Risk-averse real driving emissions optimization considering stochastic influences

A. Wasserburger\textsuperscript{a}, C. Hametner\textsuperscript{a} and N. Didcock\textsuperscript{b}

\textsuperscript{a}Christian Doppler Laboratory for Innovative Control and Monitoring of Automotive Powertrain Systems, TU Wien, Vienna, Austria; \textsuperscript{b}AVL List GmbH, Graz, Austria

\begin{abstract}
Optimization of vehicle powertrains is usually based on specific drive cycles and is performed on testbeds under reproducible conditions. However, in real-world operation, energy consumption and emissions differ significantly from the values obtained in testbed environments, which also implies breaching legislative thresholds. Therefore, in order to close the gap between testbed and real world, it is necessary to take random effects, like varying road and ambient conditions or various traffic situations, into account during the engine calibration process. In this article a stochastic optimization approach based on risk measures, that quantify the prevalent uncertainties, is presented. Rather than optimizing a deterministic value for one specific scenario described by a drive cycle, the distribution of possible outcomes is shaped in a way that it reflects the risk aversion and preferences of the decision maker. Simulation results show that incorporating randomness in the optimization process yields substantially more robust and reliable results.
\end{abstract}

\section{1. Introduction}

In light of growing public awareness regarding the detrimental effects of the transportation sector on the environment, the automotive industry is faced with great challenges. Consumers demand low fuel consumption and emissions while also expecting driving comfort and performance—see e.g. Langouët \textit{et al.} (2008). Apart from customer needs, legal requirements are becoming increasingly restrictive as well. Nonetheless, it has become a well-known fact that consumption and emission values may deviate heavily from the manufacturer specifications, which have been obtained in a controlled testing environment. In order to meet these challenges, future powertrain optimization has to take the stochastic nature of real-world driving into account. In this article it is shown how more robust results can be achieved by means of stochastic optimization appropriately quantifying the involved risk of unfavourable outcomes.

Today, powertrain calibration and energy efficiency management systems are typically based on optimization of characteristic drive cycles as in Karbowski \textit{et al.} (2006). Drive cycles are supposed to constitute representative speed–time profiles that are used to design, optimize and evaluate the performance of vehicles with respect to certain legal regulations—see e.g. Tong, Hung, and Cheung (1999). A well-known example is the New European Drive Cycle (NEDC), which was considered to represent the average use of cars in European cities. Legislation defines such drive cycles in order to provide a common basis for establishing emission standards. For manufacturers it is obviously convenient to test their vehicles by re-enacting the defined drive cycles on testbenches in their laboratories instead.
of undertaking expensive and not reproducible real-world test drives en masse—see e.g. Lyons et al. (1986). Evidently, when the whole powertrain calibration is based on one drive cycle, the appropriateness of the considered cycle is paramount. Unfortunately, traffic characteristics differ from region to region and even within the same city. Consequently, many different cycles for different places and procedures have been suggested for how to obtain these profiles, for example in Tong, Hung, and Cheung (1999), Lyons et al. (1986) and Ergeneman, Sorusbay, and Gökatan (1997).

Nevertheless, the obvious pitfall of drive cycle based optimization is the risk of overfitting the calibration to the specific cycle—see Eriksson and Nielsen (2014). This leads to calibrations that perform well on the cycle but potentially poorly when the driver deviates from it and the ambient conditions in the real world differ from the testbed environment. In fact, while legislative type-approval emission values have decreased significantly in the last decades, real driving emissions (RDE) did not diminish as much, as stated in Mock et al. (2012). For example in Ekberg, Eriksson, and Sivertsson (2016) it is demonstrated that the allowed deviations from the prescribed time–speed profile of the NEDC alone allow for significant differences in optimal fuel consumption. Therefore, it is reasonable to expect quite different optimal calibrations from heterogeneous drive cycles. Furthermore, in real operation, various drivers following the same route can exhibit quite a range of different driving patterns that lead to disparate emission and fuel consumption profiles. In Mensing, Trigui, and Bideaux (2011) the authors construct eco-drive cycles, which cover the same distance and contain the same stops, but with less fuel consumption while not changing the drivetrain configuration.

In this work optimal calibration of engines is considered in particular, even though the presented methods are not at all limited to that application. The engine calibration process generally consists of the optimal specification of engine tuning parameters, like exhaust gas recirculation or air flow, that minimize certain objectives while adhering to constraints like maximum emissions on a certain drive cycle. However, the optimal values depend on the engine’s speed and torque. These mappings from the operating space (i.e. all possible combinations of engine speed and torque) to the optimal tuning parameters are called engine maps, which are stored in the engine control unit (ECU) of the vehicle. For drivability and various other technical reasons, the parameters are not allowed to change abruptly in transient operation when the engine is moved from one operating point to another. This means that the engine maps have to fulfill certain smoothness conditions. In a nutshell, the calibration process therefore usually consists of the three following steps—see e.g. Langouët et al. (2008) and Isermann and Sequenz (2016).

1. Select a finite number of weighted operating points in the operating space that are supposedly representative for the drive cycle.
2. Optimize the engine parameters on each operating point.
3. Fill the engine maps under certain smoothness conditions based on the optimal values at the defined operating points and drive cycle-constraints.

Because of the smoothing step, the final mapping usually differs from the optimal results at the selected operating points yielding suboptimal configurations. Engine calibration is a prime example of the great influence which the definition of the drive cycle exerts on the outcome. Taking all this into account, the industry seeks to incorporate stochastic elements in the drive cycle generation as in Eriksson and Nielsen (2014) and Schwarzer, Ghorbani, and Rocheleau (2010). Furthermore, identifying a worst case drive cycle in terms of emissions and basing the calibration on that cycle is also problematic because it is unclear how to define and identify a worst case. Under varying stochastic influence, different cycles could lead to the highest emissions. Even if only a limited number of cycles is considered, the notion of worst case cycle is obscure since the emissions depend on the ECU calibration. With different calibrations different cycles can be the worst ones. To sum up, as long as the calibration depends on one single drive cycle, the above mentioned problems remain.
The more holistic approach applied in this article is stochastic optimization wherein the engine calibration is not obtained from optimizing some criterion for a single drive cycle but for a family of various cycles. The proposed workflow including the related uncertainties that can be considered in addition is illustrated in Figure 1. Besides the cycle and ambient conditions, the underlying models introduce other sources of uncertainty into the calibration process. In particular, the extrapolation quality of the models is uncertain when evaluated under varying conditions. For example in Lee et al. (2018) it is shown that, even with a very sophisticated Hardware-in-the-Loop approach, inaccuracies of 5–10% of the model output remain. In this article, however, the focus lies on cycle related uncertainties. Nevertheless, model inaccuracies could be considered in the proposed calibration process in a similar manner.

In general, stochastic programming is applied when random events influence the problem at hand. A detailed introduction to stochastic optimization can be found in Birge and Louveaux (2011) and Shapiro, Dentcheva, and Ruszczyński (2009). Stochastic programs usually deal with objective functions $G$ that are not only a function of the decision variable $u$ but also of a random variable $Y$. Clearly, a minimum (or maximum) of such a function $G(u, Y)$ with respect to the decision variable $u$ is not defined without knowing the realization of $Y$. Therefore, it is necessary to consider deterministic reformulations of the stochastic program. $G(u, Y)$ is a random variable itself and has a distribution which can be shaped by the choice of the decision variables. One possibility is the so-called horse-tail matching described in Cook and Jarrett (2018). In this approach, the deterministic optimization goal is the difference between the cumulative distribution function of $G(u, Y)$ and a desired target. Another obvious way of relating a stochastic to a deterministic program is to think about statistical

![Workflow Diagram](image-url)

**Figure 1.** A sample of drive cycles is used in order to incorporate the influence which the driver, traffic or route exert on the possible real-world velocity profile. These cycles are simulated considering varying ambient conditions in order to obtain a sample of trajectories through the engine’s operating space. These trajectories are then used to perform a stochastic optimization, which could also include further engine model related uncertainties. Uncertainties which are not explicitly considered in this work are greyed out.
parameters as in Solonen and Haario (2012). The most basic option is the expected value $E[G(u, Y)]$. The expected value, being a deterministic quantity, provides a simple way to transform the stochastic problem to a deterministic one. The resulting solution $u^*$ is such that the average of many realizations of $G(u^*, Y)$ will be better than the average of many realizations with $u \neq u^*$. Minimizing the expected value is for example considered in Moura et al. (2011), where stochastic dynamic programming is deployed on a family of drive cycles for optimal power management in hybrid electric vehicles. In Kolmanovsky, Siverguina, and Lygoe (2002), a Markov chain drive cycle generator is used as a basis for a stochastic dynamic programming algorithm that minimizes the expected value of a weighted sum of fuel consumption and emissions based on the given transition probabilities of the Markov process. Such an approach yields engine maps that are not necessarily optimal for one individual drive cycle but perform better on average when various drive cycles are evaluated.

However, returning to the question of fuel consumption and emission bounds, controlling worst case outcomes might be more interesting than optimizing the average performance. For even if the calibration has minimal average emissions, the emissions can still be very high in certain scenarios and have an unacceptable variance. Therefore, it makes sense to consider statistics that describe the relevant characteristics of the objective value distribution in a more accurate way than the expected value does. In particular, the right tail of the distribution which contains the worst case scenarios is of interest.

Suitable statistics regarding the right tail of probability distributions can be found in the area of risk management. Financial institutions like banks and insurance companies use statistics called ‘risk measures’ in order to quantify and control their financial exposure in order to ensure financial stability and solvency—see e.g. Eberlein et al. (2007) and Mitra and Ji (2010). Risk measures also play a role in portfolio optimization where the investor's goal is usually to maximize return while minimizing risk—see e.g. Gambrah and Pirvu (2014). In this work, these statistics are used in order to develop a method to control and minimize the risk of ‘too high’ emissions and fuel consumption given various drive cycles. This will lead to significantly more robust calibrations.

The article is structured as follows: in Section 2, the notions of risk measurement and optimization of risk are introduced. In Section 3, the considered optimization problem is described in detail. A comprehensive discussion of the results is presented in Section 4.

2. Risk measures in stochastic optimization

The necessary statistical fundamentals are presented in the following. As mentioned above, the goal of this work is to consider drive cycles as realizations of a random variable and optimize the shape of the distribution rather than one exclusive cycle alone. This involves characterizing the uncertainty by means of a real number so that the stochastic problem can be transformed into a deterministic one, which can be solved with standard methods. This quantification of unfavourable outcomes and how the risk aversion of the decision maker determines an adequate measurement procedure will be discussed in this section.

2.1. Risk measures

Risk, in its most general form, can be defined as ‘the possibility of deviation from the expected’ (Kolman 1999). Risk management, therefore, strives to take uncertainties about future adverse effects into account in decision making. In the finance industry, risks originate for example from credit defaults. Managing credit risk involves holding an adequate capital buffer. ‘Adequate’ here means: enough to ensure solvency but as little as possible in order to minimize capital cost. Risk management is also applied when it comes to environmental issues. For example in flood control, the optimal dimensions of a dyke have to be determined such that costs are minimal but flood protection is adequate for most weather events that occur randomly. In engine calibration the question is: ‘How can a calibration be
found that performs well not only in one defined setting but in a broad variety of possible situations?  

All this boils down to the problem of how to quantify risk. Most modern risk measures are based on the distribution of the considered quantity (loss, flood height, cost, . . .)—see e.g. Eberlein et al. (2007). A risk measure is defined here as an operator $\rho$ that maps a random variable to a real number. A trivial example is the expected value. Some other well-known measures that will be considered in this work are presented in the following definition.

**Definition 2.1:** Let $X$ be a continuous random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\alpha \in (0,1)$.

(i) The *Value-at-Risk (VaR)* at the confidence level $\alpha$ is defined as

$$\text{VaR}_\alpha (X) := \inf \{ x \in \mathbb{R} : \mathbb{P} [X \leq x] \geq \alpha \}.$$  

(ii) The *Conditional Value-at-Risk (CVaR)* at the confidence level $\alpha$ is defined as

$$\text{CVaR}_\alpha (X) := \mathbb{E} [X | X \geq \text{VaR}_\alpha (X)].$$  

(iii) The *Entropic Value-at-Risk (EVaR)* at the confidence level $\alpha$ is defined as

$$\text{EVaR}_\alpha (X) := \inf_{z > 0} \left\{ z^{-1} \ln \left( \frac{M_X(z)}{(1 - \alpha)} \right) \right\},$$

where $M_X(z) = \mathbb{E}[e^{zX}]$ is the moment-generating function.

It is worth noting that the exact formulations of Definition 2.1 generally differ slightly from author to author—see e.g. Mitra and Ji (2010) and Acerbi and Tasche (2002). VaR is definitely the most popular risk measure and many regulatory requirements in the finance sector are based upon VaR. The concept is easy to interpret as the VaR is simply the $\alpha$-quantile of the distribution of $X$, which means that realizations of $X$ exceed $\text{VaR}_\alpha$ with probability $1-\alpha$. In other words, if $X$ describes costs, $\text{VaR}_\alpha$ is the smallest cost of the $(1-\alpha) \times 100\%$ worst realizations or, equivalently, the highest cost of the $\alpha \times 100\%$ best realizations. Figure 2 illustrates the discussed risk measures based on a $\chi^2$-distribution. The area below the density function left of the VaR, i.e. the probability that the random variable takes on a smaller value, is $\alpha = 0.9$. The area to the right of $\text{VaR}_\alpha$ is $1-\alpha = 0.1$.

CVaR is the expected value of $X$ under the condition that $X$ exceeds the corresponding VaR. Therefore, the inequality $\text{VaR}_\alpha \leq \text{CVaR}_\alpha$ holds trivially. While VaR only gives a threshold above which costs are considered ‘bad’, CVaR also provides information about ‘how bad it gets in case something bad happens’, which is one source of criticism of the use of VaR in risk management. CVaR is sometimes also referred to as *Expected Shortfall* or *Average Value-at-Risk*. The last name is motivated through the following equality (which can also serve as a definition):

$$\text{CVaR}_\alpha (X) = \frac{1}{1-\alpha} \int_{\alpha}^{1} \text{VaR}_\beta (X) \ d\beta.$$  

Therefore, CVaR can be seen as an average VaR over all confidence levels above $\alpha$.

EVaR is a relatively new risk measure and an important representative of the class of $g$-entropic risk measures—see e.g. Ahmadi-Javid (2011) and Ahmadi-Javid (2012). It can be motivated through Chernoff’s bound: for any constant $a \in \mathbb{R}$ and random variable $X$ with existing moment-generating
Figure 2. Density function of a \( \chi^2 \)-distribution and the corresponding statistics. The vertical lines indicate the values of the distribution's expected value, \( \text{VaR}_\alpha \), \( \text{CVaR}_\alpha \), and \( \text{EVaR}_\alpha \), respectively. The confidence level is \( \alpha = 0.9 \). The area right of \( \text{VaR}_\alpha \) has an area of \( 1 - \alpha = 0.1 \). The relation \( \text{VaR}_\alpha \leq \text{CVaR}_\alpha \leq \text{EVaR}_\alpha \) is clearly visible.

function, the following inequality holds for every \( z > 0 \):

\[
\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[e^{zX}]}{e^{za}} = e^{-za} M_X(z).
\]

Setting \( e^{-za} M_X(z) = 1 - \alpha \) and solving for \( a \) one obtains \( a(z, \alpha) = z^{-1} \ln(M_X(z)/(1 - \alpha)) \) and \( \mathbb{P}[X \geq a(z, \alpha)] \leq 1 - \alpha \) for every \( z > 0 \). From the last inequality the relation \( \mathbb{P}[X \leq a(z, \alpha)] \geq \alpha \) follows, which shows that \( a(z, \alpha) \) is an upper bound for \( \text{VaR}_\alpha \). Minimizing this upper bound \( a(z, \alpha) \) over \( z > 0 \) yields the definition of \( \text{EVaR} \). It can be shown that \( \text{CVaR}_\alpha \leq \text{EVaR}_\alpha \) always holds.

An important class of risk measures with favourable and reasonable properties are the so-called coherent risk measures—see e.g. Artzner et al. (1999).

**Definition 2.2:** Let \( X \) and \( Y \) denote two random variables. A risk measure \( \rho \) is coherent if it satisfies the following conditions:

(i) translation invariance: \( \rho(X + c) = \rho(x) + c \) for any \( c \in \mathbb{R} \);  
(ii) monotonicity: if \( X \leq Y \) a.s., then \( \rho(X) \leq \rho(Y) \);  
(iii) positive homogeneity: \( \rho(\lambda X) = \lambda \rho(X) \) for any \( \lambda \geq 0 \);  
(iv) sub-additivity: \( \rho(X + Y) \leq \rho(X) + \rho(Y) \).

Translation invariance reflects the fact that increasing costs by a fixed amount increases the risk by the same amount. This means that shifting a distribution to the right or left has to do the same thing to the risk measure. Monotonicity requires higher costs to yield a higher risk. Homogeneity, similarly to monotonicity, implies that scaling the random variable also scales the risk by the same value. The last property, sub-additivity, is best interpreted in a finance setting, where it means that diversification of a portfolio cannot increase risk. It is pretty reasonable that a good risk measure should fulfill these properties. While \( \text{CVaR} \) and \( \text{EVaR} \) are coherent, \( \text{VaR} \) lacks sub-additivity, which is another disadvantage of \( \text{VaR} \), as stated in Ahmadi-Javid (2011) and Artzner et al. (1999). The expected value is obviously coherent.

Which measure should be used in an optimization basically depends on the risk aversion of the decision maker. Based on preferences and the characteristics of the considered problem, different
measures make more sense than others. If one is interested in achieving good results on average, then one should minimize the expected value of the distribution. However, doing so does not mean that high losses or costs are impossible. It just means that there are enough small costs to outweigh the rare extreme costs. If the user is interested in exercising more control over the right tail of the distribution, where the extreme outcomes occur, other risk measures should be considered. Optimizing VaR implies shifting the \( \alpha \times 100\% \) of the probability mass as far to the left as possible but the \((1 - \alpha) \times 100\% \) of the mass right of VaR are disregarded which means that there is no control over extreme events yet again. In contrast to that, as mentioned before, CVaR explicitly takes the shape of the tail into account. Minimizing CVaR means minimizing costs in the case of extreme events. Therefore, controlling CVaR can be seen as an optimization of the worst cases. Evidently, the resulting distribution will exhibit a larger expected value compared to the result of an expected value minimization but its tail will be thinner. Essentially, the choice of the risk measure to be optimized influences the shape of the distribution. In general, it will not be possible to get both a minimal average and a thin tail, so the risk measure of choice really depends on the priorities of the decision maker.

### 2.2. Optimization of risk measures

Suppose the random variable \( X \) that is to be optimized is a function \( X = f(u, Y) \) of a decision vector \( u \in \mathbb{R}^n \) and a random variable \( Y \) with a known distribution. The minimization problem is

\[
\min_{u \in \mathcal{U}} \rho \left( G(u, Y) \right)
\]

Several methods have been proposed to solve problem (1)—see e.g. Gambrah and Pirvu (2014), Ahmadi-Javid (2012), Rockafellar and Uryasev (2000) and Sarykalin, Serraino, and Uryasev (2008). For simple functions \( G \), the distribution of \( X \) and certain risk measures can be calculated and optimized analytically. However, for more complex models this will not be the case and the risk measures and the involved quantities have to be approximated using Monte Carlo methods. In Kreinin et al. (1998) the authors show that quasi-Monte Carlo techniques based on low-discrepancy Sobol sequences generally increase the accuracy of estimations.

The approximation quality of statistics like VaR naturally depends on the significance level \( \alpha \). The estimate is obtained by sorting the simulated samples and determining the value which is exceeded by \((1 - \alpha) \times 100\% \) of all realizations. Hence, if \( \alpha \) is close to one, the estimate only depends on very few realizations that might vary a lot if the distribution has a somewhat heavy tail. Therefore, in that case it is convenient to optimize risk measures without having to estimate VaR. For CVaR there exists an equivalent optimization program that does not include the approximation of VaR but instead depends on the expected value, which can conveniently be approximated by the sample mean. The following result is taken from Rockafellar and Uryasev (2000). For \( \rho = \text{CVaR}_\alpha \), an equivalent problem for (1) is given by

\[
\min_{u \in \mathcal{U}, \xi \in \mathbb{R}} \left\{ \xi + \frac{1}{1 - \alpha} \mathbb{E} \left[ (G(u, Y) - \xi)^+ \right] \right\},
\]

with \( x^+ = \max(0, x) \). The equivalence is meant in the following sense: let \((u^*, \xi^*)\) be the optimal solution of (2). Then \( u^* \) is also the optimal solution of (1) and \( \xi^* \) is the corresponding VaR. Note that minimizing CVaR generally also implies a small VaR since CVaR is an upper bound of VaR. Conversely, minimizing VaR can lead to an undesirably large CVaR.

In the case of \( \rho = \text{EVaR}_\alpha \), (1) can be written as

\[
\min_{u \in \mathcal{U}, t > 0} \left\{ t \ln M_{G(u, Y)}(t^{-1}) - t \ln (1 - \alpha) \right\}.
\]

Similarly to CVaR, EVaR also controls VaR and CVaR to some extent, as it poses an upper bound for both statistics.
3. Stochastic drive cycle optimization

In this section the stochastic cost minimization problem will be defined and the solution procedure will be sketched out.

3.1. Optimization formulation

Let $\mathcal{O} \subset \mathbb{R}^2$ be the operating space of an engine which contains every physically realizable (drivable) combination of engine speed and torque. Every such pair $x \in \mathcal{O}$ is called an operating point. The engine calibration is defined by a vector of $n$ tuning parameters which depend on the actual operating point. The parameter vector at operating point $x$ is denoted by $u(x) \in \mathbb{R}^n$. This means that, for each point in the operating space, the value of each tuning parameter has to be specified. This information is stored in the mappings

$$
\Psi_i : \mathcal{O} \to \mathbb{R} \quad x \mapsto u_i(x) \quad i = 1, \ldots, n.
$$

These functions $\Psi_i$ are called engine maps. The ultimate goal will be to determine these maps in an optimal fashion.

The minimization objective considered in this article is the accumulated costs stemming from random drive cycles. The costs arise from fuel consumption and exhaust gas emissions during the completion of one cycle. The costs are calculated as the price per unit times the amount of fuel consumption or emissions, respectively. However, an upper bound $S > 0$ is assumed under which emissions are free. $S$ is therefore the amount of permitted cumulated emissions for the given drive cycle. Emissions below $S$ do not incur costs, but every surplus above $S$ is penalized. The price of fuel consumption per unit is denoted by $p_{\text{fuel}}$, the price of excess emissions per unit is called $p_{\text{emission}}$. The prices are exogenous variables. On the contrary, the amounts of fuel consumed and exhaust gases emitted are clearly dependent on the operating point $x$ and the engine parameter vector $u$. These amounts, denoted by $f_{\text{fuel}}(x, u)$ and $f_{\text{emission}}(x, u)$, are obtained through data-driven modelling of engine measurement data. Note that the cost structure assumed here is just one example, and in general arbitrary cost formulations are possible.

For a given continuous-time drive cycle, a discrete representation in the operating space is defined by the pair $\mathcal{D} = (\mathcal{X}, \mathcal{W})$ with a discrete subset $\mathcal{X} = \{x_1, \ldots, x_N\} \subset \mathcal{O}$ and a corresponding vector of weights $\mathcal{W} = [w_1, \ldots, w_N] \in \mathbb{R}^N$, rating the importance of each operating point $x_i$. Finally, by denoting $u^i = u(x_i)$, the cost $C$ of the drive cycle $\mathcal{D}$ takes the form

$$
C(\mathcal{D}) = \sum_{i=1}^{N} w_i f_{\text{fuel}}(x_i, u^i) p_{\text{fuel}} + \varphi \left( \max \left\{ 0, \sum_{i=1}^{N} w_i f_{\text{emission}}(x_i, u^i) - S \right\} \right) p_{\text{emission}}. \tag{3}
$$

The penalty function $\varphi$ determines how excess emissions are evaluated. Clearly, $\varphi$ has to be a monotonically increasing, non-negative function with $\varphi(0) = 0$. Possible choices are for example $\varphi(y) = y$ or $\varphi(y) = y^2$.

The maps $\Psi_i$ are required to fulfill certain smoothness constraints. These could take the form of conditions on the first and/or second directional derivative of the maps. By virtue of difference approximation of the derivative this yields linear constraints that couple the operating points. This means that the optimal choice of $u^i$ is not independent of $u^k$. However, for the moment it suffices to condense all constraints on the engine tuning parameters as $(u^1, \ldots, u^N) \in \mathcal{U}$.

As mentioned before, the goal is not to optimize equation (3) for a single drive cycle $\mathcal{D}$ but rather for random drive cycles $\mathcal{D}_Y$ with an underlying random variable $Y$. When the drive cycle is random, so is the cost function $C(\mathcal{D}_Y)$. This random variable is now optimized with respect to the risk measure $\rho$. In other words, rather than optimizing the cost for an individual cycle, the probability distribution
of the costs is shaped by optimizing a statistical parameter. Therefore, the considered optimization problem is formulated as

\[
\min_u \rho(C(D_Y)) \\
\text{s.t. } (u^1, \ldots, u^N) \in \mathcal{U}.
\] (4)

The operator \( \rho \) stands either for the expected value or for another adequate risk measure as defined in Definition 2.1 and has a tremendous effect on the solution. The actual choice of \( \rho \) is based on the preferences and risk aversion of the decision maker as discussed in Section 2.

### 3.2. Optimization

In order to evaluate the objective function, a sample of drive cycles represented in the operating space of the engine as \( \mathcal{D} = (\mathcal{X}, \mathcal{W}) \) is needed. The sample used in this work is based on measurements of a real-world drive where detailed data on velocity, covered distance, road gradient and road curvature were recorded. Altering the velocity data, while keeping the other signals the same, leads to different drive cycles that still cover the same distance on the same track. The variations in velocity involve faster and slower driving, earlier braking and random fluctuations. Generally though, arbitrarily many real-world drives could be incorporated in the sample generation. However, the obtained speed profiles cannot be used for the optimization directly, because the engine maps \( \Psi_i \) depend on the operating space \( \mathcal{O} \) and not on velocity. Therefore, in order to obtain the trajectories in the engine's operating space, the profiles defined by distance, velocity, road gradient and road curvature were simulated in a virtual car simulation environment. Moreover, ambient conditions like wind were altered in the simulations as well, which also results in different operating space trajectories that represent the same real-world route. This workflow is also illustrated in Figure 1.

The obtained operating space trajectories, depicted in Figure 3, all have their own set of operating points, which is not compatible with the formulation of problem (4). Therefore, in order to obtain common decision variables, the operating points are projected on predefined operating points (nodes). Each point of a trajectory is projected onto the nearest node based on the Euclidean metric as illustrated in Figure 4.

![Figure 3. A family of trajectories through the operating space. All trajectories are obtained through simulation of a measured track with varying velocity profiles and ambient conditions. Therefore, each trajectory represents the exact same route but with different driving behaviour and surroundings. The bold bounding line indicates the operating space \( \mathcal{O} \).](image-url)
Figure 4. The circles show the considered nodes $x_i$. Each operating point of the trajectory represented by a small dot is projected onto the nearest node, indicated by a star. The bold line indicates the operating space $O$.

Figure 5. Two exemplary sets of weights. The weights of one trajectory are represented by circles, while the weights of the other trajectory are represented by crosses. The weights $w_i$ count the number of projections on the respective node $x_i$. The bold line indicates the operating space $O$.

Additionally, by means of the projection, the vector of weights $W$ that is used for the evaluation of the cycle-cost (3) is obtained for each trajectory. The weights for the corresponding sample trajectory can be seen in Figure 5: the weight $w_i$ of each node $x_i$ is simply the number of operating points that are projected on $x_i$. If a trajectory spends a lot of time near a certain node, the node will have a higher weight and the calibration at this node becomes more important. The nodes are arranged as a regular grid on the operating space but this is not a restriction and the choice of nodes can take any form. For example, a-priori knowledge could be exploited concerning areas of the operating space that are especially relevant for fuel consumption and emissions. The number of nodes evidently determines the size of optimization problem (4). If the trajectories do not fill the entire operating space, this will give some nodes the weight zero. Optionally, instead of assigning the weight zero to those nodes, a minimal weight can be defined, so that the entire operating space is considered in the optimization at least to some extent. Note that the sum of weights is not necessarily equal for different trajectories.
This is because the original discrete representation of the trajectories do not contain the same amount of operating points since for example a more aggressive driver needs less time to finish the track and has therefore a shorter trajectory in the operating space.

The functions $f_{\text{fuel}}$ and $f_{\text{emission}}$ are obtained through local quadratic model networks, a modelling technique that has found broad application in engine calibration. The model architecture is described in Hametner and Jakubek (2013). The considered models in this work are static but the risk-averse optimization is in principle not limited to the static case. For modelling of dynamic local model networks, see Hametner, Mayr, and Jakubek (2014). Generally, it has to be pointed out that the modelling and functional forms do not interfere with the applicability of the presented approach, as long as a suitable optimization solver is available. Fast model evaluation is naturally favourable, though. The modelling is based on real measurement data from a diesel engine. Emission here refers to nitrogen oxide. Some examples of measurements from the testbed can be seen in Figure 6.

Six engine parameters are considered for optimization: air mass flow, boost pressure, exhaust gas recirculation, main timing, rail pressure and swirl position. Furthermore, the prices of fuel, excess emissions and the threshold $S$ are set to fixed values. The penalty $\varphi$ is chosen to be the identity function.

There are two kinds of constraint: upper and lower bounds on the decision variables and smoothness constraints. The upper and lower bounds again depend on the operating point. In terms of smoothness, gradient constraints between neighbouring nodes $x_j$ and $x_k$ of the form

$$\frac{|u_j^i - u_k^i|}{\|x_j - x_k\|_2} \leq k_{\text{max}}^i$$

are considered, where $k_{\text{max}}^i$ is the maximum admissible slope of the $i$th engine map $\Psi_i$. The nonlinear constraint (5) is transformed into two linear constraints where the absolute value in the numerator is

![Figure 6. Nitrogen oxide and fuel consumption measurements from a diesel engine testbed.](image)
replaced by a plus and a minus sign, respectively. Therefore, the smoothness constraint can be written in the practical linear form $Au \leq k$ with an appropriate matrix $A$, a decision vector $u$ containing all $u_i$ and a vector $k$ containing the maximum admissible slopes for each engine map. Note that the inclusion of other types of constraint is also possible in the presented approach. For example, more complex data hulls of measured feasible inputs can be necessary to assess the physical feasibility of decision variables as in Didcock et al. (2018) and can therefore pose additional constraints on the optimization.

Problem (4) can now be solved using an adequate solver. The set of different trajectories, which are ultimately defined by the corresponding vector of weights $W$, represent the random variable $Y$. For each realization of this random variable, i.e. for each trajectory, the cost function (3) has to be evaluated for a fixed decision vector $u$. Afterwards, the risk measure $\rho$ is approximated, which yields the current evaluation of the objective function.

4. Results and discussion

In this section the results of the stochastic drive cycle optimization will be presented and discussed. The considered risk measures are expected value, $\text{VaR}_\alpha$, $\text{CVaR}_\alpha$ and $\text{EVAR}_\alpha$, with $\alpha = 0.90$. The minimum weight of unused nodes is set to $w_{\text{min}} = 1$. For each choice of risk measure $\rho$, the optimization yields an individual calibration. The various results can be compared by simulating the calibrations on a set of validation cycles. This validation set was obtained in the same way as the cycles that were used for the optimization but with different random permutations of the underlying real drive cycle. The resulting samples of cost function realizations were then used to estimate the risk measures corresponding to the respective calibration. Furthermore, the deterministic cost function based on a single cycle was optimized in order to demonstrate the effects of using risk measures in the optimization. Here, one could suspect that the selected cycle for the single-cycle optimization was simply not representative for the route and the whole set of trajectories. Therefore, the average cycle of the set was constructed by averaging the weights of each node. The resulting cycle therefore incorporates the information of all trajectories and can also be used to perform a single-cycle optimization.

In Figure 7, kernel density estimations, expected values and CVaR of the cost distributions for all calibrations are depicted. Additionally, the values of the considered statistics are listed in Table 1 for each calibration. From Figure 7 it is evident that the various calibrations obtained through the optimization of different risk measure objectives impact the shape of the cost distribution: the first two panels show the cost distribution stemming from the non-stochastic single-cycle optimizations. These distributions exhibit a much larger variance and heavier right tails compared to the rest of the calibrations. The results from the stochastic optimization approach are clearly superior: minimizing the expected value yields a calibration that shifts the mass of the cost distribution to the left resulting in the smallest expected value of costs. The approaches with the highest risk aversion are CVaR and EVaR and the results are very similar to one another as both risk measures primarily consider the far right tail of the cost distribution. The mass of the distribution is shifted to the right but therefore the right tail gets thinner, which reflects the higher risk aversion of these risk measures. In other words, this means that higher risk aversion leads to higher average costs but lower costs in the worst cases. Finally, the calibration obtained through minimizing VaR is revealed to be a compromise solution where the distribution's mass is moved to the left but the tail gets heavier in return.

Additionally, Table 1 shows that consistent results are obtained when evaluated on the validation cycles: the smallest value of each risk measure is achieved through the calibration that was supposed to minimize the respective risk measure. Moreover, the results listed in Table 1 reveal that using the averaged representative cycle leads to a suboptimal calibration in every respect. All risk measures obtained by this calibration are worse than the ones obtained through minimization of the expected value. Additionally, this approach is not feasible for a higher risk aversion since the values of CVaR and VaR are significantly worse than obtained through minimization of the respective risk measure.
Figure 7. Kernel density estimations of the cost distribution for each calibration, which was evaluated on the same set of validation cycles. The solid line indicates the expected value while the dashed line indicates CVaR of each distribution. For the other statistics, refer to Table 1. It can be seen that using only one cycle leads to a larger variance and more high costs (top two panels). Minimizing the expected value moves the mass of the distribution (and the expected value itself) to the left (third panel) and minimizing CVaR and EVaR yields a thinner right tail (panels four and five). VaR optimization is a compromise solution and lies in between (bottom panel).

The reason for this result is that the average of the cycle specification in the operating space is simply not the same as the average in outcome. Moreover, the possibility of emission costs surpassing the threshold level $S$ cannot be accounted for in a one-cycle optimization. Similarly, just stacking the different cycles together into one large cycle will yield different, inferior results because the high emissions of some parts of the cycle can be outweighed and therefore disregarded, which is not possible in the stochastic optimization approach, where the distribution of the costs rather than the aggregated cost and all possible threshold violations are considered.
Table 1. Evaluated risk measures of the resulting cost distributions depicted in Figure 7. Each column represents one calibration defined by the optimization objective. The minimal value of each statistic among all calibrations is printed in bold. All validated results are consistent, which means that each risk measure indeed achieves its minimum through the respective calibration. In the case of EVaR, the minimum is also attained by the CVaR calibration.

| Objective | Single cycle | Average cycle | Expected value | CVaR | EVaR | VaR |
|-----------|--------------|---------------|----------------|------|------|-----|
| Expected value | $3.0080 \times 10^4$ | $2.9713 \times 10^4$ | $2.9660 \times 10^4$ | $3.0056 \times 10^4$ | $3.0080 \times 10^4$ | $2.9760 \times 10^4$ |
| VaR$_{0.9}$ | $3.3470 \times 10^4$ | $3.2667 \times 10^4$ | $3.2330 \times 10^4$ | $3.2240 \times 10^4$ | $3.2284 \times 10^4$ | $3.2014 \times 10^4$ |
| CVaR$_{0.9}$ | $3.5184 \times 10^4$ | $3.4411 \times 10^4$ | $3.4003 \times 10^4$ | $3.3476 \times 10^4$ | $3.3484 \times 10^4$ | $3.3660 \times 10^4$ |
| EVaR$_{0.9}$ | $3.5890 \times 10^4$ | $3.5137 \times 10^4$ | $3.4748 \times 10^4$ | $3.4186 \times 10^4$ | $3.4186 \times 10^4$ | $3.4407 \times 10^4$ |

The results presented in Figure 7 and Table 1 show the important role of considering and defining the desired risk aversion. It cannot be deduced that either the expected value or the CVaR minimization is superior to the other because both calibrations have strengths and weaknesses. It all boils down to the question of what cost structure is more favourable, which might differ from case to case. In applications where one wants to make sure no thresholds are exceeded, one should minimize CVaR with an adequately chosen significance level $\alpha$. On the other hand, if surpassing certain values is not very critical, one can focus on the average outcome and minimize the expected value.

The differences in threshold adherence can best be seen in Figure 8: the bar chart illustrates what proportion of the validation cycles experiences a violation of the emission threshold. This point of view focuses exclusively on the emission values and not on the statistics of the total cost as before. Nevertheless, the results show the same structure in terms of risk aversion. Optimizing individual cycles leads to calibrations which have a large probability of violating emission boundaries. On the contrary, stochastic optimization reduces that risk significantly. As expected value optimization does not focus on extreme events, the number of excesses is quite high compared to the risk-averse CVaR minimization. The presented probabilities obviously also depend on the parameters of the cost function $C$ and emission threshold violations can also be steered through higher emission costs in the objective. However, without including stochastic variations this will not necessarily help the single-cycle optimizations, because the resulting calibration will typically spend the whole number of allowed emissions anyway and therefore the violation risk is independent of the emission penalty.

Instead of changing the risk measure, the risk aversion can also be adjusted by altering the confidence level $\alpha$. A higher value of $\alpha$ implies focusing on the few worst cases, while a smaller value of $\alpha$
Table 2. Evaluated CVaR for different confidence levels $\alpha$. Each column represents one calibration obtained through minimizing CVaR with the given confidence level. The minimal value of each statistic among all calibrations is printed in bold. Most validated results are consistent. The expected value is minimal for the case with lowest risk aversion ($\alpha = 70\%$), which shows again that low probabilities of high costs come at the price of slightly larger average costs.

| Confidence level, $\alpha$ (%) | 75       | 80       | 85       | 90       | 95       |
|--------------------------------|----------|----------|----------|----------|----------|
| CVaR$_{0.75}$                 | $3.2169 \times 10^4$ | $3.2174 \times 10^4$ | $3.2180 \times 10^4$ | $3.2234 \times 10^4$ | $3.2308 \times 10^4$ |
| CVaR$_{0.80}$                 | $3.2523 \times 10^4$ | $3.2509 \times 10^4$ | $3.2512 \times 10^4$ | $3.2551 \times 10^4$ | $3.2611 \times 10^4$ |
| CVaR$_{0.85}$                 | $3.2945 \times 10^4$ | $3.2913 \times 10^4$ | $3.2914 \times 10^4$ | $3.2948 \times 10^4$ | $3.3011 \times 10^4$ |
| CVaR$_{0.90}$                 | $3.3539 \times 10^4$ | $3.3480 \times 10^4$ | $3.3477 \times 10^4$ | $3.3476 \times 10^4$ | $3.3510 \times 10^4$ |
| CVaR$_{0.95}$                 | $3.4419 \times 10^4$ | $3.4326 \times 10^4$ | $3.4314 \times 10^4$ | $3.4250 \times 10^4$ | $3.4236 \times 10^4$ |
| Expected value                | $2.9824 \times 10^4$ | $2.9915 \times 10^4$ | $2.9934 \times 10^4$ | $3.0056 \times 10^4$ | $3.0172 \times 10^4$ |

uses a broader definition of ‘worst case’. The results of CVaR optimizations with different confidence levels $\alpha$ are listed in Table 2. The figures can be interpreted in the same way as before: higher risk aversion (large $\alpha$) yields better CVaR values for larger confidence levels but worse results for smaller confidence levels and also worse expected values. This means that increasing $\alpha$ again shifts the mass of the distribution to the right but decreases the probability of large costs in return.

5. Conclusions

This article presents a novel approach for drive cycle based engine calibration. Instead of basing the optimization on just one drive cycle, it is proposed rather to perform stochastic optimization on a sample of random drive cycles. For this purpose, the random objective function is transformed into a deterministic value by applying certain risk measures that describe various properties of the probability distribution of the considered random variable that has to be optimized. The choice of the appropriate risk measure is problem dependent and is also affected by the risk aversion of the decision maker. While minimizing the expected value yields good results on average, more risk-averse measures such as CVaR and EVaR focus primarily on the right tail of the distribution where the worst case scenarios occur. Generally, more risk-averse approaches yield higher average costs but a lower probability of extreme costs compared to less risk-averse measures. In contrast to single-cycle optimization, the stochastic approach considers the possibility of varying drive cycles and therefore reduces the risk of exceeding drive cycle constraints when testing a calibration in the real world complying with RDE legislation.

Acknowledgement

The authors acknowledge the TU Wien University Library for financial support through its Open Access Funding Program.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Financial support by the Christian Doppler Research Association is gratefully acknowledged.

References

Acerbi, Carlo, and Dirk Tasche. 2002. “Expected Shortfall: A Natural Coherent Alternative to Value at Risk.” Economic Notes 31 (2): 379–388. doi:10.1111/1468-0300.00091

Ahmadi-Javid, Amir. 2011. “An Information-Theoretic Approach to Constructing Coherent Risk Measures.” In Proceedings of the 2011 IEEE International Symposium on Information Theory (ISIT), 2125–2127. Piscataway, NJ: IEEE. doi:10.1109/ISIT.2011.6033932.
https://www.researchgate.net/publication/288267481_Discrepancies_between_type-approval_and_real-world_fuel_consumption_and_CO2_values.

Moura, Scott Jason, Hosam K. Fathy, Duncan S. Callaway, and Jeffrey L. Stein. 2011. “A Stochastic Optimal Control Approach for Power Management in Plug-In Hybrid Electric Vehicles.” IEEE Transactions on Control Systems Technology 19 (3): 545–555. doi:10.1109/TCST.2010.2043736

Rockafellar, R. Tyrrell, and Stanislav Uryasev. 2000. “Optimization of Conditional Value-at-Risk.” Journal of Risk 2 (3): 21–41. doi:10.21314/JOR.2000.038

Sarykalin, Sergey, Gaia Serraino, and Stan Uryasev. 2008. “Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization.” In INFORMS TutOrials in Operations Research, 270–294. Catonsville, MD: Institute for Operations Research and the Management Sciences (INFORMS). doi:10.1287/educ.1080.0052

Schwarzer, Volker, Reza Ghorbani, and Richard Rocheleau. 2010. “Drive Cycle Generation for Stochastic Optimization of Energy Management Controller for Hybrid Vehicles.” In Proceedings of the IEEE International Conference on Control Applications (CCA 2010), 536–540. Piscataway, NJ: IEEE. doi:10.1109/CCA.2010.5611150.

Shapiro, Alexander, Darinka Dentcheva, and Andrzej Ruszczyński. 2009. Lectures on Stochastic Programming: Modeling and Theory. Philadelphia, PA: SIAM. https://www2.isye.gatech.edu/people/faculty/Alex_Shapiro/SPbook.pdf.

Solonen, Antti, and Heikki Haario. 2012. “Model-Based Process Optimization in the Presence of Parameter Uncertainty.” Engineering Optimization 44 (7): 875–894. doi:10.1080/0305215X.2011.617817

Tong, H. Y., W. T. Hung, and C. S. Cheung. 1999. “Development of a Driving Cycle for Hong Kong.” Atmospheric Environment 33 (15): 2323–2335. doi:10.1016/S1352-2310(99)00074-6