Transportation Research Part C 139 (2022) 103692

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Physics-augmented models to simulate commercial adaptive cruise control (ACC) systems

Yinglong He [a], Marcello Montanino [b,*], Konstantinos Mattas [c], Vincenzo Punzo [b], Biagio Ciuffo [c]

[a] University of Cambridge, Cambridge CB3 0HA, UK
[b] University of Naples Federico II, Via Claudio, 21, 80125 Napoli, Italy
[c] European Commission – Joint Research Centre, Ispra (VA), Italy

ARTICLE INFO

Keywords:
Adaptive cruise control (ACC)
Car-following
Vehicle dynamics
Linear controller
Constant Time Headway (CTH)
Intelligent Driver Model (IDM)
Gipps' model

ABSTRACT

This paper investigates the accuracy and robustness of car-following (CF) and adaptive cruise control (ACC) models in reproducing measured trajectories of commercial ACCs. To this aim, a general modelling framework is proposed, in which ACC and CF models have been incrementally augmented with physics-based extensions: namely, perception delay, linear or nonlinear vehicle dynamics, and acceleration constraints. This framework has been applied to the Intelligent Driver Model (IDM), Gipps' model, and to three basic ACC algorithms. These are linear controllers which are coupled with a constant time-headway spacing policy, and with two other policies derived from the traffic flow theory: the IDM desired distance function, and Gipps' equilibrium distance-speed function. The ninety models resulting from the combination of the five base models with the aforementioned extensions, have been assessed and compared through a vast calibration and validation experiment against measured trajectory data of vehicles driven by ACC systems. Overall, the study has shown that physics-based extensions provide limited improvements to the accuracy of existing models. In addition, if an investigation against measured data is not carried out, it is not possible to argue which extension is the most suited for a specific model. The linear controller with Gipps' spacing policy has resulted the most accurate model, while the IDM the most robust to different input trajectories. Eventually, all models have failed to capture the behaviour of some car brands – just as models fail with some human drivers. Therefore, the choice of the "best" model is independent of the car brand to simulate.

1. Introduction

Accurate simulation of mixed traffic, including human-driven vehicles (HVs) and automated vehicles (AVs) is crucial to develop cooperative, connected, and automated mobility (CCAM) technologies and assess their impacts on safety, traffic efficiency and related externalities (Alonso Raposo et al., 2019). Car-following models (CF) and adaptive cruise control algorithms (ACC) have been proposed to mimic longitudinal dynamics of HVs and AVs, respectively.

There exists a vast literature on modelling HVs through behavioural CF models. Since the true underlying human driving process is

* Corresponding author.
E-mail addresses: yh496@cam.ac.uk (Y. He), marcello.montanino@unina.it (M. Montanino), konstantinos.mattas@ec.europa.eu (K. Mattas), vincenzo.punzo@unina.it (V. Punzo), biagio.ciuffo@ec.europa.eu (B. Ciuffo).

https://doi.org/10.1016/j.trc.2022.103692
Received 14 July 2021; Received in revised form 12 April 2022; Accepted 12 April 2022
Available online 27 April 2022

unknown, several formulations coexist in the field literature, though based on completely different assumptions (e.g., Wiedemann, 1974; Gipps, 1981; Bando et al., 1995; Treiber et al., 2000). Although theoretical properties, advantages, and drawbacks of these models have been confronted in the literature, and are very well known, it has been never demonstrated the "dominance" of one model over all others i.e., one model outperforming all others in any situation, in terms of accuracy. In fact, previous studies showed that residual errors from calibration against HV trajectories vary from one trajectory to another more than from one model to another (see e.g., Brockfeld et al., 2004; Ranjitkar et al., 2004). Thus, there are trajectories which are reasonably reproduced by most of the models, and others which all models fail to mimic at an acceptable degree of accuracy (Punzo and Simonelli, 2005; Punzo et al., 2021; another investigation also showed that, given a specific model, variation in output errors is due to variation in input trajectories more than to variation in parameters; Punzo et al., 2015). After all, these results are not unexpected, since the complex and heterogeneous mechanisms behind human driving behaviours are only partially captured by CF models. These are simple stimulus–response relationships which describe the driver-vehicle unit motion phenomenologically i.e., without disentangling the human decision-making process from the vehicle control loop (Rothery, 2001; Saifuzzaman and Zheng, 2014). Therefore, CF models do not include lower-level dynamics concerning vehicle and road characteristics like e.g., mechanical actuation delays or motor constraints (with few exceptions, see e.g., Ni, 2011).

Conversely, a number of ACC algorithms have been proposed to mimic longitudinal dynamics of AVs in a traffic simulation, just as CF models have been applied to simulate HVs (see e.g., Darbha et al., 1994; Shladover et al., 2012; Milanés and Shladover 2014, Gunter et al., 2019; Wang et al., 2021; for a review, see Yu et al., 2021). Since ACC algorithms represent only the upper-level logic of complicated controllers, they are usually found coupled with lower-level dynamics at different levels of detail. They include linear and nonlinear vehicle dynamics (Xiao et al., 2017), acceleration constraints (He et al., 2020a) and perception delays (Gunter et al., 2020).

All that said, CF models and ACC algorithms are conceptually equivalent. CF models can be thought of as upper-level controllers (like ACC algorithms), if we assume that the model output is just the desired control signal (acceleration or speed)before the actuation. Conversely, we may think of ACC models as simple CF models. Therefore, from a modelling viewpoint there is no reason to prefer CF models to simulate HVs, and ACC algorithms to simulate AVs. After all, when simulating a mixed traffic, we retain almost the same level of ignorance about the true dynamics of HVs and AVs. We do not know the true logics that inspire human driver behaviours, just as we do not know the ones that are coded in commercial AV controllers (unless manufacturers disclosed their proprietary algorithms). In addition, they both share a high degree of heterogeneity, which is due to the heterogeneity of drivers' performances in CF models, and to the heterogeneity of controller settings among AV manufacturers.[1]

Therefore, once the conceptual equivalence of CF models and ACC algorithms is acknowledged, a number of questions arise. First, one may wonder how much the accuracy of CF models can be improved by augmenting them with perception delays and detailed lower-level dynamics such as actuation lags and vehicle constraints, which are usually applied when simulating AVs (see e.g., Treiber et al., 2006; Makridis et al., 2019; He et al., 2020a). Not considering perception delays and detailed lower-level dynamics has been pointed out as a possible harm to the accuracy and validity of such models (for example, Wang, 2018, showed that both perception delay and vehicle dynamics are key in CF and ACC models to realistically reproduce string instability of traffic. Another example is Treiber et al., 2006, who incorporated some overlooked aspects into CF models e.g., delays and estimation errors, which have significant influence on vehicle behaviour. Eventually, Makridis et al., 2019, demonstrated that coupling traditional CF models and lower-level dynamics produces a more realistic traffic flow behaviour at both microscopic and macroscopic levels).

Second, what is the ability of ACC algorithms from the field literature to capture the actual dynamics of commercial ACC vehicles? Since ACC algorithms are not usually calibrated on observed AVs trajectories, with few exceptions (e.g., Gunter et al., 2019, Wang et al., 2021), the extent to which the dynamics they describe are close to those of the (undisclosed) ACC algorithms in commercial vehicles, is almost unknown. In addition, since ACCs from the literature are coupled with several formulations of lower-level dynamics at different levels of detail, a systematic investigation is necessary to assess the contribution of each physics-based extension to the accuracy of simulation results.

Eventually, are there any significant differences in the ability to capture AVs driving behaviours between the two classes of models – CF and ACC – once they are augmented with the same physics-based extensions?

To answer these questions, this paper develops a comparison framework that allows confronting different models, each one augmented with physics-based extensions of increasing complexity. To this end, all possible combinations of five base models (CF and ACC) with different extensions, including perception delay, linear or nonlinear vehicle dynamics, and acceleration constraints, are calibrated against measured vehicle trajectory data from commercial AVs.[2]

As base models, two CF models, namely the IDM (Treiber et al., 2000) and Gipps' model (Gipps, 1981), and three ACC algorithms have been tested. Applied ACCs consist of a linear controller (e.g., Li et al., 2017) coupled with three different spacing policies. These are a Constant Time Headway (CTH) spacing policy (e.g., Zheng et al., 2016), and other two policies derived from the traffic flow theory: the IDM desired distance function, and Gipps' equilibrium distance-speed function. To the best of our knowledge, the approach of coupling a linear control law with spacing policies derived from the car-following theory is novel in the literature.

As a result, 90 different models (i.e., 5 base models per 18 variants each) have been calibrated, each one against open-source

[1] In a modelling perspective, the only true difference between HVs and AVs is that the logic of AVs is not expected to be stochastic in time, opposite to HVs, see e.g., the stochastic variability of human drivers' reaction times which has motivated a number of studies e.g., Treiber and Kesting, 2018; Tian et al., 2019.

[2] On the contrary of HV trajectories, AV ones should not be stochastic in time. Therefore, model calibration is not expected to be affected by the noise due to this stochasticity.

trajectories gathered on 7 platooning experiments with four high-end vehicles controlled by ACC systems (Audi, Tesla, Mercedes, and BMW), following a platoon leader (Makridis et al., 2021). Models calibrated against a vehicle trajectory in a platoon, have been cross-validated against trajectories from other platoons. While the calibration sets a common and fair basis for model comparison, the cross-validation is necessary to avoid the comparison to be biased by model overfitting.

The proposed methodology has enabled two analyses. The first has aimed at investigating the impact of physics-based extensions on model accuracy, independently of the specific model. The second analysis has been a quantitative comparison of the prediction capability and robustness of the five model classes. As a result, it has been also possible to verify whether the consolidated habit of applying linear control laws – and not CF models – to simulate AVs has significance in practice.

The remainder of the paper is organized as follows. Section 2 describes the unified modelling framework adopted in this study, which includes five base models and three types of physics-based extensions. Section 3 describes the field platoon data applied for model calibration and validation. Section 4 provides the methodology and the design of experiments. Sections 5 and 6 discuss the calibration and validation results, respectively. Eventually, Section 7 draws the study conclusion.

2. Modelling framework

The modelling framework applied in this work is presented in Fig. 1. The framework is applied to both behavioural CF models and ACC algorithms. Through the framework, we have also clarified the equivalence between behavioural CF models and ACC algorithms from a modelling standpoint. Indeed, both types of models are augmented with lower-level dynamics of increasing complexity.

The five base models studied here are the IDM and Gipps' CF model, and three ACC algorithms that adopt a linear control law coupled with a CTH spacing policy (L-CTH), and with two new spacing policies derived from the IDM and Gipps' models (L-IDM and L-Gipps, respectively). Each base model has been formulated in 18 variants, which account for all possible combinations of a base model formulation with physics-based enhancements, including perception delay, vehicle dynamics and acceleration constraints, at increasing level of complexity.

The base model returns the acceleration command $a_{n,cmd}(t)$ to be executed by the ego vehicle n at time t as a function of model parameters and of the following inputs: the ego vehicle speed $v_n(t-\tau_p)$, the inter-vehicle (bumper-to-bumper) spacing $s_n(t-\tau_p) = x_{n-1}(t-\tau_p) - x_n(t-\tau_p) - l_{n-1}$ and the relative speed $\Delta v_n(t-\tau_p) = v_{n-1}(t-\tau_p) - v_n(t-\tau_p)$ from a preceding vehicle $n-1$, where x_i denotes the longitudinal position of vehicle i, l_i the vehicle length and τ_p a perception delay (for τ_p, see Section 2.2.3). The acceleration command is then transformed into a control signal acting on the dynamical system of the ego vehicle (Section 2.2.1). Acceleration constraints are eventually applied, providing the vehicle acceleration ($a_n(t)$) (these constraints are due, for instance, to limited motor acceleration capabilities; see Section 2.2.2).

Ego vehicle kinematics are obtained by coupling the CF/ACC model with the equation of motion. In numerical integration, a ballistic integration scheme is typically adopted (Treiber et al., 2006; Treiber and Kanagaraj, 2015):

$$\begin{cases} v_n(t) = v_n(t-\Delta t) + a_n(t)\cdot\Delta t \\ x_n(t) = x_n(t-\Delta t) + \dfrac{v_n(t)+v_n(t-\Delta t)}{2}\cdot\Delta t \end{cases} \tag{1}$$

where Δt is the numerical integration step. The nomenclature used in the paper is reported in Table 1.

2.1. Base models

In the proposed modelling framework, the base model defines the acceleration command ($a_{n,cmd}$) to the ACC vehicle in response to perceived kinematic states, e.g., s_n and Δv_n. The five base models studied in this work are presented in the next sub-sections.

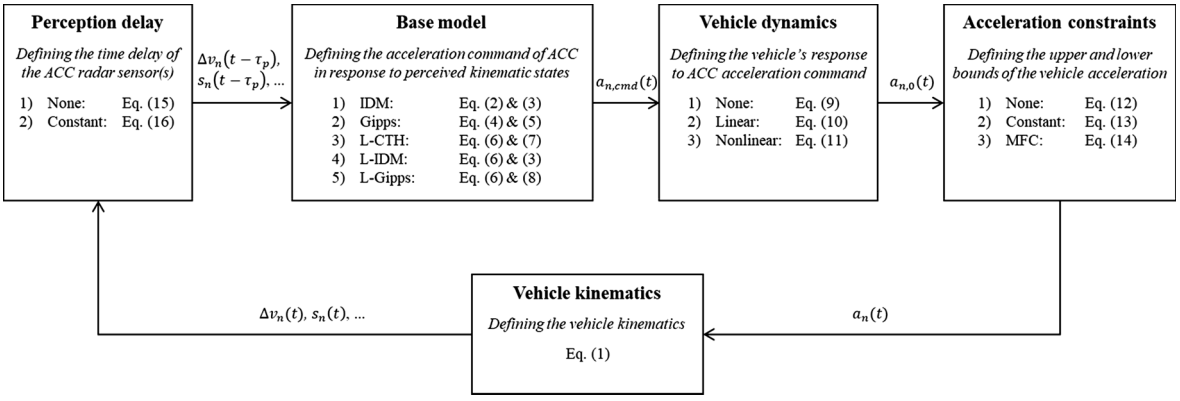


Fig. 1. Unified modelling framework.

Table 1
Nomenclature.

Variables			
t	Time (s)	$\hat{a}_{min}$	The following vehicle estimate of the preceding vehicle deceleration rate (m/s^2, negative)
a_n	Acceleration (m/s^2) of the ego vehicle	k_s	Factor (s^{-2}) for the deviation from the desired spacing
v_n, v_{n-1}	Speed (m/s) of vehicles n (ego) and $n-1$ (preceding)	k_v	Factor (s^{-1}) for the speed difference between vehicles n and $n-1$
x_n, x_{n-1}	Position (m) of vehicles n (ego) and $n-1$ (preceding)	k_0	Factor (s^{-1}) for the deviation from the set speed
s_n	Inter-vehicle spacing (m), $x_{n-1}-x_n-l_{n-1}$	τ_p	Perception delay (s)
$s_{n,des}$	Desired inter-vehicle spacing (m)	τ_a	Actuation lag (s)
Δv_n	Speed difference (m/s),$v_{n-1}-v_n$	f_0, f_1, f_2	Road load coefficients (N, kg/s, and kg/m, respectively)
Δs_n	Spacing error (m), $s_{n,des}-s_n$	r_w	Wheel radius (m)
$a_{n,cmd}$	Acceleration command (m/s^2) of the ACC system	m	Vehicle mass (kg)
$v_{n,cmd}$	Speed command (m/s) of the ACC system	g	Gravitational acceleration (9.81 m/s^2)
$a_{n,0}$	Unconstrained vehicle acceleration (m/s^2)	φ	Equivalent inertial mass factor
F_t, T_t, a_t	Tractive force (N), torque (Nm), and acceleration (m/s^2)	a_{lb}	Lower acceleration bound (m/s^2)
F_r	Total resistance force including aerodynamics, rolling, and grade resistances	a_{ub}	Upper acceleration bound (m/s^2)
α	Road slope (rad)	***Acronyms***	
a_{ap}	Acceleration potential (m/s^2)	ACC	Adaptive Cruise Control
a_{dp}	Deceleration potential (m/s^2, negative)	IDM	Intelligent Driver Model
		CTH	Constant Time Headway
Parameters		MFC	Microsimulation free-flow acceleration boundary model
Δt	Time step (s), 0.1 s	L-CTH	Linear ACC with the spacing policy of Constant Time Headway
l_{n-1}	Length (m) of the preceding vehicle	L-IDM	Linear ACC with the spacing policy of IDM-desired
v_0	Set speed (m/s) of the ACC system	L-Gipps	Linear ACC with the spacing policy of Gipps-equilibrium
s_0	Minimum spacing (m) at standstill condition	GoF	Goodness-of-Fit function
t_h	Time headway (s)	MoP	Measure of Performance
δ	Exponent factor of the IDM model	RMSE	Root Mean Square Error
θ	Delay factor (s) in Gipps' model	NRMSE	Normalized RMSE
a_{max}	Maximum acceleration (m/s^2)	CV	Coefficient of Variation
a_{min}	Comfort deceleration rate (m/s^2, negative)	IQR	Interquartile range

2.1.1. IDM model

The Intelligent Driver Model (IDM) is developed by Treiber et al. (2000) and is widely investigated in the traffic flow theory literature (see e.g., Ward and Wilson, 2011; Treiber and Kesting, 2011; Montanino et al., 2021; Montanino and Punzo, 2021):

$$a_{n,cmd}(t) = a_{max} \cdot \left(1 - \left(\frac{v_n(t)}{v_0}\right)^{\delta} - \left(\frac{s_{n,des}(t)}{s_n(t)}\right)^2\right) \tag{2}$$

where $a_{n,cmd}$ is the acceleration command to be applied by the ego vehicle, i.e., transformed into control signals of the throttle and brake pedals; a_{max} indicates the maximum acceleration rate; v_0 is the desired speed, which represents the set speed of the onboard ACC system in this work; $s_{n,des}$ indicates the desired inter-vehicle spacing; δ is an exponent factor.

The desired spacing policy ($s_{n,des}$) in the IDM model is calculated as:

$$s_{n,des}(t) = s_0 + \max\left(0, t_h v_n(t) - \frac{v_n(t)\Delta v_n(t)}{2\sqrt{-a_{max}a_{min}}}\right) \tag{3}$$

where s_0 is the minimum desired spacing; t_h denotes the desired time headway; a_{min} is the comfort deceleration rate (negative).

2.1.2. Gipps' model

Gipps' model (Gipps, 1981) specifies a safe inter-vehicle spacing and plans the vehicle speed for following instants (for a review see Ciuffo et al., 2012). Therefore, the acceleration command ($a_{n,cmd}$) is derived from the planned and the current speeds:

$$a_{n,cmd}(t) = \left(v_{n,cmd}(t+t_h) - v_n(t)\right)/t_h \tag{4}$$

where t_h is an "apparent" reaction time, and $v_{n,cmd}$ denotes the safe speed planned by Gipps' model that considers two driving regimes. The former is a free-flow regime where a vehicle accelerates to its desired speed, following a pre-defined acceleration pattern (a generalized acceleration pattern can be found in Ciuffo et al., 2012). The latter is a car-following regime where the vehicle plans a speed value which allows for a safe stop even in the occurrence of a hard deceleration of the preceding vehicle. The vehicle chooses the minimum between these two speed values:

$$v_{n,cmd}(t+t_h) = \min \left(\begin{array}{c} v_n(t) + 2.5 a_{max} t_h \cdot \left(1 - \frac{v_n(t)}{v_0}\right)\left(0.025 + \frac{v_n(t)}{v_0}\right)^{0.5}, \\ a_{min} \cdot \left(\frac{t_h}{2} + \theta\right) + \sqrt{{a_{min}}^2 \cdot \left(\frac{t_h}{2} + \theta\right)^2 - a_{min} \cdot \left(2(s_n(t) - s_0) - t_h v_n(t) - \frac{v_{n-1}(t)^2}{\widehat{a}_{min}}\right)} \end{array} \right) \tag{5}$$

where a_{max} indicates the maximum acceleration rate; a_{min} is the comfort deceleration rate (negative); $\widehat{a}_{min}$ is the following vehicle estimate of the preceding vehicle deceleration rate (negative); s_0 is the minimum desired spacing; v_0 is the desired speed; θ is an additional delay inserted to avoid braking always at the maximum rate.

2.1.3. Linear control law

Linear control laws are commonly used in studies of ACC and cooperative ACC, as these models can significantly simplify theoretical analyses (Zheng et al., 2016; Li et al., 2017; He et al., 2020c). In this work, we have studied three control algorithms, which couple the linear control law with different spacing policies. The adopted formulation of the linear control law is (VanderWerf et al., 2001; Shladover et al., 2012):

$$\begin{cases} a_{n,cmd}(t) = \min\{k_v \cdot \Delta v_n(t) - k_s \cdot \Delta s_n(t), k_0 \cdot (v_0 - v_n(t))\} \\ \Delta s_n(t) = s_{n,des}(t) - s_n(t) \end{cases} \tag{6}$$

where $s_{n,des}$ is the desired inter-vehicle spacing; Δs_n denotes the error between the desired ($s_{n,des}$) and the actual (s_n) spacing values; k_s is the gain factor for the spacing error; k_v is the gain factor for the speed difference between the preceding vehicle and the ego vehicle; k_0 is the gain factor for the error between the ACC set speed (v_0) and ego vehicle speed (v_n).

The desired spacing is defined through a spacing policy. Many researchers have studied spacing policies from microscopic or macroscopic perspectives, aiming to improve e.g., car-following safety, driver acceptance, stability performance (see e.g., Zhang et al., 2019; Wu et al., 2020). In this study, for the sake of comparison, three spacing policies are coupled with the linear control law in Eq. (6).

First, a constant time headway (CTH) spacing policy calculates the desired spacing as a linear function of the ego vehicle speed (v_n), in which the slope and the intercept of the linear function are the time headway (t_h) and the standstill spacing (s_0), respectively, i.e.,

$$s_{n,des}(t) = s_0 + t_h v_n(t) \tag{7}$$

In addition, two new spacing policies have been tested in this work. They are here referred to as IDM-desired and Gipps-equilibrium spacing policies. These two spacing policies for a linear controller are a novel contribution to the ACC field literature. The former is given by Eq. (3) as per the IDM. The latter implements the distance-speed function at equilibrium of Gipps' model, which has been derived in Punzo and Tripodi (2007):

$$s_{n,des}(t) = s_0 + (t_h + \theta) v_n(t) - 0.5 v_n(t)^2 \left(\frac{1}{a_{min}} - \frac{1}{\widehat{a}_{min}}\right) \tag{8}$$

For simplicity, the linear control law i.e., Eq. (6), coupled with the spacing policies described in Eq. (7), (3), and (8), are hereafter abbreviated as L-CTH, L-IDM, and L-Gipps, respectively.

2.2. Physics-based extensions

Most of the studies on behavioural CF models do not consider lower-level dynamics, with some exceptions (e.g., Ni, 2011; He et al., 2020a). On the contrary, most of ACC studies account for these dynamics at different levels of detail. More specifically, there are three essential aspects influencing vehicle behaviour – namely, vehicle dynamics (e.g., actuation lag), acceleration constraints, and perception delay – that are often considered in ACC control design but not captured by behavioural CF models.

Two types of vehicle dynamics, i.e., linear and nonlinear are applied in the literature, striking a balance between simplicity and accuracy (Xiao et al., 2017). Linear vehicle dynamics introduce a first-order lag between the acceleration command generated by the ACC and the actual acceleration of the ego vehicle. They only represent the response of the driveline system (VanderWerf et al., 2001; Jia and Ngoduy, 2016). Nonlinear vehicle dynamics can simulate both internal (e.g., driveline) and external (e.g., resistances of aerodynamic, rolling, and slope) factors (e.g., Zheng et al., 2017).

Acceleration constraints are specified to avoid unrealistic accelerations beyond vehicle practical capabilities. In most models, the upper and lower acceleration bounds are either set to constant values or not implemented. For example, the IDM only has a constant upper bound for acceleration (a_{max}), and no lower bound, thus potentially leading to unrealistic deceleration (e.g., in cut-in manoeuvres; Kesting et al., 2010). Similar issues are observed with Gipps' model.

In the real world, upper and lower acceleration bounds are not constant and are affected by many factors such as the engine, the braking system, and the road loads (Mannering and Washburn, 2020). Consequently, some CF models try to account for adaptive acceleration constraints i.e., varying with the speed (see e.g., Rakha, 2009, Fadhloun and Rakha, 2020, and the microsimulation free-

flow acceleration (MFC) model, Makridis et al., 2019; for a review see He et al., 2020a).

Perception delay represents the dynamics of human senses or automated sensors in data gathering and processing. It plays an important role in traffic properties e.g., string stability (Xiao and Gao, 2011). For example, experimental studies exhibit that linear ACC laws with perception delay can accurately capture the string instability of ACC platoons in field tests (e.g., Gunter et al., 2020). Also CF models have been enhanced to include a perception delay (e.g., Treiber et al., 2006).

In the following subsections, the formulations of the physics-based extensions studied in this paper are presented.

2.2.1. Vehicle dynamics (VD)

Longitudinal vehicle dynamics define how the vehicle responds to the acceleration command generated by the ACC (Li et al., 2017). The types of vehicle dynamics applied in this study can be categorized as follows:

1) None: no vehicle dynamics are applied, i.e., the resulting ego vehicle acceleration ($a_{n,0}$, unconstrained) coincides with the ACC acceleration command (Li et al., 2017):

$$a_{n,0}(t) = a_{n,cmd}(t) \tag{9}$$

2) Linear: it captures the actuation lag (τ_a) of a driveline system, and it is expressed as (VanderWerf et al., 2001; Li et al., 2017):

$$\tau_a \dot{a}_{n,0}(t) + a_{n,0}(t) = a_{n,cmd}(t) \tag{10}$$

3) Nonlinear: it reproduces nonlinear features, including internal driveline system and external road loads. The corresponding vehicle dynamics are depicted by the following equations, which are not only widely adopted in the literature (e.g., Li et al., 2017; Kühlwein 2016; Küng et al., 2019; Fiori et al., 2021), but also extensively validated through chassis dynamometer tests in our previous works (e.g., Tsiakmakis et al., 2017; Makridis et al., 2019; He et al., 2020a):

$$\begin{cases} \tau_a \dot{a}_t(\mathrm{t}) + a_t(\mathrm{t}) = a_{n,cmd}(\mathrm{t}) \\ F_t(t) = \dfrac{T_t(t)}{r_w} = m a_t(\mathrm{t}) \\ F_r(t) = f_0 \cos\alpha(t) + f_1 v_n(t) + f_2 v_n(t)^2 + mg\sin\alpha(\mathrm{t}) \\ a_{n,0}(t) = \dfrac{F_t(t) - F_r(t)}{\varphi m} \end{cases} \tag{11}$$

where F_r is the total resistance force, which includes aerodynamics, rolling and grade resistances; f_0, f_1, and f_2 are the road load coefficients (He et al., 2020b); α is the road slope; g is the gravitational acceleration; F_t, T_t, and a_t are, respectively, the tractive force, the tractive torque, and the acceleration acting on the wheels; r_w is the wheel radius; τ_a is the actuation lag; φ is the equivalent inertial mass factor; m represents the operating mass, and it is equal to the sum of the unladen mass (m_0) and of the mass of vehicle loads (m_{load}) (e.g., passengers and measuring instruments).

2.2.2. Acceleration constraints (AC)

Acceleration constraints define the maximum capabilities of vehicle propulsion and braking systems in various conditions. Three types of acceleration constraints are studied:

1) None: no physical constraint is applied to the ego vehicle acceleration:

$$a_n(t) = a_{n,0}(t) \tag{12}$$

2) Constant: it assumes that the ego vehicle acceleration is always bounded by constant values:

$$a_n(t) = \max\left\{a_{lb}, \min\left\{a_{n,0}(t), a_{ub}\right\}\right\} \tag{13}$$

where a_{lb} and a_{ub} are, respectively, the lower and upper bounds of the ego vehicle acceleration. They have been set to -7 and 5 m/s^2 in this work, basing on the maximum values of acceleration recorded in the car-following database OpenACC (Makridis et al., 2021).

3) Microsimulation free-flow acceleration (MFC) boundary model: the MFC model serves to define dynamic bounds for the ego vehicle acceleration, in which both lower and upper bounds change with the vehicle speed $v_n(t)$. The constraints are described by:

$$a_n(t) = \max\left\{a_{dp}(v_n(t)), \min\left\{a_{n,0}(t), a_{ap}(v_n(t))\right\}\right\} \tag{14}$$

where a_{ap} and a_{dp} are nonlinear curves representing acceleration and deceleration potentials of the ego vehicle, respectively. Their formulations depend on the vehicle powertrain type and are elaborated in previous studies for internal combustion engine vehicles, ICEVs (Makridis et al., 2019), and for electrified vehicles, EVs (He et al., 2020a).

2.2.3. Perception delay (PD)

Perception delay is a significant phenomenon affecting AV dynamics, and it is caused by sensors and/or V2X technologies. To capture characteristics of ACC radar sensors, the parameter τ_p is introduced in the modelling framework. It represents the time delay of sensor perception. Consequently, in all state variables and inputs in Eq. (2), (3), (5)–(8), i.e., in $v_n(t)$, $v_{n-1}(t)$, $\Delta v_n(t)$, $s_n(t)$, $s_{n,des}(t)$, and $\Delta s_n(t)$, the time instant t is replaced with the delayed time instant, $t-\tau_p$. This study provides two options for the perception delay:

1) None:

$$\tau_p = 0 \tag{15}$$

2) Constant:

$$\tau_p > 0 \tag{16}$$

2.3. Model summary

Different combinations of physics-based extensions can have different effects on model performances. Therefore, the full list of tested models is summarized in Table 2. Ninety models are derived by combining the five base models with three vehicle dynamics formulations, three acceleration constraint functions, and two perception delay types, in a full factorial design. In the table, model variants relative to the same base model are sorted by increasing level of complexity.

3. Field platoon data

Vehicle trajectory data used in this study were collected in platooning experiments carried out on the Rural Road of the AstaZero proving ground (Sweden), as indicated by the black curve in Fig. 2 (a). The test track is 5.7-km long, with slopes ranging from −3 to 3%, as presented in Fig. 2 (b).

The tested platoon was composed by five high-end vehicles from different manufacturers (Audi, Tesla, Mercedes, and BMW), whose longitudinal dynamics were fully controlled by the onboard commercial ACC systems. Trajectory data were acquired with the RT-Range S multiple target ADAS measurements solution, supplied by Oxford Technical Solutions Company. The system has a sampling rate of 100 Hz (trajectory data were down sampled to 10 Hz), and guarantees precisions of 0.02 m/s and 0.02 m for speed and position measurements, respectively (Makridis et al., 2021).

Table 3 lists the vehicle compositions (platoon leader and four followers) in 7 platoons (some platoons shared the same composition). Data analysis has shown that the platoon was in car-following regime during the whole experiment, i.e., the follower vehicle dynamics were influenced more by the preceding vehicle trajectory, vehicle characteristics and onboard ACC system setup, than by road geometry or slope.

Fig. 3 illustrates vehicle trajectory data of 7 platoons tested on the AstaZero proving ground. Fig. 3 (a)-(g) provide the speed profiles of the platoon leader and four following vehicles. Fig. 3 (h)-(n) give the corresponding spacing profiles. In the tests, the platoon leader (always an Audi A8), was instructed to keep a predefined speed value (via the onboard ACC system), in order to prevent speed fluctuations caused by human driving. Speed perturbations were then introduced in the platoon by changing the set speed (v_0) of the leading vehicle ACC system. For safety purposes, the ACC set speed lied between 13.9 and 16.7 m/s in curves, and between 25 and 27.8 m/s in straight sections.

4. Methodology and design of experiments

In order to appraise the impact of physics-based extensions on model ability to reproduce observed ACC vehicle dynamics, all models listed in Table 2 have been calibrated against ACC trajectory data.

Calibration of models allows us to compute the optimal parameters values, that is the values which maximise model performances or, better, minimize the error distance between observed and simulated vehicle kinematics. Therefore, calibration is necessary to compare models on a common and fair basis.

However, depending on the specific model, calibration may result in the overfitting of parameters to data: good model performances in calibration might be just the result of the overfitting of a "poor" model to data. Consequently, comparison of models only based on their calibration performances can be misleading, since biased by parameters compensation and overfitting (for a comprehensive discussion on this topic see section 4 "Error compensation and overfitting" in Punzo and Montanino, 2020). To have an unbiased and fair comparison of models, *cross-validation* is necessary (see Punzo and Simonelli, 2005). In a cross-validation, a model is simulated using input trajectories which are different from those used to calibrate the model. Therefore, cross-validation performances truly tell the ability of a model to capture car-following dynamics.

Table 2
Configurations of investigated models.

Model groups	Model ID [b]	ACC controller	Spacing policy	Perception delay (PD) [a]	Vehicledynamics (VD) [a]	AccelerationConstraints (AC) [a]	Calibration parameters
IDM-based models	1	IDM	IDM-desired	None	None	None	$\delta, v_0, s_0, t_h, a_{max}, a_{min}$
	2	IDM	IDM-desired	None	None	Constant	$\delta, v_0, s_0, t_h, a_{max}, a_{min}$
	3	IDM	IDM-desired	None	None	MFC	$\delta, v_0, s_0, t_h, a_{max}, a_{min}$
	4	IDM	IDM-desired	None	Linear	None	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	5	IDM	IDM-desired	None	Linear	Constant	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	6	IDM	IDM-desired	None	Linear	MFC	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	7	IDM	IDM-desired	None	Nonlinear	None	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	8	IDM	IDM-desired	None	Nonlinear	Constant	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	9	IDM	IDM-desired	None	Nonlinear	MFC	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	10	IDM	IDM-desired	Constant	None	None	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p$
	11	IDM	IDM-desired	Constant	None	Constant	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p$
	12	IDM	IDM-desired	Constant	None	MFC	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p$
	13	IDM	IDM-desired	Constant	Linear	None	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	14	IDM	IDM-desired	Constant	Linear	Constant	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	15	IDM	IDM-desired	Constant	Linear	MFC	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	16	IDM	IDM-desired	Constant	Nonlinear	None	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	17	IDM	IDM-desired	Constant	Nonlinear	Constant	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	18	IDM	IDM-desired	Constant	Nonlinear	MFC	$\delta, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
Gipps-based models	19	Gipps	*inherent*	None	None	None	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}$
	20	Gipps	*inherent*	None	None	Constant	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}$
	21	Gipps	*inherent*	None	None	MFC	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}$
	22	Gipps	*inherent*	None	Linear	None	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_a$
	23	Gipps	*inherent*	None	Linear	Constant	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_a$
	24	Gipps	*inherent*	None	Linear	MFC	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_a$
	25	Gipps	*inherent*	None	Nonlinear	None	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_a$
	26	Gipps	*inherent*	None	Nonlinear	Constant	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_a$
	27	Gipps	*inherent*	None	Nonlinear	MFC	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_a$
	28	Gipps	*inherent*	Constant	None	None	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p$
	29	Gipps	*inherent*	Constant	None	Constant	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p$
	30	Gipps	*inherent*	Constant	None	MFC	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p$
	31	Gipps	*inherent*	Constant	Linear	None	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	32	Gipps	*inherent*	Constant	Linear	Constant	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	33	Gipps	*inherent*	Constant	Linear	MFC	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	34	Gipps	*inherent*	Constant	Nonlinear	None	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	35	Gipps	*inherent*	Constant	Nonlinear	Constant	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	36	Gipps	*inherent*	Constant	Nonlinear	MFC	$\theta, v_0, s_0, t_h, a_{max}, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
L-CTH-based models	37	Linear contr.	CTH	None	None	None	$k_s, k_v, k_0, v_0, s_0, t_h$
	38	Linear contr.	CTH	None	None	Constant	$k_s, k_v, k_0, v_0, s_0, t_h$
	39	Linear contr.	CTH	None	None	MFC	$k_s, k_v, k_0, v_0, s_0, t_h$
	40	Linear contr.	CTH	None	Linear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_a$
	41	Linear contr.	CTH	None	Linear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_a$
	42	Linear contr.	CTH	None	Linear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_a$
	43	Linear contr.	CTH	None	Nonlinear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_a$
	44	Linear contr.	CTH	None	Nonlinear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_a$
	45	Linear contr.	CTH	None	Nonlinear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_a$
	46	Linear contr.	CTH	Constant	None	None	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p$
	47	Linear contr.	CTH	Constant	None	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p$
	48	Linear contr.	CTH	Constant	None	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p$
	49	Linear contr.	CTH	Constant	Linear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p, \tau_a$
	50	Linear contr.	CTH	Constant	Linear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p, \tau_a$
	51	Linear contr.	CTH	Constant	Linear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p, \tau_a$
	52	Linear contr.	CTH	Constant	Nonlinear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p, \tau_a$
	53	Linear contr.	CTH	Constant	Nonlinear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p, \tau_a$
	54	Linear contr.	CTH	Constant	Nonlinear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \tau_p, \tau_a$
L-IDM-based models	55	Linear contr.	IDM-desired	None	None	None	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}$
	56	Linear contr.	IDM-desired	None	None	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}$
	57	Linear contr.	IDM-desired	None	None	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}$
	58	Linear contr.	IDM-desired	None	Linear	None	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	59	Linear contr.	IDM-desired	None	Linear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	60	Linear contr.	IDM-desired	None	Linear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	61	Linear contr.	IDM-desired	None	Nonlinear	None	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	62	Linear contr.	IDM-desired	None	Nonlinear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	63	Linear contr.	IDM-desired	None	Nonlinear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_a$
	64	Linear contr.	IDM-desired	Constant	None	None	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p$
	65	Linear contr.	IDM-desired	Constant	None	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p$
	66	Linear contr.	IDM-desired	Constant	None	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p$
	67	Linear contr.	IDM-desired	Constant	Linear	None	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	68	Linear contr.	IDM-desired	Constant	Linear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$

(*continued on next page*)

Table 2 (*continued*)

Model groups	Model ID [b]	ACC controller	Spacing policy	Perception delay (PD) [a]	Vehicledynamics (VD) [a]	AccelerationConstraints (AC) [a]	Calibration parameters
	69	Linear contr.	IDM-desired	Constant	Linear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	70	Linear contr.	IDM-desired	Constant	Nonlinear	None	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	71	Linear contr.	IDM-desired	Constant	Nonlinear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
	72	Linear contr.	IDM-desired	Constant	Nonlinear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, a_{max}, a_{min}, \tau_p, \tau_a$
L-Gipps-based models	73	Linear contr.	Gipps-equilibrium	None	None	None	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}$
	74	Linear contr.	Gipps-equilibrium	None	None	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}$
	75	Linear contr.	Gipps-equilibrium	None	None	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}$
	76	Linear contr.	Gipps-equilibrium	None	Linear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_a$
	77	Linear contr.	Gipps-equilibrium	None	Linear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_a$
	78	Linear contr.	Gipps-equilibrium	None	Linear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_a$
	79	Linear contr.	Gipps-equilibrium	None	Nonlinear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_a$
	80	Linear contr.	Gipps-equilibrium	None	Nonlinear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_a$
	81	Linear contr.	Gipps-equilibrium	None	Nonlinear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_a$
	82	Linear contr.	Gipps-equilibrium	Constant	None	None	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p$
	83	Linear contr.	Gipps-equilibrium	Constant	None	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p$
	84	Linear contr.	Gipps-equilibrium	Constant	None	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p$
	85	Linear contr.	Gipps-equilibrium	Constant	Linear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	86	Linear contr.	Gipps-equilibrium	Constant	Linear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	87	Linear contr.	Gipps-equilibrium	Constant	Linear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	88	Linear contr.	Gipps-equilibrium	Constant	Nonlinear	None	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	89	Linear contr.	Gipps-equilibrium	Constant	Nonlinear	Constant	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$
	90	Linear contr.	Gipps-equilibrium	Constant	Nonlinear	MFC	$k_s, k_v, k_0, v_0, s_0, t_h, \theta, a_{min}, \widehat{a}_{min}, \tau_p, \tau_a$

[a] Model extensions that capture characteristics of perception delay (PD), vehicle dynamics (VD), and acceleration constraints (AC). As shown in Fig. 1, each extension can use different submodels, e.g., vehicle dynamics (VD) can adopt Eq. (9), (10), or (11).

[b] All possible combinations of the acceleration model and the extensions, namely, the base model (i.e., ID = 1, 19, 37, 55, and 73, not considering any of the extensions) and their extended variants (i.e., IDs = 2–18, 20–36, 38–54, 56–72, and 74–90, respectively).

It is also worth noting that, when the leader's trajectory that feeds the model is different from that used in calibration, a collision can occur (this may happen when model calibration returns string unstable parameters, see e.g., Montanino and Punzo, 2021; Montanino et al., 2021). The number of *collisions* occurring in cross-validation simulations is therefore another important indication of model robustness, and transferability of calibrated parameters.

These are the basic concepts that have inspired the methodology and the analyses in this study.

4.1. Methodology

A general formulation of the model calibration problem is described as follows (Punzo et al., 2021):

$$\begin{aligned} & MoP^{sim} = F(\boldsymbol{\beta}) \\ & minimize \ f\left(MoP^{obs}, MoP^{sim}\right) \\ & subject \ to: LB_{\beta} \le \boldsymbol{\beta} \le UB_{\beta}, \boldsymbol{G}(\boldsymbol{\beta}) \le 0 \end{aligned} \tag{17}$$

where $\boldsymbol{\beta}$ is a vector of model parameters to be calibrated (the number of parameters can change depending on the model variant; see Table 2); $F(\bullet)$ is a CF/ACC model variant, that is a function of $\boldsymbol{\beta}$; LB_{β} and UB_{β} represent the lower and the upper bound for the parameters in $\boldsymbol{\beta}$, respectively; $\boldsymbol{G}(\bullet)$ is a vector of constraint functions; $\boldsymbol{MoP^{obs}}$ and $\boldsymbol{MoP^{sim}}$ represent observed and simulated Measure of Performance, respectively;$\boldsymbol{f}(\bullet)$ is a goodness-of-fit (*GoF*) function; and $f(MoP^{obs}, MoP^{sim})$ is the optimization objective function. The aim of CF model calibration is to minimize $f(MoP^{obs}, MoP^{sim})$.

Basing on the findings in Punzo et al. (2021), the Normalized Root Mean Square Error of spacing, speed, and acceleration, i.e., *NRMSE*(*s*,*v*,*a*), has been adopted as *GoF* function in model calibration. In that paper, *NRMSE*(*s*,*v*,*a*) was proved to be the most preferable objective function to be used in a CF calibration problem, basing on theoretical arguments and wide empirical evidence. This GoF function has been here preferred to the *NRMSE*(*s*,*v*) since the acceleration data applied here were not noisy.[3] The formulation of *NRMSE*(*s*,*v*,*a*) is the following:

[3] In addition, the acceleration data applied here are also "internally consistent" (see Punzo et al., 2021), since they have been derived from gathered speeds.

$$\begin{aligned} &GoF:\ f = \mathrm{NRMSE}(Y) \\ &MoP:\ Y = \{s, v, a\} \end{aligned} \tag{18a}$$

$$\begin{cases} \mathrm{NRMSE}(s,v,a) = \gamma_s \mathrm{NRMSE}(s) + \gamma_v \mathrm{NRMSE}(v) + \gamma_a \mathrm{NRMSE}(a) \\ \mathrm{NRMSE}(Y_i) = \mathrm{RMSE}(Y_i) \Big/ \sqrt{\dfrac{1}{T-\tau_p} \sum_{\tau_p+\Delta t}^{T} \left(Y_i^{\mathrm{obs}}(t)\right)^2} \\ \mathrm{RMSE}(Y_i) = \sqrt{\dfrac{1}{T-\tau_p} \sum_{\tau_p+\Delta t}^{T} \left(Y_i^{\mathrm{sim}}(t) - Y_i^{\mathrm{obs}}(t)\right)^2} \end{cases} \tag{18b}$$

where γ_s, γ_v, and γ_a are weight factors, which are assumed equal to 1 in this study (i.e., model accuracy on each *MoP* is equally desirable); T is the total number of measurement instants in an experimental trajectory; τ_p is the model perception delay; Δt is the simulation time step.

In order to assess the ability of the whole calibration setting (optimization algorithm and *GoF*(*MoPs*) function) to find robust solutions in terms of *GoF* value, 10 replications of each model calibration experiment have been performed. By comparing *GoF* value among replications, we have verified whether the algorithm was able to converge to the same minimum at each calibration replication. The optimization algorithm used in this study is the Genetic Algorithm coded in Python (https://pypi.org/project/geneticalgorithm/).

Eventually, to test the robustness of the chosen calibration *GoF*, against alternative *GoF*s, models have been calibrated also by means of the following *GoF*s: *RMSE*(*s*), *RMSE*(*v*), Theil's *U*(*s*,*v*), Theil's *U*(*s*,*v*,*a*), and *NRMSE*(*s*,*v*) (for their definitions, see Punzo et al., 2021).

As mentioned, *cross-validation* consists in simulating the model using input trajectories other than those applied to calibrate it. To evaluate performances in validation experiments the *NRMSE*(*s*,*v*,*a*) has been applied, coherently with the calibration experiments.

4.2. Design of experiments

The calibration and validation experiments have been conducted for the 90 models listed in Table 2, against trajectory data of 4 following vehicles, in 7 platoons.

To allow a better understanding of the design of experiments, an index is introduced to represent both calibration and validation experiments, (*Veh*, P_{cal}, P_{val}), where *Veh* ∈ {Tesla, BMW, Audi A6, Mercedes} is the vehicle type; P_{cal} and P_{val} ∈ {P1, P2, P3, P4, P5, P6, P7} denote the ID of the platoon data used in calibration and validation, respectively. For example, the experiment with index (*Tesla*, *P2*, *P6*) means that model parameters have been calibrated against the trajectory of the vehicle Tesla in the platoon P2, and that resulting optimal parameters have been adopted to simulate the trajectory of the same vehicle in platoon P6.

Therefore, experiments with $P_{cal} = P_{val}$ denote *calibration* experiments, which are 28 per model, i.e., 4 vehicles × 7 platoons. A total of 2520 calibration experiments (28 × 90 models) have been performed.

Experiments with $P_{cal} \neq P_{val}$ denote *cross-validation*, which are 168 per model, i.e., 28 sets of calibrated parameters (4 vehicles × 7 platoons) × 6 simulations (where 6 is the number of platoons in which the same vehicle is observed, other than P_{cal}).

Each simulation has been performed following the guidelines provided in Punzo et al. (2021), i.e.:

- Time step consistency: the time step of car-following model simulation is consistent with the resolution of trajectory data (it is assumed equal to 0.1 s in this work).
- Data internal consistency: since available data consisted in GPS positions, ego vehicle and preceding vehicle speeds and accelerations have been derived using the same integration scheme applied for the model numerical integration (see Eq. (1)).

Table 4 lists the lower and upper bounds of calibration parameters, shared among the 90 models (for parameters definition, see also Table 1). Please note that the road load coefficients have been estimated by chassis dynamometer tests (He et al., 2020b), and are not included in the calibration.[4]

5. Calibration results

5.1. Robustness of the calibration setting

In CF model calibration, the strong nonlinearity of model response surface, the large-dimensional solution space and the stochasticity of some optimization algorithms, makes the algorithm likely to be trapped in a local minimum, thus resulting in poor reproducibility of calibration results.

[4] We have also verified that, if road load coefficients are calibrated (with bounds centered on the values obtained from dynamometer tests ± 20%), calibration (and validation) results do not change. In general, if dynamometer test data are unavailable, one could calibrate these parameters together with those listed in Table 4 (assuming bounds centered on values available from the literature), and verify the degree of model overfitting through the validation methodology proposed here.

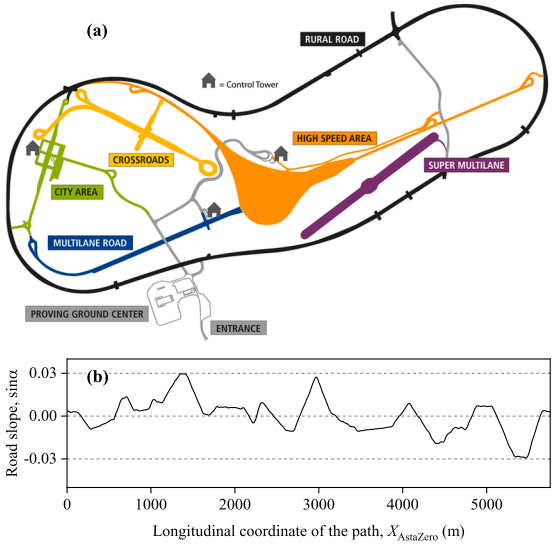


Fig. 2. AstaZero proving ground in Sweden: (a) path layout and (b) road slope. (https://www.astazero.com).

Table 3
Composition of platoons.

Vehicle position	Platoons						
	P1	P2	P3	P4	P5	P6	P7
Leader	Audi A8	Audi A8	Audi A8	Audi A8	Audi A8	Audi A8	Audi A8
Follower 1	Audi A6	Audi A6	Tesla	Tesla	Tesla	Mercedes	Mercedes
Follower 2	BMW	BMW	BMW	BMW	BMW	BMW	BMW
Follower 3	Mercedes	Mercedes	Audi A6	Audi A6	Audi A6	Tesla	Tesla
Follower 4	Tesla	Tesla	Mercedes	Mercedes	Mercedes	Audi A6	Audi A6

To address this concern, we have run 10 replications of each calibration experiment, and we have adopted the coefficient of variation (CV) of the $NRMSE(s,v,a)$ calculated at the optimum as a measure of relative variability across the replications. The analysis of CV distribution of all calibration experiments (i.e., 2520 calibration experiments) has allowed us to assess the ability of a calibration setting to find robust solutions in terms of *GoF* value. When the CV distribution results highly concentrated around the zero, i.e., when the standard deviation of the *GoF* values among replications is very close to 0 in each experiment, calibration results do not change when running the same experiment multiple times.

CV distributions are shown in Fig. 4, which shows CV distributions grouped by model class. In each subplot, the distribution of 504 CV values is presented, each value measuring the variation across 10 replications of each of the 504 calibration experiments (18 models $\times$ 4 vehicle types $\times$ 7 platoon trajectories).

Focusing on the CV variability range, findings suggest that *GoF* values at optimum, for all five model classes, are highly reproducible (CVs are always less than 3%). To infer on zero-value concentration, data have been also fitted by an exponential distribution with a decay rate parameter λ, which measures how rapidly the density declines as the CV value increases. IDM-based models stand out prominently in terms of calibration results reproducibility, showing a λ value which is almost one order of magnitude higher than the λ values of the other model classes.

Concerning the choice of the objective function, we have verified that model calibrations using the $NRMSE(s,v,a)$ have returned the lowest sum of relative errors on spacing, speed and acceleration, compared to calibrations experiments with other objective functions, i.e., $RMSE(s)$, $RMSE(v)$, Theil's $U(s,v)$, Theil's $U(s,v,a)$, and $NRMSE(s,v)$. The relative error measures the degradation in a *MoP* – relative to its optimum – of a model calibrated using a certain objective function. For more details on the methodology to compare objective functions, the reader may refer to Punzo et al. (2021).

This result confirms the recommendation provided in that paper about calibration settings, and extends the validity of the recommendation to a considerably larger number of models (90 in this study) and to a larger dataset of ACC trajectories.

5.2. Comparison of base models across vehicles and platoons

Fig. 5 compares the errors of the five base models after calibration against the trajectories of the 4 vehicles in each of the 7 platoons. Each sub-plot shows errors variability across the 4 vehicles in a specific platoon (sorted according to Table 3). For a given vehicle, the first and the last bar (white and gold bars), show the minimum and maximum $NRMSE(s,v,a)$ values achieved after calibration among all

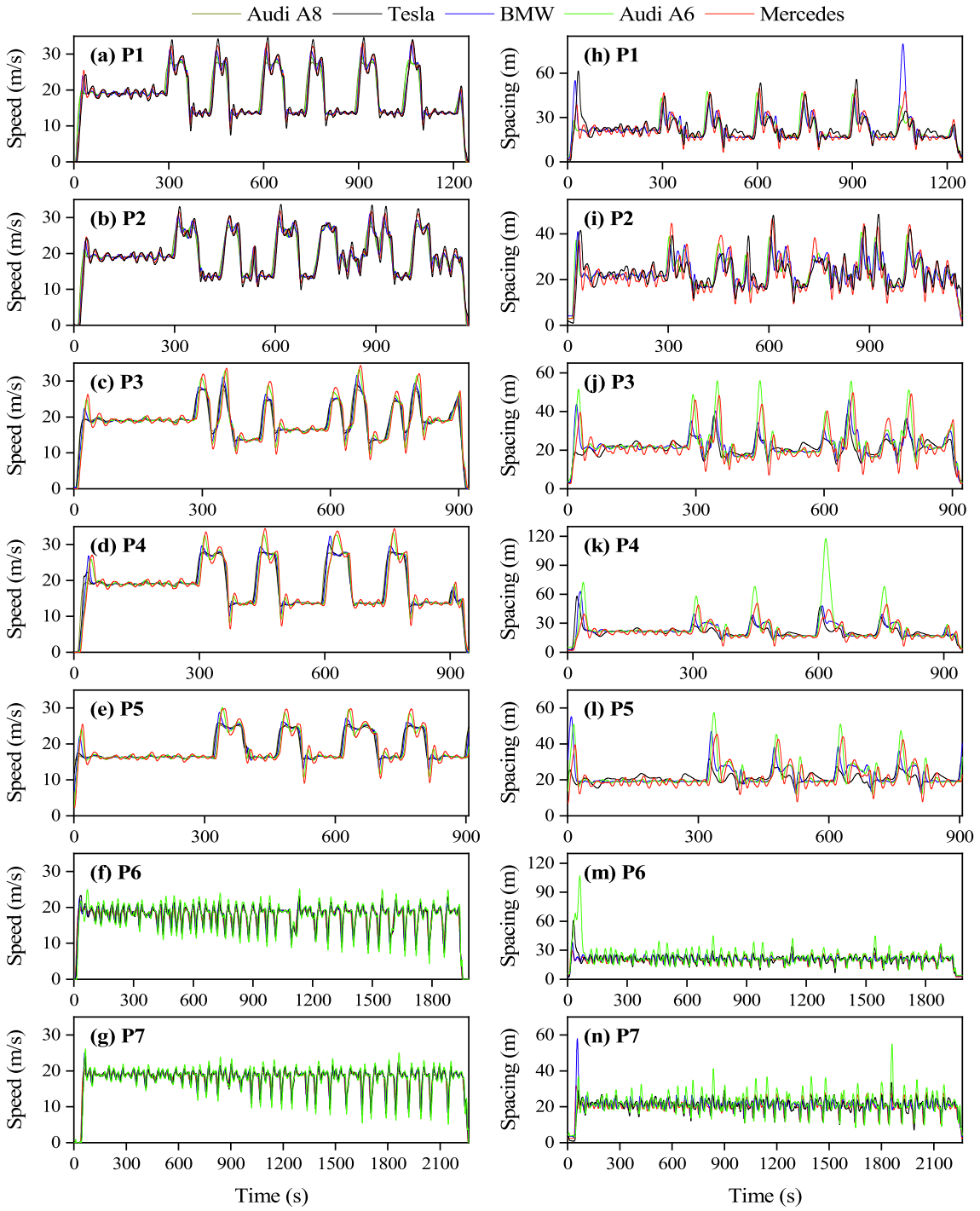


Fig. 3. Field data of seven platoons tested on the AstaZero proving ground: (a)-(g) speed (m/s) and (h)-(n) spacing (m).

Table 4
Bounds of calibration parameters.

Parameter [unit]	Lower bound	Upper bound
δ	0.1	10
v_0 [m/s]	30	35
s_0 [m]	1	5
t_h [s]	0.1	3
a_{max} [m/s^2]	0.5	5
a_{min} [m/s^2]	−5	−0.5
$\hat{a}_{min}$ [m/s^2]	−5	−0.5
θ [s]	0	3
τ_a [s]	0.3	0.8
τ_p [s]	0.1	0.8
k_s [s^{-2}]	0.01	5
k_v [s^{-1}]	0.01	5
k_0 [s^{-1}]	0.01	5

ninety models. They are the best/worst values that can be achieved with either base models or their variants. The five bars in the middle of each group tell the *NRMSE* obtained from the calibration of the five base models.

Given a vehicle, model performances sensibly vary among platoons, depending on the observed trajectories of the vehicle itself and of its immediate leader (see, for instance, the Audi A6 across the 7 platoons, that exhibits high calibration errors for all models in P4 and P6, which are mainly driven by the spacing peaks shown in Fig. 3(k) and (m)). In general, the ability to reproduce observed AV longitudinal dynamics vary more with the trajectory data (vehicle and platoon) than with the model (see Section 5.3 for more detailed explanation).

Results have revealed that, given a vehicle, no base model is able to perform consistently among all trajectories i.e., all models fail to capture the full spectrum of the observed AV behaviours.

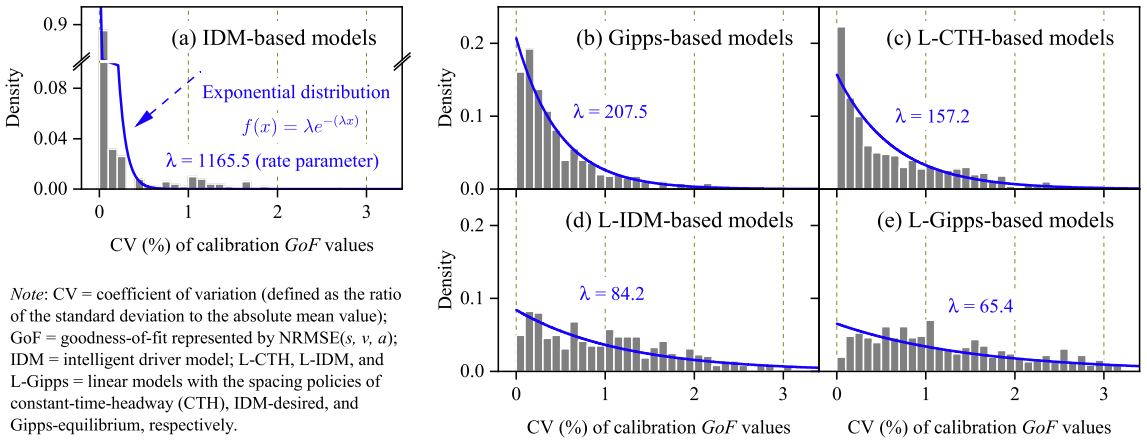
5.3. Comparison of model variants after calibration

Fig. 6a shows a box-whisker plot of calibration errors of the ninety model variants (i.e., $GoF_{i,ID}$).

To compute the boxplot of each model variant, 28 calibration experiments have been run (i.e., one for each of the 28 trajectories). The objective function value of a calibrated model in the i-th calibration experiment – i.e., the calibration error of model *ID* for trajectory i – has been indicated with $GoF_{i,ID}$. In Fig. 6a, the horizontal segment and the square within each box represent the median and the mean value, respectively, of the 28 calibration errors (i.e., of $GoF_{i,ID}$, $i \in [1, \cdots, 28]$). The bottom and top edges of the box represent the 1st (Q1) and 3rd (Q3) quartile, respectively, of the 28 errors. The whisker is drawn up to the largest error that falls within a distance of 1.5 × IQR from Q3, and it is drawn down to the lowest error that falls within the same distance from Q1. Outliers are all the errors outside the whiskers and are marked with black diamonds.

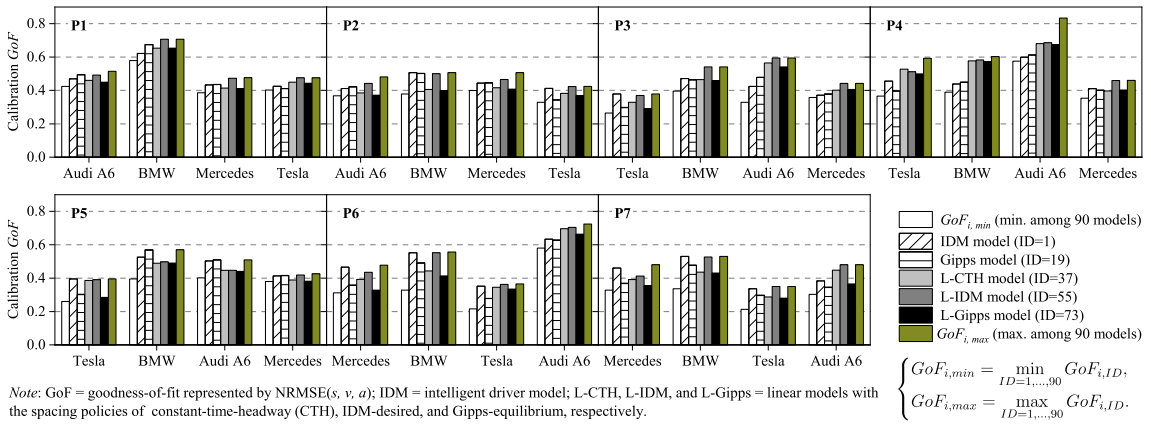
Fig. 6b is based on the same calibration errors of Fig. 6a but normalized over the minimum and maximum calibration error of each trajectory i among all 90 models, $\widetilde{GoF}_{i,ID}$:

$$\widetilde{GoF}_{i,ID} = \frac{GoF_{i,ID} - GoF_{i,min}}{GoF_{i,max} - GoF_{i,min}} \tag{19a}$$



Note: CV = coefficient of variation (defined as the ratio of the standard deviation to the absolute mean value); GoF = goodness-of-fit represented by NRMSE(s, v, a); IDM = intelligent driver model; L-CTH, L-IDM, and L-Gipps = linear models with the spacing policies of constant-time-headway (CTH), IDM-desired, and Gipps-equilibrium, respectively.

Fig. 4. Distribution of the coefficient of variation (CV) of *NRMSE*(s,v,a): (a) IDM-based models, (b) Gipps-based models, (c) L-CTH-based models, (d) L-IDM-based models, and (e) L-Gipps-based models.



Note: GoF = goodness-of-fit represented by NRMSE(s, v, a); IDM = intelligent driver model; L-CTH, L-IDM, and L-Gipps = linear models with the spacing policies of constant-time-headway (CTH), IDM-desired, and Gipps-equilibrium, respectively.

Fig. 5. Variability of calibration errors of the five base models, across the 28 vehicle trajectories. For the i-th trajectory, the calibration *GoF* values of the five base models ($ID = 1$, 19, 37, 55, and 73 in Table 2), as well as the minimum and maximum *GoF* values among all 90 models ($GoF_{i,\min}$ and $GoF_{i,\max}$), are shown.

$$\text{with}: \begin{cases} i \in [1, \cdots, 28], \\ ID \in [1, \cdots, 90], \\ GoF_{i,min} = \min\limits_{ID=1,\cdots,90} GoF_{i,ID}, \\ GoF_{i,max} = \max\limits_{ID=1,\cdots,90} GoF_{i,ID} \end{cases} \tag{19b}$$

The normalization in Fig. 6b amplifies differences among models by normalizing the calibration errors over the minimum and maximum calibration performances of the 28 experiments. A 0 value of a $\widetilde{GoF}_{i,ID}$, means that the model *ID* attained the minimum calibration error among all 90 variants for the experiment i (conversely, a value of 1 indicates the maximum error).

Boxplots are grouped into the five base model classes, and are coloured accordingly. Data, i.e., calibration errors used to draw Fig. 6 are the same of Fig. 5, but ordered by model and not by vehicle/platoon, in order to highlight model performances and relative differences.

Fig. 6a allows comparing the absolute calibration performances of the 90 model variants over the 28 experiments. Overall, *the addition of model extensions does not sensibly improve calibration performances of most of the base models*. The variability of performances among the 28 calibration experiments – as measured by box-whiskers plot amplitudes – is clearly higher than that of models. This result can be quantified by calculating the calibration error variance explained by the variability of model formulations vs. that explained by the variability of trajectories. To this end we have computed the first-order sensitivity indices of model formulations and trajectories to model errors (Saltelli et al., 2010, Punzo et al., 2015). A first order sensitivity index provides the first order effect of an input factor X_i on the unconditional variance of an output Y, and it is given by:

$$S_{X_i} = \frac{VAR_{X_i}\left(E_{X_{\sim i}}(Y|X_i)\right)}{VAR(Y)} \tag{20}$$

In our case, the two analysis factors, X_i, are the model formulation and the vehicle trajectory, while the output Y is the model calibration error. Resulting indices are: $S_{model} = 6.2\%$ and $S_{trajectory} = 81.9\%$ (the rest of the error variance being explained by higher order interaction effects). These results clearly show that, when AVs are investigated, the observed experimental trajectories (i.e., immediate leader and modelled vehicle) affect model accuracy more than the model formulations themselves. This result is analogous to that observed for car-following models of HVs (Punzo and Simonelli, 2005), and shows that *car-following models and ACC algorithms are not able to capture the full spectrum of AV behaviours* too. A further confirmation comes from Fig. 7, which enables a more detailed analysis.

The figure shows the calibration errors of each model for the 28 experiments – the same in Fig. 6a – but keeping track of the experiment, and of the car brand, across the models (see the series). Results sensibly change from one experiment to another; see the vertical span of series. Moreover, the trajectories of one car brand (i.e., Tesla) are reproduced better than those of other car brands, by all models. Conversely, some trajectories of Audi and BMW, which exhibit unstable behaviours – see platoons P1, P4 and P6 in Fig. 3 – are poorly fitted by all models. However, the ranking of car brands in terms of calibration errors does not sensibly vary from one model class to another (see the vertical order of colours which is almost invariant from one model class to another). Thus, the model class which is the best (worst) at simulating a car brand is also the best (worst) at simulating all other car brands. Therefore, *the choice of the "best" model is independent of the car brand to simulate*.

Concerning the impact of physics-based extensions on model accuracy in calibration, a comparison can be made e.g., through the *Q3 values* in Fig. 6 (i.e., the top edges of boxes). The figure reveals that linear vehicle dynamics improve model performances of all base models (Fig. 6b). To this aim, in each model class the reader can compare the fourth model to the first one, that is the base model (for

instance, in the IDM case, model ID 4 vs. model ID 1). With regards to nonlinear vehicle dynamics, they are not as beneficial as linear ones (e.g., compare ID 25 with ID 22), except for the IDM-based models (i.e., ID 7 vs. ID 4). As to the L-CTH, performances of models including nonlinear dynamics deteriorate in comparison to those of the base models (see ID 43 vs. ID 37).

Perception delay also improves modelling accuracy of all base models. Such increase has the same order of magnitude of that produced by linear dynamics (see e.g., ID 46 vs. IDs 40). A possible explanation is that the actuation lag of linear dynamics and the perception delay produce similar delayed vehicle dynamics in simulation. However, the combined effect of linear dynamics and perception delay is less than the sum of the single contributions (see e.g., ID 13 vs. IDs 4 and 10). For all model classes the impact of acceleration constraints on model accuracy is negligible.

The left half of Table 5 reports the same values of the *median calibration errors* shown in Fig. 6a (GoF_{cal}), their percentage variation from the base models (ΔGoF_{cal}), and the calibration errors (*RMSEs*; see Eq. 18) on spacing, speed and acceleration. These data allow us to investigate both the individual and the combined effect of physics-based extensions on *absolute* model accuracy. In the table, given a base model, variants are grouped in triplets. The first triplet is made by the base model and by two acceleration-constrained variants. Each other triplet refers to a physics-based extension or to a combination of physics-based extensions (models within each triplet differ only in the presence of acceleration constraints i.e., the order is the same as in Table 2).

Among the extensions, the impact of acceleration constraints on model accuracy in calibration is negligible, as anticipated. This is the effect of the least square calibration which minimizes higher deviations the most (e.g., unphysical accelerations), independently of the presence of the (constraints) extension. After all, the accelerations in the experimental data never exceed 2 m/s^2 or -4 m/s^2 so that the calibration alone prevents the achievement of unphysical accelerations. On the contrary, in cross-validation, when parameters calibrated on a trajectory are applied to other trajectories, a model response that activates constraints is more likely.

In general, physics-based extensions do not sensibly increase the accuracy of L-CTH (-1.9% of error reduction). The highest impact of the extensions occurs for Gipps' model with the addition of one parameter only, that is the perception delay, which gives an error reduction of -10.6%. As to the IDM, coherent reductions are observed among all extensions in a range from -3.5 to -6.4%. On the contrary of L-CTH, some error reductions are observed for the L-IDM and L-Gipps base models when adding both perception delay and linear vehicle dynamics (-5.5% and -4.6%, respectively).

Among the models, Gipps-based and L-Gipps-based models show the best median calibration performances ($GoF_{cal} = 0.39$ for both model classes)[5]. We recall here that both the models consider Gipps' safety assumption (i.e., the following vehicle maintains a speed which ensures a safe stop even in case of a sudden and severe deceleration of the preceding vehicle). While the assumption is inherent in Gipps-based models, in L-Gipps-based models it is implicitly considered by adopting Gipps' equilibrium speed-distance function as spacing policy.

If we focus on linear controllers, L-Gipps-based models ($GoF_{cal} = 0.39$) perform better than a linear controller with a CTH spacing policy ($GoF_{cal} = 0.42$) – which is the algorithm customarily applied in the field literature to study ACCs – and than a linear controller with an IDM desired-distance spacing policy, which is the worst performing model class ($GoF_{cal} = 0.45$).

6. Validation results

6.1. Model robustness to collisions

Model validation experiments consist in simulating a trajectory with a model calibrated on a different trajectory. When models calibrated on a trajectory are used to simulate a different trajectory, they can yield a collision, i.e., a negative inter-vehicle spacing.[6] Since observed vehicle dynamics in all platoons are collision-free, the emergence of a collision in simulation is evidence that calibrated parameters are string unstable and that, though optimal for the specific trajectories used in calibration, they cannot be transferred to simulate different trajectories.

Therefore, the frequency of collisions measured in validation experiments is an index of model robustness and parameter transferability (Zhu et al., 2018). Fig. 8 shows the collision frequency of each model over 168 validation experiments (4 vehicles × 7 platoons × 6 simulations/trajectories; see Section 4.2). In order to evaluate also the impact of different calibration settings on parameter robustness and transferability, we have plotted six curves, each one corresponding to a calibration performed with a different *GoF*(*MoP*).

Fig. 8 allows us to compare model variants in terms of robustness. *IDM-based models are the most robust*, outperforming all other models *regarding collision frequency* (which is almost zero). Concerning all other models, the higher the model complexity, the higher the collision frequency. Gipps-based models have been the second best. As to the linear controllers, the Gipps-equilibrium spacing policy has yielded more robust results than other spacing policies.

Concerning the *GoF* adopted, when *NRMSE*(*s*,*v*,*a*) and Theil's *U*(*s*,*v*,*a*) have been applied in calibration, the least number of

[5] Please, note that the model variant with the least GoF_{cal} in a class, does not necessarily coincide with the model variant which returns the least ΔGoF_{cal} in the same class. In fact, the median is not a linear operator.

[6] In model calibration, if a vehicle collision occurs, the *GoF* value is typically set to a very large number, in order to exclude the corresponding set of parameter values from the space of feasible solutions (this approach is customary in the literature and it is equivalent to adding a nonlinear constraint in the optimization problem; moreover, it avoids performing an extra model simulation, which is necessary to evaluate the nonlinear constraint). Consequently, the solution returned by the optimization algorithm does not produce vehicle collisions when the model is simulated receiving as input the same immediate leader's trajectory adopted for model calibration.

collisions have been observed in validation, for most of the ninety models.

This result extends the validity of the recommendations given in Punzo et al. (2021). In that paper, it has been shown that these two *GoFs* are the most preferable for CF model calibration in terms of accuracy, among the Pareto efficient ones (they also outperform objective functions with a single *MoP*). Fig. 8 shows that the two mentioned *GoFs* also guarantee the highest robustness and transferability of calibrated parameters.

Furthermore, when a single *MoP* is applied, plotted results show that, also in terms of robustness and transferability of calibrated parameters, spacing is preferable to speed (concerning the dualism between spacing and speed in calibration of car-following models see Punzo and Montanino, 2016). In fact, the *RMSE*(v) displays the worst performances, since it sensibly increases the collision frequency in L-CTH-based and L-IDM-based models (and in most of the Gipps-based models).

6.2. Comparison of model variants after validation

Table 5 reports the calibration (left half) and validation (right half) errors of the 90 models computed. Errors in both the halves are expressed in terms of *NRMSE*(s,v,a) (1st column), of percentage variation of the *NRMSE*(s,v,a) relative to the base model (2nd column), and of *RMSE* on spacing, speed, and acceleration (4th, 5th and 6th column). While the *RMSEs* provide the residuals from observed dynamics in their units of measurement, the *NRMSE*(s,v,a) represents a normalized score summing up errors on spacing, speed and acceleration.

Table rows are grouped by base model class. All entries in the table represent *median values* of the corresponding distributions (e.g., *NRMSE*(s,v,a) values in the first column are consistent with the median values reported in Fig. 6a). Concerning validation results, it is worth noting that only validation experiments which were collision-free were used to compute the statistics in the table right half. In the rightmost column the number of collisions that occurred in model validation are reported for each model. These numbers are consistent with the collision frequencies depicted in Fig. 8 (as given by the *NRMSE*(s,v,a) curve).

As expected, validation errors are in general higher than calibration ones, since, in cross-validation parameters calibrated in an experiment are used to reproduce trajectories of other experiments. However, a particularly high decrease of performances relative to those obtained in calibration is a clear symptom of model overfitting i.e., unrobust model. The comparison of errors between calibration and validation (see $\Delta GoF_{cal/val}$) shows that *the IDM is the most robust model*. The IDM's error increase ranges from 11% to 16% across all model variants (21–33% for Gipps and L-CTH, 24–34% for L-IDM, 15–26% for L-Gipps). This result is consistent with the one obtained in terms of collisions (see the previous section).

Concerning the extensions, the *comparative analysis of calibration and validation* results in Table 5, clearly suggests that *the choice of the "best" physics-based extension is dependent on the model class*.

The comparison of calibration and validation results relative to the IDM base model performances (see columns ΔGoF_{cal} and ΔGoF_{val}), shows that linear vehicle dynamics, which provide the best results in calibration, exhibit poor performances in validation, both alone and in combination with the perception delay. This suggests that such extension is prone to overfitting with the IDM. On the contrary, *nonlinear dynamics* (which have similar performances to linear dynamics in calibration) *are the most robust extension in validation, for the IDM*.

Concerning Gipps' model, the perception delay is by far the most robust extension, either alone or in combination with linear dynamics.

Among linear controllers the L-Gipps controller is by far the best performing model in validation. Concerning the median absolute validation error (i.e., column GoF_{val}), the *L-Gipps controller achieve the least median validation error value among all 90 models*. The most robust extension for such model is the one combining *linear dynamics, perception delay and MFC*.[7]

7. Conclusion

In this paper, a framework has been presented to compare behavioural CF models and ACC algorithms, when they are augmented with detailed lower-level dynamics such as perception delay, vehicle dynamics, and vehicle/motor constraints. The framework combines calibration and cross-validation using data from multiple trajectory datasets. All possible combinations of five base models (CF and ACC) and four physics-based extensions, i.e., 90 different model variants, have been therefore calibrated and validated against ACC trajectory data from 7 platooning experiments of four high-end vehicles controlled by ACC systems.

The five base models applied, have been the IDM and Gipps' behavioural CF models, a linear controller with a CTH spacing policy and two new ACC algorithms. These last two have been derived here by coupling a linear controller with two spacing policies from the IDM and Gipps' model. To the best of our knowledge, the approach of enhancing controllers by means of car-following theory is novel in the field literature.

The improvement in model accuracy obtained by coupling the base models with the physics-based extension has been quantified through model calibrations and cross-validations. This approach has enabled two analyses. The first one is an investigation of the impact of physics-based extensions on model accuracy. The second analysis is a quantitative comparison of the five model classes.

Concerning the impact of the *physics-based extensions*, it has been shown that *their addition does not sensibly improve calibration performances of most of the base models*. The exception is Gipps' model for which the addition of the only perception delay gives an error reduction around 10%.

[7] Please, note that the model variant with the least GoF_{val} in a class, does not necessarily coincide with the model variant which returns the least ΔGoF_{val} in the same class. In fact, the median is not a linear operator.

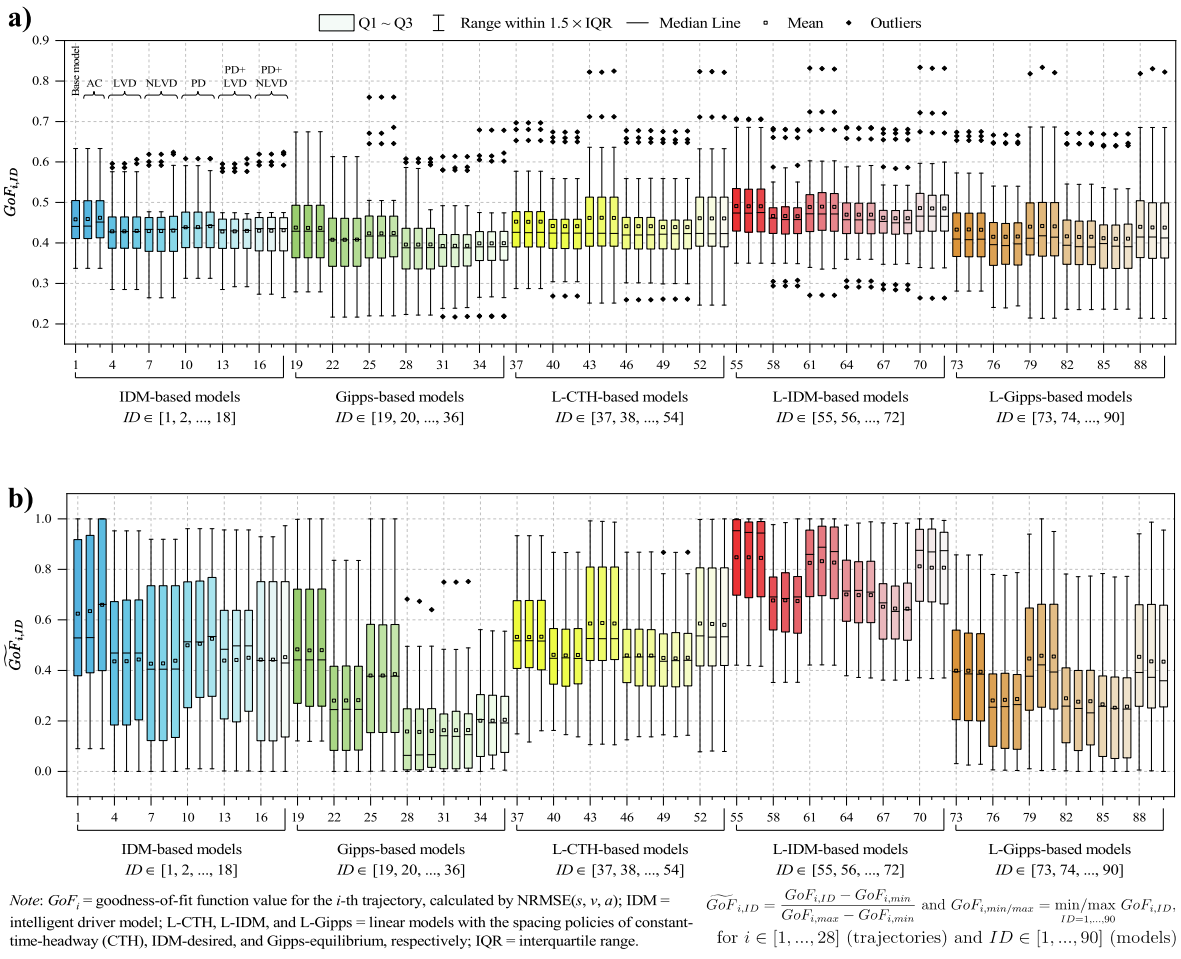


Fig. 6. Variability of (a) absolute calibration errors and (b) min–max normalized calibration errors among models. Each box plot represents the distribution of the 28 absolute / min–max normalized calibration errors of a specific model variant.

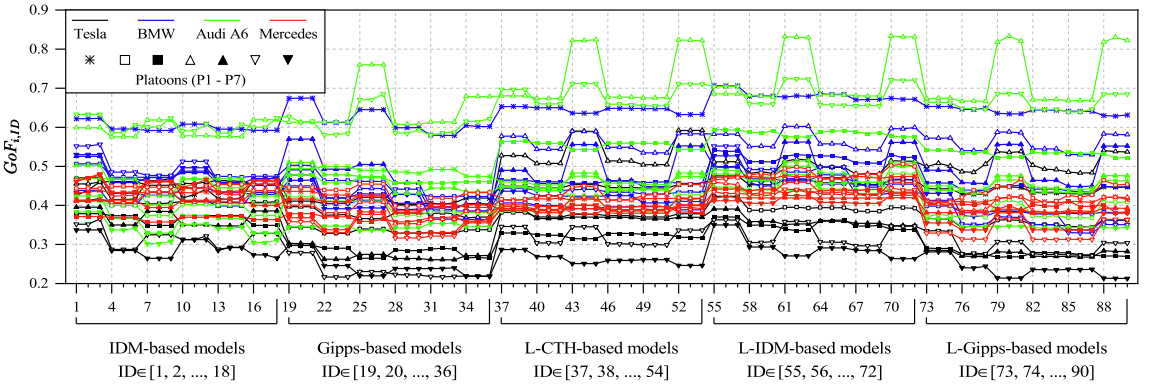


Fig. 7. Absolute calibration errors per car brand and platoon. Each one of the 28 series provides calibration errors of the 90 models for an experiment (i.e., for a specific vehicle and platoon). Colours identify car brands, markers identify platoons.

Table 5
Median calibration and validation errors of model variants, relative to the base models.

	Main extensions	ID		GoF_{cal} [1]	Δ GoF_{cal} [%] [2]	RMSE (s) [m]	RMSE (v) [m/s]	RMSE (a) [m/s^2]		GoF_{val} [1]	Δ $GoF_{cal/val}$ [%] [3]	Δ GoF_{val} [%] [2]	RMSE (s) [m]	RMSE (v) [m/s]	RMSE (a) [m/s^2]	# of collisions in validation
IDM-based models	Base	1	28 calibration experiments	0.44	–	2.97	0.51	0.16	168 collision-free validation experiments [4]	0.51	15.9	–	3.36	0.56	0.17	0/168
		2		0.44	0.0	2.97	0.51	0.16		0.51	15.9	0.0	3.38	0.56	0.17	0/168
		3		0.45	0.0	2.97	0.51	0.16		0.50	11.1	0.0	3.35	0.56	0.17	0/168
	LVD	4		0.43	−6.3	2.88	0.48	0.15		0.49	14.0	−0.9	3.43	0.56	0.17	0/168
		5		0.43	**−6.4**	2.88	0.48	0.15		0.49	14.0	−0.9	3.43	0.56	0.17	0/168
		6		0.43	−4.9	2.88	0.48	0.15		0.50	16.3	−0.7	3.43	0.56	0.17	0/168
	NVD	7		0.43	−5.0	2.63	0.46	0.15		0.48	11.6	−4.0	3.21	0.53	0.17	0/168
		8		0.43	−5.0	2.63	0.46	0.15		0.48	11.6	−4.5	3.21	0.53	0.17	0/168
		9		0.43	−4.7	2.63	0.46	0.15		0.48	11.6	**−4.6**	3.25	0.53	0.17	0/168
	PD	10		0.44	−3.6	2.93	0.49	0.15		0.51	15.9	−0.3	3.46	0.57	0.17	0/168
		11		0.44	−3.6	2.93	0.49	0.15		0.51	15.9	−0.3	3.46	0.56	0.17	0/168
		12		0.44	−3.5	2.92	0.49	0.15		0.51	15.9	−0.3	3.51	0.57	0.17	0/168
	PD + LVD	13		0.43	−5.9	2.87	0.48	0.15		0.50	16.3	−0.4	3.50	0.56	0.17	0/168
		14		0.43	−6.0	2.87	0.48	0.15		0.50	16.3	−0.2	3.51	0.56	0.17	0/168
		15		0.43	−4.6	2.87	0.48	0.15		0.50	16.3	−0.3	3.49	0.56	0.17	0/168
	PD + NVD	16		0.44	−5.0	2.62	0.46	0.15		0.49	11.4	−3.3	3.35	0.53	0.17	0/168
		17		0.44	−5.0	2.63	0.46	0.15		0.49	11.4	−3.2	3.34	0.53	0.17	0/168
		18		0.44	−4.1	2.63	0.46	0.15		0.49	11.4	−3.7	3.44	0.53	0.17	0/168
Gipps-based models	Base	19	28 calibration experiments	0.43	–	2.56	0.49	0.16	122 collision-free validation experiments [4]	0.54	25.6	–	3.59	0.56	0.17	10/168
		20		0.43	0.0	2.52	0.49	0.16		0.54	25.6	0.0	3.61	0.56	0.17	14/168
		21		0.43	0.0	2.54	0.49	0.16		0.54	25.6	0.0	3.60	0.56	0.17	15/168
	LVD	22		0.41	−5.8	2.74	0.43	0.15		0.51	24.4	−3.8	3.52	0.54	0.17	14/168
		23		0.41	−5.7	2.68	0.43	0.15		0.52	26.8	−4.1	3.53	0.54	0.17	14/168
		24		0.41	−5.6	2.72	0.43	0.15		0.52	26.8	−3.9	3.53	0.54	0.17	15/168
	NVD	25		0.42	−3.0	2.61	0.46	0.15		0.51	21.4	−1.2	3.51	0.54	0.17	13/168
		26		0.42	−2.9	2.59	0.46	0.15		0.51	21.4	−1.2	3.50	0.54	0.17	13/168
		27		0.42	−2.9	2.59	0.46	0.15		0.51	21.4	−1.2	3.46	0.54	0.17	13/168
	PD	28		0.39	**−10.6**	2.20	0.42	0.15		0.50	28.2	−6.6	3.16	0.51	0.17	15/168
		29		0.39	−10.6	2.17	0.42	0.15		0.50	28.2	**−7.0**	3.16	0.51	0.16	13/168
		30		0.39	−10.6	2.18	0.42	0.15		0.50	28.2	−7.0	3.16	0.51	0.16	16/168
	PD + LVD	31		0.39	−9.0	2.37	0.40	0.15		0.52	33.3	−6.3	3.39	0.52	0.16	29/168
		32		0.39	−9.0	2.31	0.39	0.15		0.51	30.8	−6.9	3.37	0.52	0.16	26/168
		33		0.39	−9.0	2.24	0.40	0.15		0.51	30.8	−6.9	3.36	0.52	0.17	27/168
	PD + NVD	34		0.39	−8.2	2.39	0.41	0.15		0.50	28.2	−4.7	3.41	0.52	0.17	30/168
		35		0.39	−8.1	2.35	0.41	0.15		0.50	28.2	−4.5	3.40	0.52	0.17	33/168
		36		0.39	−8.1	2.30	0.41	0.15		0.50	28.2	−4.8	3.41	0.53	0.17	34/168
L-CTH-based models	Base	37	28 calibration experiments	0.43	–	2.53	0.47	0.16	97 collision-free validation experiments [4]	0.52	20.9	–	3.11	0.58	0.19	34/168
		38		0.43	0.0	2.57	0.47	0.16		0.52	20.9	−0.1	3.11	0.58	0.19	33/168
		39		0.43	0.0	2.54	0.47	0.16		0.52	20.9	0.1	3.11	0.58	0.19	33/168
	LVD	40		0.42	−1.4	2.55	0.46	0.15		0.55	31.0	−1.4	3.26	0.57	0.20	45/168
		41		0.42	−1.6	2.56	0.46	0.15		0.53	26.2	−1.4	3.24	0.57	0.20	45/168
		42		0.42	−1.6	2.54	0.46	0.15		0.53	26.2	−1.6	3.28	0.57	0.20	45/168
	NVD	43		0.42	0.4	2.63	0.48	0.16		0.56	33.3	0.9	3.36	0.62	0.19	55/168
		44		0.42	0.8	2.69	0.48	0.16		0.56	33.3	0.8	3.32	0.62	0.19	55/168
		45		0.42	0.6	2.64	0.48	0.16		0.56	33.3	0.8	3.34	0.61	0.20	55/168
	PD	46		0.42	−1.6	2.52	0.46	0.15		0.54	28.6	−1.3	3.29	0.57	0.20	49/168

(*continued on next page*)

Table 5 (*continued*)

	Main extensions	ID		GoF_{cal} [1]	Δ GoF_{cal} [%] [2]	RMSE (s) [m]	RMSE (v) [m/s]	RMSE (a) [m/s^2]		GoF_{val} [1]	Δ $GoF_{cal/val}$ [%] [3]	Δ GoF_{val} [%] [2]	RMSE (s) [m]	RMSE (v) [m/s]	RMSE (a) [m/s^2]	# of collisions in validation
		47		0.42	**−1.9**	2.54	0.46	0.15		0.54	28.6	−1.7	3.26	0.57	0.19	50/168
		48		0.42	−1.9	2.54	0.46	0.15		0.54	28.6	−1.3	3.30	0.57	0.20	49/168
	PD + LVD	49		0.42	−1.6	2.55	0.46	0.15		0.52	23.8	−1.4	3.22	0.56	0.19	49/168
		50		0.42	−1.7	2.57	0.46	0.15		0.52	23.8	−1.8	3.22	0.56	0.19	51/168
		51		0.42	−1.8	2.57	0.46	0.15		0.52	23.8	**−1.8**	3.23	0.57	0.19	52/168
	PD + NVD	52		0.42	1.0	2.70	0.49	0.17		0.56	33.3	1.9	3.32	0.61	0.19	63/168
		53		0.42	0.7	2.68	0.49	0.17		0.56	33.3	1.5	3.41	0.61	0.19	60/168
		54		0.42	0.9	2.68	0.48	0.16		0.56	33.3	1.3	3.33	0.61	0.19	62/168
L-IDM-based models	Base	55	28 calibration experiments	0.47	–	2.79	0.54	0.18	76 collision-free validation experiments [4]	0.59	25.5	–	3.28	0.62	0.22	57/168
		56		0.47	0.0	2.81	0.54	0.18		0.59	25.5	−0.2	3.23	0.62	0.21	59/168
		57		0.47	0.0	2.79	0.54	0.18		0.59	25.5	−0.1	3.30	0.63	0.21	56/168
	LVD	58		0.46	−5.1	2.78	0.50	0.17		0.57	23.9	−2.5	3.21	0.62	0.20	65/168
		59		0.46	−4.8	2.80	0.50	0.17		0.57	23.9	**−3.6**	3.22	0.61	0.20	67/168
		60		0.46	−5.1	2.79	0.50	0.17		0.57	23.9	−2.8	3.24	0.62	0.20	68/168
	NVD	61		0.47	−0.8	2.91	0.52	0.18		0.62	31.9	2.7	3.65	0.68	0.22	79/168
		62		0.47	−0.3	3.00	0.53	0.18		0.62	31.9	2.5	3.71	0.69	0.21	79/168
		63		0.47	−0.6	2.93	0.53	0.18		0.62	31.9	2.7	3.56	0.68	0.21	80/168
	PD	64		0.46	−2.8	2.77	0.50	0.17		0.58	26.1	−2.2	3.28	0.62	0.20	71/168
		65		0.46	−3.0	2.77	0.50	0.17		0.58	26.1	−2.9	3.24	0.62	0.20	70/168
		66		0.46	−3.4	2.77	0.50	0.17		0.59	28.3	−2.4	3.32	0.63	0.20	70/168
	PD + LVD	67		0.45	−5.5	2.79	0.50	0.17		0.57	26.7	−3.3	3.30	0.62	0.20	74/168
		68		0.45	**−5.5**	2.78	0.49	0.17		0.57	26.7	−2.0	3.32	0.63	0.20	74/168
		69		0.45	−5.3	2.79	0.49	0.17		0.58	28.9	−2.7	3.32	0.64	0.20	74/168
	PD + NVD	70		0.47	−1.0	2.86	0.52	0.17		0.63	34.0	3.3	3.57	0.69	0.21	81/168
		71		0.47	−1.2	2.95	0.52	0.17		0.62	31.9	2.7	3.57	0.68	0.22	81/168
		72		0.47	−1.1	2.88	0.52	0.17		0.62	31.9	3.0	3.57	0.69	0.22	84/168
L-Gipps-based models	Base	73	28 calibration experiments	0.41	–	2.40	0.45	0.16	83 collision-free validation experiments [4]	0.47	14.6	–	2.98	0.55	0.17	35/168
		74		0.41	0.0	2.39	0.44	0.16		0.48	17.1	0.1	3.02	0.55	0.18	32/168
		75		0.41	−0.1	2.36	0.44	0.16		0.47	14.6	−0.1	2.94	0.55	0.17	38/168
	LVD	76		0.40	−3.7	2.45	0.43	0.15		0.47	17.5	−1.5	3.08	0.54	0.18	56/168
		77		0.39	−3.7	2.50	0.44	0.15		0.48	23.1	−2.7	3.01	0.53	0.18	54/168
		78		0.40	−3.5	2.52	0.43	0.15		0.48	20.0	−1.8	3.08	0.55	0.18	56/168
	NVD	79		0.41	1.1	2.62	0.46	0.15		0.51	24.4	1.0	3.23	0.57	0.18	49/168
		80		0.42	1.6	2.60	0.46	0.16		0.50	19.0	1.4	3.23	0.57	0.18	51/168
		81		0.41	1.5	2.67	0.46	0.15		0.51	24.4	1.7	3.28	0.56	0.18	57/168
	PD	82		0.39	−3.7	2.54	0.44	0.15		0.49	25.6	−1.7	3.12	0.55	0.17	56/168
		83		0.39	−3.6	2.37	0.44	0.15		0.47	20.5	−2.1	3.03	0.55	0.17	55/168
		84		0.39	−3.7	2.47	0.43	0.15		0.49	25.6	−2.2	3.09	0.55	0.17	60/168
	PD + LVD	85		0.40	−4.0	2.44	0.43	0.15		0.50	25.0	−2.2	3.21	0.54	0.17	62/168
		86		0.39	−4.4	2.46	0.43	0.15		0.48	23.1	−3.0	3.09	0.54	0.17	61/168
		87		0.39	**−4.6**	2.43	0.43	0.14		0.48	23.1	**−4.0**	3.10	0.54	0.17	57/168
	PD + NVD	88		0.41	1.5	2.63	0.46	0.16		0.50	22.0	1.7	3.21	0.57	0.18	53/168
		89		0.41	1.3	2.66	0.46	0.16		0.50	22.0	2.3	3.24	0.57	0.18	57/168
		90		0.41	1.0	2.69	0.46	0.15		0.50	22.0	1.7	3.25	0.57	0.18	57/168

[1] GoF = median(NRMSE(s, v,a)).

[2] ΔGoF = median[(GoF_{ID} -GoF_{Base})/GoF_{Base}].

[3] $\Delta GoF_{cal/val}$ = (GoF_{val} -GoF_{cal})/GoF_{cal}.

[4] # of experiments which are collision-free for all 18 models in the class.

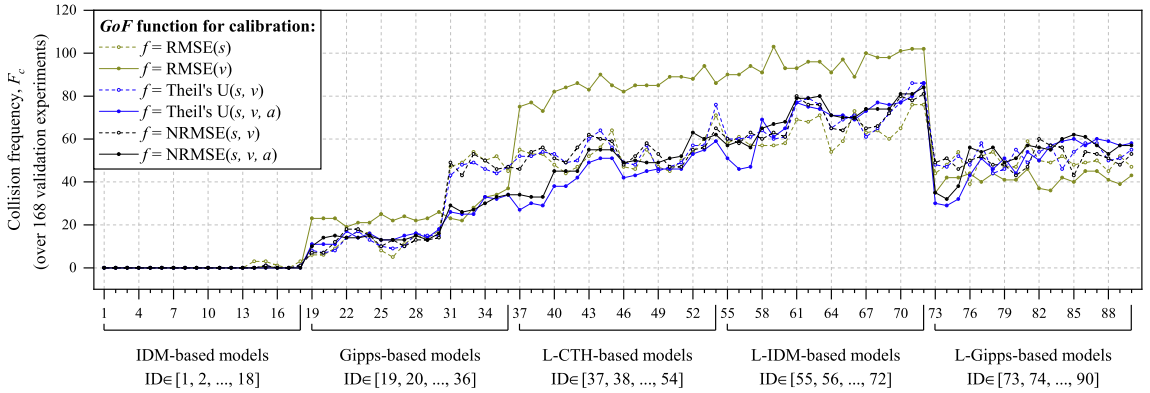


Fig. 8. Collision frequency, for each model variant, over 168 validation experiments.

Overall, in our experimentation we have shown that the variance of calibration errors among model variants and experiments is mostly explained by the experimental trajectories (82%) (the variance explained by the model formulation is around 6%). In fact, the ability to reproduce an observed follower's trajectory depends on the experiment – i.e., on the combination of the trajectory of the vehicle in front and the brand of the following car – much more than on the model variant. Thus, depending on the experimental trajectories, either all models perform (more or less) satisfactorily, or all perform (more or less) poorly. This result suggests that all models fail to capture the behaviour of some car brands, just as models fail with some human drivers. As a conclusion, *the choice of the "best" model is independent of the car brand to simulate*.

Calibration results show that IDM and Gipps' CF models – which are usually applied to simulate human-driven vehicles – reproduce ACC vehicle dynamics more accurately than a linear controller with a CTH spacing policy – which is customarily applied to simulate ACCs. Across models, *Gipps-based models* and the *linear controllers with Gipps' equilibrium speed-spacing function* were the best performing models in calibration. In particular, we have shown that Gipps' distance policy is better than the customarily applied CTH policy at reproducing AVs trajectories (when plugged into a linear controller). It is no coincidence that Gipps' model was originally conceived to interpret safe human drivers' car-following behaviours.

With the aim of model and extension comparison, validation results are clearly more significant than calibration ones, as they tell the predictive capability of a model. To this regard, *the linear controller with the Gipps' spacing policy shows the least median validation errors among all 90 models*. Concerning model robustness, as measured by the spread between calibration and validation results, the *IDM is the most robust base model*, indeed. Its robustness to input trajectories variability is demonstrated also by the remarkable capability of avoiding collisions in validation experiments. This result, and the consideration that the IDM spacing policy does not perform satisfactorily when coupled with a linear controller, suggests that the robustness and prediction capability of the IDM are mainly due to the controller formulation (rather than to the model spacing policy).

Concerning the impact of physics-based extensions, the *comparative analysis of calibration and validation* results suggests that *the choice of the "best" extension is dependent on the model class.* For instance, the extension of the L-Gipps controller (i.e., the model class achieving the least median validation error) that provides the most robust results, is the one combining linear vehicle dynamics, perception delay, and the MFC. Differently, the extension combining nonlinear dynamics and the MFC is the most robust one for the IDM.

Results have also shown that using the *NRMSE*(*s*,*v*,*a*) as objective function in CF/ACC model calibration sensibly improves model robustness and transferability of calibrated parameters. This result is a confirmation of the recommendation provided in Punzo et al. (2021) about calibration settings and extends its validity to a considerable larger number of models (90 in this study) and to a larger dataset of ACC trajectories. In addition, it proves that calibrating with the *NRMSE*(*s*,*v*,*a*) enhances also model parameter robustness and transferability.

In conclusion, this study shows that physics-based extensions provide limited improvements to the accuracy of existing models, and that it is not possible to predict what extension is the most suited for a specific model without an investigation against observed data. Since different levels of accuracy are exhibited by all models depending on the car brand to simulate, mixed traffic simulations aiming at evaluating safety and traffic efficiency of AVs, must carefully consider such limitations in order to provide meaningful results.

CRediT authorship contribution statement

Yinglong He: Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization. **Marcello Montanino:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration. **Konstantinos Mattas:** Validation, Formal analysis, Resources, Data curation, Writing – original draft. **Vincenzo Punzo:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Biagio Ciuffo:** Conceptualization, Methodology, Validation, Resources, Writing – original draft, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The views expressed here are purely those of the authors and may not, under any circumstances, be regarded as an official position of the European Commission. The authors are grateful to Giovanni Albano, Aikaterini Anesaidou and Gaetano Zaccaria for the support provided during the work, as well as to Michail Makridis for the inspiring discussions. The authors would like to thank all the staff in AstaZero proving ground in Sweden for their support during the experiments. All the data used in the paper are part of the JRC openACC database openly available online. This research was partially funded by the Italian program PON AIM - Attraction and International Mobility, Linea 1 (AIM1849341-2, CUP E61G18000540007) and by the Italian program MIUR – Departments of Excellence 2018-2022.

References

Alonso Raposo, M., Ciuffo, B., Alves Dies, P., Ardente, F., Aurambout, J.P., Baldini, G., Baranzelli, C., Blagoeva, D., Bobba, S., Braun, R., 2019. The Future of Road Transport - Implications of Automated, Connected, Low-Carbon and Shared Mobility. Publications Office of the European Union, Luxembourg.
Bando, M., Hasebe, K., Nakayama, A., Shibata, A., Sugiyama, Y., 1995. Dynamical model of traffic congestion and numerical simulation. Phys. Rev. E 51 (2), 1035–1042.
Brockfeld, E., Kühne, R.D., Wagner, P., 2004. Calibration and validation of microscopic traffic flow models. Transp. Res. Rec. 1876, 62–70.
Ciuffo, B., Punzo, V., Montanino, M., 2012. Thirty years of Gipps' car-following model: applications, developments, and new features. Transp. Res. Rec. 2315, 89–99.
Darbha, S., Hedrick, J.K., Chien, C.C., Ioannou, P., 1994. A comparison of spacing and headway control laws for automatically controlled vehicles. Veh. Syst. Dyn. 23, 597–625.
Fadhloun, K., Rakha, H., 2020. A novel vehicle dynamics and human behavior car-following model: Model development and preliminary testing. Int. J. Transp. Sci. Technol. 9, 14–28.
Fiori, C., Marzano, V., Punzo, V., Montanino, M., 2021. Energy consumption modeling in presence of uncertainty. IEEE Trans. Intell. Transp. Syst. 22 (10), 6330–6341.
Gipps, P., 1981. A behavioral car-following model for computer-simulation. Transp. Res. Part B-Methodol. 15, 105–111.
Gunter, G., Gloudemans, D., Stern, R.E., McQuade, S., Bhadani, R., Bunting, M., Monache, M.L.D., Lysecky, R., Seibold, B., Sprinkle, J., Piccoli, B., Work, D.B., 2020. Are commercially implemented adaptive cruise control systems string stable? IEEE Trans. Intell. Transp. Syst. 1–12.
Gunter, G., Stern, R., Work, D.B., 2019. Modeling adaptive cruise control vehicles from experimental data: model comparison. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). Presented at the 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pp. 3049–3054.
He, Y., Makridis, M., Mattas, K., Fontaras, G., Ciuffo, B., Xu, H., 2020a. Introducing electrified vehicle dynamics in traffic simulation. Transp. Res. Rec. 0361198120931842.
He, Y., Wang, C., Zhou, Q., Li, J., Makridis, M., Williams, H., Lu, G., Xu, H., 2020b. Multiobjective component sizing of a hybrid ethanol-electric vehicle propulsion system. Appl. Energy 266, 114843.
He, Y., Zhou, Q., Makridis, M., Mattas, K., Li, J., Williams, H., Xu, H., 2020c. Multiobjective co-optimization of cooperative adaptive cruise control and energy management strategy for PHEVs. IEEE Trans. Transp. Electrif. 6, 346–355.
Jia, D., Ngoduy, D., 2016. Platoon based cooperative driving model with consideration of realistic inter-vehicle communication. Transp. Res. Part C: Emerg. Technol. 68, 245–264.
Kesting, A., Treiber, M., Helbing, D., 2010. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. Philos. Trans. Roy. Soc. A: Math., Phys. Eng. Sci. 368, 4585–4605.
Kühlwein, J., 2016. The impact of official versus real-world road loads on CO2 emissions and fuel consumption of European passenger cars. Tech rep, The International Council on Clean Transportation, Berlin.
Küng, L., Bütler, T., Georges, G., Boulouchos, K., 2019. How much energy does a car need on the road? Appl. Energy 256, 113948.
Li, S.E., Zheng, Y., Li, K., Wu, Y., Hedrick, J.K., Gao, F., Zhang, H., 2017. Dynamical modeling and distributed control of connected and automated vehicles: challenges and opportunities. IEEE Intell. Transp. Syst. Mag. 9, 46–58.
Makridis, M., Fontaras, G., Ciuffo, B., Mattas, K., 2019. MFC free-flow model: introducing vehicle dynamics in microsimulation. Transp. Res. Rec. 2673, 762–777.
Makridis, M., Mattas, K., Anesiadou, A., Ciuffo, B., 2021. OpenACC. An open database of car-following experiments to study the properties of commercial ACC systems. Transp. Res. Part C: Emerg. Technol. 125, 103047.
Mannering, F.L., Washburn, S.S., 2020. Principles of Highway Engineering and Traffic Analysis. John Wiley & Sons.
Milanés, V., Shladover, S.E., 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. Transp. Res. Part C: Emerg. Technol. 48, 285–300.
Montanino, M., Monteil, J., Punzo, V., 2021. From homogeneous to heterogeneous traffic flows: Lp String stability under uncertain model parameters. Transp. Res. Part B: Methodol. 146, 136–154.
Montanino, M., Punzo, V., 2021. On string stability of a mixed and heterogeneous traffic flow: A unifying modelling framework. Transp. Res. Part B: Methodol. 144, 133–154.
Ni, D., 2011. Multiscale modeling of traffic flow. Mathematica Aeterna 1 (1), 27–54.
Punzo, V., Montanino, M., 2020. A two-level probabilistic approach for validation of stochastic traffic simulations: impact of drivers' heterogeneity models. Transp. Res. Part C: Emerg. Technol. 121, 102843.
Punzo, V., Montanino, M., 2016. Speed or spacing? Cumulative variables, and convolution of model errors and time in traffic flow models validation and calibration. Transp. Res. Part B: Methodol. 91, 21–33.
Punzo, V., Montanino, M., Ciuffo, B., 2015. Do we really need to calibrate all the parameters? Variance-based sensitivity analysis to simplify microscopic traffic flow models. IEEE Trans. Intell. Transp. Syst. 16 (1), 184–193.
Punzo, V., Simonelli, F., 2005. Analysis and comparison of microscopic traffic flow models with real traffic microscopic data. Transp. Res. Rec. 1934, 53–63.
Punzo, V., Tripodi, A., 2007. Steady-state solutions and multi-class calibration of Gipps' microscopic traffic flow model. Transp. Res. Rec. 1999, 104–114.
Punzo, V., Zheng, Z., Montanino, M., 2021. About calibration of car-following dynamics of automated and human-driven vehicles: Methodology, guidelines and codes. Transp. Res. Part C: Emerg. Technol. 128, 103165.
Rakha, H., 2009. Validation of Van Aerde's simplified steadystate car-following and traffic stream model. Transp. Lett. 1, 227–244.
Ranjitkar, P., Nakatsuji, T., Asano, M., 2004. Performance evaluation of microscopic traffic flow models with test track data. Transp. Res. Rec. 1876, 90–100.
Rothery, R.W., 2001. Car-following models. In: Gartner, N.H., Messer, C.J., Rathi, A.K. (Eds.), Revised Monograph on Traffic Flow Theory, 4. Transportation Research Board, USA, Chapter, pp. 1–42.

Saifuzzaman, M., Zheng, Z., 2014. Incorporating human-factors in car-following models: A review of recent developments and research needs. Transp. Res. Part C: Emerg. Technol. 48, 379–403.
Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Commun. 181 (2), 259–270.
Shladover, S.E., Su, D., Lu, X.-Y., 2012. Impacts of cooperative adaptive cruise control on freeway traffic flow. Transp. Res. Rec. 2324, 63–70.
Tian, J., Zhang, H.M., Treiber, M., Jiang, R., Gao, Z.-Y., Jia, B., 2019. On the role of speed adaptation and spacing indifference in traffic instability: Evidence from car-following experiments and its stochastic model. Transp. Res. Part B: Methodol. 129, 334–350.
Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E 62, 1805–1824.
Treiber, M., Kesting, A., 2018. The Intelligent Driver Model with stochasticity – New insights into traffic flow oscillations. Transp. Res. Part B: Methodol. 117, 613–623.
Treiber, M., Kesting, A., 2011. Evidence of convective instability in congested traffic flow: A systematic empirical and theoretical investigation. Proc.-Soc. Behav. Sci. 17, 683–701.
Treiber, M., Kanagaraj, V., 2015. Comparing numerical integration schemes for time-continuous car-following models. Physica A 419, 183–195.
Treiber, M., Kesting, A., Helbing, D., 2006. Delays, inaccuracies and anticipation in microscopic traffic models. Physica A 360, 71–88.
Tsiakmakis, S., Fontaras, G., Ciuffo, B., Samaras, Z., 2017. A simulation-based methodology for quantifying European passenger car fleet CO2 emissions. Appl. Energy 199, 447–465.
VanderWerf, J., Shladover, S., Kourjanskaia, N., Miller, M., Krishnan, H., 2001. Modeling effects of driver control assistance systems on traffic. Transp. Res. Rec. 1748, 167–174.
Ward, J.A., Wilson, R.E., 2011. Criteria for convective versus absolute string instability in car-following models. Proc. Roy. Soc. A 467 (2132), 2185–2208.
Wang, M., 2018. Infrastructure assisted adaptive driving to stabilise heterogeneous vehicle strings. Transp. Res. Part C: Emerg. Technol. 91, 276–295.
Wang, Y., Gunter, G., Nice, M., Monache, M.L.D., Work, D., 2021. Online parameter estimation methods for adaptive cruise control systems. IEEE Trans. Intell. Veh. 6 (2), 288–298.
Wiedemann, R., 1974. Simulation des strassenverkehrsflusses. PhD dissertation thesis. University of Karlsruhe.
Wu, C., Xu, Z., Liu, Y., Fu, C., Li, K., Hu, M., 2020. Spacing policies for adaptive cruise control: a survey. IEEE Access 8, 50149–50162.
Xiao, L., Gao, F., 2011. Practical string stability of platoon of adaptive cruise control vehicles. IEEE Trans. Intell. Transp. Syst. 12, 1184–1194.
Xiao, L., Wang, M., van Arem, B., 2017. Realistic car-following models for microscopic simulation of adaptive and cooperative adaptive cruise control vehicles. Transp. Res. Rec. 2623, 1–9.
Yu, H., Jiang, R., He, Z., Zheng, Z., Li, L., Liu, R., Chen, X., 2021. Automated vehicle-involved traffic flow studies: A survey of assumptions, models, speculations, and perspectives. Transp. Res. Part C: Emerg. Technol. 127, 103101.
Zhang, X., Huang, Y., Guo, K., Peng, T., Sun, S., Li, W., 2019. Integrated spacing policy considering micro- and macroscopic characteristics. Automot. Innov. 2, 102–109.
Zheng, Y., Li, S.E., Li, K., Borrelli, F., Hedrick, J.K., 2017. Distributed model predictive control for heterogeneous vehicle platoons under unidirectional topologies. IEEE Trans. Control Syst. Technol. 25, 899–910.
Zheng, Y., Li, S.E., Wang, J., Cao, D., Li, K., 2016. Stability and scalability of homogeneous vehicular platoon: study on the influence of information flow topologies. IEEE Trans. Intell. Transp. Syst. 17, 14–26.
Zhu, M., Wang, X., Tarko, A., Fang, S., 2018. Modeling car-following behavior on urban expressways in Shanghai: A naturalistic driving study. Transp. Res. Part C: Emerg. Technol. 93, 425–445.