A probabilistic approach for the calibration of incomplete microscopic traffic models

Yiolanda Englezou, Stelios Timotheou and Christos G. Panayiotou

Abstract— **Digital Twins (DTs) are steadily gaining popularity for the study of large systems in real time. In order to build an efficient and reliable DT, it is crucial to perform calibration prior to its use, that is, to use real data to estimate unknown parameters in the DT, that are of great importance for the actual physical process. This work studies the calibration of the Intelligent Driver Model (IDM) to infer driver behaviour, a crucial task when building a DT of the traffic network. We introduce a statistical model to calibration which takes into account the model uncertainty, while also taking into account possible correlations between individual vehicles that have similar characteristics. In contrast with other works in the literature, we assume that vehicles belong to a pre-specified group and develop a Bayesian approach to derive the posterior distributions of the parameters that characterise each group's behaviour. We apply the proposed approach to the IDM and derive probability density functions of the unknown model parameters and the model uncertainty. The proposed probabilistic approach is validated using realistic SUMO micro-simulations of a highway stretch. We present estimation results that show that the proposed approach can accurately derive the posterior distributions of the calibration parameters. In addition, we compare the proposed approach with a literature methodology and show that our approach higher quality posterior distributions of the parameters of interest than the literature approach.**

I. INTRODUCTION

Digital Twins (DTs) were developed from the Internet of Things as a way to formulate highly efficient simulation models from physical processes, that allow remote monitoring and control of the system [1]. A DT is continuously updated as real data are acquired in (near) real-time, which are typically used in combination with synthetic data generated from simulators. A simulator is an implementation of a complex mathematical model that maps input variables to an output, and brings physical realism in high spatiotemporal spaces [2].

This work is supported by the European Union (i. ERC, URANUS, No. 101088124, ii. Horizon 2020 Teaming, KIOS CoE, No. 739551, and iii. MSCA, BITS, No. 101003435), and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation, and Digital Strategy. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Y. Englezou, S. Timotheou and C. G. Panayiotou are with the KIOS Research and Innovation Center of Excellence, and the Department of Electrical and Computer Engineering, University of Cyprus, {englezou.yiolanda, timotheou.stelios, christosp}@ucy.ac.cy

In order to build an efficient and reliable DT, it is crucial to perform *calibration* prior to its use, that is, to use real data to estimate unknown parameters in the DT, that are of great importance for the actual physical process [3]. Calibration is the process that improves the representativeness of the simulator, or mathematical model, with respect to a physical phenomenon. Calibration parameters are present only in the simulator and have meaning in the physical experiment, however their values are either unknown or unmeasured during the running of the physical experiment. Hence, calibration is defined as the process of determining plausible values of the unknown parameters to match the observations obtained from the physical experiment [4].

Traffic flow simulation has been widely used in transportation planning, analysis and safety studies, and several microscopic and macroscopic models have been proposed over the years [5], [6]. Microscopic traffic models simulate single vehicles and the dynamic variables of such models represent microscopic properties like the position, speed and acceleration of individual vehicles. Macroscopic models assume a sufficiently large number of vehicles within a road segment such that each stream of vehicles can be treated as flowing in a tube. However it is well known that no model is perfect [7], hence calibration is a crucial procedure for the application of traffic flow models.

Several papers over the years have investigated microscopic traffic model calibration, such as car-following models. Such models control a vehicles' behaviour with respect to the front, or preceding, vehicle in the same lane. In order to realistically represent vehicle behaviour there is a need to collect data in order to efficiently calibrate the car-following model of interest. In this work we focus on the calibration of the Intelligent Driver Model (IDM). The calibration task is typically formulated as a least squares approach to minimise the difference between the simulated and observed measurements [8]. The Maximum Likelihood Estimation (MLE) approach was also proposed, that aims to estimate a point 'best' estimate of the unknown parameters of interest [9], [10]. More recently, the Bayesian inference paradigm was used as an alternative, aiming to derive the probability density function of the unknown parameters [11], [12]. To improve the estimation accuracy of the Bayesian framework, [13] used a hierarchical model formulation for multiple individual vehicles, while [14] extended this hierarchical framework to take into account the autocorrelation per individual driver.

Despite the extensive study of the efficient calibration of car-following models, the potential of the Bayesian paradigm for the calibration and validation of car-following models has not been thoroughly explored in the literature. In this work we aim to build upon the literature and derive probability density functions of the unknown calibration parameters in the IDM, while taking into account the correlation parameters between individual vehicles that might be classified to the same group of vehicle behaviour and characteristics. In this work, we develop a general framework for the statistical calibration of the IDM utilising the Bayesian paradigm, by accepting the fact that the mathematical model, or simulator, is not perfect and take into account model uncertainty, i.e. the discrepancy between real-life measurements and simulator outputs. Finally, the proposed approach is applied to the IDM, utilising realistic synthetic data for a 4km-long highway stretch, obtained from the SUMO microsimulator, and it is compared with a literature methodology.

The paper is organised as follows. In the next section (Section II) we outline the IDM, which will be used to validate the proposed approach. In Section III we introduce a general statistical calibration model, while in Section IV we propose the Bayesian framework to derive the probability density function of the unknown calibration parameters of interest. In Section V we validate our proposed methodology using synthetic data from the SUMO microsimulator and present estimation results of the proposed approach. Section VI summarises the main results of this work and outlines future work.

II. Problem Statement

Assume a set of vehicles $\mathcal{D} = \{1, \dots, D\}$ that traverse in a road segment which we observe for T hours and $\mathcal{T} = \{1, \dots, t, t+1, \dots, K\}$ the set of all time-steps. The IDM specifies a vehicle's $i \in \mathcal{D}$ acceleration given the vehicle's speed, u_{it}, distance relative to its leading vehicle, s_{it}, and speed difference $\Delta u_{it} = u_{it} - u_{lt}$, where $l \in \mathcal{D}$ is the leading vehicle and u_{lt} the speed of the leading vehicle at time t [10]. The acceleration of the IDM is defined as

$$a_{it+1} = a_{\text{max}} \left[1 - \left(\frac{u_{it}}{u_0} \right)^{\delta} - \left(\frac{s^*(u_{it}, \Delta u_{it})}{s_{it}} \right)^2 \right],$$
$$s^*(u_{it}, \Delta u_{it}) = s_0 + \max \left(0, u_{it} T_{\text{safe}} + \frac{u_{it} \Delta u_{it}}{2\sqrt{a_{\text{max}} \beta}} \right), \quad (1)$$

$\forall i \in \mathcal{D}$, $t \in \mathcal{T}$, where u_0 is the desired speed, s_0 is the minimum gap of vehicle i and the precending vehicle, T_{safe} is the minimum travel time interval between vehicle i and the leading vehicle, a_{max} is the maximum vehicle acceleration, β the comfortable breaking deceleration and δ a constant that represents the rate at which a vehicle's acceleration is changing when the vehicle is approaching the desired velocity [13].

In this work we propose a Bayesian calibration framework aiming to derive probability density functions of the unknown model parameters, or calibration parameters, $\boldsymbol{\theta} = [u_0, T_{\text{safe}}, a_{\text{max}}, \beta, \delta, s_0]^{\text{T}}$, as we observe $\mathbf{x}_{it} = [s_{it}, u_{it}, \Delta u_{it}]^{\text{T}}$ and $y_{it} = \{a_{it}\}$, $\forall t \in \mathcal{T}$ and $\forall i \in \mathcal{D}$, whilst acknowledging correlation parameters between groups of vehicles and model uncertainty, i.e. the discrepancy between the model output, and real-life measurements. These parameters depend on the vehicles, i.e. each individual vehicle $i \in \mathcal{D}$ has its own set of parameters, denoted by $\boldsymbol{\theta}_i$.

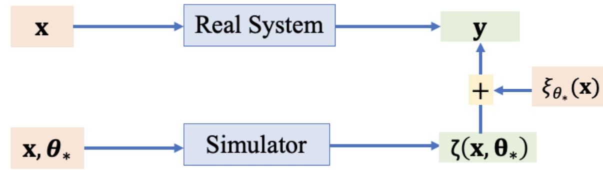

Fig. 1: The proposed statistical calibration concept.

III. A Statistical Model for Calibration

For the physical process under study, let y_{it} denote the response at a specific time $t \in \mathcal{T}$ obtained using $\mathbf{x}_{it}$. We assume a dynamic calibration model given as

$$y_{it} = g(\mathbf{x}_{it}) + \varepsilon_{it} = \eta(\mathbf{x}_{it}, \boldsymbol{\theta}_{i*}) + \delta_{\boldsymbol{\theta}_{i*}}(\mathbf{x}_{it}) + \varepsilon_{it},$$
$$\mathbf{x}_{it+1} = \delta(\mathbf{x}_{it}, y_{it}) \quad (2)$$

$\forall i \in \mathcal{D}$ and $\forall t \in \mathcal{T}$. The true value of the real process when the inputs take values $\mathbf{x}_{it}$ is denoted by $g(\mathbf{x}_{it})$. Function $\eta(\cdot, \cdot)$ denotes the simulator output when the inputs take values $\mathbf{x}_{it}$ and $\boldsymbol{\theta}_i$. The discrepancy function, $\delta_{\boldsymbol{\theta}_{i*}}(\cdot)$, encodes the difference between the simulator evaluated at the 'true' $\boldsymbol{\theta}_{i*}$, $\eta(\mathbf{x}_{it}, \boldsymbol{\theta}_{i*})$, and mean $g(\mathbf{x}_{it})$ of the real system [3]. We assume $\varepsilon_{it} \sim N(0, \sigma_{\varepsilon}^2)$ are independent. The objective of this work is to derive probability density functions of the unknown calibration parameters, $\pi(\boldsymbol{\theta}_i | \mathbf{y}_i)$ and the unknown discrepancy between the simulator and real-life measurements, $\pi(\delta_{\theta_i}(\mathbf{x}_i) | \mathbf{y}_i)$.

The basic idea of model (2) (also represented graphically in Figure 1) is that a real process can be represented by a simulator $\eta(\cdot, \cdot)$, which must be run under the 'true' calibration parameters $\boldsymbol{\theta}_*$. However, we need to take into account the model uncertainty, $\delta_{\theta}(\cdot)$, as a simulator is usually built under certain assumptions that might not fully represent real-life [15].

For the IDM given by (1), we have that $\boldsymbol{\theta}_i = [u_0^{(i)}, T_{\text{safe}}^{(i)}, a_{\text{max}}^{(i)}, \beta^{(i)}, \delta^{(i)}, s_0^{(i)}]^{\text{T}}$, $\mathbf{x}_{it} = [s_{it}, u_{it}, \Delta u_{it}]^{\text{T}}$, $y_{it} = \{a_{it}\}$ and $u_{it+1} = u_{it} + a_{it} \Delta t$, $\forall i \in \mathcal{D}$ and $\forall t \in \mathcal{T}$, where Δt denotes the model update time interval. Hence, the calibration model (2) of the IDM becomes:

$$y_{it} = \eta(u_{it}, s_{it}, \Delta u_{it}, \boldsymbol{\theta}_{i*}) + \delta_{\boldsymbol{\theta}_{i*}}(u_{it}, s_{it}, \Delta u_{it}) + \varepsilon_{it},$$
$$u_{it+1} = u_{it} + a_{it} \Delta t. \quad (3)$$

IV. Bayesian Calibration

In this section, we develop a general Bayesian calibration approach that can be applied to calibrate dynamic models, e.g. microscopic car-following models. Prior distributions of car-following model parameters are considered to derive posterior distributions of these parameters using Bayes' theorem. The proposed solution approach is formulated such that it takes into account the model uncertainty.

Assume that $\eta(\mathbf{x}, \boldsymbol{\theta}_*)$ is a known mathematical model, e.g. the IDM as described in (1), with fixed but unknown calibration parameters $\boldsymbol{\theta}_*$, and $\mathbf{x} = [\mathbf{x}_1^{\mathrm{T}}, \ldots, \mathbf{x}_D^{\mathrm{T}}]^{\mathrm{T}}$ denotes the input settings. Prior distributions are required for the unknown parameters $\boldsymbol{\theta}$, denoted as $\pi(\boldsymbol{\theta})$, to derive the posterior distributions of interest. In order to make appropriate use of observations of the physical system it is important to take into account the model uncertainty [15].

We represent prior uncertainty about the discrepancy function $\delta_{\boldsymbol{\theta}_*}(\mathbf{x})$, using a Gaussian process (GP) model [16]. A GP model is fully defined by its mean and variance function. We assume the GP prior

$$\delta_{\theta_*}(\mathbf{x}) \sim \mathrm{GP}\left[\mathbf{f}^{\mathrm{T}}(\mathbf{x})\boldsymbol{\beta}, \sigma^2\kappa(\mathbf{x}, \mathbf{x}'; \boldsymbol{\phi})\right], \tag{4}$$

where $\mathbf{f}(\mathbf{x}) = [f_0(\mathbf{x}), f_1(\mathbf{x}), \ldots, f_{q-1}(\mathbf{x})]^{\mathrm{T}}$ is a q-vector of known regression functions, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots \ \ldots, \beta_{q-1})^{\mathrm{T}} \in \mathcal{W}$ is also a q-vector which contains unknown regression coefficients, $0 \leq \kappa(\mathbf{x}, \mathbf{x}'; \boldsymbol{\phi}) \leq 1$ is the correlation function, $\boldsymbol{\phi} \in \Phi$ is the vector of correlation parameters and $\sigma^2 > 0$ is the constant variance. We select constant regression functions $\mathbf{f}(\mathbf{x})$ whose coefficients $\boldsymbol{\beta}$ are to be inferred from the data [17]. Under prior (4), a collection of function evaluations $\boldsymbol{\delta}_\theta = [\delta_\theta(\mathbf{x}_1), \ldots, \delta_\theta(\mathbf{x}_D)]^{\mathrm{T}}$ has a multivariate normal distribution,

$$\boldsymbol{\delta}_\theta \sim \mathrm{N}\left[\mathbf{F}\boldsymbol{\beta}, \sigma^2\mathbf{K}(\boldsymbol{\phi})\right], \tag{5}$$

where $\mathbf{F} = [\mathbf{f}(\mathbf{x}_1)\ \mathbf{f}(\mathbf{x}_2)\ldots\mathbf{f}(\mathbf{x}_D)]^{\mathrm{T}}$ is the $D \times q$ model matrix and $\mathbf{K}(\phi)$ is the $D \times D$ correlation matrix with $\mathbf{K}(\boldsymbol{\phi})_{kl} = \kappa(\mathbf{x}_s, \mathbf{x}_l, \boldsymbol{\phi})$, $\mathbf{x}_s, \mathbf{x}_l \in \mathcal{X}$, $s = \{i, t\}$, $l = \{i', t'\}$.

Combining (2) with (5) we get the likelihood distribution

$$\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2, \phi, \sigma_\varepsilon^2 \sim \mathrm{N}_{D,K}(\eta(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{F}\boldsymbol{\beta}, \sigma^2\mathbf{K}(\boldsymbol{\phi}), \sigma_\varepsilon^2\mathbf{I}_K),$$

where $\mathbf{y} = [\mathbf{y}_1^{\mathrm{T}}, \ldots, \mathbf{y}_D^{\mathrm{T}}]^{\mathrm{T}}$ is the $D \times K$ matrix of all vehicles acceleration and all time-steps, and $\mathbf{I}_K$ is the $K \times K$ identity matrix. We set $\mathbf{M}_y = \eta(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{F}\boldsymbol{\beta}$ and the likelihood function is:

$$\begin{aligned}&\pi_l(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\phi}, \sigma_\varepsilon^2)\\ &= \frac{\exp\left\{\frac{1}{2}\mathrm{tr}[\sigma_\varepsilon^{-2}\mathbf{I}_k^{-1}(\mathbf{y} - \mathbf{M}_y)^{\mathrm{T}}\sigma^{-2}\mathbf{K}(\boldsymbol{\phi})^{-1}(\mathbf{y} - \mathbf{M}_y)]\right\}}{(2\pi)^{KD/2}|\mathbf{I}_K|^{K/2}|\mathbf{K}(\boldsymbol{\phi})|^{D/2}}.\end{aligned} \tag{6}$$

Following, prior distributions are assumed for the unknown parameters $\boldsymbol{\theta}$. We assume that vehicles with common characteristics belong to the same group, i.e. they have a similar driving behaviour, and more than one groups exist. We assume there exist G groups of parent set of parameters and we choose multivariate normal distributions for each set of group-related parameters, i.e. $\tilde{\boldsymbol{\theta}}_c \sim N(\boldsymbol{\mu}_{\tilde{\theta}_c}, \mathbf{S}_{\tilde{\theta}_c})$, $c = \{1, \ldots, G\}$, with mean vector $\boldsymbol{\mu}_{\tilde{\theta}_c}$ and variance covariance matrix $\mathbf{S}_{\tilde{\theta}_c}$. Note that, each vehicle $i \in \mathcal{D}$ has a different set of parameters $\boldsymbol{\theta}_i$, as also mentioned before, obtained through $N(\tilde{\boldsymbol{\theta}}_c, \mathbf{S}_{\tilde{\theta}_c})$, where the mean and variance are associated with the vehicles's group.

A common selection for $\boldsymbol{\beta}$ and σ^2 of the GP prior (4), is to use conjugate prior distributions [16]. Hence we assume that $\boldsymbol{\beta}|\sigma^2$ follows a normal distribution and σ^2 an inverse-gamma distribution, $\boldsymbol{\beta}|\sigma^2 \sim N(\boldsymbol{\beta}_0, \sigma^2\mathbf{S})$ and $\sigma^2 \sim \mathrm{IG}(\delta_1, \delta_2)$, where $\boldsymbol{\beta}_0$ is the q known mean vector, $\mathbf{S}$ is a known symmetric, positive definite $q \times q$ matrix and $\delta_1, \delta_2 > 0$ are known hyperparameters. For the noise variance σ_ε^2 and the vector of correlation parameters $\boldsymbol{\phi}$ we choose exponential prior distributions, $\sigma_\varepsilon^2 \sim \mathrm{Exp}(\lambda_{\sigma_\varepsilon^2})$ and $\boldsymbol{\phi} \sim \mathrm{Exp}(\boldsymbol{\lambda}_\phi)$, where $\lambda_{\sigma_\varepsilon^2}$ and $\boldsymbol{\lambda}_\phi = [\lambda_s, \lambda_u, \lambda_{\Delta u}]^{\mathrm{T}}$ are known hyperparameters. The exponential distribution is a right-skewed distribution in which small values are more likely to occur than higher values, an appropriate assumption both for the error variance and the correlation parameters, as vehicles with common characteristics are highly correlated, hence due to the choice of the squared exponential correlation function, small values of $\boldsymbol{\phi}$ are more likely to occur. The joint prior density is given by

$$\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\phi}, \sigma_\varepsilon^2) = \pi(\boldsymbol{\theta})\pi(\boldsymbol{\beta}|\sigma^2)\pi(\sigma^2)\pi(\boldsymbol{\phi})\pi(\sigma_\varepsilon^2), \tag{7}$$

and the posterior distribution is given by

$$\begin{aligned}\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\phi}, \sigma_\varepsilon^2|\mathbf{y}) &\propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{\beta})\pi(\sigma^2)\pi(\boldsymbol{\phi})\pi(\sigma_\varepsilon^2)\\ &\quad \times \pi_l(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\phi}, \sigma_\varepsilon^2).\end{aligned} \tag{8}$$

Note that the values used for vehicle acceleration, denoted by $\mathbf{y}$, are obtained through the IDM and the parameters $\mathbf{x}$ are measured through sensors mounted on the vehicles.

The challenge here is that (8), is not available in closed-form and the integrals cannot be solved analytically. Hence in order to obtain a sample from the posterior distribution, $\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\phi}, \sigma_\varepsilon^2|\mathbf{y})$, we employ sampling techniques based on Markov Chain Monte Carlo (MCMC) methods [18]. The idea of MCMC is in a sense to by-pass the mathematical operations rather than to implement them. MCMC methods construct a Markov chain with steady state distribution equal to the posterior density of interest. A widely used MCMC algorithm that is relatively simple, is the Metropolis-Hastings algorithm [19] and it will be used for the purposes of this work (see [20] for more information).

The marginal posterior distribution of the unknown calibration parameters $\boldsymbol{\theta}$, is obtained by numerically integrating-out all nuisance parameters:

$$\pi(\boldsymbol{\theta}|\mathbf{y}) = \int_{\sigma_\varepsilon^2}\int_{\boldsymbol{\phi}}\int_{\sigma^2}\int_{\boldsymbol{\beta}} \pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\phi}, \sigma_\varepsilon^2|\mathbf{y})d\boldsymbol{\beta}d\sigma^2 d\boldsymbol{\phi}d\sigma_\varepsilon^2. \tag{9}$$

Once a sample from $\pi(\boldsymbol{\theta}|\mathbf{y})$, is obtained we can also derive $\pi(\delta_\theta(\mathbf{x})|\mathbf{y})$ through (2) and the sample from $\pi(\boldsymbol{\theta}|\mathbf{y})$.

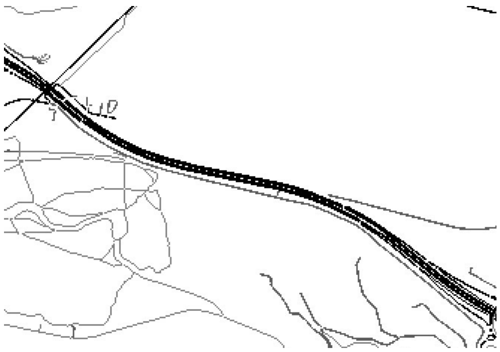

Fig. 2: SUMO representation of the highway stretch.

V. Simulation Results

To evaluate the performance of the proposed probabilistic approach we examine a simulation study of a 4km-long highway stretch using the SUMO microscopic simulator [21], shown in Figure 2. The highway stretch has two lanes in each direction of traffic and the speed limit is 120 km/h. To create traffic within the network, we consider the IDM. We observe the network for $T = 2$ hours with $D = 1,500$ vehicles passing through the specific network.

We utilise the proposed calibration framework and obtain estimation results that are compared with the ground truth used during the simulation. The behaviour of each vehicle $i \in \mathcal{D}$ is described through parameters $\boldsymbol{\theta}_i = [u_0, T_{\text{safe}}, a_{\max}, \beta, s_0]$ (we treat δ as a constant with value set to 4 [10]) and assume a-priori $G = 3$ groups of parent set of distributions for the unknown calibration parameters:

- Group 1 (G1): *Aggressive vehicles*
- Group 2 (G2): *Normal vehicles*
- Group 3 (G3): *Submissive vehicles*

As mentioned, each group has its own set of parameters $\tilde{\boldsymbol{\theta}}_c$, $c = \{1, 2, 3\}$ for which we assume that $\log \tilde{\boldsymbol{\theta}}_c \sim N(\boldsymbol{\mu}_{\tilde{\boldsymbol{\theta}}_c}, \mathbf{S}_{\tilde{\boldsymbol{\theta}}_c})$, to avoid negative values that do not have physical meaning. For each vehicle $i \in \mathcal{D}$ that belongs in group c, we assume its parameters follow $\boldsymbol{\theta}_i \sim \log N(\tilde{\boldsymbol{\theta}}_c, \mathbf{S}_{\tilde{\boldsymbol{\theta}}_c})$. Our aim is to utilise the obtained data and prior information of all unknown parameters to derive the posterior density of the parameters that describe each group.

We evaluate the calibration results in the form of posterior distributions, as shown in Figure 3. Column 1 in Figure 3 corresponds to Group 1, column 2 to Group 2 and column 3 to Group 3, while each row represents a different calibration parameter, i.e. u_0, T_{safe}, $a_{\max}$, β and s_0, respectively. In Figure 3, the red line depicts the true distribution of $\boldsymbol{\theta}$, i.e. the distribution assumed when running the simulation experiment, the grey lines represent the per vehicle estimated posterior distributions of $\boldsymbol{\theta}$, the blue line depicts the mean posterior distribution and the green line the median posterior distribution of all vehicles in the group. As shown, the estimated average posterior densities are very similar to the true densities for all parameters and the three groups, with the mean values of the estimated distributions being very close to the true mean, while the estimated standard deviations are larger than the standard deviations of the true distributions.

Par.	MAE			D_{KL}		
	G1	**G2**	**G3**	**G1**	**G2**	**G3**
u_0	0.13	0.18	0.17	4.6×10^{-4}	9×10^{-4}	1.8×10^{-3}
T_{safe}	0.03	0.24	0.02	0.15	0.11	0.13
$a_{\max}$	0.11	0.06	0.12	0.09	0.23	0.56
β	0.03	0.06	0.15	0.19	0.43	0.67
s_0	0.02	0.01	0.09	0.36	0.10	0.05

TABLE I: MAE and KL divergence, D_{KL}, of the estimated posterior distributions of the calibration parameters for each of the three vehicles' groups.

We further evaluate the performance of the proposed approach in terms of the Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{D}\sum_{i=1}^{D} |\theta_p^{\text{true}} - \hat{\theta}_{p,i}|, \tag{10}$$

where θ_p^{true} is set to the mean value of the 'true" distribution used in SUMO for $p = 1, \ldots, P$ and $\hat{\theta}_{p,i}$ is the mean value of the estimated posterior densities for each calibration parameter marginalised over each vehicle $i \in \mathcal{D}$. In addition, we calculate the Kullback-Leibler (KL) divergence [22], between the derived posterior probability distribution $\pi(\boldsymbol{\theta}|\mathbf{y})$ and the actual probability distribution assumed during the simulation experiment denoted as $f(\boldsymbol{\theta}|\mathbf{y})$. The KL divergence is given by

$$D_{\text{KL}}[\pi(\boldsymbol{\theta}|\mathbf{y})||f(\boldsymbol{\theta}|\mathbf{y})] = \int_{\boldsymbol{\theta}\in\Theta} \pi(\boldsymbol{\theta}|\mathbf{y}) \log\left[\frac{\pi(\boldsymbol{\theta}|\mathbf{y})}{f(\boldsymbol{\theta}|\mathbf{y})}\right] d\boldsymbol{\theta}. \tag{11}$$

Note that a KL divergence equals to zero indicates that the two compared distributions are identical.

Table I illustrates the MAE and KL-divergence, D_{KL}, of the proposed calibration approach. As shown, the MAE is very low for all parameters and all three groups of vehicles indicating than the mean values of the true and estimated posterior distributions are very close, as also shown in Figure 3. In addition, D_{KL} is approaching zero for u_0 and all groups, supporting that the true and estimated distributions are actually of the same family. Parameters T_{safe}, $a_{\max}$, β and s_0 also yield small values of D_{KL} indicating that the true and estimated distributions are very similar. The only exception is the posterior distributions of $a_{\max}$ and β for Group 3 that appear to slightly differ from the true distributions, however the mean of the distributions still remains close to the true mean. This result might be affected by the population of Group 3, which is 70% smaller than Group 2 and 40% smaller than Group 1, meaning that the number of vehicles might not be sufficient to 'learn' the specific parameters with high accuracy.

Following, we compare the proposed calibration approach with a two-level hierarchical approach proposed by [13], that

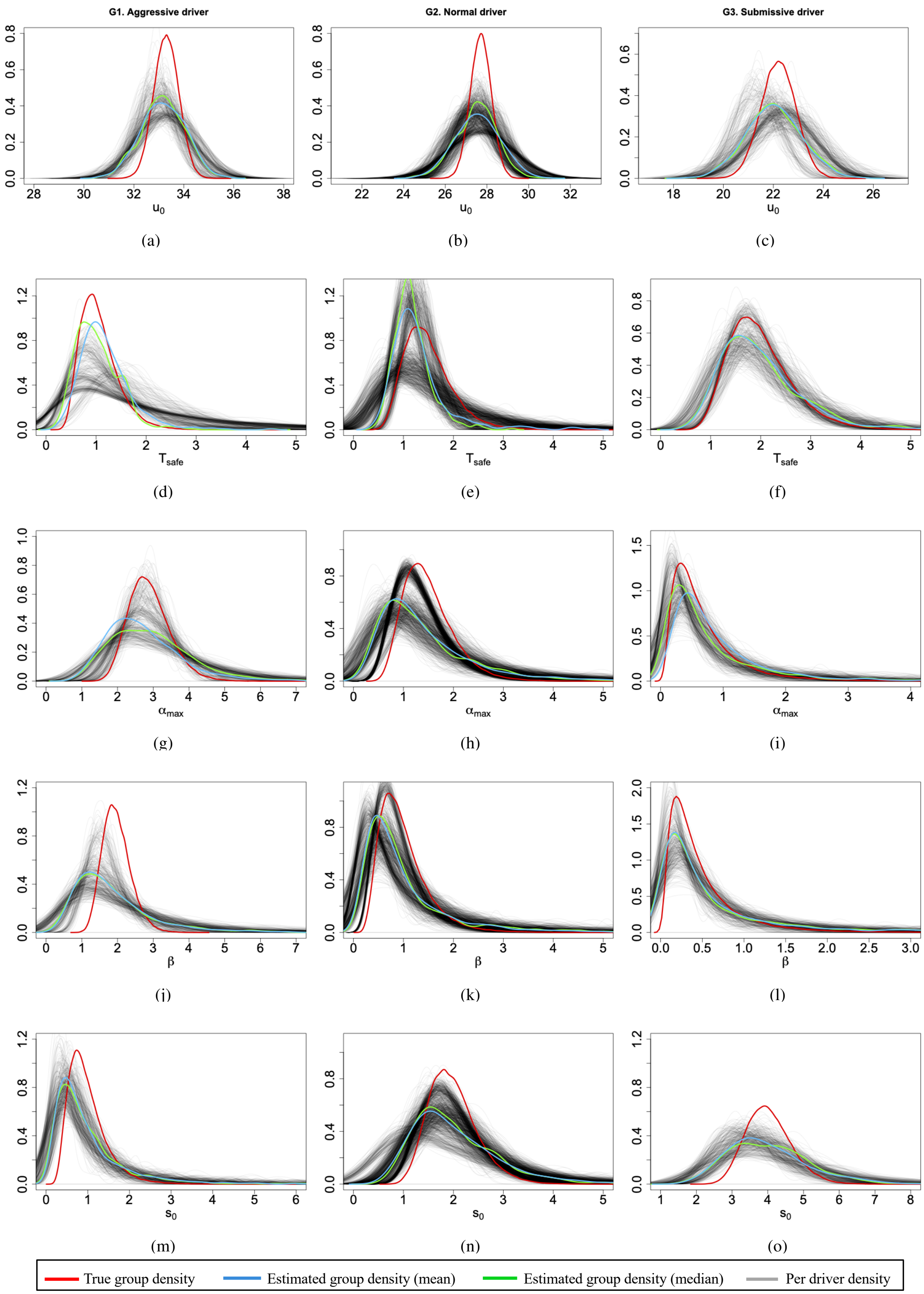

Fig. 3: Estimated posterior distributions of the unknown calibration parameters, $\boldsymbol{\theta}$. The red line depicts the true distribution of $\boldsymbol{\theta}$, the grey lines represent the per vehicle estimated posterior distributions of $\boldsymbol{\theta}$, the blue line depicts the mean posterior distribution of all vehicles in the group and the green line the median posterior distribution of all vehicles in the group. Each column corresponds to one group of vehicles, while each row corresponds to a different calibration parameter.

	Proposed Approach			Two level hierarchical approach		
Par.	**MAE**	D_{KL}	st. dev.	**MAE**	D_{KL}	st. dev.
u_0	0.16	0.001	1.03	2.64	0.004	2.5
T_{safe}	0.10	0.13	0.54	0.27	0.18	1.01
$a_{\max}$	0.10	0.30	1.07	0.88	0.47	1.24
β	0.08	0.43	1.13	0.26	0.44	1.02
s_0	0.12	0.17	0.94	0.86	0.32	1.17

TABLE II: MAE, KL divergence, D_{KL}, and standard deviation (st. dev.) of the estimated posterior distributions of the calibration parameters for the proposed and the literature two-level hierarchical approach.

assumes that each vehicle has an individual set of calibration parameters with their mean and variance drawn from one parent distribution shared across all vehicles. In Table II we present the MAE, KL divergence, D_{KL}, and standard deviation of the estimated posterior distributions of our proposed approach, averaged across the three different groups of vehicles, and the literature two-level hierarchical approach. As shown, the two level hierarchical approach consistently yields larger MAE compared to the proposed approach for all parameters and the standard deviation is around two times larger than the standard deviation of the proposed approach. As shown in Figure 3, the true distributions (red lines) have small standard deviation, hence using a distribution with much larger standard deviation to obtain the driver's behaviour will introduce more variability in the samples and the output of the model might not represent reality. The KL-divergence of the two-level hierarchical approach is also higher than the proposed approach, however the resulted values indicate that the estimated posterior distributions of both approaches and the true distribution assumed in the simulation experiment are of the same family.

VI. Conclusions and Future Work

In this paper we have introduced a novel statistical model for the calibration of car-following models and in particular the IDM. The proposed approach considers model uncertainty and possible correlations between individual vehicles that have similar driving characteristics. The proposed probabilistic approach was validated using realistic SUMO micro-simulations of a highway stretch. Results showed that the proposed approach yields high-accuracy posterior distributions of the unknown calibration parameters of interest. Further, we compared the proposed approach with a state-of-the-art approach and showed that our proposed approach yields posterior distributions with mean value closer to the true mean and significantly smaller standard deviation than the literature approach.

Future plans include the validation of this approach in real-life datasets and the calibration of other microscopic models. Further we aim to extend this approach to learn the number of group of vehicles as an unknown parameter rather than using a pre-specified number of groups. Finally, we aim to compare the proposed approach with other state-of-the-art approaches that are not necessarily under the Bayesian paradigm.

References

[1] A. Rasheed, O. San, and T. Kvamsdal, "Digital twin: Values, challenges and enablers from a modeling perspective," *IEEE Access*, 2020.

[2] F. Tao, H. Zhang, A. Liu, and Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics*, 2019.

[3] M. C. Kennedy and A. O'Hagan, "Bayesian calibration of computer models," *Journal of Royal Statistical Society Series B*, vol. 63, pp. 425–464, 2002.

[4] G. Han, T. J. Santner, and J. J. Rawlinson, "Simultaneous determination of tuning and calibration parameters for computer experiments," *Technometrics*, vol. 51, pp. 464–474, 2009.

[5] D. Chowdhury, L. Santen, and A. Schadschneider, "Statistical physics of vehicular traffic and some related systems," *Physics Reports*, vol. 329, pp. 199–329, 2000.

[6] F. Kessels, *Traffic Flow Modelling*. Springer, 2019, ch. Introduction to Traffic Flow Modelling.

[7] V. Punzo, Z. Zheng, and M. Montanino, "About calibration of car-following dynamics of automated and human-driven vehicles: Methodology, guidelines and codes," *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103165, 2021.

[8] V. Punzo, B. Ciuffo, and M. Montanino, "Can results of car-following model calibration based on trajectory data be trusted?" *Transportation Research Part C: Emerging Technologies*, vol. 2315, pp. 11–24, 2012.

[9] S. P. Hoogendoorn, S. J. L. Ossen, and J. W. C. Van Lint, "Advanced calibration of car-following models," in *Proceedings of the 11th World Conference on Transport Research*, Berkeley, USA, 2007.

[10] M. Treiber and A. Kesting, *Traffic Flow Dynamics. Data, Models and Simulation*. Springer, 2013.

[11] C. P. I. Van Hinsbergen, H. W. C. Van Lint, S. P. Hoogendoorn, and H. J. Van Zuylen, "Bayesian calibration of car-following models," in *Proceedings of the 12th IFAC Symposium on Transportation Systems*, Redondo Beach, USA, 2009.

[12] M. Rahman, D. Chowdhury, T. Khan, and P. Bhavsar, "Improving the efficacy of car-following models with a new stochastic parameter estimation and calibration method," *IEEE Transactions of Intelligent Transportation Systems*, vol. 16, pp. 2687–2699, 2015.

[13] F. Abodo, A. Berthaume, A. Zitzow-Childs, and L. Bobadilla, "Strengthening the case for a bayesian approach to car-following model calibration and validation using probabilistic programming," in *Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Auckland, NZ, 2019.

[14] C. Zhang and L. Sun, "Bayesian calibration of the intelligent driver model," *aeXiv:2210.03571v1*, 2022.

[15] J. Brynjarsdóttir and A. O'Hagan, "Learning about physical parameters: the importance of model discrepancy," *Inverse Problems*, vol. 30, p. 114007, 2014.

[16] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge: MIT Press, 2006.

[17] M. Sederlin, X. Ma, and J. Jin, "A hybrid modelling approach for traffic state estimation at signalized intersections," *IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 3604–3609, 2021.

[18] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. Boca Raton: Chapman and Hall/CRC Press, 2013.

[19] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, "Introducing Markov chain Monte Carlo," in *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds. Boca Raton: CRC Press, 1996, ch. 1.

[20] V. Kyriacou, Y. Englezou, C. Panayiou, and S. Timotheou, "Bayesian traffic state estimation using extended floating car data," *IEEE Transactions of Intelligent Transportation Systems*, vol. 24, pp. 1518 – 1532, 2022.

[21] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wiessner, "Microscopic traffic simulation using SUMO," *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2575–2582, 2018.

[22] D. V. Lindley, "On the measure of information provided by an experiment," *The Annals of Statistics*, vol. 27, pp. 986–1005, 1956.