Representative surrogate problems as test functions for expensive simulators in multidisciplinary design optimization of vehicle

Representative surrogate problems as test functions for expensive simulators in multidisciplinary design optimization of vehicle...

Article *in* Structural and Multidisciplinary Optimization · May 2016

DOI: 10.1007/s00158-016-1410-9

Ramses Sala
University of Florence

Niccolò Baldanzini
University of Florence

Marco Pierini
University of Florence

Authors accepted manuscript. The final publication is available at Springer via
http://dx.doi.org/10.1007/s00158-016-1410-9

Representative Surrogate Problems as test functions for expensive simulators in multidisciplinary design optimization of vehicle structures

Ramses Sala[1], **Niccolò Baldanzini**, **Marco Pierini**
Dipartimento di Ingegneria Industriale, Università degli Studi di Firenze,
Via di Santa Marta 3, 50139 Firenze, Italy

Abstract
A large variety of algorithms for multidisciplinary optimization is available, but for various industrial problem types that involve expensive function evaluations, there is still few guidance available to select efficient optimization algorithms. This is also the case for multidisciplinary vehicle design optimization problems involving, e.g., weight, crashworthiness, and vibrational comfort responses. In this paper, an approach for the development of Representative Surrogate Problems (RSPs) as synthetic test functions for a relatively complex industrial problem is presented. The work builds on existing sensitivity analysis and surrogate data generation methods to establish a novel approach to generate surrogate function sets, which are accessible (i.e. not resource demanding) and aim to generate statistically representative instances of specific classes of industrial problems. The approach is demonstrated through the construction of RSPs for multidisciplinary optimization problems that occur in the context of structural car body design. As a "proof of concept" the RSP approach is applied for the selection of suitable optimization algorithms, for several problem formulations and for a meta-optimization (i.e. an optimization of the optimization algorithm parameters) to increase optimization efficiency. The potential of the approach is demonstrated by comparing the efficiency of several optimization algorithms on an RSP and an independent simulation-based vehicle model. The results corroborate the potential of the proposed approach and significant performance gains in optimization efficiency are achieved. Although the approach is developed for the particular application presented, the approach is described in a general way, to encourage readers to use the gist of the concept.

Keywords: Multidisciplinary Design Optimization, test problems, benchmarking, meta-optimization, vehicle design.

Introduction
In the design process of complex structures the search for design solutions that are effective compromises between conflicting structural objectives is a challenging task. Within the scope of automotive vehicle design, weight reduction can be conflicting with the structural requirements for disciplines fields such as NVH (Noise, Vibration and Harshness) and crashworthiness. Early feasibility studies of design optimization methods applied to automotive structures involving crashworthiness analysis on sub-structures such as those by Yang et al. (1994) and Schramm and Pilkey (1996) were published about a decade after the first numerical crashworthiness simulation by Haug et al. (1986). These feasibility studies were later followed by basic studies of Multidisciplinary Design Optimization (MDO) of full vehicle structures with respect to crash, NVH and lightweight criteria in the works of, for example, Yang et al. (2001) and Sobieszczanski-Sobieski et al. (2001). Since then the investigations and showcase studies applying various types of optimization methods on vehicle design problems have increased strongly in quantity (e.g. Baldanzini et al. (2001); Baldanzini and Scippa (2004); Durgun and Yildiz (2012); Mihaylova et al. (2012); and Yildiz and Solanki (2012)).

The application of metaheuristic and nature-inspired search algorithms to solve non-convex problems, which are related to MDO of structures with complex responses has gained increasing interest in research over the last decades (see the review articles by Venkayya (1978), Sobieszczanski-Sobieski and Haftka (1997) and, Simpson et al. (2008)). A great variety of optimization methods can be applied to non-convex MDO problems. Recent reviews on derivative-free and biologically inspired algorithms are for example given in the works of Rios and Sahinidis (2012) and Tang and Wu (2009) respectively. But which of these algorithms

[1] ramses.sala@unifi.it

should be chosen for a particular structural optimization problem? According to the “no free lunch” theorem (Wolpert and Macready 1995; 1997) all optimization or search methods perform equivalent, when averaged over all possible problems. Nevertheless certain optimization methods can perform better than others on specific problem types. So although it is not a theoretically legitimate goal to find an overall best algorithm, it is a legitimate goal to rank algorithm performance in order to find efficient algorithms for particular problem types, or to tune an algorithm for a specific type of application. In order to benchmark[2] (i.e. compare the performance and efficiency of) different optimization methods, several commonly used test functions were developed in the optimization community, Examples are the Rosenbrock function (Rosenbrock 1960), the Rastrigin’s function (Rastrigin 1974) for single objective problems, and the ZDT functions by Zitzler, Deb and Thiele (2000) for multi-objective optimization problems. Such test functions, sometimes also named “artificial landscapes”, are often expressed as simple closed form expressions, which require little computational effort such that millions of function evaluations can be achieved in a small amount of time on modern computers. It remains however a challenge to relate such standard analytical test functions to particular real world problems, and vice versa.

An alternative to analytical test problems could be the use of simulation-based structural optimization benchmarks, based on standardized problem instances. The need for more complex realistic system benchmark problems is expressed in Alimoradi et al. (2010), and a relatively recent initiative to start an open benchmark database for simulation-based multidisciplinary optimization problems with engineering relevance is presented in Varis and Tuovinen (2012). For the optimization of vehicle design problems, involving crashworthiness and NVH responses, no relevant open-source benchmark problems are available yet. Although vehicle models are made publicly available by the vehicle modeling laboratory of the National Crash Analysis Center (NCAC), none of these or other models are to the knowledge of the authors used for any standardized simulation-based optimization benchmark problems. Even if standardized simulation-based benchmark optimization problems of full vehicle models would become available in the near future, the hardware and software resources required for the computationally expensive simulations remain a big hurdle to perform the large amount of function evaluations required to obtain statistically significant performance comparisons of optimization methods and algorithms for these problem types. These difficulties also exist for other structural optimization benchmark problems that involve resource demanding simulations.

Multidisciplinary full vehicle optimizations, including crashworthiness and NVH simulation responses, can be expensive in terms of hardware and software resources (solver licenses), modeling effort and computation time. For such MDO studies in an industrial context, the computational budget is restricted to approximately 280-500 function evaluations for an optimization run on a problem, that could have more than 50 design variables, and which could require about 600 CPU hours per function evaluation if several crashworthiness load cases are involved Duddeck (2008). Although MDO using metaheuristic search algorithms is commonly applied in an industrial automotive context, the literature provides nearly no significant performance comparisons, or guidelines for efficient optimization, of more than two relevant algorithms applicable for vehicle design problems involving crashworthiness. A notable exception is the work of Duddeck (2008) in which several benchmark studies for NVH and crashworthiness related problems were presented, together with a list of search algorithm requirements on such optimization problems. Investigations in another study published by Sala et al. (2014a) indicated that the optimization algorithm performance for similar vehicle optimization problems on different vehicle models are significantly correlated, such that optimization algorithm performance characteristics can be attributed to a problem type, and not only a particular problem instance. The results of that study and additional investigations, however indicated that although meta-model or response surface based benchmarks can be effectively used for mass, and low-order eigenfrequency responses, they seem however less suited for problems involving highly nonlinear and non-smooth crashworthiness responses. The non-smoothness of the responses is due to the highly nonlinear

[2] In the automotive industry the word “benchmarking” often refers to comparative test for quality assurance of components, systems or vehicles. In the scope of this paper the word benchmarking is however used in a computing context, and refers to the act of assessing the relative performance of computational algorithms w.r.t. each other under comparable conditions.

phenomena involved such as local buckling, plasticity, contact and fracture instabilities. Under such conditions, small changes to the design variables of a system can cause bifurcations or abrupt changes in the topological behavior of the system and the resulting structural response. In a high dimensional design variable space, this highly non-smooth behavior is difficult to represent by traditional meta-models. Even the use of methods that are generally considered as suitable for nonlinear responses (such as Radial Basis Functions (RBF) and Kriging), results in smooth behavior between the construction points. Such methods can capture the "global" nonlinearities with a limited number of construction points which is useful for approximation models. However, to represent local non-smooth behavior (which can affect the optimization algorithm performance on the simulation-based model) these meta-models require a number of construction points proportional to the number of local "peaks" and "valleys" of the response "landscape" to be represented. For high dimensional models with local non-smoothness such as observed in the parameter studies presented later in this work, the construction of meta-models with representative non-smooth behavior would require a density of construction points which would currently be infeasible to achieve because of high computational costs.

For optimization problems involving crashworthiness response, and more generally for MDO involving expensive simulators with complex responses the current state of the art could be paradoxically stated as: the problems for which performance matters the most, because they are expensive and restricted to a limited evaluation budget, are also the ones for which it is too expensive to compare algorithms, tune the optimization parameters or develop specialized optimization methods.

This paper describes an approach that could contribute to extricate this paradoxical situation. The paper is structured as follows: the second section introduces the objectives and general idea of the approach to generate representative surrogate responses and optimization problems; in the third section the properties of the vehicle design optimization application example are described; the fourth section provides a brief overview of the most important methods applied for the response characterization and a summary of the characterization of the vehicle simulation responses; in the fifth section the surrogate response formulation is presented; a demonstration and result corroborations follow in section six where an RSP is used to compare optimization algorithms and an meta-optimization of the optimization algorithm parameters is performed; in the last section the approach and obtained results are discussed, followed by an outlook and conclusions.

2 The concept of Representative Surrogate Problems

In scientific literature there is much attention for the development of new meta-heuristics, while there is relatively few attention for the analysis of the problems, and their characteristics (see also Shan and Wang (2010)). The general idea of the presented approach is to construct synthetic and computationally affordable test problems based on characteristics of real world complex structural optimization problems. In the proceeding of this work these synthetic test problems will be called Representative Surrogate Problems (RSP) Note that (unlike conventional meta-model or surrogate modeling methods) the involved surrogate models and responses in this context are not intended to be used as an interpolation or approximation model of the targeted simulation responses, rather they aim to serve as a representative artificial response landscape with similar typical characteristics as the simulation-based response in a statistical sense. A RSP does not fit particular problem data, but is constructed to fit or satisfy selected characteristics of a problem type or class. The RSP approach can also be regarded as an adaptation and extension of surrogate data generation methods for time series such as proposed by Prichard and Theiler (1994) for applications with multiple correlated multivariate responses. Apart from an oral conference presentation by the authors Sala et al. (2014b), in which preliminary results of this work were discussed, this or similar approaches to construct synthetic test problems based on particular real world problems did not receive attention yet in the optimization literature.

2.1 Representative surrogate functions

In this communication a surrogate function for an individual response is denoted as a Representative Surrogate Function (RSF). An RSF is intended as a representative relationship between a model response w.r.t. its design variables. In the general case this could also be a parameterized meta-model (e.g. Kriging or

RBF based meta-models), in this work the authors however use a function representation for the RSFs is inspired by the Sobol-Hoeffding function decomposition (Hoeffding 1949; Sobol 1990).

$$T(\boldsymbol{x}) \sim \theta(\boldsymbol{x}) = \theta_0 + \sum_{i=1}^{d} \theta_i(x_i) + \sum_{1 \le i < j \le d} \theta_{i,j}(x_i, x_j) + \cdots +. \theta_{i,j,\ldots,d}(x_1, x_2, \ldots, x_d) \quad (1)$$

In equation 1, $T(\boldsymbol{x})$ and $\theta(\boldsymbol{x})$ are functions of dimension d, which can be decomposed in a series of summands of increasing interaction order. The design variable vector is denoted by the symbol $\boldsymbol{x}$, and it has elements x_i in the normalized domain of the d-dimensional unit hypercube $K^d = \{\mathrm{x} | 0 \le x_i \le 1; i = 1, \ldots, d\}$. The expression $T(\boldsymbol{x})$ refers to the targeted simulation-based response function, and $\theta(\boldsymbol{x})$ is refers to the surrogate function. In the scope of this work the symbol $\sim$ will refer to similarity according to criteria to be defined by the modeler (in the application example in section 5 a particular set of such criteria will be defined and enforced as constraint expressions). Theoretically an exact decomposition $T(\boldsymbol{x})$ exists, but in the case of expensive black box functions, and function decomposition based on a limited number of samples this is of little practical relevance. The aim is to find a parameterized truncated series expansion or another computationally affordable expression that can represent the characteristic behavior of the individual simulation responses, which is not necessarily limited to an approximation of the particular response. Depending on the response type the summands that are part of the decomposition of equation 1 (truncated in “interaction order”) could be either represented by simple analytical functions or by series expansions over the corresponding variable subset. These “second” series expansions can again be truncated in “resolution”, according to the data obtained from the response characterization. The choice for the truncation, basis functions and resulting representativeness, of such an expansion is dependent on the information obtained from the response characterization. The characteristic behavior or similarity criteria of the response output w.r.t. the design variables could involve for example the degree of nonlinearity, and the variance decomposition distribution of first and higher order interactions. The function series representation enables parameterized control over such response characteristics, whereas in data-fitting based meta-models such as RBF and Kriging surrogates, the response characteristics can only be controlled indirectly.

2.2 Representative Surrogate Systems

When more responses are involved in the optimization problems such as the case in MDO, the solution of the problem is not only dependent on the individual response characteristics but also on the relationship and structure between the different responses. A set of RSFs (superscript r in equation 2) combined with defined structure or relation between the involved responses, is denoted as an RSF-Set or Representative Surrogate System (RSS).

$$T^r(\boldsymbol{x}) \sim \theta^r(\boldsymbol{x}) = \theta_0^r + \sum_{i=1}^{d} \theta_i^r(x_i) + \sum_{1 \le i < j \le d} \theta_{i,j}^r(x_i, x_j) + \cdots + \theta_{i,j,\ldots,d}^r(x_1, x_2, \ldots, x_d) \quad (2)$$

2.3 Representative Surrogate Problems

An RSP can be defined by choosing an optimization formulation involving objectives and constraints that are depending on RSS responses. An example of a single objective optimization problem subjected to nonlinear inequality constraints could be expressed as:

$$\min f\left(\theta^r(\boldsymbol{x})\right) \text{ subject to: } g_w(\theta^r(\boldsymbol{x})) \le 0 \quad (3)$$

Where index w refers to the number of constraints. Once an RSS is established it is straight forward to test different optimization formulations on a given set of responses.

2.4 RSP construction

As can be seen from the previous definitions, the most challenging part of the RSP approach is to obtain RSFs and an RSS that is representative for the responses of interest. The general structure of the approach is: to apply parameter study and other existing sensitivity analysis methods (see section 4) to identify and quantify characteristics of the involved simulation-based responses that are common over a set of problem instances (different vehicle models in the application example). These characteristics such as nonlinearity types, sensitivity index distributions, and inter-response correlations can be used to define a constraint satisfaction problem (CSP) based on the combination of suitable basis functions with free parameters, the domain of the parameters and the constraint set that enforce the selected function characteristics (See section 5). Using the solutions of this CSP problem as a parameter set for the given basis functions will result in an RSS with a selection of similar response characteristics as the simulation-based calibration responses. The responses of the resulting RSS can be used to define a synthetic optimization problem.

The activities to construct an RSP could be summarized by the following steps:

1. response characterization
2. construction of the RSFs and the RSS by defining and solving a CSP
3. combining the optimization formulation with the resulting RSS to define an RSP
4. corroboration of the RSP.

Since the RSS and RSP are not approximative surrogates, the validation or corroboration of them can only be done indirectly by comparing the characteristics, or the performance of operators such as optimization algorithms between them, and an independent model or optimization problem instance.

2.5 Applications and general remarks

The resulting synthetic problem or RSPs could be used as a test or "toy functions" to compare, select, tune and develop efficient performing optimization algorithms and optimization frameworks for the related class of real-world optimization problems. Once established, they have a computational cost orders of magnitude less than the real problem instance. In addition they also improve the accessibility of problem types which are normally only available to a limited community because simulation-based function evaluations often require modeling expertise, solver licenses and considerable computation resources to be used in an optimization. Furthermore such RSPs could be made publicly available to serve as standardized benchmark problems, enabling an increased comparability, and reproducibility between performance studies on particular type of applications. In section 5 a schematic overview of the approach for the example case study on a multidisciplinary car body design application is provided.

Although the function characterization necessary for the formulation of an RSP requires much more function evaluations than a typical optimization run of a single problem instance in an industrial design environment, the cost of such investigations can be seen as an investment that provides increased insight into the typical response structure for similar problem types. The investment to apply the approach could pay off for practitioners that deal with many similar optimization problem instances that involve expensive simulators (such as vehicle design problems), in particular for those who aim to select or develop specialized algorithms for particular complex problem types. In the case where conventional meta-models are able to represent the response characteristics, they can replace computationally expensive "black box" simulation responses, with computationally affordable "black box" meta-models. Although this can be practical, the additional insight for a systematic analysis of the problem is rather limited. For the systematic development of optimization strategies for difficult problems, it would be useful to analyze problems by their characteristics. The nature of the proposed approach enables the investigation on the influence of different response characteristics on the performance of optimization algorithms or strategies. Such additional insight could be a further justification for the required investment in the response characterization.

3 Description of the industrial application example of multidisciplinary vehicle design optimization

3.1 Solvers and vehicle models

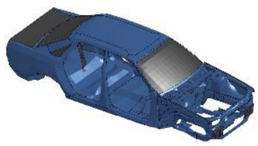
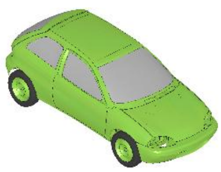
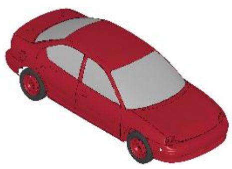
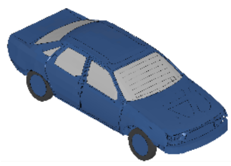
The objectives and constraints of the example optimization problems are based on the response of numerical simulations using the Finite Element Method (FEM). The vehicle models used in this work are based on the models available from the NCAC finite element model archive (NCAC 2014). For the presented investigation and case studies, the models displayed in table 1 are selected. The Metro model (A) is selected because the low mesh resolution, and forthcoming low computational cost which enabled a larger number of

The final publication is available at Springer via http://dx.doi.org/10.1007/s00158-016-1410-9

function evaluations for the response characterization. The Neon model (B) has a much higher FEM mesh resolution, and a more detailed model structure. The Taurus model (C) has a computational cost between that of models A and B, which allows the large number function evaluations required for the corroboration of the approach. The number of design variables is different for each of the vehicle models because different design car body construction concepts are used and the vehicles used are modeled with a different level of detail. Although this work is dealing with “similar” vehicle-related optimization problems, it is rather typical in industry that there are some differences between the different problem instances. The differences in geometry, FEM mesh resolution, number of design variables, and car body concepts, enable the assessment of the robustness of the response characterization for the different vehicle models and presented benchmark approach. Typical full vehicle crashworthiness models applied in industry today have about 1-10 million elements, and require computation times in the order of magnitude of 100 CPU hours, for a single 100ms crash event. The models used for the response characterization had lower mesh resolution and required less computation time (see Table 1). These models are less accurate in representing the exact behavior of a particular vehicle model, however in this work the identification of typical response characteristics w.r.t. the design variables is prioritized over the accuracy required in a detailed analysis of a particular vehicle design. The response characterization results in section 5 did not indicate any dependency of the response characteristics depending on the mesh, although the used models differed in mesh resolution for an order of magnitude.

As design variables for the optimization problems, the scaling factors on the thickness of BIP components have been parameterized. Components appearing on both sides of the vehicle are scaled symmetrically. In table 1 the parts with variable thickness are colored in the pictures of the modal analysis models, while constant parts are displayed in gray, the same design variables have been used for the crash simulation.

Table 1 Overview of the used vehicle models

	model A	model B	model C
	Metro	Neon	Taurus
Modal analysis models (in color the parts with variable thickness)			
Crashworthiness models			
Nr. of elements Crash model	16k	271k	28k
Total CPU time[3] [hr.] for a crash simulation of 100ms	0.4	30	1.1
Nr. of design variables	32	72	50

For vehicle models A, B and C the total nominal mass of the parameterized components accounts for 75%, 90% and 90% of the total BIP mass respectively. The design variables of all vehicle models are the parameterized sheet thickness of selected BIP components. The design variables are normalized to be in

[3] Approximate CPU time per simulation using a single logical core of a HP Z600 with 2 Intel Xeon E5520 processors, and 24GB DDR3 Memory.

the unit hypercube domain and scale the nominal part thickness by a scaling factor varying between 0.5 and 2.

In the following sections of this paper the results of four response types are discussed in further detail:

(1) Vehicle body mass (Mass)
(2) First (free-free) natural torsion eigenfrequency (NTF1)
(3) Deformation between A- and B-Pillars during crash (ABP. Def.)
(4) Peak acceleration during crash[4] (P. acc.)

The mass for each vehicle model design, is calculated by summation over the lumped element masses of a predefined set of elements representing the components of the vehicle Body In Prime (BIP). For the FEM based linear modal analysis using LS-DYNA-implicit (version 971) codes is used. The eigenmodes are distinguished using the Modal Assurance Criterion (MAC) with respect to the dynamic behavior of the nominal vehicle model configuration. The crash load case is a frontal crash configuration against a rigid wall at 64 km/h. For the crashworthiness simulations, nonlinear transient dynamic analysis explicit FEM (LS-DYNA 971) is used (see Wu and Gu (2012) for a general theoretical overview and Hallquist (2006) for details about the used software implementation). Fig. 1 illustrates the typical phenomena involved in the load cases.

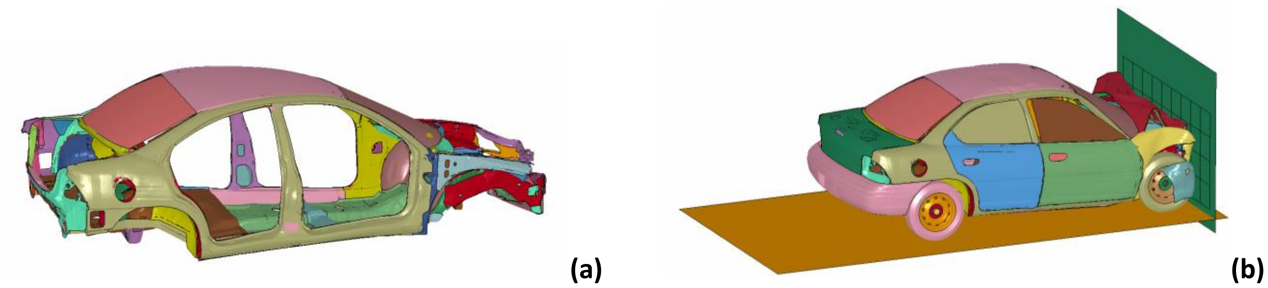


Fig. 1 a Simulation response: scaled deformation plot of the first natural torsional eigenfrequency of vehicle model B; **b** Simulation response deformation during a crash simulation of vehicle model B

For the response characterization and to perform structural optimizations, a simulation workflow is created such that the NVH, crashworthiness, and mass criteria of each design variant can be determined "automatically" as a function of the design variables.

4 Simulation response characterizations

4.1 Local parameter studies

Local one-factor-at-a-time (OFAT) and two-factor-at-a-time (TFAT), parameter studies have been performed to investigate and quantify the degree and type of nonlinearity of the response functions, as a function of the design variables. For these parameter studies, one or two variables have been changed in fixed steps over the entire domain, while all other design variables are fixed to their nominal value, hence only first and second-order effects are investigated. The term local in this context refers to the fact that these parameter studies have only been performed at a single location w.r.t. the other design variables. It has to be noted that for other responses or design variable types (such as parameterized ply orientation in the case of composite materials), or other design variable ranges the relationship type between the design variables and response could be different. Fig. 2 shows a representative sample of the response characterization w.r.t. change of a single design variable, while keeping the others fixed to the nominal value. In the present work a characterization of the responses of first-order changes of all the design variables is performed for vehicle models A and B

[4] The peak acceleration results are based on SAE 60 Hz low pass filtered acceleration values of an accelerometer element located at the center of the vehicle on the tunnel.

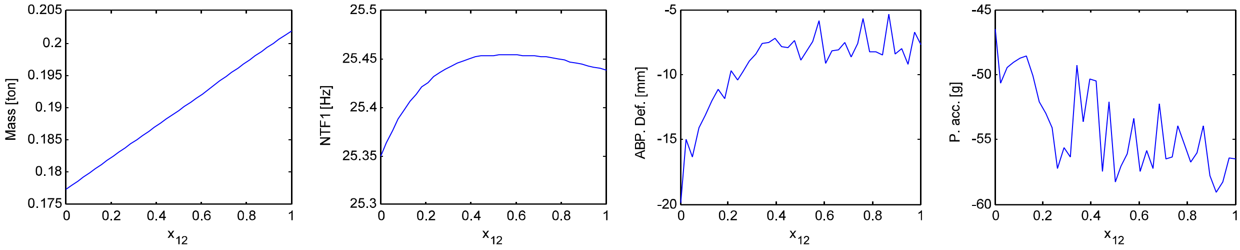


Fig. 2 Overview of different types of nonlinearities in OFAT parameter studies, for four different response types w.r.t the variation of one design variable

For the sake of brevity, only a few results are displayed. However, the full set of parameter study results indicate very similar nonlinearity characteristics w.r.t the design variables for each response type. The relative importance or amplitude of the first and second-order effects varied over the different design variables, but the "shape" of the relation between the responses and the design variables identified characteristic types of nonlinearity for each response type. The results indicated linear behavior for the vehicle mass response, mildly nonlinear behavior for the natural modes, and highly nonlinear behavior of the deformation and peak acceleration responses during the crashworthiness load cases

To investigate the type of interactions, similar investigations are performed for the using TFAT parameter studies on a subset of the design variables. The subset is defined based on the global sensitivity analysis results described in section 4.2. Fig. 3 shows an example of the results.

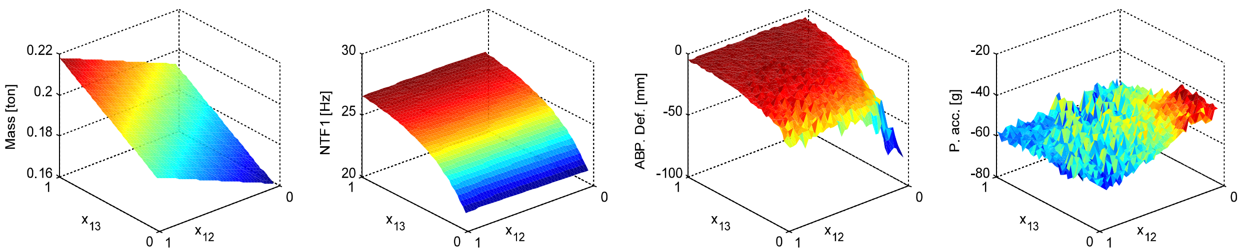


Fig. 3 Overview of different types of nonlinearities in TFAT parametric studies

Further analysis is performed to quantify the nonlinearity and variations among parameters. For the responses with nonlinear and non-smooth first-order and second-order effects, the results are analyzed using one and two-dimensional spectral wavenumber decomposition using the Fast Fourier Transform (FFT). Fig. 4 shows examples of wavenumber decomposition analysis results for the peak acceleration response, of vehicle model A.

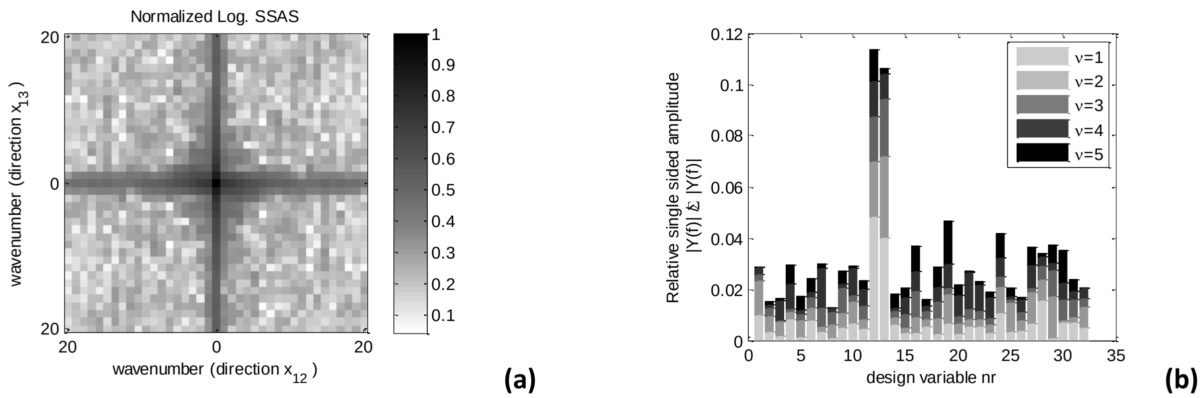


Fig. 4 Fourier analysis on the OFAT, and TFAT parameter study results. **a**: the normalized single sided amplitude spectrum (SSAS) for 2d wavenumber decomposition; **b**: the 1D SSAS of low wavenumbers ($\boldsymbol{\nu}$) for all design variables

For all of the investigated design variables and vehicle models the results indicate that the low wavenumber "trends" are of predominant importance. Although it is difficult to discover common trends in the distribution of individual amplitude contributions per wavenumber, the amplitude contribution averaged over all design variables, is decreasing with increasing wavenumber.

4.2 Global sensitivity analysis and variance decomposition

Using existing global sensitivity analysis (GSA) methods, and variance based variable screening methods, the first and second-order variance contributions or sensitivity indices of the model output w.r.t. the optimization design variables are estimated for the two vehicle models. First-order sensitivity indices are defined as: $S_i = V_i/VAR(Y)$ where $V_i = VAR(E_{X\sim i}(Y|X_i))$ represents the variance (VAR) of the expected value $E()$ of response or model output Y conditioned w.r.t. design variable X_i. Analogously second-order indices can be defined as: $S_{ij} = VAR(E_{X\sim ij}(Y|X_{ij}))$. For an introduction and further theory of GSA methods the reader is referred to Sobol (2001) and Saltelli et al. (2010). The used implementations for the sensitivity index estimation and variable screening are described in Ratto and Pagano (2010). For GSA of the response model output w.r.t. the model input, 2000 pseudo random samples of design variable combinations are used for vehicle model A, and 1000 for vehicle model B. Fig. 5 shows the sensitivity distributions[5] for the four different response types of two vehicle models.

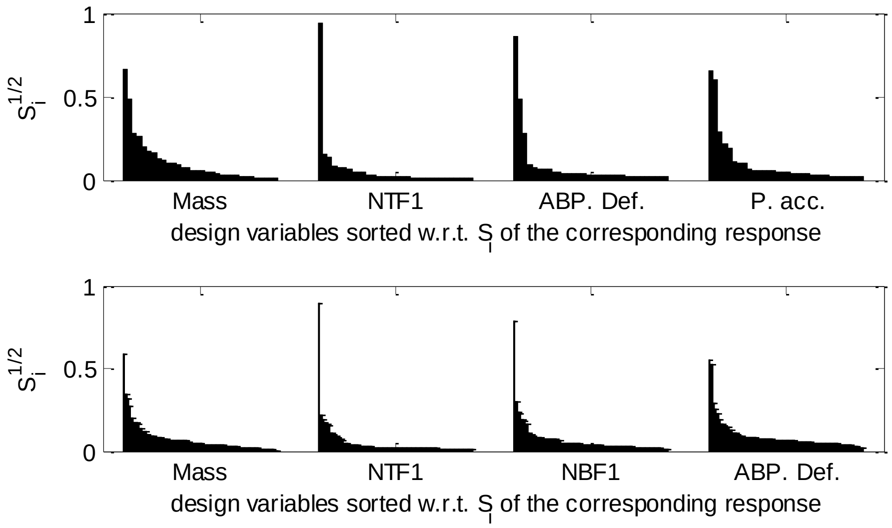


Fig. 5 Sensitivity distributions for responses and two vehicle models, for vehicle model A (top) and B (bottom). The variables are independently sorted in descending order of relevance, within each sub-figure

[5] The distribution of the first order sensitivity indices S_i are expressed in terms of $\sqrt{S_i}$ since this is in the opinion of the authors more intuitive for visualization (in a similar manner as standard deviation can be preferred over variance in particular diagrams).

The resulting estimates of the first-order sensitivity indices show characteristic distributions for all of the investigated simulation responses. For the Mass, NTF1 and “ABP. Def.” responses, a small fraction of the design variables have a high contribution to the total response variance. Similar results are obtained for both investigated vehicle models (A and B). It should be noted that in Fig. 5 the variables are sorted in descending order of relevance according to variance contributions. Their ordering for the different response types is however different, such that variables important for one response are not necessarily important for another response. This is visualized in Fig. 6, where for each vehicle model a unique ordering according to the mass response is used. The relation of the variable importance between the different simulation responses is further dealt with in section 4.3.

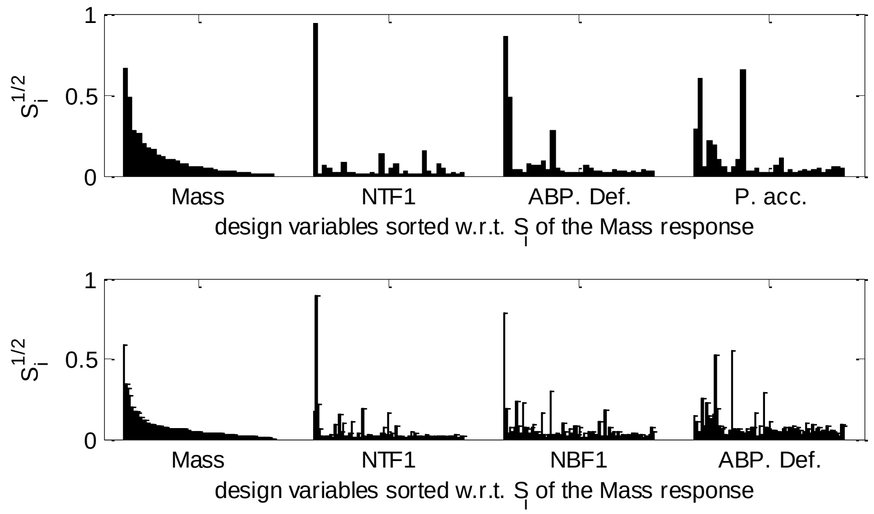


Fig. 6 First-order sensitivities sorted by mass influence, for vehicle model A (top) and B (bottom)

A natural property of the sensitivity indices or Sobol indices resulting from a GSA is that the variance contributions should sum up to unity. Combining the explained variance of a linear regression model together with the previously mentioned sensitivity analysis methods for the estimation of first and second-order sensitivity indices, an overall estimation of the variance decomposition can be obtained for each of the simulation responses (Fig. 7).

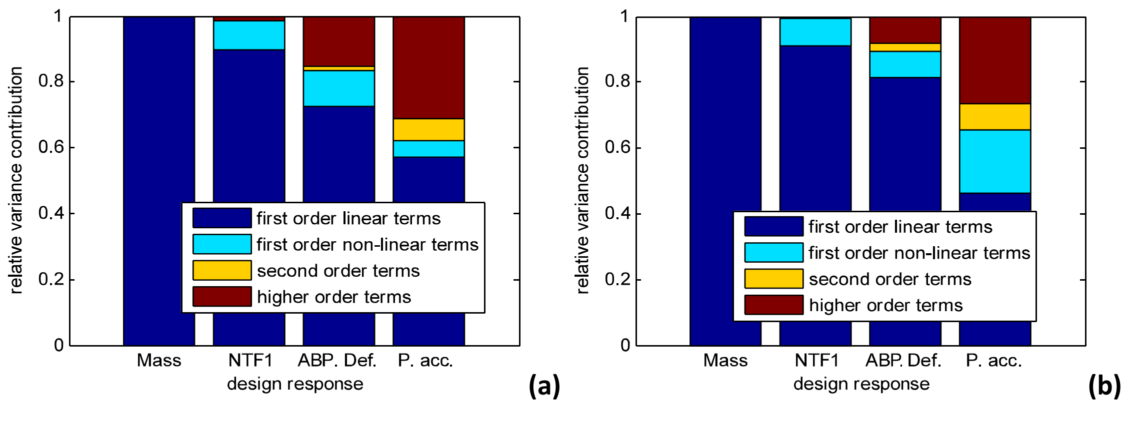


Fig. 7 Variance decomposition per response for: **a**: vehicle model A; **b**: vehicle model B

4.3 Simulation response correlations

Previous sections dealt with the analysis of the individual simulation responses with w.r.t. the design variable changes. In this section, a basic analysis of the structure between the different simulation responses of the system is presented. The structure between the different simulation responses and between the sensitivity distributions of the responses is investigated using the normalized covariance (see equation 10).

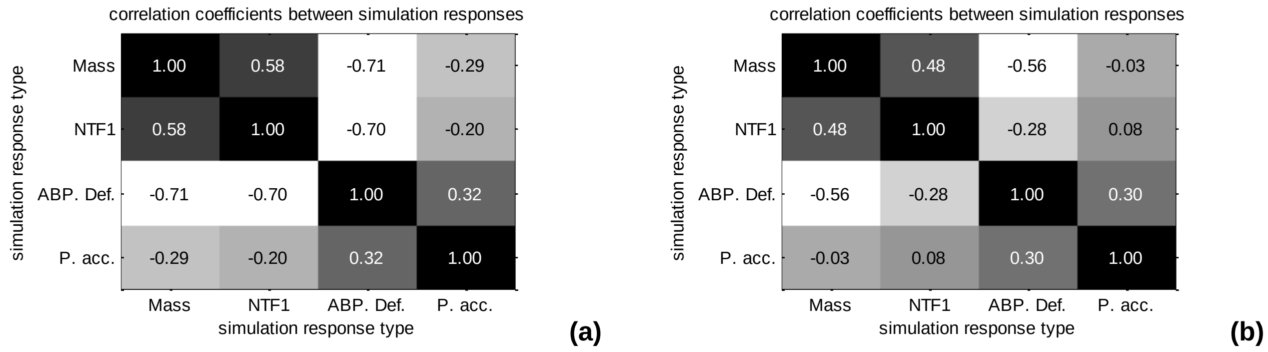


Fig. 8 Linear correlation coefficients between the simulation responses for: **a**: vehicle model A; **b** and B (right)

Fig. 8 shows the matrix of normalized covariance's also called (linear) correlation coefficients (Rodgers and Nicewander 1988) between the simulation responses, based on a quasi-random sampled design variable values for each of the vehicle models (A and B). Besides correlations among the design responses also the correlations between the linear first-order effects, of the different simulation responses are assessed.

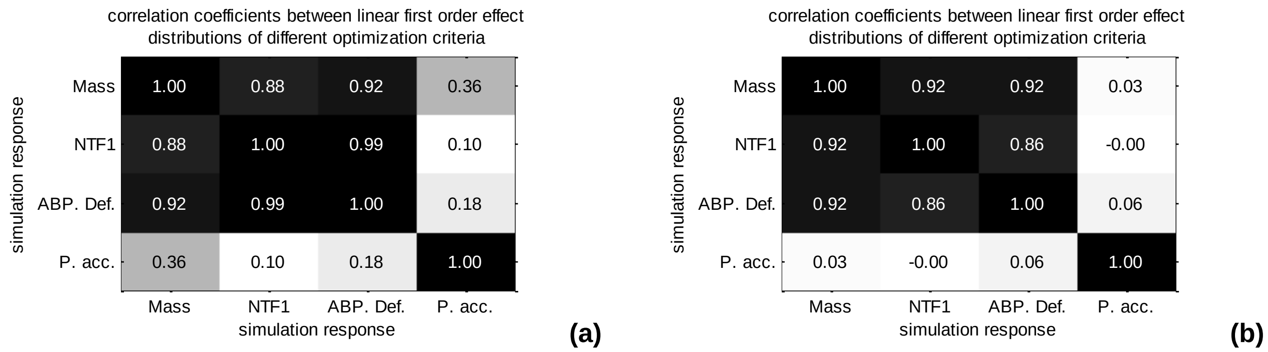


Fig. 9 Linear correlation coefficients between linear first-order effect distributions of different simulation responses for: **a**: vehicle model A; **b**: vehicle model B

As an example Fig. 9 shows the correlation coefficient matrix between the distributions of the linear first-order sensitivity index estimates (based on linear regression models) for each of the simulation responses.

4.4 Discussion and summary of the response characterization

Of course, no general results can be obtained or claimed by investigating only two vehicle models with these approximate response characterization techniques. Nonetheless comparing the results between the two vehicle models, common trends, and a band of mutual differences between the investigated response characteristics could be qualitatively estimated. Although, for other applications and response types, possibly more problem instances and other characterization techniques might be required for a useful estimate, the applied characterization methods are by no means specific to the presented application, and could be used to investigate other response types.

The response characterization performed here required a large investment in terms of computational effort. This investment, could however be worthwhile if a significant increase in optimization efficiency for other instances of related problems can be obtained with the RSP approach. The computational investment for this particular example is done in scope of a proof of concept. The main goals of this section were however to show the application of different analysis techniques that can be used for the response characterization and to provide an overview of similarities and differences in response characteristics of different simulation model instances of similar type.

The final publication is available at Springer via http://dx.doi.org/10.1007/s00158-016-1410-9

5 Construction of representative surrogate problems for the vehicle design optimization example

In this section, an example implementation of the RSP approach is presented. As stated in the general description of the approach, the aim of an RSP is to mimic selected function characteristics of the simulation-based vehicle model responses of interest in a statistical sense, and not to approximate a particular response or dataset such as is usually done in the context of meta-models or surrogate models. Fig. 10 displays a schematic workflow of the steps used to construct the RSP in this application example.

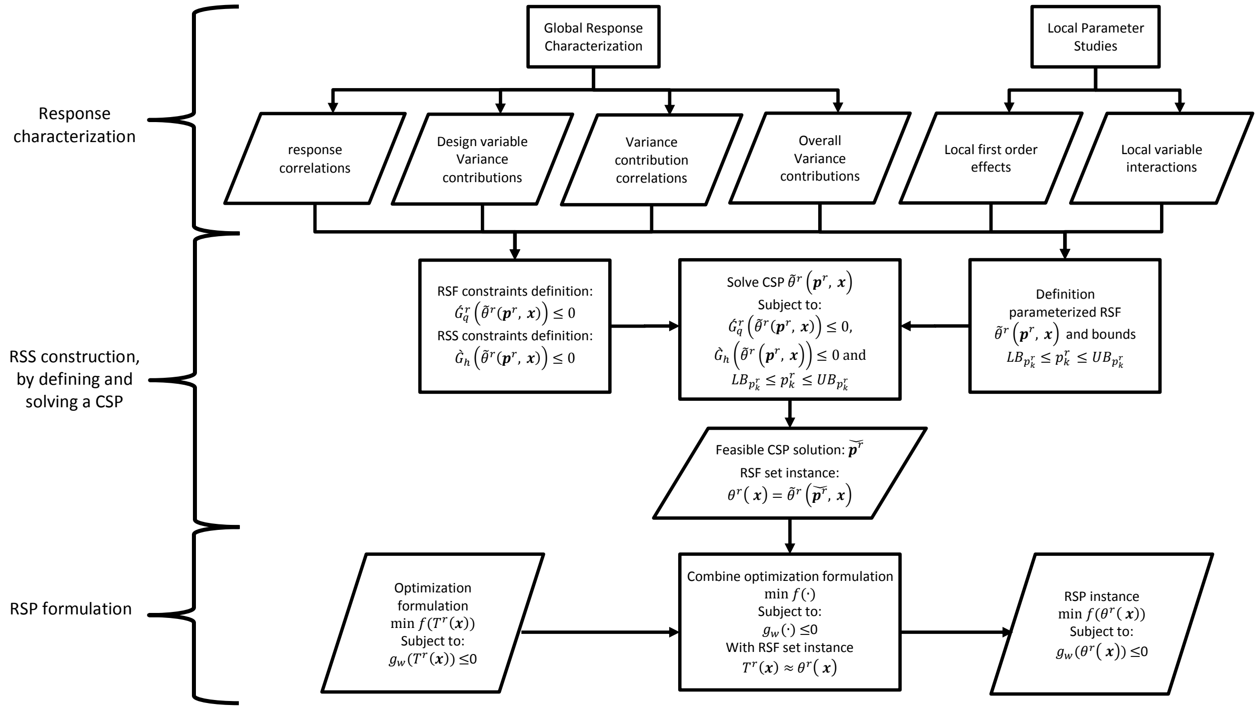


Fig. 10 Schematic flow diagram for the construction of the RSP for the car body design application case study

The results of the simulation response characterization gave an indication of common features and differences between the corresponding responses of the different vehicle models. For this example the selected characteristics for a single response are:

- The type or "shape" of the response nonlinearity w.r.t the design variables
- The distribution of the first and second-order sensitivity indices w.r.t the design variables
- The distribution of the total variance contribution of the first, second and higher order effects

The selected characteristics between the different responses are:

- The correlation coefficients (or normalized covariance) of the responses
- The correlation coefficients of the first-order sensitivity indices between the responses

For this example, an RSF formulation for each response type superscript *r*), composed of a series expansion truncated to include interaction effects up to the second-order followed by a combined higher order is used to represent the behavior of the characteristic responses.

$$\theta^r(\boldsymbol{x}) = \sum_{i=1}^{d} \theta_i^r(x_i) + \sum_{1\le i<j\le d} \theta_{i,j}^r(x_i, x_j) + \theta_{i,j,\dots,d}^r(x_1, x_2, \dots, x_d) \quad (4)$$

Superscript r is the index over the different types of outputs or pseudo responses (in this example: 1 Mass, 2 NTF1, 3 "ABP. Def.", 4 "P acc."). All operations considered are invariant to addition with a constant, which is therefore omitted at this stage. For each of the responses the choices for the summands of the representing basis functions, and the parameter bounds are summarized in Table 2. Each of the basis function summands has free parameters which are the variables for the constructed CSP. For the RSS with the general set of free parameters $\boldsymbol{p}^r$ the following notation is used: $\tilde{\theta}^r(\boldsymbol{p}^r, \boldsymbol{x})$. The general set of parameter bounds is expressed as $LB_{p_k^r} \leq p_k^r \leq UB_{p_k^r}$. The particular free parameters for each summand in the RSF of each response type are listed in column 5 of table 2.

The general set of constraints on the CSP that relate to the separate RSFs is expressed as:

$$\acute{G}_q^r\left(\tilde{\theta}^r\left(\boldsymbol{p}^r,\, \boldsymbol{x}\right)\right) \leq 0 \tag{5}$$

And the general set of constraints working on the combined set of RSFs is expressed as:

$$\grave{G}_h\left(\tilde{\theta}^r\left(\boldsymbol{p}^r,\, \boldsymbol{x}\right)\right) \leq 0 \tag{6}$$

The CSP problem can be relaxed by defining tolerances for each of the constraints, or by using lower and upper bounds for the quantities of interest, in the presented example lower and upper bounds are used instead of tolerances, this is however at the cost of doubling the number of constraints in the CSP. These nonlinear constraint functions will be defined later in this section.

5.1 Selection of the basis functions

As mentioned in section 2 the choice for basis functions and the series truncation is dependent on the respective response characterization results. For the provided example the choice for the basis function types is based on the local OFAT and TFAT parameter studies on a subset of the design variables. To represent the “Mass” and “NTF1” responses w.r.t. the design variables (see Fig. 3 and 4), linear basis functions and a subset of quadratic polynomials are selected respectively. For these response types second and higher order interaction terms are omitted, since nearly all the variance of the responses can be explained without these (see also Fig. 7). Based on the parameter study results (see Fig. 2, 3 and 4) a composition of linear functions and harmonic series (expressed as complex exponentials in table 2) is selected to represent the first and second-order characteristic nonlinear relation between the design variables and the “ABP. Def” and “P. acc” simulation responses.

An analysis of higher than second-order effects of the simulation responses requires a high amount of function evaluations and was computationally infeasible to the authors. The performed response characterization could however provide an indication for the total magnitude of the variance contribution of unexplained third and higher order effects (Fig. 7). Based on the unsmooth behavior observed in the local sensitivity analysis, the assumption is made that these higher order effects can be represented by a single non-smooth field with higher order interactions. To represent such a non-smooth field, functions that generate reproducible isotropic uniform distributed noise are used. “These functions denoted by operator $W(\boldsymbol{x})\colon \mathbb{R}^n \to M \subseteq \mathbb{R}$ serve as a multivariate random map to pseudo random but reproducible values of set M, where M is a set with a uniform distribution in the open interval $(-1,1]$. ”The magnitude of this uniform noise field is scaled by a factor q which is chosen such that variance contribution of this term matches the “explained” variance by higher order terms in the response characterization (see also Fig. 7).

Table 2 Overview on the summands for the RSFs

Response type	r	int.	representing summand formulation	function/parameter constraints
Mass	1	1	$\tilde{\theta}_i^r(a_i^1, x_i) = a_i^1 x_i$	$0 \le a_i^1 \le 1$
		2	$\tilde{\theta}_{ij}^r(x_i, x_j) = 0$	
NTF1	2	1	$\tilde{\theta}_i^r(a_i^2, b_i^2, x_i) = a_i^2(x_i - b_i^2)^2 - a_i^2(b_i^2)^2$	$-1 \le a_i^2 \le 0$ $0.7 \le b_i^2 \le 1.5$
		2	$\tilde{\theta}_{ij}^r(x_i, x_j) = 0$	
ABP. Def.	3	1	$\tilde{\theta}_i^r(b_i^r, u_i^r, x_i) = a_i^r x_i + u_i^r \sum_{n=1}^{m} c_{in}^r e^{2\pi\, inx_i}$	$0 \le a_i^3 \le 1$, $0 \le u_i^3 \le 1$ $\theta_i^r(.) \mapsto \mathbb{R}$ $c_{in}^r \sim c_{kn}^{ref\,r}$, $m = 10$
		2	$\tilde{\theta}_{ij}^r(u_{ij}^r, x_i, x_j) = u^r v_{ij}^r \sum_{n_1}^{m_1} \sum_{n_2}^{m_2} c_{i,j,n_1,n_2}^r\, e^{2\pi i n_1 x_i} e^{2\pi i n_2 x_j}$	$0 \le u^r \le 1$ $\theta_{i,j}^r(.) \mapsto \mathbb{R}$, $v_{ij}^r = S_i^r \otimes S_j^r$ $c_{i,j,n_1,n_2}^r \sim c_{k,l,n_1,n_2}^{ref\,r}$, $m_1, m_2 = 10$
		d	$\tilde{\theta}_{i,j,\dots,d}^r(q^r, x_1, x_2, \dots, x_d) = q^r * W(x_1, x_2, \dots, x_d)$	$q^3 \ge 0$
P. acc.	4	1	$\tilde{\theta}_i^r(b_i^r, u_i^r, x_i) = a_i^r x_i + u_i^r \sum_{n=1}^{m} c_{in}^r e^{2\pi\, inx_i}$	$0 \le a_i^4 \le 1$, $0 \le u_i^4 \le 1$ $\theta_i^r(.) \mapsto \mathbb{R}$ $c_{in}^r \sim c_{kn}^{ref\,r}$, $m = 10$
		2	$\tilde{\theta}_{ij}^r(u_{ij}^r, x_i, x_j) = u^r v_{ij}^r \sum_{n_1}^{m_1} \sum_{n_2}^{m_2} c_{i,j,n_1,n_2}^r\, e^{2\pi i n_1 x_i} e^{2\pi i n_2 x_j}$	$0 \le u^r \le 1$ $\theta_{i,j}^r(.) \mapsto \mathbb{R}$, $v_{ij}^r = S_i^r \otimes S_j^r$ $c_{i,j,n_1,n_2}^r \sim c_{k,l,n_1,n_2}^{ref\,r}$, $m_1, m_2 = 10$
		d	$\tilde{\theta}_{i,j,\dots,d}^r(q^r, x_1, x_2, \dots, x_d) = q^r * W(x_1, x_2, \dots, x_d)$	$q^4 \ge 0$

The Fourier series coefficients of the RSFs for responses 3 and 4 (referred to by the symbol c with the corresponding sub and superscripts) are not part of the set of free parameters of the CSP. For the first-order terms, for each variable i the complex Fourier series coefficients c_{in}^r with index n over the frequencies are of similar structure $c_{in}^r \sim c_{kn}^{ref\,r}$ to the coefficients of a reference set $c_{kn}^{ref\,r}$. The coefficients of the reference set can be obtained by performing the discrete Fourier transform on gridded data points on the design variables after subtracting the linear trend (such as done in section 4). For the presented case study "similar" Fourier coefficient structures are obtained using the Iterative Adjusted Amplitude Fourier Transform (IAAFT) algorithm described by Schreiber and Schmitz (1996) and implemented by Venema (2003). The IAAFT method (denoted by operator H) can generate various discrete series or fields (depending on the random seed "z" that have the same amplitude distribution and autocorrelations, as the provided input data (the various calibration fields), up to a specified tolerance "t" (in the example 0.005).

$$c_{in}^{sur\,r} = H(c_{kn}^{ref\,r}, t, z) \tag{7}$$

The resulting series and fields are later scaled by the factors u_i^r which are part of the variables set of the CSP. In this context the selected similarity criteria are: amplitude distribution and autocorrelation.

For the responses with considerable nonlinear second-order interactions (3 and 4), the correlation coefficient between the inner product of the first-order sensitivity indices, and the second-order sensitivity index estimate is high and significant for the calibration vehicle models. This indicates that for the application example the variables with high first-order effects are also the variables involved in the most important second-order interactions in terms of variance contribution. In order to reduce the number of free variables in the CSP the relative second-order sensitivity index distribution controlled by variable v_{ij}^r (see also table 2) is defined dependent of the first-order sensitivity distribution as: $v_{ij}^r = S_i^r \otimes S_j^r$ where $\otimes$ denotes the outer vector product. The amplitudes of the resulting nonlinear fields is scaled by variable u^rwhich is constrained such that the total variance contribution of the fields corresponds to the second-order contributions estimated in the response characterization (See Fig. 7).

Besides the selection of the basic functions and parameter bounds, also function and additional parameter constraints are defined to enforce response characteristics.

5.2 RSF constraints

The choice for the targeted sensitivity index distributions is made using the global sensitivity results presented in section 4. Sorting the sensitivity indices for each response, obtained in the function characterization on descending order, a fit for the sensitivity index distribution can be made. The distributions of all of the responses in this case study could be approximately described by a two-term exponential fit model (see also Fig. 5). The related function constraints are defined as upper ($Z_j^{\mathrm{UB}\,r}$) and lower bounds ($Z_j^{\mathrm{LB}\,r}$) on the ordered set first-order sensitivity indices. The set of upper and lower bounds is based on the fit model on the sorted set of sensitivity indices from the calibration models. This can be expressed for the general case as:

$$\acute{G}_q^r\left(\tilde{\theta}^r\left(\boldsymbol{p}^r,\,\boldsymbol{x}\right)\right)=Z_j^r-Z_j^{UB\,r}\quad \text{for q=1:d and j=d and}$$

$$\acute{G}_q^r\left(\tilde{\theta}^r\left(\boldsymbol{p}^r,\,\boldsymbol{x}\right)\right)=Z_j^{LB\,r}-Z_j^r\quad \text{for q=d+1:2d and j=q-d} \tag{8}$$

Where Z_k^r contains the sensitivity index estimates S_i^r for each response *r* in descending order over index *i* using the sorting transformation $Z_k^r{=}\beta(S_i^r)$. The sensitivity indices for each RSF $S_i^r = Q\left(\theta^r(\boldsymbol{p}^r,\,\boldsymbol{x}_w)\right)$ are estimated using the method described in Plischke (2010) denoted by operator Q based on a set of pseudo random samples $\boldsymbol{x}_w$.

5.3 RSS constraints

Following the described approach up to this point for each of the design responses (mass, frequency, deformation, peak acceleration) would lead to function formulations that could be representative for each simulation responses individually, but would not take into account the coupling structure between the responses. In the applied approach, the coupling between the responses is accounted for by applying constraints on the correlations between the function responses, and the correlations between the sensitivity distributions for each of the responses.

For a set of w design evaluation vectors the matrix of results (Y) for each design is defined as:

$$Y_{wr} = \theta^r(\boldsymbol{x}_w) \tag{9}$$

The linear correlation coefficients between the column vectors of the responses are given by:

$$\rho_{tv}^{\mathrm{Y}} = R(Y_{wt}\,, Y_{wv}) \tag{10}$$

where $R()$ is the operator that results in the correlation coefficient between two vectors defined as:

$$\rho = R(A,B) = \frac{\mathrm{cov}(A,B)}{[\mathrm{cov}(A,A)\mathrm{cov}(B,B)]^{1/2}} \tag{11}$$

The similarity of the obtained correlation coefficient matrices of the test function can be defined by choosing lower ($\rho_{tv}^{\mathrm{S_iLB}}$) and upper bounds ($\rho_{tv}^{\mathrm{S_iUB}}$) for each of the upper diagonal matrix entries. The upper and lower bounds are based on the values obtained in the response characterization of the calibration models.

$$\grave{G}_h\left(\tilde{\theta}^r\left(\boldsymbol{p}^r, \boldsymbol{x}_w\right)\right) = \rho_{tv}^{Y} - \rho_{tv}^{Y\,\mathrm{UB}} \quad \text{for t=1:(N-1), v=(t+1):N \& h=t+N(v-1)-v(v-1)/2}$$

$$\grave{G}_h\left(\tilde{\theta}^r\left(\boldsymbol{p}^r, \boldsymbol{x}_w\right)\right) = \rho_{tv}^{Y\,\mathrm{LB}} - \rho_{tv}^{Y} \quad \text{for t=1:(N-1), v=(t+1):N \& h= t+N(v-1)-v(v-1)/2+N(N-1)/2} \tag{12}$$

where N is the number of responses (4 in this example). A similar approach is used for the correlation between the first-order sensitivity indices S_i^r of all combinations responses.

$$\rho_{tv}^{\mathrm{S_i}} = R(S_i^t, S_i^v) \tag{13}$$

Also here lower and upper bounds for the correlation coefficients are defined based on the results of the response characterization of the calibration models. The corresponding constraints are defined as:

$$\grave{G}_h\left(\tilde{\theta}^r\left(\boldsymbol{p}^r, \boldsymbol{x}_w\right)\right) = \rho_{tv}^{\mathrm{S_i}} - \rho_{tv}^{\mathrm{S_iUB}} \quad \text{for t=1:(N-1), v=(1+t):N \& h=t+N(v-1)-v(v-1)/2 + N(N-1)}$$

$$\grave{G}_h\left(\tilde{\theta}^r\left(\boldsymbol{p}^r, \boldsymbol{x}_w\right)\right) = \rho_{tv}^{\mathrm{S_iLB}} - \rho_{tv}^{\mathrm{S_i}} \quad \text{for t=1:(N-1), v=(1+t):N \& h= t+N(v-1)-v(v-1)/2+3N(N-1)/2} \tag{14}$$

These function and problem constraints are selected to achieve representativeness of the surrogate problem to the calibration problems. The selection of the function formulation for each of the responses, and the corresponding free parameters, combined with the parameter constraints, function constraints, sensitivity index constraints, and correlation constraints define a CSP.

5.4 CSP solution

The general notation used previously for the set of parameterized basis functions to represent the responses can now be linked to the corresponding parameters from table 2 :

$$\tilde{\theta}^r\left(\boldsymbol{p}^r, \boldsymbol{x}\right) = \theta^r(a_i^1, a_i^2, b_i^2, a_i^3, u_i^3, u^3, a_i^4, u_i^4, u^4, x_i) \tag{15}$$

The CSP remains parameterized in the number of design variables d over index i, which can be chosen according to the target problem dimensionality. After fixing the number of design variables (set to 50 for the case study that will be presented in section 6), the CSP can be solved. For this application with relatively few problem instances in the training set the upper and lower bounds for the constraints are based on a simple “averaging approach”, where the minimum and maximum values from the calibration model characterization results are set as the lower and upper bounds of the respective constraint values. The total number of constraints K for the CSP criteria in this example scales with d according to $K=2Nd+2N(N-1)$ where N is the number of responses. Each of the constraint equations of the CSP is dependent on the set $\boldsymbol{x}_w$. In this example a, fixed set of 10^5 pseudo random samples in the domain of the design variables (the unit hyper cube) is used for $\boldsymbol{x}_w$ and during the CSP solution procedure it can be considered as a constant. Solutions to the flexible CSP problem could be obtained using various methods. For the presented example a standard Interior point method that handles nonlinear constraints (the MATLAB 2013a “fmincon” function) is used, with the full set of constraints as separate nonlinear constraints. An auxiliary objective function defined as:

$$\vartheta\left(\boldsymbol{p}^r\right) = \sum_q \delta\left(\acute{G}_q^r(\boldsymbol{p}^r)\right) + \sum_h \delta\left(\grave{G}_h\left(\boldsymbol{p}^r\right)\right) \tag{16}$$

where operator $\delta(\cdot)$ is an indicator function defined as:

$$\delta(X) \coloneqq \begin{cases} X \text{ if } X > 0 \\ 0 \text{ if } X \leq 0 \end{cases} \tag{17}$$

Combining this auxiliary objective function, the parameter bounds and constraint sets from equations 8, 12 and 14 the CSP can be solved. For the example this is done using successive optimization runs with decreasing constraint violation tolerances ranging from 1 at the start, to 1E-6 in the final run. For the successive optimizations, the final value of the previous run is used as the initial value of the next optimization. Each feasible solution to the CSP represents a parameter set which when combined with the basis functions forms a response set with representative characteristics with respect to the selected criteria.

Up to this point, the constraints are all based on relative measures (sensitivity indices, and correlation coefficients, which are invariant w.r.t. the addition of constants and scaling by multiplication). The absolute range of the response of the surrogate functions can be controlled by applying the corresponding offset χ^r and scaling factors ψ^r to the resulting RSS from the CSP solution.

$$\theta^r(\boldsymbol{x}) = \chi^r + \psi^r * \tilde{\theta}^r\left(\boldsymbol{p}^r, \boldsymbol{x}_w\right) \tag{18}$$

Optimization algorithms are however typically programmed to be scale-free. Therefore this last step is not necessary to obtain results, and the results are not affected by the choice of the offset and scaling factors.

6 Application examples and “proof of concept” corroboration

6.1 Benchmarks of optimization efficiency

As a first application example of the RSP approach, its use in benchmarking optimization performance for particular problem types is considered. The performance of several optimization algorithms is estimated on two RSP formulations, after which the results are compared with performance results based on simulation workflow based problems. The two different optimization problem formulations used for the comparisons are:

1. Objective: Minimization of the vehicle mass, subjected to crashworthiness constraints (max peak acceleration at the tunnel, and A-B-pillar deformation)
2. Objective: Maximization of the 1st natural torsion frequency, subjected to mass constraints

Both RSPs are based on a single RSS (obtained as described in section 5), and the results are compared with the optimization performances on the corresponding problem formulations of a full vehicle simulation workflow (vehicle model C) which was not part of the original calibration data set. The number of design variables RSS is set to 50 according to the targeted validation vehicle model. The comparison for the optimization efficiency is made for the following algorithms:

1. Interior point (IP) algorithm
2. Sequential quadratic programming (SQP)
3. Genetic Algorithm (GA)
4. Non-dominated Sorting Genetic Algorithm, (NSGA2)
5. Differential Evolution (DE)
6. Particle Swarm Optimization (PSO)
7. Simulated Annealing (SA)
8. Fire Fly Algorithm (FFA)

For further information on the methods and implementations the reader is referred to the corresponding references[6]. Algorithms 3, 5, 6, 7 and 8 are metaheuristic search algorithms which are commonly used for

---

[6] 1) The Interior Point (IP) algorithm is commonly used to solve convex problems. The implementation used is included in MATLAB 2013a in the “fmincon” function option 1. For a description of the algorithm see Boyd and Vandenberghe (2009).

2) Sequential Quadratic Programming (SQP) approaches are generally used to solve smooth nonlinear problems, by sequential steps of the Newton method. In this work the implementation included in MATLAB 2013a in the “fmincon” function (option 3) is used. A description of the algorithm is given in Fletcher (2010).

3) Genetic Algorithms (GA) are a class of evolutionary algorithms, which are inspired by the genetic process of reproduction in biological life. The application of such algorithms is proposed in Rechenberg (1973), and a detailed description can be found in the work of Goldberg and Holland (1988). In this work the “ALGA” implementation included in MATLAB 2013a is used.

4) The Non-dominated Sorting Genetic Algorithm (NSGA-2) is a multi-objective evolutionary algorithm developed by Deb et al. (2000). The variant of the algorithm used in this work is Reference-point based NSGA-II implemented by Lin (2011).

problem types involving non-convex nonlinear responses, whereas the IP and SQP algorithms are typically used for nonlinear convex problems, and NSGA2 is a multi-objective optimization algorithm. Although the application of NSGA2 is unconventional for single objective problems, preliminary investigations showed reasonable performances for the type problems of interest. Since optimization formulation 2 does not include the highly nonlinear crashworthiness responses, also algorithms 1 and 2 are included in the comparison.

For each formulation, the optimization algorithm is repeated on the same optimization problem with default optimization algorithm parameters (except for the random seed and the initial population), such that performance statistics can be obtained. To compare the optimization efficiency for each problem, the results can be expressed in terms of Relative Objective Improvement (ROI) which is defined as:

$$\mathrm{ROI} = \frac{F_{REF} - F_{min_{\mathrm{k}}}}{F_{REF} - F_{min_{known}}} \qquad (19)$$

where $F_{min_{\mathrm{k}}}$ is the minimum feasible objective after k function evaluations, F_{REF} is the objective value of the reference design, and $F_{min_{known}}$ is the best "known" feasible objective value found for the given problem formulation, during all optimization runs of all algorithms. The ROI expresses the ratio of the objective improvement at a given number of samples to the maximum known achievable improvement of the problem formulation.

Fig. 11 shows the performance expressed in averaged relative objective improvement for optimization runs up to 250 function evaluations per optimization run, 100 repetitions per optimization for the corresponding RSP, and 20 repetitions per optimization on the independent full vehicle C simulation model. A higher number of repetitions would allow more accurate estimates, especially of the distribution or percentiles but, this was unfeasible due to the involved computational cost. The total number of function evaluations per optimization run is limited to 250, due to the high computational cost for the validation runs. As indicated in Duddeck (2008), it is common in an industrial environment to apriori limit the number of function evaluations to a number too small to reach convergence, a true optimization up to convergence is rather exceptional when dealing with problems that involve computationally expensive simulations (see also Knowles et al. (2005).

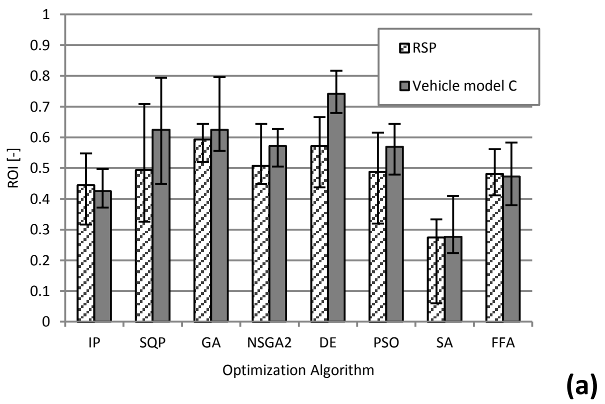
---

5) Differential Evolution (DE) is another evolutionary algorithm used for optimization (Storm and Price (1997)). The implementation used in this work is an adaptation of the code by Buehren (2008), combined with a penalty approach to enforce nonlinear constraint handling.

6) Particle Swarm Optimization (PSO) algorithms are nature inspired meta-heuristics that mimic the movement of groups of organisms such as bird flocks or fish schools. In this work the implementation by Birge (2006) is applied combined with a penalty factor approach to handle nonlinear constraints. A description for the algorithm principles can be found in chapter 8 of Yang (2010a).

7) Simulated Annealing (SA) is an optimization approach inspired by the thermodynamic process used in metallurgic annealing heat treatment (Kirkpatrick et al (1983)). In Yang (2010b) a description of the algorithm is provided together with an implementation of the algorithm that is used in this work.

8) Fire Fly inspired optimization algorithms are population based algorithms inspired by the behavior of fire flies. A description of the algorithm and implementation used in this work is provided in Yang (2010b).



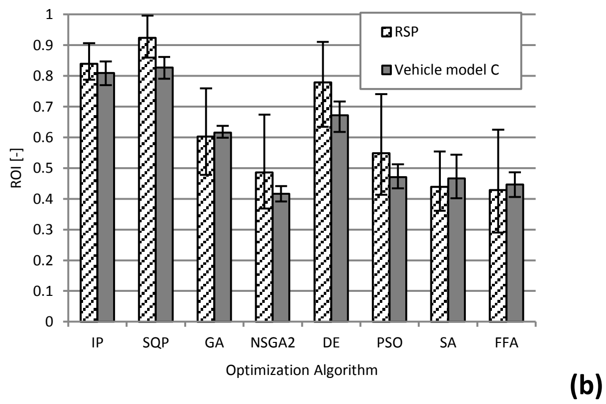


Fig. 11 comparison of average optimization efficiency for eight optimization algorithms on two different optimization problem formulations, **a**: formulation 1; **b**: formulation 2)

The results in Fig. 11 show a similar trend in relative algorithm performances between the optimizations run on the RSP and the optimization runs on the simulation workflow of vehicle model C. The error bars are estimates of the 20% and 80% percentiles. The similarity between the performance prediction and results can be quantified by the correlation coefficients R and the corresponding significance by the p-values between the vectors of optimization algorithm performance results obtained with the RSP and simulation workflow, which are R=0.910, p=0.0012 and R=0.964, p=0.0001 respectively. Thus, it can be concluded that in this corroboration example, the RSP approach offers a statistically significant prediction of the optimization efficiency of the tested algorithms applied to both problem formulations using the independent corroboration vehicle model. Application of the RSP approach to benchmark the algorithms and selecting the most efficient algorithms leads to optimization efficiency increases of 32% and 16% in terms of ROI for the respective optimization formulations (1 and 2) with respect to “the average” performance over the investigated algorithms. The computation cost of such a benchmark study without the RSP approach, comparing eight algorithms, 100 algorithm run repetitions, of 250 function evaluations, each requiring about 1 CPU hour (if a computationally cheap model is used) would require $2 * 10^5$ CPU hours. Whereas the RSP approach for the same study would take about 5 CPU hours[7] (including optimization algorithm overhead), thus saving several orders of magnitude in computation time. Even including the total function evaluation cost for the formulation of the RSP requiring about $1.5 * 10^4$ function evaluations, and a total of about $1.8 * 10^5$ CPU hours, the application of the RSP approach would already be worthwhile the computational investment, if a benchmark study was to be made. To justify the computational effort and endeavor of such a comparison, the computation cost of the comparison or RSP calibration (using reduced resolution simulation models) should be compared against the cost the industrial size problem which can be about $2 * 10^5$ CPU hours for a single optimization run. For the investigated examples the difference in efficiency between the algorithms in terms of CPU time is larger than the difference in terms of ROI. If CPU time savings in the order of 20% can be made by selecting a suitable optimization algorithm, the investment of the comparison pays off after about five industrial scale optimizations problems. For this particular example in the field of automotive engineering the increase in efficiency can however also translate in the improved mechanical performance due to the tight time constraints between design freezes in the vehicle development process.

Although the corroboration shows a significant correlation between the relative performances, such a resemblance cannot be guaranteed for any arbitrary vehicle. Nevertheless, it seems reasonable to assume that the results can be relevant for vehicle models with a similar structural concept, optimization parameters, and response criteria as the two calibration vehicle models and the third vehicle model used for the proof of concept. Furthermore, a single RSS can be used to construct several RSPs for different optimization formulations, and thus provides information and flexibility beyond single benchmark comparison results. For application-oriented practitioners the RSS and the derived RSP approach can answer more detailed

[7] CPU time for a RSP function evaluation is about 2.5E-2 [s] for the four responses in the example, using a MATLAB 2013a implementation on an Dell T3500 workstation with an Intel Xeon X5650 processor and 12 GB of RAM. The runtime of the optimizations using the RSP is dominated by the overhead of the optimization algorithm and optimization history saving.

questions than published benchmarks. For the community interested in optimization method development and comparisons, several standardized problems can be defined, and made available in order to provide access to reproducible representative surrogate problems of problem types which would be otherwise difficult to assess.

6.2 Meta-optimization by means of an RSP

A further example application of the RSP approach, regards the tuning of the parameters of an optimization algorithm to increase the optimization efficiency for problem types of interest. An optimization of the optimization parameters (or meta-optimization) is performed for the DE algorithm on an RSP. In the inner loop of the optimization, DE optimization runs with a maximum of 200 iterations are performed on the RSP. The objective of the inner loop is the minimization of the RSP mass response, with nonlinear constraints on the first natural mode and test function peak acceleration of the RSS. The design variables for the outer optimization are the optimization algorithm meta-parameters of the inner optimization (3 DE parameters: population size (π), crossover probability (μ), and step size (φ)). In the outer loop for each parameter-setting-vector -evaluation) a set of 50 inner loop optimization run repetitions is executed on the RSP: The 80% percentile of the minimum feasible pseudo mass determined after 200 function evaluations is set as the objective for the outer optimization. In the outer optimization loop a GA algorithm is used (with default settings) for 500 iterations to minimize objective thus finding statistically efficient performing optimization parameters for the inner optimization. A total of $5 * 10^6$ function evaluations on the RSP are performed for this case study.

The increase in optimization efficiency due the optimization meta-parameter tuning based on the RSP approach can be visualized by comparing the difference in optimization efficiency between the DE algorithm with "default" settings (π=30, $\mu = 0.7$ and $\varphi = 0.8$) and "optimized" (π=10, $\mu = 0.72$ and $\varphi = 0.87$) settings. Fig. 12 shows the plots of the best feasible objective history for standard parameter settings, and optimized parameter settings, on both the RSP and the full vehicle simulation workflow based optimization problem during respectively 50 and 15 optimization runs. The error bars indicate the 20% and 80% percentiles.

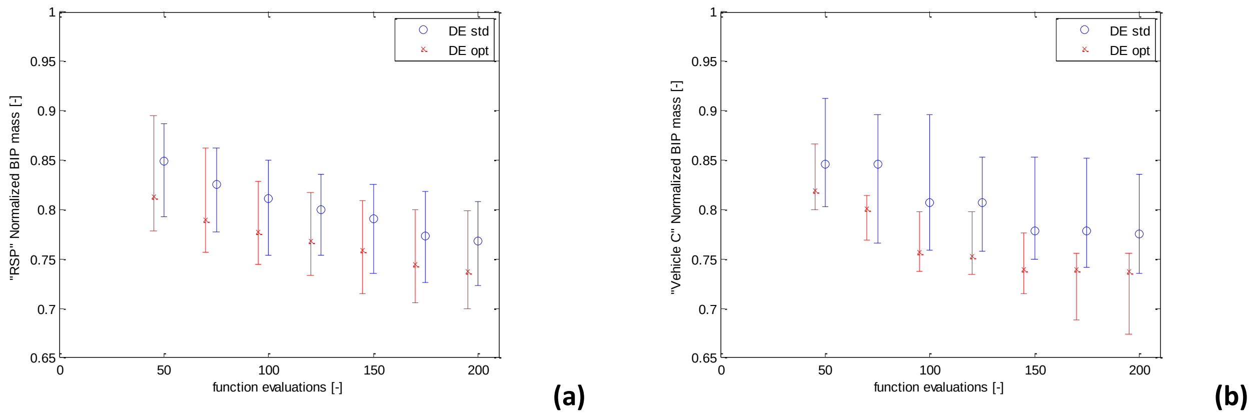


Fig. 12 Comparison of best feasible objective history for standard and optimized optimization parameter settings for: **a** the RSP; **b** the validation vehicle model C

For both the RSP as well as for the full vehicle simulation workflow based problems, the optimization performance is significantly improved by the tuning of the optimization meta-parameters. Since the full vehicle optimization problem (vehicle model C) was not part of the calibration set for the RSP, these results confirm the usefulness of the RSP approach for this problem type. In the corroboration example, the RSP approach based parameter tuning leads on average to additional performance gains of about 4% in terms of normalized BIP mass, for the fixed function evaluation budget.

Since optimizations of the full vehicle simulation workflow are orders of magnitude more computationally expensive than on the test functions, the number of repeated optimization runs on the vehicle simulation workflow is limited to 12 and hence, the resulting statistics are estimates only. The results have a significant common trend regarding the means, but the percentile statistics between optimization on the RSP and real problem are not quantitatively identical. Surprisingly the performance of the optimized settings is even better on the real problem than predicted. Although further refinements of the approach could possibly increase the general accuracy of the efficiency predictions, this accuracy is at the same time also capped by the nature of the approach. A surrogate problem, representative for a class of problems inherently has variability in efficiency prediction accuracy similar to the efficiency variation within the class of problems targeted. If additional information on the specific target problem is available prior to the simulation run, such data could be augmented to the RSP for increased performance estimation accuracy.

7 Discussion and outlook

The comparative study on optimization efficiency is unique in the literature for this type of problems. However the authors wish to emphasize, that these results are provided, to proof the concept of the RSP approach, and should not be regarded beyond this context. Although the results can be of practical significance, the main message of the presented results is not that algorithm A is "X" percent better than algorithms B, but that such relative optimization performances for this particular type and problem formulation can be estimated with significant accuracy using the RSP approach, based on calibration data from similar problem types. The particular benchmark results should be relativized by the fact that many different implementations and variations of the compared optimization algorithms exist, which could perform different as the implementations used. Besides that, the response characterization of computationally expensive problems leads to more insight of the problem structure. The main message should be that RSP based performance benchmarks could aid to the selection and tuning of optimization algorithms, in order to increase the efficiency for computationally expensive simulation-based optimization problems.

The authors furthermore would like to highlight: that suitable parameterized benchmark problems are of greater general value than published benchmark results. This is underlined by the fact that the optimization efficiency of algorithms on particular multidisciplinary problem instances can be dependent on problem and formulation properties such as the number of design variables, number and type of design responses, the choice for the objective and constraints, constraint limits (feasibility fraction), and the available function evaluation budget. Flexible parameterized benchmark problems, such as those constructed with the RSP approach, could be useful for the "practitioner" audience since the RSP problem instance parameters can be adopted to resemble the particular problem of interest. For the algorithm "developer" audience standardized RSP instances could be defined for industrially relevant problems, in order to make complex optimization problem types (in terms of simulation expertise, hardware and software resources) more accessible.

Likewise as many other works in the literature dealing with vehicle optimization, the presented study deals only with a subset of all relevant vehicle design objectives and criteria. It should be noted that to design a car suitable for production, more crash scenarios, NVH criteria, as well as structural requirements from other disciplines such as drive dynamics, and structural durability should be considered.

The vehicle models used for the response characterization are of lower mesh resolution than typical industrial models. Although lower mesh resolution models have a much lower accuracy to represent the response of a particular vehicle model, it is assumed that the most general crashworthiness response features can be represented by the models. The vehicle models used for the characterization for the RSP calibration differed in mesh resolution in about an order of magnitude while still significantly consistent response characteristics could be identified, this observation supports the previous assumption and emphasizes the robustness of the approach.

The authors are aware that the "unsmooth" parameter study results obtained from the crashworthiness simulations could also be modeled by a stochastic model with a conventional meta-model as a backbone. In reality, the response criteria for a particular vehicle design seem non-unique due to the uncertainties in a crash experiment. Small design variable perturbations can also trigger such chaotic dynamic response behavior. Depending on the crash simulation solver settings even non-unique solutions can be obtained for the same crash event, only by using another CPU configuration during the numerical solution such as also indicated by Blumhardt (2001) and Duddeck (2008). Indeed for this application type, further work is necessary to take into account aspects relevant for robust design optimization.

In the presented example case studies, the application of the RSP approach is calibrated and corroborated using similar problems (similar in terms of response types, design variable types, design variable range and optimization formulation), on different vehicle models. For other design variables, design criteria, or other applications, the relationship between the design variables and responses might be different, and other basis functions might be more suitable. Although the approach is developed for the presented vehicle design application, the general idea of the approach, the presented response characterization techniques, and the concepts to construct and solve a CSP to incorporate response characteristics for an RSP are however not limited to this particular problem type, and they could be of interest to be applied and tested on a wider range of problems and applications.

For a true generalization, there remain however still open issues. The approach is developed based on empirical investigations, a supporting theory on: which response characteristics affect the representativeness of surrogates on the performance of optimization algorithms, would have to be developed. Although the application example is rather complex, the approach should be tested on other problem types, to investigate the limitations and applicability in a more general scope. The involved response characterization requires many function evaluations and can be of considerable computation cost, but the results could provide valuable insights. If the response characteristics can be represented by meta-models with sufficient accuracy, the approach could be applied indirectly by means of meta-models. Its additional value would be the characterization of the meta-models responses. Other future work regarding the development of the approach could involve an extension to accurately represent the shape, type, and distribution of the Pareto-fronts between different responses, for RSP problems targeting multi-objective optimization problems. Due to the background of the authors, this paper is written and formulated with an “engineering” mindset. Therefore, the authors like to invite practitioners and researchers from other fields to test, develop, generalize or criticize the approach, and in particularly to establish a more theoretical basis in addition to empirical experience on which it is leaning in its present form.

Further investigations on the RSP approach in the context of vehicle design problems could be: to extend or modify the approach to deal with problems with more different design variable types, and to widen the considered design response types and load cases.

As a general outlook on the applications of the RSP approach for the optimization of systems with expensive simulators, future work can involve investigations of additional aspects or response characteristics that influence the optimization efficiency. A first point of interest for future investigations, is the comparison of different optimization frameworks (such as for example Collaborative Optimization, or Analytical Target Cascading). A second point is taking into account the available computational resources to find an efficient optimization strategy. Simulation solvers can be constrained in the number of available parallel licenses, or by the available hardware infrastructure (number and type of nodes, processors memory etc.). Aspects such as the parallelization and scalability of a single function evaluation, combined with the ability of different optimization algorithms, to use parallel function evaluations (using for example a population-based approach) can be explored. Therefore, the RSP approach could aid to find efficient optimization strategies for a particular problem, by enabling a meta-simulation of the optimization process which could take into account a particular resource environment.

8 Summary and conclusions

An approach is presented that could be used to construct computationally affordable synthetic test problems (RSPs), based on response characteristics of computationally expensive real world industrial optimization problems. The approach is developed and tested for the application of multidisciplinary vehicle design problems, involving vibrational comfort and crashworthiness responses, but the applied strategy and used methods are not limited or specific to the application example. The approach is presented in a general way to facilitate the use and testing of the concept to other application fields. A composition of existing analysis methods (parameter studies, sensitivity analysis, Fourier analysis and correlation analysis) is used to identify and quantify typical response characteristics of the simulation responses, with respect to the design variables. Based on the response characterization results, basis functions to represent the responses are selected. The combination of: the basis functions, the function parameters, parameter bounds and the formulation of constraints that enforce selected response structure characteristics, formulate a CSP. Each feasible solution of the CSP provides a set of parameters for which the set of basis functions have response characteristics which are "representative" w.r.t. the selected criteria. These surrogate response functions can be used to formulate surrogate response based optimization problems. The proof of concept and corroboration with an independent vehicle model indicated that for this relatively complex application such RSPs can be used as benchmarks to compare optimization efficiency of different optimization algorithms, and to improve the efficiency of an optimization algorithm by tuning of the optimization meta-parameters. The response characterization required for the RSP construction is computationally expensive, but once made it can provide valuable insights in the problem structure, which could pay off for applications in which many instances of similar problems. No general theory of this novel approach has been formulated yet, neither are the limits of applicability to other problems known. The presented results on are however remarkable and encourage further investigations of the concept. The approach is a step towards systematic analysis of industrially relevant complex black box optimization problems. The authors encourage creative interpretation, application, and critiques of the approach, such that further improvements, in the optimization of complex industrially relevant problems can be achieved.

Acknowledgements

This work is performed in the scope of the GRESIMO and ENLIGHT projects, targeting environmentally friendly mobility solutions. The authors have been partially funded by the European Community's 7th Framework program by means of: an ITN fellowship in the GRESIMO project as part of the People program (Marie Curie Actions) grant agreement no. 290050, and a contribution to the activities in the ENLIGHT project grant agreement no. 314567. Furthermore, the authors are thankful for the publicly available finite element vehicle models used in this work. These models have been developed by the National Crash Analysis Center (NCAC) of The George Washington University under a contract with the FHWA and NHTSA of the US DOT.

References

Alimoradi A, Foley CM, and Pezeshk S (2010) Benchmark Problems in Structural Design and Performance Optimization: Past, Present, and Future—Part I. Structures Congress 2010: pp. 455-466. doi: 10.1061/41131(370)40

Baldanzini N, Caprioli D, and Pierini M (2001) Designing the dynamic behavior of an engine suspension system through genetic algorithms. Journal of Vibration and Acoustics 123.4: 480-486.

Baldanzini N, Scippa A (2004) Shape and size optimization of an engine suspension system. Proceedings of the 2004 International Conference on Noise and Vibration Engineering, ISMA, 2004.

Birge B (2006) Particle Swarm Optimization Toolbox, retrieved from MATLAB file exchange in January 2014 http://www.mathworks.com/matlabcentral/fileexchange/7506-particle-swarm-optimization-toolbox

Blumhardt R (2001) FEM-crash simulation and optimisation. International Journal of Vehicle Design 26.4: 331-347.

Boyd S and Vandenberghe L (2009) Convex optimization. Cambridge University Press, 2009. https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf

Buehren M (2008) Differential evolution, MATLAB file exchange http://www.mathworks.com/matlabcentral/fileexchange/18593-differential-evolution

Deb K, Agrawal S, Pratap A and Meyarivan T (2000) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. Lecture Notes in Computer Science 1917 (2000): 849-858.

Duddeck F (2008) Multidisciplinary optimization of car bodies. Structural and Multidisciplinary Optimization, 35(4), 375–389. doi:10.1007/s00158-007-0130-6

Fletcher R (2010) The sequential quadratic programming method. Nonlinear Optimization. Springer Berlin Heidelberg, 2010. 165-214. DOI: 10.1007/978-3-642-11339-0_3

Goldberg DE and Holland JH (1988) Genetic algorithms and machine learning. Machine learning, 3(2), 95-99

Hallquist, JO (2006) LS-DYNA theory manual. Livermore Software Technology Corporation

Haug E, Scharnhorst T, DuBois P (1986) FEM-Crash, Berechnung eines Fahrzeugaufpralls. VDI-Tagung: Berechnung im Automobilbau, Würzburg, Germany, (VDI-Berichte 613), 479-505

Hoeffding W (1948) A class of statistics with asymptotically normal distribution. The Annals of Mathematical Statistics (1948): 293-325

Durgun I, and Yildiz AR (2012) Structural design optimization of vehicle components using cuckoo search algorithm. Materials Testing 54.3: 185-188.

Kirkpatrick S, Gelatt CD, and Vecchi MP (1983) Optimization by simulated annealing Science 220.4598 (1983): 671-680.

Knowles J, and Hughes EJ (2005). Multiobjective optimization on a budget of 250 evaluations. In Evolutionary Multi-Criterion Optimization (pp. 176-190). Springer Berlin Heidelberg.

Lin S (2011) NGPM A NSGA-II Program in Matlab, MATLAB file exchange. http://www.mathworks.com/matlabcentral/fileexchange/31166-ngpm-a-nsga-ii-program-in-matlab-v1-4

Mihaylova P, Pratellesi A, Baldanzini N, & Pierini M (2012) Optimization of the global static and dynamic performance of a vehicle body by means of response surface models. ASME 2012 11th Biennial Conference on Engineering Systems Design and Analysis. American Society of Mechanical Engineers, 2012.

NCAC Finite Element Model Archive. retrieved from January 2014, from http://www.ncac.gwu.edu/vml/models.html

Plischke E (2010). An effective algorithm for computing global sensitivity indices (EASI). Reliability Engineering & System Safety, 95(4), 354-360.

Prichard D and Theiler J (1994) Generating surrogate data for time series with several simultaneously measured variables. Physical Review Letters 73.7: 951.

Rainer S and Price K (1997) Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization 11.4: 341-359.

Rastrigin LA (1974) Systems of extremal control. Theoretical Foundations of Engineering Cybernetics Series, (In Russian) Nauka, Moscow.

Ratto M and Pagano A (2010) Using recursive algorithms for the efficient identification of smoothing spline ANOVA models. AStA Advances in Statistical Analysis 94.4: 367-388.

Rechenberg I (1973). Evolutionsstrategie. Stuttgart: Holzmann-Froboog. ISBN 3-7728-0373-3.

Rios LM and Sahinidis NV (2012) Derivative-free optimization: A review of algorithms and comparison of software implementations. Journal of Global Optimization (2012): 1-47.

Rodgers LJ and Nicewander WA (1988). Thirteen ways to look at the correlation coefficient. The American Statistician, 42(1), 59-66.

Rosenbrock HH (1960) An automatic method for finding the greatest or least value of a function. The Computer Journal 3.3: 175-184.

Sala R, Pierini M and Baldanzini N (2014a) Optimization efficiency in multidisciplinary vehicle design including NVH criteria. Proceedings of the Leuven Conference on Noise and Vibration Engineering (ISMA); 2014 September 15-17; Leuven, Belgium

Sala R, Pierini M and Baldanzini N (2014b) The development and application of tailored test problems for meta-simulation of multidisciplinary optimization of vehicle structures. Presentation at the (XI) World Congress on Computational Mechanics; 2014 July 20-25; Barcelona, Spain

Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M and Tarantola S (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index Computer Physics Communications 181.2: 259-270.

Schramm U and Pilkey WD (1996), Review: optimal design of structures under impact loading. Shock and Vibration, Vol. 3, pp. 69-81.

Schreiber T and Schmitz A (1996) Improved surrogate data for nonlinearity tests. Physical Review Letters 77.4: 635.

Shan S and Wang GG (2010). Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. Structural and Multidisciplinary Optimization,41(2), 219-241.

Simpson TW, Toropov V, Balabanov V and Viana FA (2008) Design and analysis of computer experiments in multidisciplinary design optimization: a review of how far we have come or not 12th AIAA/ISSMO multidisciplinary analysis and optimization conference. Vol. 5. 2008.

Sobieszczanski-Sobieski J, and Haftka RT (1997) Multidisciplinary aerospace design optimization: survey of recent developments, Structural Optimization 14.1: 1-23.

Sobieszczanski-Sobieski J, Kodiyalam S, and Yang RJ (2001) Optimization of car body under constraints of noise, vibration, and harshness (NVH), and crash. Structural and Multidisciplinary Optimization 22.4: 295-306.

Sobol IM (1990) On sensitivity estimation for nonlinear mathematical models Matematicheskoe Modelirovanie 2.1 (1990): 112-118.

Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and Computers in Simulation 55.1-3: 271-280.

Tang WJ, and Wu QH (2009) Biologically inspired optimization: a review. Transactions of the Institute of Measurement and Control 31.6 (2009): 495-515.

Varis T and Tuovinen T (2012) Open Benchmark database for multidisciplinary optimization problems. Proceedings of the International Conference on Modeling and Applied Simulation

The final publication is available at Springer via http://dx.doi.org/10.1007/s00158-016-1410-9

Venema V (2003) IAAFT implementation in MATLAB retrieved January 2014 from http:\\www.meteo.uni-bonn.de\victor

Venkayya VB (1978) Structural optimization: a review and some recommendations International Journal for Numerical Methods in Engineering 13.2: 203-228.

Wolpert DH and Macready WG (1995) No free lunch theorems for search. Vol. 10. Technical Report SFI-TR-95-02-010, Santa Fe Institute

Wolpert DH and Macready WG (1997) No free lunch theorems for optimization Evolutionary Computation, IEEE Transactions on 1.1 (1997): 67-82.

Wu SR and Gu L (2012) Introduction to the Explicit Finite Element Method for Nonlinear Transient Dynamics. John Wiley & Sons.

Yang RJ, Gu L, Tho CH, Sobieski J (2001) Multi-disciplinary optimization of a full vehicle with high performance computing. In: Conf. of the American Inst. of Aeronautics and Astronautics, pp 688–698, AIAA Paper No. AIAA- 2001- 1273

Yang RJ, Tseng L; Nagy L; Cheng J (1994) Feasibility study of crash optimization. In: Gilmore BJ, Hoetzel DA, Dutta D, Eschenauer HA (eds.) Advances in design automation, DE-Vol 69-2 pp, 549-556. ASME

Yang XS (2010a) Nature-inspired metaheuristic algorithms. Luniver press, 2010. http://web.info.uvt.ro/~dzaharie/cne2013/proiecte/tehnici/FireflyAlgorithm/Yang_nature_book_part.pdf

Yang XS (2010b), Engineering optimization: an introduction with metaheuristic applications. John Wiley & Sons

Yildiz AR, and Solanki KN (2012) Multi-objective optimization of vehicle crashworthiness using a new particle swarm based approach The International Journal of Advanced Manufacturing Technology 59.1-4: 367-376.

Zitzler E, Deb K and Thiele L (2000) Comparison of multiobjective evolutionary algorithms: Empirical results. Evolutionary computation 8.2: 173-195.