Data driven uncertainty quantification in macroscopic traffic flow models

Alexandra Würth, Mickaël Binois, Paola Goatin, Simone Göttlich

To cite this version:

Alexandra Würth, Mickaël Binois, Paola Goatin, Simone Göttlich. Data driven uncertainty quantification in macroscopic traffic flow models. 2021. hal-03202124v1

HAL Id: hal-03202124

https://hal.science/hal-03202124v1

Preprint submitted on 19 Apr 2021 (v1), last revised 17 Nov 2021 (v2)

Data driven uncertainty quantification in macroscopic traffic flow models

Alexandra Würth[1] Mickaël Binois[1] Paola Goatin[1] Simone Göttlich[2]

April 19, 2021

Abstract

We propose a Bayesian approach for parameter uncertainty quantification in macroscopic traffic flow models from cross-sectional data. We validate the results comparing the error metrics of both first and second order models. While involving more parameters to be calibrated, second order models globally show better performances in reconstructing traffic quantities of interest.

2020 Mathematics Subject Classification: 35L65, 35R30, 62F15, 76A30

Key words: Macroscopic traffic flow models; Godunov scheme; loop detector traffic data; Bayesian calibration; parameter estimation; optimization.

1 Introduction

Macroscopic traffic flow models, consisting in hyperbolic partial differential equations based on the mass conservation principle, are employed since several decades to describe the spatio-temporal evolution of traffic aggregate quantities such as density and mean velocity on road networks. Compared to microscopic approaches, they offer the advantage of being computationally less expensive, and therefore adapted to large road networks. Moreover, their analytical properties make them suitable for solving optimal control problems motivated by traffic management issues. Last but not least, they involve a small number of parameters, thus reducing calibration cost. Yet, the inherent simplification of the dynamics induced by the models, their non-linearity and the data noise are all sources of challenging difficulties when dealing with parameter identification.

In this work, we focus in particular on the comparison between first order models, that are models consisting only in the mass conservation equation, here represented by the celebrated Lighthill-Whitham-Richards (LWR) model [35, 46], and second order ones, including a second equation accounting for the speed evolution.

Classically, macroscopic traffic models are calibrated either by fitting the so-called fundamental diagram, i.e., the density-flow or density-speed mapping described by the model flux

---

[1] Inria Sophia Antipolis - Méditerranée, Université Côte d'Azur, Inria, CNRS, LJAD, 2004 route des Lucioles - BP 93, 06902 Sophia Antipolis Cedex, France. E-mail: {`alexandra.wuerth, mickael.binois, paola.goatin`}`@inria.fr`

[2] University of Mannheim, Department of Mathematics, B 6, 28-29, 68159 Mannheim, Germany. E-mail: `goettlich@uni-mannheim.de`

function (see e.g., [6, 14, 15, 16, 17]), or by minimizing some error measure of the simulation output, against either data provided by loop detectors at fixed locations [41, 48, 51] or trajectory data [43, 52]. Also, data information can be leveraged to improve existing models or design new ones, as proposed by [13, 15, 16, 17, 27, 39]. Nevertheless, up to our knowledge few works have been devoted to evaluate the inherent uncertainty of both models and data and its impact on model-based predictions [3, 20, 28]. Yet, this is a fundamental aspect to improve decision and control strategies based on mathematical models. To this end, we propose here to follow a *Bayesian approach*, which allows to evaluate the parameter probability distribution given the observed data [32].

Statistical calibration techniques have been recently used for pedestrian models [9, 22, 23, 24] against microscopic trajectory data. In this paper, we only consider aggregate measurements provided by detectors at fixed locations. Without the ability to closely follow the microscopic dynamics, we are forced to use non-intrusive methods, combined with comparatively coarser grained data. Consequently, following Kennedy-O'Hagan [33], we introduce a bias term to better account for possible discrepancies between the mathematical models and reality, which also need to be estimated. This generic framework has been applied in a variety of fields, e.g., ranging from physics [29] to engineering [2, 31] or biology [44]. See also, e.g., [5] for a recent review of the methods and [4] for a discussion on the model discrepancy.

The article is organized as follows. In Section 2, we introduce the class of Generic Second Order Models, which will be the object of the study. Section 3 describe the data sets considered for model calibration, whose statistical calibration handling is presented in Section 4. The calibration results are discussed in Section 5 and some perspectives are drawn in Section 6.

2 Macroscopic traffic flow models

Generic Second Order traffic flow Models (GSOM in short) were introduced in [34] and consist in 2×2 hyperbolic systems of the form

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0, \\ \partial_t w + v \partial_x w = 0, \end{cases} \qquad t > 0, \;\; x \in \mathbb{R}, \tag{2.1}$$

where the average speed of vehicles is a function of the density $\rho = \rho(x,t)$ and a Lagrangian vehicle property $w = w(x,t)$, namely $v = \mathcal{V}(\rho, w)$ for some speed function $\mathcal{V}$ satisfying the following hypotheses [16]:

$$\mathcal{V}(\rho, w) \geq 0, \quad \mathcal{V}(0, w) = w, \quad \mathcal{V}(\rho, 0) = 0, \tag{2.2a}$$

$$\frac{\partial^2 Q}{\partial \rho^2}(\rho, w) < 0 \text{ for } w > 0, \quad \text{where } Q(\rho, w) := \rho \mathcal{V}(\rho, w) \tag{2.2b}$$

$$\frac{\partial \mathcal{V}}{\partial w}(\rho, w) > 0, \tag{2.2c}$$

$$\forall w > 0 \quad \exists\, R_w > 0 : \quad \mathcal{V}(R_w, w) = 0. \tag{2.2d}$$

As in [7, 16], we observe that (2.2b) implies $\dfrac{\partial \mathcal{V}}{\partial \rho}(\rho, w) < 0$ for $w > 0$, if $\mathcal{V}$ is a C^2 function in ρ. We also remark that in (2.2d) we can have $R_w = R$ for all $w > 0$.

Setting $y := \rho w$ for $\rho > 0$, system (2.1) can be rewritten in conservative form as

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0, \\ \partial_t y + \partial_x(y v) = 0, \end{cases} \qquad x \in \mathbb{R}, \ t > 0, \tag{2.3}$$

with $v = \widetilde{\mathcal{V}}(\rho, y) := \mathcal{V}\left(\rho, \dfrac{y}{\rho}\right)$.

System (2.1), respectively (2.3), is strictly hyperbolic for $\rho > 0$, with eigenvalues

$$\lambda_1(\rho, w) = \mathcal{V}(\rho, w) + \rho \mathcal{V}_\rho(\rho, w), \qquad \lambda_2(\rho, w) = \mathcal{V}(\rho, w), \tag{2.4}$$

with the first characteristic field being genuinely non-linear and the second linearly degenerate. The associated Riemann invariants are

$$z_1(\rho, w) = \mathcal{V}(\rho, w), \qquad z_2(\rho, w) = w.$$

Since shock and rarefaction curves of the first family coincide, the system belongs to the Temple class [49]. Notice that, setting $\mathcal{V}(\rho, w) = w - p(\rho)$ for a suitable "pressure" function p, system (2.3) corresponds to the celebrated Aw-Rascle-Zhang (ARZ) model [1, 53]. We also remark that taking $w = const$ we recover the classical Lighthill-Whitham-Richards (LWR) model [35, 46].

In the present setting, we are interested in the Initial Boundary Value Problem (IBVP) for (2.3), namely

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0, \\ \partial_t(\rho w) + \partial_x(\rho w v) = 0, \end{cases} \qquad x \in]x_{in}, x_{out}[\subset \mathbb{R}, \ t > 0, \tag{2.5a}$$

$$(\rho, w)(x, 0) = (\rho_0, w_0)(x), \qquad x \in]x_{in}, x_{out}[, \tag{2.5b}$$

$$(\rho, w)(x_{in}, t) = (\rho_{in}, w_{in})(t), \qquad t > 0, \tag{2.5c}$$

$$(\rho, w)(x_{out}, t) = (\rho_{out}, w_{out})(t), \qquad t > 0, \tag{2.5d}$$

with values in an invariant domain of the form

$$\Omega_{\mathcal{V}} := \Big\{ (\rho, w) \in \mathbb{R}^2_+ \colon w \in [w_{min}, w_{max}] \Big\}$$

for some $0 \le w_{min} \le w_{max} < +\infty$.
Solutions to (2.5) have to be intended in the weak sense as in [8, Definition 2.2]. In particular, we remark that boundary conditions (2.5c) and (2.5d) may not be satisfied in the strong sense, i.e., we may have

$$(\rho_{in}, w_{in})(t) \neq \lim_{x \to x_{in}+} (\rho, w)(x, t),$$
$$(\rho_{out}, w_{out})(t) \neq \lim_{x \to x_{out}-} (\rho, w)(x, t).$$

Note also that general IBVP well-posedness results provided in the literature, see e.g., [8, Theorem 2.3], hold under the hypothesis of strict hyperbolicity, which is not satisfied by (2.5a) at $\rho = 0$. The well-posedness of (2.5) will make the object of a separate study.

Remark 1. *In the implementation, we derive the initial and boundary data directly from real traffic data. Since the quantity w is not directly provided by the measured data, we need to invert the velocity function $\mathcal{V}$. This gives us a new function $\mathcal{W}$ which reads as*

$$\begin{aligned} \mathcal{W} : \Omega_{\mathcal{W}} &\to [w_{min}, w_{max}] \\ (\rho, v) &\mapsto \mathcal{W}(\rho, v) \end{aligned} \tag{2.6}$$

where the domain is defined as $\Omega_{\mathcal{W}} = \{(\rho, v) \colon \rho \in [0, R_{w_{max}}], v \in [\mathcal{V}(\rho, w_{min}), \mathcal{V}(\rho, w_{max})]\}$. We note that providing the initial data and boundary conditions for both the density and the empty-road velocity, the second order GSOM model is equipped with more real data compared to the first order LWR model.

In this paper, we consider a speed function of the form

$$\mathcal{V}(\rho, w) = w \left(1 - \exp\left(\frac{C}{V} \left(1 - \frac{R}{\rho} \right) \right) \right), \tag{2.7}$$

which is derived from Newell-Franklin [19, 40] and satisfies (2.2). In (2.7), the parameters to be identified are $\theta = (V, C, R, w_{min}, w_{max})$, with $V > 0$ is the maximal speed, $R > 0$ is the maximal density, $C > 0$ is the wave propagation speed in congestion, w_{min} and w_{max} are respectively the minimum and maximum vehicle property. In this setting, the variable w plays the role of an *empty-road velocity* of the driver [16]. Notice that, fixing $w = V$ in (2.7), we recover the first order LWR model. In this case, the calibration parameter set reduces to $\theta = (V, C, R)$.

2.1 Numerical solution

To efficiently compute approximate solutions of (2.5), we use the Harten-Lax-van Leer (HLL) Riemann solver [26]. Given a (possibly non-uniform) spatial discretization $\{x_0, \dots, x_M\}$ of the interval $]x_{in}, x_{out}[$ (with $x_0 = x_{in}$ and $x_M = x_{out}$), we set the cell sizes $\Delta x_j := x_j - x_{j-1}$ for $j = 1, \dots, M$, and a suitable time step Δt satisfying the Courant-Friedrichs-Lewy (CFL) stability condition

$$\Delta t \leq \frac{\min_j \Delta x_j}{\max_{(\rho,w)\in\Omega_{\mathcal{V}}} \{ |\lambda_1(\rho, w)|, |\lambda_2(\rho, w)| \}}. \tag{2.8}$$

Denoting by

$$U := \begin{pmatrix} \rho \\ \rho w \end{pmatrix} \qquad \text{and} \qquad F(U) := \begin{pmatrix} \rho v \\ \rho w v \end{pmatrix}$$

the vectors of the conserved quantities and fluxes respectively, we approximate the initial data (2.5b) as

$$\mathbf{U}_j^0 := \frac{1}{\Delta x_j} \int_{x_{j-1}}^{x_j} U^0(y)\, dy.$$

Approximate solutions are then computed iteratively by the finite volume formula

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \frac{\Delta t}{\Delta x_j} \left(\mathbf{F}_j^n - \mathbf{F}_{j-1}^n \right), \qquad j = 1, \dots, M, \tag{2.9}$$

where the numerical fluxes $\mathbf{F}_j^n := F_{HLL}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n)$ are computed using the HLL Riemann solver:

$$F_{HLL}(\mathbf{U}_L, \mathbf{U}_R) := \begin{cases} F(U_L) & \text{if } S_L \geq 0, \\ F(U_{HLL}) & \text{if } S_L < 0 \leq S_R, \\ F(U_R) & \text{if } S_R < 0, \end{cases} \tag{2.10}$$

where $S_L < S_R$ are the approximate wave speeds. Following [12, 16], we employ the following definition of characteristic speeds

$$\begin{aligned} S_L &:= \min\{\lambda_1(\mathbf{U}_L), \lambda_1(\mathbf{U}_R)\}, \\ S_R &:= \min\{\lambda_2(\mathbf{U}_L), \lambda_2(\mathbf{U}_R)\}, \end{aligned} \tag{2.11}$$

and we set

$$F(U_{HLL}) := \frac{S_R F(\mathbf{U}_R) - S_L F(\mathbf{U}_L) - \big(F(\mathbf{U}_R) - F(\mathbf{U}_L)\big)}{S_R - S_L}.$$

Note that, since $\lambda_2(\rho, w) = v \geq 0$ for all $(\rho, w) \in \Omega_{\mathcal{V}}$, the case $S_R < 0$ in (2.10) never occurs.

Boundary conditions (2.5c) and (2.5d) are taken into account by (2.9) setting

$$\begin{aligned} \mathbf{U}_0^n &:= \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} U_{in}(s)\, ds, \\ \mathbf{U}_{M+1}^n &:= \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} U_{out}(s)\, ds, \end{aligned}$$

where $t^n = n\Delta t$ and $n = 1, \ldots, N$.

When considering the scalar LWR equation, we use the classical Godunov scheme [21] in its supply-demand implementation [11]. We note that for the first order LWR model the boundary conditions are given in terms of flows F_{in} and F_{out}, instead of densities. Thus, the IBVP (2.5) reads

$$\begin{aligned} &\partial_t \rho + \partial_x(\rho v) = 0, && x \in]x_{in}, x_{out}[,\ t > 0, \\ &\rho(x, 0) = (\rho_0)(x), && x \in]x_{in}, x_{out}[, \\ &q(x_{in}, t) = F_{in}(t), && t > 0, \\ &q(x_{out}, t) = F_{out}(t), && t > 0. \end{aligned}$$

Remark 2. *Since initial and boundary conditions provided by data can lie outside the domain of the inverse mapping $\mathcal{W}$, we will perform a projection following [16, Section 3.3]:*
Given a density $0 \leq \rho \leq R_{w_{max}}$ and any speed $v > 0$ (can be outside of the domain $\Omega_{\mathcal{W}}$) we consider

$$\tilde{\mathcal{V}}(\rho, v) = \min\Big\{\max\{v, \mathcal{V}(\rho, w_{min})\}, \mathcal{V}(\rho, w_{max})\Big\} \tag{2.12}$$

where $\mathcal{V}(\rho, w_{min})$ (resp. $\mathcal{V}(\rho, w_{max})$) is the lower (resp. upper) bound of the speed curves. This allows to define $\tilde{D}_{\mathcal{W}} := \{(\rho, v) \colon \rho \in [0, R_{w_{max}}], v \geq 0\}$.

3 Data and simulation setting

In this section, we describe the two sensor data sets that will be used for the calibration tests in Section 5.

One data set was provided by the *Direction Interdépartementale des Routes Méditerranée* (DIRMED) and covers 3 months, from September 1st to November 30th, 2015. The considered set of data (DirMed data) comes from 135 loop detectors placed on the freeways in the North of Marseille (A7, A51, A50 and A55). Given the 6 minute averages of the traffic flow and speed measured by the loop detectors, the traffic density is a derived quantity.
In this paper, we focus on a 1.25 km long rampless road stretch on the highway A50 from Aubagne to Marseille. The road consists of three lanes and the speed limit is 90 km/hour. The data for this highway part were collected by 4 loop detectors (IDs 305, 304, 303 and 302). For the simulation, we divide the chosen road stretch into 4 segments. Each segment contains one detector. All the details about the considered traffic data and the loop detector locations are available in the git repository `https://gitlab.inria.fr/acumes/dduq-traffic`.
The fundamental diagrams in Figure 1 illustrate the collected flow and speed data together with their corresponding densities for the inner loop 303. Most of the data points are located in the free flow phase which is visible by the clear functional relationship between density and flow or speed. The free flow phase corresponds roughly to speed values greater than 60 km/h. The more widely distributed data points in the higher density region correspond to congested regimes.

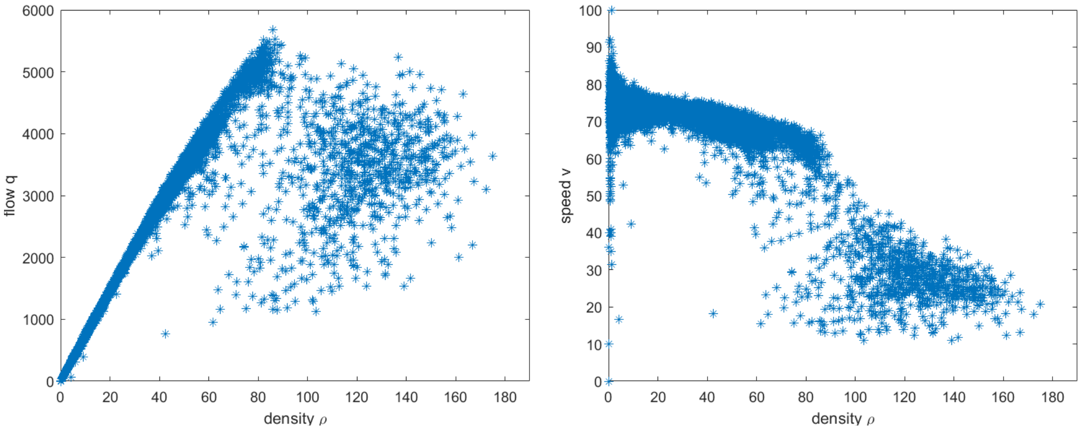


Figure 1: Fundamental flow and speed diagrams for DirMed data corresponding to loop 303. Data from 09/01 to 11/30/2015.

For the calibration and validation in Section 5, we will focus on Tuesday 11/10 for a five hours time slot between 7am and 12pm. The density evolution is illustrated in Figure 3a.

The other data set is referred to as the RTMC data set [38], which is provided by the Minnesota Department of Transportation (MnDOT). Analogously to the DirMed data set, the MnDOT data are 6 minute averages obtained by single loop detectors measuring the traffic flow and the occupancy (fraction of time that a vehicle is occupying a detector) along several (interstate) highways (while in the DirMed data set, speeds were measured instead of the occupancy). For the tests, we divide a 1.1 km long road stretch on the northbound

direction of the interstate highway I-35W again into 4 segments. Each of them contains one loop detector ($S60, S61, S1708, S62$). The rampless road stretch has five lanes and the speed limit is 55 miles/hour ($\approx$ 90 km/hour).
The fundamental flow and speed diagram for the inner loop $S1708$ is depicted in Figure 2. In contrast to Figure 1, the congested area is equipped with more data points, which are again widely spread. However, the functional relationship in the free flow region is still visible, especially in the flow diagram. We emphasize that both the maximum density and the maximum flow value are higher than in the DirMed data case, since the road has five lanes (instead of three).

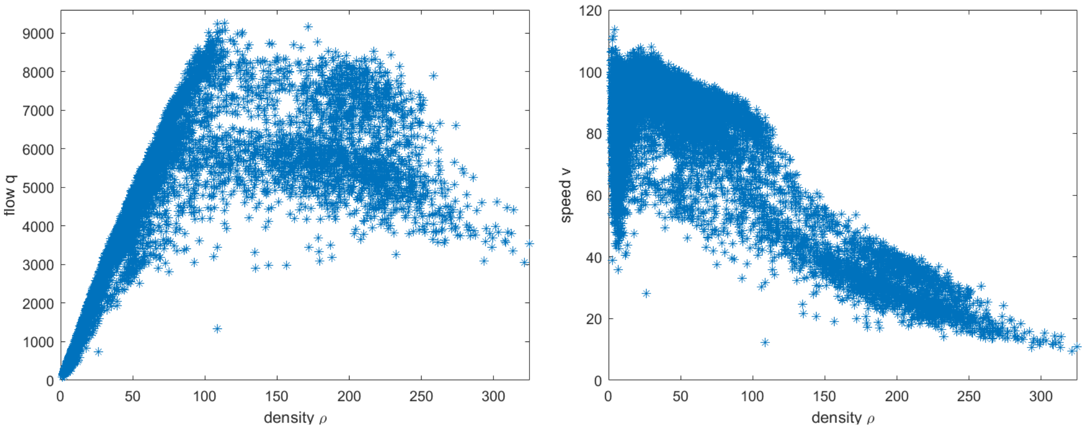
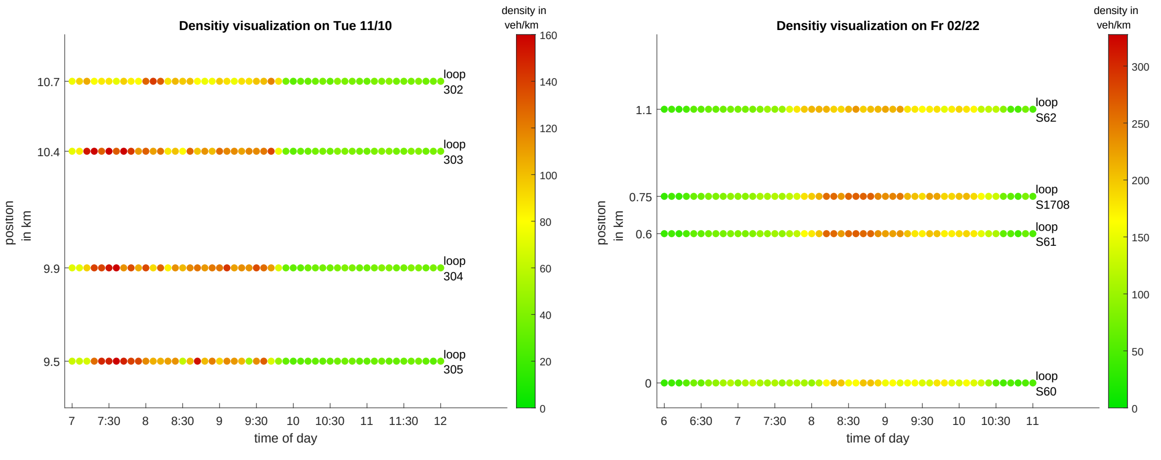


Figure 2: Fundamental flow and speed diagrams for RTMC data corresponding to loop $S1708$. Data from weekdays from 01/01 to 04/12/2013.

In our analysis, we will consider the data on Friday 02/22/2013 in a five hours time slot between 6am and 11am. The density evolution is illustrated in Figure 3b. Figure 3 emphasizes that we take both congestion (orange and red colored densities) and free flow (green colored densities) phases into account. The rush hour for the DirMed data case takes place at the



(a) DirMed data

(b) RTMC data

Figure 3: Density visualization for two data scenarios in a 5 hours time slot.

beginning of the considered time period, between 7:30am and 9:30am, whereas the congested phase for the RTMC data starts after two hours, at 8:00am.

Remark 3. *In order to reduce the influence of the piecewise constant initial data, we run the traffic model through an initialization phase of 6 minutes (see [16]). Thus, the validation of the data will be finally done on a 4 hours 54 minutes time slot.*

4 Calibration approaches

After presenting the mathematical model in Section 2 and describing the data sets in Section 3, we now present our calibration procedure to fuse the two. Instead of a least-squares approach on a family of flow rate curves, as generally performed in this context, see, e.g., [13, 16] , we follow the statistical framework proposed by [29, 33]. The benefits are two-fold. First, when relying on a Bayesian estimation procedure, uncertainty quantification is directly available in the form of a posterior distribution on the calibration parameters, rather than scalar values. Looking at the fundamental diagrams of Figures 1 and 2, it is clear that a single parameter θ value is insufficient to properly fit the cloud of points. Second, accounting for and estimating a bias term between the mathematical model and data enables alleviating the shortcomings of the former.

4.1 Kennedy-O'Hagan calibration framework

We denote by P the real process under study (e.g., the space-time evolution of the density ρ), F the so-called "field" where P is physically observed, both depending on x and t. Denote by M the mathematical model, which depends on (x, t, θ), with θ the additional calibration parameter(s), and by (X, T) all the (x, t) points where observations have been recorded. Denote by $y^F(x, t)$ the field observations under conditions x, t of the real output $y^P(x, t)$.

It is generally assumed that P and F are related by

$$y^F(x,t) = y^P(x,t) + \varepsilon, \text{ where } \epsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2).$$

Kennedy-O'Hagan (KOH) [33] proposed to additionally take into account inadequacy between the mathematical model with optimal parameters θ^* and reality, via an additional discrepancy (or bias) term $b(x, t, \theta)$: $y^P(x,t) = y^M(x,t,\theta^*) + b(x,t,\theta^*)$ resulting in

$$y^F(x,t) = y^M(x,t,\theta^*) + b(x,t,\theta^*) + \epsilon.$$

In [33], the authors further proposed a Bayesian framework to estimate the best calibration parameter θ^*, along with σ_ϵ and $b(\cdot,\cdot,\cdot)$. The outcome is a posterior probability distribution instead of a single optimal value. In KOH, b is modeled by a Gaussian process (GP), which amounts to assume a multivariate normal distribution for the errors. KOH also models y^M with a Gaussian process as y^M is computationally expensive in their setup. This is not necessary here since the mathematical model takes only a few seconds to evaluate, thus corresponding to the framework described in [29].

To follow [29], besides defining priors on the calibration parameters, one needs to be able to estimate the posterior. Bayes' rule expresses it as

$$\pi(\theta^* \mid y^F) = \frac{\mathcal{L}(y^F \mid \theta^*) \times \pi(\theta^*)}{\pi(y^F)},$$

which says that the posterior is given by the product of the likelihood and the prior, divided by the marginal likelihood. The marginal likelihood (a.k.a. evidence) is a normalizing constant that does not depend on θ. Analytical expressions of posterior distributions are seldom available, leading to a variety of estimation techniques. In this paper, we use a standard Markov Chain Monte Carlo (MCMC) method, the Metropolis algorithm, see Algorithm 1. More details can be found in [29]. Independent samples are obtained by thinning out the output chain, that is, keeping only one out of p samples, to avoid autocorrelation (see, e.g., [30]).

Algorithm 1 Metropolis algorithm with a symmetric proposal distribution

```
initialize $\theta_1^*$ (sample from the prior distribution $\pi(\theta^*)$)
for each $i$ in $\{2, ...N\}$ do
    generate $\hat{\theta}^*$ from a symmetric distribution, e.g., $\hat{\theta}^* \sim \mathcal{N}(\theta_{i-1}^*, \sigma_p^2 \mathbf{I})$
    compute the ratio $\alpha := \min\left\{1, \frac{\pi(\hat{\theta}^*|y^F)}{\pi(\theta_{i-1}^*|y^F)}\right\}$
    generate a uniform random number $u \sim \mathcal{U}([0,1])$
    if $u \leq \alpha$ then
        $\theta_i^* = \hat{\theta}^*$
    else
        $\theta_i^* = \theta_{i-1}^*$
    end if
end for
Return $\theta^* = (\theta_1^*, \dots, \theta_N^*)$
```

4.2 Gaussian process regression

We rely on Gaussian process regression for estimating the bias term [29, 33]. Given a set of observations $\mathbf{b}_n = (b(x^{(1)}), \dots, b(x^{(n)}))$ of the bias at $(x^{(1)}, \dots, x^{(n)})$ (dropping t and θ^* for notational simplicity), this GP assumption amounts to consider that $\mathbf{b}_n$ is a realization of a (zero-mean) multivariate normal distribution: $\mathbf{b}_n \sim \mathcal{N}(\mathbf{0}_n, \mathbf{K}_n)$. Here, $\mathbf{K}_n$ denotes the covariance matrix between the observed biases, that is, $\mathbb{C}ov\left[b(x), b(x')\right] = k(x, x')$ where $k(\cdot, \cdot)$ is a positive definite function, typically from a parametric family, such as the Gaussian kernel: $k_G(x, x') = \sigma^2 \exp(-(x - x')^2/l^2)$, among others like Matérn kernels, see e.g., [47]. The (hyper-)parameters σ^2 and l are, respectively, the process variance and the length-scale. Extension to the multivariate case is generally obtained via a product of univariate covariance kernels.

GP predictive equations of $b(\cdot)$ at n' new locations $\mathcal{X}$, denoted $Y(\mathcal{X}) \mid \mathbf{b}_n$, are the so-called “kriging” equations:

$$Y(\mathcal{X}) \mid \mathbf{b}_n \sim \mathcal{N}(m_n(\mathcal{X}), s_n^2(\mathcal{X}, \mathcal{X})), \text{with} \tag{4.1}$$

$$m_n(\mathcal{X}) := \mathbb{E}[Y(\mathcal{X})|\mathbf{b}_n] = \mathbf{k}_n(\mathcal{X})^\top \mathbf{K}_n^{-1} \mathbf{b}_n, \tag{4.2}$$

$$s_n^2(\mathcal{X}, \mathcal{X}) := \mathbb{C}ov[Y(\mathcal{X}), Y(\mathcal{X})|\mathbf{b}_n] = k(\mathcal{X}, \mathcal{X}) - \mathbf{k}_n(\mathcal{X})^\top \mathbf{K}_n^{-1} \mathbf{k}_n(\mathcal{X}). \tag{4.3}$$

where $\mathbf{k}_n(\mathcal{X}) := (k(\mathcal{X}^{(j)}, x^{(i)}))_{1 \leq j \leq n', 1 \leq i \leq n}$ and the predictive variance is the diagonal of $s_n^2(\mathcal{X}, \mathcal{X})$. Equations (4.1) describe the best (minimizing Mean Square Prediction Error) linear unbiased predictor (BLUP).

To estimate the hyperparameters of the kernel function, we maximize the likelihood, the probability density of the observations given the parameters: $\mathbb{P}[Y(\mathbf{X}_n) = \mathbf{b}_n]$. As $\mathbf{b}_n \sim \mathcal{N}(0, \mathbf{K}_n)$, the likelihood $\mathcal{L}$ is given by the MVN density. Taking the logarithm, this gives:

$$\log \mathcal{L}(l_1, \dots, l_d, \sigma^2) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\mathbf{K}_n| - \frac{1}{2} \mathbf{b}_n^\top \mathbf{K}_n^{-1} \mathbf{b}_n.$$

Rewriting $\mathbf{K}_n = \sigma^2(\mathbf{C}_n + g\mathbf{I}_n)$ with $\mathbf{C}_n$ the correlation matrix (foreshadowing the need to account for noise in observations with $g \geq 0$), we can compute the optimal variance $\hat{\sigma}^2$ differentiating the resulting expression

$$\log \mathcal{L}(l_1, \dots, l_d, \sigma^2) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2} \log |\mathbf{C}_n + g\mathbf{I}_n| - \frac{1}{2\sigma^2} \mathbf{b}_n^\top (\mathbf{C}_n + g\mathbf{I}_n)^{-1} \mathbf{b}_n,$$

so that

$$\frac{\partial \log \mathcal{L}}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \mathbf{b}_n^\top (\mathbf{C}_n + g\mathbf{I}_n)^{-1} \mathbf{b}_n = 0$$

which gives $\hat{\sigma}^2 = \frac{\mathbf{b}_n^\top (\mathbf{C}_n + g\mathbf{I}_n)^{-1} \mathbf{b}_n}{n}$.
Plugging in this $\hat{\sigma}^2$ in the likelihood, we obtain the concentrated likelihood $\tilde{\mathcal{L}}$:

$$\log \tilde{\mathcal{L}}(l_1, \dots, l_d, g) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \hat{\sigma}^2 - \frac{1}{2} \log |\mathbf{C}_n + g\mathbf{I}_n| - \frac{n}{2}. \tag{4.4}$$

Solving for the remaining hyperparameters is not possible in closed-form, therefore numerical optimization methods must be used (usually relying on the gradient of the concentrated log-likelihood). We refer to [25, 45] for further details on GP modeling.

Bayesian calibration via MCMC estimation can be time consuming. Several simplifications have been proposed, such as relying on optimization and on modularization, as in e.g., [36, 25]. Without bias estimation, the resulting estimation procedure reverts to the least squares approach. Due to the grid structure of the data (regular measurement at the same loop locations), the computational cost of fitting GPs (in $\mathcal{O}(n^3)$) can be reduced by exploiting the resulting Kronecker structure (see, e.g., [10] for a detailed application example).

5 Validation and comparison

In this section, we present our calibration results for the two data sets introduced in Section 3. Our quantity of interest (y) will be the flow observations since we expect flow measurements be more precise compared to the speed or density data.

5.1 Direct approach

We start with considering a calibration approach leading to a single optimized value instead of a posterior probability distribution, which will be referred to as the *direct approach*. It consists in the maximization of the concentrated log-likelihood function (4.4). Since the observed bias

$$b(x, t, \theta) = y^F(x, t) - y^M(x, t, \theta) \qquad \text{for all } (x, t)\text{-combinations in } (X, T)$$

depends on the unknown calibration parameter θ, the process variance $\hat{\sigma}^2$ in (4.4) depends also on the calibration parameter θ. This results in solving the optimization problem

$$\max_{l_1, l_2, g, V, C, R, w_{min}, w_{max}} \log \tilde{\mathcal{L}}(l_1, l_2, g) \tag{5.1}$$

for 8 parameters. Following the principle of modularization [36], we solve this expanded optimization problem by a 2-level approach leading to solve two nested maximization problems. In the so-called *inner-level*, we maximize the concentrated log-likelihood function dependent on θ, thus obtaining the hyperparameters $\hat{l}_1(\theta), \hat{l}_2(\theta), \hat{g}(\theta)$. These hyperparameters are then inserted into (4.4) and the concentrated likelihood is maximized with respect to θ, giving the optimized calibration parameter θ^*. We call the second step the *higher-level*. The inner optimization level is executed by the local optimization solver `fmincon`. For the higher-level we compare both `fmincon` and `pso` from MATLAB and we choose the one which leads to a higher value for the likelihood function. Since our quantity of interest are the flow observations, the optimization is done on the flow errors. The initial guesses and boundary intervals are listed in Table 1.

		inner-level			higher-level				
		l_1	l_2	g	V	C	R	w_{min}	w_{max}
DirMed	initial guess	1	0.52	3.3	90	30	350	40	115
	lower bound	0.1	0.1	0.005	70	10	180	1	70
	upper bound	5	1.2	5	120	70	450	80	300
RTMC	initial guess	1	0.52	3.3	90	30	500	40	115
	lower bound	0.1	0.1	0.005	70	10	250	1	70
	upper bound	5	1.1	5	120	70	750	80	300

Table 1: Initial guesses, upper and lower bounds for the direct approach.

The optimization results for the calibration parameters are summarized in Table 2. We observe that the two traffic flow models (LWR and GSOM) lead to different optimal calibration parameters although we use mostly the same initial guesses and bounds. Nevertheless, the results of the non-convex optimization problem depend crucially on the initial guess. However, there are two aspects to emphasize: the speed parameters obtained by the GSOM model are closer to the maximum speed level (90 km/h) and none of the parameters lie on the boundary in the second order model (unlike the maximum density of 450 (DirMed data) and the maximum speed close to 70 (RTMC data) for the LWR model). We note that, in both traffic flow models, the maximum density for the DirMed data is smaller since there are only three lanes instead of five.

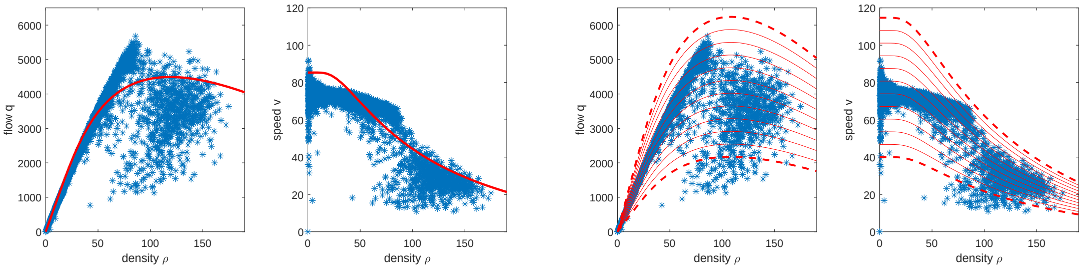
	LWR			GSOM				
	V^*	C^*	R^*	V^*	C^*	R^*	w^*_{min}	w^*_{max}
DirMed	85.34	17.93	450.00	89.88	28.11	349.97	40.00	114.70
RTMC	70.12	29.53	463.99	92.71	29.63	488.23	40.00	113.19

Table 2: Optimization results for the direct approach.

Remark 4. *In contrast to [15, 16], we do not set a-priori the value of the maximum density R. Here, it is considered as one of the parameters to calibrate.*

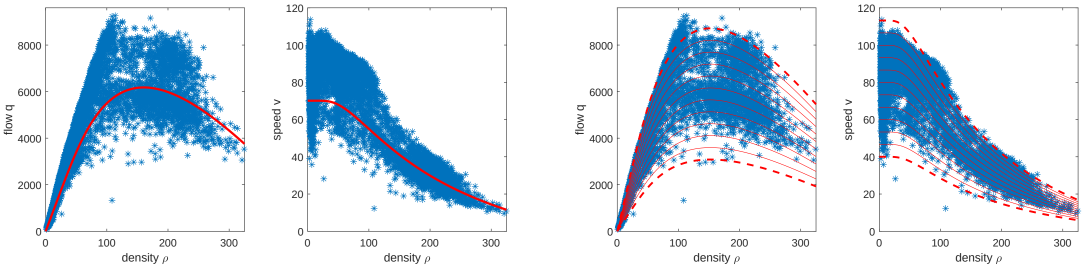
The fundamental diagrams in Figures 4 and 5 report the same data as Figures 1 and 2, where we superpose the flow and speed curves obtained corresponding to the optimal calibration parameters from Table 2 for the function (2.7). Figure 4b and Figure 5b show that the GSOM model can capture the spread of the data in the congested part, unlike the LWR model. In both cases, the fit looks better for RTMC data.

Figure 6 depicts the flow and speed profiles for the DirMed scenario and Figure 7 shows the flow and density profiles for the RTMC scenario. We observe that for the GSOM model the numerical solution (blue squares) follows the profile of the measured speed and density data (red stars) better than the flow values. Even jumps and drop regions are well captured,



(a) LWR model (b) GSOM model

Figure 4: Fundamental flow and speed diagrams for DirMed data corresponding to loop 303. In red, equilibrium curve for LWR (Figure 4a) and family of flow and speed curves for GSOM (Figure 4b), based on the optimal calibration parameter of Table 2.



(a) LWR model (b) GSOM model

Figure 5: Fundamental flow and speed diagrams for RTMC data corresponding to loop $S1708$. In red, equilibrium curve for LWR (Figure 5a) and family of flow and speed curves for GSOM (Figure 5b) based on the optimal calibration parameter of Table 2.

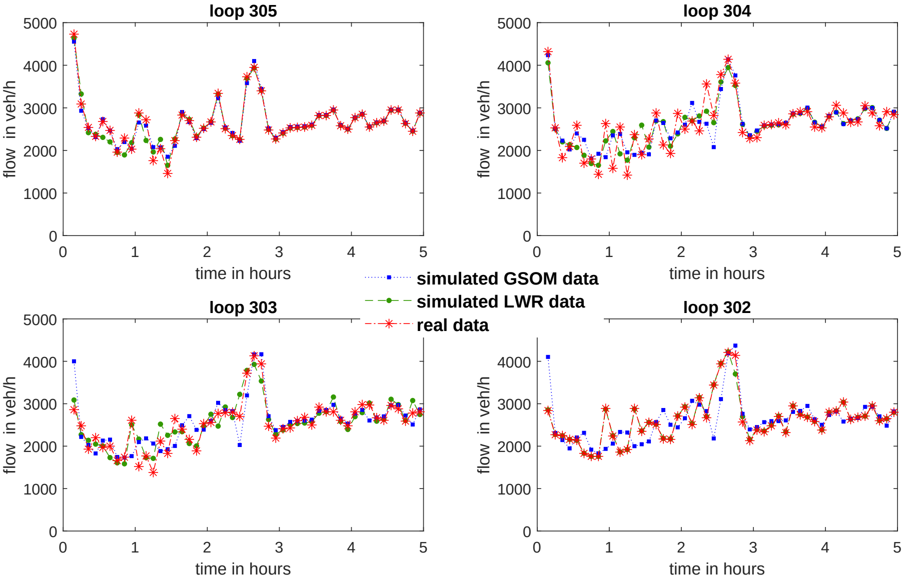
though being more difficult to predict. At this point, we want to emphasize that although the optimization is done only on the flow errors, the GSOM is able to recover the density and speed quantities much better than the LWR model. However, by comparing the flow profiles of the two traffic flow models, we observe a better performance for the LWR model. The simulated GSOM flow data have more outliers, especially for the DirMed data case on loop 302.

In order to compare the predictive accuracy of the results between the two traffic flow models numerically, we define an error metric $\mathbf{E}$ which consists of the sum of normalized flow, speed and density errors. Denoting y^F_{flow} (resp. y^F_{speed}, y^F_{density}) the measured flow (resp. speed, density) data and y^M_{flow} (resp. y^M_{speed}, y^M_{density}) as the simulated flow (resp. speed, density) data, including the correction by their predictive kriging means (4.2), the error function is defined as the sum of

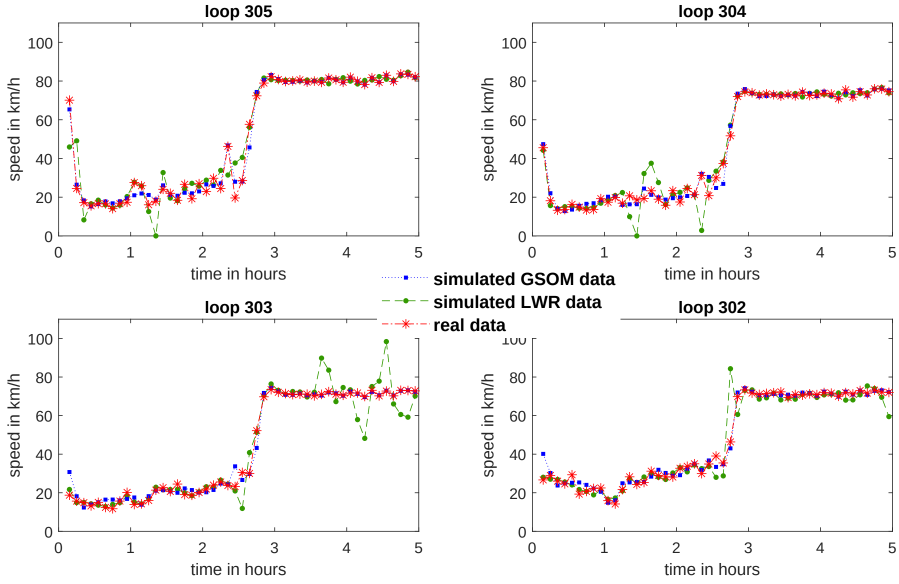
$$E^k = \frac{1}{T_f \cdot (x_{out} - x_{in})} \frac{1}{\Delta q} \sum_{(x,t)\in(X,T)} |y^F_k(x,t) - y^M_k(x,t,\theta^*)|,$$

for $k \in \{\text{flow}, \text{speed}, \text{density}\}$. Thus

$$\mathbf{E} = E^{\text{flow}} + E^{\text{speed}} + E^{\text{density}}. \tag{5.2}$$



(a) flow profile



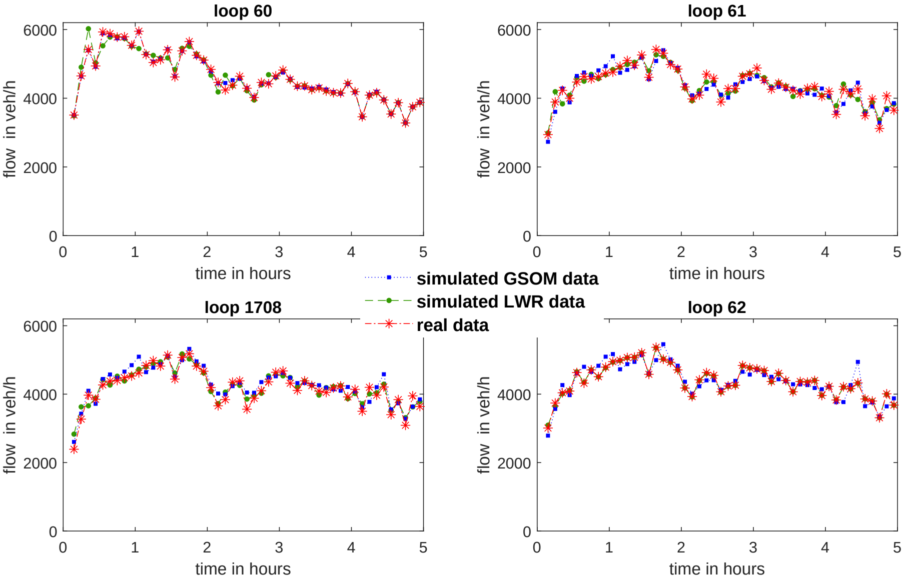
(b) speed profile

Figure 6: Comparison of real and simulated flow and speed profiles for LWR and GSOM models: direct approach for the DirMed data.

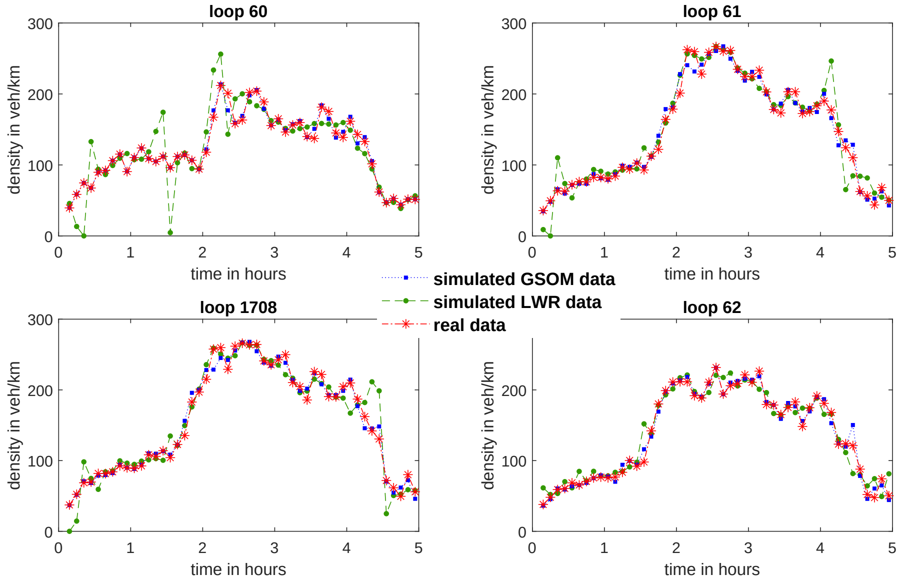
Above, T_f stands for the time horizon (4 hours 54 minutes) and Δq, (resp. Δv, $\Delta\rho$) represents normalization constants to overcome the biases induced by the different physical units. For the choice of the normalization constants, we follow [15] and define Δq (resp. Δv, $\Delta\rho$) as the ranges of the real measured traffic flow (resp. speed, density) data. Thus, we obtain

$$\Delta q = Q^{\max} - Q^{\min},\ \Delta v = V^{\max} - V^{\min},\ \Delta\rho = \mathrm{R}^{\max} - \mathrm{R}^{\min}$$

where $Q^{\max}$ and $Q^{\min}$ (resp. $V^{\max}$ and $V^{\min}$, $\mathrm{R}^{\max}$ and $\mathrm{R}^{\min}$) stands for the maximum and minimum observed traffic flow (resp. speed, density) in the considered data set.



(a) flow profile



(b) density profile

Figure 7: Comparison of real and simulated flow and density profiles for LWR and GSOM models: direct approach for the RTMC data.

Table 3 shows the error values for both data sets and both traffic flow models. The GSOM

	LWR				GSOM			
	E^{flow}	E^{speed}	E^{density}	**E**	E^{flow}	E^{speed}	E^{density}	**E**
DirMed	**1.1457**	1.5533	4.0071	6.7061 (+35%)	1.9889	**1.2459**	**1.7348**	**4.9696**
RTMC	**0.8974**	1.7818	2.3153	4.9944 (+98%)	1.2170	**1.0761**	**0.0868**	**2.3799**

Table 3: Error results for the direct approach. In bold, the lowest flow, speed, density and total errors per data scenario.

model has the smallest total error **E** for both data scenarios, the error in the LWR model being at least 35% higher. As we already observed in Figures 6 and 7, the main error contributions for the LWR model come from the speed and density errors. These values are strikingly higher than the corresponding errors from the GSOM model, which in turn performs worst on the flow variable. We refer to Section 5.3 for further investigation on this aspect.

Appendix A presents the results of a more intuitive approach to determine the optimal calibration parameter. However, the direct approach outperforms this rather naive approach since every single error quantity in Table 3 is lower than the corresponding one from Table 9. This supports our decision to equip the model with a discrepancy term.

5.2 Bayesian approach

The main aim of this work is to apply the the Bayesian approach introduced in Section 4 to obtain posterior probability distributions of the model parameters. In the following, we introduce the formulas used in the Metropolis Algorithm 1.

- The prior distribution $\pi(\theta_{\mathrm{VCR}})$ for the calibration parameters $\theta_{\mathrm{VCR}} = (V, C, R)$ is given by a multivariate normal distribution, i.e.,

  $$\pi(\theta_{\mathrm{VCR}}) \propto \frac{1}{\sqrt{|\Sigma_{\theta_{\mathrm{VCR}}}|}} \exp\Big\{-0.5 \left(\theta_{\mathrm{VCR}} - \mu_{\theta_{\mathrm{VCR}}}\right)^\top \Sigma^{-1}_{\theta_{\mathrm{VCR}}} \left(\theta_{\mathrm{VCR}} - \mu_{\theta_{\mathrm{VCR}}}\right)\Big\}$$

  with mean $\mu_{\theta_{\mathrm{VCR}}}$ and covariance matrix $\Sigma_{\theta_{\mathrm{VCR}}}$.

- For the parameter $w_m = (w_{\min}, w_{\max})$, we use a uniform prior, i.e.,

  $$w_{min} \sim \mathcal{U}([a_{w_{min}}, b_{w_{min}}]), \qquad w_{max} \sim \mathcal{U}([a_{w_{max}}, b_{w_{max}}]).$$

  Since we assume the parameters of the prior to be independent, the prior joint probability density function $\pi(w_m)$ is finally given by the product of the univariate functions, i.e.,

  $$\pi(w_m) = \frac{1}{b_{w_{min}} - a_{w_{min}}} \cdot \mathbb{1}_{[a_{w_{min}}, b_{w_{min}}]}(w_{min}) \cdot \frac{1}{b_{w_{max}} - a_{w_{max}}} \cdot \mathbb{1}_{[a_{w_{max}}, b_{w_{max}}]}(w_{max}).$$

- The proposal distribution for $\hat{\theta}_{\mathrm{VCR}}$ is defined as a multivariate normal distribution, i.e.,

  $$\hat{\theta}_{\mathrm{VCR}} \sim \mathcal{N}(\theta_{\mathrm{VCR}_{i-1}}, \Sigma^p_{\mathrm{VCR}})$$

  with covariance matrix Σ^p_{VCR}.

- The proposal for the calibration parameter $\hat{w}_m$ is given by a multivariate normal distribution specified by

  $$\hat{w}_m \sim \mathcal{N}(w_{m_{i-1}}, \Sigma^p_m) \text{ with covariance matrix } \Sigma^p_m.$$

- The sampling model for y^F reads as

  $$\mathcal{L}(y^F \mid \theta^*) \propto \frac{1}{\sqrt{|\mathbf{K}_n|}} \exp\Big\{-0.5 \left(b(\theta^*)^\top \mathbf{K}_n^{-1} b(\theta^*)\right)\Big\}.$$

  Inserting the process variance $\hat{\sigma}^2$ from the covariance matrix $\mathbf{K}_n$ appearing in the exponential yields the following simplification:

  $$\mathcal{L}(y^F \mid \theta^*) \propto |\mathbf{K}_n|^{-1/2}$$

Thus, the posterior distribution $\pi(\theta^* \mid y^F)$ is computed by means of Bayes Theorem, i.e.,

$$\pi(\theta^* \mid y^F) \propto L\left(y^F \mid \theta^*\right) \times \pi(\theta_{\text{VCR}}) \times \pi(w_m).$$

The choice of the mean vectors, covariance matrices and bounds for the uniform distributions can be found in Table 4.

	$\mu_{\theta_{\text{VCR}}}$	$\text{diag}(\Sigma_{\theta_{\text{VCR}}})$	$\text{diag}(\Sigma^p_{\text{VCR}})$		$[a_{w_{min}}, b_{w_{min}}]$	$[a_{w_{max}}, b_{w_{max}}]$	$\text{diag}(\Sigma^p_m)$
			LWR	GSOM			
DirMed	$(90, 30, 350)$	$(100, 100, 2500)$	$(6, 1.5, 90)$	$(6, 3, 100)$	$(10, 60)$	$(70, 200)$	$(2, 2)$
RTMC	$(90, 40, 450)$	$(56.3, 225, 5652)$	$(8, 4, 100)$	$(4, 2, 80)$	$(10, 60)$	$(70, 200)$	$(2, 2)$

Table 4: Prior and proposal specifications applied in the Metropolis Algorithm 1 for the two data sets. The choices do not differ between the LWR and GSOM model except for the proposal variance Σ^p_{VCR}.

Remark 5. *We note that the choice of the bounds for the uniform distributions (see column 5 and 6 in Table 4) on the parameters w_{min} and w_{max} can be extended to a larger range. However, preliminary tests did no show a significant difference in the outcome.*

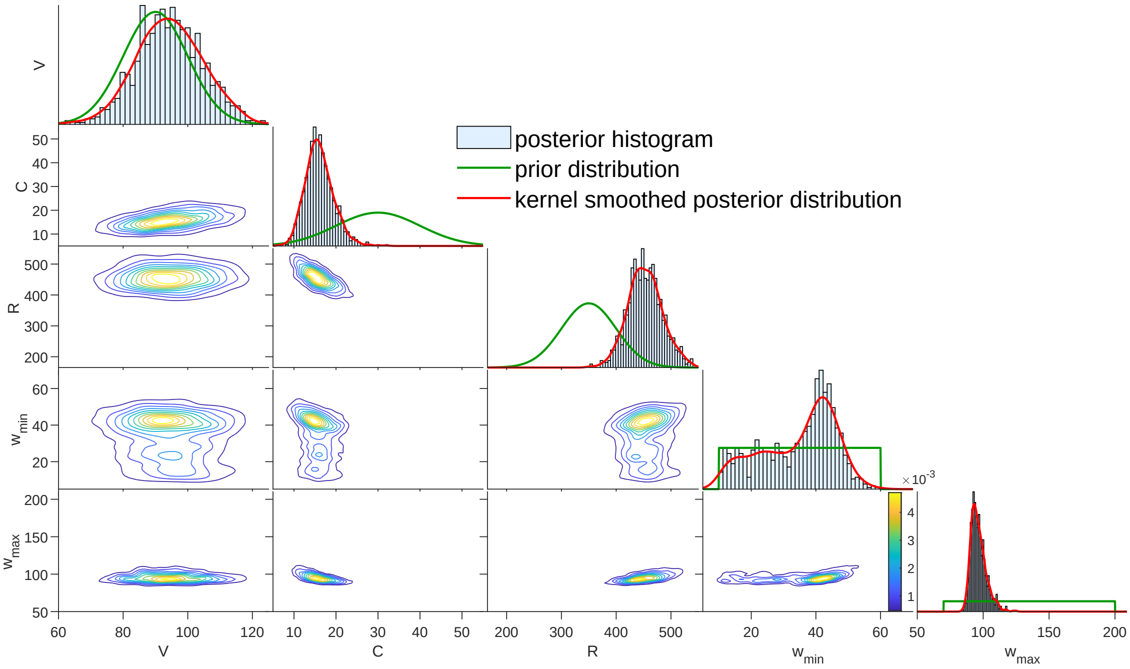
To apply the Metropolis Algorithm 1, we set the number of iterations to $N = 10^5$. It is a common approach to remove the first MCMC outputs in order to reduce the dependence of the proposal distribution on the initial guess. We set this *burn-in phase* to 10% of the N-iterations. Next, we reduce the sample chain to a number $\hat{N}$ minimizing autocorrelations. To this end, we use the multivariate effective sample size (ESS) function, `multiESS` in the `R` package `mcmcse` [18]. The effective sample sizes are listed in Table 5.

	LWR	GSOM
DirMed	720	979
RTMC	1407	842

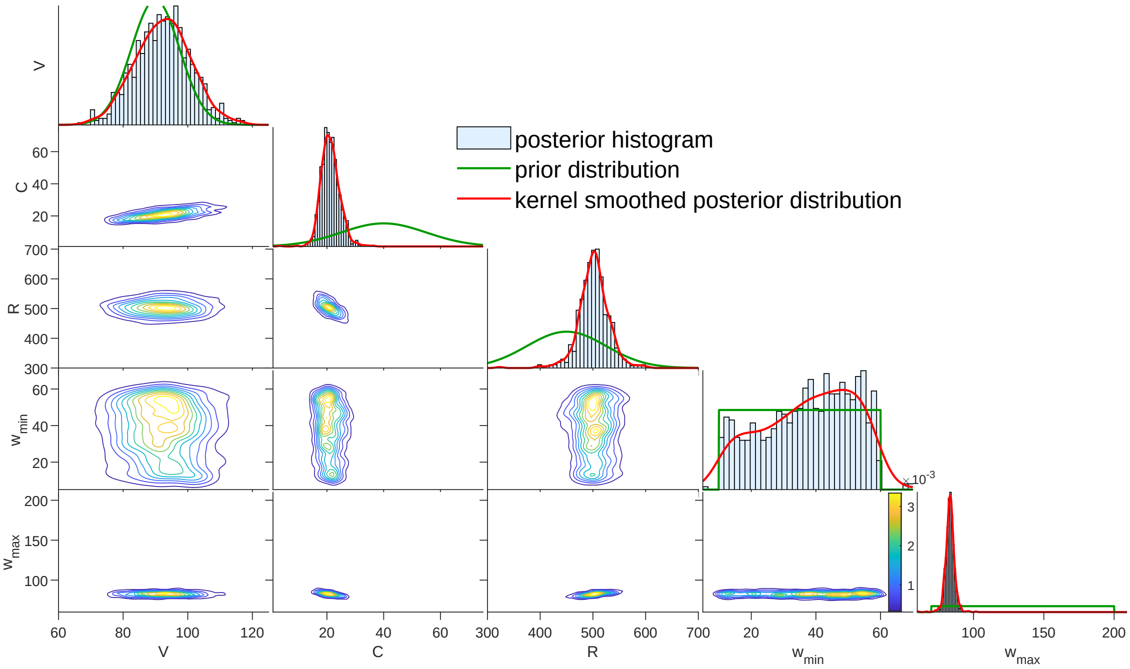
Table 5: Effective sample sizes $\hat{N}$ for the two data sets.

For a graphical representation of the results of the MCMC method, we consider both the histograms and the two-dimensional density contour plots, which are smoothed by a kernel density estimator (see Figure 8). In the histogram graphics, we additionally add the probability density of the prior distribution for the calibration parameters (green line) and the kernel smoothed posterior distribution which is computed by the MATLAB command `fitdist` (red line). This operator fits a kernel probability distribution object to the sample data. The parameters for the kernel distribution object in MATLAB are chosen to be “normal” by default.

We observe that, for both data sets, the speed calibration parameter V has the largest variance among the V, C, R parameters (despite being the most intuitive one to specify a priori) and the kernel smoothed posterior distribution is close to the prior distribution. In comparison, the posterior distributions for C and R are more peaked and shifted in one direction compared to the prior ones. In general, we detect a negative correlation between C and R, visible by the diagonal shape of the $C - R$ contours. We conclude that lower C values correspond to higher R values and vice-versa, whereas the V parameter seems uncorrelated with the other ones. Comparing the histograms for the minimum and maximum empty-road velocity, we see that w_{min} and w_{max} can be shifted or spread along their range. For both



(a) DirMed data



(b) RTMC data

Figure 8: Histograms and 2-dimensional density contour plots for the GSOM model.

data scenarios, the w_{max} parameter takes the more important role whereas w_{min} seems to be less determinant. Note that this may be an effect of the projection in Remark 2. However, this is not always the case since we also observed RTMC and DirMed scenarios which behave in the opposite way. Examples can be found in the pdf-file of the git repository.

For a performance comparison between the two traffic flow models and also between the two calibration approaches, we again define an error metric for the simulated MCMC outputs. In formulas, this reads for the flow error

$$E^{\text{flow}}_{\text{MCMC}} = \frac{1}{T_f \cdot (x_{out} - x_{in})} \frac{1}{\Delta q} \sum_{(x,t)\in(X,T)} \left(\left| y^F_{\text{flow}}(x,t) - \mathbb{E}[y^M_{\text{flow}}(x,t)] \right| \right)$$

$$\text{with } \mathbb{E}[y^M_{\text{flow}}(x,t)] = \mathbb{E}\left[\mathbb{E}[y^M_{\text{flow}}(x,t) \mid \theta^*] \right] \approx \frac{1}{\hat{N}} \sum_{i=1}^{\hat{N}} y^M_{\text{flow}}(x,t,\theta^*_i) \,,$$

where $y^M_{\text{flow}}(x,t,\theta^*_i)$ is defined as the output of the simulation code evaluated at the i^{th} optimal calibration parameter θ^*_i (for $i \in \{1,\dots,\hat{N}\}$) at time t and loop position x. This error metric both quantifies the accuracy of the model and is coherent with the one used for the direct approach.

Analogously, we define the speed (resp. density) error $E^{\text{speed}}_{\text{MCMC}}$ (resp. $E^{\text{density}}_{\text{MCMC}}$) by using the measured and simulated speed (resp. density) values and by using the normalization constant Δv (resp. $\Delta\rho$). Again, we emphasize that we correct all simulated traffic quantities y^M by their predictive means (4.2).
Finally, the total MCMC cost error is given by the sum of the previously defined errors, i.e.,

$$\mathbf{E}_{\text{MCMC}} = E^{\text{flow}}_{\text{MCMC}} + E^{\text{speed}}_{\text{MCMC}} + E^{\text{density}}_{\text{MCMC}} .$$

The error values for the two traffic flow models are listed in Table 6. We clearly see that the total error $\mathbf{E}_{\text{MCMC}}$ is at least 37% higher for the first order model. Additionally, the speed and density predictions of the GSOM model are more accurate for all considered scenarios. As observed with the direct approach, we remark again that although the calibration is based only on the flow data, the corrected simulated speed and density outputs lead to good prediction accuracy results in the second order model, whereas the flow error is lower for the first order LWR model.

	LWR				GSOM			
	$E^{\text{flow}}_{\text{MCMC}}$	$E^{\text{speed}}_{\text{MCMC}}$	$E^{\text{density}}_{\text{MCMC}}$	$\mathbf{E}_{\text{MCMC}}$	$E^{\text{flow}}_{\text{MCMC}}$	$E^{\text{speed}}_{\text{MCMC}}$	$E^{\text{density}}_{\text{MCMC}}$	$\mathbf{E}_{\text{MCMC}}$
DirMed	**1.0560**	1.2720	4.3358	6.6639 (+37%)	1.9956	**1.1756**	**1.6801**	**4.8513**
RTMC	**0.8961**	2.5066	2.2182	5.6209 (+78%)	1.2588	**0.8643**	**1.0287**	**3.1518**

Table 6: Time-space error results for the Bayesian approach. In bold, the lowest flow, speed, density and total errors per data scenario.

Remark 6. *In [15, 16], the authors consider only the error on the speed and the density. For these quantities, they come to the same conclusion, namely that GSOM outperforms the LWR model. However, the flow error is neglected in their evaluations, although they construct their data-fitted model on the measured flow-density fundamental diagram.*

So far, we considered the results of the direct approach and the Bayesian approach separately. To complete the analysis, we also need to compare their performance among each other. We have seen that the Bayesian approach provides us with more information than the direct single output approach. Regarding the numerical error computations, we can compare the entries of Tables 3 (direct approach) and 6 (Bayesian approach). We observe that the total error of the Bayesian approach $\mathbf{E}_{\text{MCMC}}$ is slightly lower for the DirMed data but higher for the RTMC data. For the single error quantities (flow, speed and density error) we cannot detect a clear pattern on the performance accuracy regarding the two approaches. The flow errors lie in a similar range whereas the speed and density errors can be visibly higher for both the direct and Bayesian approach.

5.3 Benchmark on synthetic data

In the previous analysis, we observed that the second order GSOM model outperforms the LWR model on real data regarding the speed, density and total error for both statistical approaches (direct and Bayesian), but not on flow reconstruction. To understand which scenario characteristics are decisive for the flow performance, we perform a study on so-called *synthetic data*, which should be reconstructed more easily by the simulation code than real data. The procedure is the following:

1. Choose a data scenario (e.g., DirMed scenario) to have initial and boundary data.
2. Given these data, choose one arbitrary reasonable set of parameters (e.g., $(V, C, R) = (90, 20, 400)$, $w_{\min} = 40, w_{\max} = 130$) and run both the LWR and GSOM simulations.
3. Compute the synthetic flow and speed data from the simulation output.
4. Construct perturbed data by adding noise to the new synthetic data. This noise accounts for the observational noise and it is assumed to be a normally distributed random variable with zero mean. The noise variance is taken of increasing order $\gamma = 0\%$, 1%, 10%, 20% or 50% of the range of the synthetic flow and speed data.
5. Derive the perturbed density data by the functional relationship: density = flow/speed.
6. Use the initial and boundary conditions of these perturbed data to calibrate both models by the direct approach.
7. Compare the single traffic quantity errors and the total error of both models (computed with respect to the noisy data to match the real setup).

Remark 7. *In step 6, we use both the initial calibration parameters (V, C and R are the same for both models) and the parameters obtained by the global* `pso` *solver (V, C and R are different for the models). The following conclusions, however, are independent of the choice.*

The results are reported in Table 7. For low noise data ($\gamma = 0\%, 1\%$), the second order model performs notably better with respect to all error quantities. Increasing data noise, we observe that the GSOM model still performs better regarding all error quantities for $\gamma \in \{10\%, 20\%\}$. However, for higher γ values, the flow accuracy of the second order model deteriorates faster compared to the LWR model. At $\gamma = 50\%$ the flow error is finally higher.

We conclude that:

1. The higher the noise, the worse the performance on the flow error.
2. The other error quantities (speed, density and total error) are always lower in the second order model for every choice of γ.

Fan [15] comes to a similar conclusion, suggesting that the GSOM model can be beaten by other models due to the oscillatory behaviour of the boundary conditions.

	LWR				GSOM			
γ	E^{flow}	E^{speed}	E^{density}	**E**	E^{flow}	E^{speed}	E^{density}	**E**
0%	0.1417 (+165%)	0.1388	0.2909	0.5714 (+147%)	**0.0534**	0.0793	0.0988	0.2315
1%	0.2082 (+178%)	0.2304	0.3904	0.8289 (+119%)	**0.0748**	0.1099	0.1942	0.3789
10%	0.7897 (+22%)	0.7063	2.1694	3.6654 (+112%)	**0.6482**	0.5337	0.5446	1.7265
20%	0.9957 (+10%)	1.1221	2.6550	4.7727 (+57%)	**0.9044**	0.9841	1.1530	3.0415
50%	**1.5266** (−1%)	1.4391	3.2298	6.1955 (+25%)	1.5360	1.2575	2.1558	4.9493

Table 7: Flow error comparison for synthetic data corresponding to the calibration parameters $(V, C, R) = (90, 20, 400), (w_{\min}, w_{\max}) = (40, 130)$. In bold, the lowest flow error for every choice of γ.

6 Conclusion

We have applied a Bayesian calibration technique for parameter identification and uncertainty quantification in macroscopic road traffic models, exploiting different loop detector data sets. The study has highlighted the globally better performances of second order compared to first order models, except for flow errors, which may suffer from high noise levels.

More generally, the proposed approach results in better reconstruction performances than direct calibration techniques commonly used in practice, which moreover do not consider parameter uncertainty. In particular, our results point out the benefit of introducing a bias term to compensate model limitations in reproducing real data.

From the traffic modeling point of view, further investigations should include more complex situations including on- and off-ramps and road junctions, traffic lights, etc. Also, time or space dependence could be considered for some parameters [42], as well as local variations of the bias on the road possibly depending on the traffic regimes [31].

Regarding the calibration techniques, in this paper we have applied a modularized version of KOH calibration [36, 50], alleviating some of the shortcomings of the original approach. Given the flexibility of this framework, it can be improved in many ways: either by reducing identifiability issues with orthogonality constraints as in [44], or by increasing the scalability with the deployment of deep GPs in a variational inference scheme [37]. Refining the model comparison, e.g., via efficient computation of Bayes factors, is another topic for further research.

A Least square approach

A commonly used approach to calibrate the optimal parameters is the minimization of a least square cost function taking both the real data and simulated data into account, see e.g., [48, 41, 51].

Accordingly, θ will be calibrated based on the cost function

$$C(\theta) = \sum_{(x,t)\in(X,T)} \left| y^F(x,t) - y^M(x,t,\theta) \right|^2 .$$

Thus, the optimal calibration parameter θ^* is given by

$$\theta^* = \operatorname*{argmin}_{\theta\in\Theta} C(\theta).$$

The bounds for the five dimensional parameter space Θ and the initial guesses for the optimization solver `fmincon` from MATLAB are those defined in the right columns of Table 1. The optimization results for the calibration parameters are summarized in Table 8.

	LWR			GSOM				
	V^*	C^*	R^*	V^*	C^*	R^*	w^*_{min}	w^*_{max}
DirMed	93.03	18.38	431.47	91.48	16.28	450	2.874	91.14
RTMC	70.04	32.74	439.19	89.11	36.4	415.87	46.45	70

Table 8: Optimization results for the least squares approach.

Comparing the errors reported in Table 9 with those of Tables 3 and 6, we conclude that both the Bayesian approach and direct approach greatly outperform this basic calibration procedure, thus evidencing the benefit of introducing a bias term.

	LWR				GSOM			
	E^{flow}	E^{speed}	E^{density}	$\mathbf{E}$	E^{flow}	E^{speed}	E^{density}	$\mathbf{E}$
DirMed	**1.2318**	10.8871	27.9798	40.0988 (+349%)	2.3614	**2.9826**	**3.578**	**8.9224**
RTMC	**1.5918**	10.4583	17.3069	29.3571 (+296%)	2.4984	**2.2846**	**2.6247**	**7.4077**

Table 9: Time-space error results for the least squares approach. In bold, the lowest flow, speed, density and total errors per data scenario.

References

[1] A. Aw and M. Rascle. Resurrection of "second order" models of traffic flow. SIAM J. Appl. Math., 60(3):916–938, 2000.

[2] M. J. Bayarri, J. O. Berger, M. C. Kennedy, A. Kottas, R. Paulo, J. Sacks, J. A. Cafeo, C.-H. Lin, and J. Tu. Predicting vehicle crashworthiness: Validation of computer models for functional and hierarchical data. Journal of the American Statistical Association, 104(487):929–943, 2009.

[3] E. Bertino, R. Duvigneau, and P. Goatin. Uncertainty quantification in a macroscopic traffic flow model calibrated on GPS data. Math. Biosci. Eng., 17(2):1511–1533, 2020.

[4] J. Brynjarsdóttir and A. O'Hagan. Learning about physical parameters: The importance of model discrepancy. Inverse problems, 30(11):114007, 2014.

[5] M. Carmassi, P. Barbillon, M. Chiodetti, M. Keller, and E. Parent. Bayesian calibration of a numerical code for prediction. Journal de la société française de statistique, 160(1):1–30, 2019.

[6] N. Chiabaut, C. Buisson, and L. Leclercq. Fundamental diagram estimation through passing rate measurements in congestion. IEEE Transactions on Intelligent Transportation Systems, 10(2):355–359, 2009.

[7] F. A. Chiarello, J. Friedrich, P. Goatin, and S. Göttlich. Micro-macro limit of a nonlocal generalized Aw-Rascle type model. SIAM J. Appl. Math., 80(4):1841–1861, 2020.

[8] R. M. Colombo and A. Groli. On the initial boundary value problem for Temple systems. Nonlinear Anal., 56(4):569–589, 2004.

[9] A. Corbetta, A. Muntean, and K. Vafayi. Parameter estimation of social forces in pedestrian dynamics models via a probabilistic method. Math. Biosci. Eng., 12(2):337–356, 2015.

[10] I. Crandell, A. J. Millican, S. Leman, E. Smith, W. N. Alexander, W. J. Devenport, R. Vasta, R. Gramacy, and M. Binois. Anomaly detection in large-scale wind tunnel tests using Gaussian processes. In 33rd AIAA Aerodynamic Measurement Technology and Ground Testing Conference, page 4131, 2017.

[11] C. F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Res. Part B, 28(4):269–287, 1994.

[12] S. F. Davis. Simplified second-order Godunov-type methods. SIAM J. Sci. Statist. Comput., 9(3):445–473, 1988.

[13] M. L. Delle Monache, K. Chi, Y. Chen, P. Goatin, K. Han, J.-M. Qiu, and B. Piccoli. A three-phase fundamental diagram from three-dimensional traffic data. Axioms, 10(1):17, 2021.

[14] G. Dervisoglu, G. Gomes, J. Kwon, R. Horowitz, and P. Varaiya. Automatic calibration of the fundamental diagram and empirical observations on capacity. In Transportation Research Board 88th Annual Meeting, volume 15, pages 31–59. Citeseer, 2009.

[15] S. Fan. Data-fitted generic second order macroscopic traffic flow models. ProQuest LLC, Ann Arbor, MI, 2013. Thesis (Ph.D.)–Temple University.

[16] S. Fan, M. Herty, and B. Seibold. Comparative model accuracy of a data-fitted generalized Aw-Rascle-Zhang model. Netw. Heterog. Media, 9(2):239–268, 2014.

[17] S. Fan, Y. Sun, B. Piccoli, B. Seibold, and D. B. Work. A collapsed generalized Aw-Rascle-Zhang model and its model accuracy, 2017.

[18] J. M. Flegal, J. Hughes, D. Vats, and N. Dai. mcmcse: Monte Carlo Standard Errors for MCMC. Riverside, CA, Denver, CO, Coventry, UK, and Minneapolis, MN, 2020. R package version 1.4-1.

[19] R. Franklin. The structure of a traffic shock wave. Civil Engineering Pulb. Wks. Rev, 56:1186–1188, 1961.

[20] S. Gerster, M. Herty, and E. Iacomini. Stability analysis of a hyperbolic stochastic galerkin formulation for the aw-rascle-zhang model with relaxation. arXiv preprint arXiv:2102.09359, 2021.

[21] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. Mat. Sb. (N.S.), 47 (89):271–306, 1959.

[22] S. N. Gomes, A. M. Stuart, and M.-T. Wolfram. Parameter estimation for macroscopic pedestrian dynamics models from microscopic data. SIAM J. Appl. Math., 79(4):1475–1500, 2019.

[23] S. Göttlich and S. Knapp. Artificial neural networks for the estimation of pedestrian interaction forces. In L. Gibelli, editor, Crowd Dynamics, Volume 2: Theory, Models, and Applications, pages 11–32. Springer International Publishing, Cham, 2020.

[24] S. Göttlich and C. Totzeck. Optimal control for interacting particle systems driven by neural networks. arXiv preprint arXiv:2101.12657, 2021.

[25] R. B. Gramacy. Surrogates: Gaussian process modeling, design, and optimization for the applied sciences. CRC Press, 2020.

[26] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. SIAM Rev., 25(1):35–61, 1983.

[27] M. Herty, A. Fazekas, and G. Visconti. A two-dimensional data-driven model for traffic flow on highways. Netw. Heterog. Media, 13(2):217–240, 2018.

[28] M. Herty, A. Tosin, G. Visconti, and M. Zanella. Reconstruction of traffic speed distributions from kinetic models with uncertainties. arXiv preprint arXiv:1912.03706, 2019.

[29] D. Higdon, M. Kennedy, J. C. Cavendish, J. A. Cafeo, and R. D. Ryne. Combining field data and computer simulations for calibration and prediction. SIAM Journal on Scientific Computing, 26(2):448–466, 2004.

[30] P. D. Hoff. A first course in Bayesian statistical methods, volume 580. Springer, 2009.

[31] J. Huang, R. B. Gramacy, M. Binois, and M. Libraschi. On-site surrogates for large-scale calibration. Applied Stochastic Models in Business and Industry, 36(2):283–304, 2020.

[32] M. Iglesias and A. M. Stuart. Inverse problems and uncertainty quantification. SIAM News, pages 2–3, 2014.

[33] M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3):425–464, 2001.

[34] J.-P. Lebacque, S. Mammar, and H. H. Salem. Generic second order traffic flow modelling. In Transportation and Traffic Theory 2007. Papers Selected for Presentation at ISTTT17Engineering and Physical Sciences Research Council (Great Britain) Rees Jeffreys Road FundTransport Research FoundationTMS ConsultancyOve Arup and Partners, Hong KongTransportation Planning (International) PTV AG, 2007.

[35] M. J. Lighthill and G. B. Whitham. On kinematic waves. II. A theory of traffic flow on long crowded roads. Proc. Roy. Soc. London Ser. A, 229:317–345, 1955.

[36] F. Liu, M. Bayarri, J. Berger, et al. Modularization in Bayesian analysis, with emphasis on analysis of computer models. Bayesian Analysis, 4(1):119–150, 2009.

[37] S. Marmin and M. Filippone. Variational calibration of computer models. arXiv preprint arXiv:1810.12177, 2018.

[38] Minnesota Departement of Transportation. Mn/Dot Traffic Data. Website: `http://data.dot.state.mn.us/datatools/`.

[39] S. Mollier, M. L. Delle Monache, C. C. de Wit, and B. Seibold. Two-dimensional macroscopic model for large scale traffic networks. Transportation Research Part B: Methodological, 122:309–326, 2019.

[40] G. Newell. A theory of traffic flow in tunnels. Theory of Traffic Flow, pages 193–206, 1961.

[41] D. Ngoduy and S. Hoogendoorn. An automated calibration procedure for macroscopic traffic flow models. IFAC Proceedings Volumes, 36(14):263–268, 2003. 10th IFAC Symposium on Control in Transportation Systems 2003, Tokyo, Japan, 4-6 August 2003.

[42] M. Pereira, P. B. Baykas, B. Kulcsár, and A. Lang. Parameter and density estimation from real-world traffic data: A kinetic compartmental approach. arXiv preprint arXiv:2101.11485, 2021.

[43] B. Piccoli, K. Han, T. L. Friesz, T. Yao, and J. Tang. Second-order models and traffic data from mobile sensors. Transportation Research Part C: Emerging Technologies, 52:32–56, 2015.

[44] M. Plumlee. Bayesian calibration of inexact computer models. Journal of the American Statistical Association, 112(519):1274–1285, 2017.

[45] C. E. Rasmussen and C. Williams. Gaussian Processes for Machine Learning. MIT Press, 2006.

[46] P. I. Richards. Shock waves on the highway. Operations Res., 4:42–51, 1956.

[47] M. L. Stein. Interpolation of spatial data: some theory for kriging. Springer Science & Business Media, 2012.

[48] G. Strofylas, K. Porfyri, I. Nikolos, A. Delis, and M. Papageorgiou. Using synchronous and asynchronous parallel differential evolution for calibrating a second-order traffic flow model. Advances in Engineering Software, 125:1–18, 2018.

[49] B. Temple. Systems of conservation laws with invariant submanifolds. Trans. Amer. Math. Soc., 280(2):781–795, 1983.

[50] R. Tuo and C. Jeff Wu. A theoretical framework for calibration in computer models: parametrization, estimation and convergence properties. SIAM/ASA Journal on Uncertainty Quantification, 4(1):767–795, 2016.

[51] P. Wagner. Fluid-dynamical and microscopic description of traffic flow: a data-driven comparison. Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 368(1928):4481–4495, 2010.

[52] D. B. Work, S. Blandin, O.-P. Tossavainen, B. Piccoli, and A. M. Bayen. A traffic model for velocity data assimilation. Appl. Math. Res. Express. AMRX, pages 1–35, 2010.

[53] H. M. Zhang. A non-equilibrium traffic model devoid of gas-like behavior. Transportation Res. Part B, 36(3):275–290, 2002.