Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2025) 184-197

ECCOMAS
Proceedia

UNCECOMP 2025
6th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Science and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Rhodes Island, Greece, 15-18 June 2025

UNCERTAINTY QUANTIFICATION FOR THE SIMULATION OF AUTOMOTIVE COMPONENT LOADS USING STOCHASTIC SURROGATE MODELS

Paul Straehle[1,2], **Steve Wolff-Vorbeck**[1], **Thomas Leyendecker**[1] **and Carsten Proppe**[2]

[1] Robert Bosch GmbH
Robert-Bosch-Campus 1, 71272 Renningen, Germany
e-mail: {paul.straehle, steve.wolff-vorbeck, thomas.leyendecker}@de.bosch.com

[2] Institute of Engineering Mechanics, Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
e-mail: carsten.proppe@kit.edu

Abstract. *To design reliable automotive components, it is important to understand the expected operational loads. Early in development, virtual load generation helps derive these loads realistically. Many loads depend on vehicle speeds, making velocity simulation an essential part of virtual load generation. The velocity profiles are determined using a stochastic simulator, which employs a stochastic process with variable parameters to model the variation in driving behavior and generate random velocity profiles for a given route and driver. The velocity profiles are propagated to component loads using use case specific system simulations. The derived velocities and loads are subject to model and data uncertainty. To address this, we propose a framework that models input parameters as random variables in a Bayesian framework, allowing for the handling of uncertainties. To enable the identification of these uncertainties and sampling from the posterior distribution, we apply Polynomial Chaos Expansion and Generalized Lambda Models to derive stochastic surrogate models as a substitute for the velocity simulation. We show that the parameter distributions derived with the surrogate provide valid descriptions of the velocity simulation uncertainty and validate the results with respect to measured velocities. As a result, our study yields a broad probabilistic framework for data-driven reliability prediction of automotive components and enables the consideration of load estimation uncertainties in an early stage of component development.*

Keywords: Uncertainty Quantification stochastic surrogate model, Bayesian inference, velocity simulation

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2025.
doi: 10.7712/120225.12349.21214

1 Introduction

The established approach to estimate the loads on automotive components is to perform test drives under certain assumptions of the actual usage of the car by a customer. From these assumptions, a test route is derived which is then driven in a specifi manner by a test driver. The loads measured on this route are then extrapolated to the targeted lifetime of the component. However, due to the high uncertainties associated with the assumptions and the very limited number of kilometers driven, this approach can only give a limited view into the loads acting on the component over its actual lifetime. In addition, existing physical components (e.g. prototypes or components from previous generations) are required on which the measurements can be carried out. This aspect makes the approach unfeasible for the development of new, innovative components as currently developed for electric and fuel cell electric vehicles. In this case, load measurements must be "virtualized", meaning that they are fully generated by simulations. This requires the modeling of realistic profile of the speed of the vehicle in the driving direction, referred to in the following as the velocity of the vehicle, along an arbitrary route taking into account influence such as traffi and driving behavior.

In the following, the velocity simulation [1, 2] is a stochastic simulator $\mathcal{M}_v(s, \boldsymbol{r}(s), \boldsymbol{\theta}_v)$ and considers various random influence to generate random velocity profile for a given route. It produces a stochastic process

$$\{V_s = \mathcal{M}_v(s, \boldsymbol{r}(s), \boldsymbol{\theta}_v) : s \in S\} \tag{1}$$

where s is a position along the route contained in the set of positions S. Moreover, we consider different properties $\boldsymbol{r}(s)$, such as the legal speed limit, curvature, slope, etc., of the road as a function of the position s along the route and $\boldsymbol{\theta}_v$ is the parameter vector describing driver behavior and vehicle properties. To assess component loads the simulated velocities are fed into a system simulation (i.e., transfer function) $\mathcal{M}_g(v(s), \boldsymbol{p}(s), \boldsymbol{\theta}_g)$ which simulates component loads based on the velocity as well as other inputs $\boldsymbol{p}(s)$ and parameters $\boldsymbol{\theta}_g$. Due to the random velocity profiles the component load is thus also described by a stochastic process

$$\{T_s = \mathcal{M}_g(V_s, \boldsymbol{p}(s), \boldsymbol{\theta}_g) : s \in S\}. \tag{2}$$

In Section 2 we describe a comprehensive Uncertainty Quantificatio (UQ) framework to handle uncertainties of the velocity simulation in the context of virtual load derivation. To enable the quantificatio of uncertainty of the velocity simulation we incorporate a Generalized Lambda Model (GLaM) based on the model derived in [3], where the authors used the Generalized Lambda Distribution (GLD) to surrogate the response distribution of a stochastic simulator. In Section 3, the fittin and validation of the stochastic surrogate model is described. Furthermore, in Section 4 a Sobol' indices based sensitivity analysis [4] of the velocity simulation parameters is performed to identify the most important parameters. The posterior distribution of the most important parameters is identifie within a Bayesian framework, see Section 5. Finally, in Section 6 an evaluation of the simulation results under consideration of the identifie uncertainties is performed.

2 Uncertainty Quantification framework

The UQ framework provides the structure and connects the needed parts of a model to identify uncertainties in the velocity simulation. An overview of the framework is given in Fig. 1. We focus on a use case where we want to identify the uncertainties in the parameter vector $\boldsymbol{\theta}_v$

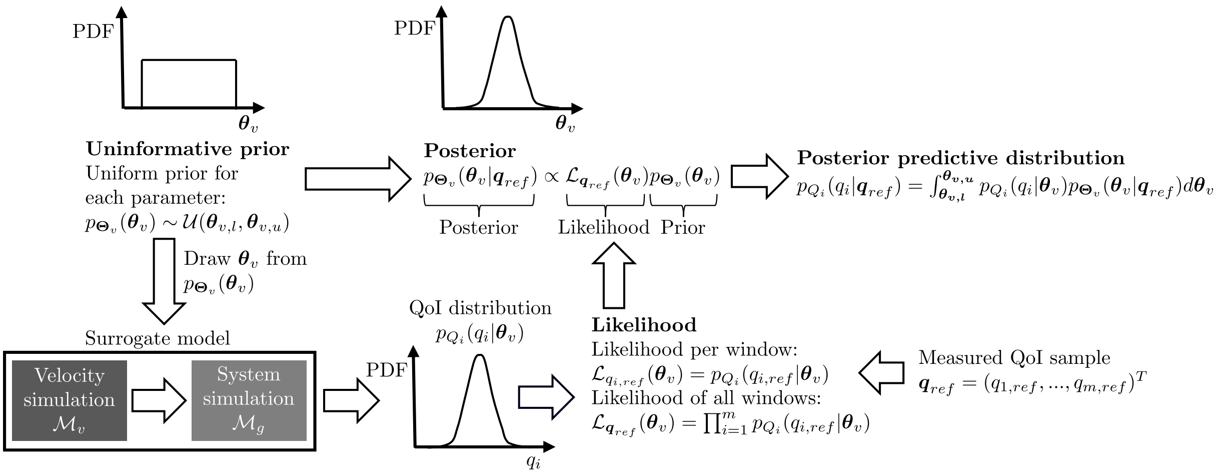


Figure 1: UQ framework

describing the driving behavior of a single driver on a set of routes. The available dataset consists of velocity measurements on $216\,\mathrm{km}$ of city, country, and motorway roads. We consider one specifi Quantity of Interest (QoI) which is given by the torque load on the drivetrain of an electric vehicle. However, we note that the method developed in the following can be applied to any other load case. Thus, the system simulation $\mathcal{M}_g(v(s), \boldsymbol{p}(s), \boldsymbol{\theta}_g)$ is a model of the vehicle and its drivetrain which, under consideration of the vehicle dynamics, outputs the respective torque in the drivetrain component. As outlined in [2] the torque series are split into windows of length and shift $s_{w,l} = s_{w,s} = 5000\,\mathrm{m}$ to asses the component load. For each window i the gear damage sum $d_{G,i}$ as it could be used for gear dimensioning according to [5] is calculated. As we only investigate the influenc of the load generation on the damage sum $d_{G,i}$ and do not consider a particular component and its properties the damage sum is normalized with the mean damage of all windows $\bar{d}_G$ and the window length $s_{w,l}$

$$d'_{G,i} = \frac{d_{G,i}}{s_{w,s}\bar{d}_G}. \tag{3}$$

Further, due to the exponential form of the damage sum we consider its logarithm as the QoI q_i, i.e.,

$$q_i = \log d'_{G,i}. \tag{4}$$

As the focus is on the uncertainty in the velocity simulation we obtain a measured sample of the QoI by applying the system simulation to measured velocities instead of using a measured torque series directly. This way the reference damage $q_{i,ref}$ for each window is obtained from the measured velocities on the given set of routes. If there were a second set of velocity measurements from the same driver on the same set of routes the damage per window would change as a result of random changes in driving behavior, traffic etc. It is thus necessary to model the damage per window as a random variable with associated Probability Density Function (PDF)

$$Q_i \sim p_{Q_i}(q_i) \tag{5}$$

and, for multiple windows $\boldsymbol{Q} = (Q_1, ..., Q_m)^T$, obtain the joint distribution $p_{\boldsymbol{Q}}(\boldsymbol{q})$. As the properties of the damage distribution can change in each window depending on the driving and road situation in the window, the distributions Q_i are not identically distributed and the damage distributions between windows are assumed as independent. Using the stochastic velocity simulation

$\mathcal{M}_v(s, \boldsymbol{r}(s), \boldsymbol{\theta}_v)$ in combination with the deterministic torque simulation $\mathcal{M}_g(v(s), \boldsymbol{p}(s), \boldsymbol{\theta}_g)$ the damage distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ can be estimated as a function of the velocity simulation parameters $\boldsymbol{\theta}_v = (\theta_{v,1}, ..., \theta_{v,k})^T$. The parameters of the torque simulation $\boldsymbol{\theta}_g$ are assumed to be known exactly and are not considered. The measured sample $\boldsymbol{q}_{ref} = (q_{1,ref}, ..., q_{m,ref})^T$ is drawn from the true, unknown distribution $\tilde{p}_{\boldsymbol{Q}}(\boldsymbol{q})$. The velocity simulation parameters $\boldsymbol{\theta}_v$ are define as a random variable $\boldsymbol{\Theta}_v$ in a Bayesian interpretation of probability, meaning that we aim to fin the posterior distribution $p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v|\boldsymbol{q}_{ref})$ using the sample $\boldsymbol{q}_{ref}$ such that the true damage distribution $\tilde{p}_{\boldsymbol{Q}}(\boldsymbol{q})$ is well approximated and the uncertainty in the simulated distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ is expressed through the uncertain parameters $\boldsymbol{\Theta}_v$.

For the uncertain velocity simulation parameters a uniform, uninformative prior

$$p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v) \sim \mathcal{U}(\boldsymbol{\theta}_{v,l}, \boldsymbol{\theta}_{v,u}) \tag{6}$$

is defined The lower $\boldsymbol{\theta}_{v,l}$ and upper $\boldsymbol{\theta}_{v,u}$ bounds are set to threshold values estimated from a database with velocity simulation parameters from numerous drivers to ensure a physically plausible meaning for them. Table 1 lists all parameters in the vector $\boldsymbol{\theta}_v$ and their bounds. Based on the prior a Design of Experiment (DoE) is performed using Latin Hypercube Sampling (LHS) [6]. The DoE then provides the data needed to fi a stochastic surrogate model, see Section 3.

Parameters $\boldsymbol{\theta}_v$	$\boldsymbol{\theta}_{v,l}$	$\boldsymbol{\theta}_{v,u}$	Description
$\bar{a}_{x,max,dr}$	1	5	Mean of maximum driver acceleration in $\mathrm{m/s^2}$
$\bar{a}_{x,min,dr}$	-5	-1	Mean of maximum driver deceleration in $\mathrm{m/s^2}$
$o_{v,ci}$	0.5	1.3	Velocity offset on city roads
$o_{v,co}$	0.5	1.3	Velocity offset on country roads
$o_{v,mo}$	0.5	1.3	Velocity offset on motorways
$\bar{\sigma}_{\zeta,log,ci}$	-5	-2	Mean log standard deviation of velocity fluctuation on city roads
$\bar{\sigma}_{\zeta,log,co}$	-6	-2	Mean log standard deviation of velocity fluctuation on country roads
$\bar{\sigma}_{\zeta,log,mo}$	-7	-3	Mean log standard deviation of velocity fluctuation on motorways
$\bar{v}_{max,dr}$	110	150	Mean maximum driver speed in $\mathrm{km/h}$

Table 1: Velocity simulation parameters $\boldsymbol{\theta}_v$ and their bounds

Based on the simulated damage distribution per window $p_{\boldsymbol{Q}_i}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ we defin the likelihood of the measurement per window i by

$$\mathcal{L}_{q_{i,ref}}(\boldsymbol{\theta}_v) = p_{Q_i}(q_{i,ref}|\boldsymbol{\theta}_v). \tag{7}$$

Due to the independence of the windows, the likelihood of all windows being their product

$$\mathcal{L}_{\boldsymbol{q}_{ref}}(\boldsymbol{\theta}_v) = \prod_{i=1}^{m} p_{Q_i}(q_{i,ref}|\boldsymbol{\theta}_v), \tag{8}$$

and the log-likelihood $\ell_{\boldsymbol{q}_{ref}}(\boldsymbol{\theta}_v)$ follows in the same manner as the sum of their logarithms.

Through Bayes' rule, the posterior distribution of the velocity simulation parameters is obtained

$$p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v|\boldsymbol{q}_{ref}) \propto \mathcal{L}_{\boldsymbol{q}_{ref}}(\boldsymbol{\theta}_v)p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v). \tag{9}$$

Evaluation of the damage distribution under consideration of the posterior gives the posterior predictive distribution of the damage per window given by

$$p_{Q_i}(q_i|\boldsymbol{q}_{ref}) = \int_{\boldsymbol{\theta_{v,l}}}^{\boldsymbol{\theta_{v,u}}} p_{Q_i}(q_i|\boldsymbol{\theta}_v)p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v|\boldsymbol{q}_{ref})d\boldsymbol{\theta}_v. \tag{10}$$

3 Stochastic surrogate model

A surrogate model approximates a simulation model to reduce costly evaluations of the full simulation. As the used velocity simulation in combination with the torque simulation and the damage evaluation is a stochastic simulator this requires a stochastic surrogate model which allows for faster evaluations of the damage distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ by the approximate surrogate distribution $\hat{p}_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$. Different stochastic surrogate models are described in [3, 7, 8]. For our problem, we choose the GLaM [3] which is built on the GLD [9].

A GLD can well approximate most unimodal distributions and has four parameters $\{\lambda_l : l = 1, ..., 4\}$. The lambda parameters are given by a function of the parameters $\boldsymbol{\theta}_v$ and thus the GLaM gives the approximate QoI distribution per window as a function of $\boldsymbol{\theta}_v$

$$\hat{Q}_i(\boldsymbol{\theta}_v) \sim \hat{p}_{Q_i}(q_i|\boldsymbol{\theta}_v) = \mathrm{GLD}(\lambda_1(\boldsymbol{\theta}_v), \lambda_2(\boldsymbol{\theta}_v), \lambda_3(\boldsymbol{\theta}_v), \lambda_4(\boldsymbol{\theta}_v)). \tag{11}$$

Moreover as the distributions per window are assumed as independent they can be combined to the joint distribution

$$\hat{p}_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v) = \prod_{i=1}^{m} \hat{p}_{Q_i}(q_i|\boldsymbol{\theta}_v). \tag{12}$$

Each lambda parameter and its dependence on $\boldsymbol{\theta}_v$ is modeled with a Polynomial Chaos Expansion (PCE)

$$\begin{aligned} \lambda_l(\boldsymbol{\theta}_v) &\approx \lambda_l^{PC}(\boldsymbol{\theta}_v;\boldsymbol{c}) = \sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{A}}_l} c_{l,\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_v), \quad l = 1, 3, 4 \\ \lambda_2(\boldsymbol{\theta}_v) &\approx \lambda_2^{PC}(\boldsymbol{\theta}_v;\boldsymbol{c}) = \exp\left(\sum_{\boldsymbol{\alpha}\in\boldsymbol{\mathcal{A}}_2} c_{2,\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}_v)\right) \end{aligned} \tag{13}$$

where $\lambda_2(\boldsymbol{\theta}_v)$ is transformed to $\log(\lambda_2(\boldsymbol{\theta}_v))$ to ensure positivity.

The model is built on an experimental design of $N = 1000$ samples without replications, i.e., for each sample from the parameter distribution $p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v)$ one sample from the conditional damage distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ is drawn. Using the Maximum Likelihood Estimation (MLE) procedure described in [3], the PCE coeffici nts $c_{l,\boldsymbol{\alpha}}$ for the lambda parameters are fitted The polynomial degrees for the four lambdas are $p_1 = 2, p_2 = p_3 = p_4 = 1$ with $q_1 = q_2 = q_3 = q_4 = 1$ for the full polynomial basis in the truncation sets $\{\boldsymbol{\mathcal{A}}_l : l = 1, ..., 4\}$.

For validation, a second DoE with $N = 100$ samples is done where for each parameter sample $\boldsymbol{\theta}_v$ we run 1000 repetitions to approximate $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$. Fig. 2 compares how the distributions

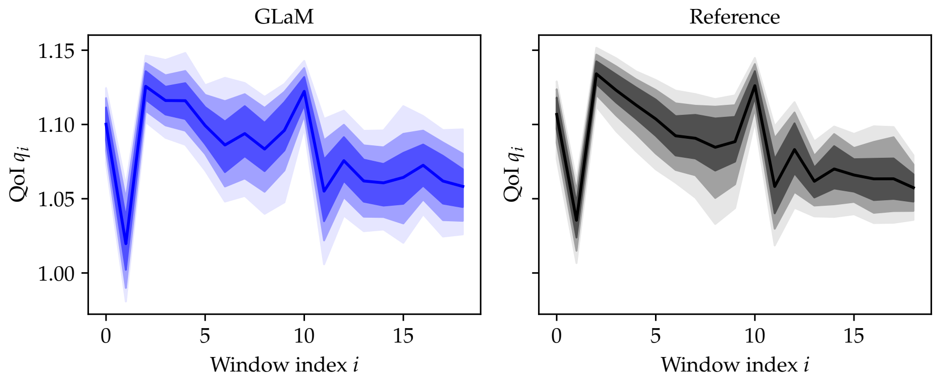


Figure 2: Comparison of reference $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ and GLaM $\hat{p}_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ distributions for different windows along a route with fi ed parameter vector $\boldsymbol{\theta}_v$. The 3 color shades indicate regions that contain $50, 75$, and 90% of the distributions.

$p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ and $\hat{p}_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ deviate for different windows along a route with the parameter vector $\boldsymbol{\theta}_v$ being fi ed.

Between the GLaM distribution $\hat{p}_{Q_i}(q_i|\boldsymbol{\theta}_v)$ and the distribution $p_{Q_i}(q_i|\boldsymbol{\theta}_v)$ the empirical Wasserstein distance (e.g. [10]) $d_{WS}(Q_i(\boldsymbol{\theta}_v), \hat{Q}_i(\boldsymbol{\theta}_v))$ is calculated for each of the 100 parameter samples and each window i. The empirical Wasserstein distance is computed with the 1000 samples obtained from $p_{Q_i}(q_i|\boldsymbol{\theta}_v)$ through repeated sampling and an equal amount of samples drawn from $\hat{p}_{Q_i}(q_i|\boldsymbol{\theta}_v)$. Each computed Wasserstein distance is normalized by the standard deviation of the empirical distribution $\sigma(Q_i(\boldsymbol{\theta}_v))$

$$d(Q_i(\boldsymbol{\theta}_v), \hat{Q}_i(\boldsymbol{\theta}_v)) = \frac{d_{WS}(Q_i(\boldsymbol{\theta}_v), \hat{Q}_i(\boldsymbol{\theta}_v))}{\sigma(Q_i(\boldsymbol{\theta}_v))}. \tag{14}$$

The standard deviation $\sigma(Q_i(\boldsymbol{\theta}_v))$ can be seen as the Wasserstein distance between the distribution $p_{Q_i}(q_i|\boldsymbol{\theta}_v)$ and the mean $\mathbb{E}[Q_i(\boldsymbol{\theta}_v)]$, see [3] for more details. The normalized Wasserstein distance can thus be interpretedas the ratio between the Wasserstein distances $d_{WS}(Q_i(\boldsymbol{\theta}_v), \hat{Q}_i(\boldsymbol{\theta}_v))$ and $d_{WS}(Q_i(\boldsymbol{\theta}_v), \mathbb{E}[Q_i(\boldsymbol{\theta}_v)])$, i.e., how much improvement is obtained by using the stochastic surrogate distribution $\hat{p}_{Q_i}(q_i|\boldsymbol{\theta}_v)$ compared to a degenerate distribution at the true mean.

For each window the mean of the normalized Wasserstein distance for all parameter samples from $p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v)$ is taken

$$\epsilon_i = \mathbb{E}[d(Q_i(\boldsymbol{\Theta}_v), \hat{Q}_i(\boldsymbol{\Theta}_v))]. \tag{15}$$

Fig. 3 depicts the distribution of the error measure ϵ_i across all windows. For most windows, an error $\epsilon_i < 0.5$ is observed where the mean error is $\bar{\epsilon} = 0.55$. When the GLaM is slightly modifie by assuming normality of the distributions $\hat{p}_{Q_i}(q_i|\boldsymbol{\theta}_v)$, i.e., $\lambda_3 = \lambda_4 = 0.13$ the mean error measure increases slightly to $\bar{\epsilon} = 0.58$.

A second metric for the validation is the accuracy with which the likelihood of a measurement can be predicted with the surrogate model. To calculate the likelihood a reference sample is define as the mean of the prior predictive distribution (see also (10) for the posterior predictive distribution)

$$\boldsymbol{q}_{ref,Pr} = \mathbb{E}[\boldsymbol{Q}(\boldsymbol{\Theta}_v)] \text{ with } p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v) \sim \mathcal{U}(\boldsymbol{\theta}_{v,l}, \boldsymbol{\theta}_{v,u}). \tag{16}$$

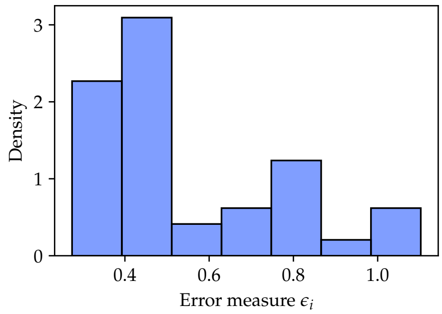
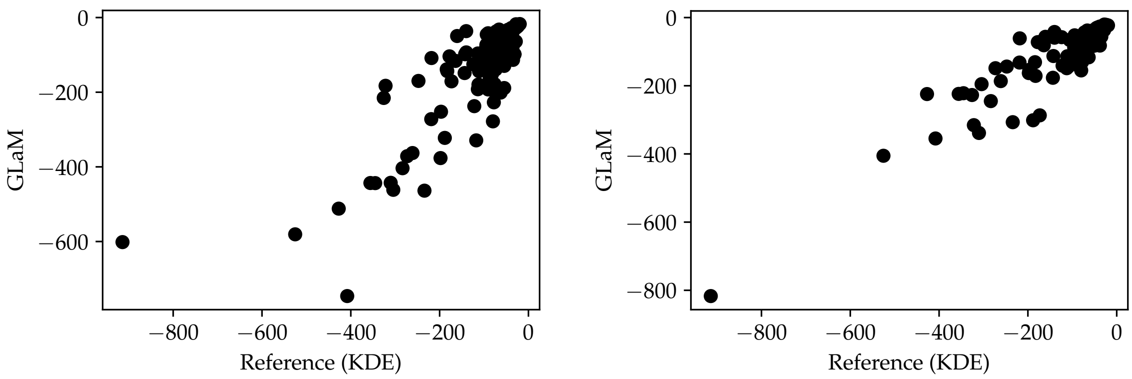


Figure 3: Distribution of error measure ϵ_i

The log-likelihood is thus calculated according to (8). To calculate the likelihood with the reference distribution a Kernel Density Estimation (KDE) is applied to the repeated samples from $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ to approximate the distribution. Fig. 4a depicts the log-likelihood for each of the 100 parameter samples derived with the KDE approximation of $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ and the surrogate model $\hat{p}_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$. The coefficien of determination equals $R^2 = 0.51$. When we simplify the GLaM by using the normality assumption and setting $\lambda_3 = \lambda_4 = 0.13$ the prediction is improved with $R^2 = 0.81$, see Fig. 4b. As shown in the analysis of the normalized Wasserstein distance there is only a slightly better approximation of the distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$ when using the full GLD. Additionally, through the normality assumption, the model is regularized and less prone to overfitting contributing to the improved prediction of the log-likelihood. Due to these reasons, all further analyses are carried out with the simplifie GLaM which assumes a normal distribution of the surrogate $\hat{p}_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$.



(a) Full GLaM ($R^2 = 0.51$)

(b) GLaM with $\lambda_3 = \lambda_4 = 0.13$ ($R^2 = 0.81$)

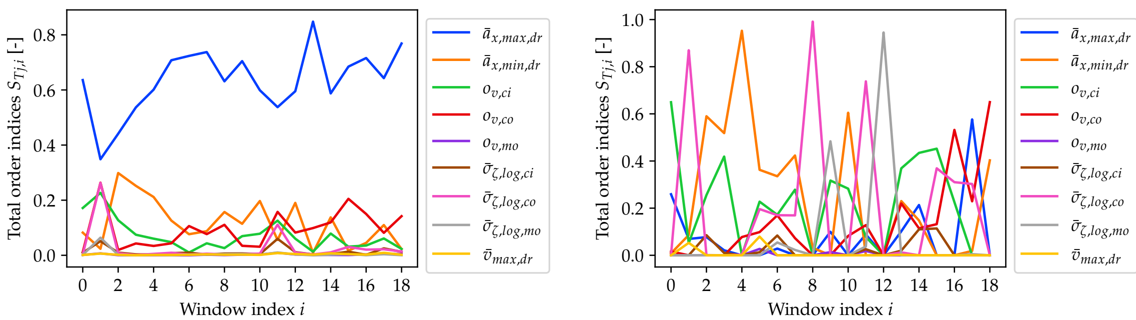
Figure 4: Comparison of log-likelihood $\ell_{\boldsymbol{q}_{ref,Pr}}(\boldsymbol{\theta}_v)$ between reference solution obtained by KDE approximation of repeated samples and the GLaM surrogate model

4 Sensitivity analysis

Based on the derived GLaM surrogate a global sensitivity analysis using Sobol' indices [4] of the input parameters is conducted. In [11] different approaches to global sensitivity analysis of stochastic simulators are described. We focus on sensitivity analysis based on deterministic QoIs derived from the damage distribution $\hat{p}_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_v)$.

For each window i the Sobol' indices of the mean $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$ and standard deviation $\sigma(Q_i|\boldsymbol{\theta}_v)$ are estimated by Monte Carlo simulations [12, 13]. Mean and standard deviation are directly obtained from the description of the GLD in the GLaM. The parameter space for the Monte Carlo simulations is define by the prior $p_{\boldsymbol{\Theta}_v}(\boldsymbol{\theta}_v)$ from (6) where $N = 10^4$ samples are drawn. Fig. 5 depicts the total order Sobol' indices $S_{Tj,i}$ for the parameters $\theta_{v,j}$ along a route calculated for $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$ and $\sigma(Q_i|\boldsymbol{\theta}_v)$, see Tab. 1 for a description of the parameters $\theta_{v,j}$. For this route, the mean $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$ of most windows is strongly influence by $\bar{a}_{x,max,dr}$. The standard deviation $\sigma(Q_i|\boldsymbol{\theta}_v)$ of the windows is influence by a greater variety of parameters. To analyze the overall sensitivity of the QoIs to the parameters the mean firs and total order Sobol' indices

$$
\begin{aligned}
\bar{S}_j &= \frac{1}{m}\sum_{i=1}^{m} S_{j,i} \\
\bar{S}_{Tj} &= \frac{1}{m}\sum_{i=1}^{m} S_{Tj,i}
\end{aligned}
\tag{17}
$$

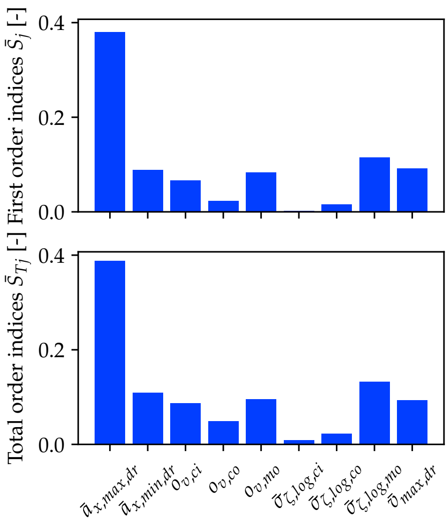


(a) $\mathbb{E}[Q_i]$ (b) $\sigma(Q_i)$

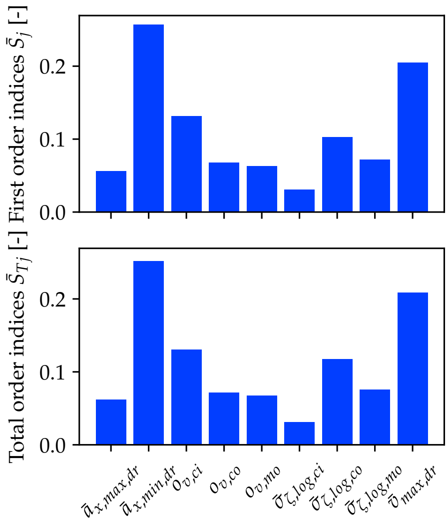
Figure 5: Total order Sobol' indices $S_{Tj,i}$ along a route for $\mathbb{E}[Q_i]$ and $\sigma(Q_i)$

are defined Overall, the mean $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$ is mainly sensitive to $\bar{a}_{x,max,dr}$ while the standard deviation $\sigma(Q_i|\boldsymbol{\theta}_v)$ is primarily sensitive to $\bar{a}_{x,min,dr}$ and $\bar{v}_{max,dr}$, see Fig. 6. For both QoIs there is little difference between the mean firs $\bar{S}_j$ and total order $\bar{S}_{Tj}$ indices, indicating no significan couplings of the velocity simulation parameters $\boldsymbol{\theta}_v$. The parameters $\bar{a}_{x,max,dr}$ and $\bar{a}_{x,min,dr}$ influenc the acceleration and deceleration behavior of the driver which is strongly related to the torque in the drivetrain and thus the investigated gear damage sum. Both parameters are therefore expected to strongly influenc the mean $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$ and standard deviation $\sigma(Q_i|\boldsymbol{\theta}_v)$ of the damage distribution. The parameter $\bar{v}_{max,dr}$ influence the mean speed of the driver (especially on motorways) which changes the duration and frequency of the driver accelerating and decelerating and thus the torque and gear damage sum.

To globally analyze the importance of the parameters the sensitivity of the log-likelihood $\ell_{\boldsymbol{q}_{ref,Pr}}(\boldsymbol{\theta}_v)$ to the velocity simulation parameters $\boldsymbol{\theta}_v$ is calculated. The likelihood is calculated with the reference sample $\boldsymbol{q}_{ref,Pr}$ define as the mean of the prior predictive distribution, see (16). The Sobol' indices thus give an indication of how sensitive the likelihood of the reference sample is to changes in the parameters. The firs S_j and total order S_{Tj} Sobol' indices show a strong influenc of the three parameters $\bar{a}_{x,max,dr}$, $\bar{a}_{x,min,dr}$ and $o_{v,ci}$, see Fig. 7a, the firs two of



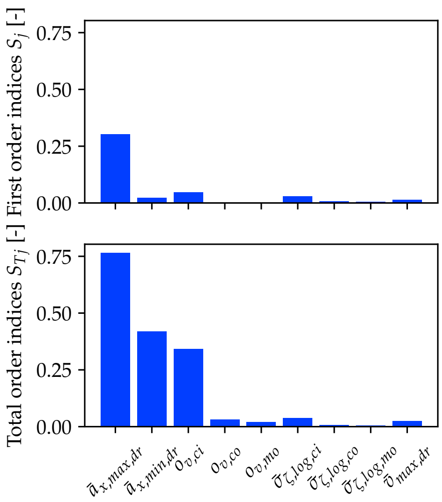
(a) $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$



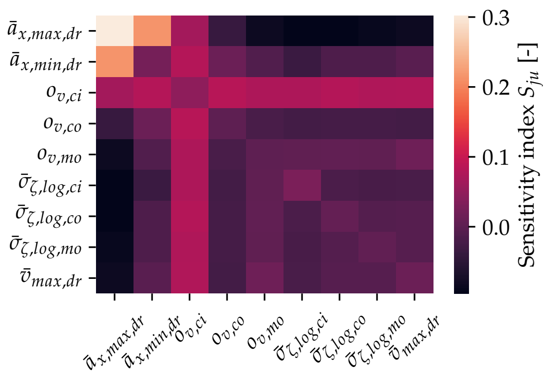
(b) $\sigma(Q_i|\boldsymbol{\theta}_v)$

Figure 6: Mean firs $\bar{S}_j$ and total order $\bar{S}_{Tj}$ indices for $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$ and $\sigma(Q_i|\boldsymbol{\theta}_v)$

which were also the strongest influenc factors in the sensitivity analysis of mean $\mathbb{E}[Q_i|\boldsymbol{\theta}_v]$ and standard deviation $\sigma(Q_i|\boldsymbol{\theta}_v)$. The offset $o_{v,ci}$ influence the mean velocity driven in the city and has similar effects on the gear damage sum as the parameter $\bar{v}_{max,dr}$ mentioned before but with a focus on city streets. The values of the first and second-order indices are significantl different, indicating higher-order coupling effects. Fig. 7b depicts the second-order Sobol' indices S_{ju} with a strong coupling effect between $\bar{a}_{x,max,dr}$ and $\bar{a}_{x,min,dr}$.



(a) First S_j and total order S_{Tj} Sobol' indices



(b) Second order Sobol' indices S_{ju}

Figure 7: Sobol' indices for log-likelihood $\ell_{\boldsymbol{q}_{ref,Pr}}(\boldsymbol{\theta}_v)$

5 Bayesian parameter inference

A Bayesian framework is applied to identify the posterior distribution $p_{\mathbf{\Theta}_v^*}(\boldsymbol{\theta}_v^*|\boldsymbol{q}_{ref})$ of the parameters in

$$\mathrm{A}_v = \{\bar{a}_{x,max,dr}, \bar{a}_{x,min,dr}, o_{v,ci}\} \tag{18}$$

with the most influenc on the log-likelihood. These parameters $\boldsymbol{\theta}_v^* = (\theta_{v,j} : j \in \mathrm{A}_v)^T$ are modeled as a random variable $\mathbf{\Theta}_v^*$. The Markov Chain Monte Carlo (MCMC) algorithm with slice sampling (e.g. [14]) is employed to sample from the posterior distribution in (9), with the measured values of the QoI for each window $\boldsymbol{q}_{ref}$. As a result the joint posterior distribution $p_{\mathbf{\Theta}_v^*}(\boldsymbol{\theta}_v^*|\boldsymbol{q}_{ref})$ is obtained. A KDE of the distribution marginalized to one and two dimensions is depicted in Fig. 8. The black diamonds and lines indicate the default values of the parameters $\boldsymbol{\theta}_{v,0}$ identifie with a heuristic from the measured velocities of the driver (e.g. $\bar{a}_{x,max,dr}$ and $\bar{a}_{x,min,dr}$), as done in [1], or taken as a guess value from experience (e.g. $o_{v,ci}$). For $\bar{a}_{x,min,dr}$ and $o_{v,ci}$ these values are within the posterior distributions. The posterior distribution of $\bar{a}_{x,max,dr}$ is bounded to the left due to the lower bound of the prior distribution $\boldsymbol{\theta}_{v,l}$ to keep the parameter values within a physically plausible range.

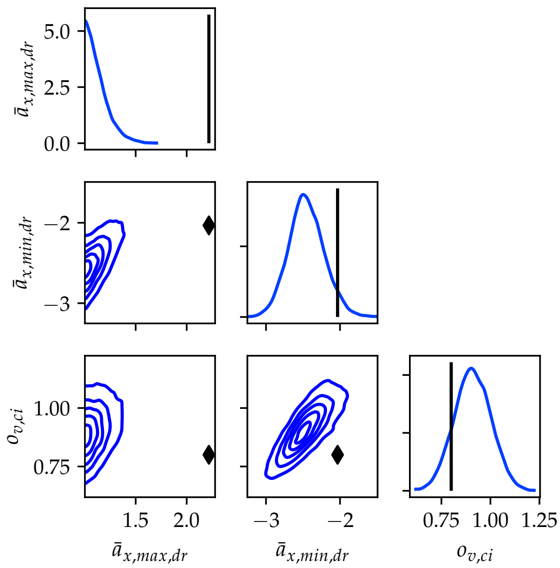


Figure 8: Joint posterior distribution $p_{\mathbf{\Theta}_v^*}(\boldsymbol{\theta}_v^*|\boldsymbol{q}_{ref})$ marginalized over 1 and 2 dimensions. Black diamonds and lines indicate the default values of the parameters $\boldsymbol{\theta}_{v,0}$

6 Evaluation of results

For evaluation of the identifie posterior distribution $p_{\mathbf{\Theta}_v^*}(\boldsymbol{\theta}_v^*|\boldsymbol{q}_{ref})$ the posterior predictive $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ distribution in (10), is evaluated and compared to the distribution obtained with the default parameter vector $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$. The posterior predictive $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ is evaluated for the parameters in A_v while the other parameters are set to their default values. It should be noted that the default parameter vector $\boldsymbol{\theta}_{v,0}$ is partly also a function of the measured sample $\boldsymbol{q}_{ref}$ as some of the parameters in it are derived with heuristics from the measured velocities which are used to calculate $\boldsymbol{q}_{ref}$. Both distributions $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ and $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$ are obtained by sampling $N = 1000$ times with the full simulation model.

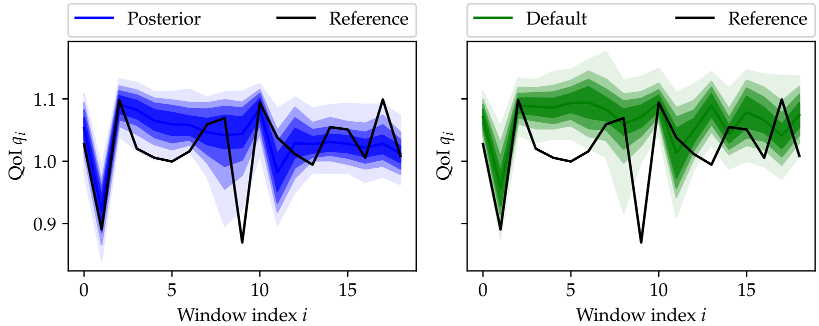


Figure 9: Comparison of posterior predictive distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ and distribution obtained with default parameters $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$ for multiple windows along a route. The center line is the mean of the distributions, the different color shades indicate the $50\ \%, 75\ \%, 90\ \%$, and $99\ \%$ credible intervals.

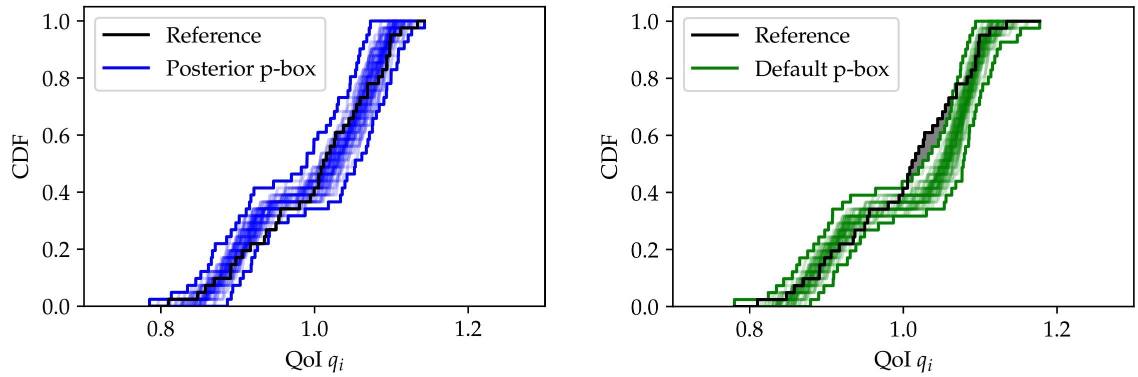
Fig. 9 depicts the $\gamma = 50\ \%, 75\ \%, 90\ \%$ and $99\ \%$ credible intervals as well as the mean of $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ and $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$ compared with $\boldsymbol{q}_{ref}$ for the windows along a route. The credible intervals are define from the quantiles $[q_{(1-\gamma)/2}, q_{(1+\gamma)/2}]$ such that γ is the probability of a value being in the interval. Overall the intervals are shifted towards the measured reference solution when using the posterior parameter distribution $p_{\boldsymbol{\Theta}_v^*}(\boldsymbol{\theta}_v^*|\boldsymbol{q}_{ref})$ compared to the default parameters $\boldsymbol{\theta}_{v,0}$. When evaluating the $99\ \%$ credible interval for all windows of all routes the posterior predictive distribution covers $83\ \%$ of all measured windows in $\boldsymbol{q}_{ref}$ compared to $76\ \%$ of the default distribution.

The density $p_{\tilde{Q}}(\tilde{q})$ describes the distribution of the QoI across multiple windows. In the considered use case this describes how often a certain load and damage sum $d_{G,i}$ occur in the considered set of routes. The available windows $i \in 1, ..., m$ are a sample from the set of all windows obtained from all routes driven over the lifespan of a vehicle. The uncertainty regarding the measured sample from the distribution $p_{\tilde{Q}}(\tilde{q})$ is thus reducible through more measurements, i.e., it is epistemic. From the simulation single samples from the window distributions $p_{Q_i}(q_i)$ for $i \in 1, ..., m$ can be combined into the distribution $p_{\tilde{Q}}(\tilde{q})$. Repeating this sampling multiple times leads to multiple empirical distributions for $p_{\tilde{Q}}(\tilde{q})$ which can be combined into a probability-box (p-box) giving a quantificatio regarding the epistemic uncertainty caused by the size of the route sample. The p-boxes are obtained by taking 1000 samples from all windows, i.e. by "driving" 1000 times over the given set of routes. Fig. 10 compares the p-boxes obtained from the default distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$ and the posterior predictive distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$. In the case of the posterior predictive distribution the p-box not only includes the uncertainty regarding the route set but also regarding the simulation parameters. When taking samples from $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ the simulated p-box fully encloses the measured sample from $\tilde{p}_{\tilde{Q}}(\tilde{q})$, while for the default distribution $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$ there is a significan deviation to the measurement.

Moreover, Fig. 11 compares the mean

$$\bar{\tilde{q}} = \mathbb{E}[\tilde{Q}] = \frac{1}{m} \sum_{i=1}^{m} q_i \tag{19}$$

obtained from the measurement $\boldsymbol{q}_{ref}$ and the distribution $p_{\bar{\tilde{Q}}}(\bar{\tilde{q}})$ of it derived from the 1000 samples taken from $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ and $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$ where each sample is the mean $\bar{\tilde{q}}$ of a single draw from $\boldsymbol{q} \sim \boldsymbol{Q}$. As q_i is the logarithm of the normalized damage sum (see (4)) the mean



(a) Obtained with $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{q}_{ref})$ (b) Obtained with $p_{\boldsymbol{Q}}(\boldsymbol{q}|\boldsymbol{\theta}_{v,0})$

Figure 10: Simulated p-boxes of the distribution $p_{\tilde{Q}}(\tilde{q})$ compared to the measured reference sample

$\bar{\tilde{q}}$ of it is linearly related to the accumulated log damage sum across all windows $d_{log,sum} = \sum_{i=1}^{m} \log(d_{G,i})$

$$\begin{aligned}
\bar{\tilde{q}} &= \frac{1}{m}\sum_{i=1}^{m} \log\left(\frac{d_{G,i}}{s_{w,s}\bar{d}_G}\right) \\
\bar{\tilde{q}} &= \frac{1}{m}\left(\sum_{i=1}^{m} \log(d_{G,i}) - m(s_{w,s}\bar{d}_G)\right) \\
d_{log,sum} &= m\bar{\tilde{q}} + m(s_{w,s}\bar{d}_G) = \sum_{i=1}^{m} \log(d_{G,i})
\end{aligned} \tag{20}$$

making it especially relevant for the presented gear dimensioning use case. By using the posterior distribution of the parameters $p_{\Theta_v^*}(\boldsymbol{\theta}_v^*|\boldsymbol{q}_{ref})$ the distribution of the mean $p_{\bar{\tilde{Q}}}(\bar{\tilde{q}})$ includes the measurement with higher probability.

7 Conclusion

In this contribution, we presented a UQ framework which allows to quantify parametric uncertainties in the virtual estimation of automotive component loads derived from simulated velocities. The framework is applied to a use case where the piecewise gear damage sum is calculated from the torque load in the drivetrain. The torque series and gear damage sum are derived from the simulated velocity on a given set of routes. It is demonstrated that the GLaM presented in [3] is suitable to build a surrogate for the stochastic velocity simulation. With the help of the GLaM a global sensitivity analysis is performed and Sobol' indices are derived for the different parameters of the velocity simulation. With the sensitivity analysis, the set of relevant parameters for the use case can be significantl reduced. Gear damage sums derived from measured velocities of a single driver are incorporated into a Bayesian framework to inversely quantify the uncertainty of the relevant velocity simulation parameters for the driver regarding the gear damage sum. When forwarding the identifie parameter uncertainties through the simulation chain an improved match with the measured gear damage sums is observed taking into account the parametric uncertainties.

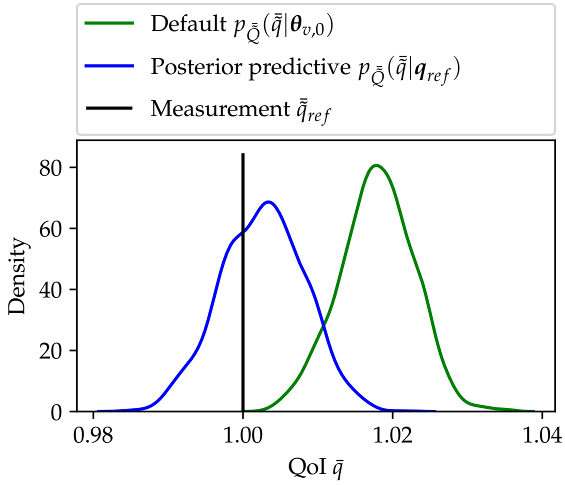


Figure 11: Mean distribution $p_{\bar{Q}}(\bar{\bar{q}})$

References

[1] K. Sandmann, T. Leyendecker, M. Burger, and M. Speckert. *Ableitung von feldrelevanten Lastkollektiven mittels Stochastischer Verkehrssimulation*. 2018. URL: https://publica.fraunhofer.de/handle/publica/403187.

[2] P. Strähle, S. Wolff-Vorbeck, T. Leyendecker, and C. Proppe. "Modeling Velocity Fluctuations for the Derivation of Automotive Component Loads". In: *ECCOMAS 2024*. 9th European Congress on Computational Methods in Applied Sciences and Engineering. 2024. URL: https://www.scipedia.com/public/Strahle_et_al_2024a.

[3] Xujia Zhu and Bruno Sudret. "Emulation of Stochastic Simulators Using Generalized Lambda Models". In: *SIAM/ASA Journal on Uncertainty Quantification* 9.4 (Jan. 2021), pp. 1345–1380. ISSN: 2166-2525. DOI: 10.1137/20M1337302. URL: https://epubs.siam.org/doi/10.1137/20M1337302 (visited on 11/06/2024).

[4] I.M Sobol. "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates". In: *Mathematics and Computers in Simulation* 55.1 (Feb. 2001), pp. 271–280. ISSN: 03784754. DOI: 10.1016/S0378-4754(00)00270-6. URL: https://linkinghub.elsevier.com/retrieve/pii/S0378475400002706 (visited on 05/22/2024).

[5] ISO 6336-6. *Calculation of load capacity of spur and helical gears — Part 6: Calculation of service life under variable load*. Version Second edition 2019-11. Geneva, 2019.

[6] M. D. McKay, R. J. Beckman, and W. J. Conover. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code". In: *Technometrics* 21.2 (May 1979), p. 239. ISSN: 00401706. DOI: 10.2307/1268522. URL: https://www.jstor.org/stable/1268522?origin=crossref (visited on 01/17/2025).

[7] Xujia Zhu and Bruno Sudret. "Stochastic polynomial chaos expansions to emulate stochastic simulators". In: *International Journal for Uncertainty Quantification* 13.2 (2023), pp. 31–52. ISSN: 2152-5080. DOI: 10.1615/Int.J.UncertaintyQuantification.2022042912. URL: https://www.dl.begellhouse.com/journals/

52034eb04b657aea,629ce0e12931a3e6,49d88de9790d7920.html (visited on 01/29/2025).

[8] Nora Lüthen, Stefano Marelli, and Bruno Sudret. "A spectral surrogate model for stochastic simulators computed from trajectory samples". In: *Computer Methods in Applied Mechanics and Engineering* 406 (Mar. 2023), p. 115875. ISSN: 00457825. DOI: 10.1016/j.cma.2022.115875. URL: https://linkinghub.elsevier.com/retrieve/pii/S0045782522008313 (visited on 01/29/2025).

[9] Marshall Freimer, Georgia Kollia, Govind S. Mudholkar, and C Thomas Lin. "A study of the generalized Tukey lambda family". In: *Communications in Statistics - Theory and Methods* 17.10 (Jan. 1988), pp. 3547–3567. ISSN: 0361-0926, 1532-415X. DOI: 10.1080/03610928808829820. URL: http://www.tandfonline.com/doi/abs/10.1080/03610928808829820 (visited on 01/27/2025).

[10] L. V. Kantorovich. "Mathematical Methods of Organizing and Planning Production". In: *Management Science* 6.4 (July 1960), pp. 366–422. ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.6.4.366. URL: https://pubsonline.informs.org/doi/10.1287/mnsc.6.4.366 (visited on 10/03/2024).

[11] X. Zhu and B. Sudret. "Global sensitivity analysis for stochastic simulators based on generalized lambda surrogate models". In: *Reliability Engineering & System Safety* 214 (Oct. 2021), p. 107815. ISSN: 09518320. DOI: 10.1016/j.ress.2021.107815. arXiv: 2005.01309[stat]. URL: http://arxiv.org/abs/2005.01309 (visited on 01/20/2025).

[12] Andrea Saltelli. "Making best use of model evaluations to compute sensitivity indices". In: *Computer Physics Communications* 145.2 (May 2002), pp. 280–297. ISSN: 00104655. DOI: 10.1016/S0010-4655(02)00280-1. URL: https://linkinghub.elsevier.com/retrieve/pii/S0010465502002801 (visited on 01/24/2025).

[13] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index". In: *Computer Physics Communications* 181.2 (Feb. 2010), pp. 259–270. ISSN: 00104655. DOI: 10.1016/j.cpc.2009.09.018. URL: https://linkinghub.elsevier.com/retrieve/pii/S0010465509003087 (visited on 01/24/2025).

[14] Radford M. Neal. "Slice sampling". In: *The Annals of Statistics* 31.3 (June 1, 2003). ISSN: 0090-5364. DOI: 10.1214/aos/1056562461. URL: https://projecteuclid.org/journals/annals-of-statistics/volume-31/issue-3/Slice-sampling/10.1214/aos/1056562461.full (visited on 01/28/2025).