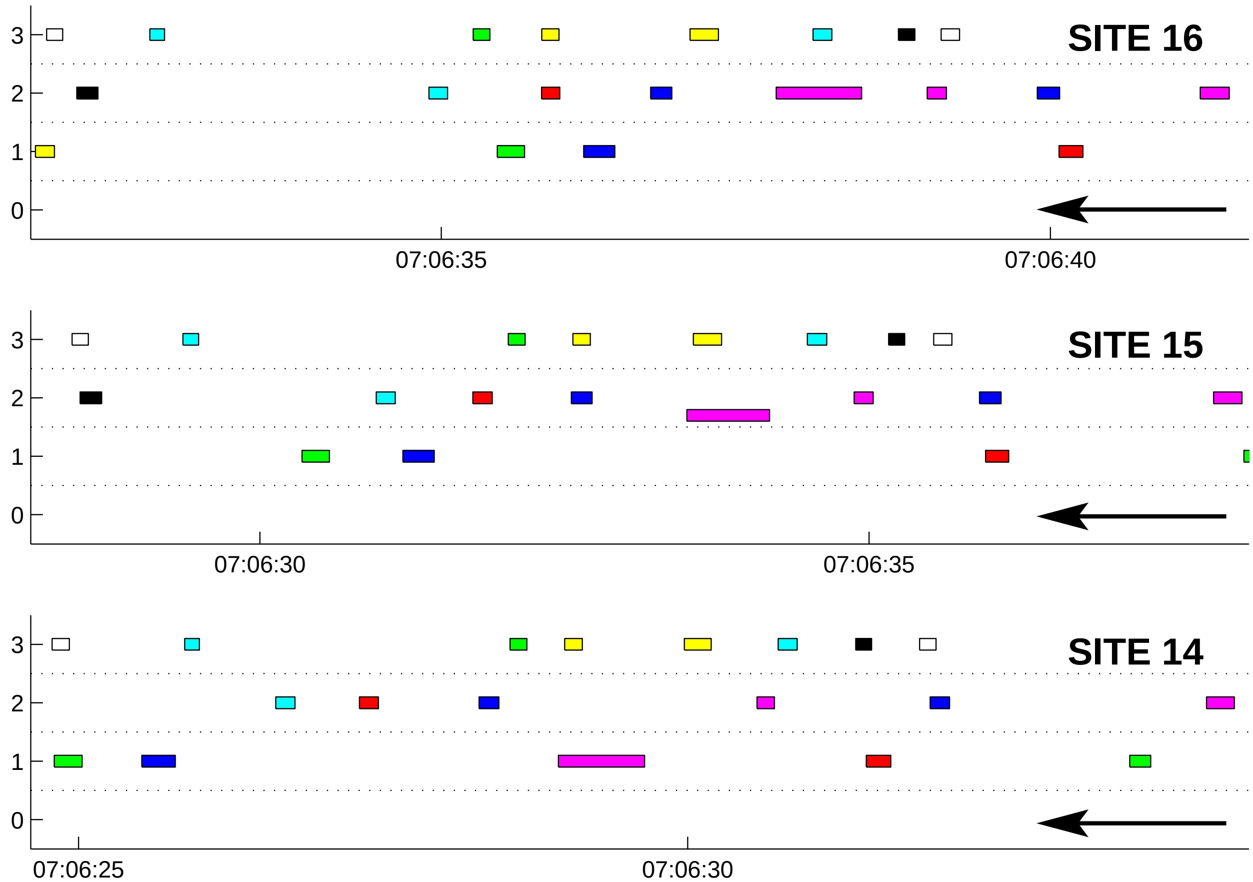



**FIGURE 1 : Visualisation of a short section of Individual Vehicle Data from three inductance loop sites labelled according to figure 2. Since the time axis runs to the right, vehicles apparently drive to the left in each pane of the picture, thus in effect providing a 'helicopter view' of the traffic at each of the sites. The colouring 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. Note that the signal processing system has also detected that the vehicle is not yet fully in-lane at site 15.**

The Individual Vehicle Data include the (1) arrival time, (2) speed, (3) lane number and (4) length of each vehicle as it passes each of the sites. Because the inductance loop sites are separated by only 100m, it is possible to apply pattern matching techniques and *re-identify* the data and track individual vehicles as they drive down the highway, see Wilson (*6*) — an idea due originally to Coifman and Cassidy (*7*). Figure 1 illustrates the re-identification in visual terms.

One drawback of our method is that the error rate of the automatic re-identification algorithms is not known exactly. However, as part of the 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). In contrast, in very congested or strongly dynamic traffic conditions, the re-identification process fails entirely. This study is therefore restricted to traffic conditions which correspond to the uncongested branch of the fundamental diagram. It thus provides complementary information to that which may be

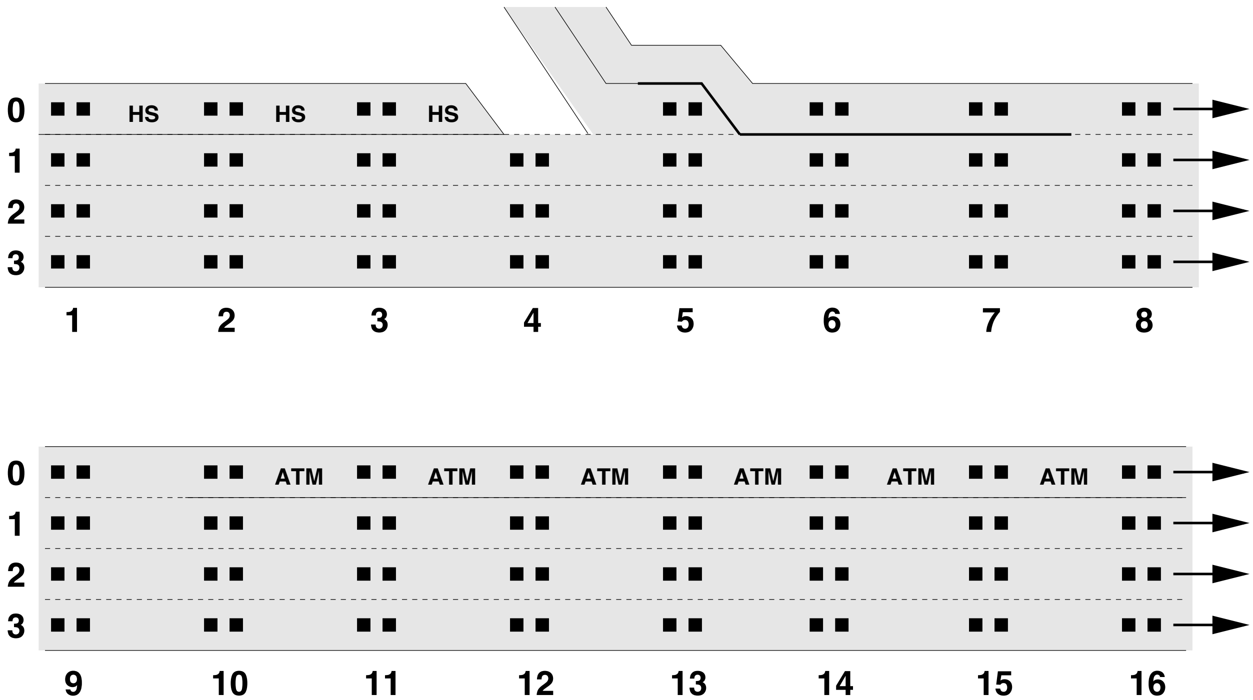



**FIGURE 2 : Geometry of the instrumented section showing the location of double loop detectors. Note that in the UK, slower vehicles drive on the left. HS denotes the hard shoulder (emergency lane) and ATM (Active Traffic Management) denotes hard shoulder which can be activated as an ordinary running lane in peak periods, although such periods are filtered out of the data used here. 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 detectors 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 approximately 500 meters section between detectors 11 and 16.**

obtained from NGSIM video data, which is mostly very congested.

For the purposes of this study, we have constrained ourselves to a particularly clean stretch of data (with relatively few equipment outages etc.) taken from the period 1 October to 30 November 2008 inclusive. Furthermore, we have considered only lane changes recorded between the 6 most downstream sites (a section of approximately 500 meters in length) so as to reduce the impact of the on-ramp which features upstream of this location. Figure 2 depicts the highway geometry. The effect on lane-changing of the speed control and hard shoulder running systems is of considerable interest, but in this initial study we simplify matters as far as possible and filter out all times at which these control systems were in operation. Intervals affected by road works and accidents have also been removed from the data set.

Since our purpose is to understand the lane-changing rate in terms of macroscopic traffic variables, we must aggregate the Individual Vehicle Data. As a first approach, we break each day starting from midnight into 1440 disjoint 1-minute intervals, so that there are 87840 such intervals from the 61 days of data altogether. For each minute, we use the re-identification to count the number of lane changes in the 500 meters section. Moreover, we aggregate the Individual Vehicle Data to calculate the flow and the harmonic-average (i.e., space-average) speed at each of

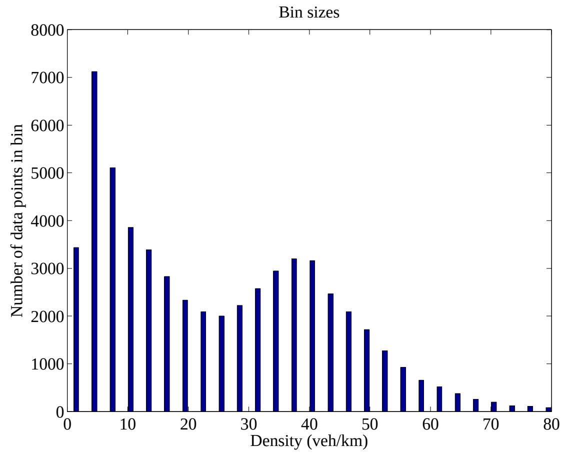



**FIGURE 3 : The number of aggregation intervals counted according to the observed traffic density summed over the 3-lane carriageway. Here aggregation intervals are collected according to a bin-width of 3 veh/km.**

the detectors, and any other macroscopic quantities of interest, such as the speed variance (within lane) or speed differences between lanes. Note that the 1-minute data which the loop system reports in standard operation does not provide all quantities of interest, and in particular the time-average speed is reported so that density may not be computed correctly. Finally, note that the lane changes themselves may also be disaggregated depending on the pair of lanes and pair of detectors between which they occurred — however, these are fine details, beyond the scope of this introductory paper.

Note that in essence we apply the Edie traffic theory (section **??**) for the trivial set-up where each aggregation area $A$ is a 1-minute $\times$ 500 meters rectangle. Care must therefore be taken, particularly in the neighbourhood of travelling features such a stop-and-go waves, to ensure that each aggregation area does indeed correspond to homogeneous traffic conditions. To achieve this, we adopt a selection criterion which rejects minutes for which the spread of space-average speeds across the six detectors is too large. We have found that a 5% threshold for the standard deviation of the space-average speeds as a proportion of their mean successfully removes portions of the data in which there are clear spatiotemporal features. Unfortunately, this method also removes some minutes in low flow conditions which do appear homogeneous, due to the statistical sampling effect. However, the size of our data set is such that a great many aggregation intervals remain even though the filtering procedures are quite crude. The final set contains 57224 aggregation intervals (each corresponding to a 1-minute $\times$ 500 meters rectangle), corresponding to about two thirds of the original data.

In figure 3 we count the number of aggregation periods according to the traffic density recorded in each. Note that due to the size of the data set, we are able to group the data with a very fine bin-width of 3 veh/km, yet maintain statistically large numbers of aggregation intervals in each bin.

## 4 DATA ANALYSIS AND RESULTS

We now analyse the lane-changing data from the aggregation intervals grouped together according to the traffic density. Based on the theoretical discussion in section 2, we develop and compute two measures of lane-changing: the first counts the total number of lane changes in each aggregation interval, whereas the second tries to normalise this measure by the traffic flow.

### 4.1 Number of lane changes, $N$

In order to see best which relative dependency proposed in section 2 would fit best, we express the number of lane changes as function of the density. To this end, we classify all aggregation intervals in bins with similar density (bin-width 3 veh/km).

For aggregation intervals with similar traffic conditions (e.g., with similar density), we may view lane-changing as a stochastic process with a given rate parameter $\lambda$ (that is consequently a function of the given density, so we may write $\lambda = \lambda(\rho)$). If we suppose that lane changes are independent of each other, then the time between consecutive lane changes will be exponentially distributed, and the number of lane changes in any given aggregation interval will be distributed according to a Poisson process with rate $\lambda$. Thus the probability of $n$ events in an aggregation interval of length $T$ (in our case 1 minute) is given by

$$f(n) = \frac{(\lambda T)^n}{n!} \exp(-\lambda T). \tag{14}$$

Since we have many aggregation intervals with common traffic conditions, the counts for lane changes for such intervals may be used to estimate $\lambda$ according to equation 14. Formally, we use the Maximum Likelihood Estimator (MLE) in which $\lambda T$ is simply the expected value of $f(n)$ according to the supplied the data. However, the use of MLE library routines also enables the calculation of error bars by using the Fisher information and the asymptotic normality of the estimator (**?** ).

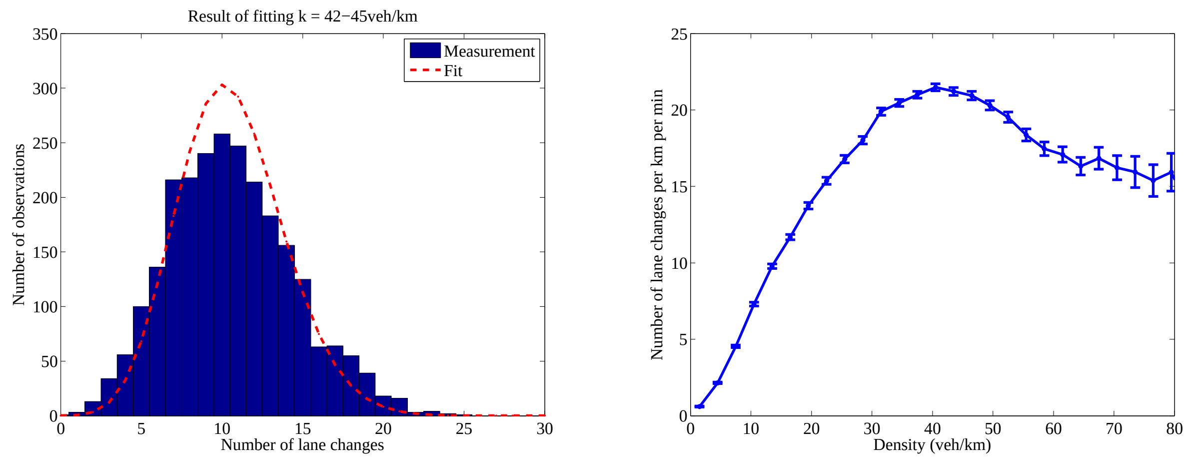
Figure 4a shows the number of lane changes for aggregation intervals with a density of 42 to 45 veh/km. The histogram shows the distribution of data, and the line shows the fit according to the Poisson distribution with the MLE choice of $\lambda$. The fit is acceptable, but not perfect for two key reasons:

- Firstly, it is probable that lane changes are not perfectly independent of each other but rather are weakly correlated. For example, it is likely that an overtaking manouvre is correlated with a subsequent manouvre where the overtaking vehicle 'pulls in'.
- Secondly, we have disaggregated the lane-changing data according to density, but it is likely that there are many other causal factors, so that we should write $\lambda = \lambda(\rho, \ldots)$. In this case, the data of Figure 4a would constitute many different but superimposed Poisson distributions.

In the light of these limitations, the fit of the simplest possible Poisson model is quite remarkable.

Figure 4b shows the result of the MLE fit for $\lambda$ as the density varies, as well as 95% confidence intervals for the fit in each density bin. Note that the confidence decreases for the very high density bins, since there is a relatively low number of such aggregation intervals in the cleaned data set. By dividing the number of lane changes per time by the distance, we obtain the number of lane changes $n_A$ as proposed in equation 3. Note that $n_A$ first increases with density, but then reaches a maximum and starts to decrease. To clarify: this means that once there are more vehicles on the road, there are *less* lane changes, even when counted in absolute number.

Reflecting this in the light of the theoretical considerations in section 2, we have to remark that the number of lane changes per km per min is not constant. In case the speeds are independent




(a) The number of lane changes for the density bin k=42-45 veh/km and the result of the MLE fit of a Poisson distribution thereof. Although there are some limiting assumptions, the fit is reasonably good.

(b) Number of lane changes per km per min as function of density, as well as the error bounds from the fit. The number of lane changes first increases more than proportional to the density and then gradually the derivative decreases and finally the number of lane changes will decrease with increasing density.

**FIGURE 4 : The results of the fitting procedure**

this rate would increase quadratically, as shown for instance in in equation 1. For very small densities the increase in quadratically, but then it flattens, meaning the vehicles influence each other in their speeds.

## 4.2 Relative number of lane changes

Since the absolute number is not constant and section 2 shows that other, relative measures might give the number of lane changes as more independent variable, this section analyses the relative number of lane changes. The procedure for fitting these is different compared to the number of lane changes in an area. The Poisson distribution shows the number of events in a certain time interval, and thus can be used to fit directly the number of lane changes. When it comes to a lane-changing rate, equations 5 to 11, one needs to fit a lane-changing *rate* and the number of observations can not be fitted directly.

To find the rate, we convert first of all the observations to a rate, using one of the equations 5, 7 or 10 - 11. Let us discuss $n_D$ as example. Equation 5 gives the lane-changing rate: the number of lane changes per distance. The inverse, $1/n_D$ gives the average distance between two lane changes, given the assumption that the lane changes are independent. If the number of lane changes is zero, the average distance between lane changes is infinity. In practice, this means that the number is higher than the the total distance covered in the measurement interval, $D$. We can draw the cumulative distribution of the average distance between lane changes. However, for the time intervals in which the number of lane changes is zero ($N$ = 0), the average is unknown, but larger than $D$ (note this value varies for different intervals). Figure 5 shows this cumulative distribution function which raises one step at any point between $D$ and infinity. As soon as the maximum measurable rate is reached, two cumulative distributions are drawn: one which raises in infinity, the other one which immediately takes the step. Figure 5a shows the density bin 9-12

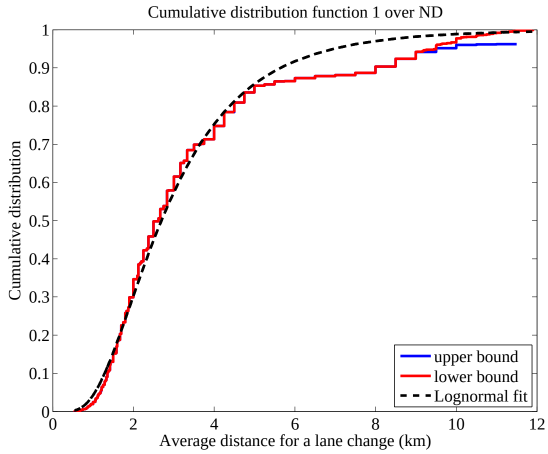



**FIGURE 5 : Two cumulative distribution functions for the measured average distance between two lane changes and the fit**

veh/km, which is a density bin where there are relatively high number of observations with no lane change at all.

We calibrated this rate for all derived lane changing rates, $n_A$, $n_D$, $n_T$, $n_{TT}$, $n_{DD}$, $n_{DT}$, $n_{Dk}$, and $n_{Tk}$. Figure 6 shows how the number of lane changes per vehicle kilometer ($n_D$) and the number of lane changes per vehicle minute ($n_T$) changes as function over time in figure a and b respectively. First, the number of lane changes increases sharply with the density, but the number starts to decrease at around 25 veh/km.

The number of lane changes is relative to the distance travelled is quite constant. Figure 6a shows that the lane change rate, although decreasing, is about once every 1 km. If $n_D$ is constant or decreasing as function of density, the further derivatives – methods relative to time and number of other cars as proposed in section – decrease even more. Therefore, further divisions by time spent, distance covered or density are considered not useful as measures to express the number of lane changes.

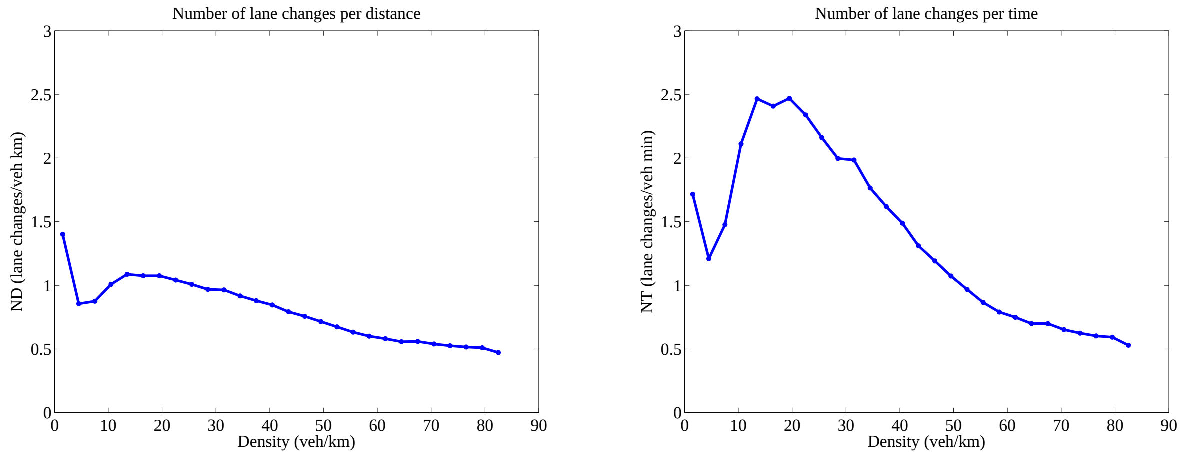
## 5 CONCLUSION AND DISCUSSION

This paper introduces a methodology to quantify lane changes, and applied this on a real-life data set. First, several possibilities and relative dependencies were tested. Testing them with motorway data, It was found that the number of lane changes per veh km travelled was quite flat for different traffic conditions and it is therefore advised to use this measure as quantification for the number of lane changes.

Using motorway data of the number of lane changes, the relationship between the number of lane changes and density was shown, thus showing a “fundamental diagram of lane-changing”. The proposed lane change measurement can in future be used to quantify the effect of lane changes on the capacity, using for instance the impact of lane changes as described by Jin (*9*).

We are aware that this paper is just a first step in quantifying lane changes. First of all, it is just a statistical description of the number of lane changes, whereas a more detailed analysis of the underlying traffic processes lacks. This is work in progress.

Furthermore, the analyses assume that the lane change rate is equal for all vehicles, and that the (rectangular) areas in space-time are homogeneous. Both assumptions are incorrect. It




(a) Lane changes per veh km travelled (b) Lane changes per veh min travelled

**FIGURE 6 : Relative number of lane changes**

will be studied what will be the influence of these hypotheses and what is the consequence if we relax these.

## References

[1] Laval, J. A. and L. Leclercq, Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. *Transportation Research Part B: Methodological*, Vol. 42, No. 6, 2008, pp. 511–522.

[2] 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.

[3] Edie, L., Discussion of traffic stream measurements and definitions. In *Proceedings of the Second International Symposium on the Theory of Traffic Flow*, OECD, Paris, France, 1965.

[4] British Highway Agency. `http://www.highways.gov.uk/knowledge/1334.aspx`, 2010.

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

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

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

[8] Hogg, R. V., A. Craig, and J. W. McKean, *Introduction to Mathematical Statistics*, chap. 6, ????

**Acknowledgement:** Research sponsored by the Netherlands Organisation for Scientific Research (NWO) and HTAS SPITS. Thanks to IDRIS Diamond Consulting Services (*10*) and the Highways Agency for providing the data.

[9] Jin, W. L., Macroscopic Characteristics of Lane-Changing Traffic. In *89th Annual Meeting of the Transportation Research Board*, 2010.

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