IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 July 2024, accepted 23 August 2024, date of publication 29 August 2024, date of current version 10 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3451622

RESEARCH ARTICLE

Experimenting With an Efficient Driver Behavior Dynamical Model Applicable to Simulated Lane Changing Tasks

MIROSLAV JIRGL[1], ONDREJ MIHÁLIK[1], SABRINA BOUJENFA[2], ZDENĚK BRADÁČ[3], AND PETR FIEDLER[3]

[1]Faculty of Electrical Engineering and Communication, Brno University of Technology, 616 00 Brno, Czech Republic
[2]Thyssenkrupp Presta AG, 9492 Eschen, Liechtenstein
[3]CEITEC, Brno University of Technology, 616 00 Brno, Czech Republic

Corresponding author: Miroslav Jirgl (jirgl@vut.cz)

This work was supported in part by the Internal Science Fund of Brno University of Technology under Grant FEKT- S-23-8451, and in part by the Technology Agency of the Czech Republic Communication Systems under Grant TN02000067/005.

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Committee for Biomedical Research at the Faculty of Electrical Engineering and Communication, Brno University of Technology, under Application No. 04b/2020.

ABSTRACT We test an approach to modelling the car driver behaviour during simulated lane changing tasks, aiming to obtain a sufficiently precise model in the simplest possible form, namely, with a small number of parameters. Various applications of such models are available in the literature. Based on a recent review of the research to date, the cybernetic single-loop transfer function models employing McRuer's theory are applied. The purpose of the presented method is to evaluate the optimal structure of the transfer function via cross-validation as a technique known from machine learning. The experiments utilize a driving simulator with in-house developed software; this configuration facilitates acquiring the data at the desired sampling frequency and in a manner that ensures the repeatability of the test process scenarios. Using the cross-validation results, we evaluate the second-order model with a derivative state and a reaction delay component as an optimal structure for approximating the measured data, which originated from a set of measurements on 92 active drivers. Even though more complex driving tasks could require high-order models, driver's control action during our specific experiment is described through only four parameters. The parameters are jointly determined by the current driver's mental state and the testing conditions defined in our scenario. Since the parameters are related to his/her dynamical behaviour, they allow easier mutual comparison of the drivers than complex models with many parameters. The results are verified via establishing a relationship to the multi-loop model presented in the recent literature. The larger dataset enables evaluating the confidence intervals of the drivers' parameters which is inconvenient with 4 to 10 drivers commonly presented in the relevant sources.

INDEX TERMS Cross-validation, driver behavior, identification, model, simulator, steering control.

I. INTRODUCTION

The mental and physiological performance of pilots and drivers embodies a major transport safety-related problem and, as such, has been subjected to multiple expert analyses, simulations, and discussions. Previous research, including projects designed by the authors of this article, focused to a great extent on military pilots, exposing procedures that determine the condition and performance of a pilot by monitoring their responses to pre-defined situations (scenarios) and evaluating the obtained dynamic capabilities [1], [2]. Further investigation then showed that these capabilities are sensitive to variations in relevant mental and physiological factors, such as experience or tiredness [3], [4]. The findings

The associate editor coordinating the review of this manuscript and approving it for publication was Mouquan Shen.

were later applied to car drivers too, as is also obvious from the issues examined herein. Our approach is based on searching for and identifying an effective mathematical model to define human driver behaviour during simulated lane changing task.

The human behaviour and the modelling techniques in general are set out in, for example, sources [5], [6], and an insight into recent driver-oriented approaches is proposed in articles [7], [8], [9]. According to these studies, the most popular modelling methods can be classified into three main categories:

- artificial neural networks (ANN),
- hidden Markov models (HMM),
- cybernetic approach (based on the feedback control theory).

The current HMM and ANN models are aimed mainly at predicting the driver's future actions based on the present states of the driver and the vehicle. This concept finds use in assistance systems and driver training. Lee et al. [10] trained two ANNs to predict the steering and pedal operation in professional drivers; the predictions were then successfully employed to guide less experienced drivers via haptic assistance. Further, Cai et al. [9] reported that Bi-directional Long Short-Term Memory (Bi-LSTM) ANNs outperform the cybernetic models despite requiring measurement of the driver's future actions, a process that may be unavailable in many applications.

In addition to the prediction capabilities, the cybernetic approach enables a mathematical description and analysis of the entire control loop, which comprises the vehicle and the driver. Such a capacity then allows an offline simulation and analysis of the feedback loop under user-defined conditions [11], [12], [13], [14]. Perhaps the most obvious advantage of the cybernetic approach lies in its ability to analyse the control loop via methods well known from system and automatic control theories.

Considering the previous research, longitudinal movement control involving the speed or relative position control of a car is usually represented by car-following models. The Gipps' kinematics-based model and Wiedemann physiology-psychology model constituted perhaps the most common modelling approaches in the long-term perspective [15], [16]. Later, models based on fuzzy logic, neural networks, and machine learning were widely researched and applied [17], [18]. A combination of the traditional and machine learning-based approaches was recently studied and published in diverse papers and articles, such as [19]. The remaining set of researched and presented models (including our approach) are focused on lateral movement control.

Lateral control of the vehicle movement is, for instance, applied by Zhang et al. [20], who uses simulation to assess the stability of the loop under non-linear tyre characteristics; similarly, Nash and Cole simulations facilitated investigating the effects of a limited dynamical range of human sensory dynamics [12], or the influence of the longitudinal position of the centre of vehicle's mass on the human-vehicle dynamics [14]. Mathematical models of human dynamics are necessary for design of human-machine shared driving control, because incorrect driver model could lead to conflicts between the human and automated controller [21].

The cybernetic approach exploits a mathematical description of the human-machine interaction as a standard feedback control loop and analyzes its dynamic properties. The various cybernetic methods historically gave rise to different human behavior modeling options involving linear and non-linear streams. The non-linear techniques most often employ model predictive control (MPC), which exhibits an improved performance and accuracy [12]. The concept was applied in various projects, such as that outlined in report [22], whose authors discuss the output of a controller during an aggressive, non-linear steering manoeuvre. Another example is exposed in [23], here, an optimal linear preview control with only modest excursions from the equilibrium state is extended into the general large-lateral-motion area. All the presented sources set out the benefits of employing non-linear techniques in critical or strongly non-linear applications; however, as reported in [24], identifying such a complex non-linear model may involve cumbersome optimization approaches that lack appreciable advantages over linear transfer-function models when the vehicle operates in the linear region. Another comparison centered on linear and non-linear MPC is available in [20]. The article proposes that non-linear MPC yields a smaller path tracking error during high-speed lane change manoeuvres; the indicated improvement, however, is only 7.66%. The application of a linear quadratic regulator (LQR) and MPC methods in driver steering control modeling is compared in [25]; the report demonstrates that, for a driver model controlling a linear time-invariant vehicle without constraints, the results of LQR and MPC controllers are equivalent. In [26], by extension, the non-linear vehicle model is linearized to allow calculating an LQR for the equilibrium point. A further non-linear modeling option lies in ANNs; the authors of [27] nevertrheless conclude that a shallow ANN comprising a linear activation function suffices to facilitate lane changing and double-lane changing tasks. Such a model is then equivalent to a linear feedback controller.

Although non-linear controllers are undoubtedly more versatile than linear ones, they require relatively wide and precise knowledge of the parameters of human perception dynamics [11]; moreover, they expect the driver to know the controlled vehicle in detail, are relatively complex, and their decisions cannot be easily interpreted. Based on the above-referenced literature, we can conclude that a linear model should provide a sufficiently accurate approximation of the driver's control actions if less critical or well-linearizable situations are investigated. Due to such reasons, this study focuses on the linear models. In particular, numerous studies are accessible that employ linear state feedback models with

delay [27], [28], single-loop transfer-function models [29], [30], or diverse combinations of these approaches [7], [31], [32], [33], [34], [35].

Intending to maximize the interpretability of the model, we decided to restrict our study to McRuer's models based on transfer functions [29] and state feedback models. The approaches that employ transfer function models are built on the principle referred to as Crossover law; see, for example, articles [36], [37]. This law, formulated by D. T. McRuer, embodies one of the most famous theories in the field; its indisputable advantage lies in the simplicity and usability for a wide spectrum of human activities [29].

The article presents a method for finding a suitable model structure, namely, a transfer function, to describe the driver behavior during specific tasks. The benefits of this study, which adds to the literature, are set out as follows:

1) **Model selection:** Selecting the structure of the model should be dictated by the data rather than by the experimenter's intuition [29]. Various structures are objectively ranked utilizing the cross-validation technique [38], an approach that relies on a set of measured data. The method selects the simplest model with a minimal number of parameters, namely, an instrument that delivers a sufficiently accurate approximation of the human control actions during a particular scenario. The actual simple character of the model facilitates an easier interpretation of the results and a reduction of the errors arising from the uncertainties of the established parameters [38]; in contrast, increasing the complexity does not necessarily improve the prediction accuracy during our lane-changing task.
2) **Sample size:** The identification procedure optimizes the model's parameters from the data obtained via testing the 92 drivers on a driving simulator and evaluates the average model of a concrete driver during our lane change scenario. The dataset is significantly larger than the small datasets presented in the literature, as these usually subsume 4–10 drivers, [13], [24], [27], [34], [35], [39]. Thus, we can expect the results presented herein, the histograms and the Gaussian mixture model in particular, to better approximate the true probability density functions of the drivers.

The article is organized as follows: The second and third chapters discuss data acquisition via a car driving simulator; chapter IV presents a method for selecting a single-loop model structure to suit our use case (based on McRuer's theory), and relationship between the selected single-loop model and the multi-loop model; and the final portion, chapter V, focuses on statistical analysis.

II. CAR DRIVING SIMULATOR

The data that characterize a driver's behaviour during specific situations embody an integral part of the scientific discipline that focuses on driver modelling. Acquiring the data generated during vehicle driving can provide complex information about the driver's habits and learned routines. Such data are nevertheless markedly influenced by the environment and the current situation; all of the relevant conditions need to be monitored (measured) and considered, presenting a relatively challenging task. An alternative option rests in using a car driving simulator. Simulators in general comprise multiple scenarios that may be difficult to execute in the real world, enabling us to create a safe test environment. The importance of simulators is increasing, especially as regards car driving and testing or flight training [40], [41], [42]. In our experiments, the main advantages lie in data acquisition at a sufficient sampling rate and in the ability to simulate car driving under various pre-defined conditions or situations, with an emphasis on realistic interpretation. Naturally, the simulators are not capable of simulating an absolutely accurate real-world model devoid of side factors. This apparent drawback nevertheless facilitates excluding all undesirable effects and defining stable measuring conditions to guarantee repeatability of the testing process.

Driving simulators are often designed commercially; however, to comply with industrial requirements, customized simulators are available, and these devices then suit specific research aims and objectives. To collect relevant data from test subjects, commercial systems need to be additionally equipped with external sensor technologies, which are rarely integrated in the course of the manufacturing phase [43], [44], [45].

Due to the above reasons, we developed an in-house car driving simulator (CDS) software. The procedures were performed in Unreal Engine 4 (UE4), a game engine that is used primarily for developing computer or console games but also embodies a convenient framework (based on C++) for research applications, such as [46], [47], [48], and [49].

The features include, for instance, advanced graphics to deliver high-level graphical realism, an implemented advanced PhysX vehicle model, custom C++ functions and plugins, multiple assets and object models, and VR (virtual reality) development compatibility. The functional block diagram of the car driving simulator (CDS) is shown in Figure 1.

In addition to UE4 as the core of the simulator application, several related elements and functions enable the simulator to form an entire unit; such aspects include a PhysX-based car model; the definition of the custom testing scenarios, environment, and levels (base maps); and storing the measured and virtual data in a .csv file.

The hardware is embodied in a stationary platform (Figure 2). The literature research shows that researchers often combine UE4 simulation environment with Logitech G920 [50] or G29 [49], [51] steering wheel and pedals. In accordance with these studies, the drivers in our study control the car by means of a Logitech G920, which exhibits a maximum rotation of 900°, and the resolution corresponds to about 0.1°. The car driving control actions rely on visual perception of the simulated scene, mediated by a 49'' Samsung CHG90 QLED gaming display.

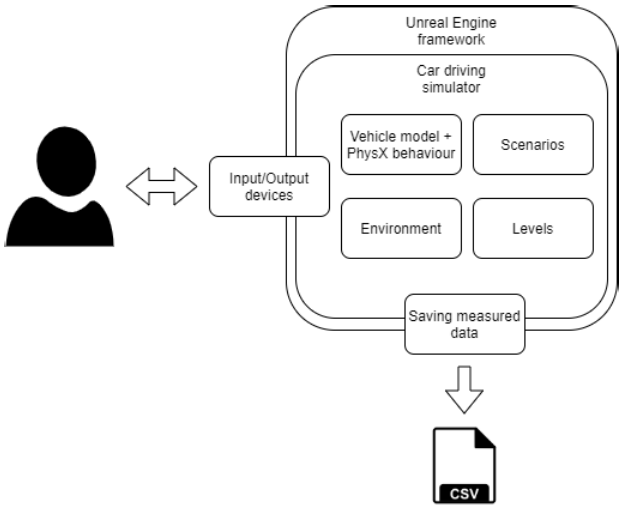

FIGURE 1. The functional block diagram of the car driving simulator (CDS) software.

FIGURE 2. The used simulator platform.

The simulation environment has been in the past enhanced based on feedback of multiple experienced drivers. Taking advantage of highly realistic model of vehicle dynamics, combined with excellent physical model of interaction between vehicle and environent models the resulting level of realism is more than sufficient for the intended purposes. Drivers with all levels of experience have confirmend that adaptation to the simulated vehicle (adaptation to pre-set steering wheel ratio, steering geometry and overal steering feel using the Logitech platform) is fast and easy.

More detailed information about the CDS software and the technical parameters of the hardware can be found in [52].

III. MEASURING THE DRIVER'S RESPONSES

Measuring the responses to various stimuli under different conditions is an indispensable step for state-of-the-art approaches to human driver behaviour assessment, including our scheme. Considering the above-described reasons, the sufficient sampling rate for the data acquisition and the possibility of implementing the desired testing scenarios in particular, we developed our CDS. Several scenarios are available, such as calibration, step response, long-distance drive, sudden obstacle, and Moose test. All of these options allow us to acquire the driving data, involving the steering wheel angle (drivers' control actions), pedal depression rate, car speed, distance from the centre of the lane, and time stamp. A more detailed discussion is presented in source [52].

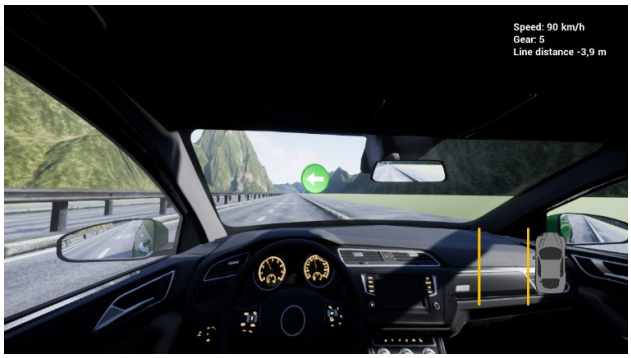

FIGURE 3. The driver's view in the step response scenario.

The scenarios correspond to real-life situations but are maximally simplified to exclude undesired side factors. Thus, we can produce almost ideal testing conditions at a high repeatability. This approach is limited in that the deduced conclusions are valid only under ideal (simulated) conditions and need to be validated in a real-world situation; choosing such a solution, however, is fully justifiable through several aspects. In this context, let us stress above all the fact that the driver behaviour can be modelled using a simple control loop without influencing disturbances, a capability which then leads to easier designing of the behavioural model and simpler identification of its parameters. Respecting the laws of control theory, the obtained model structures are presumably valid also in real-world situations, where the influencing factors will act as disturbances or additional feedback loop(s) (e.g., a haptic feedback) [7].

A. TESTING SCENARIO

For the purposes of the experiment, the results of the step response scenario were considered. The task of the driver rests in responding to the required changes of the driving lane; the requirements are signalized by a green arrow in front of the driver as well as by an implemented lane assistant (Figure 3). From the cybernetics perspective, these changes represent step variations of the input signal (the desired value). All of the measurements are performed at a pre-defined car speed, namely, 90 km/h in the case described herein. The constant speed is maintained by the speed limiter; thus, the driver is only expected to respond to lane changes via the steering wheel. Due to the simplicity of the task, most of the influencing factors can be ignored, and the whole car-driver interaction is characterizable using a simple single-loop feedback structure.

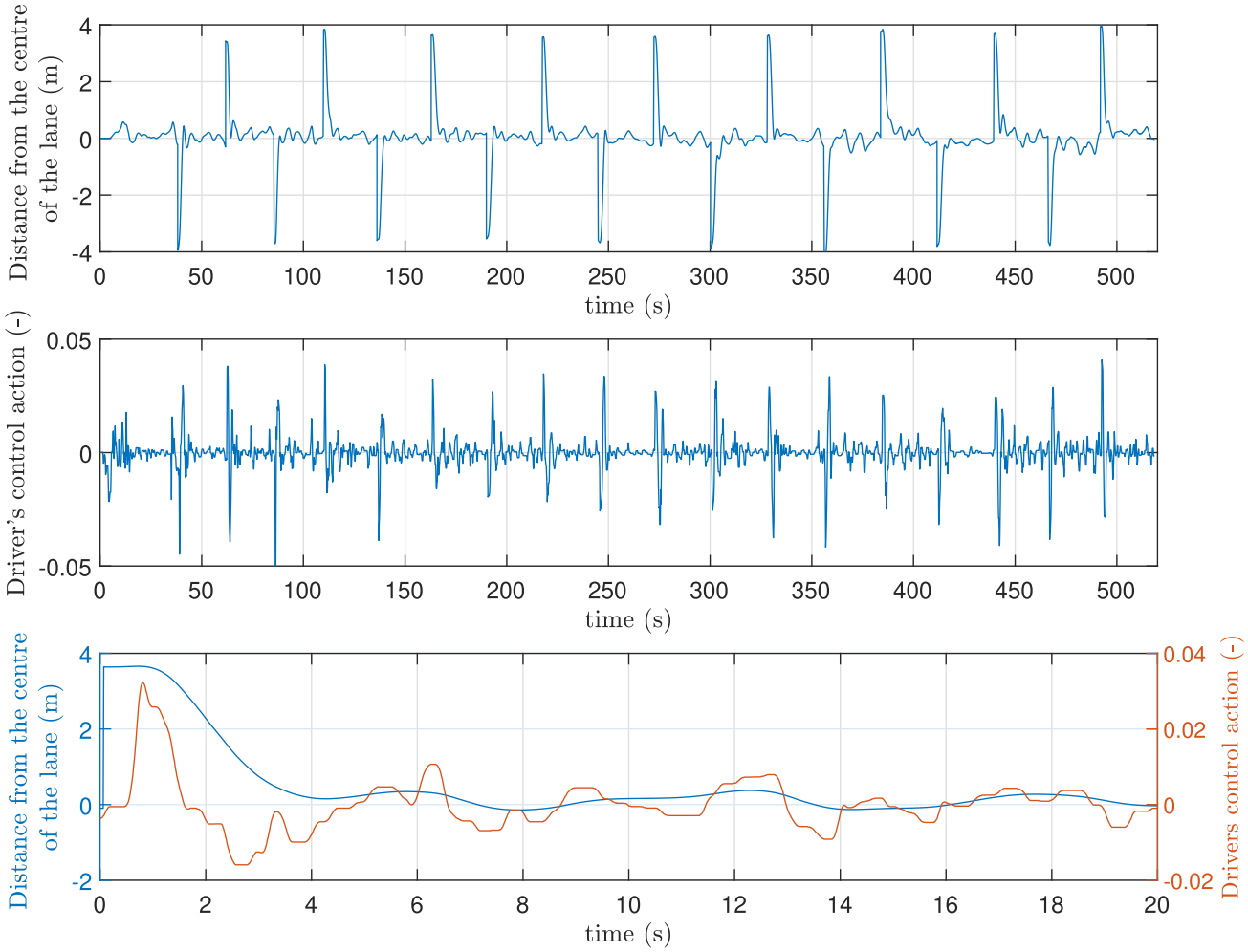

FIGURE 4. An example of the measured data for the driver No. 1.

Step response embodies one of the most important scenarios implemented herein because the measured data enable the creation of a dynamic driver behavioural model and allow us to identify its parameters. The parameters provide broad and complex information about the driver's abilities and limitations, such as the reaction delay or control dynamics. An example of the measured (acquired) data from the step response scenario under the pre-defined conditions is shown in Figure 4; the data are related to the steering wheel rotation (the driver's control actions), $u(t)$, and the distance from the centre of the desired lane (the resulting control error involving the car's response), $e(t)$. The relevant subject is driver No. 1. The detailed response patterns are exposed at the bottom of the figure.

B. SAMPLING AND DATA ACQUISITION

Running the simulation under Windows 10 inevitably entails non-uniformities in the time stamps. The sampling time fluctuated during the measurements, typically attaining values between 8 ms and 11 ms, as illustrated in the histogram in Figure 5. Sampling period of 10 ms was also used, e.g., in [35].

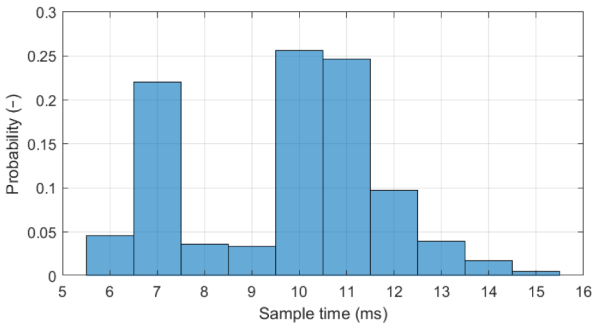

FIGURE 5. The histogram of the evaluated sampling time. The sampling period is much lower than the time constants of human dynamics.

Non-uniform sampling does not degrade signals when performed properly. In this context, let us note that the well-known Nyquist-Shannon-Kotelnikov theorem was expanded to cover non-uniform sampling by Landau [53], who proved that aliasing is avoided if the *average* sampling period $\overline{T}_s$ is linked to the maximal frequency f_{max} via the formula

$$\overline{T}_s < \frac{1}{2f_{max}}. \quad (1)$$

The constraint is, in fact, identical to that of uniform sampling. Thus, the sole disadvantage of the non-uniformities lies in the requirement for signal resampling. As our paper presents an offline identification approach, and the computational time does not constitute an important aspect, the added computational burden is negligible.

We compared the reconstruction methods available, outlining a part of the results within a previously published paper [54]. The band-limited reconstruction technique turned out to be the most accurate. Subsequent experiments in this direction revealed that the method of Cubic Smoothing Splines (CSS) [55] yields nearly indistinguishable results but requires a much smaller overhead than band-limited reconstruction.

All applicable experiments and observations could be of interest within the discussion, but their description would lead us too far from the principal subject; therefore, we can merely state that we utilized the Matlab function *csaps* to resample the CSS [56]. In this manner, non-uniformly sampled signals from the CDS are conveniently transformed into a uniformly sampled time series suitable for the subsequent analyses via the standard signal processing methods.

C. TESTING PROCEDURE AND TESTED GROUP OF DRIVERS

The research was approved by the Ethical Committee for Biomedical Research at the Faculty of Electrical Engineering and Communication, Brno University of Technology, under approval No. 04b/2020.

The measurement involved a group of 92 active drivers that followed a structured testing procedure. In the initial phase, each participant filled out a pre-measuring form with questions concerning their age, driving experience (mileage), subjective self-assessment of the current fatigue condition, and other aspects. All the data were anonymous and are stored together with unique IDs; thus, no personal information can be retrieved, and reverse identification of the tested subject is not feasible. A brief introduction containing indispensable instructions is completed with the calibration scenario, the aim being to adapt the user to the car dynamics and control elements. This scenario lasts approximately 5 minutes, and no data are logged. Subsequently, the main testing scenario is launched. In the case of the step response scenario, a predefined number of lane changes are applied. The overall duration of the testing scenario is 8–9 minutes, and all of the described data are recorded.

For the sake of clarity, the detailed discussion of the identification methods and its results in Section IV are presented for the first ten drivers; the information on this subset is summarised in Table 1. Subsection V then analyses the whole test group of 92 drivers.

TABLE 1. Information concerning the first ten tested drivers.

Driver ID	Age	Driving licence (years)	Mileage per year (km/year)	Tiredness selfassessment (rel. 1-10)
1	24	4	5,000	3
2	23	6	10,000	1
3	26	8	10,000	1
4	27	9	40,000	5
5	29	9	2,000	4
6	60	42	25,000	6
7	25	7	25,000	2
8	25	7	3,000	5
9	33	15	10,000	2
10	46	28	14,000	6

IV. DYNAMICAL DRIVER BEHAVIOR MODELS

As presented in the Introduction, this article focuses on the lane changing experiment and the description of acquired data using the simplest single-loop models, which exhibit a minimal number of parameters. The models are based on McRuer's Crossover law [29]. According to this theory, the human operator (at the control level, [57]) acts as a biological feedback controller having a transfer function, $F_R(s)$, and adapting his or her behaviour to the dynamics of the controlled element, $F_C(s)$. Knowing the dynamic model, $F_C(s)$, enables us to define the structure of the appropriate model of the controller, $F_R(s)$, to satisfy the open-loop transfer function, $F_0(s)$, in the form

$$F_0(s) = F_R(s)F_C(s) = \frac{K_R K_C}{s} e^{-\tau s}, \tag{2}$$

where s denotes the Laplace operator, K_R and K_C are the static gains of the human controller and the controlled element, respectively, and τ represents the reaction delay of the human operator [29].

The defining equation (2) describes the manner in which the human element adapts to the controlled dynamics. However, several experiments have proved that the prescription is valid only in the area closest to the crossover region and does not consider a broader frequency range, the dynamics of the neuromuscular (NM) system, and other relevant factors [29]. Thus, more complex models must be utilized to provide a sufficient approximation of the human behaviour [7], [29]. The complexity of the models can differ, respecting, above all, the desired precision; further, no unambiguous rule (based on a mathematical tool) is available to choose a suitable structure.

Using linear system theory, the lateral motion of a car driving at a constant speed can be approximated in a simplified way (within the range considered) as a double integrator having a certain gain, K_C, reading

$$F_C(s) = \frac{Y(s)}{U(s)} = \frac{K_C}{s^2}. \tag{3}$$

Here the $U(s)$ stands for the Laplace transform of the steering wheel angle $u(t)$, and the $Y(s)$ denotes the Laplace transform of the car lateral position $y(t)$.

The model (3) roughly approximates an actual non-linear model under small steering wheel angles, and becomes very accurate for the fast drive on the highway. During our testing scenario, where the constant speed of 90 km/h is held by the cruise control, the heading angle varies in the range of $\pm 10°$. For this reason, the longitudinal velocity decreases by a very

small factor of 1 – cos(10°) which is approximately 1.5% or 1.4 km/h. The authors are aware of possible lags representing the dynamics of the steering mechanism or delays caused by the sampling, screen refresh rate (120 Hz), and other factors. Presumably, however, the impact of the dynamics is virtually negligible compared to that of the main dynamics (embodied in the double integrator), especially in the middle-frequency area. Thus, in the text below, the controlled element dynamics are considered according to (3).

In order to comply with Crossover law see [29], the model structure describing the dynamics of the human controller needs to be a transfer function in the form of (4). We then have

$$F_{\mathrm{R}}(s) = \frac{U(s)}{E(s)} = K_{\mathrm{R}}\, s\, \mathrm{e}^{-\tau s}, \tag{4}$$

where $E(s)$ is the Laplace transform of the lateral control error $e(t)$, which is the lateral distance from the centre of the desired lane. The model (4) describes the behaviour of the human (driver) as a pure derivative controller with a gain K_{R} and a reaction delay τ. This structure compensates a pole in the controlled element. Such a derivative system, however, would predict a Dirac impulse in the driver's response to a step change of the input signal. This type of response cannot be physically realized and is therefore not observable in the measured data (Figure 4).

Although model (4) agrees with the fundamental theory, it is only a theoretical formula valid in the middle-frequency regions or in the area adjacent to the crossover frequency. In a model reflecting a broader frequency range, the base formula must be extended with further dynamics, and the resulting formula must involve one or more poles. Such pole(s) would account for the NM dynamics, then the dynamics of the steering wheel, and those of other related factors or components. Although there are instances when the number of poles is determined according to the author's experience [31], [58], McRuer suggests that the selection of the model structure should not be dictated by the experimenter's intuition but by the data [29].

Embodying suitable forms to approximate the measured drivers' responses, the following four model structures were investigated:

$$F_{\mathrm{R}}(s) = \frac{K_{\mathrm{R}} s}{Ts+1}\, \mathrm{e}^{-s\tau}, \tag{5a}$$

$$F_{\mathrm{R}}(s) = \frac{K_{\mathrm{R}} s}{T^2 s^2 + 2\xi T s + 1}\, \mathrm{e}^{-s\tau}, \tag{5b}$$

$$F_{\mathrm{R}}(s) = \frac{K_{\mathrm{R}} s}{(T^2 s^2 + 2\xi T s + 1)(T_2 s + 1)}\, \mathrm{e}^{-s\tau}, \tag{5c}$$

$$F_{\mathrm{R}}(s) = \frac{K_{\mathrm{R}} s (T_3 s + 1)}{(T^2 s^2 + 2\xi T s + 1)(T_2 s + 1)}\, \mathrm{e}^{-s\tau}. \tag{5d}$$

These transfer functions describe the dynamical relationship between the drivers' inputs (the distance from the centre of the desired lane), $e(t)$, and their control actions, namely, the angle of the steering wheel with respect to its idle position, $u(t)$. The angle is expressed in relative values between −1 and 1, correspondingly to the endpoints of the rotation (see Section II). Each of the models includes at least one derivative component-related zero and the reaction delay, τ. Thus, the complexity of the models differs especially in the number of poles or, possibly, in using two zeros in the case of the model (5d).

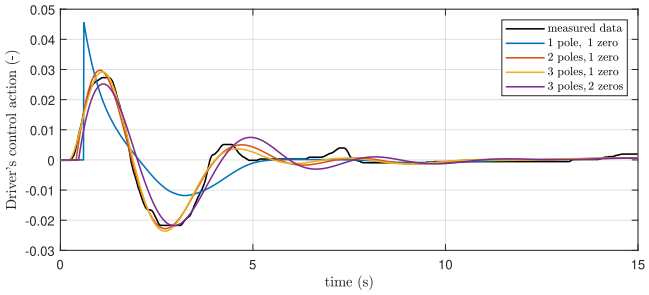

FIGURE 6. Approximating the selected driver's control action with the models in (5).

Approximating one selected driver's control action via the models specified in (5) is illustrated in Figure 6. The proposed comparison of the outputs and the measured transient response is associated with the data indicated in the bottom part of Figure 4. The presented models were obtained using the `tfest` function described in Section IV-B.

A. CROSS-VALIDATION METHOD

Choosing the right complexity of a model embodies one of the problems of statistical learning; this issue, however, can be resolved via various common and convenient model selection methods [38], [59]. Our research relies on the leave-one-out cross-validation procedure. The approach, despite being routinely employed in statistics, is rather scarce in the literature concerned with human driver modelling; Nash and Cole [13], for example, use the last 30 seconds from each experiment to run hold-out validation. As already mentioned above, a recorded drive consists of K-step responses, and we therefore split the signal into segments containing the individual responses. With this type of data, cross-validation may be performed in the manner outlined below.

We set the first response aside to facilitate the validation, and all the remaining $K-1$ responses are utilized in identifying the first model, whose quality is then assessed using the validation step response as described in the following parts of the text. At the next stage, the second response is set aside, and the other $K-1$ responses are employed to identify a new model. The process continues until each response has been used for the validation.

In each model structure (5), we will identify a set of K models denoted by $\{F_{\mathrm{R}k}(s)\}_{k=1}^{K}$. To define the process exactly, we introduce the notation below. The output of a model is related to its input via the formula

$$\hat{u}_{l,k}(t) = \mathcal{L}^{-1}\{F_{\mathrm{R}k}(s)\mathcal{L}\{e_l(t)\}\}, \qquad k, l = 1, 2, \ldots, K, \tag{6}$$

where the $\mathcal{L}$ and $\mathcal{L}^{-1}$ denote the forward and inverse Laplace transform, respectively; the index l signifies that the system responds to lth input signal $e_l(t)$; and the index k refers to the kth model.

To train (identify) the kth model, we leave out one (kth) experiment for validation. The remaining $K-1$ responses then define training mean squared error (MSE)

$$E_{\mathrm{t},k} = \frac{1}{N(K-1))} \sum_{\substack{l=1 \\ l\neq k}}^{K} \sum_{n=0}^{N-1} \left[u_l(nT_\mathrm{s}) - \hat{u}_{l,k}(nT_\mathrm{s})\right]^2, \quad (7)$$

and the validation MSE

$$E_{\mathrm{v},k} = \frac{1}{N} \sum_{n=0}^{N-1} [u_k(nT_\mathrm{s}) - \hat{u}_{k,k}(nT_\mathrm{s})]^2. \quad (8)$$

Equations (7) and (8) conceal the fact that the modelled output $u_{l,k}(nT_\mathrm{s})$, and therefore the errors, are functions of the free parameters T, K_R, τ and possibly also ξ, T_2, T_3, depending on the model structure selected from (5).

The training MSE, $E_{\mathrm{t},k}$, is necessary for obtaining the parameters through the optimization procedure involved in the identification methods. The training MSE, $E_{\mathrm{t},k}$, tends to decrease whenever we increase the model complexity; however, the validation MSE, $E_{\mathrm{v},k}$, may rise once the model order has reached too high. Thus, the training MSE, $E_{\mathrm{t},k}$, does not embody a good measure for selecting the model complexity [38].

B. IDENTIFYING THE MODEL PARAMETERS

The parameter identification is enabled by diverse approaches. Our research relies primarily on the `tfest` function in the Matlab System Identification Toolbox. The `tfest`, in the first phase, roughly estimates the model via the Instrumental Variable (IV) method and then refines its parameters through non-linear minimisation of the squared error of the output error (OE) model [60].

Nonlinear minimisation, however, may entail multiple local minima, and the algorithm may yield a sub-optimal solution. This condition occurred in the second- and third-order models (5). We sought the transport delay, τ, by calling the `tfest` for all the integer multiples of the sampling period, T_s, between 0 and 1.2 seconds. The entire identification was repeated from the starting values of $T = \xi = 0.3$ and $K = 0.01$, allowing us to alleviate the problem at the cost of time.

The input (measured) data for each driver include several requirements for a step change of the lane to induce relevant driver responses (Figure 4), thereby suggesting several ways of applying the identification function, namely:

- global minimisation of the error (7),
- obtaining parameters from each response individually,
- various modifications of these approaches.

In the global approach, the `tfest` seeks the minimum of the training error (7). The $K-1$ training responses are passed to the function as a multi-experiment `iddata` structure,

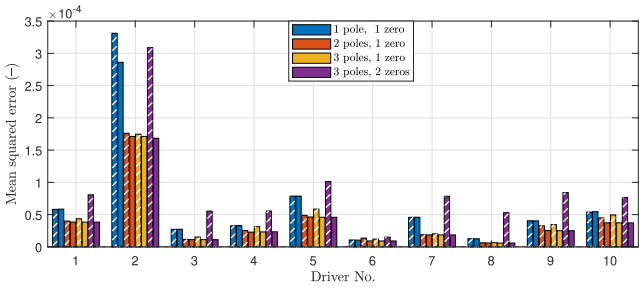

FIGURE 7. The cross-validation MSE of models obtained via global minimisation (plain bars) and using parameter averaging (hatched).

as described in the documentation [60]. The process is repeated K times to deliver a model for each cross-validation fold. The whole dataset facilitates the identification process; in such a case, however, the reaction delay differences in the individual responses cannot be fully accounted for by one model. A considerable variance is observable in the measured data, making the authors concerned that this phenomenon might affect the estimation of the parameters. Identifying one model from each response appears to be a suitable solution to the issue because it allows the model delay adjustment for each step response. Thus, $K-1$ individual delays τ, gains K_R, time constants T, damping ξ, and other parameters are obtained. As we aim to create one model for each driver, constituting all $K-1$ training responses, the individual models need to be merged; this is accomplished by averaging where the arithmetic mean of each model parameter is computed.

The third alternative lies in using FFT-based techniques to compute the model output in (7). This option enabled us to construct custom cost functions to verify the multiple hypotheses that certain parameters should be shared by all the training responses while others should vary for each response, similarly to the situation in parameter averaging. The cost function was minimised via standard iterative optimisation methods. The Quasi-Newton method with Levenberg-Marquardt damping was evaluated as preferable due to its fast convergence and high accuracy [61], [62]. Different propositions as to which parameters are independent or common for all the training folds were tested, involving, for instance, a recipe where all parameters were common but the transport delay varied with the responses. However, we did not reveal appreciable improvement in the prediction accuracy (the rates often differed by fractions of per cent). The conclusion lies in that all model parameters should be common for all the training responses; therefore, the `tfest` estimation using global minimisation appears to be the best option in terms of the model accuracy and the simplicity of the practical implementation. A slightly more detailed discussion is provided in the next subsection.

C. CROSS-VALIDATION RESULTS

As there are K models validated on K responses, the overall cross-validation error embodies the mean of validation errors

from individual folds:

$$\overline{E}_{\mathrm{v}} = \frac{1}{K}\sum_{k=1}^{K} E_{\mathrm{v},k}. \tag{9}$$

MSE are shown in Fig. 7, and can be used to compare the quality of different models for the same driver. However, MSE also depends on the amplitude of the measured driver's responses, and therefore when we compare models between individual drivers it is more convenient to adopt as the measure of model quality a relative measure, such as model efficiency (ME) or variance accounted for (VAF). The cross-validation ME is defined by the formula

$$\mathrm{ME} = 1 - \frac{\sum_{k=1}^{K}\sum_{n=0}^{N-1}[u_k(nT_\mathrm{s}) - \hat{u}_{k,k}(nT_\mathrm{s})]^2}{\sum_{k=1}^{K}\sum_{n=0}^{N-1}[u_k(nT_\mathrm{s}) - \overline{u}_k]^2}. \tag{10}$$

with

$$\overline{u}_k = \frac{1}{N}\sum_{n=0}^{N-1} u_k(nT_\mathrm{s}) \tag{11}$$

being the mean value of measured steering wheel angle in kth data fold. Since the experiment takes place on a straight motorway, all means $\overline{u}_k$ are virtually nil; otherwise, its second integral would grow until the car would exit the motorway. Hence ME will tally with VAF, which was adopted by [11], [12], [24],

$$\mathrm{VAF} = 1 - \frac{\sum_{k=1}^{K}\sum_{n=0}^{N-1}[u_k(nT_\mathrm{s}) - \hat{u}_{k,k}(nT_\mathrm{s})]^2}{\sum_{k=1}^{K}\sum_{n=0}^{N-1} u_k^2(nT_\mathrm{s})}. \tag{12}$$

The values obtained through these equations depend on the structure of the model and reflect the structure's generalisation power. The cross-validation ME may be employed to objectively guide the experimenter in selecting a proper model. The ME of a model being trained always increases as the model order rises. It might be less known in human driver modelling that the cross-validation ME does not always increase with the model order and carries a potential to rise when the order reaches excessively high. In machine learning, this scenario is commonly referred to as over-fitting, where the model fits the training data with an excellent precision but lacks the prediction capability with unobserved or new data [38].

Models with increasing complexity are compared in Figure 8. The hatched bars correspond to models obtained by averaging the parameters from the individual responses, and the plain bars denote models identified by means of global minimisation; both methods were presented in Section IV-B. This usage applies to any bar figure in the article.

The first-order model (5a) is evidently too simple to account for human actions, being outperformed by more complex models in terms of the ME. Furthermore, it is assumed that the dynamical model of the human operator contains a direct feed-through signal path, as illustrated by the instantaneous step rise (Figure 6, blue). This results in the driver's ability to transmit infinitely large frequencies with a gain, K_R/T; such a capacity, however, is not realistic.

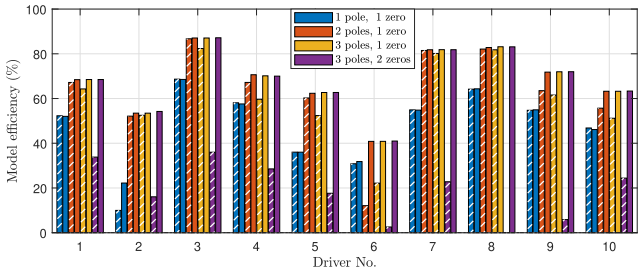

FIGURE 8. The cross-validation ME of models obtained via global minimisation (plain bars) and using parameter averaging (hatched). The model involving two zeros yields a negative ME for driver No. 8 when identified with parameter averaging.

Nevertheless, the model ensures a decent approximation of the human actions, with the ME values close to those of the more complex models.

The second-order model (5b) appears to maximise the cross-validation MEs, suggesting that the model exhibits the highest prediction accuracy. Our data invariably yield $\xi < 1$, a model with a damped oscillatory character (Figure 6, yellow). The model will be described in detail in the sections below.

The first third-order model (5c) expands the previous model by filtering its input (or output) through a low pass filter with the time constant T_2. We will not attempt to infer whether such an assumption is justified, physically or otherwise; the importance of this time constant may be conveniently assessed by the prediction error. As suggested by the MEs in Figure 8, introducing the constant does not have an appreciable effect. The response in Figure 6 (yellow) also documents that the model is essentially the same as model (5b) because the added pole always exhibits a very small time constant (not illustrated in the figure).

To assure the reader that a further increase in the complexity will not enhance the overall benefit, we test model (5d), obtained from model (5c) by adding another zero. As illustrated in Figure 8 (magenta), there is no appreciable improvement in the cross-validation error; it then follows that the second-order model (5b) has the optimal number of poles and zeros. In the remaining part of the article, model (5b) will be referred to whenever we write F_R.

Based on the results presented in Figure 8, we can also conclude that parameter averaging performs best with simple models (first order) and that its applicability deteriorates as the model complexity increases. The effect is prominent in Figure 8 (hatched bars), where both third-order models often provide clearly worse predictions than the second-order ones. This rise in the error rate can be fully attributed to the averaging method because the effect does not occur when the custom functional is used (plain bars). The parameter averaging impairs the resulting model, presumably due to the large variance in the model parameters induced when identifying a complex model from a sole step response. This is especially pronounced in both of the third-order models. In contrast, differences in the cross-validation errors are less

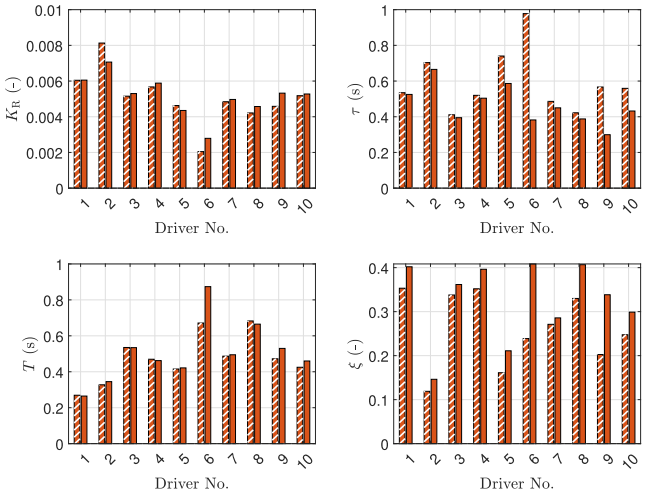

FIGURE 9. Comparing the identified parameters of model (5b) in ten individual drivers, obtained via global minimisation (plain bars) and using the averaging of parameters (hatched).

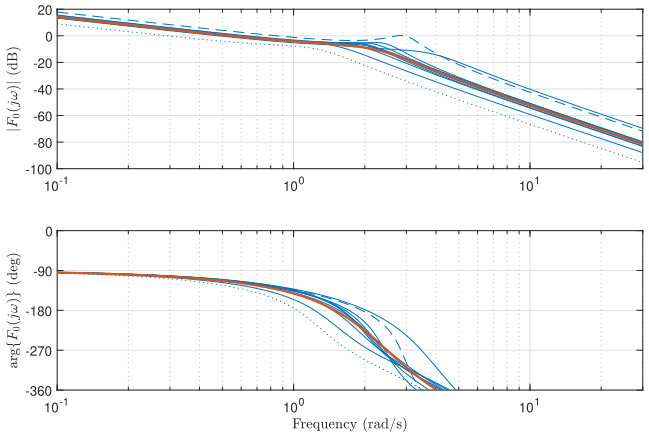

FIGURE 10. The open-loop frequency responses $F_0(j\omega)$ in the individual drivers, associated with the controlled element according to (3), where $K_C = 85$.

pronounced in the first- and second-order models. Since the model averaging seldom produces models more accurate than the global application of the `tfest`, we find the latter method preferable.

D. IDENTIFIED PARAMETERS OF THE TRANSFER FUNCTION MODEL

The parameters of the transfer functions in the form of (5b) were identified using the aforementioned `tfest`-based procedures: global minimisation (multi-experiment) and model averaging. The results from the first ten drivers are summarised in Figure 9. Broader analysis of the results from more drivers is in Section V.

While the model structure respects the control laws and is thus common for the individual drivers, the values of the identified parameters define the differences in the dynamic behaviour and are individual to some extent. The parameters therefore have a direct impact on the resulting control action. The impact of the parameters on the dynamic behaviour of a driver is usually demonstrated through the frequency response (Bode diagram).

The open-loop frequency response (corresponding to (2)) of the individual drivers, associated with the controlled element according to (3), is presented in Figure 10 (blue lines). The gain of the controlled element, K_C, was identified from the measured data ($K_C = 85$).

Most of the frequency responses are very similar and approach the evaluated average (red line). The average crossover frequency is about 0.5–0.6 rad/s; however, two (or three) more pronounced deviations are observable.

The first outlier, the dashed blue line, relates to the youngest participant, driver No.2. We can observe the highest resonance peak as well as the highest crossover frequency (around 0.7 rad/s), both of which affect the ability to deliver a fast response to sudden changes. This characteristic also relates to the values of the identified parameters, namely, the highest value of the gain K_R and the low values of the time constant T and damping ξ. Conversely, the driver's reaction delay, τ, is objectively the highest. Similar behaviour can be identified in driver No.1, the second youngest driver of the tested group; the similarity lies in a higher value of the gain, K_R, and the low value of the time constant, T. However, in contrast to driver No.2, the damping value ξ ranges among the highest ones.

The second case, the dotted blue line, represents the frequency response of a sixty-year-old driver (No. 6). The characteristics contain an almost nonexistent resonance peak, and the crossover frequency is very low (about 0.3 rad/s) in contrast to those of the other drivers, driver No. 2 in particular. In this behavioural pattern, the responses of driver No. 6 are very slow but, at the same time, no overshoots are present.

E. RELATION TO MULTI-LOOP MODEL

In order to clarify the relationship between our model and multi-loop control, we provide a possible transformation of the transfer functions obtained in the previous section. The comparison will be accomplished via decomposing the model in (5b) into a multi-loop model which is most common model in the recent literature [28], [34], [63], [64]. In a multi-loop model, it is typically assumed that a human driver is capable of perceiving the vehicle's lateral position, $y(t)$, and heading angle, $\psi(t)$. Assuming a constant velocity v_0, kept by the speed limiter in this testing scenario, the relation between the vehicle's lateral position, $y(t)$, and heading angle, $\psi(t)$, can be evaluated as (the approximate equality applies to small values of $\psi(t)$) [65]

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = v_0 \sin\psi(t) \approx v_0\psi(t). \tag{13}$$

with $v_0 = 25$ m/s (90 km/s) in our experiments. Thus, there are two feedback loops in the first block diagram, Figure 11 (a).

Rearranging the inner feedback loop, as shown in Figure 11 (a), and applying block algebra will yield, through

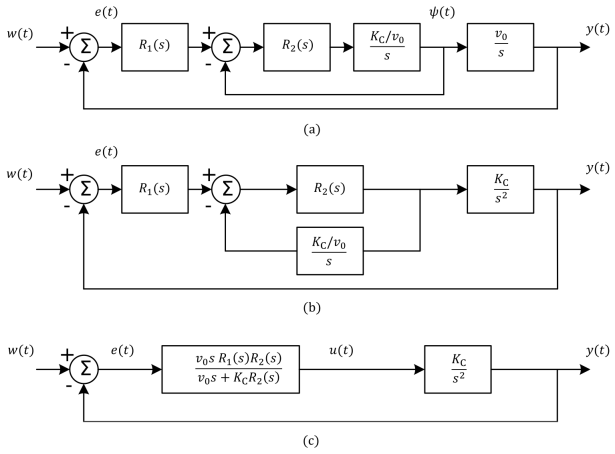

FIGURE 11. The multi-loop model (a) and its modification into a single-loop model (c) through (b) using block algebra rules.

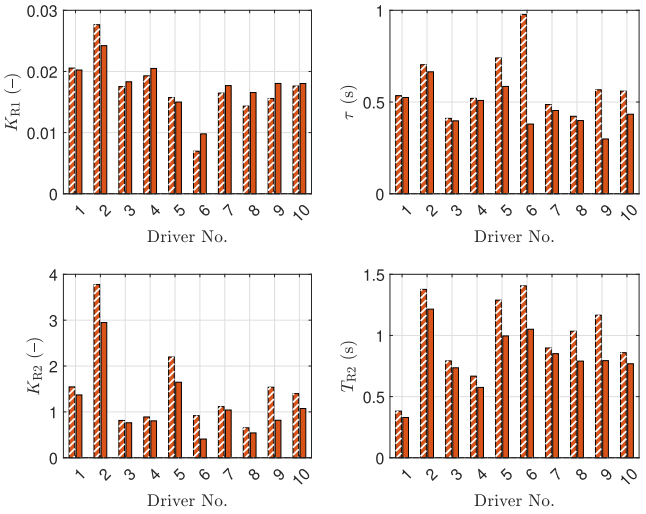

FIGURE 12. Comparing the identified parameters of model (16) in ten individual drivers, obtained via global minimisation (plain bars) and using parameter averaging (hatched).

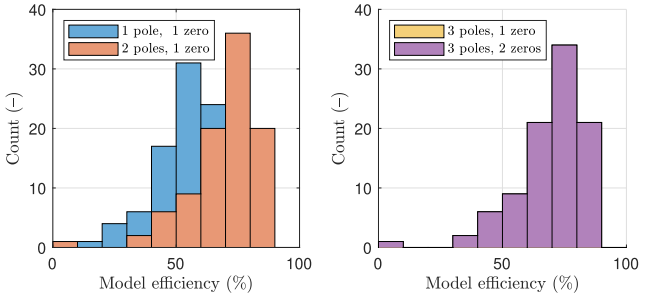

FIGURE 13. The cross-validation model efficiency obtained via global minimisation. The first order model has the lowest efficiency. All second- and third order models exhibit virtually the same prediction capabilities.

Figure 11 (b), the single-loop model exposed in the last diagram, Figure 11 (c). The dynamics of the single-loop model $F_R(s)$ are determined by the feedback $R_1(s)$ and $R_2(s)$ of the multi-loop model. In the diagram, the input of the $R_1(s)$ is the lateral distance $e(t)$ from the centre of the desired lane, and its output is the required lateral velocity. The deviation from the required lateral velocity enters the second controller, $R_2(s)$, whose output is the steering wheel angle $u(t)$. The transfer functions of these feedback paths, however, cannot be uniquely determined if only the $F_R(s)$ is known. Multiple solutions to the inverse problem are outlined below. As an example, we provide the following three different admissible options:

$$R_1(s) = \frac{K_R K_C e^{-\tau s}}{v_0(T_A s + 1)}, \qquad R_2(s) = \frac{v_0}{T_B K_C}; \tag{14}$$

$$R_1(s) = \frac{K_R K_C e^{-\tau s}}{v_0(T_B s + 1)}, \qquad R_2(s) = \frac{v_0}{T_A K_C}; \tag{15}$$

$$R_1(s) = K_{R1} e^{-\tau s} = \frac{K_R K_C}{v_0} e^{-\tau s},$$

$$R_2(s) = \frac{K_{R2}}{T_{R2} s + 1} = \frac{v_0}{2\xi K_C T \left(\frac{T}{2\xi} s + 1\right)}. \tag{16}$$

Here, T_A and T_B stand for the time constants of the single-loop human controller

$$F_R(s) = \frac{K_R s e^{-\tau s}}{(T_A s + 1)(T_B s + 1)}, \tag{17}$$

if the controller is overdamped, i.e. when $\xi > 1$. We may, of course, construct other pairs, $R_1(s)$ and $R_2(s)$, but these would entail more poles and zeros. Regardless of the shortcomings, we selected the simplest possible decomposition of the $F_R(s)$, as direct identification of the $R_1(s)$ and $R_2(s)$ with the scenario employed in this study is not practicable. Furthermore, the analysis of the experimental data excludes models (14) and (15) because all instances of the model $F_R(s)$ identifi ed herein yielded $\xi < 1$. Thus, (16) is the only valid solution from the aforementioned three pairs. This model structure is very similar to the model advanced by Schnelle et al. [34]. The evaluated parameters of the $R_1(s)$ and $R_2(s)$ structures according to (16) are presented for ten drivers in Figure 12 and results from more drivers are presented in Section V.

V. STATISTICAL ANALYSIS

The presented procedure was then verified by being applied to a statistically more significant group of drivers. This group included 92 active drivers (10 drivers from the previous subsections and 82 other participants), encompassing a relatively large spectrum of participants, namely, men and women aged between 18 and 60 that cover various mileages per year.

A. ANALYSIS OF THE TRANSFER FUNCTION MODEL

The cross-validation model efficiency (ME) obtained via global minimisation is presented using histograms, Fig. 13. It is obvious from the graphs that the conclusions drawn from 10 participants also translate to the above-mentioned larger group comprising 92 persons. The first-order model yields on

TABLE 2. The centres and covariance matrices of the two components of the GMM.

$\boldsymbol{\Sigma}_1$	K_R	τ	T	ξ
K_R	$2.6 \cdot 10^{-6}$	$-3.3 \cdot 10^{-5}$	$-1.4 \cdot 10^{-4}$	$-7.0 \cdot 10^{-6}$
τ	$-3.3 \cdot 10^{-5}$	$1.2 \cdot 10^{-2}$	$3.1 \cdot 10^{-4}$	$-3.6 \cdot 10^{-3}$
T	$-1.4 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$	$1.1 \cdot 10^{-2}$	$2.0 \cdot 10^{-3}$
ξ	$-7.0 \cdot 10^{-6}$	$-3.6 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	$5.2 \cdot 10^{-3}$
$\boldsymbol{\mu}_1$	$5.8 \cdot 10^{-3}$	0.39	0.44	0.28
$\boldsymbol{\Sigma}_2$	K_R	τ	T	ξ
K_R	$7.2 \cdot 10^{-7}$	$-3.7 \cdot 10^{-6}$	$-1.1 \cdot 10^{-4}$	$2.4 \cdot 10^{-5}$
τ	$-3.7 \cdot 10^{-6}$	$7.1 \cdot 10^{-3}$	$-2.4 \cdot 10^{-3}$	$-4.0 \cdot 10^{-3}$
T	$-1.1 \cdot 10^{-4}$	$-2.4 \cdot 10^{-3}$	$2.8 \cdot 10^{-2}$	$-3.2 \cdot 10^{-3}$
ξ	$2.4 \cdot 10^{-5}$	$-4.0 \cdot 10^{-3}$	$-3.2 \cdot 10^{-3}$	$1.3 \cdot 10^{-2}$
$\boldsymbol{\mu}_2$	$4.4 \cdot 10^{-3}$	0.32	0.67	0.43

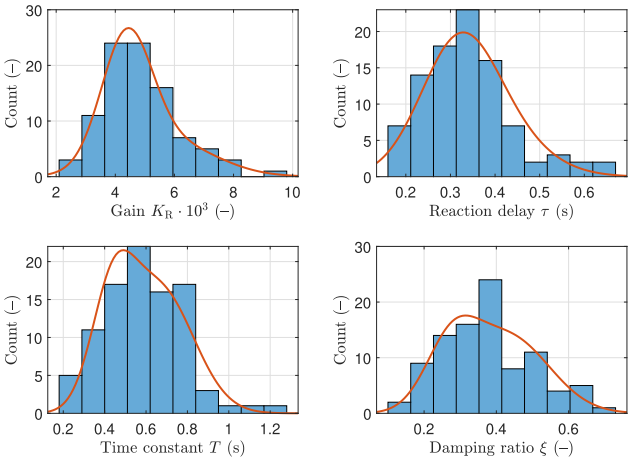

FIGURE 14. Histograms of the identified parameters of the transfer function model (5b) obtained via global error minimisation on 92 drivers. The red line displays the PDF of the GMM (scaled to outline the same area as the histograms).

average efficiency of 54 %, while the others offer 69 %; this observation confirms that the second-order model is optimal even for a larger group of participants. The third-order option does not improve the cross-validation accuracy and entails a higher complexity, a property which then hinders the actual intepretability, as we employ more parameters to describe the same behaviour.

Once it has been established that our dataset can be best characterised via the second-order model, we launch an analysis of its parameters. These are presented in Fig. 14, where we utilize histograms to visualise extensive data groups in a better manner than otherwise possible with bar graphs.

Yet even the histograms capture only the distribution for the individual parameters, and their interactions remain concealed. Thus, to improve recording the distribution of the drivers' parameters, we trained a Gaussian mixture model (GMM) [66]. Apparently, the drivers are separated into two clusters, and we therefore model the probability of the parameters $\mathbf{x} = [K_\mathrm{R}\ \tau\ T\ \xi]$ as a mixture of two normal distributions:

$$p(\mathbf{x}) = 0.34\,\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + 0.66\,\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2). \qquad (18)$$

One thousand GMMs were trained, starting from different initialisation values, using the Matlab `fitgmdist` function; the model with the highest likelihood is presented. Their values' means, $\boldsymbol{\mu}$, and covariance matrices, $\boldsymbol{\Sigma}$, are listed in Table 2. The probability density functions (PDFs) of the individual variables are scaled and shown with the histograms in Figure 14. The GMM (18) can find use in the future research activities, such as the Monte Carlo simulation of drivers, where dynamical models of drivers are randomly generable based on this distribution.

The evaluated parameters have a direct impact on the dynamical behaviour of a driver; this effect can be demonstrated through the open-loop frequency response (corresponding to (2)). The gain K_R influences the value of the cross-over frequency related to time to reach the desired value. The ξ component corresponds with the damping of the oscillations in the controller. The reaction delay, τ, embodies the standard metrics related to human information processing and control and is responsible for the phase shift in the frequency response. In general, its presence degrades the dynamical properties of the control loop. The graph in Fig. 15 exposes the frequency responses of the control loop for the statistical set of drivers controlling the same system defined as (3) with $K_\mathrm{C} = 85$. The dispersion of the characteristics corresponds with the histograms in Fig. 14. This effect follows the main advantage of the simple model form, when only 4 parameters can fully represent the dynamical properties of the human behaviour during a defined control task, i.e. they have the same informative value as the frequency responses and can be then simply processed by machine-learning or statistical tools. Such a capability stands in contrast to more advanced structures which find application in more complex tasks. To provide an example, the red curve represents the control functions of an average driver (the mean of the frequency responses). Then, the abilities of the other drivers can be relatively compared within the statistical group or to the average driver. This step could embody an easy path to evaluating the driver performance and may be utilized in, for instance, deciding about the driving competences; such an application, however, still needs to be subjected to an in-depth and precise research.

B. ANALYSIS OF THE MULTI-LOOP MODEL

Provided that we adopt the transformation of the single-loop model $F_\mathrm{R}(s)$ into a multi-loop one, as discussed in Section IV-E, the time constant of the inner feedback loop $R_2(s)$ is uniquely determined by the original transfer-function model $F_\mathrm{R}(s)$. The evaluated parameters of the $R_1(s)$ and $R_2(s)$ structures according to (16) are presented for 92 drivers using histograms in Figure 16.

The time constant $T_{\mathrm{R}2}$ attains values in a broader range, with 95 % confidence $T_{\mathrm{R}2} \in \langle 0.32\,\mathrm{s}, 1.51\,\mathrm{s}\rangle$. Schnelle et al. [34] performed tests on 10 drivers but report

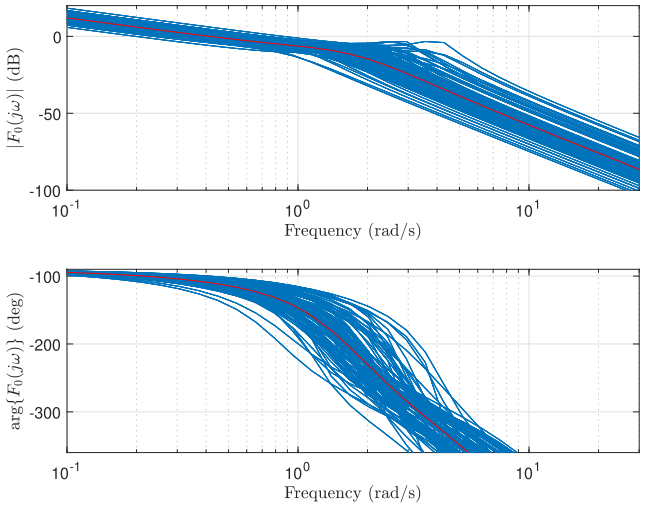

FIGURE 15. The open-loop frequency responses $F_0(j\omega)$ for the statistical group of 92 drivers, associated with the controlled element according to (3), where $K_C = 85$.

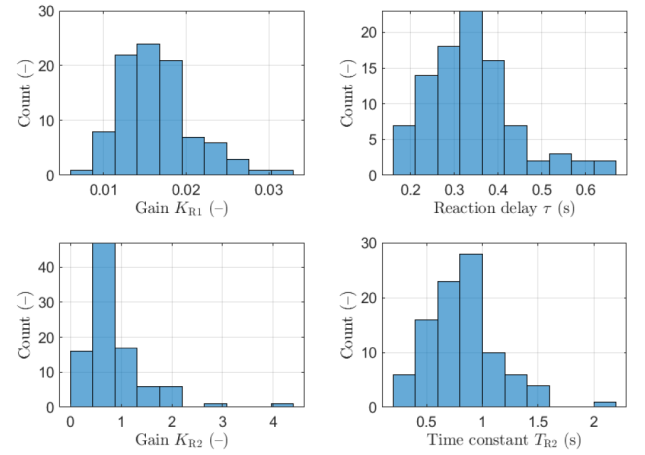

FIGURE 16. Histograms of the identified parameters of the multi-loop model (16) from 92 drivers, obtained via global error minimisation.

only four values of this constant. Only one of their values lies outside our confidence interval.

The obtained interval suggests that the time constant is not directly related to the so-called NM constant discussed in articles [29], [31], [58], [63] and that its interpretation is markedly hypothetical. Presumably, the quantity corresponds to the time constant of a system comprising both the steering wheel and the NM system, involving the driver's arms. Similar ambiguity in the identification of the NM constant was reported by Nash and Cole [13], Pick and Cole [67], and through an extensive review of human pilots' dynamics proposed by Xu et al. [6].

The transport delay of $\tau = 0.15$ s suggested by Hess et al. [58] or Nash and Cole [13] appears to be manifestly smaller than our values, and lies outside our 95 % confidence interval $\tau \in \langle 0.20\,\text{s},\ 0.61\,\text{s}\rangle$. This discrepancy, in fact, may be true, as the setting of the referenced experiments differs from ours, especially as regards the forcing function. In our study, the drivers responded to step changes, but the respondents in the research by Hess et al. focused on the curvature of the road; Nash and Cole project, by extension, employed forcing functions and disturbances in the form of low-pass coloured noise. In both of the studies, the drivers could predict further changes in the required lateral position, thus exhibiting smaller delays. The low-pass forcing functions used in the two reports are simpler to predict by means of band-limited extrapolation: An unobserved part of a band-limited signal may be predicted in a small future interval, whose length increases when the band-limit or the signal noise level decrease [68], [69]. Furthermore, the employment of the fixed-based simulator induces higher delays in the drivers' responses, in accordance with the conclusions outlined in [24].

VI. CONCLUSION

The aim of the research was to identify an efficient model to accurately approximate the driver responses during the simulated experiments, thus facilitating comparison between the drivers within the test group.

According to the researched literature, the complex, non-linear models find wide use in critical applications, including modelling the reactions during racing scenarios where a wheel slip occurs. In contrast, our testing setup focuses on step response experiments in the form of a lane changing task during a simulated highway drive under precisely defined conditions. Thus, based on the literature, we focused on linear models, which involve transfer functions or state feedback controllers.

Utilizing the cross-validation method, we successfully determined that, in our setting, a non-complex cybernetic model is suitable for observing and analyzing the driver reactions. The second-order transfer-function model provides the best prediction capability in terms of the ME and characterises human actions in standard lane-changing situations more effectively than the higher-order models. Such a capability is then potentially very beneficial in reducing the computational burden, allowing markedly simpler interpretation of the four relevant parameters that have a definite impact on the drivers' frequency responses. The order is markedly lower than those specified in the literature; the discrepancy arose probably due to different experimental setups and forcing functions.

The transfer function controller is closely related to the state feedback controller. The parameters of the former can be transformed into the time constants and state feedback gains of the two states (or outputs) of the vehicle: the lateral displacement and its derivative proportional to the vehicle heading angle. Such an interpretation ensures an easier comparison with the recent literature; however, due to the mathematical equivalence in our settings, these two models predict the same approximation of the steering wheel angle.

The identified parameters of the driver-related transfer functions are summarised in Figures 9 and 10, explaining the dynamic properties of ten randomly selected drivers. The

parameters of the whole dataset comprising 92 drivers are displayed in Figure 14. Referring to the relevant figures, according to the global identification procedure, the reaction delay lies in the interval $\tau \in \langle 0.20\,\text{s},\ 0.61\,\text{s}\rangle$ with a 95 % confidence. In our study, each driver model involves an oscillatory component having the damping ratio $\xi \in \langle 0.18, 0.63\rangle$, time constant $T \in \langle 0.26\ \text{s}, 1.00\ \text{s}\rangle$, and gain $K \in \langle 2.8 \cdot 10^{-3}, 8.1 \cdot 10^{-3}\rangle$. The credibility of the presented confidence intervals is supported by the fact that we analysed 92 drivers, a dataset significantly broader than those in several recent studies on control-theory steering models. The authors, as mentioned in the introduction, typically reduce their analyses to only five drivers, a number that precludes a reasonable statistical analysis.

Moreover, strong interactions are at play between some of the above-mentioned parameters, whose covariance is captured by the two-component GMM. The parameters' variance may be caused by the different attributes of the individual participants, such as the control attitude, experience, and current mental state. Thus, our further research could examine, via machine learning methods, the relationship between the four parameters and the drivers' mileage, age, mental state, and other factors recorded in the pre-measurement questionnaire. The machine learning methods will potentially gain in efficiency with the reduced set of strong features proposed herein than by using multiple weak features extracted from high-order models with a lower ME.

REFERENCES

[1] M. Jirgl, J. Boril, and R. Jalovecky, "Statistical evaluation of pilot's behavior models parameters connected to military flight training," *Energies*, vol. 13, no. 17, p. 4452, Aug. 2020.

[2] J. Bořil, M. Jirgl, and R. Jalovecký, "Use of flight simulators in analyzing pilot behavior," in *Proc. 12th IFIP Int. Conf. Artif. Intell. Appl. and Innov.*, 2016, pp. 255–263.

[3] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neurosci. Biobehavioral Rev.*, vol. 44, pp. 58–75, Jul. 2014, doi: 10.1016/J.NEUBIOREV.2012.10.003.

[4] M. Chai, S.-W. Li, W.-C. Sun, M.-Z. Guo, and M.-Y. Huang, "Drowsiness monitoring based on steering wheel status," *Transp. Res. Part D, Transp. Environ.*, vol. 66, pp. 95–103, Jan. 2019, doi: 10.1016/J.TRD.2018.07.007.

[5] C. C. Macadam, "Understanding and modeling the human driver," *Vehicle Syst. Dyn.*, vol. 40, nos. 1–3, pp. 101–134, Jan. 2003, doi: 10.1076/VESD.40.1.101.15875.

[6] S. Xu, W. Tan, A. V. Efremov, L. Sun, and X. Qu, "Review of control models for human pilot behavior," *Annu. Rev. Control*, vol. 44, pp. 274–291, Oct. 2017, doi: 10.1016/J.ARCONTROL.2017.09.009.

[7] M. Mulder, D. M. Pool, D. A. Abbink, E. R. Boer, P. M. T. Zaal, F. M. Drop, K. van der El, and M. M. van Paassen, "Manual control cybernetics: State-of-the-art and current trends," *IEEE Trans. Hum.-Mach. Syst.*, vol. 48, no. 5, pp. 468–485, Oct. 2018, doi: 10.1109/THMS.2017.2761342.

[8] S. I. Noubissie Tientcheu, S. Du, and K. Djouani, "Review on haptic assistive driving systems based on drivers' steering-wheel operating behaviour," *Electronics*, vol. 11, no. 13, p. 2102, Jul. 2022, doi: 10.3390/ELECTRONICS11132102.

[9] J. Cai, H. Jiang, and J. Wang, "Implementation of the human-like lane changing driver model based on Bi-LSTM," *Discrete Dyn. Nature Soc.*, vol. 2022, no. 1, pp. 1–17, May 2022, doi: 10.1155/2022/9934292.

[10] H. Lee, H. Kim, and S. Choi, "Driving skill modeling using neural networks for performance-based haptic assistance," *IEEE Trans. Hum.-Mach. Syst.*, vol. 51, no. 3, pp. 198–210, Jun. 2021, doi: 10.1109/THMS.2021.3061409.

[11] C. J. Nash, D. J. Cole, and R. S. Bigler, "A review of human sensory dynamics for application to models of driver steering and speed control," *Biol. Cybern.*, vol. 110, nos. 2–3, pp. 91–116, Apr. 2016, doi: 10.1007/S00422-016-0682-X.

[12] C. J. Nash and D. J. Cole, "Modelling the influence of sensory dynamics on linear and nonlinear driver steering control," *Vehicle Syst. Dyn.*, vol. 56, no. 5, pp. 689–718, May 2018, doi: 10.1080/00423114.2017.1326615.

[13] C. J. Nash and D. J. Cole, "Identification and validation of a driver steering control model incorporating human sensory dynamics," *Vehicle Syst. Dyn.*, vol. 58, no. 4, pp. 495–517, Apr. 2020, doi: 10.1080/00423114.2019.1589536.

[14] C. J. Nash and D. J. Cole, "A simulation study of human sensory dynamics and driver–vehicle response," *J. Dyn. Syst., Meas., Control*, vol. 144, no. 6, Mar. 2022, Art. no. 061002.

[15] P. G. Gipps, "A behavioural car-following model for computer simulation," *Transp. Res. Part B, Methodol.*, vol. 15, no. 2, pp. 105–111, Apr. 1981.

[16] J. Han, X. Wang, and G. Wang, "Modeling the car-following behavior with consideration of driver, vehicle, and environment factors: A historical review," *Sustainability*, vol. 14, no. 13, p. 8179, Jul. 2022, doi: 10.3390/SU14138179.

[17] X. Wang, R. Jiang, L. Li, Y. Lin, X. Zheng, and F.-Y. Wang, "Capturing car-following behaviors by deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 910–920, Mar. 2018, doi: 10.1109/TITS.2017.2706963.

[18] A. Bennajeh, S. Bechikh, L. B. Said, S. Aknine, X. Zheng, and F.-Y. Wang, "A fuzzy logic-based anticipation car-following model: A historical review," *Trans. comput. collective intell.*, vol. 14, no. 13, pp. 200–222, 2018.

[19] D. Yang, L. Zhu, Y. Liu, D. Wu, and B. Ran, "A novel car-following control model combining machine learning and kinematics models for automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 1991–2000, Jun. 2019, doi: 10.1109/TITS.2018.2854827.

[20] Z. Zhang, H. Zhang, and W. Zhao, "A human-vehicle game stability control strategy considering drivers' steering characteristics," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 7306–7316, Nov. 2021, doi: 10.1109/TITS.2020.3008153.

[21] C. Dai, C. Zong, D. Zhang, M. Hua, H. Zheng, and K. Chuyo, "A bargaining game-based human–machine shared driving control authority allocation strategy," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 10, pp. 10572–10586, Oct. 2023, doi: 10.1109/TITS.2023.3277015.

[22] S. D. Keen and D. J. Cole, "Application of time-variant predictive control to modelling driver steering skill," *Vehicle Syst. Dyn.*, vol. 49, no. 4, pp. 527–559, Apr. 2011, doi: 10.1080/00423110903551626.

[23] M. Thommyppillai, S. Evangelou, and R. S. Sharp, "Car driving at the limit by adaptive linear optimal preview control," *Vehicle Syst. Dyn.*, vol. 47, no. 12, pp. 1535–1550, Dec. 2009, doi: 10.1080/00423110802673109.

[24] C. J. Nash and D. J. Cole, "Identification of a driver model incorporating sensory dynamics, with nonlinear vehicle dynamics and transient disturbances," *Vehicle Syst. Dyn.*, vol. 60, no. 8, pp. 2805–2824, Aug. 2022, doi: 10.1080/00423114.2021.1931696.

[25] D. J. Cole, A. J. Pick, and A. M. C. Odhams, "Predictive and linear quadratic methods for potential application to modelling driver steering control," *Vehicle Syst. Dyn.*, vol. 44, no. 3, pp. 259–284, Mar. 2006, doi: 10.1080/00423110500260159.

[26] G. Braghieri, A. Haslam, M. Sideris, J. Timings, and D. Cole, "Quantification of road vehicle handling quality using a compensatory steering controller," *J. Dyn. Syst., Meas., Control*, vol. 139, no. 3, pp. 1–18, Mar. 2017, doi: 10.1115/1.4035009.

[27] J. Cao, H. Lu, K. Guo, and J. Zhang, "A driver modeling based on the preview-follower theory and the jerky dynamics," *Math. Problems Eng.*, vol. 2013, pp. 1–10, Jun. 2013.

[28] E. Donges, "A two-level model of driver steering behavior," *Human Factors, J. Human Factors Ergonom. Soc.*, vol. 20, no. 6, pp. 691–707, Dec. 1978, doi: 10.1177/001872087802000607.

[29] D. T. McRuer and E. S. Krendel, "Elementary concepts for the quantitative description of human operator dynamics in control systems," in *Mathematical Models of Human Pilot Behavior*, 1st ed., London, U.K.: AGARD, Jan. 1974, ch. 2, pp. 5–15.

[30] S. S. Tohidi and Y. Yildiz, "A control theoretical adaptive human pilot model: Theory and experimental validation," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 6, pp. 2585–2597, Nov. 2022, doi: 10.1109/TCST.2022.3164237.

[31] R. A. Hess, "Structural model of the adaptive human pilot," *J. Guid. Control*, vol. 3, no. 5, pp. 416–423, Sep. 1980, doi: 10.2514/3.56015.

[32] Z. Wang, X. Zhou, H. Shen, and J. Wang, "Algebraic driver steering model parameter identification," *J. Dyn. Syst., Meas., Control*, vol. 144, no. 5, May 2022, Art. no. 051006, doi: 10.1115/1.4053431.

[33] K. van der El, D. M. Pool, M. R. M. van Paassen, and M. Mulder, "A unifying theory of driver perception and steering control on straight and winding roads," *IEEE Trans. Hum.-Mach. Syst.*, vol. 50, no. 2, pp. 165–175, Apr. 2020, doi: 10.1109/THMS.2019.2947551.

[34] S. Schnelle, J. Wang, H. Su, and R. Jagacinski, "A driver steering model with personalized desired path generation," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 47, no. 1, pp. 111–120, Jan. 2017, doi: 10.1109/TSMC.2016.2529582.

[35] A. Mihály and P. Gáspár, "Identification of a linear driver model based on simulator experiments," in *Proc. IEEE 9th IEEE Int. Symp. Appl. Comput. Intell. Informat. (SACI)*, May 2014, pp. 13–18, doi: 10.1109/SACI.2014.6840057.

[36] M. Havlikova, S. Sediva, Z. Bradac, and M. Jirgl, "A man as the regulator in man-machine systems," *Adv. Electr. Electron. Eng.*, vol. 12, no. 5, Dec. 2014, Art. no. 469475.

[37] D. M. Pool, P. M. T. Zaal, H. J. Damveld, M. M. van Paassen, and M. Mulder, "Pilot equalization in manual control of aircraft dynamics," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, Oct. 2009, pp. 2480–2485.

[38] T. Hastie, R. Tibshirani, and J. H. Friedman, "Model assessment and selection," in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., New York, NY, USA: Springer, 2009, ch. 7, p. 219259.

[39] A. M. C. Odhams and D. J. Cole, "Identification of the steering control behaviour of five test subjects following a randomly curving path in a driving simulator," *Int. J. Vehicle Auto. Syst.*, vol. 12, no. 1, p. 44, 2014, doi: 10.1504/IJVAS.2014.057863.

[40] M. Akamatsu, P. Green, and K. Bengler, "Automotive technology and human factors research: Past, present, and future," *Int. J. Veh. Technol.*, vol. 2013, pp. 1–27, Sep. 2013, doi: 10.1155/2013/526180.

[41] C. Guo, C. Sentouh, J.-C. Popieul, J.-B. Haué, S. Langlois, J.-J. Loeillet, B. Soualmi, and T. N. That, "Cooperation between driver and automated driving system: Implementation and evaluation," *Transp. Res. Part F, Traffic Psychol. Behaviour*, vol. 61, pp. 314–325, Feb. 2019, doi: 10.1016/J.TRF.2017.04.006.

[42] D. J. Allerton, "The impact of flight simulation in aerospace," *Aeronaut. J.*, vol. 114, no. 1162, pp. 747–756, Dec. 2010, doi: 10.1017/S0001924000004231.

[43] H.-P. Schöner and B. Morys, "Dynamische fahrsimulatoren," in *Handbuch Fahrerassistenzsysteme*. Wiesbaden, Germany: Springer, 2015, pp. 139–154.

[44] W. König, "Nutzergerechte entwicklung der mensch-maschine-interaktion von fahrerassistenzsystemen," in *Handbuch Fahrerassistenzsysteme*. Wiesbaden, Germany: Springer, 2009, pp. 33–42.

[45] E. Altendorf, M. Baltzer, M. Kienle, S. Meier, T. Weigerber, M. Heesenn, and F. Flemisch, "H-mode 2D," in *Handbuch Fahrerassistenzsysteme*. Wiesbaden, Germany: Springer, 2015, pp. 1124–1136.

[46] J. P. Espineira, J. Robinson, J. Groenewald, P. H. Chan, and V. Donzella, "Realistic LiDAR with noise model for real-time testing of automated vehicles in a virtual environment," *IEEE Sensors J.*, vol. 21, no. 8, pp. 9919–9926, Apr. 2021, doi: 10.1109/JSEN.2021.3059310.

[47] H. Jeon, Y. Kim, M. Choi, D. Park, S. Son, J. Lee, G. Choi, and Y. Lim, "CARLA simulator-based evaluation framework development of lane detection accuracy performance under sensor blockage caused by heavy rain for autonomous vehicle," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 9977–9984, Oct. 2022, doi: 10.1109/LRA.2022.3192632.

[48] C. Brogle, C. Zhang, K. L. Lim, and T. Bräunl, "Hardware-in-the-loop autonomous driving simulation without real-time constraints," *IEEE Trans. Intell. Vehicles*, vol. 4, no. 3, pp. 375–384, Sep. 2019, doi: 10.1109/TIV.2019.2919457.

[49] C. Merenda, C. Suga, J. Gabbard, and T. Misu, "Effects of vehicle simulation visual fidelity on assessing driver performance and behavior," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 1679–1686, doi: 10.1109/IVS.2019.8813863.

[50] S. Ansari, F. Naghdy, H. Du, and Y. N. Pahnwar, "Driver mental fatigue detection based on head posture using new modified reLU-BiLSTM deep neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10957–10969, Aug. 2022.

[51] G. Silvera, A. Biswas, and H. Admoni, "DReyeVR: Democratizing virtual reality driving simulation for behavioural interaction research," 2022, *arXiv:2201.01931*.

[52] D. Michalík, M. Jirgl, J. Arm, and P. Fiedler, "Developing an unreal engine 4-Based vehicle driving simulator applicable in driver behavior analysis—A technical perspective," *Safety*, vol. 7, no. 2, p. 25, Apr. 2021, doi: 10.3390/SAFETY7020025.

[53] H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, no. 1, pp. 37–52, 1967, doi: 10.1007/BF02395039.

[54] O. Mihálik, "Reconstruction of non-uniformly sampled signals using Gerchberg-Papoulis method," in *Proc. 26th Conf. student EEICT*, 2020, pp. 136–140.

[55] T. Hastie, R. Tibshirani, and J. H. Friedman, "Basis expansion and regularisation," in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., New York, NY, USA: Springer, 2009, ch. 5, pp. 139–218.

[56] MathWorks Inc. *Curve Fitting Toolbox: User's Guide*. Accessed: Jan. 19, 2024. [Online]. Available: https://www.mathworks.com/help/pdf_doc/curvefit/curvefit.pdf

[57] M. Havlíkova, S. Šedivá, K. Stibor, and Z. Bradáč, "A driver as the regulator in man-machine system," in *Proc. 12th IFAC Conf. Program. Devices Embedded Syst.*, 2013, pp. 342–347.

[58] R. A. Hess and A. Modjtahedzadeh, "A control theoretic model of driver steering behavior," *IEEE Control Syst. Mag.*, vol. 10, no. 5, pp. 3–8, Aug. 1990, doi: 10.1109/37.60415.

[59] L. Ljung, "Model structure selection and model validation," in *System Identification: Theory for the User*, 2nd ed., Upper Saddle River, NJ, USA: Prentice & Hall PTR, 1999, pp. 491–519.

[60] L. Ljung. *System Identification Toolbox: Reference*. Accessed: Jan. 19, 2024. [Online]. Available: https://www.mathworks.com/help/pdf_doc/ident/ident_ref.pdf

[61] J. Nocedal and S. J. Wright, "Line search methods," in *Numerical Optimization*, 2nd ed., New York, NY, USA: Springer, 2006, ch. 3, pp. 30–65.

[62] M. Bierlaire, "Descent methods and line search," in *Optimization: Principles and Algorithms*, 2nd ed., Lausanne, Switzerland: EPFL Press, 2018, ch. 1, pp. 245–290.

[63] R. A. Hess, "Analyzing manipulator and feel system effects in aircraft flight control," *IEEE Trans. Syst. Man, Cybern.*, vol. 20, no. 4, pp. 923–931, Aug. 1990, doi: 10.1109/21.105091.

[64] K. Van Der El, D. M. Pool, M. van Paassen, and M. Mulder, "Identification and modeling of driver multiloop feedback and preview steering control," in *Proc. IEEE Int. Conf. Syst. Man, Cybern. (SMC)*, Oct. 2018, pp. 1227–1232.

[65] D. T. McRuer, R. W. Allen, D. H. Weir, and R. H. Klein, "New results in driver steering control models," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 19, no. 4, pp. 381–397, Aug. 1977.

[66] T. Hastie, R. Tibshirani, and J. Friedman, "Model inference and averaging," in *Series in Statistics*. Cham, Switzerland: Springer, 2009, pp. 1–34.

[67] A. J. Pick and D. J. Cole, "Dynamic properties of a driver's arms holding a steering wheel," *Proc. Inst. Mech. Engineers, Part D, J. Automobile Eng.*, vol. 221, no. 12, pp. 1475–1486, Dec. 2007.

[68] G. de Villiers and E. R. Pike, "Beyond the 2WT theorem," in *The Limits of Resolution*, 1st ed., Boca Raton, FL, USA: CRC Press, 2016, ch. 2, pp. 331–362.

[69] A. Devasia and M. Cada, "High precision numerical implementation of bandlimited signal extrapolation using prolate spheroidal wave functions," in *Proc. Trans. Eng. Technol., Special Issue World Congr. Eng. Comput. Sci.*, 2013, pp. 387–402.

MIROSLAV JIRGL was born in Brno, Czech Republic, in 1988. He received the Ph.D. degree in cybernetics, control, and measurement from Brno University of Technology, Brno, in 2017.

From 2018 to 2021, he was with the Central European Institute of Technology, Brno. Currently, he is an Assistant Professor with the Department of Control and Instrumentation, Faculty of Electrical Engineering and Communication, Brno University of Technology. His research interests include man-machine systems, dynamical systems and their control, and systems modeling.

ONDREJ MIHÁLIK was born in Slovakia, in 1996. He received the B.S. and M.S. degrees in automation and measurement from Brno University of Technology, Brno, Czech Republic, in 2017 and 2019, respectively, where he is currently pursuing the Ph.D. degree in cybernetics, control, and measurement, with a focus on anomaly detection.

From 2019 to 2021, he was with the Central European Institute of Technology, Brno. Since 2021, he has been with the Eaton European Innovation Center, Prague, Czech Republic. His main research interests include modern signal processing and machine learning methods.

SABRINA BOUJENFA received the B.Sc. and M.Sc. degrees in mechanical engineering from the University of Siegen, Germany, in 2013 and 2016, respectively. She is currently pursuing the Ph.D. degree with the Faculty of Electrical Engineering and Communication, Brno University of Technology, focusing on sensor-based analysis of the state of consciousness.

Alongside the Ph.D. degree, she is currently the Technical Project Manager of Thyssenkrupp Presta AG, Liechtenstein.

ZDENĚK BRADÁČ was born in 1973. He received the Ph.D. degree in technical cybernetics from Brno University of Technology, Brno, Czech Republic, in 2004.

He is currently an Associate Professor with the Department of Control and Instrumentation, Faculty of Electrical Engineering and Communication, Brno University of Technology. His research interests include HMI systems, fault-tolerant systems, information systems safety, and security.

PETR FIEDLER was born in 1974. He received the Ph.D. degree in technical cybernetics from Brno University of Technology, Brno, Czech Republic, in 2005.

He is currently an Associate Professor with the Department of Control and Instrumentation, Faculty of Electrical Engineering and Communication, Brno University of Technology. His research interests include resilient systems and embedded systems for real-time control and sensing.