

Correlations in Spatiotemporal Headway Dynamics of Road Traffic Using Extremely Accurate Microscopic Empirical Data

Maximilian Kumm · Michael Schreckenberg

University of Duisburg-Essen, Physics of Transport and Traffic, Duisburg, Germany

E-mail: maximilian.kumm@uni-due.de, E-mail: michael.schreckenberg@uni-due.de

Received: 9 March 2020 / Last revision received: 25 March 2020 / Accepted: 26 March 2020

DOI: [10.17815/CD.2020.81](https://doi.org/10.17815/CD.2020.81)

Abstract As we recently showed [1] by using empirical data there is a certain behavior referring to the development of the headway between two consecutive human driven vehicles. Following on from this, we investigate correlations of the change in temporal headway over two subsequent road segments as the main goal of the present work and found a strongly correlated behaviour for increasing temporal headways. In this way a strong improvement for short-term prediction algorithms of conventional road users should be achieved. A stationary infrared-based sensor system was developed for this purpose, which has been mounted at reflector posts next to an urban street over a distance of about 50m. Due to its good accuracy, we are able to resolve vehicle following times down to 25 milliseconds and to determine speeds more precisely. In 45 hours of measurement the system detected over 20,000 passing vehicles.

Keywords Sensor system · road traffic · urban traffic · headway development · empirical headway distribution · empirical speed distribution · correlations in headway dynamics · connected and automated vehicle (CAV) · single lane driving · vehicle merging

1. Introduction and Related Work

Safety, energy efficiency, and comfort are three of the most important issues related to modern and upcoming road traffic [2]. In the course of ever increasing automation, the energy consumption per vehicle will further continue to decrease [3]. A system that will consist only of connected and automated vehicles (CAVs) is comparatively easy to handle with regard to these points, since individual driving behaviour can be electronically

coordinated either by an external source or by the vehicles themselves. Hence, different situations become predictable leading to an efficient, crash-free use of the road network. In this way, impending traffic disruptions can be counteracted in advance through targeted, collective action using *vehicle-to-vehicle* (V2V) [4] or *vehicle-to-infrastructure* (V2I) [5] communication. Even an increase in traffic flow and road capacity is expected due to a higher influence of automated vehicles as simulations show [6–8].

Before the phase of fully automated traffic occurs in the distant future, however, there will be the much more challenging phase of partially automated traffic between automated and human driven vehicles. These new circumstances call for novel approaches with regards to *intelligent transportation systems* (ITS) aiming at a decrease in individual travel time finally [9]. Especially urban situations, which require frequent interaction between machines and human beings, will be difficult to handle. Thus, it becomes necessary to get into the behaviour of conventional road users and make it as predictable as possible for upcoming technologies. In order to capture different traffic scenarios with human driven vehicles involved, the on-board sensors of an automated vehicle can be supplemented by V2I sensor systems, especially if the direct view is obstructed through obstacles like buildings or other vehicles. The collected data yield short-term predictions and thus enable the CAV to operate efficiently, while disturbing the ongoing traffic as little as possible [1]. It is particularly important to minimize the risk of any dangerous situations including crashes. If a situation should ever develop contrary to the calculated scenario, the driving strategy of the automated vehicle must of course change in an appropriate way. Even continual changes may become necessary to meet the required conditions in terms of safety, energy efficiency and comfort.

The actual prediction can be realized by using different motion models [10], taking into account real-time information of the locations and speeds of the relevant road users. With the help of a dedicated algorithm, the CAV can plan and finally follow the most promising trajectory [2, 11]. However, it should be noted that such a driving manoeuvre and the associated success depends crucially on the quality of the used model as this is seen as a key feature in terms of robotic architectures [12].

In this paper we provide microscopic empirical data which is the basis for a realistic prediction model. The use of the Pearson correlation coefficient in the context of vehicle-vehicle interaction over subsequent road sections leads to a quantitative estimation of human behaviour in terms of driving motor vehicles. The recorded traffic data should help to select the right model approach and to calibrate it correctly. To the best of the authors' knowledge, there is no comparable work that deals with microscopic correlations regarding vehicle following behaviour and temporal headways. However, there are publications that deal with observables of static origin only like the distribution of headways (e.g. [1, 13]). In the mid-1970s, Cowan was the first to come up with an empirically confirmed result [14]. His famous so-called M3-distribution is still well known in transportation science today. On the other hand there are approaches analyzing spatiotemporal patterns in traffic and their correlation within an entire road network [15, 16]. For these purposes microscopic models are usually used, mostly based on classic models like *cellular automata* (CA) [17] or *car-following* models [18] in order to obtain macroscopic observables finally. However, this does not mean that correct microscopic results

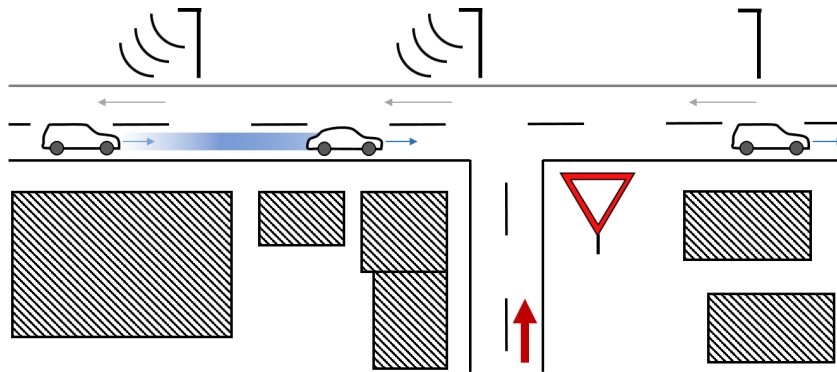


Figure 1 Exemplary application of the results in this work. T-intersection with common road traffic on its major street. A connected and automated vehicle (red arrow) is approaching from below and tries to merge into the ongoing traffic (blue marked headway). The question is how the marked headway has developed, when reaching the intersection point. Due to the view obstructed by the buildings (hatched boxes), an infrastructure sensor system is required to provide the headway data from the major street to the CAV. [1]

also can be derived by this procedure as no classical model is able to reproduce real world vehicle-following behaviour in an appropriate way [19,20].

2. Application

The actual application of this work is illustrated in Fig. 1. An connected and automated vehicle (red arrow) is approaching a T-intersection without traffic lights. As the CAV is driving on the minor road, it must give way to conventional traffic on the major road before turning right. Due to the infrastructure sensors installed along the major road, the CAV is informed about the microscopic traffic situation already before it reaches the intersection area. A large set of previously collected traffic data from there, allows a statistical analysis regarding certain patterns in human driving behaviour. Based on this historical data, it is now possible to make a statistical short-term prediction of the vehicles on the major road considering the current vehicle configuration. The CAV can therefore adjust its driving strategy in advance in such a way that it can merge into a suitable gap (*headway*) among the ongoing traffic, preferably without stopping at all. Nevertheless, it has to take into account possible changes of the target gap during its approach to the intersection and has to adapt its speed according to arisen circumstances continuously. Even stopping may be necessary, if the merging manoeuvre cannot be completed safely, i.e. crashes cannot be almost 100% excluded. This strategy is expected to reduce stopping manoeuvres compared to conventional road user, but indeed is not able to totally avoid them. All in all, this technology yields great advantages not only in terms of safety, but as well in terms of energy efficiency and passenger comfort. The present paper firstly deals with the collection of necessary traffic data and secondly provides a mathematical statistical analysis using correlations with regard to the spatio-temporal prediction of headways between consecutive human driven vehicles.

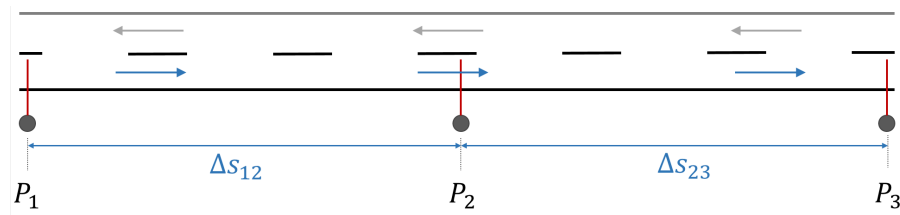


Figure 2 Schematic function of the measuring system. Three measuring units including the lidar sensors are attached to reflector posts at positions P_1 , P_2 , and P_3 next to an urban street. The red vertical lines represent the lidar beams caused by each measuring unit to detect passing vehicles (only vehicles from the bottom lane are detected). The distance between any two reflector posts is called Δs_{ij} . For the distance between adjacent reflector posts applies: $\Delta s_{12} \approx \Delta s_{23} \approx 23.5$ m.

3. Measurement

With the premise of obtaining extremely accurate microscopic data, the measuring system from our previous work [1] has been improved by replacing the infrared-based sensors by even more reliable lidar sensors. In this new hardware configuration the current sensor value can be read by the microcontroller at least every 25 ms and is then saved together with a time stamp to the internal flash memory. Beside the higher sampling rate, these new lidar sensors also deliver a much better detection rate, and a lower susceptibility to external interferences at the same time. Another difference compared to the previous work is, that only three of the four available sensors are needed here with regard to the current application. Three sensors mean two resulting sectors between which the addressed correlation is investigated to keep things as simple as possible. In [1] we were mainly interested in the change of the temporal headway $\Delta\tau$ over *different* distances, where the consideration of several sectors is more expedient.

The measurement has taken place on an urban single-lane road in Duisburg, Germany with a speed limit of 50 km/h, where there is no influence on passing vehicle e.g. through traffic lights, bus stops, speed bumps, or speed cameras. As a consequence, all road users are able to choose their personal driving speed freely taking into account the speed limit and the current traffic situation. During the four days of measurement in September and October 2019, there have been similar weather conditions with dry roads and a mix of cloudy and sunny periods. A total of exactly $N = 20,018$ vehicles were detected within 45 hours of measurement in order to enable a statistical data processing afterwards. The traffic on the concerning road is dominated by cars. Only a small proportion of the passing vehicles are made up of trucks and buses. Please refer to [1], if there is interest in the traffic volume according to individual vehicle classes, since another section of the same road was studied there.

3.1. Measuring System

Three measuring units including one lidar sensor each make up the measuring system. As Fig. 2 and 3 show, they have been attached to reflector posts at positions P_1 , P_2 , and P_3 . The distance between any two reflector posts i and j at positions P_i and P_j is called Δs_{ij} .

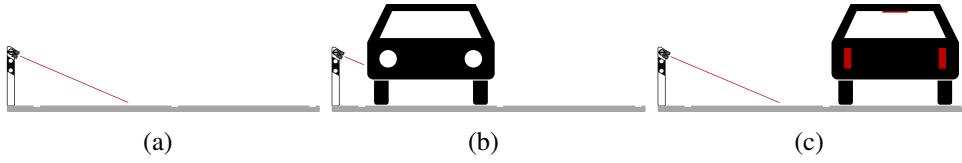


Figure 3 Detection of passing vehicles. The three lidar sensors have been attached to three neighboring reflector posts next to the street (compare to Fig. 2). The red line visualizes the corresponding infrared laser beam. To avoid the detection of vehicles from the opposite lane, the lidar sensors are aligned down to the lane of interest. Subfig. (a): no vehicle at all \rightarrow IDLE (no detection); Subfig. (b): a vehicle on the lane of interest \rightarrow DETECTION; Subfig. (c): a vehicle on the opposite lane \rightarrow IDLE (no detection).

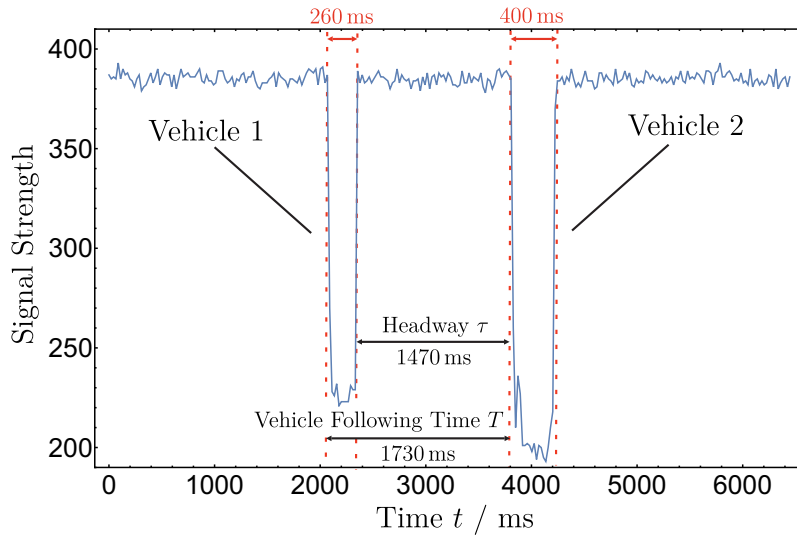


Figure 4 Typical signal shape of a stationary single measuring unit caused by two passing vehicles. The difference between the temporal headway $\tau = 1470\text{ms}$ and the vehicle following time $T = 1730\text{ms}$ is illustrated. The corresponding signal durations d (260ms and 400ms) yielding the respective vehicle length, when considering the current speed. It is used to calculate the temporal headway τ from the vehicle following time T during the evaluation (see Eq. 4).

In the present study the distance between adjacent reflector posts is about the same. It applies $\Delta s_{12} \approx \Delta s_{23} \approx 23.5\text{m}$.

To avoid the detection of passing vehicles from the opposite lane, the sensors are attached at a certain height and aligned down to the lane of interest (see Fig. 3). All three measuring units are identical and are running independently of each other during the whole measurement. Since we are investigating a single-lane road, it is not necessary to consider the exact lateral position of passing vehicles, only the longitudinal position is relevant here. However, during the evaluation, the recorded data of the single units is finally merged. For this purpose it is necessary to synchronize the system times of all units involved before the actual measurement takes place. This has been achieved by using three extremely accurate clock modules, which must be set to the same system time first. After the measurement has been successfully completed, the clock modules have to be checked electronically for the exact amount of their temporal deviation. As this deviation shows

i, j	$\Delta\bar{t}_{ij}^{(c)}$ [ms]	$\sigma_{ij}^{(c)}$ [ms]	Δs_{ij} [m]
1,2	1614	20	23.31 ± 0.28
2,3	1639	11	23.67 ± 0.16
Mean	1627	16	23.49 ± 0.22

Table 1 Calibration process driving at a defined vehicle speed $v^{(c)} = 52$ km/h. The average transit time between adjacent reflector posts i and j at positions P_i and P_j (see Fig. 2) is called $\Delta\bar{t}_{ij}^{(c)}$ and the corresponding standard deviation $\sigma_{ij}^{(c)}$, respectively. According to Eq. 1 the specified sector length Δs_{ij} is finally calculated.

a linear behaviour in dependence of time, it is thus possible to eliminate the resulting systematic error in retrospective. To ensure that the flowing traffic behaves as usual, the measuring units are not recognizable for passing drivers, because the lidar sensors are mounted on the back side of the reflector posts. Every time a vehicle passes one of the relevant positions P_i on the lane of interest, a characteristic signal drop occurs. This is due to the fact that the lidar sensors in this setup are actually used for distance measurements. The internal electronics determine the time delay between the transmission of a infrared laser signal and its reception after reflecting off of a target. This information is calculated together with the known speed of light and is finally translated into a distance. A low voltage output corresponds to a low distance (a passing car, compare to Fig. 3 (a)) and a high voltage output to a high distance (the road, compare to Fig. 3 (b)). Please refer to Fig. 4 for a typical example of the signal shape.

3.2. Calibration Method

After the measuring system has been successfully installed and has taken up its function, it now has to be calibrated. This step is necessary because of the following two reasons. First, to determine the exact distance of the reflector posts yielding the exact speed of passing vehicles and second, to get an idea about the error in measurement. Achieving this goal has been done by passing the concerning road segment multiple times with a car driving at a GPS confirmed *calibration speed* of $v^{(c)} = 52$ km/h. In order to make sure that $v^{(c)}$ is both constant and reproducible, the car's cruise control has been activated during the calibration process. As a result the average *transit times* $\Delta\bar{t}_{12}^{(c)}$ and $\Delta\bar{t}_{23}^{(c)}$ for the respective road sectors Δs_{12} and Δs_{23} are determined. The corresponding standard deviations $\sigma_{12}^{(c)}$ and $\sigma_{23}^{(c)}$ serve as a quantity for the error in measurement. Additionally, the small resulting values of the standard deviation confirm that the constant speed using the cruise control of the calibration vehicle can be seen as reproducible. A total of ten calibration runs were carried out per measurement day. A summary of all corresponding parameters can be seen in Tab. 1.

3.3. Data Processing

From the averaged *transit times* $\Delta\bar{t}_{ij}^{(c)}$ in the previous Sec. 3.2, the sector lengths $\Delta s_{ij} = \{\Delta s_{12}, \Delta s_{23}\}$ are determined (see Tab. 1):

$$\Delta s_{ij} = v^{(c)} \cdot \Delta\bar{t}_{ij}^{(c)}. \quad (1)$$

To obtain finally the average speed v_{ij}^n (short: v) of a particular vehicle n within the sector Δs_{ij} using the corresponding *transit time* Δt_{ij}^n , we write:

$$v_{ij}^n = v^{(c)} \cdot \frac{\Delta\bar{t}_{ij}^{(c)}}{\Delta t_{ij}^n} = \frac{\Delta s_{ij}}{\Delta t_{ij}^n}. \quad (2)$$

The *vehicle following time* T_i^n (short: T) for two consecutive vehicles n and $n+1$ is calculated as follows:

$$T_i^n = t_i^{n+1} - t_i^n, \quad (3)$$

where t_i^n describes the *signal start time* of vehicle n at position P_i (see Fig. 2). The *temporal headway* τ_i^n (short: τ) is derived from this by subtracting the corresponding signal duration d_i^n (short: d) from the preceding vehicle n :

$$\tau_i^n = T_i^n - d_i^n = t_i^{n+1} - t_i^n - d_i^n. \quad (4)$$

Consequently Eq. 4 describes the temporal headway *in front of* vehicle n . In Fig. 4 the difference between T and τ is illustrated with an exemplary signal sequence. Finally, the *change in temporal headway* $\Delta\tau_{ij}^n$ (short: $\Delta\tau$) between two consecutive vehicles n and $n+1$ over the defined distance Δs_{ij} is regarded:

$$\Delta\tau_{ij}^n = \tau_j^n - \tau_i^n. \quad (5)$$

Hence, a positive value of $\Delta\tau$ corresponds with an increase, whereas a negative one corresponds with a decrease in temporal headway τ .

Last but not least, the correlation behaviour between the change in temporal headway $\Delta\tau_{ij}^n$ within the adjacent road sectors Δs_{12} and Δs_{23} is presented. It has been empirically shown [1] that $\Delta\tau_{ij}^n$ is sufficiently normally distributed to justify the following approach. The *Pearson correlation coefficient* c (first found by Francis Galton in 1888 [21]) is a statistical variable measuring the *linear* correlation between two series $\Delta\tau_{12}$ and $\Delta\tau_{23}$ of length N . For a detailed mathematical view please refer to [22]. From this the following N value pairs

$$\{\Delta\tau_{12}^*, \Delta\tau_{23}^*\} = \dots, \{\Delta\tau_{12}^{n*}, \Delta\tau_{23}^{n*}\}, \{\Delta\tau_{12}^{n+1*}, \Delta\tau_{23}^{n+1*}\}, \dots, \{\Delta\tau_{12}^{N*}, \Delta\tau_{23}^{N*}\} \quad (6)$$

are formed. However, in order to apply the Pearson correlation coefficient the right way and to obtain values between $c = -1$ (perfectly anti correlated behaviour) and $c = 1$

(perfectly correlated behaviour), it has been necessary to *normalize* the data series $\Delta\tau_{12}$ and $\Delta\tau_{23}$ with its elements $\Delta\tau_{12}^n$ and $\Delta\tau_{23}^n$, respectively, first:

$$\Delta\tau_{ij}^{n*} = \frac{\Delta\tau_{ij}^n - \langle\Delta\tau_{ij}\rangle}{\sigma(\Delta\tau_{ij})}, \quad (7)$$

where $\langle\Delta\tau_{ij}\rangle$ describes the mean, and $\sigma(\Delta\tau_{ij})$ the standard deviation of the series $\Delta\tau_{ij}$. To calculate the Pearson correlation coefficient between the normalized series $\Delta\tau_{12}^*$ and $\Delta\tau_{23}^*$ (each of length N) finally, the following formula is applied:

$$c = c(\Delta\tau_{12}^*, \Delta\tau_{23}^*) = \langle\Delta\tau_{12}^* \cdot \Delta\tau_{23}^*\rangle = \frac{1}{N} \sum_{n=1}^N \Delta\tau_{12}^{n*} \cdot \Delta\tau_{23}^{n*}. \quad (8)$$

Please note that $N + 1$ detected vehicles lead to N temporal headways τ to be included in the calculation as the last vehicle $n = N + 1$ has no vehicle ahead.

In order to obtain a more detailed view on the underlying correlation behaviour, the $\Delta\tau$ -data is splitted up into subsets according to the related headway $\tau = \tau_1$ at position P_1 . For this purpose the following τ -intervals

$$\left(\tau_-(k), \tau_+(k)\right] = \left(\tau_0 + k \cdot \tau_{\text{range}}, \tau_0 + (k + 1) \cdot \tau_{\text{range}}\right] \quad (9)$$

are used, where $k = 0, 1, 2, \dots$ indicates the respective interval number, $\tau_-(k)$ the lower and $\tau_+(k)$ the upper boundary. In the present work all resulting intervals are of the same range $\tau_{\text{range}} = 500$ ms. The starting value of the first interval ($k = 0$) is called $\tau_0 = 500$ ms. Hence, this leads to the following τ -intervals to be considered:

$$(500, 1000] \text{ ms}, (1000, 1500] \text{ ms}, (1500, 2000] \text{ ms}, \dots \quad (10)$$

Each of these contains N_k elements. It is obvious that for the used dataset applies $N > N_k$. If we now want to calculate the Pearson coefficient c of the concerning headway values $\tau = \tau_1$ within a particular interval k , Eq. 8 changes to

$$c = \frac{1}{N_k} \sum_{n=1}^{N_k} \left(\Delta\tau_{12}^{n*} \cdot \Delta\tau_{23}^{n*} \mid \tau_1 \in (\tau_-(k), \tau_+(k)) \right). \quad (11)$$

In other words, the data is first sorted according to different temporal headways $\tau = \tau_1$ detected at position P_1 . All those vehicles whose temporal headways τ_1 fall within the relevant interval from Eq. 10 are used for further processing. The change in their temporal headway $\Delta\tau_{12}$ in sector Δs_{12} is then compared with that ($\Delta\tau_{23}$) in sector Δs_{23} (see Fig. 2). From that the correlation coefficient according to Pearson is finally calculated as shown in Eq. 11. All in all, this method leads to a much more accurate prediction of the most probable headway behaviour, finally.

Another method to show the relationship of the two series $\Delta\tau_{12}$ and $\Delta\tau_{23}$ with length N (or with subsets of these according to Eq. 9) is to plot them in a common *scatter diagram*. In this way, the (linear) correlation between two variables can be visualized. This is achieved

by entering all N (N_k) pairs of elements $\{\Delta\tau_{12}^n, \Delta\tau_{23}^n\}$ ($\{\Delta\tau_{12}^n, \Delta\tau_{23}^n\} \in (\tau_-(k), \tau_+(k))$) as points in a common diagram using Cartesian coordinates. It is important to see that despite the same symmetric shape of the *marginal distributions*, the scatter diagram can vary in appearance depending just on the correlation coefficient c . For values around $c = 0$ (no or very less linear correlation) the set of dots in the diagram looks circular, whereas for increasing positive (or negative) values of c the set of dots becomes more and more ellipsoidal. The difference is that for correlated data ($c > 0$), the slope of the main axis of this ellipse becomes positive and for anti-correlated data ($c < 0$) negative.

4. Results

The following part is splitted up into two subsections. Beside the empirically determined speed and temporal headway distribution in Sec. 4.1, the main goal of this work is presented in Sec. 4.2. The latter includes the visualization of the correlated $\Delta\tau$ -data using scatter diagrams. Additionally, a sample of the underlying symmetrical marginal distributions is illustrated. Last but not least, the development of the respective Pearson correlation coefficient for different temporal headways τ is shown and quantified by using a non-linear fit function.

4.1. Speed and Headway Distribution

In order to give the reader further information about the properties of the recorded vehicle data, the empirical speed (Fig. 5(a)) and the temporal headway distribution (Fig. 5(b)) are shown in advance. Due to the fact that there are three measuring units installed along the street, *two* average speeds v per vehicle are available, one for each road sector Δs_{12} and Δs_{23} . For more details regarding the averaging process, please refer to Sec. 3.3. As the two single speed statistics are not exhibiting significant statistically deviation among themselves, they are displayed in a common histogram. The same applies for the temporal headway τ . Because this observable is measured at three positions P_1 , P_2 , and P_3 , there are even *three* values available for each passing vehicle contributing to the respective histogram.

4.2. Correlations

This chapter is devoted to the empirically measured correlation between the change in temporal headway $\Delta\tau_{12}$ and $\Delta\tau_{23}$ within the two adjacent road sectors Δs_{12} and Δs_{23} (see Fig. 2; for the formal definition of $\Delta\tau$ please refer to Eq. 5). As explained in Sec. 3.3, the vehicle data are divided into the intervals from Eq. 10 according to different temporal headways τ (measured at P_1), first. Subsequently, the associated changes in temporal headway $\Delta\tau_{12}$ and $\Delta\tau_{23}$ are plotted over each other in a scatter diagram (Fig. 6). This is done for the first nine τ -intervals. Based on this visualization, an increasing trend of the correlation between the two variables $\Delta\tau_{12}$ and $\Delta\tau_{23}$ can be seen with increasing τ . This is

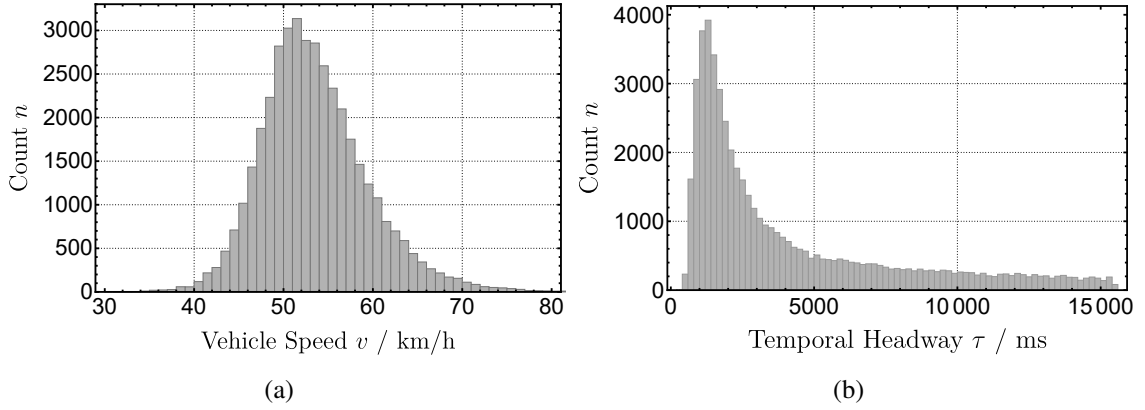


Figure 5 Histograms of the empirically determined speed distribution (Subfig. (a), bin width 1 km/h) and temporal headway distribution (Subfig. (b), bin width 200 ms). The histograms include the data of all four measurement days from the urban road specified in Sec. 3. Please note that due to visualization purposes a cutoff in Fig. 5(b) has taken place for headways $\tau > 15000$ ms.

	Interval of τ	μ [ms]	σ [ms]	γ	κ	$\bar{\mu}$	$\bar{\sigma}$	\bar{v}
$\Delta\tau_{12}$	$500 < \tau \leq 1000$	10	67	0.82	8.67	8.61	48.17	4.15
	$1000 < \tau \leq 1500$	4	84	0.07	8.66	3.84	61.62	4.49
	$1500 < \tau \leq 2000$	-7	108	-0.13	7.79	-5.64	87.34	6.02
$\Delta\tau_{23}$	$500 < \tau \leq 1000$	9	62	0.46	5.50	6.48	48.68	4.80
	$1000 < \tau \leq 1500$	1	81	-0.28	5.41	1.82	62.57	4.79
	$1500 < \tau \leq 2000$	0	102	-0.04	5.16	0.93	82.61	5.78

Table 2 Properties of the histograms shown in Fig. 7. The underlying data has been analyzed for its mean μ , standard deviation σ , skewness γ , and kurtosis κ , where μ and σ are also the fit parameters of the respective normal distribution. To differentiate, the fit parameters of the generalized Student's t-distribution are written with a bar on top: $\bar{\mu}$, $\bar{\sigma}$, and \bar{v} .

also confirmed by the calculated Pearson coefficient in the top left corner of each scatter plot.

From this correlation analysis, it can be concluded that a small temporal headway τ between successive vehicles tend to remain almost constant, but is subject to individual, more or less disordered fluctuations. This may be due to the fact that the closer a vehicle approaches the vehicle in front, the less it is able to choose its personal speed. As a consequence, the vehicle will adapt the speed from its preceding vehicle without caring what the precise headway is, i.e. there will be a continuous switching between slight enlargement and reduction of τ leading to a small correlation coefficient c , finally. This is also in line with the three-phase traffic theory [23], in which the temporal headway (*time gap*) is constantly changing to a certain extent. In terms of larger headways τ (large c), where there is no disturbing vehicle directly ahead, another effect is observed. The driver is now able to maintain the selected state of motion over greater distances without constantly switching between slight acceleration and decelerating. As a result, a

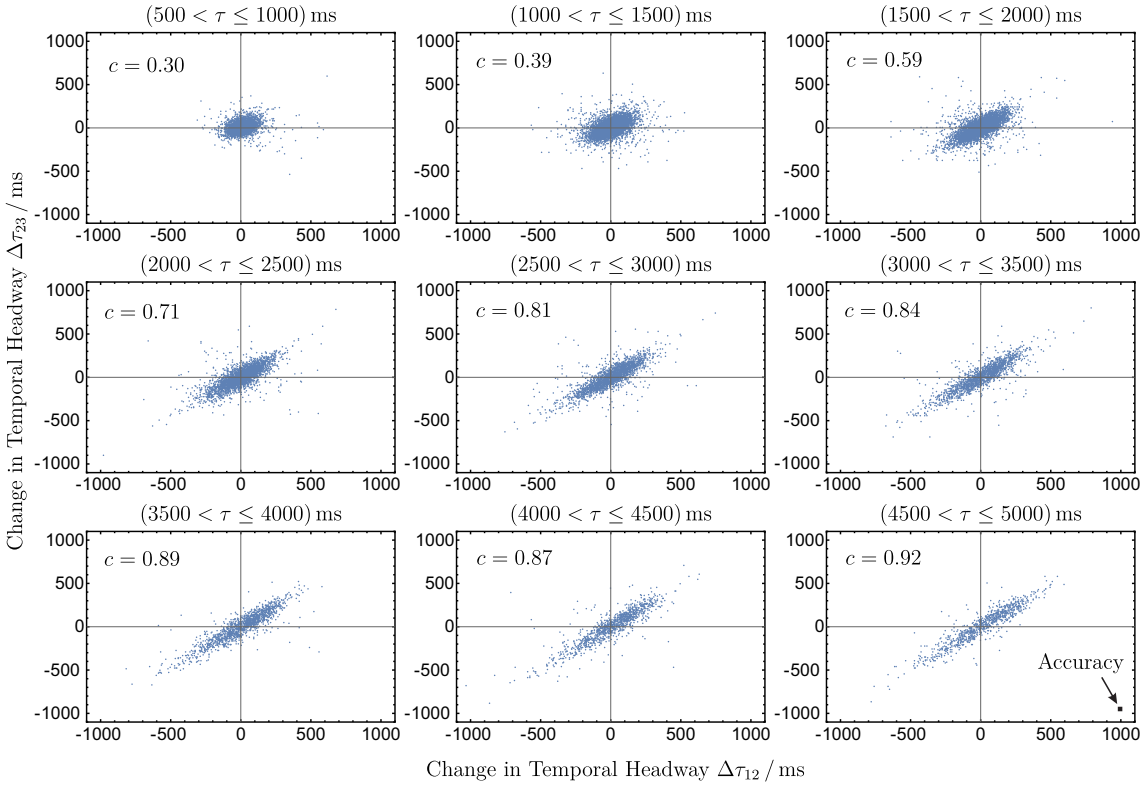


Figure 6 Scatter diagrams visualizing the correlation between $\Delta\tau_{12}$ and $\Delta\tau_{23}$ according to different temporal headways τ . Each point corresponds to exactly one vehicle. To get an idea about the related error in measurement of all nine shown diagrams, a black $(25 \text{ ms} \times 25 \text{ ms})$ -box named "Accuracy" has been added to the bottom right one. The actual accuracy of the microcontroller is around 1 ms. However, due to the fact that the sensor can only be read out every 25 ms, the error box must be adjusted accordingly. The associated Pearson correlation coefficient c is displayed in the top left corner of each plot. As one can already see here, c shows an increasing trend with increasing temporal headway τ . In order to quantify this fact, the related fit function $c(\tau)$ can be seen in Fig. 8. The axis labeling refers to all nine plots. All times are given in milliseconds (ms).

systematic change in temporal headway takes place, i.e. the gap becomes systematically larger or smaller. This in turn corresponds to a high positive correlation c between the two values of $\Delta\tau_{12}$ and $\Delta\tau_{23}$, which in this case are predominantly either both positive or both negative.

At this point the associated marginal distributions of $\Delta\tau_{12}$ and $\Delta\tau_{23}$ for the first three scatter plots from Fig. 6 (columnwise) should be presented to give the reader an idea about their shape. They are shown in Fig. 7. The best probability density function (PDF) to fit these histograms turned out to be the *generalized Student's t-distribution* (dashed blue line) taking into account the *heavy tails* of the underlying data. For comparison, the *normal distribution* is plotted to the data as well (red line). Please refer to App. A in order to see the exact mathematical form of the PDFs. All related parameters can be found in Tab. 2. The histograms are symmetrically (skewness $\gamma \approx 0$) distributed around zero and the mean value μ within the chosen measure of error $\sigma_{ij}^{(c)}$ known from the calibration

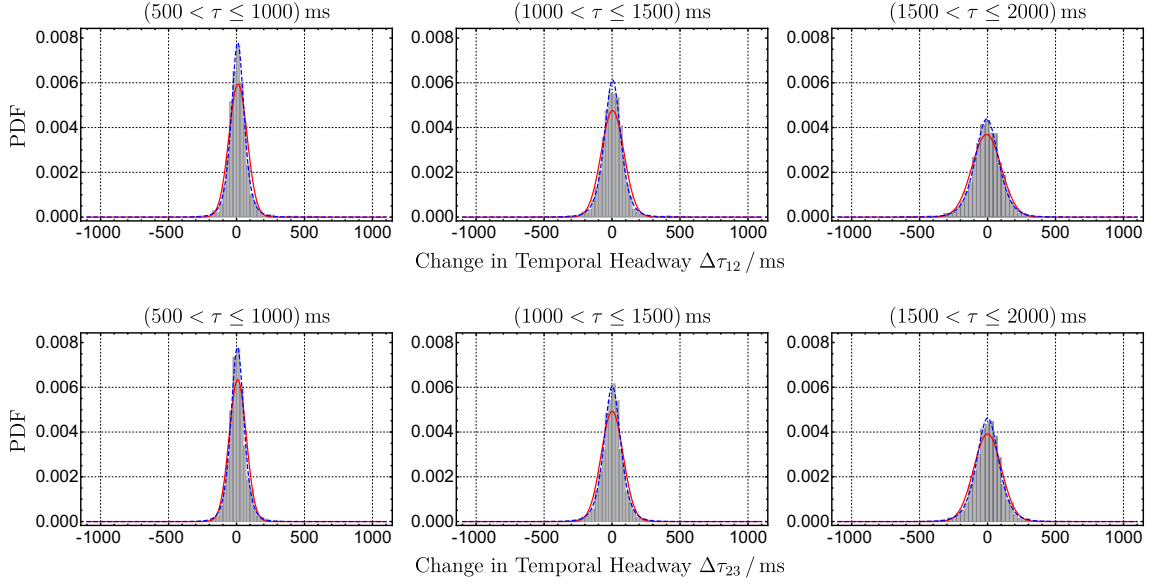


Figure 7 Marginal distributions (columnwise) of the first three scatter plots in Fig. 6. Each plot shows the empirically determined histogram of the change in temporal headway $\Delta\tau_{12}$ (first row of the plots) and $\Delta\tau_{23}$ (second row of the plots) within the first sector Δs_{12} and second sector Δs_{23} , respectively. The best probability density function (PDF) to fit the underlying histograms turned out to be the *generalized Student's t-distribution* (plotted in dashed blue). For comparison, the *normal distribution* is plotted to the data in red additionally. All related statistical parameters are indicated in Tab. 2.

process (see Tab. 1). As expected, the width of the $\Delta\tau$ -distributions (standard deviation σ) in Fig. 7 increases with the temporal headway τ , which probably can be explained due to a greater scope for action of the respective vehicle. In addition to the scatterplots in Fig. 6, it is also of interest to quantify the development of the Pearson correlation coefficient c for different headways τ . This is achieved by calculating c for the first thirty intervals of τ from Eq. 10. A fit function of the form

$$c(\tau) = -a \cdot \exp(-b \cdot \tau) + c_0 \quad (12)$$

has been applied to this data, where a and b are curvature parameters and c_0 the asymptotic parameter (upper boundary). Using the assumption $c(\tau = 0) = 0$ leads us to the final and simplified expression

$$c(\tau) = c_0(1 - \exp(-b\tau)). \quad (13)$$

The resulting plot is shown in Fig. 8. Obviously, the course of the underlying data points exhibits an increasing asymptotic behaviour, which justifies the use of the fit function according to Eq. 13. Based on the present data, the value of τ , from which the correlation is assumed to be constant and saturated, can be set somewhere between 3000 ms and 5000 ms.

These information enable a statistical short-term prediction for the spatiotemporal development of the temporal headway τ . By implementing an appropriate algorithm using the

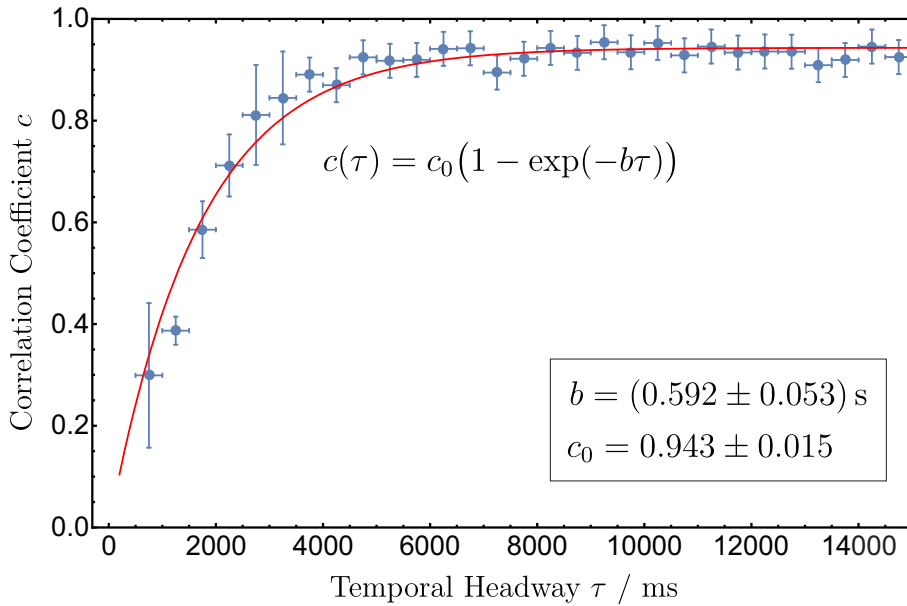


Figure 8 Pearson Correlation Coefficient c of the change in temporal headway $\Delta\tau$ between consecutive road segments according to different headways τ (detected at position P_1 , see Fig. 2). All data points are shifted to the middle of the associated interval by 250ms. A fit function (red) according to Eq. 13 is used to approach the underlying data. The associated parameters b and c_0 are listed at the bottom right corner. The resulting function is rapidly converging towards a very high value of saturation $c_0 > 0.9$. Please note that the units in terms of the fit function can be switched from seconds to milliseconds by applying $\text{s}^{-1} = 10^{-3}\text{ms}^{-1}$ to parameter b .

current vehicle configuration of a major street, an automated vehicle can adapt its speed in advance in order to merge into the ongoing traffic as efficiently as possible (see Fig. 1).

5. Conclusion and Outlook

In the coming decades, road traffic will be increasingly dominated by automated vehicles. In order to enable the smoothest possible interaction between them and conventional vehicles, a number of conditions must be fulfilled. One of the most important points in this context is a good estimation of the behaviour of human-controlled vehicles through associated algorithms. Especially in terms of lane changing and merging maneuvers at intersections (see Fig. 1), short-term predictions of conventional vehicles become indispensable with regard to a fluent traffic flow on the one hand, and individual safety, energy, and comfort conditions on the other hand. Therefore, this paper deals with a statistical analysis of the headway behaviour between consecutive human driven vehicles.

As already shown in our latest paper [1], it can be confirmed once again that the change in temporal headway $\Delta\tau$ increases with a growing temporal headway τ between consecutive vehicles (see Fig. 7). In the present work, this value was measured over a fixed distance of about 23.5m. The associated probability density function (PDF) best matching to the $\Delta\tau$ -data is a generalized Student's t-distribution, since it takes into account the heavy tails of the underlying histogram. For a simple mathematical treatment, however, even a nor-

mal distribution may be justified in order to approach the data satisfactorily.

When the road is splitted up into two adjacent sectors of the same length, in which $\Delta\tau$ is measured each, it is possible to gain even more information of the headway behaviour of two consecutive vehicles. This is done by determining the Pearson correlation coefficient c between the concerning $\Delta\tau$ -values for different temporal headways τ . In this case, the empirical distributions of $\Delta\tau$ from Fig. 7 serve as the respective marginal distributions of the related scatter diagrams shown in Fig. 6. Finally, the development of the Pearson coefficient c itself has been quantified according to different temporal headways τ (see Fig. 8). It turned out that there is a strongly positive correlated and asymptotic behaviour for large headways ($\tau > 5000$ ms). Between $\tau = 500$ ms and $\tau = 5000$ ms the correlation coefficient rises sharply from $c = 0.30$ to $c = 0.92$ and reaches its upper boundary value around there. This characteristic behavior justifies the usage of an exponential fit function as described in Eq. 13 with a boundary value of $c = c_0 = 0.943$.

Due to the high correlation values these results can be used to obtain short-term predictions and answer the question of how the investigated headway τ develops most likely in a hypothetical third sector. Because of the increasing correlation coefficient, the reliability of this forecast increases with the headway between the regarded vehicles.

Using this information, an automatic vehicle can now adapt its driving strategy in order to merge as smoothly as possible into the ongoing traffic. Nevertheless, it must always be prepared to react adequately to unforeseen events and, if necessary, revise its chosen parameters as all underlying predictions are of statistical origin.

Acknowledgements The authors thank the partners for their support within the project "MEC-View – Mobile Edge Computing basierte Objekterkennung für hoch- und vollautomatisiertes Fahren", funded by the German Federal Ministry of Economics and Energy by resolution of the German Federal Parliament.

References

- [1] Kumm, M., Schreckenberg, M.: Statistical analysis of temporal headway development through empirical data in urban traffic. In: 2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), pp. 1–9. IEEE (2019)
- [2] González, D., Pérez, J., Milanés, V., Nashashibi, F.: A review of motion planning techniques for automated vehicles. *IEEE Trans. ITS* **17**(4), 1135–1145 (2016)
- [3] Wadud, Z., MacKenzie, D., Leiby, P.: Help or hindrance? the travel, energy and carbon impacts of highly automated vehicles. *Transportation Research Part A: Policy and Practice* **86**, 1–18 (2016)
- [4] Wang, C.X., Cheng, X., Laurenson, D.I.: Vehicle-to-vehicle channel modeling and measurements: recent advances and future challenges. *IEEE Communications Magazine* **47**(11) (2009)

-
- [5] Milanes, V., Villagra, J., Godoy, J., Simo, J., Rastelli, J.P., Onieva, E.: An intelligent v2i-based traffic management system. *IEEE Transactions on ITS* **13**(1), 49–58 (2012)
- [6] Wagner, P.: Traffic control and traffic management in a transportation system with autonomous vehicles. In: *Autonomous Driving*, pp. 301–316. Springer (2016)
- [7] Bischoff, J., Maciejewski, M.: Simulation of city-wide replacement of private cars with autonomous taxis in berlin (2016)
- [8] Friedrich, B.: Verkehrliche wirkung autonomer fahrzeuge. In: *Autonomes Fahren*, pp. 331–350. Springer (2015)
- [9] Vranken, T., Sliwa, B., Wietfeld, C., Schreckenberg, M.: Performance comparison of dynamic vehicle routing methods for minimizing the global dwell time in upcoming smart cities. In: *88th IEEE Vehicular Technology Conference (VTC-Fall)*, 2018. IEEE (2018)
- [10] Lefèvre, S., Vasquez, D., Laugier, C.: A survey on motion prediction and risk assessment for intelligent vehicles. *ROBOMECH journal* **1**(1), 1 (2014)
- [11] Glaser, S., Vanholme, B., Mammar, S., Gruyer, D., Nouveliere, L.: Maneuver-based trajectory planning for highly autonomous vehicles on real road with traffic and driver interaction. *IEEE Transactions on intelligent transportation systems* **11**(3), 589–606 (2010)
- [12] Elbanhawi, M., Simic, M.: Sampling-based robot motion planning: A review. *IEEE access* **2**, 56–77 (2014)
- [13] Krbálek, M., Šleis, J.: Vehicular headways on signalized intersections: theory, models, and reality. *Journal of Physics A: Mathematical and Theoretical* **48**(1), 015101 (2014)
- [14] Cowan, R.J.: Useful headway models. *Transportation Research* **9**(6), 371–375 (1975)
- [15] Yuan, J., Mills, K.: A cross-correlation-based method for spatial–temporal traffic analysis. *Performance evaluation* **61**(2-3), 163–180 (2005)
- [16] Pan, T., Sumalee, A., Zhong, R.X., Indra-Payoong, N.: Short-term traffic state prediction based on temporal–spatial correlation. *IEEE Transactions on Intelligent Transportation Systems* **14**(3), 1242–1254 (2013)
- [17] Nagel, K., Schreckenberg, M.: A cellular automaton model for freeway traffic. *Journal de physique I* **2**(12), 2221–2229 (1992)
- [18] Chandler, R.E., Herman, R., Montroll, E.W.: Traffic dynamics: studies in car following. *Operations research* **6**(2), 165–184 (1958)

-
- [19] Ossen, S., Hoogendoorn, S.P., Gorte, B.G.: Interdriver differences in car-following: A vehicle trajectory-based study. *Transportation Research Record* **1965**(1), 121–129 (2006)
- [20] Monteil, J., OHara, N., Cahill, V., Bouroche, M.: Real-time estimation of drivers' behaviour. In: *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pp. 2046–2052. IEEE (2015)
- [21] Galton, F.: I. co-relations and their measurement, chiefly from anthropometric data. *Proceedings of the Royal Society of London* **45**(273-279), 135–145 (1889)
- [22] Fahrmeir, L., Heumann, C., Künstler, R., Pigeot, I., Tutz, G.: *Statistik: Der Weg zur Datenanalyse*. Springer-Verlag (2016)
- [23] Kerner, B.S.: *Breakdown in traffic networks: fundamentals of transportation science*. Springer (2017)

A. Appendix

Below, the probability density functions (PDFs) used in Sec. 4.2 are introduced shortly. Please note that this is not comparable to a detailed mathematical description.

A.1. Normal Distribution

The probability density function of the *normal distribution* is

$$\mathcal{N}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (14)$$

where μ is the *mean* and σ the *standard deviation*. The expression σ^2 is also known as the *variance*.

A.2. Generalized Student's T-Distribution

The *generalized Student's t-distribution* has the probability density function given by

$$f(x; \bar{\mu}, \bar{\sigma}, \bar{\nu}) = \left(\frac{\bar{\nu}}{\bar{\nu} + \frac{(x-\bar{\mu})^2}{\bar{\sigma}^2}}\right)^{\frac{1+\bar{\nu}}{2}} \cdot \frac{1}{\bar{\sigma}\sqrt{\bar{\nu}} \cdot B(\frac{\bar{\nu}}{2}, \frac{1}{2})} \quad (15)$$

where $\bar{\mu}$ is called *location parameter*, $\bar{\sigma}$ the *scale parameter*, and $\bar{\nu}$ the *degrees of freedom*. The expression $B(y, z)$ is known as the *Euler beta function* and is defined as follows:

$$B(y, z) = \int_0^1 k^{y-1} (1-k)^{z-1} dk \quad (16)$$

Please note that the related parameters of f are usually written without a bar. Here, we want to depict the difference to the parameters of the normal distribution as used above (see Tab. 2).