



## Development of the Västra Götaland operating cycle for long-haul heavy-duty vehicles

Downloaded from: <https://research.chalmers.se>, 2024-04-30 08:44 UTC

Citation for the original published paper (version of record):

Romano, L., Bruzelius, F., Jacobson, B. (2023). Development of the Västra Götaland operating cycle for long-haul heavy-duty vehicles. IEEE Access. <http://dx.doi.org/10.1109/ACCESS.2023.3295989>

N.B. When citing this work, cite the original published paper.

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

Received 27 June 2023, accepted 12 July 2023, date of publication 17 July 2023, date of current version 21 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3295989

## APPLIED RESEARCH

# Development of the Västra Götaland Operating Cycle for Long-Haul Heavy-Duty Vehicles

LUIGI ROMANO<sup>1</sup>, MICHELE GODIO<sup>2</sup>, PÄR JOHANNESSON<sup>2</sup>, FREDRIK BRUZELIUS<sup>1</sup>,  
TOHEED GHANDRIZ<sup>3</sup>, AND BENGT JACOBSON<sup>1</sup>

<sup>1</sup>Department of Mechanics and Maritime Sciences, Chalmers University of Technology, 412 96 Gothenburg, Sweden

<sup>2</sup>RISE Research Institutes of Sweden, 412 79 Gothenburg, Sweden

<sup>3</sup>Volvo AB, 40508 Gothenburg, Sweden

Corresponding author: Luigi Romano (luigi.romano@chalmers.se)

This work was supported in part by the Swedish Energy Agency, and in part by the Swedish Vehicle Research and Innovation Programme (FFI) under Grant 44929-1.

**ABSTRACT** In this paper, a complete operating cycle (OC) description is developed for heavy-duty vehicles traveling long distances in the region of Västra Götaland, Sweden. Variation amongst road transport missions is accounted for using a collection of stochastic models. These are parametrized from log data for all the influential road parameters that may affect the energy performance of heavy trucks, including topography, curvature, speed limits, and stop signs. The statistical properties of the developed OC description are investigated numerically by considering some composite variables, condensing the salient information about the road characteristics, and inspired by two existing classification systems. Two examples are adduced to illustrate the potential of the OC format, which enables ease of classification and detailed simulation of energy efficiency for individual vehicles, with application in vehicle design optimization and selection, production planning, and predictive maintenance. In particular, for the track used in the first example, a Volvo FH13 equipped with a diesel engine, simulation results indicate mean CO<sub>2</sub> emissions of around 1700 g km<sup>-1</sup>, with a standard deviation of 360 g km<sup>-1</sup>; in the second example, dealing with electrical fleet sizing, the optimal proportion shows a predominance of tractor-semitrailer vehicles (70%) equipping 4 motors and 11 battery packs.

**INDEX TERMS** Operating cycle, mission classification, road transport mission, stochastic modeling, autoregressive models, Markov models.

## NOMENCLATURE

Symbol	Description		
$C$	Random variable for curvature (m <sup>-1</sup> ).	$L_s$	Sampling length for topography (m).
$G_R$	Generator matrix for road type.	$L_{tot}$	Total road length (km).
$G_{V r_k}$	Generator matrix for speed sign.	$L_R$	Mean length vector for road type (km).
$K$	Random variable for continuous road curvature (m <sup>-1</sup> ).	$\tilde{L}_R$	Stochastic mean length vector for road type (km).
$L_h$	Stochastic mean hill length (m).	$L_{Ri}$	Mean length for road type (km).
$L_{h r_i}$	Conditional mean hill length (m).	$\tilde{L}_{Ri}$	Stochastic mean length for road type (km).
$\tilde{L}_{h r_i}$	Conditional stochastic mean hill length (m).	$L_{V r_k}$	Conditional mean length vector for speed sign (km).
$L_m$	Mission length (km).	$\tilde{L}_{V r_k}$	Conditional stochastic mean length vector for speed sign (km).
$\tilde{L}_m$	Stochastic mission length (km).	$L_{Vi r_k}$	Conditional mean length for speed sign (km).
		$\tilde{L}_{Vi r_k}$	Conditional stochastic mean length for speed sign (km).
		$M_{C r_i}$	Conditional stochastic mean log-radius (ln m).
		$M_{L r_i}$	Conditional stochastic mean log-length (ln m).

The associate editor coordinating the review of this manuscript and approving it for publication was S. Ali Arefifar<sup>1</sup>.

$\bar{N}'_C$	Stochastic expected number of curves ( $\text{km}^{-1}$ ).	$p_{Rij}$	Markov matrix entries for road type.
$\bar{N}_s$	Stochastic expected number of stops ( $\text{km}^{-1}$ ).	$\tilde{p}_{Rij}$	Stochastic Markov matrix entries for road type.
$P_K$	Stochastic curviness length ratio.	$p_{Vij r_k}$	Conditional Markov matrix entries for speed signs.
$\mathbf{P}_R$	Markov matrix for road type.	$\tilde{p}_{Vij r_k}$	Conditional stochastic Markov matrix entries for speed signs.
$\tilde{\mathbf{P}}_R$	Stochastic Markov matrix for road type.	$p_{y r_i}$	Conditional road grade length ratio.
$\tilde{\mathbf{P}}_{Ri}$	Stochastic Markov row for road type.	$p_{K r_i}$	Conditional curviness length ratio.
$\tilde{\mathbf{P}}_{V r_k}$	Conditional stochastic Markov matrix for speed signs.	$r_i$	Road type.
$\tilde{\mathbf{P}}_{Vi r_k}$	Conditional stochastic Markov row for speed signs.	$r_{\text{turn}}$	Minimum radius of curvature (m).
$\mathbf{P}_{V r_i}$	Conditional Markov matrix for speed signs.	$t_s$	Stop time (s).
$P_y$	Stochastic road grade length ratio.	$t_{\min}, t_{\max}$	Minimum and maximum recommended time (s).
$R_t$	Random variable for road type.	$\hat{v}_{ r_k}$	Conditional mean legal speed ( $\text{km h}^{-1}$ ).
$T_{\min r_i}, T_{\max r_i}$	Conditional stochastic minimum and maximum recommended time (s).	$v_{i r_k}$	Conditional legal speed ( $\text{km h}^{-1}$ ).
$\Delta T_{ r_i}$	Conditional stochastic delta recommended time (s).	$y$	Road grade (%).
$T_s$	Random variable for recommended time (s).	$\kappa$	Continuous road curvature ( $\text{m}^{-1}$ ).
$V$	Random variable for speed sign ( $\text{km h}^{-1}$ ).	$\lambda_{C r_i}$	Conditional mean log-radius ( $\text{km}^{-1}$ ).
$\mathbf{X}_{\text{GTA}}$	Multivariate random vector for GTA (km), ( $\text{km}^{-1}$ ).	$\lambda_{s r_i}$	Conditional stop signs intensity ( $\text{km}^{-1}$ ).
$\mathbf{X}_{\text{UFD}}$	Multivariate random vector for UFD (km), ( $\text{km}^{-1}$ ).	$\mu_{C r_i}$	Conditional mean log-radius (ln m).
$\mathbf{X}_{\mathcal{R}_s}$	Multivariate random vector of sOC parameters (ln m), (ln km).	$\mu_{L r_i}$	Conditional mean log-length (ln m).
$Y$	Random variable for road grade (%).	$\mu_{\mathbf{X}_{\mathcal{R}_s}}$	Vector-valued mean of random sOC parameters (ln m), (ln km).
$\mathbf{\Pi}_R$	Stochastic stationary probability vector for road type.	$\pi_R$	Stationary probability vector for road type.
$\Pi_{Ri}$	Stochastic stationary probability for road type.	$\pi_{Ri}$	Stationary probability for road type.
$\mathbf{\Pi}_{V r_k}$	Conditional stochastic stationary probability vector for speed signs.	$\pi_{V r_k}$	Conditional stationary probability vector for speed signs.
$\Pi_{Vi r_k}$	Conditional stochastic stationary probability for speed signs.	$\pi_{Vi r_k}$	Conditional stationary probability for speed signs.
$\Lambda_{C r_i}$	Conditional stochastic curve intensity ( $\text{ln km}^{-1}$ ).	$\phi_{Y r_i}$	Conditional autoregressive coefficient for topography.
$\Lambda_{s r_i}$	Conditional stochastic stop signs intensity ( $\text{ln km}^{-1}$ ).	$\sigma_{eY r_i}$	Conditional standard deviation for the innovation in the topography (%).
$\Sigma_{C r_i}$	Conditional stochastic std log-radius (ln m).	$\sigma_{C r_i}$	Conditional standard deviation of log-radius (ln m).
$\Sigma_{L r_i}$	Conditional stochastic std log-length (ln m).	$\sigma_{L r_i}$	Conditional standard deviation of log-length (ln m).
$\Sigma_{Y r_i}$	Conditional stochastic standard deviation of topography (%).	$\sigma_{Y r_i}$	Conditional standard deviation of topography (%).
$e_Y$	Error term for topography (%).	$\text{Dir}(\cdot)$	Dirichlet distribution.
$g_{Rij}$	Entry of the generator matrix for road type.	$\mathcal{E}(\cdot)$	Exponential distribution.
$g_{Vij r_k}$	Entry of the conditional generator matrix for speed signs.	$\text{Ga}(\cdot, \cdot)$	Gamma distribution.
$\bar{n}'_{C r_i}$	Conditional expected number of curves ( $\text{km}^{-1}$ ).	$\mathcal{N}(\cdot, \cdot)$	Normal distribution.
		$\mathbb{P}(\cdot)$	Probability.
		$\mathbb{E}(\cdot)$	Expectation.
		$\mathbb{V}\text{ar}(\cdot)$	Variance.
		$\mathbb{C}\text{ov}(\cdot)$	Covariance.

## I. INTRODUCTION

The optimal design of both commercial and heavy-duty vehicles is traditionally supported by the extensive use of simulations in dedicated virtual environments, which allow for a preliminary evaluation of energy performance. However, at early stages, prototypes are typically tested considering a

reduced number of transport missions, which should ideally be representative of the overall intended usage. Instead, reproducing the variation amongst road operations in an accurate way is essential to correctly estimate the energy performance in real-world scenarios, depending on the characteristics of the transportation task [1], [2], [3], [4]. Indeed, the inherent variation amongst road transport missions may excite different responses in ground vehicles, implying the need for a proper understanding of where and how they are operated, especially in view of the impellent shift of paradigm towards full automation and electrification [5], [6]. This also implies the need of evaluating in isolation the contribution of individual parameters like road properties, weather, and traffic conditions.

In this context, the conventional description in terms of a driving cycle is not suitable, since all the relevant information about the operating environment is lost in the process of synthesizing a representative speed profile [7], [8]. On the contrary, the operating cycle (OC) format, recently proposed in [9] and [10], allows to explicitly account for the influence of the surroundings on vehicular performance, allowing for ease of classification of road operations and detailed assessment of energy efficiency. Previous works on the OC representation have however been limited to considering individual transport operations and the variation within. Instead, this paper deals with the development of a complete OC description concerning long-haul heavy-duty vehicles, considering variation amongst transport missions. More specifically, the spread in road conditions – which is reflected in those amongst transport operations and, in turn, energy performance – is captured by modeling the variation in the influential parameters.

To better highlight the salient differences between the two types of representation mentioned above, the next Section I-A recapitulates the work previously done on driving and operating cycles, along with introducing the three fundamental problems identified by Pettersson: the representation, variation, and classification problems [9].

## A. PREVIOUS WORKS ON DRIVING AND OPERATING CYCLES

This Section summarizes the previous contributions to the modeling of driving and operating cycles.

### 1) LITERATURE REVIEW ON THE DRIVING CYCLE REPRESENTATION

In the past years, a great deal of research has produced a number of alternative approaches aimed at synthesizing representative transient driving cycles. In fact, whilst the initially proposed rule-based methods have been rapidly abandoned in favour of statistical techniques [11], [12], these have proliferated, leading to a plethora of variants. Statistical approaches are typically preferred when synthetically generated speed profiles are required to correlate with specific operating conditions. The considered signals may include information concerning only the vehicle's state, e.g., cruising, idling,

acceleration or braking [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], or also account for road grade [24], [25].

The scientific literature distinguishes between four different techniques to construct synthetic driving cycles: micro-trip-based, segment-based, pattern classification and modal cycle construction [26]. The first method initially generates several candidate cycles by resorting to the so-called micro-trip decomposition technique. Multiple approaches to identifying consecutive micro-trips have been proposed, either involving random processes or based on the definition of specific criteria that relate to the modal characteristics of consecutive excursions. Stochastic optimization algorithms are then employed to optimally select the final driving cycle from the initial population [27], [28]. Transient driving cycles synthesized using the micro-trip decomposition include for example those for the cities of Hong Kong [29], Pune and Chennai, India, [23], [30] and Singapore [31]. Segment-based methods synthesize driving cycles mainly starting from the available speed and acceleration signals. A major advantage of such a technique resides in that, compared to the micro-trip-based one, it allows to incorporate details about the surroundings, including road properties and traffic conditions [4]. However, since the trips are segmented based on different criteria than consecutive stops, additional constraints need to be imposed when transitioning from each segment [32], [33]. Pattern classification methods partition data into kinematic sequences by resorting to statistical techniques [34], [35], [36]. After specifying opportune criteria, the kinematic sequences are subsequently classified into heterogeneous classes. Finally, by combining different kinematic sequences based on the statistical properties of each class, an optimal driving cycle is generated. In this context, two frequent approaches employed to synthesize representative cycles from kinematic sequences are the cluster and principal component analyses [37]. Maximum likelihood estimation techniques are used to cluster the measured speed data into snippets in the modal cycle method [18], [24], [25]. Relying on a Markov property argument, supported also theoretically by the analyses conducted in [38], driving cycles are constructed starting from opportunely selected snippets. Advanced formulations, including multi-dimensional Markov chains that integrate speed signals with information concerning the acceleration and road grade, or variable passenger loads, have been proposed for example in [24], [25], [26]. A summary of these different techniques, with their advantages and disadvantages, is given in Table 1.

The above mentioned approaches are all aimed at synthesizing a unique representative transient cycle starting from a large amount of data. In fact, a single driving cycle is usually sufficient when it comes to apply conventional algorithms and routines for the purposes of vehicle design optimization and selection. In this context, the need for an accurate, individual description is referred in the literature as the *representation problem* [9]. On the other hand, reproducing variation in transport operations may lead to a more accurate prediction of the energy performance of road vehicles. Ideally, departing

from a single representation of the usage, driving cycles may be constructed to be statistically equivalent, so as to produce a meaningful spread in performance. This line of research has been pioneered by the authors of [32] and [33], who have proposed a procedure that allows for the synthesis of multiple driving cycles, departing from a single trip. More specifically, it was shown in [32] and [33] that the energy performance of electric city buses is highly sensitive to fluctuations in both the number of stops and passenger load. Indeed, by treating both quantities as random variables, they deduced that the energy consumption follows approximately a normal distribution. The variability connected to the randomness of the external surroundings, concerning for example temperature and rolling resistance, was also investigated in [39], where the inherent uncertainty was modeled by adding extensive noise to the assumed nominal value for the quantities in interest. The fundamental contribution of [32], [33], and [39] essentially explored the *variation problem*, as defined in [9]. Concerning specifically the driving cycle description, the technique introduced in [32] and [33] is renamed *stochastic generation* in Table 1.

Along with the representation and variation problems, however, Pettersson also defines the so-called *classification problem* [9], in conjunction with the need to properly qualify a road mission (or an entire transport application) using a simplified set of metrics and labels. This is mainly motivated by the fact that, rather than referring to a certain speed profile, vehicle manufacturers and operators usually describe the usage concerning the characteristics of the environment, which are easier to interpret. In this context, whilst the conventional description in terms of a driving cycle has been successful in addressing the first two problems indicated in [9], the latter still deserves particular attention. In fact, a major limitation connected with the driving cycle representation is that the information about the operating environment – including parameters that may dramatically impact energy consumption [1], [2], [3], [4], [40], [41], [42], [43], [44], [45], [46] – is often lost or accounted for only implicitly during the synthesis of a representative speed profile.

## 2) LITERATURE REVIEW ON THE OPERATING CYCLE REPRESENTATION

The OC representation is a relatively novel description developed in [9], [47], and [48] and further extended in [49] and [50]. As opposed to a driving cycle, in the OC framework, the mission properties are modeled separately from those of the surroundings, without postulating a reference speed profile. The external stimuli originating from the operating environment are then converted dynamically into the desired speed by using a driver model. This expedient allows describing road transport missions in a way that is, to a large extent, independent of both driver and vehicle.

The OC format comprises three levels of representation and describes a physical quantity via deterministic or

statistical models. More specifically, the bird's-eye view is conceived as a high-level representation, where road missions are labeled by resorting to simplified indicators. It is intrinsically suited to cope with the classification problem [9]. The stochastic operating cycle (sOC) consists of a more formal description and collects all the random parameters that are needed to reproduce the statistical properties of a transport operation. It is the mathematical tool suggested in [9] to address the variation problem. Finally, the deterministic operating cycle (dOC) may be interpreted as a single realization of an sOC, and includes a detailed representation of the operating environment. Similar to a driving cycle, it has been mainly conceived to deal with the representation problem [9].

Stochastic and deterministic models are currently available for road parameters [47], [48], weather, and traffic conditions [49]. Moreover, the authors of [50] have recently shown how to build a piece of consistent and cohesive machinery that connects all three levels of representation, by establishing a set of formal relationships between the bird's-eye view metrics and the statistical models included in the sOC. In [50], this operation was strongly inspired by two already existing classification systems, namely the Global Transport Application (GTA) [51] and the User Factor Description (UFD) adopted by Volvo and Scania, respectively.

## B. CONTRIBUTION OF THIS PAPER

So far, the application of the OC format has been limited to individual missions. As already mentioned, a single operating cycle, opportunely selected to be representative of the overall usage, may be sufficient for the purpose of design optimization, whereas the energy efficiency of road vehicles should be more comprehensively assessed by considering the intended transport application.

Therefore, from a conceptual and methodological perspective, the first contribution of this paper consists of exploring an additional dimension of the OC framework, concerning the distribution of missions – mainly with respect to the road characteristics, e.g., topography, curvature, speed limits, and stop signs<sup>1</sup> – over the entire population defining the transport application. The focus is on heavy-duty trucks traveling distances typically longer than 50 km. The investigation is systematically conducted by exploiting the statistical and hierarchical structure of the OC description. In particular, the sOC parameters (interpreted as realizations for individual road operations) are modeled in this paper as random variables when defined on the entire population of missions. By doing so, it becomes possible to capture variation in usage due to modifications in the statistical properties of the environment, ultimately resulting in the external stimuli affecting

<sup>1</sup>As explained in Section I-A2, the current OC description does contain also models for weather and traffic conditions. However, these are not included in the paper. There are multiple reasons for this, the first being that the parameters of the weather and traffic models cannot be directly estimated from log data, and thus cannot provide information about variation within the transport application. If needed for simulation purposes, traffic and weather models may eventually be incorporated starting from previous contributions by the authors [49], [50].



**TABLE 1.** Summary of the different driving cycle and operating cycle-based descriptions of a road transport mission, with their advantages and disadvantages.

Description	Technique	Main references	Advantages	Disadvantages
<i>Driving cycle</i>	Micro trip	[23], [26]–[31]	Considers different operating conditions, using the definition of micro-trips. It is relatively easy to implement.	Does not consider the effect of external conditions. Does not allow reproducing variation in usage. Unsuitable for classification purposes.
	Segmentation	[4], [32], [33]	Incorporates details about the surroundings (road grade, traffic conditions, etc.).	Requires imposing additional constraints on segmentation. Does not allow reproducing variation. Unsuitable for classification purposes.
	Pattern classification	[34]–[37]	Generates statistically optimal driving cycles based on advanced techniques.	Does not allow reproducing variation. Does not account for the effect of the surroundings. Unsuitable for classification purposes.
	Modal construction	[18], [24], [25], [38]	Incorporates details about the surroundings (road grade, variable passenger loads, etc.). It is also theoretically justified.	Does not allow reproducing variation. Unsuitable for classification purposes.
	Stochastic generation	[32], [33], [39]	Incorporates details about the surroundings (road grade, variable passenger loads, temperature and humidity, etc.). Allows reproducing variation in the operating conditions.	Unsuitable for classification purposes.
<i>Operating cycle</i>	Stochastic generation	[9], [47]–[50]	Incorporates details about the surroundings (road grade, curviness, temperature and humidity, etc.). Allows reproducing variation in the operating conditions. Allows for ease of classification based on physical quantities.	Requires parametrizing a larger number of models. It may be more computationally costly.

the vehicle's behaviour. The first part of the work is hence dedicated to the stochastic modeling of the sOC parameters. As it is traditionally in the spirit of the OC representation, which fervently endorses the principle of parsimony, these are assumed to be independent of each other and to follow simple distributions that allow for ease of interpretation.

A numerical analysis of the resulting OC description is then conducted. This operation heavily relies on the non-bijective connection between the sOC and the bird's-eye view levels of representation and builds upon some recent results established in [50]. The abstract problem of properly classifying an OC model intimately relates to the aspect of representativeness and, in fact, may specifically be reformulated in question form: “what type of usage is this OC representative of, and what can be done with such information?”. In this context, the interrogation is everything but trivial. Indeed, as opposed to the conventional driving cycle representation, for an OC, describing a transport mission directly in terms of the operating environment, the notion of representativeness should be referred to the characteristics of the surroundings. To answer this question, the statistical properties of the parametrized OC are investigated concerning some composite variables, which, apart from reducing the relatively large dimension of the original problem, condense the

salient features of the individual road missions. In this paper, the definition of the composite variables is inspired by two different classification systems in use by Volvo and Scania, respectively. Moreover, a simple methodology is proposed to synthesize a single set of representative sOC parameters, starting from their original distributions.

The second contribution of this paper is to illustrate some applications of the whole OC machinery to the processes of virtual design, selection and testing, by taking advantage of all three levels of representation comprised in the format. To this end, two different examples are adduced. The first focuses on the single-vehicle perspective and is aimed at assessing energy efficiency in a simulation environment. The second example deals with the optimal selection and production planning, based on the detailed distribution of the road missions, in the footsteps of [52], [53], [54]. The methods of data synthesis used in the paper are also implemented in MATLAB/Simulink<sup>®</sup> environment.

The remainder of the paper is organized as follows. Section II provides a general introduction to the OC description, with details about each level of representation and the relationships between them. In Section III, the fundamental theory for stochastic road and mission models is recapitulated and extended to account for variability in the transport

application. Moreover, the parametrization of the road models is carried out using log data collected from heavy-duty vehicles during real-world operations. Section IV is devoted to the numerical analysis of the developed OC description, using the composite variables inspired by the GTA and UFD systems adopted by Volvo and Scania. An example of application of the entire OC format for vehicle design, selection, and testing is then adduced in Section V. A comprehensive discussion about the methodology presented in this paper, along with its limitations and possible extensions, is presented in Section VI. Finally, the main conclusions, together with some directions for future research, are summarized in Section VII.

## II. BACKGROUND ON THE OPERATING CYCLE

The OC format consists of a mathematical framework that describes road transport missions and applications directly in terms of the operating environment. As already mentioned, it comprises three main levels of representation: the bird's-eye view, the sOC, and the dOC, arranged in a hierarchical fashion, as shown in Figure 1. These three different descriptions address the classification, variation and representation problems, respectively. In turn, each description groups the corresponding models or parameters into four common categories: road, weather, traffic and mission [9], [47], [49]. This paper limits itself to consider the road category, whereas weather and traffic have been discussed extensively in previous works [10], [49], [50]

### A. THE BIRD'S-EYE VIEW

Positioned at the top of the pyramid in Figure 1, the bird's-eye view may be regarded as a high-level description of a transport mission. It essentially consists of a collection of labels and metrics, which target either individual missions or entire applications, in order to provide an intuitive understanding of how vehicles are operated on the road. This description is discrete in nature, meaning that it assigns a countable number of labels to a given road mission, corresponding to operating classes that cannot overlap. The bird's-eye view is the main tool to address the classification problem [9]. In previous works, and also in this paper, the metrics and labels for the bird's-eye view have been borrowed from existing classification systems (see, in particular, [50]). An example with the topography parameter is adduced below considering the GTA representation adopted by Volvo, which prescribes four different levels [51]:

- 1) FLAT if slopes with a grade of less than 3% occur during more than 98% of the driving distance.
- 2) P-FLAT if slopes with a grade of less than 6% occur during more than 98% of the driving distance.
- 3) HILLY if slopes with a grade of less than 9% occur during more than 98% of the driving distance.
- 4) V-HILLY if the other criteria are not fulfilled.

In this case, the bird's-eye view labels coincide with the operating classes FLAT, P-FLAT (predominantly flat), HILLY and V-HILLY (very hilly), whilst the metrics are the values

imposed on the road grade (3%, 6% and 9%, respectively) and the probability of occurrence, identically equal to 0.98. Similarly, the User Factor Description (UFD) proposed by Scania considers only three levels:

- 1) FLAT if max 20% of the road section inclines more than 2%.
- 2) HILLY if between 20-40% of the road section inclines more than 2%.
- 3) V-HILLY if more than 40% of the road section inclines more than 2%.

Apart from the number of classes, the main difference with the previous example is that the label is determined by varying the probability of occurrence and specifying a constant value for the threshold on the road grade 2%. For completeness, further details about the GTA and UFD classification systems may be found in Appendix VII-B.

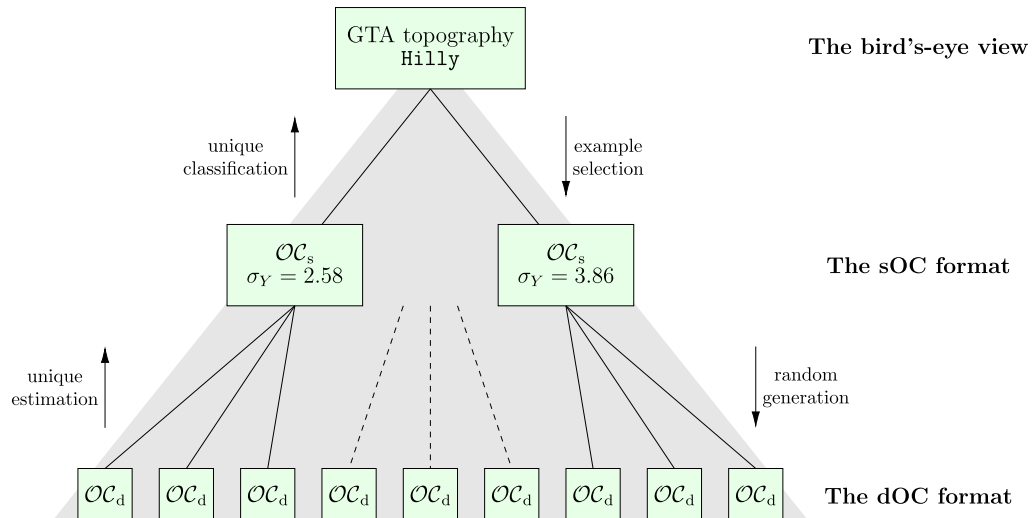
The bird's-eye-view description plays a fundamental role in the processes of optimal design and selection of road vehicles, depending on the characteristics of the intended transport application, i.e., the usage. This aspect is discussed in higher detail in Section V.

### B. THE STOCHASTIC OPERATING CYCLE

Being a description with an intermediate level of detail, the sOC condenses the statistical properties of a road mission and serves the purpose of investigating the variation problem [8], [16], [20]. The sOC makes use of a collection of stochastic models arranged in a hierarchical fashion, which are in turn equipped with their own set of stochastic parameters (mean, variance, et cetera). The modularity of the sOC description is achieved by assuming that its stochastic models are mutually independent. Some sort of realism is however preserved by introducing two sets of models: primary and secondary ones (subordinate). In this way, it becomes possible to build a composite, highly diversified structure, which guarantees a certain degree of interaction between the secondary models. At the same time, any need for considering complicated multivariate distributions is completely eliminated. Specifically, in the sOC description, primary models for the road and weather categories intimately tied in with the notions of *road type* and *season*, as explained more extensively in [50] and in Section III-A, limited to the road models. For completeness, Table 2 lists the complete set of stochastic models, plus their relative categories.

### C. THE DETERMINISTIC OPERATING CYCLE

The dOC level represents the most adequate tool to model an operating cycle when the need is on representing individual transport missions. It was specifically conceived in [9] as a virtual environment for optimal vehicle design, virtual testing and synthesis of control algorithms, with the objective of addressing the already mentioned representation problem. Its modular structure permits the addition, modification and removal of models and parameters in a very straightforward way.



**FIGURE 1.** Schematic representation of the pyramidal structure of an OC. A certain transport application includes all the missions that are equivalent according to the bird's-eye-view description (GTA system). The variation within a transport application is taken instead into account by the sOC level of representation. Finally, the dOC description reproduces the variability amongst road operations that are statistically similar. In the figure,  $OC_s$  and  $OC_d$  denote the sets of sOC and dOC parameters, respectively. The diagram assumes a single road type, and thus a unique value for the topography variance  $\sigma_Y$  (the reader is referred to Section III-B1 for clarifications).

**TABLE 2.** Stochastic models and deterministic parameters (dOC parameters) for the sOC and dOC representations. The definitions *linear* and *constant* refer to the interpolation strategy adopted for each parameter. The mathematical model of *Dirac delta* is specified when a parameter is interpreted as an event.

Model or parameter	Category	Type (for dOC)	Interp. model (for dOC)	Dim	Quantity
Speed signs	Road	Function	Constant	1	Speed limit
Altitude	Road	Function	Linear	1	Vertical coordinate
Curvature	Road	Function	Linear	1	Curvature
Ground type	Road	Function	Constant	2	Surface type, cone index
Roughness	Road	Function	Constant	2	Waviness, roughness coeff.
Stop signs	Road	Event	Dirac delta	1	Standstill time
Traffic lights	Road	Event	Dirac delta	1	Standstill time
Give way signs	Road	Event	Dirac delta	1	Standstill time
Speed bumps	Road	Event	Dirac delta	3	Length, height, angle of approach
Longitude	Road	Function	Linear	1	WGS84 longitude
Latitude	Road	Function	Linear	1	WGS84 latitude
Ambient temperature	Weather	Function	Linear	1	Temperature
Atmospheric pressure	Weather	Function	Linear	1	Pressure
Precipitation	Weather	Function	Constant	1	Precipitation amount
Wind velocity	Weather	Function	Constant	2	Velocity vector
Relative humidity	Weather	Function	Linear	1	Humidity
Traffic density	Traffic	Function	Constant	1	Density
Mission stops	Mission	Event	Dirac delta	1	Standstill time
Cargo weight	Mission	Function, event	Linear, constant	1	Payload
Power take-off	Mission	Function	Linear	1	Output power
Charging power	Mission	Function	Constant	1	Input power
Travel direction	Mission	Function	Constant	1	Driving direction

In particular, the dOC regards the sOC models as *parameters*, interpreted as discrete functions of time and position. Some parameters – like those linked to the road or weather categories – depend explicitly on either position or time; others, like the ones marked in the traffic category, on both. Additionally, the information contained in each parameter is encoded using a scalar or a vector-valued signal (see

dimensionality in Table 2), and supplemented with a suitable interpolation model.

#### D. RELATIONSHIPS BETWEEN DESCRIPTIONS

The three levels of description presented above are intimately related, and ordered hierarchically, as already illustrated in Figure 1. The connection existing between the sOC and



dOC representations is perhaps the most immediate to comprehend. Considering a fixed set of stochastic parameters, a dOC may be interpreted as an individual realization of its equivalent sOC description. In fact, a dOC may always be synthesized starting from a fully parametrized sOC. In this context, it should be emphasized that multiple dOCs generated starting from the same set of sOC parameters are statistically identical, but might exhibit significant differences in practice. On the contrary, considering a specified set of stochastic models, an sOC parametrized from a given dOC is always unique. From these reflections, it may be concluded that the relationship between the dOC and sOC levels is non-bijective.

Analogous considerations hold when approaching the higher level of the pyramid, concerning the relationship between the sOC and bird's-eye view, which are both statistical descriptions. The main difference resides in the resolution of such representations. In fact, the bird's-eye view generally encompasses an entire transport application, whereas the sOC usually targets road operations. In this context, the formal relationship subsisting between the two levels may be elucidated considering again the topography classes of the GTA system, whose thresholds are translated by the bird's-eye view into a value for the variance of the road grade. Departing from an individual set of sOC parameters, it is clear that such value may always be determined uniquely. By contrast, for a predetermined GTA class, corresponding to a continuous interval of variances, infinitely many sOCs may exist in theory. In this context, it has recently been shown in [50] that analytical relationships may be established between the sOC parameters and the set of bird's-eye view metrics and thresholds, by formulating the connection between the two representations in terms of probability and expectations. The existence of such mathematical expressions allows labeling a given sOC according to the specific classification system.

However, the attention has been restricted so far to consider individual road missions, whereas, as already explained, the bird's-eye view is mainly intended for the classification of entire transport applications. Therefore, this paper precisely investigates the distribution of the sOC parameters over the population of missions that define the application. The new dimension added to the problem permits both qualitatively and quantitatively categorizing the usage from the perspective of energy performance. The analysis is limited to the road models and focuses principally on heavy-duty trucks operating on long distances.

### III. STOCHASTIC ROAD MODELS

The distribution of the transport missions for long-haul heavy trucks is analyzed departing from the stochastic road models included in the sOC representation [47], [49], [50]. As briefly anticipated in Section II-B, these are divided into primary and secondary ones. The primary model essentially defines the road type, whereas the secondary models describe the spatial evolution of the quantities in interest along the road, and inherit their parameters from the corresponding road

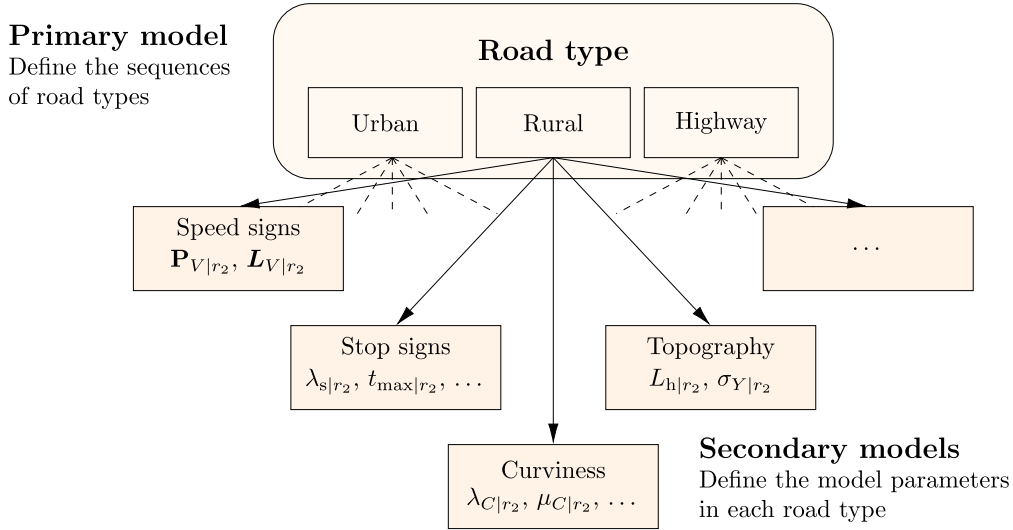
type [47]. The structure of the model is exemplified in Figure 2, where only three road types, namely urban, rural, and highway, have been considered. In this paper, the models are parametrized starting from log data collected from 33 heavy-duty vehicles operating in the area of Västra Götaland, Sweden. Hereinafter, the dataset will be consistently referred to as the OCEAN dataset. Details about the OCEAN dataset and the estimation techniques used for the following analyses are omitted from the current discussion and reported for completeness in Appendix VII-A. Instead, Sections III-A and III-B recall the stochastic models for road types, topography, curviness, stop signs, and legal speeds. Furthermore, a novel formulation for the mission length is introduced in Section III-C.

It is worth clarifying that all the road models discussed in this paper have already been validated in previous works [9], [47], [49], [50]. Moreover, whilst the mathematical structure of such models is very general and not limited to a specific geographical area, concerning more precisely the Västra Götaland region, a comparison between the actual distribution of the road properties and the parametrized descriptions is shown in [50].

In the remainder of the paper, the notation is as follows: for a generic random variable  $A : \Omega_A \mapsto \mathcal{S}_A$ , its realizations are denoted by  $a$ , unless specified otherwise. The probability and expectation operators are denoted as  $\mathbb{P}(\cdot)$  and  $\mathbb{E}(\cdot)$ , respectively; variance and covariance as  $\text{Var}(\cdot)$  and  $\text{Cov}(\cdot)$ . The set of real numbers is denoted by  $\mathbb{R}$ ; the sets of positive and negative real numbers are denoted by  $\mathbb{R}_{\geq 0}$ ,  $\mathbb{R}_{\leq 0}$  when including the zero and by  $\mathbb{R}_{> 0}$ ,  $\mathbb{R}_{< 0}$  when excluding it. The set of positive integer numbers is denoted by  $\mathbb{N}$ , whereas  $\mathbb{N}_0$  denotes the extended set of positive integers including zero, i.e.,  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ . Sequences of random variables are denoted by  $\{A_k\}_k$  (the subscript  $k$  is often dropped when the clarity allows). Finally, the acronyms PMF, PDF and CDF stand for *probability mass function*, *probability density function*, and *cumulative distribution function*, respectively.

#### A. ROAD TYPES

The road type represents the primary model for the road category. In particular, the sOC description models the sequence of road types along the vehicle's route depending on the value of the legal speed on each road segment. The two notions are intimately related, based on what was suggested previously in [55]. The stochastic model presented in this paper firstly postulates the existence of  $n_r$  different road types  $r_t \in \{r_1, \dots, r_{n_r}\}$ . These are uniquely defined starting from a sequence of  $n_r - 1$  characteristic speeds, ordered in ascending magnitude (a specific number  $n_{v|r_i}$  of speed signs is associated with each road type). The characteristic speeds mark the transition between two consecutive road types, which are in turn treated as random variable  $R_t$ , assuming values  $r_t \in \mathcal{S}_{R_t} = \{r_1, \dots, r_{n_r}\}$ , as a function of the speed sign  $V(x)$  along the road. The resulting description consists of



**FIGURE 2.** Structure of the road category in the sOC representation, with the primary (road types) and secondary (subordinate) road models.

a sequence of positions along the route, marked with the corresponding road types  $\{X_k, R_{t,k}\}_{k \in \mathbb{N}}$ .

Since the Markov property is assumed to hold [56], [57], the conditional probability of transitioning to a specific road type reads

$$\begin{aligned} \mathbb{P}(R_{t,k+1} = r_{i,k+1} | R_{t,1} = r_{i,1}, \dots, R_{t,k} = r_{i,k}) \\ = \mathbb{P}(R_{t,k+1} = r_{i,k+1} | R_{t,k} = r_{i,k}), \end{aligned} \quad (1)$$

where the generic road type  $r_i$  is an element of the road type vector  $\mathbf{R}_t = [r_1 \dots r_{n_r}]^T$ . Moreover, by modeling the locations  $X_k$  for the road types as a Poisson process, the complete model is then described by a continuous-time Markov chain, and parametrized by the entries  $p_{Rij}$  of the single-step transition matrix  $\mathbf{P}_R \in \mathbb{R}_{\geq 0}^{n_r \times n_r}$  and the  $n_r$  intensities  $\lambda_{Ri}$ , reading

$$\lambda_{Ri} = \frac{1}{L_{Ri}}, \quad (2)$$

being  $L_{Ri}$  the mean length of the road type  $r_i$ , collected in a vector  $\mathbf{L}_R = [L_{R1} \dots L_{Rn_r}]^T$ . It should be noticed that, in the construction above, no self-transitions are allowed, i.e.,  $p_{Rii} = 0$ ,  $i = 1, \dots, n_r$ , which automatically implies  $\sum_{j \neq i} p_{Rij} = 1$ ,  $i = 1, \dots, n_r$ .

Starting from the stochastic model for road types detailed above, the stationary  $\pi_R$  distribution<sup>2</sup> of the overall process may be derived as the solution of the system<sup>3</sup> [50]

$$\pi_R \mathbf{G}_R = \mathbf{0}, \quad (3a)$$

$$\sum_{i=1}^{n_r} \pi_{Ri} = 1, \quad (3b)$$

<sup>2</sup>It is supposed that there exists a unique stationary distribution.

<sup>3</sup>One of the equations in (3a) is dependent on the others, since, by construction, summing the columns of the infinitesimal generator gives the zero vector.

where the entries  $g_{Rij} = g_{Rij}(p_{Rij}, L_{Ri})$  of the generator matrix  $\mathbf{G}_R$  may be calculated as

$$g_{Rij}(p_{Rij}, L_{Ri}) = \begin{cases} \lambda_{Ri} p_{Rij} = \frac{p_{Rij}}{L_{Ri}}, & i \neq j, \\ -\lambda_{Ri} = -\frac{1}{L_{Ri}}, & i = j. \end{cases} \quad (4)$$

Equations (3) and (4) describe the stationary distribution of the road types along a road transport mission, as a function of the observed number of transitions between road types and their mean lengths. The analytical expression for the stationary vector  $\pi_R = \pi_R(\mathbf{P}_R, \mathbf{L}_R)$  is reported in Appendix VII-C1 for  $n_r = 3$ .

For an individual transport mission, the total probability and expectation of a random variable may be calculated by weighted summation over the different road types, respectively, using the total laws for probability, expectation and variance. Indeed, in the sOC representation, road segments belonging to the same road type are described using the same values for the sOC parameters of the secondary models. Denoting with  $A$  a generic random variable, the formulae are hence given as follows:

$$\begin{aligned} \mathbb{P}(A) &= \sum_{i=1}^{n_r} \mathbb{P}(A | R_t = r_i) \mathbb{P}(R_t = r_i) \\ &= \sum_{i=1}^{n_r} \mathbb{P}(A | R_t = r_i) \pi_{Ri}(\mathbf{P}_R, \mathbf{L}_R), \end{aligned} \quad (5a)$$

$$\begin{aligned} \mathbb{E}(A) &= \sum_{i=1}^{n_r} \mathbb{E}(A | R_t = r_i) \mathbb{P}(R_t = r_i) \\ &= \sum_{i=1}^{n_r} \mathbb{E}(A | R_t = r_i) \pi_{Ri}(\mathbf{P}_R, \mathbf{L}_R), \end{aligned} \quad (5b)$$

where the analytical expressions for the probabilities  $\mathbb{P}(A \mid R_t = r_i)$  and expectations  $\mathbb{E}(A \mid R_t = r_i)$  may be derived starting from the secondary models illustrated in Section III-B.

On the other hand, the sOC parameters for the road type may also be interpreted as realizations of random variables describing their distributions amongst transport applications. Starting with the probabilities  $p_{Rij}$  of the single-step transition matrix, it should be observed that they need to satisfy  $\sum_{j \neq i} p_{Rij} = 1, i = 1, \dots, n_r$ , since, by model construction, there are no self-transitions. Therefore, denoting with  $\tilde{\mathbf{P}}_{Ri}$  the vector of random variables whose realization  $\mathbf{P}_{Ri}$  collects the probabilities  $p_{Rij}$  for each row (with  $i \neq j$ ), a Dirichlet distribution is proposed in this paper, i.e.,  $\tilde{\mathbf{P}}_{Ri} \sim \text{Dir}(\alpha_{Ri})$ . The stochastic row vectors  $\tilde{\mathbf{P}}_{Ri}$  may be concisely organized into a stochastic matrix  $\tilde{\mathbf{P}}_R$ .

For a given road mission, the mean lengths  $L_{Ri}, i = 1, \dots, n_r$ , may be analogously interpreted as realizations of random variables  $\tilde{L}_{Ri}$  defined over the space of the entire transport application. In this paper, the stochastic lengths  $\tilde{L}_{Ri}$  are assumed to be independently distributed. More specifically, they are modeled as lognormal distributions, that is,  $\ln \tilde{L}_{Ri} \sim \mathcal{N}(\mu_{\tilde{L}_{Ri}}, \sigma_{\tilde{L}_{Ri}}^2)$ . The stochastic vector of mean hill lengths is accordingly denoted as  $\tilde{\mathbf{L}}_R = [\tilde{L}_{R1} \dots \tilde{L}_{Rn_r}]^T$ . For what follows, it is crucial to understand that treating the matrix  $\tilde{\mathbf{P}}_R$  and the vector of mean hill lengths  $\tilde{\mathbf{L}}_R$  as stochastic variables implies that the stationary distribution of the road type becomes a function of random variables, and thus, a random variable itself. Hereinafter, the stochastic counterpart of the stationary vector  $\pi_R$  is denoted by  $\Pi_R = \Pi_R(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$  when considering a population of random missions. In turn, it may be easily realized that the probabilities and expectations calculated according to (5) may in turn be interpreted as random variables, depending on the distribution of the primary and secondary road models.

Figure 3 illustrates the distribution of the mean lengths  $\tilde{L}_{Ri}, i = 1, 2, 3$ , for the three road types considered in the paper, namely urban, rural, and highway, along with lognormal model fitted from the OCEAN dataset.

## B. SECONDARY ROAD MODELS

This Section illustrates the secondary road models, which include road topography, curviness, stop signs, and speed signs.

### 1) ROAD TOPOGRAPHY

In the sOC description, the road topography  $Y_k$  is assumed to behave as a stationary, first-order autoregressive AR(1) model [40], [41]:

$$Y_k = \phi_{Y|r_i} Y_{k-1} + e_{Y,k}, \quad (6a)$$

$$e_{Y,k} \mid R_t = r_i \sim \mathcal{N}(0, \sigma_{eY|r_i}^2), \quad (6b)$$

where  $\phi_{Y|r_i} \in (-1, 1)$  and  $\sigma_{eY|r_i} \in (0, \infty)$  are the two characteristic parameters that depend on the given road type. According to (6), on each road type, the road grade itself

follows a normal distribution with zero mean [58], i.e.,

$$Y \mid R_t = r_i \sim \mathcal{N}(0, \sigma_{Y|r_i}^2). \quad (7)$$

Note that in the above (7) the subscript  $k$  has been dropped for the sake of notation.

Departing from (6) and (7), the conditional variance of the process may also be deduced as

$$\sigma_{Y|r_i}^2 = \frac{\sigma_{eY|r_i}^2}{1 - \phi_{Y|r_i}^2}, \quad (8)$$

where the autoregressive coefficient  $\phi_{Y|r_i}$  may be also reinterpreted as a function of the mean hill length  $L_{h|r_i}$  for the road type  $r_i$ :

$$\phi_{Y|r_i} = \sin\left(\frac{\pi}{2} - 2 \frac{L_s}{L_{h|r_i}}\right), \quad (9)$$

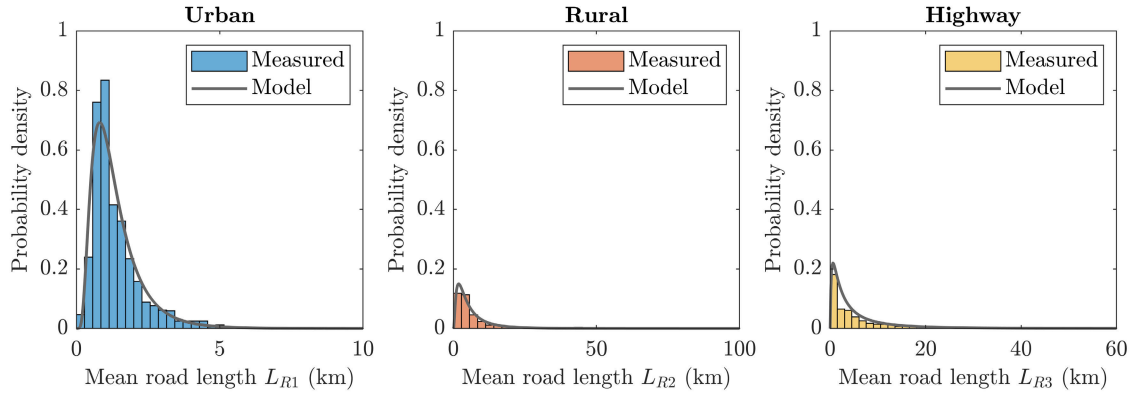
being  $L_s$  is the sampling length. Owing to these assumptions, the conditional standard deviation  $\sigma_{Y|r_i}$  and the mean hill length  $L_{h|r_i}$  condense all the information about the road topography.

An important indicator that may be used to qualify a road segment with respect to the topography model is the proportion  $p_{Y|r_i}$  between the length of the road for which the absolute value of the road grade is below a certain specified threshold  $y$ , and the total length. As shown again in [50], this may be stated mathematically in terms of a conditional probability as follows:

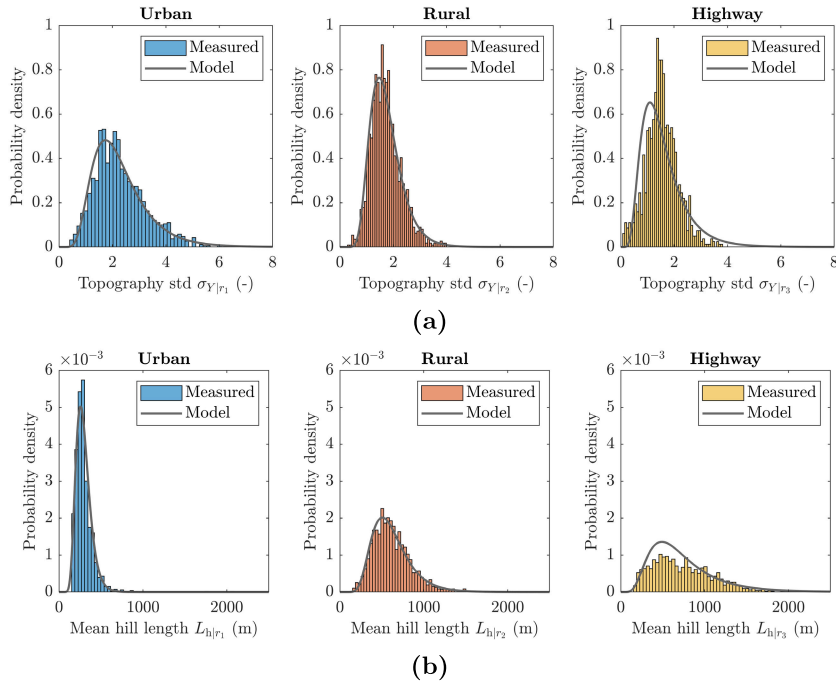
$$\begin{aligned} p_{Y|r_i}(y, \sigma_{Y|r_i}) &\triangleq \mathbb{P}(|Y| \leq y \mid R_t = r_i) \\ &= 2\Phi\left(\frac{y}{\sigma_{Y|r_i}}\right) - 1, \end{aligned} \quad (10)$$

where  $\Phi(\cdot)$  denotes the CDF of the normal distribution. Equation (10) is valid for a single road type; however, since the ratio  $p_{Y|r_i}$  may be interpreted as a conditional probability, the total probability may be calculated with the aid of (5a). Combining the example in Section II-A with (8), it may be inferred that, if a single road type is considered ( $n_r = 1$ ), the standard deviation  $\sigma_Y \equiv \sigma_{Y|r_1}$  already incorporates all the information needed to classify a road segment according to both the GTA and UFD systems. For this simplified case, the values of  $\sigma_Y \equiv \sigma_{Y|r_1}$  marking the transition between the different classes have been calculated in [50].

For a fixed road type, the characteristic parameters  $\sigma_{Y|r_i}$  and  $L_{h|r_i}$  are allowed to vary over the entire population of transport missions. More specifically, in this paper, both parameters are modeled statistically using a lognormal distribution. Denoting with  $\Sigma_{Y|r_i}$  and  $\tilde{L}_{h|r_i}$  the random variables for the road topography standard deviation and mean hill length,  $\ln \Sigma_{Y|r_i} \sim \mathcal{N}(\mu_{\Sigma_{Y|r_i}}, \sigma_{\Sigma_{Y|r_i}}^2)$ , and  $\ln \tilde{L}_{h|r_i} \sim \mathcal{N}(\mu_{\tilde{L}_{h|r_i}}, \sigma_{\tilde{L}_{h|r_i}}^2)$ . Clearly, also in this case, treating the variance of the process as a random variable, the proportion  $p_{Y|r_i}$  in (10) becomes, in turn, a random variable. Figure 4 shows the distributions of both quantities for each road type (namely urban, rural, and highway), together with the fitted lognormal models, whose parameters were estimated from the OCEAN dataset.



**FIGURE 3.** Measured distributions for road type lengths (urban, rural, highway), compared to the fitted lognormal models.



**FIGURE 4.** Measured and fitted lognormal distributions for the road topography parameters, for different road types (urban, rural, and highway): (a) topography standard deviation distributions; (b) mean hill length distributions.

## 2) ROAD CURVINESS

The sOC representation treats the curves along the road as isolated events, modeled using a sequence of locations, curvatures, and lengths  $\{X_k, C_k, L_k\}_{k \in \mathbb{N}}$ . The resulting model – referred to as the *curviness* of the road – has been introduced in [47], and is inspired from the description proposed in [59]. In particular, for each road type, the locations are assumed to follow a Poisson distribution, i.e.,

$$X_{k+1} - X_k \mid R_t = r_i \sim \mathcal{E}(\lambda_{C|r_i}), \quad (11)$$

where the intensity  $\lambda_{C|r_i} \in (0, \infty)$  should be interpreted as the mean number of curves per unit of distance along a certain road type. The curvatures  $C_k$  is modeled as a modified

lognormal distribution as follows:

$$\frac{1}{C_k} = R'_k + r_{\text{turn}}, \quad \ln R'_k \mid R_t = r_i \sim \mathcal{N}(\mu_{C|r_i}, \sigma_{C|r_i}^2), \quad (12)$$

where the parameter  $r_{\text{turn}} \in (0, \infty)$  appears because roads are constructed with a lower bounded radius. In theory,  $r_{\text{turn}}$  is not a statistical measure, but rather an inherent property of the road type. In this paper, the same value of  $r_{\text{turn}}$  is used for all the road types. Finally, the lengths of the curve  $L_k$  are modeled using a lognormal distribution:

$$\ln L_k \mid R_t = r_i \sim \mathcal{N}(\mu_{L|r_i}, \sigma_{L|r_i}^2). \quad (13)$$

Two important indicators that may be used to qualify a road based on the curviness parameter, using a reduced number of variables, relate to the expected number  $\hat{n}'_{C|r_i}$  of

curves per unit of length for which the curvature exceeds a minimum value, and to the proportion of the road  $p_{\kappa|r_i}$  for which the curvature is below a threshold  $\kappa$ . By defining with  $N'_C$  a Binomial variable describing the number of curves for which the curvature exceeds a minimum threshold  $\kappa$ , the first criterion may be formalized mathematically as a conditional expectation [50]:

$$\begin{aligned} \bar{n}'_{C|r_i}(\kappa, \lambda_{C|r_i}, \mu_{C|r_i}, \sigma_{C|r_i}) \\ \triangleq \mathbb{E}\left(\frac{N'_C}{L_{\text{tot}}} \mid R_t = r_i\right) \\ = \frac{\lambda_{C|r_i}}{2} \left[ 1 + \text{erf}\left(\frac{\ln(1/\kappa - r_{\text{turn}}) - \mu_{C|r_i}}{\sqrt{2}\sigma_{C|r_i}}\right) \right], \quad (14) \end{aligned}$$

where  $L_{\text{tot}}$  denotes the total length of the road segment. Equation (14) is the same type of relationship prescribed by the UFD system (see Appendix VII-B). In particular, it may be observed that the characteristic parameters  $\mu_{L|r_i}$  and  $\sigma_{L|r_i}$  do not appear in (14), implying that two roads which differ for the mean length and variance of the curvature are equivalent according to this criterion.

On the other hand, describing a segment in terms of curvature distribution along the road, as done by the GTA system, provides a more general relationship, involving all the characteristic parameters of the model. To this end, the curvature may be regarded as a continuous function  $K(X)$  of the coordinate along the road, interpreted in turn as a stochastic variable. Accordingly, the conditional distribution of the curvature may be derived approximately as follows [50]:

$$\begin{aligned} p_{\kappa|r_i}(\kappa, \lambda_{C|r_i}, \mu_{C|r_i}, \sigma_{C|r_i}, \mu_{L|r_i}, \sigma_{L|r_i}) \\ \triangleq \mathbb{P}(K \leq \kappa \mid R_t = r_i) \\ = 1 - \frac{\lambda_{C|r_i}}{2} \exp\left(\mu_{L|r_i} + \frac{\sigma_{L|r_i}^2}{2}\right) \\ \times \left[ 1 + \text{erf}\left(\frac{\ln(1/\kappa - r_{\text{turn}}) - \mu_{C|r_i}}{\sqrt{2}\sigma_{C|r_i}}\right) \right]. \quad (15) \end{aligned}$$

Equations (14) and (15) correspond again to a conditional expectation and probability, respectively, and thus the total expectation and probability may be calculated using (5). The resulting formulae would completely qualify a single transport mission with respect to the curviness model.

However, all the characteristic parameters also vary over the set of missions defining the transport application. Starting with the curve intensity, it is assumed that, for each individual road type, the corresponding stochastic variable is lognormally distributed, that is,  $\ln \Lambda_{C|r_i} \sim \mathcal{N}(\mu_{\Lambda_{C|r_i}}, \sigma_{\Lambda_{C|r_i}}^2)$ . The comparison between the measured distributions and the models fitted from the OCEAN dataset is shown in Figure 5.

When the log-radius described by (12) is allowed to vary over the population of missions, the corresponding random variables  $M_{C|r_i}$  and  $\Sigma_{C|r_i}$  are modeled using a normal and lognormal distribution, respectively. In formulae:  $M_{C|r_i} \sim \mathcal{N}(\mu_{M_{C|r_i}}, \sigma_{M_{C|r_i}}^2)$ , and  $\ln \Sigma_{C|r_i} \sim \mathcal{N}(\mu_{\Sigma_{C|r_i}}, \sigma_{\Sigma_{C|r_i}}^2)$ . Figure 6

illustrates the comparison between the measured distribution and the fitted models for both quantities, according to the three road types considered in the paper.

Similar models are also employed to describe the variation of the parameters for the curve length modeled as in (13). Indeed, the following distributions are assumed:  $M_{L|r_i} \sim \mathcal{N}(\mu_{M_{L|r_i}}, \sigma_{M_{L|r_i}}^2)$ , and  $\ln \Sigma_{L|r_i} \sim \mathcal{N}(\mu_{\Sigma_{L|r_i}}, \sigma_{\Sigma_{L|r_i}}^2)$ . The analytical pdf, together with the distribution extracted from the OCEAN dataset, is plotted in Figure 7 for both quantities.

Generally speaking, it may be noticed that the agreement is satisfactory for the variables relating to the urban and rural roads, whereas the data collected on highways exhibit larger discrepancies with the fitted model, and more skewed distribution. The reason for that, which partially legitimates the assumptions on the chosen distribution, should be ascribed to the method applied in estimating the curvature from log data, which relies on measurements of the yaw rate. In this context, the reaction of the driver to very large values of the curvature radius may not be captured by such measurements, implying that small curvatures are automatically filtered out.

It is worth observing that, since all the characteristic parameters of the model are regarded as random variables when considering a population of missions, the quantities  $\bar{n}'_{C|r_i}$  and  $p_{\kappa|r_i}$  in (14) and (15) also become random variables.

### 3) STOP SIGNS

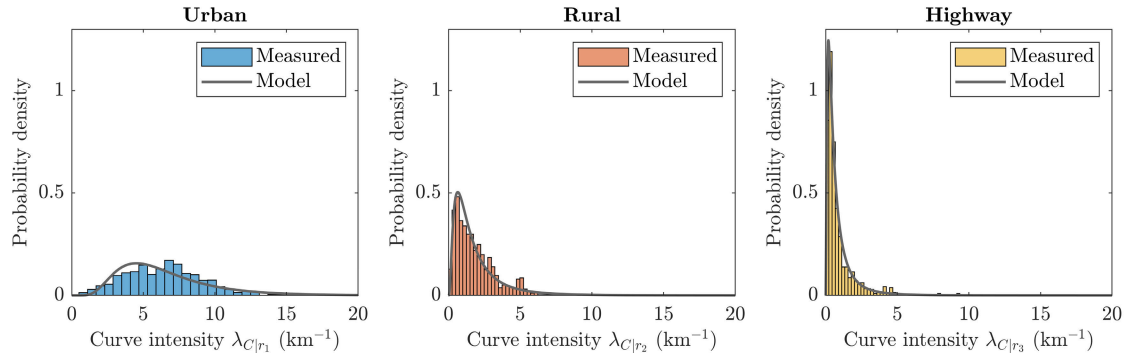
In the sOC representation, stop signs<sup>4</sup> are treated as independent events and described by the sequence  $\{X_k, T_{s,k}\}_{k \in \mathbb{N}}$ , where  $X_k$  is again the location, and  $T_{s,k}$  is interpreted as a recommended time. A Poisson model similar to that in (11), but with intensity  $\lambda_{s|r_i}$ , is used to model the distance between two consecutive stops. Moreover, for each road type, the recommended time  $T_{s,k}$  is supposed to be uniformly distributed between a minimum and maximum value, i.e.,  $T_{s,k} \mid R_t = r_i \sim \mathcal{U}(t_{\min|r_i}, t_{\max|r_i})$ . Usually, the conditional intensities  $\lambda_{s|r_i}$  are sufficient to completely qualify a road mission with respect to the stop signs model, whereas the stop times are not involved in the reduced description.

Considering the entire population of missions spanning the transport application, the corresponding random variables for the intensity of the event, and the minimum and maximum times are denoted by  $\Lambda_{s|r_i}$ ,  $T_{\min|r_i}$  and  $T_{\max|r_i}$ , respectively. As usual, the conditional intensity  $\Lambda_{s|r_i}$  is modeled using a lognormal distribution, that is,  $\ln \Lambda_{s|r_i} \sim \mathcal{N}(\mu_{\Lambda_{s|r_i}}, \sigma_{\Lambda_{s|r_i}}^2)$ .

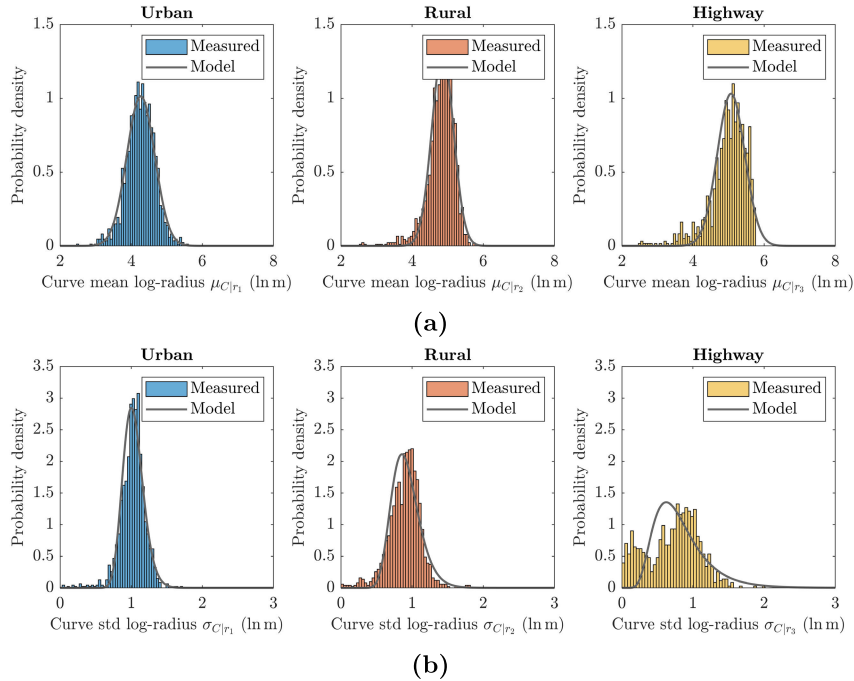
The minimum and maximum values for the recommended waiting time are instead modeled using a Gamma and modified Gamma distribution, respectively. In particular, it is first assumed that the minimum time is Gamma distributed with shape and rate parameters  $\alpha_{T_{\min|r_i}}, \beta_{T_{\min|r_i}} \in (0, \infty)$ , i.e.,  $T_{\min|r_i} \sim \text{Ga}(\alpha_{T_{\min|r_i}}, \beta_{T_{\min|r_i}})$ . Then, the stochastic model for the maximum time is constructed by adding a positive increment, that is,  $T_{\max|r_i} = T_{\min|r_i} + \Delta T_{r_i}$ , where  $\Delta T_{r_i} \sim \text{Ga}(\alpha_{\Delta T_{r_i}}, \beta_{\Delta T_{r_i}})$ . This approach ensures that any realization

<sup>4</sup>Actually, a similar model is also used for speed bumps, which are not considered in the present paper.





**FIGURE 5.** Measured and fitted lognormal distributions for the curve intensity, for different road types (urban, rural, highway).



**FIGURE 6.** Distributions of curve log-radius parameters for different road types (urban, rural, and highway): (a) curve mean log-radius; (b) curve std log-radius.

$t_{\max|r_i}$  will always be greater than the corresponding lower limit  $t_{\min|r_i}$ . Apart from obvious advantages connected with its support, the Gamma distribution is chosen in this paper for two additional reasons, the first being that it represents a generalization of the exponential distribution, and the second that, if the rate parameters  $\beta_{T_{\min|r_i}}$  and  $\beta_{\Delta T|r_i}$  accidentally coincide, then the maximum time  $T_{\max|r_i}$  becomes itself Gamma distributed.

The distributions for the stop signs parameters are compared to the proposed stochastic models in Figures 8 and 9.

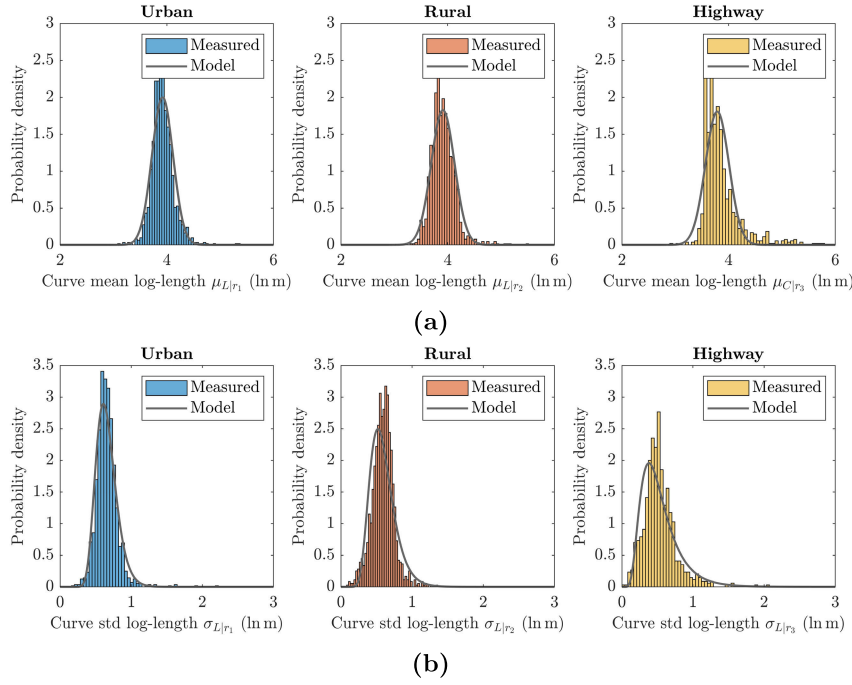
#### 4) SPEED SIGNS

Similarly as for the model introduced in Section III-A, for a given road type<sup>5</sup>  $r_k$ , the speed signs are regarded as piecewise

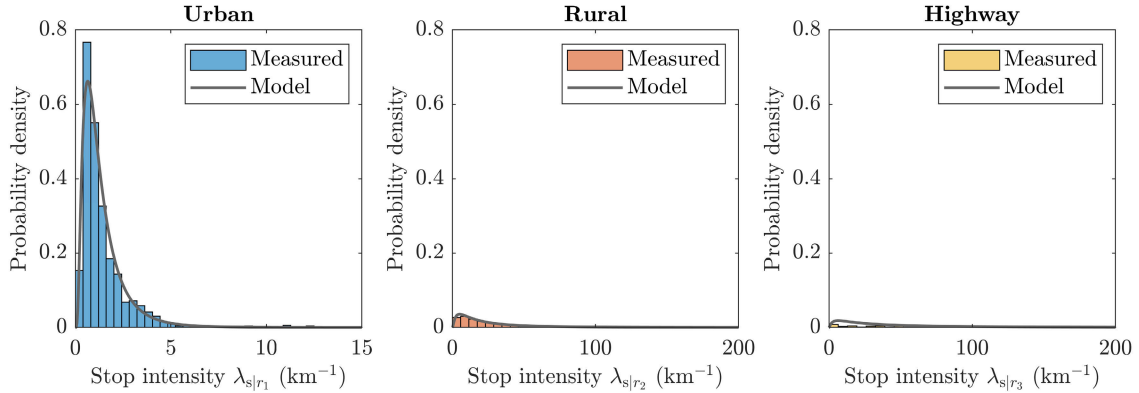
<sup>5</sup>In this section, the index  $k$  is used for the road type to avoid confusion with the indexing of the speed sign probabilities.

constant, right-side continuous functions of the position [47]. In particular, speed signs are modeled as a random process  $V = V(x)$  along with the position on the road. Accordingly, the variable  $V(x)$  assumes discrete values in the state space  $\mathcal{S}_{V|r_k} = \{v_{1|r_k}, \dots, v_{n_{v|r_k}|r_k}\}$ , where  $n_{v|r_k}$  denotes the finite number of possible speed limits for the road type  $r_k$ . Also in this case, the complete model collects a sequence of positions, marked with the corresponding values for the legal speed, i.e.,  $\{X_k, V_k\}_{k \in \mathbb{N}}$ . More specifically, the sequence of speed limits is approximated by using a Markov chain [56], [57], and assumes discrete values  $v_{i|r_k}$  organized into the speed vector  $\mathbf{v}_{|r_k} = [v_{1|r_k} \dots v_{n_{v|r_k}|r_k}]^T$ .

Accordingly, the entries of the conditional Markov probability matrix  $\mathbf{P}_{V|r_k} \in \mathbb{R}_{\geq 0}^{n_{v|r_k} \times n_{v|r_k}}$  fully characterize the discrete chain, with  $p_{vij|r_k}$  modeling the conditional probability of transitioning from state  $i$  to state  $j$ . Since no self-transition are allowed, as usual, they satisfy



**FIGURE 7.** Distributions of curve log-length parameters for different road types (urban, rural, and highway): (a) curve mean log-length; (b) curve std log-length.



**FIGURE 8.** Measured and fitted lognormal distributions for the stop intensity, for different road types (urban, rural, highway).

$\sum_{j \neq i}^{n_{v|r_k}} p_{Vij|r_k} = 1, i = 1, \dots, n_{v|r_k}$ . The speed sign locations are again modeled as in (11). For each road type, the  $n_{v|r_k}$  intensities  $\lambda_{V1|r_k}, \dots, \lambda_{Vn_{v|r_k}|r_k}$  may be deduced from the corresponding mean lengths  $L_{Vi|r_k}$ :

$$\lambda_{Vi|r_k} = \frac{1}{L_{Vi|r_k}}, \quad (16)$$

collected into a vector  $\mathbf{L}_{V|r_k} = [L_{V1|r_k} \dots L_{Vn_{v|r_k}|r_k}]^T$ . The resulting model is completely parametrized by the conditional probabilities  $p_{Vij|r_k}$  and the  $n_{v|r_k}$  mean lengths  $L_{Vi|r_k}$  (or, alternatively, the intensities  $\lambda_{Vi|r_k}$ ). Additionally, it should be observed that the speed  $V(x)$  itself behaves as continuous-time Markov chain [56], since the distance between consecutive transitions is modeled using a Poisson process. In particular, the stationary distribution  $\boldsymbol{\pi}_{V|r_k}$  of the overall process may be derived departing from its generator

matrix  $\mathbf{G}_{V|r_k}$ , and satisfies the usual set of equations

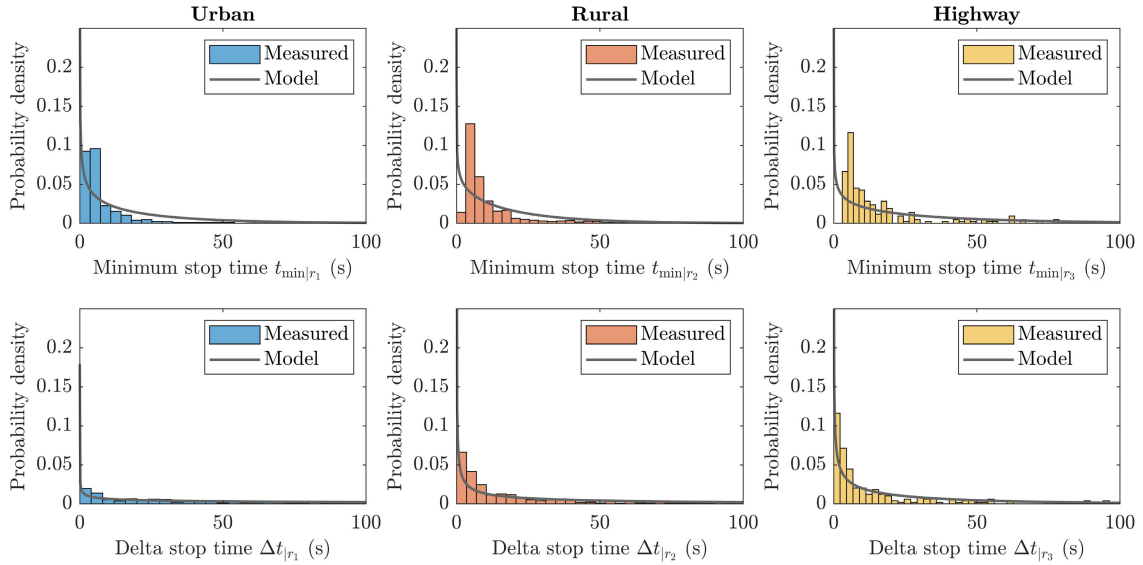
$$\boldsymbol{\pi}_{V|r_k} \mathbf{G}_{V|r_k} = \mathbf{0}, \quad (17a)$$

$$\sum_{i=1}^{n_{v|r_k}} \pi_{Vi|r_k} = 1, \quad (17b)$$

where the entries  $g_{Vij|r_k} = g_{Vij|r_k}(p_{Vij|r_k}, L_{Vi|r_k})$  of  $\mathbf{G}_{V|r_k}$  are given by

$$g_{Vij|r_k}(p_{Vij|r_k}, L_{Vi|r_k}) = \begin{cases} \frac{p_{Vij|r_k}}{L_{Vi|r_k}}, & i \neq j, \\ -\frac{1}{L_{Vi|r_k}}, & i = j. \end{cases} \quad (18)$$

Closed-form expressions for  $\boldsymbol{\pi}_{V|r_k} = \boldsymbol{\pi}_{V|r_k}(\mathbf{P}_{V|r_k}, \mathbf{L}_{V|r_k})$  are reported in Appendix VII-C2 for the cases  $n_{v|r_k} = 2$  and 3, respectively.



**FIGURE 9.** Measured and fitted gamma distributions for the stop time, for different road types (urban, rural, highway).

A parameter of fundamental importance for what follows is the conditional expectation of the legal speed over a segment. For a given road type, this may be computed as in [50]:

$$\begin{aligned} \hat{v}_{|r_k}(\mathbf{P}_{V|r_k}, \mathbf{L}_{V|r_k}) &\triangleq \mathbb{E}(V \mid R_t = r_k) \\ &= \sum_{i=1}^{n_{V|r_k}} v_{i|r_k} \pi_{v_{i|r_k}}(\mathbf{P}_{V|r_k}, \mathbf{L}_{V|r_k}). \end{aligned} \quad (19)$$

It should be emphasized that, for an individual transport mission, the expected value  $\hat{v}_{|r_k} = \hat{v}_{|r_k}(\mathbf{P}_{V|r_k}, \mathbf{L}_{V|r_k})$  derived in Eq.(19) may be expressed as a function of the single-step transition probabilities, and mean speed lengths.

Similar to what done for the road type model in Section III-A, for a given road type, each row of the transition matrix  $\mathbf{P}_{V|r_k}$  is assumed to follow a Dirichlet distribution, that is  $\tilde{\mathbf{P}}_{V|r_k} \sim \text{Dir}(\boldsymbol{\alpha}_{V|r_k})$ , where  $\tilde{\mathbf{P}}_{V|r_k}$  is a vector of random variables for the conditional transition probabilities from state  $i$  and for a given road type, and  $\boldsymbol{\alpha}_{V|r_k} \in \mathbb{R}_{>0}^{n_{V|r_k}-1}$  a vector of parameters. For ease of notation, the row vectors  $\tilde{\mathbf{P}}_{V|r_k}$  may be organized into the stochastic matrix  $\tilde{\mathbf{P}}_{V|r_k}$ .

As usual, for each road type  $r_k$ , when interpreted as random variables over the entire population of missions, the mean speed lengths  $\tilde{L}_{V|r_k}$  are supposed to be lognormally distributed, i.e.,  $\ln \tilde{L}_{V|r_k} \sim \mathcal{N}(\mu_{\tilde{L}_{V|r_k}}, \sigma_{\tilde{L}_{V|r_k}}^2)$ ,  $i = 1, \dots, n_{V|r_k}$ . Figure 10 illustrates the distributions of the mean speed lengths for each speed and road type, according to the OCEAN dataset. It should be noticed that parameters for the random variable  $\tilde{L}_{V3|r_3}$  (corresponding to a speed of  $110 \text{ km h}^{-1}$ ) could not be estimated from the available measurements. In fact, the single-step transition matrix  $\mathbf{P}_{V|r_3}$  for the highway road reduced to a deterministic two-by-two matrix.

When regarding the single-step transition probabilities and the mean speed lengths as random variables defined

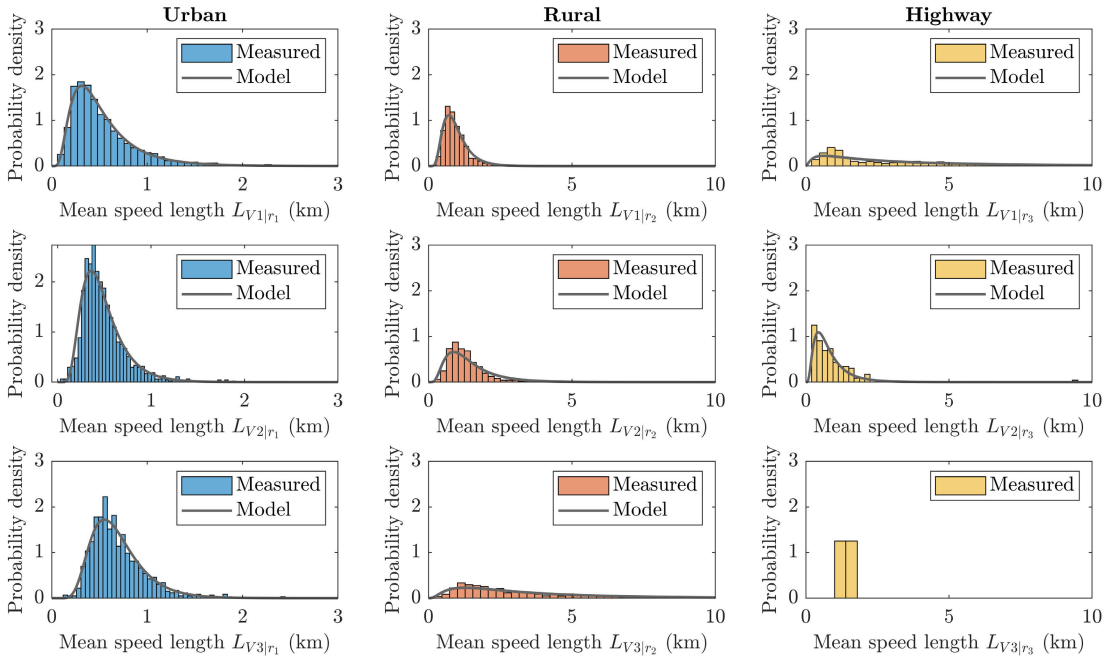
over the space of transport missions, the generic stationary distribution vector  $\boldsymbol{\pi}_{V|r_k}$  becomes a random vector  $\boldsymbol{\Pi}_{V|r_k}(\tilde{\mathbf{P}}_{V|r_k}, \tilde{\mathbf{L}}_{V|r_k})$  collecting the stochastic stationary distributions for the speed limits, conditioned to the road type  $r_k$ . The analytical expression for the random vector  $\boldsymbol{\Pi}_{V|r_k}$  as a function of  $\tilde{\mathbf{P}}_{V|r_k}$  and  $\tilde{\mathbf{L}}_{V|r_k}$  may be derived immediately from (17) and (18), by considering the single-step transition probabilities and the mean lengths, and are reported in Appendix VII-C2 for completeness. Accordingly, the conditional mean speed derived as in (19) should also be interpreted as a random variable.

All the stochastic models presented so far, together with the values for the corresponding parameters extracted from the OCEAN dataset, are summarized in Table 3.

### C. MISSION LENGTH

The mission length is not technically a road parameter; nonetheless, it plays a fundamental role in the classification of a road transport mission, according to both the GTA and UFD representations. Therefore, it is included in the present paper.

In this context, it should be mentioned that daily driving distances are traditionally modeled as lognormal, Gamma, or Weibull distributions [60]. In the present paper, the focus is not actually on the distance travelled during the entire day, but rather on the individual mission length. For the sake of simplicity, a Gamma variable  $\tilde{L}_m \sim \text{Ga}(\alpha_{\tilde{L}_m}, \beta_{\tilde{L}_m})$ , assuming values  $L_m \in \mathcal{S}_{\tilde{L}_m} \equiv \mathbb{R}_{>0}$ , is used. This specific choice is also motivated again by the fact that the sum of Gamma variables with the same rate parameter is still a Gamma variable. For the case of mission length, all the realizations are generated from a unique distribution, and thus any sequence of consecutive missions will also obey the same law. This introduces some freedom in the definition of a mission itself, which could be



**FIGURE 10.** Measured and fitted lognormal distributions for the speed mean length, for different road types (urban, rural, highway).

interpreted as a single trip associated to a specific task, as well as a collection of trips or subtasks.

The comparison between the empirical and analytical distributions is shown in Fig. 11 for  $\alpha_{\tilde{L}_m} = 1.31$  and  $\beta_{\tilde{L}_m} = 0.016 \text{ km}^{-1}$ . Accordingly, the mean value and the variance of the distribution amount approximately to  $81.72 \text{ km}$  and  $4.69 \cdot 10^3 \text{ km}^2$ , respectively.

#### IV. ANALYSIS AND CLASSIFICATION

The present Section attempts to answer the very fundamental question about the representativeness of the OC, already formulated in Section I-B. More specifically, the basic idea behind the following investigation is to understand what type of usage the parametrized OC is representative of – concerning the road properties and their distributions – and how to exploit this knowledge to design and test more energy-efficient vehicles. In this context, the question is formalized in terms of a classification problem and addressed from the perspective of two already existing descriptions, namely the GTA and UFD systems. To this end, new stochastic variables, condensing the salient information about a road transport mission according to the models detailed in Section III, are introduced and analyzed in terms of distributions and moments. These new variables may be deduced directly from the GTA and UFD descriptions, and are henceforth referred to as *composite variables*, since they may be expressed as a function of the *simple* random variables introduced before.

##### A. COMPOSITE RANDOM VARIABLES

For a given mission, the statistical indicators used by the GTA and UFD description target road properties that may

be reasonably anticipated to have a substantial impact on energy performance. Considering a whole population of transport missions, these indicators will eventually behave as random variables. Their distribution and moments may be determined, analytically (where and whether possible) or numerically, starting from those of the stochastic variables introduced in Section III.

##### 1) ROAD TOPOGRAPHY

Starting with the model for topography discussed in Section III-B1, the road may be qualified considering two main indicators. The first, as already mentioned in Section III-B1, consists of the proportion between the road length for which the grade is lower than a specified threshold  $y$  and the total length of the segment. The second relates to the hill mean length. In the first case, the analytical expression for the new random variable as functions of the conditional random variances  $\Sigma_{Y|r_i}$ ,  $i = 1, \dots, n_r$ , may be deduced by combining (5) with (10). The final formula is as follows:

$$P_y(y, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{L}_R) \triangleq \sum_{i=1}^{n_r} \left[ 2\Phi\left(\frac{y}{\Sigma_{Y|r_i}}\right) - 1 \right] \times \Pi_{Ri}(\tilde{\mathbf{P}}_R, \tilde{L}_R), \quad (20)$$

where the vector  $\Sigma_Y \triangleq [\Sigma_{Y|r_1} \dots \Sigma_{Y|r_{n_r}}]^T$  collects the conditional stochastic standard deviations. The relationship for the random mean hill length may be instead derived by noticing that the generic  $L_{h|r_i}$  can already be interpreted as a

TABLE 3. Stochastic models for road parameters.

Model	Distribution	Parameters	Value (per road type)			Unit
			Urban	Rural	Highway	
Road type						
Road type probability matrix $\tilde{\mathbf{P}}_R$	Dirichlet	$\alpha_{Ri}$	[0.85 0.10]	[0.59 0.48]	[0.14 1.18]	-
Mean road length $\tilde{L}_{Ri}$	Lognormal	$\mu_{\tilde{L}_{Ri}}$	0.15	1.48	1.18	ln km
		$\sigma_{\tilde{L}_{Ri}}$	0.59	0.93	1.30	ln km
Topography						
Road std $\Sigma_{Y r_i}$	Lognormal	$\mu_{\Sigma_{Y r_i}}$	0.74	0.50	0.34	-
		$\sigma_{\Sigma_{Y r_i}}$	0.43	0.36	0.49	-
Mean hill length $\tilde{L}_{h r_i}$	Lognormal	$\mu_{\tilde{L}_{h r_i}}$	5.63	6.36	6.47	ln m
		$\sigma_{\tilde{L}_{h r_i}}$	0.30	0.37	0.52	ln m
Curviness						
Curve intensity $\Lambda_{C r_i}$	Lognormal	$\mu_{\Lambda_{C r_i}}$	1.76	0.29	-0.64	ln km <sup>-1</sup>
		$\sigma_{\Lambda_{C r_i}}$	0.50	0.85	1.02	ln km <sup>-1</sup>
Curve mean log-radius $M_{C r_i}$	Normal	$\mu_{M_{C r_i}}$	4.27	4.86	5.07	ln m
		$\sigma_{M_{C r_i}}$	0.39	0.30	0.39	ln m
Curve std log-radius $\Sigma_{C r_i}$	Lognormal	$\mu_{\Sigma_{C r_i}}$	0.02	-0.11	-0.29	ln m
		$\sigma_{\Sigma_{C r_i}}$	0.14	0.22	0.43	ln m
Curve log log-length $M_{L r_i}$	Normal	$\mu_{M_{L r_i}}$	3.92	3.91	3.79	ln m
		$\sigma_{M_{L r_i}}$	0.20	0.22	0.22	ln m
Curve std log-length $\Sigma_{L r_i}$	Lognormal	$\mu_{\Sigma_{L r_i}}$	-0.45	-0.57	-0.74	ln m
		$\sigma_{\Sigma_{L r_i}}$	0.22	0.29	0.48	ln m
Stop signs						
Stop intensity $\Lambda_{s r_i}$	Lognormal	$\mu_{\Lambda_{s r_i}}$	-0.09	-2.94	-3.71	ln km <sup>-1</sup>
		$\sigma_{\Lambda_{s r_i}}$	0.72	1.14	1.36	ln km <sup>-1</sup>
Minimum stop time $T_{\min r_i}$	Gamma	$\alpha_{T_{\min r_i}}$	0.63	0.79	0.69	-
		$\beta_{T_{\min r_i}}$	0.03	0.04	0.02	s <sup>-1</sup>
Delta stop time $\Delta T_{ r_i}$	Gamma	$\alpha_{\Delta T_{ r_i}}$	0.64	0.43	0.53	-
		$\beta_{\Delta T_{ r_i}}$	$2.5 \cdot 10^{-3}$	$5.7 \cdot 10^{-3}$	$1.75 \cdot 10^{-2}$	s <sup>-1</sup>
Speed signs						
Speed sign probability matrix $\tilde{\mathbf{P}}_{V r_i}$	Dirichlet	$\alpha_{Vj r_i}$	[0.39 0.42] [0.36 0.37] [0.30 0.36]	[0.33 0.28] [0.48 1.03] [0.35 1.22]	-	-
Mean speed	Lognormal	$\mu_{\tilde{L}_{Vj r_i}}$	[-0.79 -0.82 -0.44]	[-0.14 0.20 1.06]	[ 1.14 -0.38]	ln km
length $\tilde{L}_{Vj r_i}$		$\sigma_{\tilde{L}_{Vj r_i}}$	[0.60 0.45 0.39]	[0.46 0.59 0.91]	[1.27 0.67]	ln km

conditional expectation, yielding

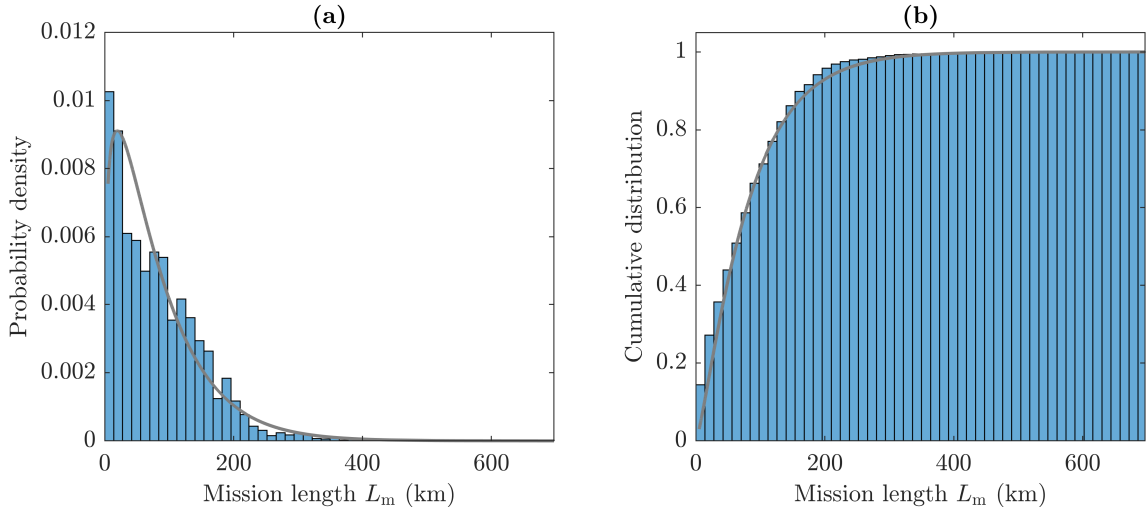
$$\tilde{L}_h(\tilde{L}_h, \tilde{\mathbf{P}}_R, \tilde{L}_R) \triangleq \sum_{i=1}^{n_r} \tilde{L}_{h|r_i} \Pi_{Ri}(\tilde{\mathbf{P}}_R, \tilde{L}_R), \quad (21)$$

in which  $\tilde{L}_h \triangleq [\tilde{L}_{h|r_1} \dots \tilde{L}_{h|r_{n_r}}]^T$  is a vector collecting the stochastic mean hill lengths. As already shown previously, both the GTA and UFD classification systems resort to an expression of the same type as in (20), whereas the mean hill

length is disregarded.<sup>6</sup> Nevertheless, the latter parameter is integrated in the analysis, since the influence of the hill length on vehicular performance may be in general non-negligible, as pointed out in [9].

<sup>6</sup>It is worth noticing that the GTA and UFD use (20) in relation to the topography, whereas this paper distinguishes more specifically between road grade length ratio for (20) and mean hill length for (21).





**FIGURE 11.** Measured distribution and the analytical PDF and CDF of the mission length: (a) probability density; (b) cumulative distribution. Model parameters:  $\alpha_{L_m} = 1.31$  and  $\beta_{L_m} = 0.016 \text{ km}^{-1}$ .

## 2) ROAD CURVINESS

For the curviness model, two different relationships may be derived based on (14) and (15), respectively. These would correspond to the criterion imposed by the UFD and GTA representations, in turn. For the expected number of curves, the stochastic variable describing the variation over the mission population may be expressed as a function of the other random parameters according to being this time  $\Lambda_C \triangleq [\Lambda_{C|r_1} \dots \Lambda_{C|r_{n_r}}]^T$ ,  $M_C \triangleq [M_{C|r_1} \dots M_{C|r_{n_r}}]^T$ ,  $\Sigma_C \triangleq [\Sigma_{C|r_1} \dots \Sigma_{C|r_{n_r}}]^T$ . It is worth observing that the analytical expression for  $\tilde{N}'_C = \tilde{N}'_C(\Lambda_C, M_C, \Sigma_C, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$  in (22), as shown at the bottom of the next page, does not depend upon the stochastic mean lengths  $M_{L|r_i}$  and standard deviations  $\Sigma_{L|r_i}$ ,  $i = 1, \dots, n_r$ . On the other hand, combining (5) and (15) yields (23) with, as usual,  $\mathbf{M}_L \triangleq [M_{L|r_1} \dots M_{L|r_{n_r}}]^T$ , and  $\Sigma_L \triangleq [\Sigma_{L|r_1} \dots \Sigma_{L|r_{n_r}}]^T$ . Comparing (23), as shown at the bottom of the next page, to (22), it may be clearly noticed that the former relationship is more complete, and involves all the characteristic parameters of the curviness model.

## 3) STOP SIGNS

With respect to the stop signs model, only the conditional intensities are assumed to play a role in the qualification of the transport mission, whereas the contribution of the minimum and maximum stop times is disregarded. This criterion, relating to the expected number of stops along the road, is present in the UFD classification system, but completely absent in the GTA representation. In particular, since the conditional intensities  $\lambda_{s|r_i}$ ,  $i = 1, \dots, n_r$ , may be interpreted as expected number of events per unit of length, the following relationship may be derived directly from the law of total expectation:

$$\tilde{N}_s(\Lambda_s, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \triangleq \sum_{i=1}^{n_r} \Lambda_{s|r_i} \Pi_{Ri}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R), \quad (24)$$

where it has been defined  $\Lambda_s \triangleq [\Lambda_{s|r_1} \dots \Lambda_{s|r_{n_r}}]^T$ .

## 4) SPEED SIGNS

Concerning the distribution of the legal speed along the road, the GTA and UFD propose different criteria, based on the expected number of transitions between speeds and the predominant speed value along the mission, respectively. Whilst the first approach leads to a rather complicated analytical expression that does not involve the values  $v_{i|r_k}$  of the speed itself, the second one cannot be consistently used to uniquely qualify a road transport mission, as discussed more extensively in [50]. Therefore, in this paper, an alternative criterion, based on the mean legal speed and borrowed from [54], is used. The latter allows for ease of classification and interpretation, since directly relates to the values  $v_{i|r_k}$  of the speed limits. Owing to the premises above, the analytical expression for the mean speed may be derived starting from (19), and reads

$$\begin{aligned} & \hat{v}(\tilde{\mathbf{P}}_{V|r_1}, \dots, \tilde{\mathbf{P}}_{V|r_{n_r}}, \tilde{\mathbf{L}}_{V|r_1}, \dots, \tilde{\mathbf{L}}_{V|r_{n_r}}, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ & \triangleq \sum_{i=1}^{n_r} \sum_{j=1}^{n_{v|r_i}} v_{j|r_i} \Pi_{Vj|r_i}(\tilde{\mathbf{P}}_{V|r_i}, \tilde{\mathbf{L}}_{V|r_i}) \Pi_{Ri}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R), \quad (25) \end{aligned}$$

where the stochastic variables  $\Pi_{Vj|r_i}$  have been introduced in Section III-B4.

Equations (20), (21), (22), (23), (24), and (25) provide the analytical expressions for the stochastic variables involved in the classification of an entire transport application. The latter may be conveniently classified by considering the expectation of each variable over the population of road missions. The composite variables, together with their interpretation in terms of expectations or probability, are summarized in Table 4.

**TABLE 4.** Composite variables and their interpretation.

Model	Composite variable	Simple variables	Equation	Interpretation
Topography	$P_y$	$\Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R$	(20)	Probability
	$\tilde{L}_h$	$\tilde{L}_h, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R$	(21)	Expectation
Curviness	$\tilde{N}'_C$	$\Lambda_C, M_C, \Sigma_C, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R$	(22)	Expectation
	$P_\kappa$	$\Lambda_C, M_C, \Sigma_C, M_L, \Sigma_L, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R$	(23)	Probability
Stop signs	$\tilde{N}_s$	$\Lambda_s, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R$	(24)	Expectation
Speed signs	$\hat{V}$	$\tilde{\mathbf{P}}_{V r_1}, \dots, \tilde{\mathbf{P}}_{V r_{n_r}}, \tilde{\mathbf{L}}_{V r_1}, \dots, \tilde{\mathbf{L}}_{V r_{n_r}}, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R$	(25)	Expectation

## B. NUMERICAL ANALYSIS AND CLASSIFICATION

The composite variables introduced in Section IV-A permit to easily classify both individual road missions and entire transport applications. In particular, the latter may be qualified depending on the mean value assumed by the quantities in (20), (21), (22), (23), (24), and (25), computed over the whole population of missions (Section IV-B1). This is the approach followed by both the GTA and UFD systems. On the other hand, whilst a single indicator may be sufficient to classify the transport application, more detailed information about the multivariate distribution of the composite variables may be useful for production planning and predictive maintenance (Section IV-B2).

### 1) CLASSIFICATION USING NUMERICAL EXPECTATIONS

Since the distributions for the simple random variables are known from Section III (Table 3), those for the composite variables may be calculated numerically, or generated by simulation. It should be emphasized that, as opposed to the simple variables, the composite variables are correlated through the stationary probabilities  $\Pi_{Ri}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$ ,  $i = 1, \dots, n_r$ , which implies the need to describe an entire transport application using a multivariate random vector, whose definition would, in turn, depend upon the specific choice of classification system.<sup>7</sup> Specifically, taking inspiration from the GTA and UFD representations, respectively, this paper considers the random vectors  $\mathbf{X}_{\text{GTA}}$  and  $\mathbf{X}_{\text{UFD}}$ ,

defined as

$$\mathbf{X}_{\text{GTA}}(\mathbf{X}_{\mathcal{R}_s}) \triangleq \begin{bmatrix} P_y(3, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ P_y(6, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ P_y(9, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \tilde{L}_h(\tilde{L}_h, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ P_\kappa(0.008, \Lambda_C, \dots, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \tilde{N}_s(\Lambda_s, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \hat{V}(\tilde{\mathbf{P}}_{V|r_1}, \dots, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \tilde{L}_m \end{bmatrix}, \quad (26)$$

and

$$\mathbf{X}_{\text{UFD}}(\mathbf{X}_{\mathcal{R}_s}) \triangleq \begin{bmatrix} P_y(2, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \tilde{L}_h(\tilde{L}_h, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \tilde{N}'_C(0.008, \Lambda_C, \dots, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \tilde{N}_s(\Lambda_s, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \hat{V}(\tilde{\mathbf{P}}_{V|r_1}, \dots, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \\ \tilde{L}_m \end{bmatrix}, \quad (27)$$

where the vector  $\mathbf{X}_{\mathcal{R}_s}$  has been defined collecting all the simple stochastic random variables for the models presented in Section III, i.e.,  $\mathbf{X}_{\mathcal{R}_s} \triangleq [\tilde{\mathbf{P}}_R \ \tilde{\mathbf{L}}_R \ \Sigma_Y \ \dots \ \tilde{\mathbf{P}}_{V|r_1} \ \dots \ \tilde{\mathbf{P}}_{V|r_{n_r}} \ \tilde{\mathbf{L}}_{V|r_1} \ \dots \ \tilde{\mathbf{L}}_{V|r_{n_r}} \ \tilde{L}_m]^T$ .

In (26), the first three stochastic components of the vector  $\mathbf{X}_{\text{GTA}} = \mathbf{X}_{\text{GTA}}(\mathbf{X}_{\mathcal{R}_s})$  are used to classify the road in respect to the topography parameter (more specifically, the road grade length ratio). Indeed, referring to the example reported in

<sup>7</sup>It is worth emphasizing that the mission length  $\tilde{L}_m$  is independent of the road models, and also from the composite variables.

$$\tilde{N}'_C(\kappa, \Lambda_C, M_C, \Sigma_C, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \triangleq \sum_{i=1}^{n_r} \frac{\Lambda_{C|r_i}}{2} \left[ 1 + \operatorname{erf} \left( \frac{\ln(1/\kappa - r_{\text{turn}}) - M_{C|r_i}}{\sqrt{2}\Sigma_{C|r_i}} \right) \right] \Pi_{Ri}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R), \quad (22)$$

$$P_\kappa(\kappa, \Lambda_C, M_C, \Sigma_C, M_L, \Sigma_L, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \triangleq 1 - \sum_{i=1}^{n_r} \frac{\Lambda_{C|r_i}}{2} \exp \left( M_{L|r_i} + \frac{\Sigma_{L|r_i}^2}{2} \right) \times \left[ 1 + \operatorname{erf} \left( \frac{\ln(1/\kappa - r_{\text{turn}}) - M_{C|r_i}}{\sqrt{2}\Sigma_{C|r_i}} \right) \right] \Pi_{Ri}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R), \quad (23)$$

Section II-A, an individual mission is labeled as **FLAT** if  $P_y(3, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) > 0.98$ , **P-FLAT** if the previous condition is not met but  $P_y(6, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) > 0.98$ , **HILLY** if both the previous criteria are not satisfied but  $P_y(9, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) > 0.98$ , and **V-HILLY** otherwise. In this context, it should be observed that, for a given road transport mission,  $y_1 \leq y_2 \implies P_y(y_1, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) \leq P_y(y_2, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$ , which follows from the fact that  $P_y(y, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$  is actually a cumulative distribution function. Therefore, the criterion imposed on the topography according to the GTA classification system is always well-defined. The fourth component of  $\mathbf{X}_{\text{GTA}}$  also relates to the topography model, but targets instead the mean hill length. Finally, the last four components are used to label the mission with respect to the curviness (for which the criterion is expressed again in terms of probability), stop signs, speed signs, and mission length parameters. As opposed to the multivariate random variable  $\mathbf{X}_{\text{GTA}}$ , the stochastic vector  $\mathbf{X}_{\text{UFD}} = \mathbf{X}_{\text{UFD}}(\mathcal{R}_s)$  includes only one relationship for the road grade length ratio. In fact, according to the example adduced in Section II-A, the label is directly assigned considering the value assumed by the probability  $P_y(2, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$  appearing in (27). The remaining components of  $\mathbf{X}_{\text{UFD}}$  relate to the mean hill length, curviness (for which the condition is formulated in terms of expectation), stop signs, speed signs, and mission length.

The difficulty of analyzing the multivariate random variables  $\mathbf{X}_{\text{GTA}}$  and  $\mathbf{X}_{\text{UFD}}$  resides in the fact that all the composite variables are correlated. Whereas a rigorous treatment may be prohibitive, and out of the scope of the present paper, approximated results may be derived following the method of propagation of uncertainty (the reader may refer to Section VII-D for further details). For the sake of brevity, the discussion is here restricted to considerations about the expected values for the vectors. Indeed, the transport application may be labeled by considering the expected values  $\mathbb{E}(\mathbf{X}_{\text{GTA}}(\mathcal{R}_s))$  and  $\mathbb{E}(\mathbf{X}_{\text{UFD}}(\mathcal{R}_s))$ , respectively. The computation of the mean clearly yields  $\mathbb{E}(\mathbf{X}_{\text{GTA}}(\mathcal{R}_s)) = [\mathbb{E}(P_y(3, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)) \dots \mathbb{E}(\tilde{\mathbf{L}}_m)]^T$ , and analogously  $\mathbb{E}(\mathbf{X}_{\text{UFD}}(\mathcal{R}_s)) = [\mathbb{E}(P_y(2, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)) \dots \mathbb{E}(\tilde{\mathbf{L}}_m)]^T$ , which implies that a transport application may be simply classified by determining the marginal distributions and the expectations of each individual random variable contained in the vectors  $\mathbf{X}_{\text{GTA}}$  and  $\mathbf{X}_{\text{UFD}}$ . In this paper, the marginal distributions for the composite variables presented in Section IV-A were estimated numerically by generating a population of 10000 synthetic sOCs (whose stochastic parameters are reported in a separate file). The results of such a process are illustrated in Figure 12, where the marginal PDFs are plotted, along with the estimated mean value (solid grey lines), and the analytical expectation computed using a first-order approximation (dashed yellow lines), as better explained in the following Section IV-C. In Figure 12, the cumulative distribution for the road grade length ratio is only plotted for  $y = 2$  in  $P_y(y, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$ , i.e., according to the UFD description. In particular, it is interesting to observe

how both the realizations  $p_y$  and  $p_k$  range between zero and one: this could already be anticipated, since the composite variables  $P_y$  and  $P_k$  denote a probability.<sup>8</sup> The mean values of the composite variables are listed in Table 5, where the corresponding GTA and UFD classes are also reported. As a general observation, it should be noticed that the notion of representativeness for the parametrized OC, evaluated in terms of composite variables, is heavily dependent on the choice of the classification system.

As a concluding remark, it is worth observing that using a single road type, i.e.,  $n_r = 1$ , implies that the composite variables are independent,<sup>9</sup> and, thus, the joint PMF becomes the product of the individual PMFs, at least for the vector  $\mathbf{X}_{\text{UFD}}$  (a more exhaustive discussion is reported in Section VII-D). This would enormously simplify the mathematical treatment.<sup>10</sup>

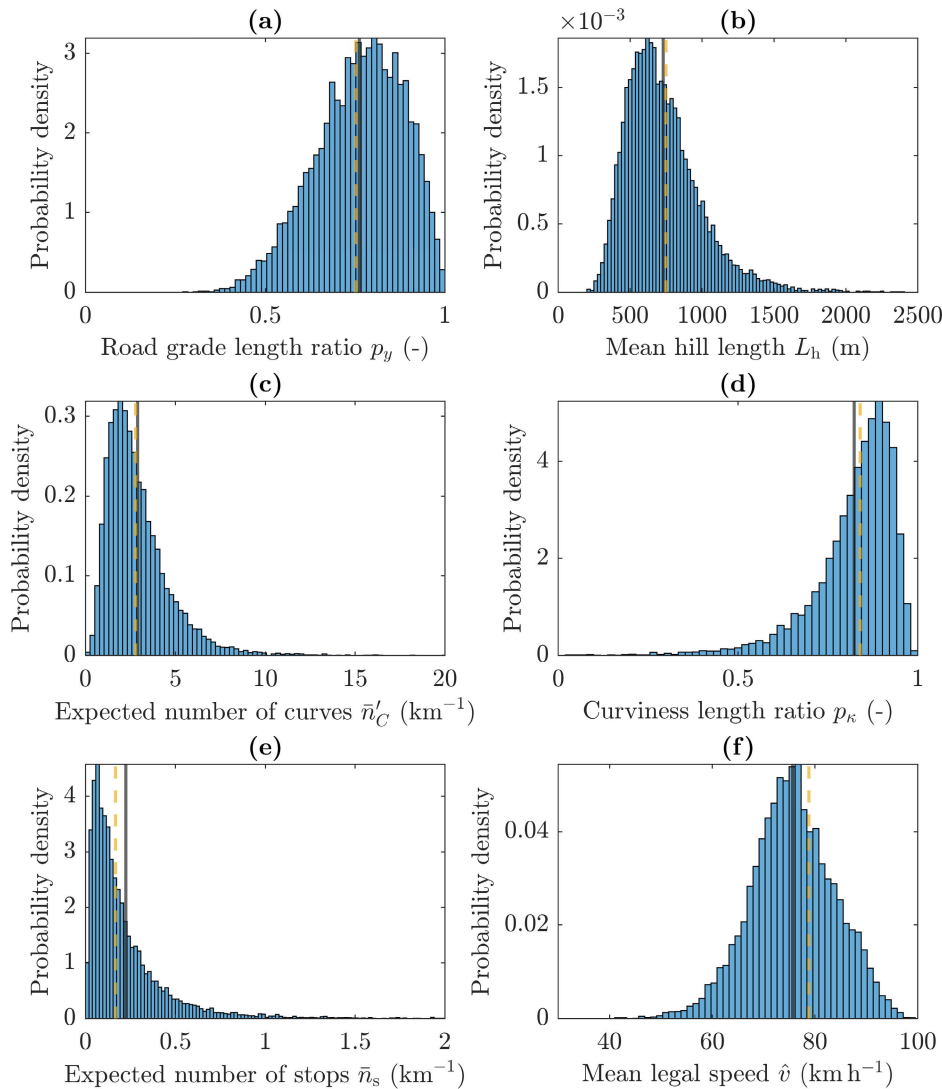
## 2) MULTIVARIATE PMF ESTIMATION

Relying on more detailed information concerning the actual distribution of the road missions, according to some specified classification system, may facilitate the processes of production planning and predictive maintenance. In this context, the joint PMF for a given transport application, opportunely parametrized in terms of sOCs, may be estimated numerically using the relationships established for the composite variables of Section IV-A, and prescribing a set of thresholds on each model in isolation. In this operation, the combinatorial nature of the problem and the resolution of the chosen classification system represent two key factors in determining the total number of possible combinations of classes. For what follows, it is worth emphasizing that estimating the joint PMF for the discrete case is, in fact, different from estimating the continuous multivariate PDF for the composite variables of Section IV-A. Specifically, in this paper, the joint PMF was estimated for the GTA and UFD classification systems considering a population of 10000 sOCs synthesized using the stochastic models detailed in Section III. The theoretical number of possible combinations may be calculated as 2304 for both the GTA and UFD descriptions, whereas the generated sOC yielded a total of 325 and 344 possible combinations, respectively. Using the GTA description, the predominant combination is classified as **HIGH** for legal speed, **P-FLAT** for road grade, **MEDIUM** for mean hill length, **LOW** for curviness, **RESIDENTIAL** for stop signs, and **L-DISTANCE** for mission length, with a probability of occurrence close to 9.8%. Instead, the corresponding labels assigned by the UFD representation are **HIGH** for speed,

<sup>8</sup>In reality, the expression for the generic probability  $p_{k|i_i}$  in (15) is approximated (see [50] for further details). Therefore, the extremely rare event of generating negative curviness length ratios might actually happen. By simulation, the probability of such an occurrence was estimated to be around 1/1000.

<sup>9</sup>Except for the road grade length ratios in (26).

<sup>10</sup>It should be noticed that joint PDFs for the composite variables relating to the curviness model would still be very complicated to determine analytically.



**FIGURE 12.** Distributions of the composite variables, with estimated (solid grey lines) and approximated (dashed yellow lines) mean values.

**TABLE 5.** Estimated mean values for the composite variables of Section IV-A, plus the mission length, along with the corresponding GTA and UFD classes.

Description	Composite variable	Expectation	Unit	GTA Class	UFD Class
Road grade length ratio	$P_y(2, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	0.761	-	P-FLAT	HILLY
	$P_y(3, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	0.893			
	$P_y(6, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	0.989			
	$P_y(9, \Sigma_Y, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	0.998			
Mean hill length	$\tilde{L}_h(\tilde{\mathbf{L}}_h, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	722	m	MEDIUM	MEDIUM
Number of curves	$\tilde{N}'_C(0.008, \mathbf{A}_C, \mathbf{M}_C, \mathbf{\Sigma}_C, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	2.944	km <sup>-1</sup>	HIGH	-
Curviness length ratio	$P_\kappa(0.008, \mathbf{A}_C, \mathbf{M}_C, \mathbf{\Sigma}_C, \mathbf{M}_L, \mathbf{\Sigma}_L, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	0.819	-	-	LOW
Number of stops	$\tilde{N}_s(\mathbf{A}_s, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	0.225	km <sup>-1</sup>	LIGHT	LIGHT
Mean legal speed	$\hat{V}(\tilde{\mathbf{P}}_{V r_1}, \dots, \tilde{\mathbf{P}}_{V r_{n_r}}, \tilde{\mathbf{L}}_{V r_1}, \dots, \tilde{\mathbf{L}}_{V r_{n_r}}, \tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$	76.18	km h <sup>-1</sup>	HIGH	HIGH
Mission length	$\tilde{L}_m$	81.72	km	L-DISTANCE	L-DISTANCE

FLAT for road grade, MEDIUM for mean hill length, HIGH for curviness, FLUID for stop signs, and L-DISTANCE for mission length, with a maximum probability of 9.4%. Again, the other probability values are reported in a separate file, and not discussed in this paper for the sake of brevity.

Instead, Figure 13 illustrates some joint PMFs obtained by only considering two variables at a time, that is, marginalizing over all the remaining dimensions of the problem. The resulting bivariate distributions are easier to visualize, given the reduced number of involved parameters, and may be used to elucidate some interesting differences between the

two classification systems. Starting with the first set of histograms in Figure 13, showing the PMF for the mission length and road grade, most of the missions, representing nearly 41.9% of the total, are labeled as L-DISTANCE/P-FLAT using the GTA criteria, with the second most populated combination of classes being the REGIONAL/P-FLAT, covering an additional 27.7%. On the other hand, the UFD description collocates the majority of the missions in the L-DISTANCE/HILLY and L-DISTANCE/FLAT combinations (27.6 and 24.6%, respectively). In both cases, the combinations with STOP&GO are extremely rare. Similar proportions could already be expected from the analyses preliminarily conducted in Sections IV-B1 and IV-C, and align with the results obtained by considering the individual expectations of the composite variables. Generally speaking, it may also be observed that, excluding the P-FLAT class, the missions are more homogeneously distributed according to the UFD representation. This may be explained by noticing that the UFD system is actually deficient in the P-FLAT class. As a consequence, the missions labeled P-FLAT by the GTA description are divided almost equally between the FLAT and HILLY classes. Moreover, some of the missions qualified as HILLY using the GTA metrics belong instead to the V-HILLY class according to the UFD representation. A similar effect, also due to the different resolution of the two classification systems, may be observed by looking at the bivariate PMF of the mission length and legal speed parameters. Indeed, with respect to the latter, the majority of the missions is labeled either MODERATE or HIGH using the thresholds prescribed by the GTA description, which misses the V-HIGH class. The UFD representation also distributes the transport missions mainly between the MODERATE and HIGH classes, but with an opposite tendency. The predominant combinations are the L-DISTANCE/MODERATE for the GTA system and L-DISTANCE/HIGH for the UFD (41.5 and 43.0% of the total number of missions, respectively). Finally, concerning the bivariate PMF for the mission length and curviness parameters, the resolution is the same for both classification systems. In this case, the substantial discrepancy between the two distributions should be ascribed to the different criteria used: whilst the GTA imposes a limit on the portion of the road for which the curvature exceeds a certain value, the UFD prescribes a sequence of thresholds on the expected number of curves, without specifying anything about the covered length. As a result, the predominant combinations are the L-DISTANCE/LOW for the GTA description (40.4%) and L-DISTANCE/HIGH for the UFD representation (38.0%). In any case, the general consideration is again as in Section IV-B1: the representativeness of the OC should be evaluated case-by-case according to the specific choice of the classification system. Additional details about the distributions plotted in Figure 13 are provided in Table 6.

### C. BUILDING A SINGLE REPRESENTATIVE sOC

A single representative sOC may be derived that embodies the salient properties of the entire transport application, using a

reduced number of parameters. The intuitive approach proposed in this paper is to consider the sOC resulting from selecting the mean values of the sOC parameters, interpreted as random variables over the population of missions. Accordingly, the resulting reference sOC may be parametrized using the vector-valued mean  $\mu_{X_{\mathcal{R}_s}} = \mathbb{E}(X_{\mathcal{R}_s}) \triangleq [\mathbb{E}(\tilde{P}_R) \ \mathbb{E}(\tilde{L}_R) \ \mathbb{E}(\Sigma_y) \ \dots \ \mathbb{E}(\tilde{P}_{V|r_1}) \ \dots \ \mathbb{E}(\tilde{L}_{V|r_{nr}}) \ \mathbb{E}(\tilde{L}_m)]^T$ . There are several reasons that legitimate such a choice. The first is that such a parametrized sOC would obviously be independent of the assumed classification system. In fact, if a representative sOC were constructed starting from the composite variables introduced in Section IV-A, it would be inherently and heavily conditioned by the structure of a specific bird's-eye-view description, which may however vary amongst vehicle manufacturers and road operators. Therefore, its practical relevance would be potentially limited to some specific application. Besides, it should be observed that the composite variables of Section IV-A incorporate many different sOC parameters. Therefore, defining a single reference sOC in terms of composite variables would ultimately determine the feasible combinations of sOC parameters, but would not provide their value explicitly. This aspect inherently relates to the non-bijective relationship between the different levels of the OC description. Finally, there is another advantage, more theoretical in nature, connected with constructing a representative sOC according to the outlined approach: the realizations for the composite variables may be interpreted as a zeroth-order approximation of their expected values over the population of road missions.<sup>11</sup> This last consideration allows to easily evaluate the representativeness of the reference sOC also in terms of a specific bird's-eye view description.

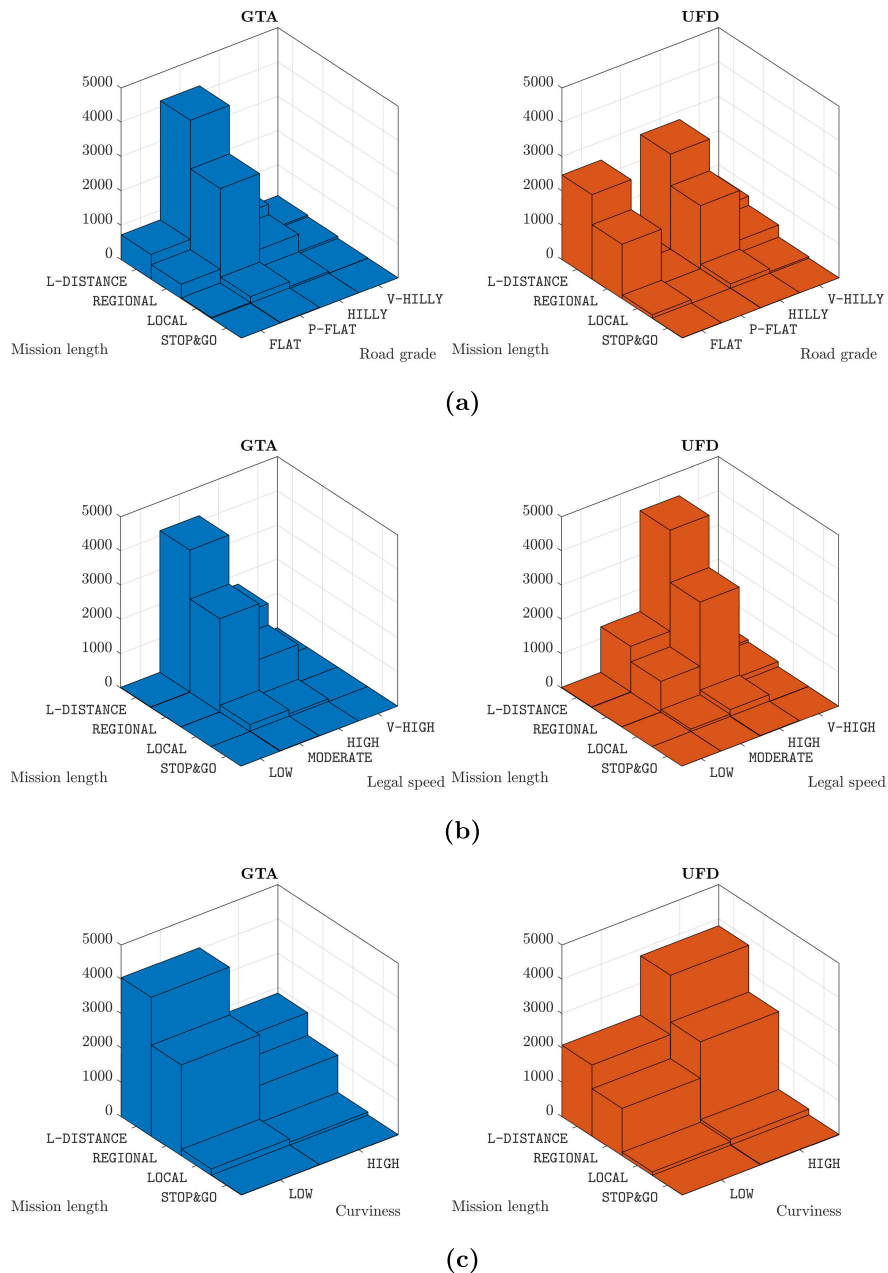
As an example, the expectations determined analytically using the zeroth-order approximation are plotted in Figure 12 (dashed yellow lines) for all the composite variables considered in this paper. Accordingly, the dashed yellow lines provide an intuitive understanding of how the reference sOC is collocated in the distribution of possible missions, and how distant it is from the mean sOC described in terms of composite variables. By looking at Figure 12, it may be generally concluded that, for the specific set of sOC parameters under consideration, the match between the two sOCs is quite satisfactory. Indeed, excluding the number of stops, the approximated expectations exhibit negligible relative errors (usually below 5%) compared to those estimated numerically. Moreover, classifying the transport application using the approximated values for the expectations yields the same combination of classes as in Section IV-B, the unique exception concerning the stop signs, for which the label FLUID is assigned.

### V. APPLICATIONS

The present Section discusses some practical applications of the whole OC edifice. The examples adduced in the following

<sup>11</sup> Since, for a generic vector-valued function  $f(\cdot)$  of random variables  $X = [X_1 \ \dots \ X_n]^T$ , its expectation may be approximated as  $\mathbb{E}(f(X)) \approx f(\mathbb{E}(X))$  using the propagation of the uncertainty technique.





**FIGURE 13.** Joint PMFs for different combination of parameters according to the GTA (left) and UFD (right) representations. (a) Mission length and road grade; (b) Mission length and legal speed; (c) Mission length and curviness. The distributions refer to a population of 10000 missions generated using the sOCs extracted from the OCEAN dataset.

should be intended as illustrative, and mainly serve the purpose of demonstrating the potential of the OC format. Two problems, of different nature, are considered in this paper: the first concerns accurate simulation of vehicular performance, in the contexts of certification or early design; the second deals with the process of optimal production planning.

#### A. ASSESSING VARIATION IN ENERGY PERFORMANCE

Departing from a fully parametrized transport application, vehicular performance may be easily assessed in a virtual environment by combining the OC with vehicle and driver

models. If a detailed study is required, a large population of road missions may be generated directly by simulating (in a statistical sense) the stochastic models introduced in Section III, permitting to accurately represent variation in usage and road characteristics. This type of analysis would be more suited when referring to a certain usage in terms of energy efficiency, relating for example to a specific geographical area. Moreover, since the stochastic parameters are allowed to vary over the population of missions that define the transport application, the influence of different road parameters – including, e.g., the composite variables

**TABLE 6.** Distribution of the road transport missions for different combinations of parameters.

Road grade/Mission length	GTA				UFD			
	STOP&GO	LOCAL	REGIONAL	L-DISTANCE	STOP&GO	LOCAL	REGIONAL	L-DISTANCE
FLAT	2	26	425	720	9	105	1590	2458
P-FLAT	13	194	2769	4185	0	0	0	0
HILLY	1	54	564	866	6	131	1836	2760
V-HILLY	1	7	65	108	2	45	397	661
Legal speed/Mission length	GTA				UFD			
	STOP&GO	LOCAL	REGIONAL	L-DISTANCE	STOP&GO	LOCAL	REGIONAL	L-DISTANCE
LOW	0	0	6	16	0	0	0	16
MODERATE	13	200	2721	4152	2	71	895	1347
HIGH	4	81	1096	1711	15	203	2766	4297
V-HIGH	0	0	0	0	0	7	156	219
Curviness/Mission length	GTA				UFD			
	STOP&GO	LOCAL	REGIONAL	L-DISTANCE	STOP&GO	LOCAL	REGIONAL	L-DISTANCE
LOW	9	189	2649	4043	3	102	1383	2075
HIGH	9	92	1174	1836	14	179	2440	3804

**TABLE 7.** Parameters for a single representative sOC, parametrized by the vector-valued mean  $\mu_{\mathcal{X}|\mathcal{R}_s}$ .

Model	Parameter	Formula	Value (per road type)			Unit
			Urban	Rural	Highway	
Road type	$L_{Ri}$	$\exp\left(\mu_{\tilde{L}_{Ri}} + \frac{\sigma_{\tilde{L}_{Ri}}^2}{2}\right)$	1.38	6.77	7.57	km
	$p_{Rij}$	$\frac{\alpha_{Rij}}{\sum_{j \neq i} \alpha_{Rij}}$	[0 0.89 0.11]	[0.55 0 0.45]	[0.11 0.89 0]	-
Speed signs	$v_{i r_k}$		[30 40 50] <sup>T</sup>	[60 70 80] <sup>T</sup>	[90 100] <sup>T</sup>	km h <sup>-1</sup>
	$L_{V i r_k}$	$\exp\left(\mu_{\tilde{L}_{V i r_k}} + \frac{\sigma_{\tilde{L}_{V i r_k}}^2}{2}\right)$	[0.54 0.49 1.44] <sup>T</sup>	[0.97 1.36 6.47] <sup>T</sup>	[7.00 1.83] <sup>T</sup>	km
	$p_{V ij r_k}$	$\frac{\alpha_{V ij r_k}}{\sum_{j \neq i} \alpha_{V ij r_k}}$	$\begin{bmatrix} 0 & 0.48 & 0.52 \\ 0.49 & 0 & 0.51 \\ 0.45 & 0.55 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0.54 & 0.46 \\ 0.32 & 0 & 0.68 \\ 0.22 & 0.78 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	-
Stop signs	$\lambda_{s r_i}$	$\exp\left(\mu_{\tilde{\lambda}_{s r_i}} + \frac{\sigma_{\tilde{\lambda}_{s r_i}}^2}{2}\right)$	1.18	0.10	0.06	km <sup>-1</sup>
	$t_{\min r_i}$	$\frac{\alpha_{T_{\min r_i}}}{\beta_{T_{\min r_i}}}$	21.00	19.75	34.50	s
	$t_{\max r_i}$	$\frac{\alpha_{T_{\min r_i}}}{\beta_{T_{\min r_i}}} + \frac{\alpha_{\Delta T r_i}}{\beta_{\Delta T r_i}}$	277.00	95.19	64.79	s
Curviness	$\lambda_{C r_i}$	$\exp\left(\mu_{\tilde{\lambda}_{C r_i}} + \frac{\sigma_{\tilde{\lambda}_{C r_i}}^2}{2}\right)$	6.59	1.92	0.89	km <sup>-1</sup>
	$\mu_{C r_i}$	$\mu_{M_{C r_i}}$	4.27	4.86	5.07	ln m
	$\sigma_{C r_i}$	$\exp\left(\mu_{\tilde{\sigma}_{C r_i}} + \frac{\sigma_{\tilde{\sigma}_{C r_i}}^2}{2}\right)$	1.03	0.92	0.82	ln m
	$r_{\text{turn}}$		12.5	12.5	12.5	m
	$\mu_{L r_i}$	$\mu_{M_{L r_i}}$	3.92	3.91	3.79	ln m
	$\sigma_{L r_i}$	$\exp\left(\mu_{\tilde{\sigma}_{L r_i}} + \frac{\sigma_{\tilde{\sigma}_{L r_i}}^2}{2}\right)$	0.65	0.59	0.54	ln m
Topography	$L_{h r_i}$	$\exp\left(\mu_{\tilde{L}_{h r_i}} + \frac{\sigma_{\tilde{L}_{h r_i}}^2}{2}\right)$	291	811	739	m
	$\sigma_{Y r_i}$	$\exp\left(\mu_{\tilde{\sigma}_{Y r_i}} + \frac{\sigma_{\tilde{\sigma}_{Y r_i}}^2}{2}\right)$	2.30	1.76	1.58	%

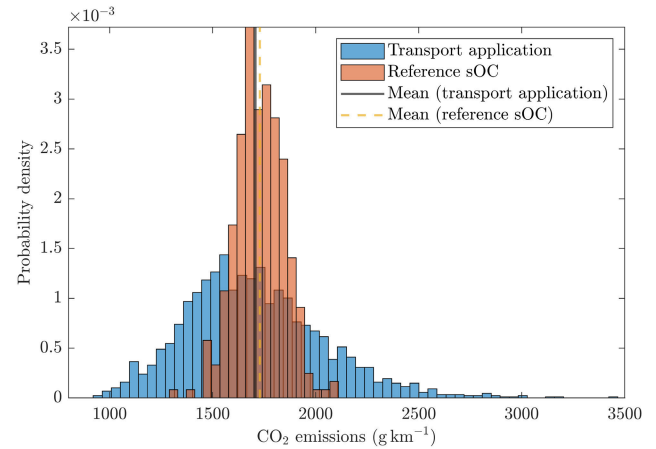
introduced in Section IV-A – may be investigated comprehensively. Such an investigation is deliberately omitted from the present paper, but may be carried out as done in [47] and [49]. On the other hand, considering a single reference sOC may be sufficient when dealing with certification and preliminary assessments of new vehicle designs. In this context, it should be clarified that a unique sOC may still describe variation in usage, and actually, multiple dOCs need to be synthesized

starting from the same sOC to accurately predict spread in performance.

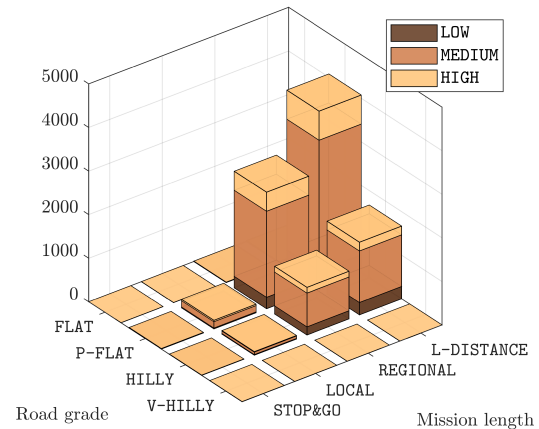
Figure 14 compares the CO<sub>2</sub> emissions estimated by generating and then simulating multiple dOCs in the VehProp environment (the reader may refer to Appendix VII-E for further details about the implementation and computational aspects). More specifically, the blue histogram shows the distribution of pollutant emissions obtained by simulating

10000 dOCs, randomly synthesized from the same population of 10000 sOCs as in Section IV. On the contrary, the orange histogram refers to a population of 300<sup>12</sup> dOCs generated from the same reference sOC parametrized in Section IV-C. In both cases, the missions were mirrored to ensure a balance in the potential (gravitational) energy. Coherently with the scope of the paper, Figure 14 has been produced considering a Volvo FH13 vehicle model, equipped with a diesel engine, an actuated, stepped gearbox with 12 forward gears and a kerb weight of 7540 kg. These specifications are, in fact, typical of heavy-duty trucks of the kind commonly used for long-haul missions. Additional information about the vehicle's configuration can be found in the Volvo datasheets.<sup>13</sup>

Along with the actual distributions of CO<sub>2</sub> emissions, the mean values are also reported in Figure 14, evaluated concerning both the entire population of stochastically generated sOCs (solid grey line), and the representative sOC parametrized in Section IV-C (dashed yellow line). The values computed numerically amount to 1708 and 1729 g km<sup>-1</sup>, respectively. The difference between the two estimates is negligible, which may be explained informally as follows. First, the existence of a function  $f_{CO_2}(\cdot)$  may be postulated describing the mean CO<sub>2</sub> emissions calculated over a population of dOCs, synthesized from the same sOC. In general, the analytical expression for such a function may not be known *a priori* but, for a fixed set of sOC parameters, it would be deterministic, and dependent solely upon the sOC parameters, the vehicle's specifications, and possibly the driver's behaviour. Then, treating the sOC parameters as random variables, and with a similar rationale as in Section IV-C, its expectation over the population of missions may be approximated as  $\mathbb{E}(f_{CO_2}(X_{\mathcal{R}_s})) \approx f_{CO_2}(\mathbb{E}(X_{\mathcal{R}_s})) \equiv f_{CO_2}(\mu_{X_{\mathcal{R}_s}})$ . This result, which may perhaps appear trivial, asserts that the average emissions may be estimated simply by simulating an adequately large number of dOCs generated from the representative sOC, parametrized by the vector-valued mean parameters  $\mu_{X_{\mathcal{R}_s}}$ . This observation allows to considerably reduce the number of needed simulations, compared to synthesizing a larger population of missions that explicitly consider the variation of sOC parameters. Moreover, since the above approximation is valid for any other function, it may be conjectured that, in order to estimate the mean value of a quantity in interest, simulating a number of dOCs originating from the representative sOC defined as in Section IV-C should always be sufficient. Clearly, it is quite evident from Figure 14 that a similar argument does not apply concerning the variances and standard deviations. Indeed, accounting for variation between road transport missions, which mathematically translates into the variation of the sOC parameters, produces a much larger spread in performance than considering a single representative sOC. In fact, the standard



**FIGURE 14.** Distribution of CO<sub>2</sub> emissions estimated in simulation by generating dOCs according to both the entire transport application (blue histogram) and the reference sOC (orange histogram).



**FIGURE 15.** Distribution of road transport missions according to the reduced GTA description of Section V-B.

deviations estimated from simulating the complete transport application and the reference sOC amount to 360 and 117 g km<sup>-1</sup>, respectively.

## B. OPTIMAL SELECTION AND PRODUCTION PLANNING

The second example discussed in this Section elucidates the application of the OC format for the purposes of optimal design selection and production planning. The idea is to exploit the information about the distribution of the road transport missions, concerning, e.g., a specific geographical area, to estimate the proportion of vehicles to be delivered to that market. The assessment would clearly be only preliminary but may provide important indications on the required volumes, with obvious advantages in the planning for the setting of the production lines. For example, Volvo aims at replacing its current fleet of trucks with battery-electric vehicles (BEVs), integrated with automated driving systems, by 2040. In this context, the sOC parametrized in this paper may be thought of as representative of the usage for the region of Västera Götaland, concerning heavy-duty rigid and tractor-semitrailer vehicles.

<sup>12</sup>This number of dOCs was empirically found already to yield sufficiently good estimates for the mean energy consumption and CO<sub>2</sub> emissions in [47] and [49].

<sup>13</sup><https://www.volvotrucks.se/sv-se/trucks/trucks/volvo-fh16/specifications/data-sheets.html>

By labeling the missions based on the internal GTA description, an optimal configuration may thus be found using *ad-hoc* algorithms for every combination of classes. Typically, this operation would involve the selection of a number of representative sOCs, the synthesis of dOCs, and their simulation in virtual environments, in conjunction with appropriate models for vehicle dynamics. Once the design has been optimized, the proportions of vehicles to be produced may subsequently be deduced from the distribution of missions, provided that the total volume has been estimated in advance, based, for example, on already placed orders or historical data. The spread in performance within the entire transport application may then be evaluated *a posteriori*, as done before in Section V-A. To simplify the problem, this paper only considers three parameters, namely the mission length, road grade, mean legal speed, and two vehicle configurations,<sup>14</sup> that is, rigid truck (RT) and tractor-semitrailer (TS). These are the same vehicle types as those in Appendix VII-A. Figure 15 illustrates the actual distribution of missions for the same population of 10000 road operations synthesized in Section IV. Owing to the premises above, optimal solutions for each combination of classes have been derived in [54] and [61] using a particle swarm optimization algorithm. Concerning the predominant combinations of classes, the corresponding proportions are listed in Table 8 for both variants, along with the number and type of electric motors (EMs), and battery packs (BPs). The interpretation for the EM and BP types in Table 8 is the same as in [54]. Exhaustive information about the optimal configurations and propulsion systems for the considered combinations of classes may be found in [61], where the results of a comprehensive sensitivity analysis are also presented.

## VI. DISCUSSION

The present paper has constituted a first attempt at a complete parametrization – concerning the road models – of an entire transport application, within the theoretical framework provided by the OC representation. The contribution is articulated in three main moments. The first coincides with the development of a set of stochastic models to describe the variation in usage amongst the road missions. The focus is on heavy-duty trucks, traveling long distances in the region of Västra Götaland, Sweden. The second is represented by the numerical analysis of the resulting OC, and its classification, according to two other simplified descriptions, namely the GTA and UFD systems. Finally, the third contribution resides in the exemplification of the potential of the OC format, relatively to the processes of vehicle design optimization, virtual testing, and production.

The present Section pauses to reflect upon the fundamental questions raised by the manuscript, along with their contextual interpretations and implications.

<sup>14</sup>It is also assumed that the volumes are known for both configurations, since the choice of a specific design may depend upon different factors other than the road properties, mission length, and mean legal speed.

## A. OC DEVELOPMENT AND PARAMETRIZATION, AND RELATED LIMITATIONS

In this paper, an OC has been developed and parametrized limited to long-haul heavy-duty vehicles. Concerning the parametrization process, data have been collected from trucks operating in the region of Västra Götaland, combining available measurements with *ad-hoc* estimation techniques. Moreover, weather and traffic models have not been considered in the present work, mainly because of the difficulties connected with the estimation of the related quantities. However, a complete OC description should opportunely be integrated with information regarding also weather and traffic, similar to what was done in [49]. Also, road parameters may be more accurately parametrized by collecting information from open-source databases, like OpenStreetMap,<sup>15</sup> in conjunction with GPS measurements. This possibility has been deliberately disregarded in the paper. Generally, it should be acknowledged that the proposed OC is only representative of a certain usage – i.e., that of heavy trucks traveling long distances –, whereas extensions may always be possible, as corroborated by the flourishing research dealing with the synthesis of conventional driving cycles.

In the very same context, it is worth recalling some observations from the introduction of this paper, to better highlight the advantages of the OC over the conventional driving cycle representation. More specifically, it appears that the OC format currently qualifies the only description capable of satisfactorily addressing all the three problems enounced in Section I, namely the representation, variation, and classification problems. Indeed, concerning the micro-trip and pattern classification methods [23], [26], [27], [28], [29], [30], [31], [34], [35], [36], [37], it is clear that the variation and classification aspects cannot be covered, since details about the operating environment are usually not included in the generated driving cycles. On the other hand, whereas some of the techniques listed in Table 1 – in particular, the modal construction and segment-based approaches [4], [18], [24], [25], [32], [33], [38] – permit the synthesis of driving cycles that account for the effect from the surroundings, these do not respond to the need for an easily interpretable description of the vehicle's usage. In fact, the influence of external factors like road grade, traffic density, and weather conditions is considered only implicitly by such driving cycles. In turn, this implies the impossibility of classifying a road transport mission based on the characteristics of the environment where it takes place. Moreover, even the segmentation and modal construction methods, which qualify perhaps amongst the most sophisticated techniques, do not allow for stochastic generation of equivalent driving cycles. As observed in [32], [33], and [39], this aspect becomes crucial when assessing variation in performance caused by fluctuations in the operating conditions, whilst also being a fundamental prerequisite to ensure robustness. From this perspective, whilst the approach pioneered in [32] and [33] seems to have addressed

<sup>15</sup><https://www.openstreetmap.org/#map=5/62.912/17.385>

**TABLE 8.** Optimal vehicles' design and proportion based on the mission distribution.

Configuration	Road grade	Speed	Distance	EM type	n. EMs	BP type	n. BPs	Proportion
<i>Rigid truck</i>								
RT	P-FLAT	MEDIUM/HIGH	REGIONAL	2	2	2	3	28%
RT	P-FLAT	MEDIUM/HIGH	L-DISTANCE	2	2	2	8	42%
RT	HILLY	MEDIUM/HIGH	REGIONAL	1	4	2	9	6%
RT	HILLY	MEDIUM/HIGH	L-DISTANCE	2	2	2	3	9%
<i>Tractor-semitrailer</i>								
TS	P-FLAT	MEDIUM/HIGH	REGIONAL/L-DISTANCE	1	4	2	11	70%
TS	HILLY	MEDIUM/HIGH	REGIONAL	1	6	2	13	6%
TS	HILLY	MEDIUM/HIGH	L-DISTANCE	1	4	2	5	9%

both the representation and variation problems adequately, again it does not seem compatible with the necessity for the classification of road missions.

Therefore, in spite of the present limitations, from a conceptual perspective, claiming that the OC format has come to maturity does not seem an incautious statement. The stochastic models introduced in this paper, in particular, have been motivated, wherever possible, based on their implications and ease of interpretation. In all the other cases, the principle of parsimony has been invoked, in line with the philosophy behind the whole OC machinery. It is also in the authors' intention to clarify that the models proposed here are nothing definitive, and other formulations may work equally well or even better on different datasets. With similar arguments, the format itself is not necessarily complete, and may always be improved or modified to accommodate needs that have not been identified so far. Whilst the possibility of building an all-embracing theory for the OC description might undoubtedly be seductive, such a colossal venture is not attempted in this paper. Besides, modularity and replaceability have been inspiring qualities since the very early conception of the OC framework.

## B. ON THE NOTION OF REPRESENTATIVENESS, AND ITS IMPLICATIONS

Since an OC describes a road transport mission directly concerning the operating environment, in this paper, the notion of representativeness has been discussed in relation to the characteristics of the surroundings. The fundamental question motivating the investigation has been formulated in the introductory Section I-B and phrased as: "What type of usage is a certain OC representative of?". For the problem under consideration, the interrogation should be clearly interpreted in the sense of road characteristics and mission length. In this paper, an answer to the previous question has been sought by considering two existing bird's-eye view descriptions: the GTA and UFD classification systems developed by Volvo and Scania. The differences and similarities between these two representations have highlighted the circumstantial nature of the notion of representativeness, and perhaps the need for a unified approach.

Adducing again the example of Volvo and Scania, the construction of their classification systems was presumably instrumental to the processes of optimal design and selection

of energy-efficient vehicles. In this context, the notion of representativeness is clearly evaluated based on an underlying metric, that is, the energy performance of road vehicles. Even so, the GTA and UFD labeled the OC parametrized in this paper differently. In Sections IV-B1 and IV-B2, this discrepancy has been ascribed to the different criteria specified by both descriptions concerning, for example, road grade and curvature. In this context, the disagreement between the two representations may reveal particularly problematic when considering the interaction with other entities (for example, for certification and urban planning purposes [62]). The wish is that vehicle manufacturers and governmental agencies could team to concord on the building of a unique description that would allow no room for misunderstanding. Scientists and academic scholars should also be aware of such a delicate aspect, and could actively participate in the process.

Such considerations about the notion of representativeness are not restricted to the scope of the operating cycle. Transient driving cycles should also be developed considering the internal systems adopted by vehicle manufacturers, to facilitate the integration with their development and production processes.

## C. CERTIFICATION AND PERFORMANCE EVALUATION

In this paper, two methods to assess the energy consumption and pollutant emissions of road vehicles have been investigated, based on the OC description. The first includes detailed information about the distribution of road missions, accounting for variation occurring over the entire population of transport operations. The second is based on a simplified approach and only considers a single representative sOC, selected opportunely and independently of the chosen classification system.

A specific method may be more suited for a certain application. For example, concerning the certification process, the simplified approach could certainly replace the conventional representation in terms of a driving cycle, if the focus is on mean quantities. This conclusion was supported in Section V by both the numerical and informal mathematical analyses conducted with respect to the pollutant emissions. Moreover, since a single sOC would still be able to reproduce variation in performance, information about the expected energy consumption and CO<sub>2</sub> emissions could be integrated more conveniently with additional indicators, e.g., variance or



standard deviation. However, considering the entire transport application would yield more accurate results in this case. In this context, it should be observed that the same result may also be achieved by considering the stochastic driving cycles proposed in [32], [33], and [39].

On the other hand, vehicle manufacturers may prefer testing vehicles considering the actual distribution of sOCs within a transport application, especially if log data are collected and stored in internal databases. To ensure consistency between the two approaches, however, a fundamental prerequisite would be the absence of any ambiguity in the translation between the corresponding descriptions. This happens again to relate to the notion of representativeness discussed previously.

#### **D. CHALLENGES AND ADVANTAGES RELATING TO OC-BASED VEHICLE DESIGN OPTIMIZATION, PRODUCTION PLANNING, AND PREDICTIVE MAINTENANCE**

The aspect of vehicle design optimization, albeit mentioned in the paper, has not been explored fully in the present study. In this context, compared to employing the conventional description in terms of driving cycles, the process of design optimization conducted using the OC format would be intuitively more robust, since it would explicitly account for variation in the operating conditions. Considering only a few missions, the computational effort would also be comparable to that of resorting to a common driving cycle (see discussion in [54]), with the additional advantage of the OC being a more accurate and reliable description of vehicle usage [63]. However, it is worth mentioning that the need for increasing robustness would imply simulating a larger population of dOCs. With these premises, a main drawback connected with using a large number of dOCs certainly consists in the fact that several road transport missions would need to be simulated whilst running an optimization routine, with high computational cost. Whereas this may be prohibitive with the current technologies, the imminent revolution of quantum computing could enable fast optimization in the next few years, possibly rendering the conventional driving-cycle-based approaches obsolete.

Concerning instead production planning and predictive maintenance, the problem is much simpler, since simulating a large number of dOCs is not particularly expensive.<sup>16</sup> Therefore, once the number of possible configurations is set, applying similar rationales as those discussed in the present paper becomes relatively straightforward when it comes to production planning. Analogously, predictive maintenance and control algorithms may be easily run and tested using the OC description, as done again in [54], where the number of stochastic models was however limited compared to the present study.

<sup>16</sup>Simulating a single dOC of 250 km takes around 1 min on a personal computer.

#### **VII. CONCLUSION**

In this paper, an OC for long-haul heavy-duty vehicles, complete with stochastic road models, has been developed using log data collected from trucks operating in the region of Västra Götaland, Sweden. The proposed description allows for the representation of energy usage in a realistic manner and captures variation between road transport missions by resorting to a statistical approach.

The representativeness of the fully parametrized OC has been evaluated and discussed concerning two existing classification systems in use by Volvo and Scania, respectively. According to both descriptions, the considered transport application is labeled as long distance, but other parameters are judged differently. A methodology has also been proposed to synthesize a single representative sOC, described in terms of mean parameters, starting from the complete population of road missions.

As an application of the developed OC, two different examples have been adduced in the paper. The first was aimed at assessing energy efficiency in simulation, considering both the actual distribution of road missions determining the application and the mean representative operation. Concerning specifically the Volvo FH13 employed in the study, it was demonstrated in simulation that the mean CO<sub>2</sub> emissions computed numerically according to the two different approaches were similar (amounting to 1708 and 1729 g km<sup>-1</sup>, respectively), whereas a large discrepancy could be observed with respect to the predicted values for the standard deviation (360 and 117 g km<sup>-1</sup>). The second example dealt instead with the process of optimal design and selection for future battery-electric vehicles, depending on the characteristics of the intended usage. The optimal proportion was characterized by a predominance of tractor-semitrailer trucks (70%), equipping 4 motors and 11 battery packs.

#### **ACKNOWLEDGMENT**

The authors gratefully acknowledge financial support from the COVER project (44929-1), funded by the Swedish Energy Agency and the Swedish Vehicle Research and Innovation Programme (FFI). In addition, they kindly acknowledge the support from the COVER project member Anders Jensen at Scania, who supplied information on the User Factor Description (UFD).

#### **DECLARATION OF INTEREST**

Declaration of interest: none.

#### **STATEMENT OF CONTRIBUTIONS**

**Luigi Romano:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft; **Michele Godio:** Data curation, Formal analysis; **Pär Johannesson:** Methodology, Supervision; **Fredrik Bruzelius:** Conceptualization, Supervision; **Toheed Ghandriz:** Data curation; **Benkt Jacobson:** Funding acquisition, Project administration, Supervision. All

the authors read and approved the original version of the manuscript.

## APPENDIX

### A. THE OCEAN DATASET

The stochastic models presented in this paper were parametrized using the OCEAN database. The database was created as part of the OCEAN project, and contains data from 33 different long-haul heavy-vehicle configurations, mainly operating in the region of Västtra Götaland, Sweden. The properties of each truck, including its identifier, maximum torque and axle configuration are listed in Table 9.

The database is organized into different directories. Each directory contains the data corresponding to every single truck (A1, A2, ..., D2). In each folder, the data are organized by log files, which correspond to the routes traveled by the trucks. In total, the database contains 1872 log files, each counting 34 signals, including GPS coordinates, yaw rate measurements, brake pedal position, et cetera. The number of files is quite heterogeneously distributed amongst the different vehicles and covers a period of time corresponding to a typical quarter.

The quantities required to fit the models presented in Section III were estimated from the available measurements using different techniques. For example, whilst the total distance could be deduced immediately from log data, the topography parameters needed to be recovered using statistical tools, like the Wafo package implemented in [64] and [65]. The road curviness was also inferred from the yaw rate, by postulating simple equilibrium equations relating the local curvature to the acceleration. Finally, the legal speed was estimated by filtering out the influence of the previous parameters from the resulting speed profiles, using *ad-hoc* algorithms. A viable alternative would be to retrieve the information about the speed signs directly from an open source database, like OpenStreetMap, using GPS signals.

### B. THE GTA AND UFD CLASSIFICATION SYSTEMS

The GTA and UFD classification systems developed by Volvo and Scania are conceived as high-level representations targeting individual road missions, as well as entire transport applications. They have been refined over the years during an iterative process, and are intended to facilitate the interaction with the customer during the selection and sales stages. To this end, they describe the operating environments using colloquial tones and statements, which may however be reformulated in terms of statistical indicators, including mathematical expectations and probabilities. Imposing limits and thresholds on these, the GTA and UFD systems build a discrete representation of the usage. Consequently, a road transport mission may be qualified by resorting to a countable number of labels, each of them relating to a specific operating class. More specifically, the formalization of both descriptions has been worked out in [50], where the analytical

expressions for the operating classes have been deduced for all the road models presented in this paper.

For the sake of brevity, the original formulation of each operating class is not discussed in this paper, whereas Table 10 lists the classes specified by both the GTA and UFD representations in mathematical form, with respect to the different stochastic models discussed in the paper. The quantities appearing in Table 10 represent the realizations of the composite variables introduced in Section IV-A, with an obvious convention for the notation. It is worth mentioning that the original classification systems are actually deficient in the mean hill length parameter and, concerning the topography, completely qualify a mission depending on the value of the road grade length ratio. Moreover, the GTA description does not include any indication about the mean speed. However, in this paper, the latter parameters have been added to allow for a fair comparison with the UFD system. In this context, the values for the speed limits have been inspired by the stochastic model for the road type. Finally, the value  $\kappa = 0.008 \text{ m}^{-1}$  appearing in both the expressions for the expected number of curves and the curviness length ratio corresponds to a speed reduction of nearly 20% when driving at  $70 \text{ km h}^{-1}$  [50].

### C. ANALYTICAL EXPRESSIONS FOR THE STATIONARY DISTRIBUTIONS OF ROAD TYPES AND SPEED SIGNS

This Section provides the analytical expressions for the stationary distributions for road types and speed signs, together with the corresponding random versions.

#### 1) STATIONARY DISTRIBUTIONS FOR ROAD TYPES

Starting with (3a) and (4), for  $n_r = 3$ , the stationary distribution  $\pi_R = \pi_R(\mathbf{P}_R, \mathbf{L}_R)$  may be derived in components as in (28), shown at the bottom of the next page. From (28), the stochastic stationary distribution  $\Pi_R = \Pi_R(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R)$ , defined over the entire population of missions, is thus given by, as in (29), shown at the bottom of the next page.

#### 2) STATIONARY DISTRIBUTIONS FOR SPEED SIGNS

For a given road type  $r_k$ , the stationary distribution  $\pi_{V|r_k}$  for the speed signs may be found from (17) and (18). For  $n_r = 3$ , the stationary distribution  $\pi_R = \pi_R(\mathbf{P}_R, \mathbf{L}_R)$  may be derived in components as in (30), shown at the bottom of the next page, and

$$\pi_{V1|r_k}(L_{V1|r_k}) = \frac{L_{V1|r_k}}{L_{V1|r_k} + L_{V2|r_k}}, \quad (31a)$$

$$\pi_{V2|r_k}(L_{V2|r_k}) = \frac{L_{V2|r_k}}{L_{V1|r_k} + L_{V2|r_k}}, \quad (31b)$$

for  $n_{V|r_k} = 2$ . Accordingly, the stochastic counterparts read as in (32), as shown at the bottom of page 31, and

$$\Pi_{V1|r_k}(\tilde{L}_{V1|r_k}) = \frac{\tilde{L}_{V1|r_k}}{\tilde{L}_{V1|r_k} + \tilde{L}_{V2|r_k}}, \quad (33a)$$

$$\Pi_{V2|r_k}(\tilde{L}_{V2|r_k}) = \frac{\tilde{L}_{V2|r_k}}{\tilde{L}_{V1|r_k} + \tilde{L}_{V2|r_k}}. \quad (33b)$$

**TABLE 9.** Truck vehicle configurations included in the OCEAN database.

Truck ID	Max torque (N m)	Axle config	Unit	Truck ID	Max torque (N m)	Axle config	Unit
A1	3550	RADD	Rigid	B3	3550	RADD	Rigid
A2	2900	RADT	Tractor	B4	?	?	?
A3	3150	RAD	Tractor	B5	?	?	?
A4	3150	RAPD	Tractor	B6	?	?	?
A5	2900	RAD	Tractor	B7	?	?	?
A6	3150	RADT	Tractor	B8	?	?	?
A7	3550	RAPD	Tractor	C1	2600	RADDT	Rigid
A8	2500	RADT	Tractor	C2	2150	RADT	Rigid
A9	2300	RAPD	Tractor	C3	3550	RADDT	Rigid
A10	2300	RAPD	Tractor	C4	?	?	?
A11	?	?	?	C5	?	?	?
A12	?	?	?	C6	?	?	?
A13	?	?	?	C7	?	?	?
A14	?	?	?	C8	?	?	?
A15	?	?	?	D1	2600	RADD	Rigid
B1	2600	RADD	Rigid	D2	?	?	?
B2	3550	RADDT	Rigid				

**D. APPROXIMATED ANALYSIS FOR VARIANCE**

The correlation between the composite variables introduced in Section IV, may be studied by resorting to approximated

analytical approaches. In particular, using the propagation of uncertainty technique, the covariance matrix  $\Sigma_{X_{\text{GTA}}X_{\text{GTA}}}$  for the GTA vector  $X_{\text{GTA}} = X_{\text{GTA}}(X_{\mathcal{R}_s})$  of composite random

$$\pi_{R1}(\mathbf{P}_R, \mathbf{L}_R) = \frac{L_{R1}(p_{R21}p_{R32} + p_{R31})}{L_{R1}(p_{R21}p_{R32} + p_{R31}) + L_{R2}(p_{R12}p_{R31} + p_{R32}) + L_{R3}(1 - p_{R12}p_{R21})}, \quad (28a)$$

$$\pi_{R2}(\mathbf{P}_R, \mathbf{L}_R) = \frac{L_{R2}(p_{R12}p_{R31} + p_{R32})}{L_{R1}(p_{R21}p_{R32} + p_{R31}) + L_{R2}(p_{R12}p_{R31} + p_{R32}) + L_{R3}(1 - p_{R12}p_{R21})}, \quad (28b)$$

$$\pi_{R3}(\mathbf{P}_R, \mathbf{L}_R) = \frac{L_{R3}(1 - p_{R12}p_{R21})}{L_{R1}(p_{R21}p_{R32} + p_{R31}) + L_{R2}(p_{R12}p_{R31} + p_{R32}) + L_{R3}(1 - p_{R12}p_{R21})}. \quad (28c)$$

$$\Pi_{R1}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) = \frac{\tilde{L}_{R1}(\tilde{p}_{R21}\tilde{p}_{R32} + \tilde{p}_{R31})}{\tilde{L}_{R1}(\tilde{p}_{R21}\tilde{p}_{R32} + \tilde{p}_{R31}) + \tilde{L}_{R2}(\tilde{p}_{R12}\tilde{p}_{R31} + \tilde{p}_{R32}) + \tilde{L}_{R3}(1 - \tilde{p}_{R12}\tilde{p}_{R21})}, \quad (29a)$$

$$\Pi_{R2}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) = \frac{\tilde{L}_{R2}(\tilde{p}_{R12}\tilde{p}_{R31} + \tilde{p}_{R32})}{\tilde{L}_{R1}(\tilde{p}_{R21}\tilde{p}_{R32} + \tilde{p}_{R31}) + \tilde{L}_{R2}(\tilde{p}_{R12}\tilde{p}_{R31} + \tilde{p}_{R32}) + \tilde{L}_{R3}(1 - \tilde{p}_{R12}\tilde{p}_{R21})}, \quad (29b)$$

$$\Pi_{R3}(\tilde{\mathbf{P}}_R, \tilde{\mathbf{L}}_R) = \frac{\tilde{L}_{R3}(1 - \tilde{p}_{R12}\tilde{p}_{R21})}{\tilde{L}_{R1}(\tilde{p}_{R21}\tilde{p}_{R32} + \tilde{p}_{R31}) + \tilde{L}_{R2}(\tilde{p}_{R12}\tilde{p}_{R31} + \tilde{p}_{R32}) + \tilde{L}_{R3}(1 - \tilde{p}_{R12}\tilde{p}_{R21})}. \quad (29c)$$

$$\pi_{V1|r_k}(\mathbf{P}_V|_{r_k}, \mathbf{L}_V|_{r_k}) = L_{V1|r_k}(p_{V21|r_k}p_{V32|r_k} + p_{V31|r_k}) \left[ L_{V1|r_k}(p_{V21|r_k}p_{V32|r_k} + p_{V31|r_k}) + L_{V2|r_k}(p_{V12|r_k}p_{V31|r_k} + p_{V32|r_k}) + L_{V3|r_k}(1 - p_{V12|r_k}p_{V21|r_k}) \right]^{-1}, \quad (30a)$$

$$\pi_{V2|r_k}(\mathbf{P}_V|_{r_k}, \mathbf{L}_V|_{r_k}) = L_{V2|r_k}(p_{V12|r_k}p_{V31|r_k} + p_{V32|r_k}) \left[ L_{V1|r_k}(p_{V21|r_k}p_{V32|r_k} + p_{V31|r_k}) + L_{V2|r_k}(p_{V12|r_k}p_{V31|r_k} + p_{V32|r_k}) + L_{V3|r_k}(1 - p_{V12|r_k}p_{V21|r_k}) \right]^{-1}, \quad (30b)$$

$$\pi_{V3|r_k}(\mathbf{P}_V|_{r_k}, \mathbf{L}_V|_{r_k}) = L_{V3|r_k}(1 - p_{V12|r_k}p_{V21|r_k}) \left[ L_{V1|r_k}(p_{V21|r_k}p_{V32|r_k} + p_{V31|r_k}) + L_{V2|r_k}(p_{V12|r_k}p_{V31|r_k} + p_{V32|r_k}) + L_{V3|r_k}(1 - p_{V12|r_k}p_{V21|r_k}) \right]^{-1}. \quad (30c)$$

**TABLE 10.** Operating classes according to the GTA and UFD classification system. The limits on the mean hill length  $L_h$  are expressed in m; those for the expected number of curves and stop signs  $\bar{n}_c$  and  $\bar{n}_s$ , respectively, in  $\text{km}^{-1}$ ; the thresholds on the mean legal speed  $\hat{v}$  are prescribed in  $\text{km h}^{-1}$ ; finally, the limits on the mission length  $L_m$  are specified in km.

Model	Variable	Operating Class	GTA Relationship	UFD Relationship
Topography	Road grade length ratio	FLAT P-FLAT HILLY V-HILLY	$p_y(3, \sigma_Y, \mathbf{P}_R, \mathbf{L}_R) > 0.98$ $p_y(6, \sigma_Y, \mathbf{P}_R, \mathbf{L}_R) > 0.98$ $p_y(9, \sigma_Y, \mathbf{P}_R, \mathbf{L}_R) > 0.98$ else	$p_y(2, \sigma_Y, \mathbf{P}_R, \mathbf{L}_R) \geq 0.8$ - $0.6 \leq p_y(2, \sigma_Y, \mathbf{P}_R, \mathbf{L}_R) < 0.8$ else
	Mean hill length	SHORT MEDIUM LONG V-LONG	$L_h(\mathbf{L}_h, \mathbf{P}_R, \mathbf{L}_R) \leq 500$ $500 < L_h(\mathbf{L}_h, \mathbf{P}_R, \mathbf{L}_R) \leq 1000$ $1000 < L_h(\mathbf{L}_h, \mathbf{P}_R, \mathbf{L}_R) \leq 1500$ else	$L_h(\mathbf{L}_h, \mathbf{P}_R, \mathbf{L}_R) \leq 500$ $500 < L_h(\mathbf{L}_h, \mathbf{P}_R, \mathbf{L}_R) \leq 1000$ $1000 < L_h(\mathbf{L}_h, \mathbf{P}_R, \mathbf{L}_R) \leq 1500$ else
Curviness	Number of curves	LOW HIGH	-	$\bar{n}_C(0.008, \lambda_C, \mu_C, \sigma_C, \mathbf{P}_R, \mathbf{L}_R) \leq 2$ else
	Curviness length ratio	LOW HIGH	$p_k(0.008, \lambda_C, \mu_C, \sigma_C, \mu_L, \sigma_L, \mathbf{P}_R, \mathbf{L}_R) > 0.8$ else	-
Stop signs	Number of stops	FLUID LIGHT RESIDENTIAL DENSE START&STOP VF-START&STOP	$\bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 0.2$ $0.2 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 0.5$ $0.5 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 1.5$ $1.5 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 2.5$ $2.5 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 5$ else	$\bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 0.2$ $0.2 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 0.5$ $0.5 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 1.5$ $1.5 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 2.5$ $2.5 < \bar{n}_s(\lambda_s, \mathbf{P}_R, \mathbf{L}_R) \leq 5$ else
		LOW MEDIUM HIGH V-HIGH	$\hat{v}(\mathbf{P}_{V r_1}, \dots, \mathbf{L}_{V r_1}, \dots, \mathbf{P}_R, \mathbf{L}_R) \leq 50$ $50 < \hat{v}(\mathbf{P}_{V r_1}, \dots, \mathbf{L}_{V r_1}, \dots, \mathbf{P}_R, \mathbf{L}_R) \leq 80$ else	$\hat{v}(\mathbf{P}_{V r_1}, \dots, \mathbf{L}_{V r_1}, \dots, \mathbf{P}_R, \mathbf{L}_R) \leq 50$ $50 < \hat{v}(\mathbf{P}_{V r_1}, \dots, \mathbf{L}_{V r_1}, \dots, \mathbf{P}_R, \mathbf{L}_R) \leq 70$ $70 < \hat{v}(\mathbf{P}_{V r_1}, \dots, \mathbf{L}_{V r_1}, \dots, \mathbf{P}_R, \mathbf{L}_R) \leq 90$ else
Mission length	Mission length	STOP&GO LOCAL REGIONAL L-DISTANCE	$L_m < 0.5$ $0.5 \leq L_m < 5$ $5 \leq L_m < 50$ else	$L_m < 0.5$ $0.5 \leq L_m < 5$ $5 \leq L_m < 50$ else

variables may be deduced approximately as

$$\Sigma_{X_{\text{GTA}} X_{\text{GTA}}} \approx \mathbf{J}_{X_{\text{GTA}}}(\mathbb{E}(X_{\mathcal{R}_s})) \Sigma_{X_{\mathcal{R}_s} X_{\mathcal{R}_s}} \mathbf{J}_{X_{\text{GTA}}}^T(\mathbb{E}(X_{\mathcal{R}_s})), \quad (34)$$

where  $\mathbf{J}_{X_{\text{GTA}}}$  is the Jacobian of the vector function  $X_{\text{GTA}}(\cdot)$  with respect to the set of simple sOC random variables  $X_{\mathcal{R}_s}$ , and  $\Sigma_{X_{\mathcal{R}_s} X_{\mathcal{R}_s}}$  denotes the covariance matrix of the simple variables. The expression for the UFD system is analogous

$$\Sigma_{X_{\text{UFD}} X_{\text{UFD}}} \approx \mathbf{J}_{X_{\text{UFD}}}(\mathbb{E}(X_{\mathcal{R}_s})) \Sigma_{X_{\mathcal{R}_s} X_{\mathcal{R}_s}} \mathbf{J}_{X_{\text{UFD}}}^T(\mathbb{E}(X_{\mathcal{R}_s})), \quad (35)$$

with a similar notation for the Jacobian matrix. It should be observed that, whilst the Jacobians appearing in (34) and (35) are different, the matrix  $\Sigma_{X_{\mathcal{R}_s} X_{\mathcal{R}_s}}$  is the same for both relationships. In particular, the matrix  $\Sigma_{X_{\mathcal{R}_s} X_{\mathcal{R}_s}}$  may be put in diagonal form as in (36), shown at the bottom of the next page,

where all the matrices are diagonal by model construction, except those for the stochastic transition matrices for the road types and conditional speed signs, which may be in turn decomposed in diagonal form as

$$\Sigma_{\tilde{\mathbf{P}}_R \tilde{\mathbf{P}}_R} = \begin{bmatrix} \Sigma_{\tilde{\mathbf{P}}_{R1} \tilde{\mathbf{P}}_{R1}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Sigma_{\tilde{\mathbf{P}}_{Rn_r} \tilde{\mathbf{P}}_{Rn_r}} \end{bmatrix}, \quad (37)$$

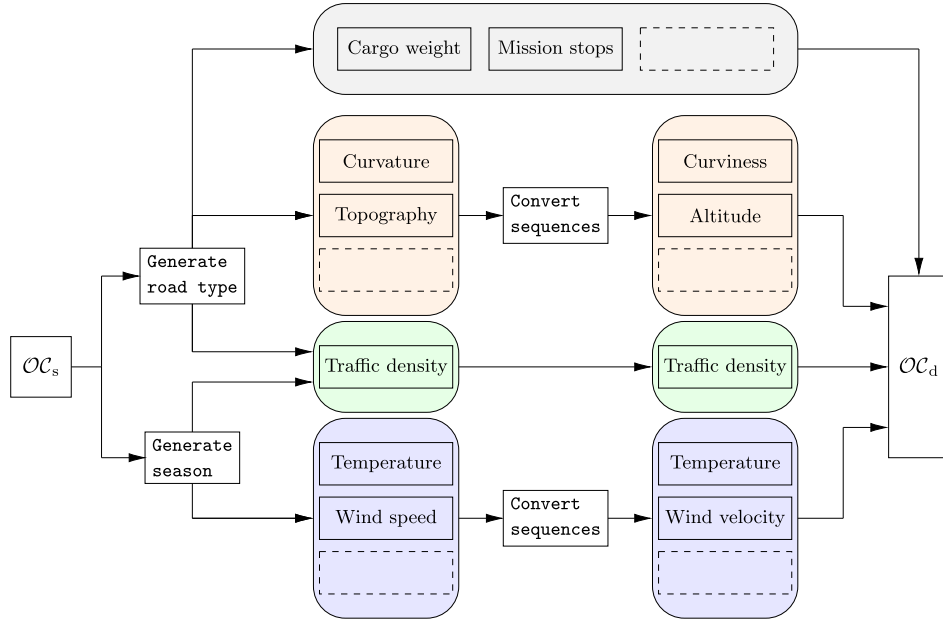
and

$$\Sigma_{\tilde{\mathbf{P}}_{V|r_k} \tilde{\mathbf{P}}_{V|r_k}} = \begin{bmatrix} \Sigma_{\tilde{\mathbf{P}}_{V1|r_k} \tilde{\mathbf{P}}_{V1|r_k}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Sigma_{\tilde{\mathbf{P}}_{Vn_v|r_k} \tilde{\mathbf{P}}_{Vn_v|r_k}} \end{bmatrix}, \quad k = 1, \dots, n_r, \quad (38)$$

$$\Pi_{V1|r_k}(\tilde{\mathbf{P}}_{V|r_k}, \tilde{\mathbf{L}}_{V|r_k}) = \tilde{L}_{V1|r_k}(\tilde{p}_{V21|r_k} \tilde{p}_{V32|r_k} + \tilde{p}_{V31|r_k}) \left[ \tilde{L}_{V1|r_k}(\tilde{p}_{V21|r_k} \tilde{p}_{V32|r_k} + \tilde{p}_{V31|r_k}) + \tilde{L}_{V2|r_k}(\tilde{p}_{V12|r_k} \tilde{p}_{V31|r_k} + \tilde{p}_{V32|r_k}) + \tilde{L}_{V3|r_k}(1 - \tilde{p}_{V12|r_k} \tilde{p}_{V21|r_k}) \right]^{-1}, \quad (32a)$$

$$\Pi_{V2|r_k}(\tilde{\mathbf{P}}_{V|r_k}, \tilde{\mathbf{L}}_{V|r_k}) = \tilde{L}_{V2|r_k}(\tilde{p}_{V12|r_k} \tilde{p}_{V31|r_k} + \tilde{p}_{V32|r_k}) \left[ \tilde{L}_{V1|r_k}(\tilde{p}_{V21|r_k} \tilde{p}_{V32|r_k} + \tilde{p}_{V31|r_k}) + \tilde{L}_{V2|r_k}(\tilde{p}_{V12|r_k} \tilde{p}_{V31|r_k} + \tilde{p}_{V32|r_k}) + \tilde{L}_{V3|r_k}(1 - \tilde{p}_{V12|r_k} \tilde{p}_{V21|r_k}) \right]^{-1}, \quad (32b)$$

$$\Pi_{V3|r_k}(\tilde{\mathbf{P}}_{V|r_k}, \tilde{\mathbf{L}}_{V|r_k}) = \tilde{L}_{V3|r_k}(1 - \tilde{p}_{V12|r_k} \tilde{p}_{V21|r_k}) \left[ \tilde{L}_{V1|r_k}(\tilde{p}_{V21|r_k} \tilde{p}_{V32|r_k} + \tilde{p}_{V31|r_k}) + \tilde{L}_{V2|r_k}(\tilde{p}_{V12|r_k} \tilde{p}_{V31|r_k} + \tilde{p}_{V32|r_k}) + \tilde{L}_{V3|r_k}(1 - \tilde{p}_{V12|r_k} \tilde{p}_{V21|r_k}) \right]^{-1}. \quad (32c)$$



**FIGURE 16.** Generation process of a deterministic operating cycle (dOC) from a stochastic one (sOC). All the secondary models are generated stochastically departing from the sequences of primary ones (in this paper, only the road properties are considered). A conversion is needed between the sOC and dOC formalisms.

being  $\Sigma_{\tilde{\mathbf{p}}_{Ri}\tilde{\mathbf{p}}_{Ri}}, i = 1, \dots, n_r$  and  $\Sigma_{\tilde{\mathbf{p}}_{Vil|r_k}\tilde{\mathbf{p}}_{Vil|r_k}}, i = 1, \dots, n_{v|r_k}, k = 1, \dots, n_r$  the covariance matrices for the Dirichlet distributions of Sections III-A and III-B4, respectively. Since the diagonal elements are always zero, i.e.,  $\tilde{p}_{Rii} = 0$  for the road types and  $\tilde{p}_{Vil|r_k} = 0$  for the speed signs, the matrices  $\Sigma_{\tilde{\mathbf{p}}_{Ri}\tilde{\mathbf{p}}_{Ri}}$  and  $\Sigma_{\tilde{\mathbf{p}}_{Vil|r_k}\tilde{\mathbf{p}}_{Vil|r_k}}$  have no elements on the principal diagonal. The nondiagonal entries may be instead evaluated as

$$\mathbb{Cov}(\tilde{p}_{Ril}, \tilde{p}_{Rim}) = -\frac{\alpha_{Ril}\alpha_{Rim}}{\sum_{l \neq m} \alpha_{Ril}(\alpha_{Ril} + \sum_{l \neq m} \alpha_{Ril})}, \quad i = 1, \dots, n_r, \quad (39)$$

and

$$\mathbb{Cov}(\tilde{p}_{Vil|r_k}, \tilde{p}_{Vim|r_k}) = -\frac{\alpha_{Vil|r_k}\alpha_{Vim|r_k}}{\sum_{l \neq m} \alpha_{Vil|r_k}(\alpha_{Vil|r_k} + \sum_{l \neq m} \alpha_{Vil|r_k})}, \quad i = 1, \dots, n_{v|r_k}, k = 1, \dots, n_r. \quad (40)$$

According to (36), (37) and (38), only the stochastic variables relating to the same Dirichlet distributions are

correlated. On the other hand, the composite variables are correlated via the stationary distributions for the road types. In (34) and (34), this correlation is accounted for by the Jacobian matrices  $\mathbf{J}_{X_{GTA}}$  and  $\mathbf{J}_{X_{UFD}}$ , respectively. It should be remarked that assuming a single road type  $r_1$  implies that the composite variables in the UFD vector  $\mathbf{X}_{UFD}$  are all independent, and thus uncorrelated. This happens because none of the original simple random variables appears in more than one scalar function in (27). The same is not true for the GTA vector  $\mathbf{X}_{GTA}$ , since its first three components would all be functions of the stochastic standard deviation  $\Sigma_Y \equiv \Sigma_{Y|r_1}$ .

#### E. IMPLEMENTATION AND COMPUTATIONAL DETAILS

The present Section describes the practical implementation of a dOC, starting from the corresponding sOC description, along with details concerning the simulation of longitudinal vehicle dynamics in VehProp. In particular, Figure 16 is a schematic illustration of the typical workflow needed to synthesize a reference dOC starting from an equivalent sOC. First, the primary models are generated over a specific mis-

$$\Sigma_{X_{\mathcal{R}_s} X_{\mathcal{R}_s}} = \begin{bmatrix} \Sigma_{\tilde{\mathbf{p}}_R \tilde{\mathbf{p}}_R} & \dots & \dots & \dots & \dots & \dots & \mathbf{0} \\ \vdots & \Sigma_{\tilde{\mathbf{L}}_R \tilde{\mathbf{L}}_R} & & & & & \vdots \\ \vdots & & \ddots & & & & \vdots \\ \vdots & & & \Sigma_{\tilde{\mathbf{p}}_{V|r_1} \tilde{\mathbf{p}}_{V|r_1}} & & & \vdots \\ \vdots & & & & \ddots & & \vdots \\ \vdots & & & & & \Sigma_{\tilde{\mathbf{L}}_{V|r_{n_r}} \tilde{\mathbf{L}}_{V|r_{n_r}}} & \vdots \\ \mathbf{0} & \dots & \dots & \dots & \dots & \dots & \mathbb{V}\text{ar}(\tilde{\mathbf{L}}_m) \end{bmatrix}, \quad (36)$$



sion distance, which may either be prescribed or simulated using an opportune distribution, as done in the present paper. The primary models are also simulated simultaneously, since they do not interact explicitly, owing to the principles of independence and parsimony upon which the OC is built. The secondary sOC models are then derived from the primary ones. More specifically, the simulation of road properties may be carried out using the *ad-hoc* WAFO package implemented in MATLAB® [64], [65]. The sequences obtained using this procedure need to be converted into the dOC formalism. For example, curvature and topography are translated into curviness and altitude.<sup>17</sup> From the signed curvature, the actual road profile and the tangent vector to the trajectory are also deduced numerically using Fresnel integrals [49]. The dOC parameters, plus their location in either space or time (or both for the traffic density), are then encoded in the dOC description and tabulated. Finally, a longitudinal vehicle model is simulated in the VehProp environment, consisting of a standard set of equations implemented in MATLAB/Simulink®, similar to those that may be found in any reference textbook [7]. At each time step, the dOC parameters are used as input to the governing equations of motion, and intermediate values are calculated by using a suitable interpolation law (as explained more extensively in [9]). Concerning a dOC of 250 km, a simulation in VehProp typically takes 1 min on a standard personal computer to be completed.

## REFERENCES

- [1] D. W. Wyatt, H. Li, and J. E. Tate, "The impact of road grade on carbon dioxide (CO<sub>2</sub>) emission of a passenger vehicle in real-world driving," *Transp. Res. D, Transp. Environ.*, vol. 32, pp. 160–170, Oct. 2014, doi: [10.1016/j.trd.2014.07.015](#).
- [2] K. M. Sentoff, L. Aultman-Hall, and B. A. Holmén, "Implications of driving style and road grade for accurate vehicle activity data and emissions estimates," *Transp. Res. D, Transp. Environ.*, vol. 35, pp. 175–188, Mar. 2015, doi: [10.1016/j.trd.2014.11.021](#).
- [3] D. Llopis-Castelló, A. M. Pérez-Zuriaga, F. J. Camacho-Torregrosa, and A. García, "Impact of horizontal geometric design of two-lane rural roads on vehicle CO<sub>2</sub> emissions," *Transp. Res. D, Transp. Environ.*, vol. 59, pp. 46–57, Mar. 2018, doi: [10.1016/j.trd.2017.12.020](#).
- [4] A. Sciarretta, *Energy-Efficient Driving of Road Vehicles*, 1st ed. Cham, Switzerland: Springer, 2020. [Online]. Available: [10.1007/978-3-030-24127-8](#).
- [5] R. C. Ramírez, I. Adin, J. Goya, U. Alvarado, A. Brazalez, and J. Mendizabal, "Freight train in the age of self-driving vehicles. A taxonomy review," *IEEE Access*, vol. 10, pp. 9750–9762, 2022, doi: [10.1109/ACCESS.2022.3144602](#).
- [6] B. Hernández, A. Alkayes, E. Azar, and A. T. Mayyas, "Mathematical model for the placement of hydrogen refueling stations to support future fuel cell trucks," *IEEE Access*, vol. 9, pp. 148118–148131, 2021, doi: [10.1109/ACCESS.2021.3123901](#).
- [7] L. Guzzella and A. Sciarretta, *Vehicle Propulsion Systems: Introduction to Modeling and Optimization*, 3rd ed. Berlin, Germany: Springer, 2013.
- [8] M. Åsbögård, L. Johansson, D. Angervall, and P. Johansson, "Improving system design of a hybrid powertrain using stochastic drive cycles and dynamic programming," in *Proc. SAE World Congr. Exhib.*, Detroit, MI, USA, 2007, pp. 1–9.
- [9] P. Pettersson, "Operating cycle representation for road vehicles," Ph.D. dissertation, Dept. Mech. Maritime Sci., Chalmers Univ. Technol., Göteborg, Sweden, 2019.
- [10] L. Romano, "Mathematical modelling of operating cycles for road vehicles," Licentiate thesis, Dept. Mech. Maritime Sci., Chalmers Univ. Technol., Göteborg, Sweden, 2021.
- [11] Z. Zou, S. Davis, K. Beaty, M. O'Keefe, T. Hendricks, R. Rehn, S. Weissner, and V. K. Sharma, "A new composite drive cycle for heavy-duty hybrid electric class 4–6 vehicles," in *Proc. SAE Tech. Paper Ser.*, Mar. 2004, pp. 1–10, doi: [10.4271/2004-01-1052](#).
- [12] M. Montazeri-Gh and M. Naghizadeh, "Development of car drive cycle for simulation of emissions and fuel economy," in *Proc. 15th Eur. Simulation Symp.*, A. Verbraesk and V. Hlupic, Eds. SCS European Council/SCS Europe BVBA, 2003.
- [13] P. Nyberg, "Evaluation, generation and transformation of driving cycles," Ph.D. dissertation, Dept. Elect. Eng., Linköping Univ. Electron. Press, Linköping, Sweden, 2015.
- [14] P. Nyberg, E. Frisk, and L. Nielsen, "Using real-world driving databases to generate driving cycles with equivalence properties," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 4095–4105, Jun. 2016, doi: [10.1109/TVT.2015.2502069](#).
- [15] P. Nyberg, E. Frisk, and L. Nielsen, "Driving cycle equivalence and transformation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 1963–1974, Mar. 2017, doi: [10.1109/TVT.2016.2582079](#).
- [16] A. Ashtari, E. Bibeau, and S. Shahidinejad, "Using large driving record samples and a stochastic approach for real-world driving cycle construction: Winnipeg driving cycle," *Transp. Sci.*, vol. 48, no. 2, pp. 170–183, May 2014, doi: [10.1287/trsc.1120.0447](#).
- [17] E. Tazelaar, J. Bruinsma, B. Veenhuizen, and P. van den Bosch, "Driving cycle characterization and generation, for design and control of fuel cell buses," *World Electric Vehicle J.*, vol. 3, no. 4, pp. 812–819, Dec. 2009, doi: [10.3390/wevj3040812](#).
- [18] J. Lin and D. A. Niemeier, "An exploratory analysis comparing a stochastic driving cycle to California's regulatory cycle," *Atmos. Environ.*, vol. 36, no. 38, pp. 5759–5770, 2002, doi: [10.1016/S1352-2310\(02\)00695-7](#).
- [19] J. Brady and M. O'Mahony, "Development of a driving cycle to evaluate the energy economy of electric vehicles in urban areas," *Appl. Energy*, vol. 177, pp. 165–178, Sep. 2016, doi: [10.1016/j.apenergy.2016.05.094](#).
- [20] T.-K. Lee and Z. S. Filipi, "Synthesis of real-world driving cycles using stochastic process and statistical methodology," *Int. J. Vehicle Des.*, vol. 57, no. 1, pp. 17–36, 2011. [Online]. Available: <http://www.inderscience.com/offer.php?id=43590>
- [21] T. Lee, B. Adornato, and Z. S. Filipi, "Synthesis of real-world driving cycles and their use for estimating PHEV energy consumption and charging opportunities: Case study for Midwest/US," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4153–4163, Nov. 2011, doi: [10.1109/TVT.2011.2168251](#).
- [22] G. Amirjamshidi and M. J. Roorda, "Development of simulated driving cycles for light, medium, and heavy duty trucks: Case of the Toronto waterfront area," *Transp. Res. D, Transp. Environ.*, vol. 34, pp. 255–266, Jan. 2015, doi: [10.1016/j.trd.2014.11.010](#).
- [23] S. H. Kamble, T. V. Mathew, and G. K. Sharma, "Development of real-world driving cycle: Case study of Pune, India," *Transp. Res. D, Transp. Environ.*, vol. 14, no. 2, pp. 132–140, Mar. 2009, doi: [10.1016/j.trd.2008.11.008](#).
- [24] E. Silvas, "Integrated optimal design for hybrid electric vehicles," Ph.D. dissertation, Dept. Mech. Eng., Technische Universiteit Eindhoven, Eindhoven, The Netherlands, 2015.
- [25] E. Silvas, K. Hereijgers, H. Peng, T. Hofman, and M. Steinbuch, "Synthesis of realistic driving cycles with high accuracy and computational speed, including slope information," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 4118–4128, Jun. 2016, doi: [10.1109/TVT.2016.2546338](#).
- [26] X. Liu, J. Ma, X. Zhao, J. Du, and Y. Xiong, "Study on driving cycle synthesis method for city buses considering random passenger load," *J. Adv. Transp.*, vol. 2020, pp. 1–21, Feb. 2020, doi: [10.1155/2020/3871703](#).
- [27] S. Shi, S. Wei, H. Kui, L. Liu, C. Huang, and M. Liu, "Improvements of the design method of transient driving cycle for passenger car," in *Proc. IEEE Vehicle Power Propuls. Conf.*, Sep. 2009, pp. 1–7, doi: [10.1109/VPPC.2009.5289594](#).
- [28] L. Liu, C. Huang, M. Liu, and S. Shi, "Study on the combined design method of transient driving cycles for passenger car in Changchun," in *Proc. IEEE Vehicle Power Propuls. Conf.*, Sep. 2008, pp. 1–14. [Online]. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4677594>
- [29] W. T. Hung, H. Y. Tong, C. P. Lee, K. Ha, and L. Y. Pao, "Development of a practical driving cycle construction methodology: A case study in Hong Kong," *Transp. Res. D, Transp. Environ.*, vol. 12, no. 2, pp. 115–128, Mar. 2007, doi: [10.1016/j.trd.2007.01.002](#).

<sup>17</sup>The sequence  $\{z_k\}_{k \in \mathbb{N}}$  is constructed according to  $z_{k+1} = z_k + \frac{y_k}{100} L_s = \sum_{i=0}^k y_i + z_0$ , being  $\{y_k\}_{k \in \mathbb{N}}$  the sequence of realisations for the variable  $Y_k$  describing the road grade.

- [30] K. S. Nesamani and K. P. Subramanian, "Development of a driving cycle for intra-city buses in Chennai, India," *Atmos. Environ.*, vol. 45, no. 31, pp. 5469–5476, Oct. 2011, doi: [10.1016/j.atmosenv.2011.06.067](https://doi.org/10.1016/j.atmosenv.2011.06.067).
- [31] S.-H. Ho, Y.-D. Wong, and V. W.-C. Chang, "Developing Singapore driving cycle for passenger cars to estimate fuel consumption and vehicular emissions," *Atmos. Environ.*, vol. 97, pp. 353–362, Nov. 2014, doi: [10.1016/j.atmosenv.2014.08.042](https://doi.org/10.1016/j.atmosenv.2014.08.042).
- [32] K. Kivekäs, J. Vepsäläinen, and K. Tammi, "Stochastic driving cycle synthesis for analyzing the energy consumption of a battery electric bus," *IEEE Access*, vol. 6, pp. 55586–55598, 2018, doi: [10.1109/ACCESS.2018.2871574](https://doi.org/10.1109/ACCESS.2018.2871574).
- [33] K. Kivekäs, J. Vepsäläinen, K. Tammi, and J. Anttila, "Influence of driving cycle uncertainty on electric city bus energy consumption," in *Proc. IEEE Vehicle Power Propuls. Conf. (VPPC)*, Dec. 2017, pp. 1–5, doi: [10.1109/VPPC.2017.8331014](https://doi.org/10.1109/VPPC.2017.8331014).
- [34] F. Guo and F. Zhang, "A study of driving cycle for electric cars on Beijing urban and suburban roads," in *Proc. IEEE Int. Conf. Power Renew. Energy (ICPRE)*, Oct. 2016, pp. 319–322, doi: [10.1109/ICPRE.2016.7871224](https://doi.org/10.1109/ICPRE.2016.7871224).
- [35] S. Ou, Y. Zhou, J. Lian, P. Jia, and B. Tian, "Development of hybrid city bus's driving cycle," in *Proc. Int. Conf. Electric Inf. Control Eng.*, Apr. 2011, pp. 2112–2116, doi: [10.1109/ICEICE.2011.5777149](https://doi.org/10.1109/ICEICE.2011.5777149).
- [36] L. Berzi, M. Delogu, and M. Pierini, "Development of driving cycles for electric vehicles in the context of the city of Florence," *Transp. Res. D, Transp. Environ.*, vol. 47, pp. 299–322, Aug. 2016, doi: [10.1016/j.trd.2016.05.010](https://doi.org/10.1016/j.trd.2016.05.010).
- [37] Z. Jing, G. Wang, S. Zhang, and C. Qiu, "Building Tianjin driving cycle based on linear discriminant analysis," *Transp. Res. D, Transp. Environ.*, vol. 53, pp. 78–87, Jun. 2017, doi: [10.1016/j.trd.2017.04.005](https://doi.org/10.1016/j.trd.2017.04.005).
- [38] S. Shi, N. Lin, Y. Zhang, C. Huang, L. Liu, B. Lu, and J. Cheng, "Research on Markov property analysis of driving cycle," in *Proc. IEEE Vehicle Power Propuls. Conf. (VPPC)*, Oct. 2013, pp. 1–5, doi: [10.1109/VPPC.2013.6671737](https://doi.org/10.1109/VPPC.2013.6671737).
- [39] J. Vepsäläinen, K. Kivekäs, K. Otto, A. Lajunen, and K. Tammi, "Development and validation of energy demand uncertainty model for electric city buses," *Transp. Res. D, Transp. Environ.*, vol. 63, pp. 347–361, Aug. 2018, doi: [10.1016/j.trd.2018.06.004](https://doi.org/10.1016/j.trd.2018.06.004).
- [40] P. Johannesson, K. Podgórski, I. Rychlik, and N. Shariati, "AR(1) time series with autoregressive gamma variance for road topography modeling," *Probabilistic Eng. Mech.*, vol. 43, pp. 106–116, Jan. 2016, doi: [10.1016/j.probengmech.2015.12.006](https://doi.org/10.1016/j.probengmech.2015.12.006).
- [41] P. Johannesson, K. Podgórski, and I. Rychlik, "Laplace distribution models for road topography and roughness," *Int. J. Vehicle Perform.*, vol. 3, no. 3, p. 224, 2017. [Online]. Available: <http://www.inderscience.com/offer.php?id=85032>
- [42] A. Donkers, D. Yang, and M. Viktorovic, "Influence of driving style, infrastructure, weather and traffic on electric vehicle performance," *Transp. Res. D, Transp. Environ.*, vol. 88, Nov. 2020, Art. no. 102569, doi: [10.1016/j.trd.2020.102569](https://doi.org/10.1016/j.trd.2020.102569).
- [43] D. Akin, V. P. Sisiopiku, and A. Skabardonis, "Impacts of weather on traffic flow characteristics of urban freeways in Istanbul," *Proc. Social Behav. Sci.*, vol. 16, pp. 88–89, Jan. 2011, doi: [10.1016/j.sbspro.2011.04.432](https://doi.org/10.1016/j.sbspro.2011.04.432).
- [44] E. Hooper, L. Chapman, and A. Quinn, "The impact of precipitation on speed-flow relationships along a UK motorway corridor," *Theor. Appl. Climatol.*, vol. 117, nos. 1–2, pp. 303–316, Jul. 2014, doi: [10.1007/s00704-013-0999-5](https://doi.org/10.1007/s00704-013-0999-5).
- [45] O. Travasset-Baro, M. Rosas-Casals, and E. Jover, "Transport energy consumption in mountainous roads. A comparative case study for internal combustion engines and electric vehicles in Andorra," *Transp. Res. D, Transp. Environ.*, vol. 34, pp. 16–26, Jan. 2015, doi: [10.1016/j.trd.2014.09.006](https://doi.org/10.1016/j.trd.2014.09.006).
- [46] B. Y. Boroujeni and H. C. Frey, "Road grade quantification based on global positioning system data obtained from real-world vehicle fuel use and emissions measurements," *Atmos. Environ.*, vol. 85, pp. 179–186, Mar. 2014, doi: [10.1016/j.atmosenv.2013.12.025](https://doi.org/10.1016/j.atmosenv.2013.12.025).
- [47] P. Pettersson, P. Johannesson, B. Jacobson, F. Bruzelius, L. Fast, and S. Berglund, "A statistical operating cycle description for prediction of road vehicles' energy consumption," *Transp. Res. D, Transp. Environ.*, vol. 73, pp. 205–229, Aug. 2019, doi: [10.1016/j.trd.2019.07.006](https://doi.org/10.1016/j.trd.2019.07.006).
- [48] P. Pettersson, S. Berglund, B. Jacobson, L. Fast, P. Johannesson, and F. Santandrea, "A proposal for an operating cycle description format for road transport missions," *Eur. Transp. Res. Rev.*, vol. 10, no. 2, pp. 1–19, Jun. 2018, doi: [10.1186/s12544-018-0298-4](https://doi.org/10.1186/s12544-018-0298-4).
- [49] L. Romano, P. Johannesson, F. Bruzelius, and B. Jacobson, "An enhanced stochastic operating cycle description including weather and traffic models," *Transp. Res. D, Transp. Environ.*, vol. 97, Aug. 2021, Art. no. 102878, doi: [10.1016/j.trd.2021.102878](https://doi.org/10.1016/j.trd.2021.102878).
- [50] L. Romano, P. Johannesson, E. Nordström, F. Bruzelius, R. Andersson, and B. Jacobson, "A classification method of road transport missions and applications using the operating cycle format," *IEEE Access*, vol. 10, pp. 73087–73121, 2022.
- [51] S. Edlund and P. O. Fryk, "The right truck for the job with global truck application descriptions," SAE Tech. Paper 2004-01-2645, Oct. 2004, doi: [10.4271/2004-01-2645](https://doi.org/10.4271/2004-01-2645).
- [52] T. Ghandriz, J. Hellgren, M. Islam, L. Laine, and B. Jacobson, "Optimization based design of heterogeneous truck fleet and electric propulsion," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Rio de Janeiro, Brazil, Nov. 2016, pp. 328–335.
- [53] T. Ghandriz, "Transportation mission based optimization of heavy vehicle fleets including propulsion tailoring," Licentiate thesis, Dept. Mech. Maritime Sci., Chalmers Univ. Technol., Göteborg, Sweden, 2018.
- [54] T. Ghandriz, B. Jacobson, L. Laine, and J. Hellgren, "Impact of automated driving systems on road freight transport and electrified propulsion of heavy vehicles," *Transp. Res. C, Emerg. Technol.*, vol. 115, Jun. 2020, Art. no. 102610, doi: [10.1016/j.trc.2020.102610](https://doi.org/10.1016/j.trc.2020.102610).
- [55] L. Ntziachristos and Z. Samaras, "EMEP/EEA air pollutant emission inventory guidebook 2016 part B.1.A.3.B.I-IV road transport 2018," Eur. Environ. Agency, Chalmers Univ. Technol., Copenhagen, Denmark, 2018.
- [56] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 4th ed. Oxford, U.K.: Oxford Univ. Press, 2020.
- [57] H. P. Hsu, *Schaum's Outline of Probability, Random Variables, and Random Processes*, 4th ed. New York, NY, USA: McGraw-Hill Education, 2020.
- [58] G. Box, G. Jenkins, G. Reinsel, G. Ljung, and G. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. San Francisco, CA, USA: Wiley, 2015.
- [59] M. Karlsson, "Load modelling for fatigue assessment of vehicles—A statistical approach," Ph.D. dissertation, Dept. Math. Sci., Chalmers Univ. Technol., Göteborg, Sweden, 2007.
- [60] P. Plötz, N. Jakobsson, and F. Sprei, "On the distribution of individual daily driving distances," *Transp. Res. B, Methodol.*, vol. 101, pp. 213–227, Jul. 2017.
- [61] T. Ghandriz, B. Jacobson, L. Laine, and J. Hellgren, "Optimization data on total cost of ownership for conventional and battery electric heavy vehicles driven by humans and by automated driving systems," *Data Brief*, vol. 30, Jun. 2020, Art. no. 105566, doi: [10.1016/j.dib.2020.105566](https://doi.org/10.1016/j.dib.2020.105566).
- [62] M. Owais and A. Alshehri, "Pareto optimal path generation algorithm in stochastic transportation networks," *IEEE Access*, vol. 8, pp. 58970–58981, 2020, doi: [10.1109/ACCESS.2020.2983047](https://doi.org/10.1109/ACCESS.2020.2983047).
- [63] P. Pettersson, B. Jacobson, F. Bruzelius, P. Johannesson, and L. Fast, "Intrinsic differences between backward and forward vehicle simulation models," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 14292–14299, 2020, doi: [10.1016/j.ifacol.2020.12.1368](https://doi.org/10.1016/j.ifacol.2020.12.1368).
- [64] P. A. Brodtkorb, P. Johannesson, G. Lindgren, I. Rychlik, J. Rydén, and E. Sjö, "WAFO—A MATLAB toolbox for the analysis of random waves and loads," in *Proc. 10th Int. Offshore Polar Eng. Conf. (ISOPE)*, Seattle, WA, USA, vol. 3, 2000, pp. 343–350.
- [65] *WAFO—A MATLAB Toolbox for Analysis of Random Waves and Loads—A Tutorial*, WAFO-Group, Math. Statist., Centre Math. Sci., Lund Univ., Lund, Sweden, 2000. [Online]. Available: <http://www.maths.lth.se/matstat/wafo>



**LUIGI ROMANO** received the M.Sc. degree in mechanical engineering from the University of Naples Federico II, Italy, in 2018. He is currently pursuing the Ph.D. degree with the Vehicle Dynamics Group, Chalmers University of Technology, Gothenburg, Sweden, within the COVER Project. After graduation, he was an Independent Consultant on the Feel-Tire Unina Project, developing model-based estimators for intelligent tire technologies. His research interests include sustainable transportation systems, vehicle state estimation and control, and tire dynamics and modeling. In 2016, he was awarded the Prize "Buon Compleanno Federico II" as the Best Mechanical Engineering Student.



**MICHELE GODIO** received the double master's degree in civil and structural engineering from the Milano's Polytechnic Institute and Paris' Ponts et Chaussées University, France, and the master's degree (advanced research) in mechanics of materials and the Ph.D. degree in mechanics of materials and structures from Paris' Ponts et Chaussées. During his studies, he spent a two-month internship in a design office, where he carried out the seismic analysis of industrial structures. In October 2019, he joined the Department of Applied Mechanics, RISE Research Institute of Sweden, as a Researcher, after a four-year postdoctoral stay with the Earthquake Engineering and Structural Dynamics Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. His research interest includes the dynamics of materials and structures by means of analytical, numerical, experimental, and monitoring approaches that he develops with his many collaborators.



**PÄR JOHANNESSON** received the Ph.D. degree in mathematical statistics from the Lund Institute of Technology, with a thesis on statistical load analysis for fatigue, and the Postdoctoral degree in mathematical statistics from the Chalmers University of Technology and PSA Peugeot Citroën, France. He is currently a Senior Researcher in mechanical reliability, fatigue, and load analysis with the RISE Research Institute of Sweden, Gothenburg, Sweden. He is also a Docent in mathematical statistics with the Chalmers University of Technology. He has more than 20 years of experience in research and development in fatigue and loads analysis with a focus on engineering statistics and reliability in industrial applications. He has published more than 20 articles in international journals. He is a Co-Editor of the handbook *Guide to Load Analysis for Durability in Vehicle Engineering*.



**FREDRIK BRUZELIUS** received the M.Sc. degree in applied mathematics from Linköping University, in 1999, and the Ph.D. degree in control theory from the Chalmers University of Technology, in 2004. He has five years of experience with Volvo, developing vehicle dynamics estimation and control strategies in cars and trucks. He is currently a Senior Researcher in vehicle dynamics with the Department of Mechanics and Maritime Sciences, Chalmers University of Technology. His research interests include vehicle state estimation and control, tire modeling and simulation, driving simulators, and vehicle testing.



**TOHEED GHANDRIZ** received the M.Sc. degree in mechanical engineering from the Blekinge Institute of Technology, Karlskrona, Sweden, in 2014, and the Ph.D. degree in vehicle dynamics from the Chalmers University of Technology, Gothenburg, Sweden. From 2014 to 2015, he was a Research and Development Mechanical Development Engineer with Dassault Systèmes, Lund, Sweden. He is currently with Volvo AB. His research interests include the optimal distribution of propulsion over axles in long combination vehicles considering energy, dynamic performance, and transport economy.



**BENGT JACOBSON** received the Ph.D. degree in machine elements from the Chalmers University of Technology, Gothenburg, Sweden, in 1993. From 2001 to 2010, he was a Technical Expert with Volvo Car Corporation, Gothenburg, with a focus on vehicle dynamics control and active safety. He is currently a Senior Professor with the Department of Mechanics and Maritime Sciences, Chalmers University of Technology, and the Head of the Vehicle Dynamics Group.

...