

Bivariate time-periodic Fokker-Planck model for freeway traffic

F. Lenz^a, D. Herde, A. Riegert, and H. Kantz

Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, 01187 Dresden, Germany

Received 26 May 2009 / Received in final form 7 October 2009

Published online 31 October 2009 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2009

Abstract Recorded data of the density of cars and their speed from a German motorway are modeled by a bivariate Fokker-Planck equation. In order to cope with the evident diurnal variation, we assume a 24 h-periodicity in the drift and diffusion coefficients of this equation. After fitting these and smoothing them by polynomials, we validate the model by comparison of the empirical densities and densities generated by the model dynamics. We show that the time dependence of the drift field is related to a saddle-node bifurcation due to which the congested traffic state becomes stable. The separatrix between the basins of attraction is used to define flowing and jamming traffic during rush hours and characterizes the traffic dynamics together with the fixed points and the centre manifold.

PACS. 02.50.Ga Markov processes – 05.10.Gg Stochastic analysis methods (Fokker-Planck, Langevin, etc.) – 45.70.Vn Granular models of complex systems; traffic flow

1 Introduction

Complex dynamical behaviour with irregular fluctuations is very often generated by the interaction of two mechanisms: nonlinear feedback loops usually destroy the validity of simple superposition principles and they also cause nonlinear input-output relations. Irregular driving forces can best be assumed to be noise which acts as additional input to the system. In such a general situation, the best modeling approach is a nonlinear stochastic model itself. In continuous time, the most suitable model class is that of Langevin equations and the equivalent Fokker-Planck equations, since here forces can be decomposed into deterministic components and potentially state dependent stochastic components. The drawback of this model class is that it describes only the subclass of Markov processes. Therefore, the attempt to describe a given data set by a Langevin- or Fokker Planck equation relies on a Markov hypothesis for the given data. Friedrich and Peinke demonstrated in a sequence of pioneering papers [1,2] that one can indeed extract the drift and diffusion coefficients of univariate and bivariate Fokker-Planck equations from experimental time series data. Successful applications of this technique include data from hydrodynamic turbulence, the stock market, the brain, etc. [2–5].

Highway traffic also displays strong irregular fluctuations which look stochastic, but there are evident nonlinear feedback loops with very strong deterministic components. In cellular automata models [6], e.g., deterministic rules for the velocity increment of an individual car as a function of the car in front of it together with stochastic fluctuations mimic real traffic quite well. Hydrodynamic models [7] describe traffic in a continuum approximation and hence contain exclusively deterministic feedback.

We base our model on data recorded at a single induction loop station of the highway ring around the German city of Cologne. Inspired by the work of Kriso et al. [8] we model these data by a Fokker Planck equation in two dimensions, where our variables are the density of cars measured in cars/km and the flux of cars measured in cars/min. The data represent the mean values of these two quantities averaged over 1 minute each. From the available data sets we choose one from a station where we have about 4 years of minute-mean data. Although the raw data is available for separate lanes of the highway, we use an effective one-lane modeling approach, averaging of the values over all three lanes.

Kriso et al. [8] show a vector field for the drift field in two dimensions, where clearly two rest states corresponding to the free flow and the congested flow can be discerned. However, as everyone can experience quite easily, traffic is a phenomenon with a very pronounced diurnal cycle. Traffic during the night differs strongly from the rush hours. Therefore we extend the analysis of [8] to a Fokker-Planck equation with time dependent coefficients, which are assumed to be periodic in time with a period of 24 h. When fitting drift and diffusion coefficients, we exclude all data taken from weekends, so that the weekly cycle is strongly suppressed and is averaged out. This is an approximation, since careful data analysis shows that also the working days differ slightly from each other and that in particular Monday and Friday are different from Tuesday to Thursday.

The main result of our paper will be a bifurcation diagram of the drift field as a function of the time of day. This will model the fact that congested traffic is hardly observable during the night hours. The separatrix between the basins of attraction of the two stable fixed points (if present) will define a natural classification into congested and free traffic flow. The centre manifold will connect our model to one-dimensional models.

^a e-mail: flenz@mpipks-dresden.mpg.de

In the next section we define the model which we want to use and the way how to fit its parameters. In Section 3 we discuss the empirically determined drift field and diffusion tensor and the bifurcation diagram, the centre manifold, and the separatrix and we validate our model in Section 4.

2 A Fokker-Planck model for traffic data

The data we analyzed were measured in Germany on the ring of motorways around Cologne. There are 43 gauging stations with induction loop detectors under the motorways which were (and are) used to detect cars. In our analysis we use the time of day φ and:

v the average speed of the cars [km/h]:

$$v := \frac{v_P q_P + v_T q_T}{q_P + q_T}$$

d the density [cars/km] given by $d := (q_P + q_T)/v$

which are derived from the data which was registered each minute on each lane of the motorway: v_P, v_T the average speed of passenger cars/trucks [km/h] and q_P, q_T the flow of passenger cars/trucks [cars/min].

We want to describe these data by a bivariate Fokker-Planck equation which is the advection diffusion equation for the distribution function or probability density function of the macroscopic variables of a Markov process:

$$\begin{aligned} \frac{\partial}{\partial t} w(\mathbf{X}, t) = & - \sum_{i=1}^m \frac{\partial}{\partial X_i} \left(D_i^{(1)}(\mathbf{X}, t) w(\mathbf{X}, t) \right) \\ & + \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2}{\partial X_i \partial X_j} \left(D_{i,j}^{(2)}(\mathbf{X}, t) w(\mathbf{X}, t) \right) \quad (1) \end{aligned}$$

where w is the probability density function (PDF), $\mathbf{D}^{(1)}$ is the drift vector, $\mathbf{D}^{(2)}$ is the diffusion tensor, \mathbf{X} is the vector in phase space and m is the dimension of the system. In our case, $m = 2$ and $\mathbf{X} = (v, d)$.

Before extracting the parameters of such an equation from the observed data, one should consider whether the data really represent a Markov process. Indeed, in [4] a test for the Markov property in terms of verifying the validity of the Chapman-Kolmogorov equation was suggested and used. Apart from the fact that it is a necessary but not a sufficient condition for the Markov property, which in addition is often violated on very short time scales (e.g., also by measurement noise on Markovian data), in our case its verification is even more difficult because of the explicit time dependence of our process. We therefore refrain from performing a test for Markovianity at this point and instead perform a model validation after we have determined the model parameters.

For Markov processes the estimation of the drift and diffusion coefficients is given by [9], p. 84:

$$\mathbf{D}^{(1)}(\mathbf{X}) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left\langle \tilde{\mathbf{X}}(t + \tau) - \tilde{\mathbf{X}}(t) \right\rangle_{\tilde{\mathbf{X}}(t) = \mathbf{X}} \quad (2)$$

$$\mathbf{D}^{(2)}(\mathbf{X}) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \left\langle \left(\tilde{\mathbf{X}}(t + \tau) - \tilde{\mathbf{X}}(t) \right) \cdot \left(\tilde{\mathbf{X}}(t + \tau) - \tilde{\mathbf{X}}(t) \right)^\top \right\rangle_{\tilde{\mathbf{X}}(t) = \mathbf{X}} \quad (3)$$

We estimate the drift and diffusion coefficients from a time series $\tilde{\mathbf{X}}(t)$ at discrete values of d and v by partitioning the data into 48×48 equally sized bins covering the phase space. Measured time series have a finite sampling rate which means one cannot take the limit of $\tau \rightarrow 0$ but one has to use the smallest τ available. In our application to traffic data this is $\tau = 1$ [min]. Due to this approximation we have to correct the diffusion term for finite time effects by the drift term¹:

$$\mathbf{D}^{(2)}(\mathbf{X}) \approx \frac{1}{\tau} \left\langle \left(\tilde{\mathbf{X}}(t + \tau) - \tilde{\mathbf{X}}(t) - \tau \mathbf{D}^{(1)}(\mathbf{X}) \right) \cdot \left(\tilde{\mathbf{X}}(t + \tau) - \tilde{\mathbf{X}}(t) - \tau \mathbf{D}^{(1)}(\mathbf{X}) \right)^\top \right\rangle_{\tilde{\mathbf{X}}(t) \approx \mathbf{X}} \quad (4)$$

2.1 Modeling periodic dependencies

In order to analyze the effect of the different times of the day on the dynamics of traffic flows we now extend our model by time-dependent Fokker-Planck coefficients. We assume a separation of timescales of:

- the fast dynamics of car density and average car velocity which is governed by a Fokker-Planck equation at each time of day;
- and the slow periodic change of the Fokker-Planck coefficients during the day.

This approximation can be seen as a model which consists of a family of Fokker-Planck equations indexed over the times of day or equivalently as only one Fokker-Planck equation which has the time of day $\varphi(t)$ as an additional variable in phase space. This is a quite general approach to model periodic influences on a system with a separation of timescales. We discretize the ‘time of day’ $\varphi(t)$ into $n_\varphi = 48$ time intervals of length $\Delta_\varphi = 0.5$ [h] in the same way as the other variables in phase space. The estimate of the Fokker-Planck coefficients is again done by equation (2) with the minor change that now also time t is subject to a condition, namely $t \in [\varphi, \varphi + \Delta_\varphi]$.

3 Results

3.1 Influence of the time of day on traffic flow

The results of the data analysis are time dependent drift vector fields $\mathbf{D}^{(1)}(v, d, \varphi)$ and time dependent diffusion tensors

¹ A more elaborate correction and a discussion of different corrections can be found in [10].

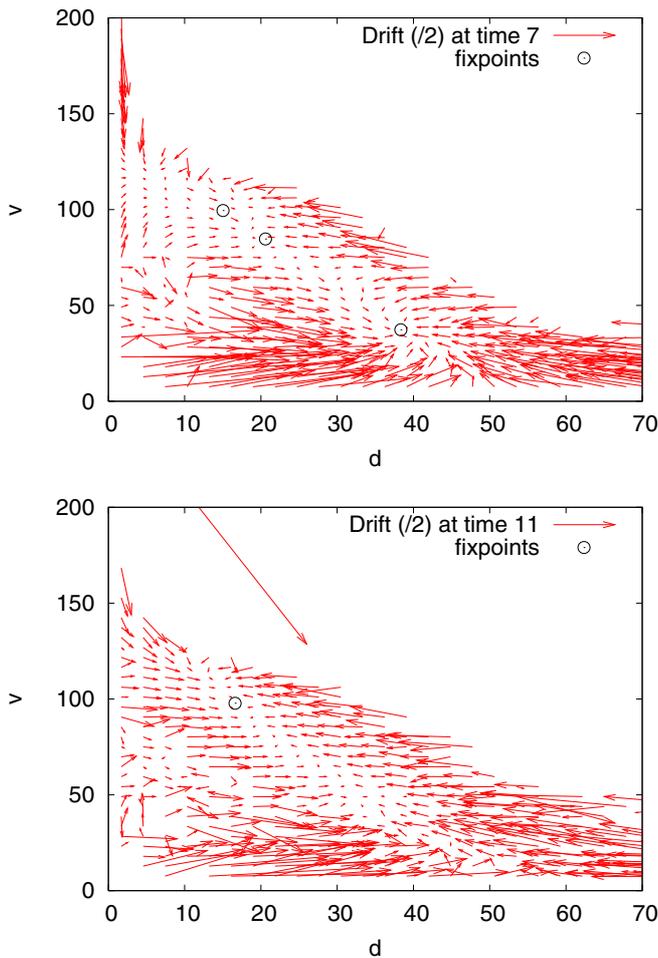


Figure 1. (Color online) Drift vector field over density d and velocity v from 7:30 to 8:00 (top) resp. from 11:30 to 12:00 (bottom).

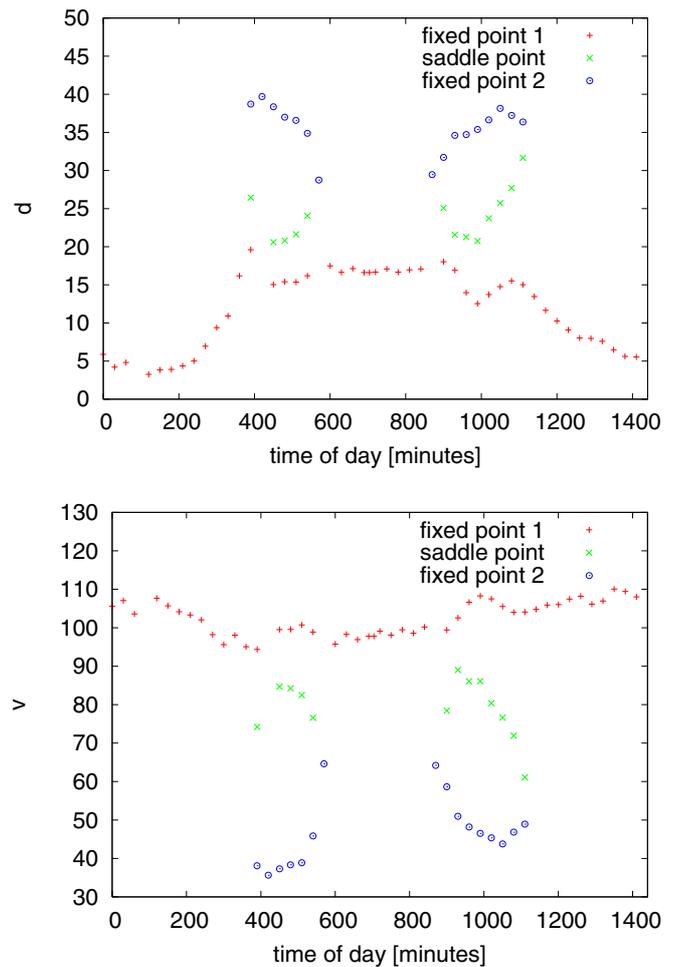


Figure 2. (Color online) The density-coordinate d (top) and the velocity-coordinate v (bottom) of fixed points depending on the time of day.

$D^{(2)}(v, d, \varphi)$. In Figure 1 we show two examples of the estimated drift coefficients as vector fields in phase space for different times of the day. The drift vector field changes during the day. If the drift field has a point in space where its modulus is zero, then this point represents a fixed point of the deterministic part of a corresponding Langevin equation. We observe that there always exists a stable fixed point at high velocities v and low densities d corresponding to free flow traffic. At rush hours there is also a second stable fixed point at low velocities and high densities corresponding to congested traffic states. At rush hours our drift coefficients are qualitatively the same as the drift coefficients estimated without a resolution of the time of day in [8].

Let us stress that these empirical vector fields are usually not gradient fields of potentials. Hence, at first sight the existence of fixed points seems to be a nontrivial result. However, the preconditions for the fixed-point theorem of Brouwer can be assumed to be satisfied: the empirically observed process is stable in the sense that the trajectory never leaves a bounded region around the origin (no speed exceeding 250 km/h, no density exceeding the one of a fully packed highway). Therefore, the deterministic dynamics induced by the drift vector field constitutes a map from a bounded domain into itself. This domain is simply connected. If in addition we assume smooth-

ness of the drift field (which is supported by the empirical data), then the Brouwer fixed point theorem applies and there exists at least one fixed point of the dynamics inside this domain.

One way to compress the information stored in the sequence of 48 of such vector fields is to study the time evolution of the fixed points. Indeed, we observe that during the night only one fixed point exists, and that the other one appears and disappears by a saddle-node bifurcation, together with an unstable fixed point. These bifurcations mark the beginnings and the ends of the rush hours in the morning and in the afternoon.

Along with the temporal change of the vector field goes a change of the region in phase space which is visited by the observed trajectories. Evidently, with any data analysis method one can only obtain information about the dynamics in those sub-regions of the phase space which is actually visited by the trajectories. Therefore, at first sight it might be surprising that we can observe the time trace of the fixed points including the saddle point which is unstable. However, it is evident that in a stochastic process with finite barriers a trajectory will always explore the neighbourhood of every stable fixed point. In our time dependent problem, the additional issue is whether the time scale of the process to relax is sufficiently fast compared to

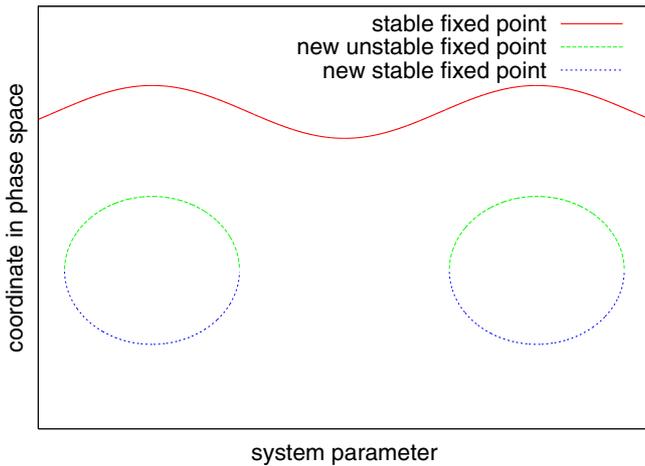


Figure 3. (Color online) Schematic plot of the tangent bifurcation at beginning and end of each rush hour.

the time scale on which the vector field changes its properties. This is verified a posteriori by the fact that within a time interval of Δ_φ the vector field changes only moderately, whereas the auto-correlations on the trajectory decay with several minutes (see Sect. 4).

3.2 Centre manifold reduction

A closer inspection of the drift field shows that it constitutes a slow and a fast variable. A typical initial condition subject to the deterministic drift force relaxes very fast to a one-dimensional manifold (see Fig. 4) and moves along this manifold to a stable fixed point. This centre manifold is an invariant manifold of the dynamics and can be constructed by an iterative algorithm: we start with a straight line through the stable fixed point(s) and evolve it forward in time using the polynomial fit to the empirical vector field. After every time unit, we extend the two ends of this line segment tangentially towards the border of our domain. After several iterations this procedure converges to a one-dimensional set which maps into itself. The result for different times of the day is shown in Figure 5. Remarkably, the manifolds thus constructed are rather robust against the change of the time of day. In particular, during the times shown the vector field undergoes forward and backward saddle-node-bifurcations. I.e., the dynamics on the centre manifold changes its character, whereas the manifold itself is almost time invariant. This manifold also nicely agrees with the empirically known *optimal velocity curve*, thereby verifying the microscopic laws of vehicle traffic which are basis to many microscopic models (e.g., [11]). However, in detail, some deviations from a strictly linear dependence of velocity on distance are visible.

The diffusion tensor as it results from our fit to data is also explicitly coordinate dependent. Therefore, the corresponding Langevin equations are driven by what is commonly called *multiplicative noise*, and only the proper noise terms together with the drift field generate the correct probability densities. However, the structure of the diffusion tensor cannot be easily interpreted and will not be discussed here any further.

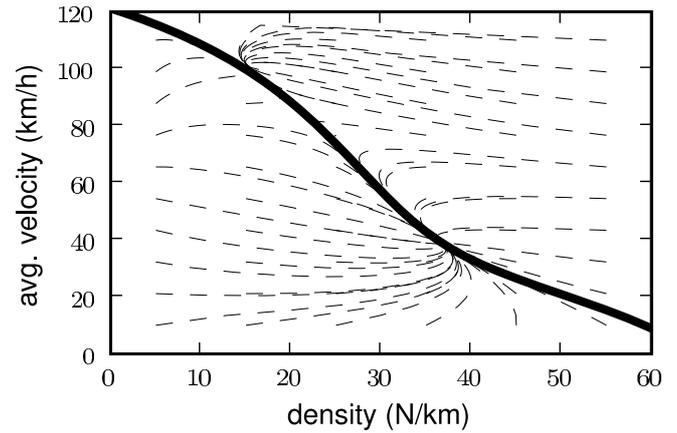


Figure 4. The centre manifold (bold curve) together with deterministic trajectories (dashed), evolving deterministically under the reconstructed drift field, on their way towards the centre manifold. Time of day: from 8:30 to 9:00.

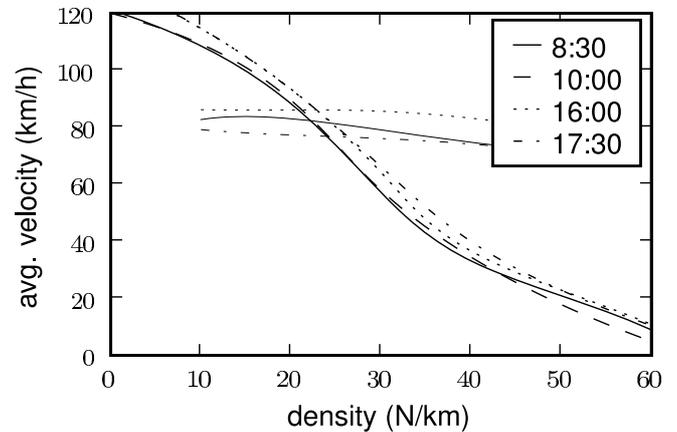


Figure 5. The centre manifolds representing four drift fields for different times of the day. Despite the fact that the dynamics *on* the centre manifold differs from time to time, the manifolds themselves are almost time invariant. The mostly horizontal curves are the corresponding places of the separatrix which exists only during rush hours.

Hence, our time dependent Fokker-Planck analysis reveals that the drift vector field does depend on time in the very essential way that there are bifurcations controlled by the time of day. These bifurcations determine whether the congested traffic is a (meta-)stable state or not. Despite this time dependence, the centre manifold which describes the mean relation between velocity and density is almost time invariant.

3.3 Finding the separatrix

In order to find the separatrix between the basins of attraction of the two stable fixed points at a given time of day we take two points A_i, B_i which evolve to different fixed points under the deterministic part of the Langevin equation in 4. We search with nested intervals on the line segment $[A_i, B_i]$ for the points C_i which are on the separatrix in Figure 5.

If we want to interpret it as the border between free flow and congested traffic states we have to consider that the position of

the separatrix also changes during the day and that it only exists during rush hours where there are two stable fixed points.

4 Validation

The drift vector fields were reconstructed from the data under the assumption that the time series data can be appropriately modeled by a Fokker-Planck equation. A Fokker-Planck equation describes a Markov process. Modeling the traffic data by a Fokker-Planck equation therefore implies for conditional probabilities $P(v', d', t'|v, d, t)$ that

$$P(v', d', t'|v, d, t; v'', d'', t'') = P(v', d', t'|v, d, t)$$

if $t'' < t < t'$, i.e., that the probability distribution at some future time depends exclusively on the probability distribution at the most recent time and not on the farther past. We did not perform any explicit check whether the data really support this property as the conditioning has too many dimensions for a statistical test of the equality. Instead, we will now validate the Fokker-Planck equation in two ways.

First, we will show that this equation generates probability densities which agree well with the observed data. To this end, we convert the Fokker-Planck equation into the corresponding Langevin equation,

$$\frac{d}{dt}\mathbf{X}(t) = \mathbf{g}(\mathbf{X}(t), t) + \mathbf{h}(\mathbf{X}(t), t)\boldsymbol{\Gamma}(t) \tag{5}$$

where $\boldsymbol{\Gamma}$ is a vector of *Gaussian white noise* with $\langle \Gamma_i(t) \rangle = 0$ and $\langle \Gamma_i(t)\Gamma_j(t') \rangle = \delta_{i,j}\delta(t - t')\forall i, j$. We first interpolate the Fokker-Planck coefficients by fitting multivariate polynomials to them. The reliability of the Fokker-Planck coefficients given by the number of data points in each bin used in the estimation is taken into account in the fitting procedure. As this is not a fit of a model but just an interpolation, we did not optimize the number of fitting parameters.

In order to be able to continue a sample path in the rare case, that it enters a region in phase space where the drift and diffusion terms cannot be determined due to absence of data, we extrapolate the field by taking the nearest known coefficients.

Then we convert the interpolated drift and diffusion coefficients to the Langevin coefficients $\mathbf{g}(\mathbf{X}, t)$ and $\mathbf{h}(\mathbf{X}, t)$ using the Itô interpretation of the Langevin equation:

$$\mathbf{g}(\mathbf{X}, t) = \mathbf{D}^{(1)}(\mathbf{X}, t) \tag{6}$$

$$\mathbf{h}^2(\mathbf{X}, t) = \mathbf{D}^{(2)}(\mathbf{X}, t). \tag{7}$$

We integrate the Langevin equation with the Euler-Maruyama approximation:

$$\mathbf{X}(t_{n+1}) = \mathbf{X}(t_n) + \Delta\mathbf{g}(\mathbf{X}(t_n), t_n) + \mathbf{h}(\mathbf{X}(t_n), t_n)(\mathbf{W}(t_{n+1}) - \mathbf{W}(t_n)) \tag{8}$$

where \mathbf{W} is the d-dimensional Wiener process and Δ is the time step. Comparing histograms in phase space of the original data and the generated data in Figures 6 and 7 show a qualitative agreement of the histograms for all times of day. This

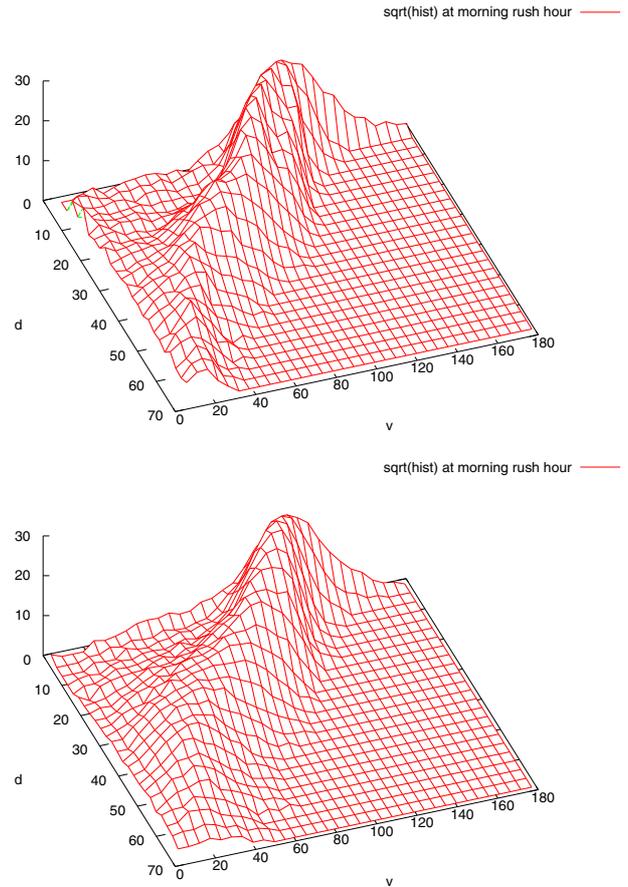


Figure 6. (Color online) Histogram from 8:00 to 8:30 of *original* (top) and *generated* (bottom) data over density d and velocity v . The square root of the number of data is taken to emphasize regions of low data density.

means that the assumptions of the model – the Markov property and a slow dependence of the Fokker-Planck coefficients on the time of day – did not induce large errors on the probability density function. Even when successful, such a model validation can only test necessary conditions but can never be sufficient in order to prove the correctness of the model.

Another, even less time consuming test consists in a comparison of the auto-correlation function of the original data and of sample paths generated with the Langevin equation. Due to the explicit time dependence (even if periodic) we do not average over all times but we compute the normalized auto-correlations of all data for a fixed time of day and average them over all working days:

$$c_\varphi(\tau) := \langle \mathbf{X}(t + \tau)\mathbf{X}(t) \rangle$$

where φ is the time of day (in this case 12–12:30 pm) and τ the delay in minutes. Figure 8 shows that, as expected [12], there is a fast decay. The deviation of the auto-correlation of the generated data is caused by the averaging over days with different dynamics.

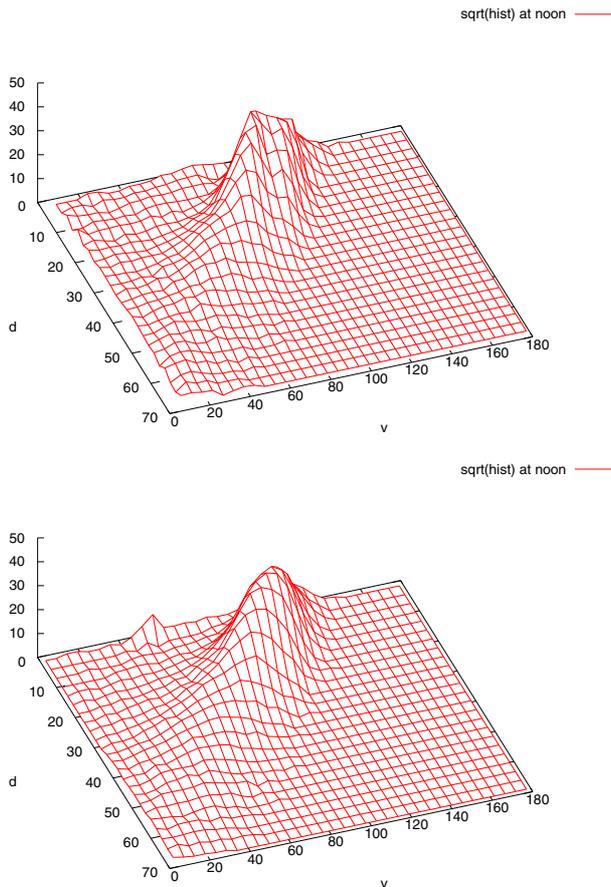


Figure 7. (Color online) Histogram from 12:00 to 12:30 of *original* (top) and *generated* (bottom) data over density d and velocity v .

5 Discussion

It is a common approach to approximate the scatterplot which one obtains when plotting the velocity values v versus the values of density d at the same times by a 1-dimensional curve. This relation between density and velocity is called the *fundamental diagram*. Our analysis in terms of a bivariate Fokker-Planck equation is in some sense a correction to this crude approximation. Indeed, from a theoretical point of view, such a curve could be a centre manifold of the deterministic part of the dynamics, which would mean that deviations from this curve are damped out very soon. Our vector field shows that this is indeed a reasonable approximation as density fluctuations are damped away mostly without affecting the velocity: this is true in the range of values where the arrows of the drift in Figure 1 are horizontal. However, in particular at times when two stable fixed points exist, there are also evident deviations, namely in the vicinity of these two stable fixed points, given by a rotational component in the drift field.

In principle, also the diffusion tensor contains information about the dynamics. The estimation of the diffusion coefficients permits a dependence of these on the state variable. Indeed, we observe that the diffusion tensor is not constant across the

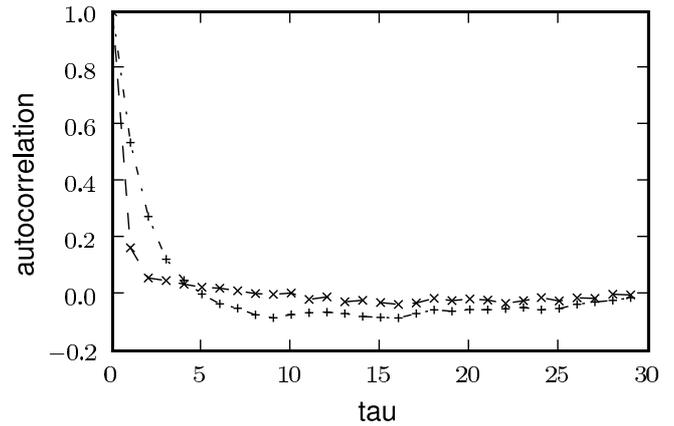


Figure 8. The normalized autocorrelations of original (\times) and generated ($+$) data at noon averaged over all days.

whole state space. However, this dependence does not lead to any conclusions, so that we do not show the numerical results.

In summary, fitting a time-periodic Fokker-Planck equation leads to a valid model for the traffic dynamics on normal working days. It does not only reproduce essential statistical features of the observed data, but supplies insight into details which would otherwise be inaccessible. So we learn about a bifurcation of the vector field, but also about the fact that a one-dimensional approach does not give a full description.

We would like to thank the ‘Landesbetrieb Strassenbau NRW’ for supplying the data and S. Albeverio and V. Jentsch for discussions. This project was supported by the DFG (project number: KA827/4).

References

1. S. Siegert, R. Friedrich, J. Peinke, Phys. Lett. A **243**, 275 (1998)
2. R. Friedrich, S. Siegert, J. Peinke et al., Phys. Lett. A **271**, 217 (2000)
3. R. Friedrich, J. Peinke, Physica D **102**, 147 (1997)
4. R. Friedrich, J. Peinke, Ch. Renner, Phys. Rev. Lett. **84**, 5224 (2000)
5. M. Siefert, J. Peinke, Int. J. Bifurcation and Chaos **14**, 2005 (2004)
6. K. Nagel, M. Schreckenberg, J. Phys. I France **2**, 2221 (1992)
7. M.J. Lighthill, G.B. Whitham, Proc. R. Soc. London, Ser. A **229**, 317 (1955)
8. S. Kriso, J. Peinke, R. Friedrich et al., Phys. Lett. A **299**, 287 (2002)
9. H. Risken, *The Fokker-Planck equation*, 2nd edn. (Springer, Berlin, 1989)
10. M. Ragwitz, H. Kantz, Phys. Rev. Lett. **87**, 254501 (2001)
11. D. Helbing, Rev. Mod. Phys. **73**, 1086 (2001)
12. L. Neubert, L. Santen, A. Schadschneider, M. Schreckenberg, Phys. Rev. E **60**, 6480 (1999)