Computer-Aided Civil and Infrastructure Engineering **26** (2011) 408–419

Reconstructing the Traffic State by Fusion of Heterogeneous Data

Martin Treiber* & Arne Kesting

Institute for Transport & Economics, Technische Universität Dresden, Falkenbrunnen, Würzburger Str. 35, D-01187 Dresden, Germany

&

R. Eddie Wilson

Department of Engineering Mathematics, University of Bristol, Queen's Building, Bristol BS8 1TR, United Kingdom

Abstract: *We present an advanced interpolation method for estimating smooth spatiotemporal profiles for local highway traffic variables such as flow, speed and density. The method is based on the "adaptive smoothing method" which takes as input stationary detector data as typically collected by traffic control centers. We generalize this method to allow for fusion with floating car data or other traffic information. The resulting profiles display transitions between free and congested traffic in great detail, as well as fine structures such as stop-and-go waves. We establish the accuracy and robustness of the method and demonstrate three potential applications: (1) compensation for gaps in data caused by detector failure; (2) separation of noise from dynamic traffic information; and (3) the fusion of floating car data with stationary detector data.*

1 INTRODUCTION

A detailed picture of speed and flow is essential for understanding flow breakdown on highways and the dynamics of congestion. In particular, highway traffic may not be understood by time series data alone but rather we must consider its structure in space and time jointly (Treiber et al., 2000; Bertini et al., 2005a,b). Let t and x denote, respectively, time and distance driven down the highway. We may thus introduce *spatiotemporal profiles* for the macroscopic variables velocity $V(x,t)$, flow $q(x,t)$, and density $\rho(x,t)$. Where the quality of data permits, these quantities may be displayed as color charts or landscapes which exhibit rich structure such as wave propagation, phase transitions, etc. To continue the analogy with particle physics, we may speak of the *traffic state*, by which we mean the classification of the traffic at any one time according to its spatiotemporal structure, e.g., as *free flow*, *synchronized flow*, *stop-and-go waves* (Kerner and Rehborn, 1996), or in terms of more detailed classifications (Helbing et al., 1999, 2009; Treiber et al., 2010), see also the searchable image database (Kesting and Treiber, 2010).

Unfortunately, highway traffic data comes from many heterogeneous sources (van Lint and Hoogendoorn, 2009). Most simply we have stationary detector data (SDD) collected by fixed infrastructure, which typically consists of inductance loops buried in the surface of the road. In their usual operation, the loops count vehicles and estimate their lengths and speeds, which are then sent to regional traffic control centers in the form of

*To whom correspondence should be addressed. E-mail: *treiber@vwi.tu-dresden.de*.

DOI: 10.1111/j.1467-8667.2010.00698.x

1-minute aggregate data. However, modern communication systems have sufficient bandwidth to carry full individual vehicle data (IVD) to the control center, and this data may lead to advances in incident detection algorithms, for example. More recent stationary detection systems operate in a similar fashion to inductance loops, but are based on magnetometers, and radar/laser/infrared devices installed on bridges.

Other traffic data is not provided at fixed points in space, for example floating-car data (FCD) from GPS devices (Fastenrath, 1997; Ivan and Sethi, 1998; Herrera and Bayen, 2008), or floating-phone data (FPD) (Caceres et al., 2008). Yet further data is produced on an event-oriented basis such as messages from the police.

Two problems with such highway traffic data sets are:

1. *Sparseness*. Each source of data individually may be insufficient to determine the traffic state. For example, the distance between consecutive stationary detectors may be too great to infer what is happening between them. This problem is compounded by detector failure. In the case of FCD, it is a relatively small proportion of vehicles that report GPS data.
2. Noise, of several different types, for example: (i) measurement error committed by detectors; (ii) sampling errors in aggregate data due to small numbers (a problem in low flow conditions); and (iii) heterogeneity of the driver/vehicle population (so that FCD for one vehicle may not be at all representative of those around it).

However, both sparseness and noise can be addressed to some degree by combining data from either similar or heterogeneous sources, and the focus of this article is a method which can fuse heterogeneous data in order to reconstruct smooth spatiotemporal profiles for local traffic variables such as the speed, generalizing the direct bridging of data gaps proposed by Chen et al. (2003).

Our method is based on the adaptive smoothing method (ASM) (Treiber and Helbing, 2002). Whereas the original ASM was restricted to SDD only, here we consider the generalized adaptive smoothing method (GASM) which is extended to cope with heterogeneous sources (Kesting and Treiber, 2008). In particular, the GASM can interpolate locally inconsistent data: for example, an FCD speed measurement need not equal a 1-minute aggregate SDD with the same (x, t) coordinates, but the GASM can nevertheless smoothly combine these disparate measurements.

Unlike the ASDA/FOTO method (Kerner et al., 2004), the GASM in its present form is not suited to online applications. Its power lies in the detailed reconstruction of the traffic dynamics from historical data which, in turn, may serve as the basis to improve actual online state estimators. Moreover, the GASM does not rely on a specific traffic theory but only on generally accepted facts about propagation velocities.

Both the ASM and GASM interpolate between data points (thereby tackling the problem of incomplete data coverage) and eliminate high-frequency noise whereas preserving most of the relevant dynamic information. However, their chief novelty is that they surpass nonspecialized smoothing methods by incorporating some traffic physics, in the form of well-understood wave propagation characteristics (Helbing et al., 2009). Specifically, it is known that in congestion, small perturbations to the traffic propagate upstream at about 15 km/hour, whereas in free-flow they propagate downstream at the speed of the vehicles (Kerner and Rehborn, 1996; Cassidy and Bertini, 1999; Schönhof and Helbing, 2007).

The chief contributions of this article are an investigation of the robustness and accuracy of the GASM and a demonstration of some potential applications, including the fusion of FCD and SDD. The organization is as follows. In Section 2, we formulate the GASM and discuss calibration and validation issues based on data from the German autobahn A9 and the English M42 motorway. Because the detector coverage on the M42 is extremely dense (loop detector spacings of 100 m), the validation (Section 2.5) will be based on real (not simulated) traffic dynamics which can be considered as completely known for our purposes. In Section 3, we propose possible applications of this method such as bridging data gaps (Section 3.1), separating noise from information (Section 3.2), and fusion of FCD and SDD (Section 3.3). Finally, the method and the results will be briefly discussed in Section 4.

2 GENERALIZED ADAPTIVE SMOOTHING METHOD

We now present the details of the GASM which performs two-dimensional interpolation to reconstruct the spatiotemporal traffic state from discrete traffic data. Speed data v_i measured at known locations x_i and times t_i are obtained from either SDD or FCD, or on an event-oriented basis, and are combined to produce a smooth velocity field as a function of continuous space and time. With minor modifications, flow and other spatiotemporal variables may also be reconstructed.

For the special case of SDD originating from K detectors at positions $\tilde{x}_k$ that send their information at common times $\tilde{t}_j$, the points (x_i, t_i) form a grid in the two-dimensional space–time plane, and the sum over all points i can be decomposed into a double sum over j

and k where $i(j, k) = jK + k$, $t_{i(j,k)} = \tilde{t}_j$, and $x_{i(j,k)} = \tilde{x}_k$. However, the formulation with a single index i is more general because it also includes arbitrarily scattered points in the xt-plane which is relevant for FCD and the fusion of heterogeneous data.

The GASM is based on two-dimensional interpolation in space and time using smoothing kernels (Section 2.1). However, in contrast to a conventional isotropic filter, the method incorporates the known characteristic velocities of information propagation in free and congested traffic (Section 2.2), by skewing the principal axes of the smoothing kernel. The switch between free and congested traffic is then managed by a nonlinear adaptive speed filter (Section 2.3). We then demonstrate the effectiveness of the GASM by comparing it with conventional smoothing with an isotropic kernel (Section 2.4). Finally, in Section 2.5, we validate the GASM using M42 data where the inductance loop system is overspecified. The approach is to apply the GASM to a subset of the inductance loop data and reconstruct the velocity field at the positions of detectors which have not been used in the interpolation. The accuracy of the GASM may then be established in comparison to the detector data which is regarded as the ground truth.

2.1 Conventional spatiotemporal interpolation

Our inputs are a set of discrete data points $\{x_i, t_i, v_i\}$, $i = 1, \ldots, n$, and the interpolation task is to derive from them a smooth velocity field $V(x, t)$ in a given spatiotemporal interval. The broad approach is to employ the convolution

$$V(x,t) = \frac{1}{\mathcal{N}(x,t)} \sum_i \phi_i(x - x_i, t - t_i)\, v_i \qquad (1)$$

where the smoothing kernels $\phi_i(x, t)$ are sufficiently localized functions that decrease with increasing $|x|$ or $|t|$, and we define the normalization factor by

$$\mathcal{N}(x,t) = \sum_i \phi_i(x - x_i, t - t_i) \qquad (2)$$

This formulation allows for different types of data points to use different kernels, but to simplify matters we shall usually assume that the kernels are identical and take the symmetric exponential form

$$\phi(x,t) = \exp\left[-\left(\frac{|x|}{\sigma} + \frac{|t|}{\tau}\right)\right] \qquad (3)$$

although a bivariate Gaussian would also be suitable. The smoothing kernel acts as a kind of low-pass filter, and the positive constants σ and τ define characteristic "widths" for the spatial and temporal smoothing, respectively, so that features with finer scales tend to be smoothed out.

Suitable values for σ and τ are of the order of half the typical distance between neighboring data points. For example, in a typical situation where an inductance loop system provides 1-minute aggregate data at 2 km intervals, we choose $\tau = 30\,\text{s}$ and $\sigma = 1\,\text{km}$ (see Section 2.5). However, for a stronger reduction of noise, larger smoothing widths may be chosen (e.g., τ up to 2 min).

2.2 Traffic-adaptive smoothing

A serious challenge in traffic data is that the typical scale of some traffic patterns, such as the wavelength of stop-and-go waves, is (at 1–2 km) similar to the spacing of stationary detectors. Consequently, important dynamical features may be lost in the interpolation process, and even entirely spurious patterns may be reconstructed (see Figure 2).

To enhance the resolution of the filter, we may use established facts concerning the propagation of information in traffic flow. The "propagation velocity" is that at which small perturbations to the traffic flow are propagated, and in traffic theories based on hyperbolic partial differential equations, it corresponds to the characteristic wave velocity given by the gradient of the equilibrium flow-density curve (the so-called "fundamental diagram"). It is well-known that:

1. In free traffic, perturbations move downstream (i.e., in the direction of traffic flow) (Kerner and Rehborn, 1996; Schönhof and Helbing, 2007) and the characteristic propagation velocity c_{free} is similar to the average speed of the vehicles constituting the perturbation.
2. In congested traffic, however, perturbations travel *against* the movement of the vehicles (i.e., upstream). Moreover, the characteristic propagation velocity c_{cong} is usually in the vicinity of -15 km/hour. This value is well established as the typical velocity of stop-and-go waves (Kerner and Rehborn, 1996; Schönhof and Helbing, 2007), but seems to apply more generally to nearly all information propagation in congested traffic.

To apply these facts, we skew the conventional isotropic smoothing kernel (3) in order to obtain the anisotropic interpolation formulae

$$V_{\text{free}}(x,t) = \frac{1}{\mathcal{N}(x,t)} \sum_i \phi\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{free}}}\right) v_i \qquad (4)$$

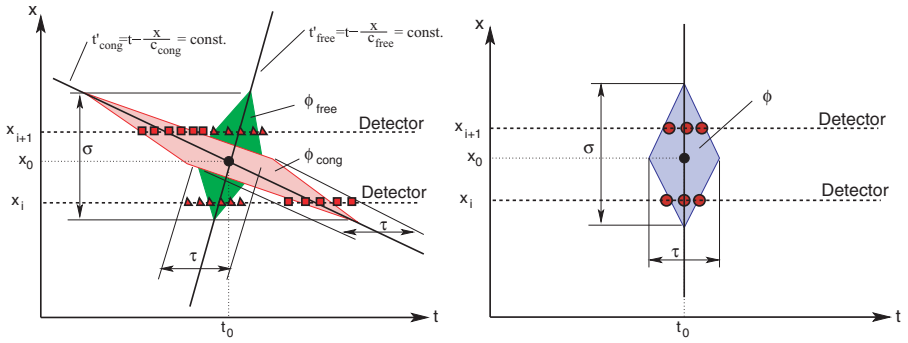

Fig. 1. (Left) Illustration of the smoothing kernels for free and congested traffic. The inclination angles capture the different characteristic velocities c_{free} and c_{cong}. In particular, perturbations propagate upstream (against the driving direction) in congested traffic. (Right) Illustration of conventional smoothing.

$$V_{\text{cong}}(x,t) = \frac{1}{\mathcal{N}(x,t)} \sum_i \phi\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{cong}}}\right) v_i \tag{5}$$

for free and congested traffic respectively. Here, $\mathcal{N}(x,t)$ is defined in analogy to (2) but using the corresponding skewed kernel. In effect, these new filters correspond to smoothing in preferred directions in the (x,t) plane (see Figure 1), based on the propagation velocities c_{free} and c_{cong}. Note that the conventional isotropic smoothing corresponds to the limit $c_{\text{free}} = c_{\text{cong}} \to \infty$.

In practice, we take $c_{\text{cong}} \approx -15\,\text{km/hour}$, whereas we find that $c_{\text{free}} \simeq +70\,\text{km/hour}$ gives good results in a highway context. The use of just two characteristic propagation velocities is consistent with kinematic wave models (Cassidy and Windower, 1995; Cassidy and Bertini, 1999) and the assumption of a piecewise linear ("triangular") fundamental diagram (Newell, 1993).

2.3 Nonlinear adaptive filter

We must now construct a single smoothing filter which combines the formulae for free and congested traffic. To this end, we define

$$V(x,t) = w(x,t)\, V_{\text{cong}}(x,t) + [1 - w(x,t)]\, V_{\text{free}}(x,t) \tag{6}$$

where the weight factor $w(x,t) = W(V_{\text{free}}(x,t), V_{\text{cong}}(x,t))$ controls the superposition of the free and congested velocity fields (4), (5). We require $W \approx 0$ at high speeds and $W \approx 1$ at low speeds, and thus we use the smooth s-shaped function

$$W(V_{\text{free}}, V_{\text{cong}}) = \frac{1}{2}\left[1 + \tanh\left(\frac{V_{\text{thr}} - \min(V_{\text{free}}, V_{\text{cong}})}{\Delta V}\right)\right] \tag{7}$$

Here, the "predictor" $\min(V_{\text{free}}, V_{\text{cong}})$ is defined such that the patterns of congested traffic are better reproduced by the resulting nonlinear filter than that for free traffic. The threshold between free and congested traffic is defined by V_{thr} whereas the transition width is determined by ΔV. Typical parameter values are given in Table 1.

Table 1
Parameters of the adaptive smoothing algorithm with typical numerical values used in this article. The spatial and temporal smoothing widths are chosen as half of the average inter-detector spacing Δx and sampling time δt, respectively

Parameter	*Value*
Smoothing width in space coordinate σ	$\Delta x/2$
Smoothing width in time coordinate τ	$\Delta t/2$
Propagation velocity of perturbations in free traffic c_{free}	70 km/h
Propagation velocity of perturbations in congested traffic c_{cong}	−15 km/h
Crossover from congested to free traffic V_{thr}	60 km/h
Transition width between congested and free traffic ΔV	20 km/h

2.4 Sensitivity analysis

For illustration, we apply the GASM to a small portion of 1-minute aggregate detector data from the Southbound A9 autobahn near Munich, Germany (Treiber et al., 2000; Treiber and Helbing, 2002). We consider data from nine detectors spread over 14 km of highway during 4 hours of a busy morning on which there were pronounced stop-and-go patterns. Here, as throughout the article, 1-minute aggregate SDD points also incorporate a flow-weighted aggregate of speed measurements across the lanes of the highway. In fact, in congested traffic, the speed variance between lanes tends to be rather small, and aggregation across lanes thus helps reduce sampling noise.

See Figure 2, which compares the performance of the GASM with the standard isotropic filter (1). Here the

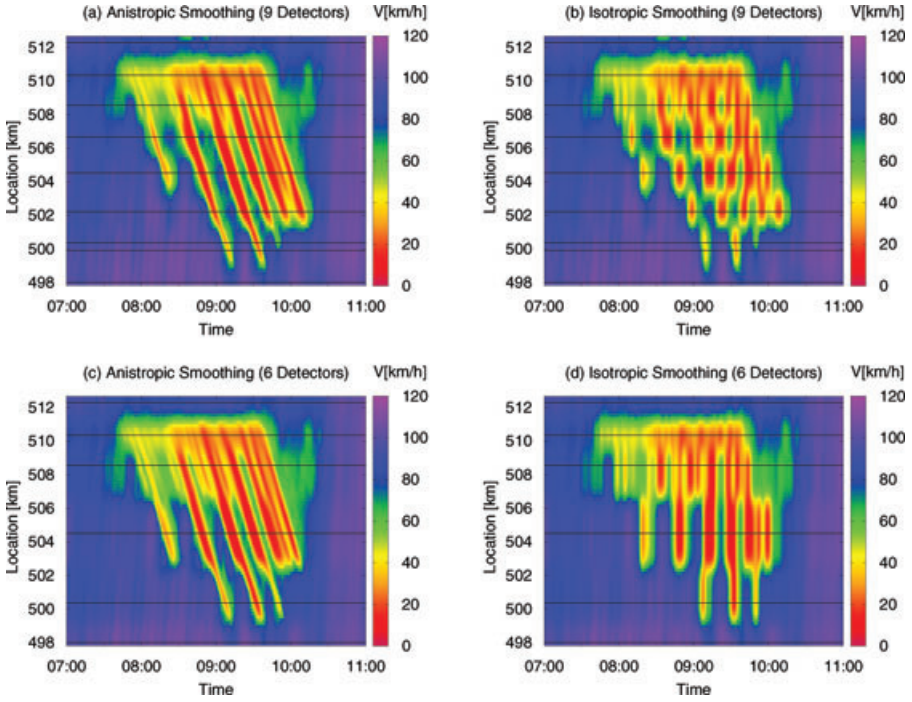

Fig. 2. Generalized adaptive smoothing method (a, c) versus conventional isotropic interpolation (b, d) applied to loop detector data from the German autobahn A9 (near Munich, direction South). In (c, d), only six out of the nine detectors are used in reconstruction. Conventional interpolation may neglect important spatiotemporal features or even identify spurious structure.

GASM uses the standard parameters from Table 1. The positions of the detectors which are used in reconstruction are indicated by horizontal lines. Compare Figures 2a and b: we may observe that the GASM is able to resolve individual stop-and-go waves when isotropic smoothing is not able to identify the pattern. As a more difficult challenge, in Figures 2c and d the smoothing algorithms are compared when data from only six out of the nine detectors is used. In this case isotropic smoothing identifies spurious patterns although the GASM continues to reconstruct stop-and-go waves correctly even though the spatial resolution of the input data is very poor.

Our experience is that the GASM performs better than isotropic smoothing in all of the data sets that we have tried. Moreover, we have found that the parameter choices in Table 1 are robust and do not need re-tuning for each new application. As an illustration, we apply the GASM to the same data as for Figure 2, but with pathological changes to the algorithm parameters. See Figure 3.

In Figure 3a, we change the propagation velocities c_{free} from 70 km/hour to 200 km/hour, and c_{cong} from -15 km/hour to -12 km/hour. In contrast, in Figure 3b, we modify the nonlinear filter, by reducing the transition width from $\Delta V = 20$ km/hour to 5 km/hour, and the crossover threshold from $V_{\text{thr}} = 60$ km/hour to 45 km/hour. To assess the robustness, these plots should be compared with the corresponding results for the GASM with standard parameters and for conventional isotropic smoothing, in Figures 2c and d, respectively.

On visual inspection, the quality of the GASM depends only weakly on its parameters and in all cases surpasses isotropic smoothing at this spatial resolution. We have found that the most sensitive parameter is the propagation velocity c_{cong} for congested traffic, because too low or too high values result in step-like artifacts in the reconstruction. However, in practice, c_{cong} varies very little from situation to situation (Kerner and Rehborn, 1996; Schönhof and Helbing, 2007).

In conclusion, we have yet to perform a formal optimization of the GASM parameters. But our experience is that the GASM reconstructs traffic patterns robustly with the parameter choices of Table 1 and stationary detector spacings up to about 3 km. Near a bottleneck, this spacing should preferably be reduced so that the stationary downstream jam front can be accurately positioned.

2.5 Validation

For validation of the GASM, we consider a 15-kilometer long section of the North-bound M42 motorway near Birmingham, England. As part of

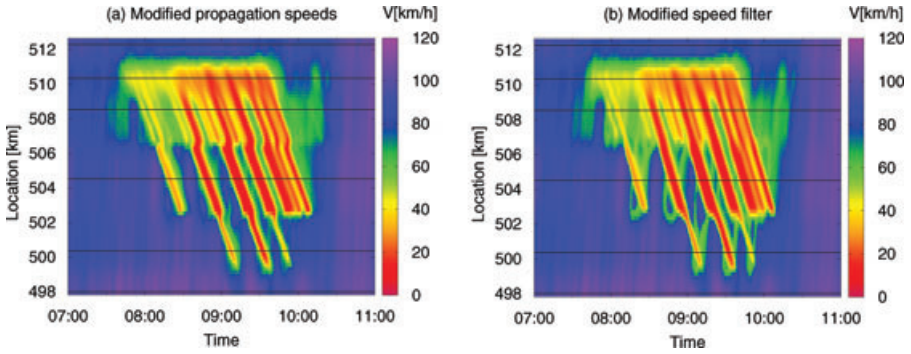

Fig. 3. Sensitivity of the generalized adaptive smoothing method to variation in its parameters (cf. Figure 2, full details in the main text).

the English Highways Agency's Active Traffic Management system (English Highways Agency's Active Traffic Management system homepage, 2009) this highway has been equipped with an almost unprecedented coverage of inductance loop detectors, with a typical nominal spacing of 100 m. In consequence, spatiotemporal patterns may be identified without any interpolation process at all. Thus in effect, the ground truth is directly available and we may use it to definitively evaluate the performance of interpolation algorithms.

As a test case, we take data from Friday January 11, 2008, and in Figure 4 we display a scatter plot of

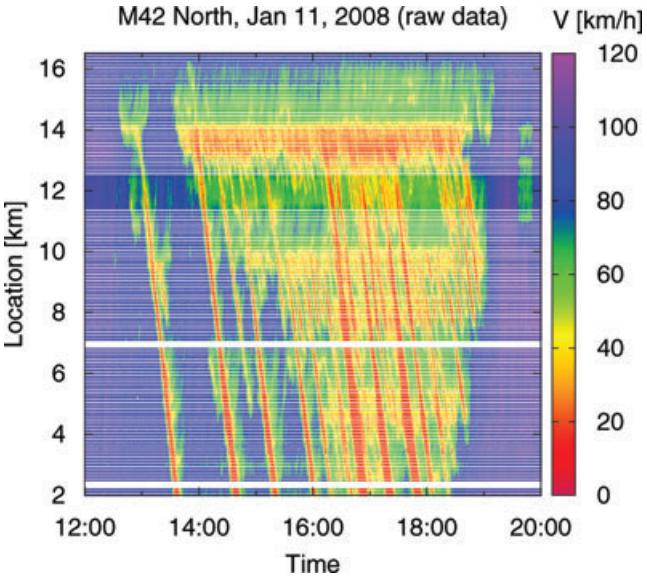

Fig. 4. Reference situation used for validation. The inter-detector spacing is 100 m (40 m in the vicinity of $x = 12$ km). The data are visualized as a spatiotemporal scatter plot. Each data point corresponds to the local speed aggregated over all lanes and over 1 min. No further data processing has been applied.

1-minute aggregate lane-average speed, showing a complex spatially extended pattern incorporating several bottlenecks and large amplitude stop-and-go waves. Close examination of this raw data largely supports the use of just two distinct propagation velocities c_{free} and c_{cong}, validating the overall GASM technique.

For a quantitative investigation, we apply the GASM with standard parameters to input data chosen from just a small selection of the available detectors. The interpolated field $V(x, t)$ is then compared to speed data at detectors which are halfway between those whose data has been used in the reconstruction. For example, at a spacing of 1 km corresponding to Figure 5a, the error measure is based on the detectors at $x = 2.5$ km, $x = 3.5$ km and so forth. Figure 5 displays the reconstructed traffic states with reduced sets of loop detectors. In summary, the most important features are identified even when the detector spacing is increased to 4 km.

Figure 6 presents RMS errors of the reconstructed velocities, averaged over all applicable test sites, as a function of the detector spacing. To assess the quality of the GASM, we compare it with conventional isotropic smoothing, that is, setting c_{free} and c_{cong} to ∞. For a given detector spacing, the quality of the GASM can be compared to that of conventional smoothing when about twice as many detectors are available. Specifically, when using the GASM, the quality of the reconstruction at a detector spacing of 2.5 km is comparable to that of isotropic smoothing when detectors are available every kilometer. To demonstrate the robustness of the GASM with respect to incorrectly specified parameter values, we also plotted the reconstruction quality for the obviously incorrect propagation velocities $c_{\text{free}} = 200$ km/hour, and $c_{\text{cong}} = -12$ km/hour (cf. Figure 3a). Remarkably, in comparison with conventional smoothing, the quality deteriorates only by a small amount.

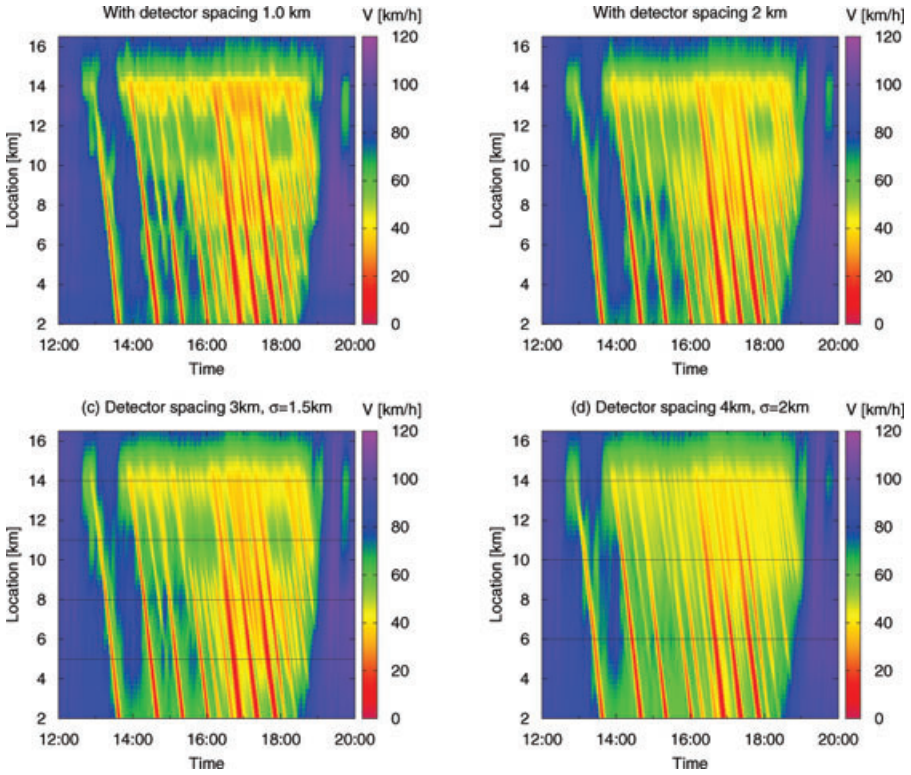

Fig. 5. Reconstruction of the reference situation of Figure 4 by the adaptive smoothing method (standard parameter set) applied on reduced data sets with detector spacings between 1 and 4 km. The locations of detectors whose data has been used in the reconstruction are indicated by horizontal lines.

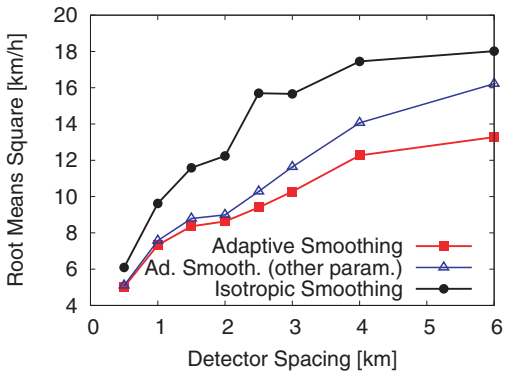

Fig. 6. Reconstruction of time series for speed using the adaptive smoothing method (lower curve), the incorrectly specified adaptive smoothing method (middle curve, parameters as in Figure 3a), and conventional smoothing (upper curve). Shown is the RMS velocity deviation with respect to the actual measurement, averaged over all detector sites that are halfway between the detectors available to the reconstruction methods. The deviation is plotted as a function of the spacing of the available detectors.

3 APPLICATIONS

We now give three illustrative examples of potential applications for the GASM: (1) Compensation for detector failure (Section 3.1); (2) Separation of noise from true traffic dynamics (Section 3.2); and (3) Fusion of FCD with SDD (Section 3.3).

3.1 Compensation for detector failure

Common operational problems with inductance loop systems include

- a temporary but simultaneous failure of all detectors covering a certain road section, usually due to a failure in the communications subsystem; or
- the permanent failure of one or a few detectors, often due to their installation not meeting specified standards.

In either case, the resulting SDD has "gaps," and we may estimate the data that is missing by interpolating the data that is extant.

Our example is of the former type and is taken from the South-bound A5 autobahn near Frankfurt,

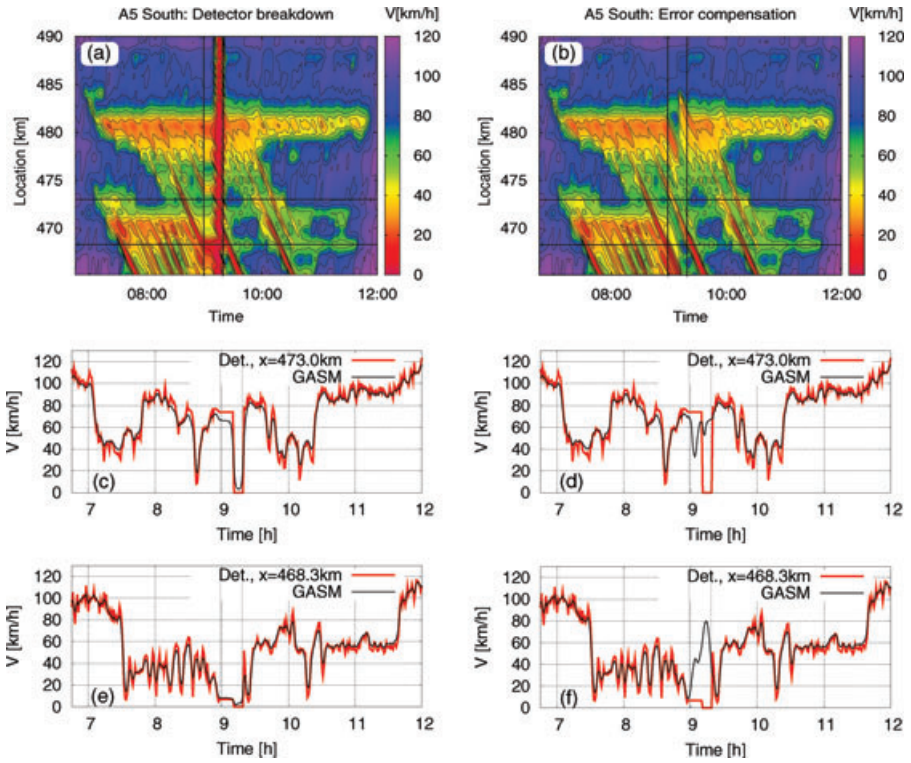

Fig. 7. Detector failure between 08:59 and 09:19 (South-bound A5 autobahn on August 6, 2001). Vertical lines in all plots indicate the period of failure. Plots (a, c, e) show results when the GASM incorporates erroneous data and (b, d, f) shows the correct results when the GASM omits the erroneous data and fills the gap by interpolation.

Germany, on the morning of August 6, 2001, see Figure 7 and (Schönhof and Helbing, 2007) for full details of the site. There is a 20 min breakdown of all detectors between 08:59 and 09:19. During this period, the detection system becomes "frozen" for 10 minutes and then records zero speed for a further 10 minutes. Figure 7a displays the resulting spurious spatiotemporal field recovered by the GASM.

However, additional error bits signal the detector failure and the correct approach is thus to eliminate the period 08:59 to 09:19 and apply the GASM to what remains, see Figure 7b. Observe that the GASM can bridge the gap in the data in a mostly natural way, although an artifact is introduced in the transition from congested to free traffic at the bottleneck at kilometer 482. This is because the GASM is well-tuned to reconstruct missing information in structures with velocities c_{free} and c_{cong}, but not to reconstruct stationary structures.

To further clarify the properties of the reconstructed state, Figures 7c–f display time series of speed at two selected detectors. Each plot shows the original detector data, and Figures 7c and e compare it with the erroneous GASM reconstruction whereas Figures 7d and f compare with the correct reconstruction with the erroneous data removed. Notice that the reconstructions in Figures 7d and f do not correspond to a simple interpolation in time, because of the way in which the GASM incorporates spatial information from other detectors.

Further investigations have shown that temporal data gaps of up to 30 min and spatial gaps up to 3 km can typically be compensated for. In the latter case, however, the result depends strongly on the position of the failed detectors relative to bottlenecks.

3.2 Elimination of noise

A significant problem in SDD is noise due to the time interval (typically 1 minute) over which raw data is aggregated by the road-side signal processing subsystems. In particular, when the traffic flow rate is low, the aggregate may be constructed from only a handful of individual vehicle observations, which themselves are prone to measurement error, and in consequence it is subject to gross statistical sampling error. The two situations of concern are dense queuing traffic and sparse high-speed traffic, and the latter case is subject to the additional

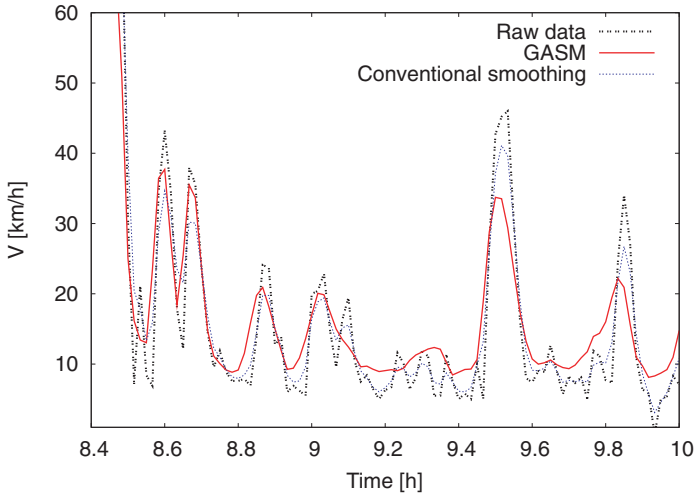

Fig. 8. Time series for a stationary detector at location $x = 473.0$ km on the German autobahn A5-South (see Figure 7 for location—but this data is taken from a different day). The original data (dashed line) is compared with the reconstruction at this position using the generalized adaptive smoothing method (thick solid), and with conventional smoothing (thin solid line).

problem that the variance in individual vehicles' true velocities may be very high.

A common real-time application is to use speed and flow thresholds from single detectors to trigger queue warnings on Variable Message Signs immediately upstream of a stop-and-go wave. It is thus crucial that stop-and-go waves are identified correctly and that alerts are not triggered by statistical outliers. To reduce the effect of sampling error, one may aggregate data over wider time intervals. However, there is not a clear separation of time scales and sharp features such as the boundaries of the waves may get lost if the aggregation window is broadened too far. Hence in general, it is nontrivial to separate noise from dynamics in detector time series.

For illustration, we again consider the South-bound A5 autobahn near Frankfurt, Germany (Schönhof and Helbing, 2007), with data now taken from July 9, 2001, on which there were a number of strong stop-and-go waves. See Figure 8, which displays the time series of speed from a single stationary detector and compares conventional temporal smoothing using the kernel $\exp(-|t|/\tau)$ with $\tau = 60$ s and the GASM with standard parameters.

In summary, the GASM is much more effective than simple temporal filters at reducing noise whilst retaining structure. In effect, the GASM uses spatial information, by blending the time series of nearby detectors, to enhance vehicle counts without broadening the temporal averaging window. The difference between noise and real traffic oscillations can thus be identified, because traffic information is correlated between detectors whereas noise is not.

3.3 Fusion of FCD

As intelligent transport systems are rolled out further across our trunk road networks, savings in infrastructure costs may be achieved if inductance loops are installed much more sparsely than is typical for the busiest highways. However, a detailed picture of traffic flow may still be derived if we also have access to some FCD, for example from high-end GPS services (such as the ITIS (ITIS Homepage, 2009) or Trafficmaster (Trafficmaster – Intelligent Driving, 2009) systems in the United Kingdom) or from data from the mobile phone networks, that is so-called FPD (Caceres et al., 2008) which is incorporated into high-end navigation services such as TMCpro in Germany (TMC-pro, 2009).

As a final application of the GASM, we consider the fusion of FCD and SDD. Our illustration uses synthetic data generated by microsimulation based on the Human Driver Model (HDM) (Treiber et al., 2006), which is a refinement of the well-known Intelligent Driver Model (Treiber et al., 2000). In the HDM, the reaction-time and multi-anticipatory effects have been calibrated so that the quantitative details (wavelength, amplitude, etc.) of spatiotemporal traffic patterns are reproduced.

Our specific example simulates a 12 km section of a single-lane highway without junctions over more than 2 hours. Rush-hour conditions are simulated by a peak in the in-flow (traffic demand) at $x = 0$ km. To induce traffic patterns, a flow-conserving bottleneck is applied at $x = 11$ km by increasing the car-following model's time-headway parameter at this point. Four stationary detectors are then simulated at $x = 2, 8, 10, 12$ km, and

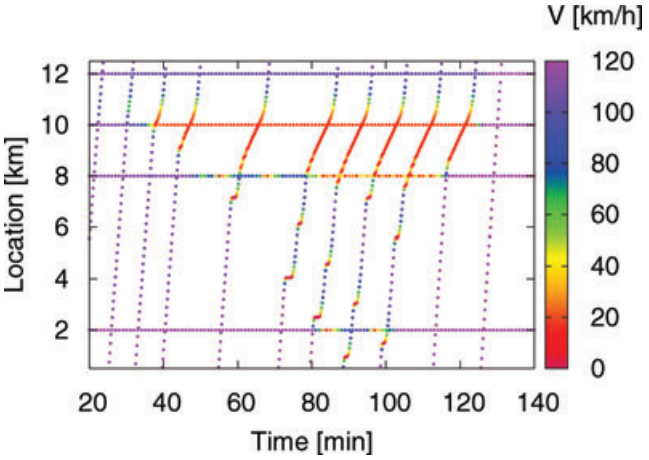

Fig. 9. Scatter plot of simulated floating car and stationary detector speed data.

just 10 vehicles out of a total of 2,750 in the simulation are selected randomly to provide speed data at 10 s resolution. These SDD and FCD are presented in the scatter plot in Figure 9.

We then reconstruct spatiotemporal velocity profiles by applying the GASM to different combinations of the SDD and FCD data. See Figure 10. Here Figure 10a displays the ground truth profile derived by using FCD from all vehicles in the simulation. Note that for the reconstruction based on SDD alone, not only interpolation but also extrapolation is required at the extreme upstream and downstream locations.

In Figure 10d SDD and FCD are combined using the same kernel function, except that each FCD point has been given double the weight of an SDD point. In practice one would need to experiment with this relative weighting and/or tune it to the particular application under consideration.

In summary, the SDD reconstruction outperforms FCD in resolving the stationary downstream front at the bottleneck at $x = 11$ km, whereas the FCD reconstruction outperforms SDD in resolving stop-and-go waves. The reconstruction using both FCD and SDD together combines the best properties of either individual reconstruction. This example shows that the incorporation of even very small amounts of FCD (one vehicle in every 200 or 300) can significantly improve the reconstruction of the traffic state.

4 DISCUSSION

In summary, in this article we have demonstrated how the GASM can solve practical problems in the analysis of highway traffic patterns. In particular, it can (1) compensate for gaps in data caused by detector failure; (2) reduce noise due to either sampling or measurement error (whilst compromising resolution less than a purely temporal filter); and (3) fuse heterogeneous data. Although we have focused on the reconstruction of the spatiotemporal velocity field, other fields (e.g., flow) may also be reconstructed: to achieve this, use alternative data in the linear formulae (4) and (5) but

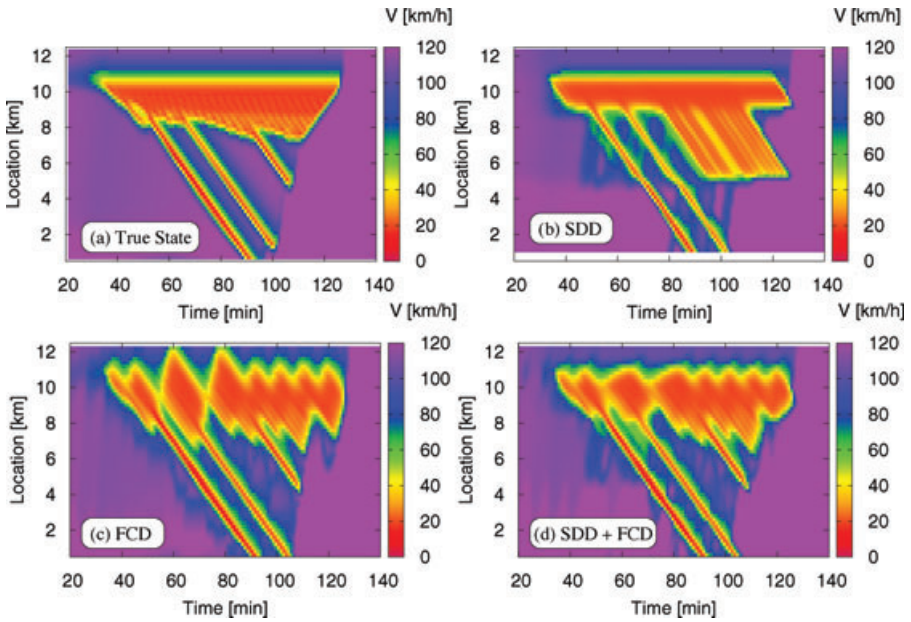

Fig. 10. Spatiotemporal speed profiles: (a) ground truth; (b) using stationary detector data only; (c) using floating car data only; (d) combining stationary detector and floating car data.

retain unchanged the dependence on velocity in the nonlinear component (6) and (7).

We have established that the performance of the GASM is quite robust to changes in its parameters. In consequence, it does not need tuning for each new scenario, and for freeway applications, the values from Table 1 may be used with confidence. A substantial benefit of the method is that it is not based on any one specific model of highway traffic and therefore may be used in the objective benchmarking of different models. On the other hand, a fine-tuning of the parameter c_{cong} may be used for research purposes to determine the propagation velocity of congested traffic patterns—and when formulated in r.m.s. error this method is related to the standard one based on the cross-correlation of time series from different stationary detector locations (Zielke et al., 2008).

However, the results here represent only a first analysis and a detailed calibration and validation study remains for future work. We have yet to carry out a formal optimization of the GASM parameters (e.g., using overspecified inductance loop data, as from the English M42 motorway), but there are other ways in which the details of the formulation that we have presented may be varied, and they should be analysed systematically using a variety of data from different highways and different countries. For example, we might experiment with different kernel functions (and note in practice these are implemented with a finite range "cut-off" for computational efficiency). Also, the relative weighting and parameterization of the kernels applied to different types of data point (individual vehicle or aggregate) needs investigation, and if possible a rigorous grounding in Bayesian statistics.

In fact, in either wholly congested or wholly free traffic, the GASM is linear and thus its performance may be compared definitively to a more general space of linear filters tuned for any one highway scenario. However, in practice, it is not so easy to generalize the nonlinear switch (6) and (7) which is key to the reconstruction of complete traffic patterns.

In principle, our method might be used to extrapolate data from the past into the near future, and thus form the basis of a short-term forecasting algorithm such as the ASDA/FOTO method (Kerner et al., 2004). In fact, stop-and-go waves are forecasted very well by this approach. The problem is the behavior of the GASM at bottlenecks where there is commonly a congestion pattern whose downstream front is stationary with respect to the highway (Helbing et al., 2009). At present, the GASM does not distinguish different congested flow regimes and so in extrapolation to the future, it will erroneously propagate this structure away from the bottleneck at velocity c_{cong}. The correction of this problem remains for future work, as does the incorporation of other aspects of traffic physics which may be useful in forecasting, such as the conservation of vehicles.

From the point of view of infrastructure design and specification, it is interesting to investigate the optimal positioning of a fixed number of stationary detectors, in order for the GASM to yield the best reconstruction of the velocity field. To answer this question we should again return to the performance of the method in an environment like the M42 motorway where there is a redundancy of detectors. We suspect that detectors should be placed more densely near bottlenecks, but the optimal configuration is unknown.

Finally, the increasing availability of FCD may allow much coarser SDD in future. A systematic investigation of the effects of average SDD spacing, SDD aggregation time, FCD penetration rate, and FCD aggregation time/event-based protocols remains for future work.

ACKNOWLEDGMENTS

The authors thank the Verkehrszentrale Hessen and the Autobahndirektion Südbayern for providing German autobahn data and the English Highways Agency for access to MIDAS motorway data sets. R.E. Wilson acknowledges the support of EPSRC Advanced Research Fellowship EP/E055567/1.

REFERENCES

Bertini, R. L., Boice, S. & Bogenberger, K. (2005a), Using ITS data fusion to examine traffic dynamics on a freeway with variable speed limits, in *IEEE Proceedings of Intelligent Transportation Systems*, 1006–11, DOI 10.1109/ITSC.2005.1520188.

Bertini, R. L., Hansen, S. & Bogenberger, K. (2005b), Empirical analysis of traffic sensor data surrounding a bottleneck on a German autobahn, *Transportation Research Record*, **1934**, 97–107.

Caceres, N., Wideberg, J. & Benitez, F. (2008), Review of traffic data estimations extracted from cellular networks, *Intelligent Transport Systems, IET*, **2**(3), 179–92.

Cassidy, M. J. & Bertini, R. L. (1999), Some traffic features at freeway bottlenecks, Transportation Research Part B: Methodological, **33**, 25–42.

Cassidy, M. J. & Windower, J. (1995), Methodology for assessing dynamics of freeway traffic flow, *Transportation Research Record*, **1484**, 73–79.

Chen, C., Kwon, J., Rice, J., Skabardonis, A. & Varaiya, P. (2003), Detecting errors and imputing missing data for single-loop surveillance systems, *Transportation Research Record*, **1855**, 160–67.

English Highways Agency's Active Traffic Management system homepage (2009), http://www.highways.gov.uk/knowledge/1334.aspx, accessed September 2010.

Fastenrath, D. (1997), Floating car data on a larger scale, in ITS-World Congress, DDG Gesellschaft für Verkehrsdaten mbH, Technical Report. Available at http://www.ddg.de/pdf-dat/ddgfcd.pdf, accessed September 2010.

Helbing, D., Hennecke, A. & Treiber, M. (1999), Phase diagram of traffic states in the presence of inhomogeneities, *Physical Review Letters*, **82**, 4360–63.

Helbing, D., Treiber, M., Kesting, A. & Schönhof, M. (2009), Theoretical vs. empirical classification and prediction of congested traffic states, *The European Physical Journal B*, **69**, 583–98.

Herrera, J. C. & Bayen, A. M. (2008), Traffic flow reconstruction using mobile sensors and loop detector data, in TRB 87th Annual Meeting Compendium of Papers DVD (paper 08-1868), Transportation Research Board of the National Academies, Washington DC.

ITIS Homepage (2009), http://www.itisholdings.com/, accessed September 2010.

Ivan, J. N. & Sethi, V. (1998), Data fusion of fixed detector and probe vehicle data for incident detection, *Computer-Aided Civil and Infrastructure Engineering*, **13**, 329–37.

Kerner, B. & Rehborn, H. (1996), Experimental features and characteristics of traffic jams, *Physical Review E*, **53**, R1297–300.

Kerner, B., Rehborn, H., Aleksic, M. & Haug, A. (2004), Recognition and tracking of spatio-temporal congested traffic patterns on freeways, *Transportation Research Part C: Emerging Technology*, **12**, 369–400.

Kesting, A. & Treiber, M. (2008), Calculating travel times from reconstructed spatiotemporal traffic data, in U. Martin, et al. (eds.), Networks for Mobility (Proceedings of the 4th International Symposium), FOVUS Universität Stuttgart, p. 41, ISBN: 978-3-921882-24-5.

Kesting, A. & Treiber, M. (2010), http://www.traffic-states.com, accessed September 2010.

Newell, G. (1993), A simplified theory of kinematic waves in highway traffic. Part I: general theory, *Transportation Research Part B: Methodological*, **27**(4), 281–87.

Schönhof, M. & Helbing, D. (2007), Empirical features of congested traffic states and their implications for traffic modeling, *Transportation Science*, **41**, 1–32.

TMC-pro (2009), http://de.wikipedia.org/wiki/TMCpro, accessed September 2010.

Trafficmaster – Intelligent Driving (2009), http://www.trafficmaster.co.uk, accessed September 2010.

Treiber, M. & Helbing, D. (2002), Reconstructing the spatio-temporal traffic dynamics from stationary detector data, Cooper@tive Tr@nsport@tion Dyn@mics, 1, 3.1–3.24, (Internet Journal, www.TrafficForum.org/journal).

Treiber, M., Hennecke, A. & Helbing, D. (2000), Congested traffic states in empirical observations and microscopic simulations, *Physical Review E*, **62**, 1805–24.

Treiber, M., Kesting, A. & Helbing, D. (2006), Delays, inaccuracies and anticipation in microscopic traffic models, *Physica A*, **360**, 71–88.

Treiber, M., Kesting, A. & Helbing, D. (2010), Three-phase traffic theory and two-phase models with a fundamental diagram in the light of empirical stylized facts, *Transportation Research Part B: Methodological* **44**(8–9), 983–1000.

van Lint, J. & Hoogendoorn, S. P. (2009), A robust and efficient method for fusing hetergeneous data from traffic sensors on freeways, *Computer-Aided Civil and Infrastructure Engineering*, **24**, 1–17.

Zielke, B., Bertini, R. & Treiber, M. (2008), Empirical measurement of freeway oscillation characteristics: an international comparison, *Transportation Research Record*, **2088**, 57–67.