HAL open science

Sequential designs for sensitivity analysis of functional inputs in computer experiments

Jana Fruth, Olivier Roustant, Sonja Kuhnt

To cite this version:

Jana Fruth, Olivier Roustant, Sonja Kuhnt. Sequential designs for sensitivity analysis of functional inputs in computer experiments. 2014. hal-00943509

HAL Id: hal-00943509

https://hal.science/hal-00943509

Preprint submitted on 18 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sequential designs for sensitivity analysis of functional inputs in computer experiments

J. Fruth[a,*], O. Roustant[b], S. Kuhnt[a,c]

[a] *Faculty of Statistics, TU Dortmund University, Vogelpothsweg 87, 44227 Dortmund, Germany*
[b] *Ecole Nationale Supérieure des Mines, FAYOL-EMSE, LSTI, F-42023 Saint-Etienne, France*
[c] *Dortmund University of Applied Sciences and Arts, Emil-Figge-Strasse 42, 44227 Dortmund, Germany*

Abstract

Computer experiments are nowadays commonly used to analyze industrial processes aiming at achieving a wanted outcome. Sensitivity analysis plays an important role in exploring the actual impact of adjustable parameters on the response variable. In this work we focus on sensitivity analysis of a scalar-valued output of a time-consuming computer code depending on scalar and functional input parameters. We investigate a sequential methodology, based on piecewise constant functions and sequential bifurcation, which is both economical and fully interpretable. The new approach is applied to a sheet metal forming problem in three sequential steps, resulting in new insights into the behavior of the forming process over time.

Keywords: Sensitivity analysis, Functional input, Sequential approach, Sheet metal forming, Springback

1. Introduction

In this work a general method is presented to study the influence of functional input and scalar parameters on a scalar output of a time-consuming computer model, where the interest lies in the behavior of the sensitivity over

*Corresponding author

Email addresses: `fruth@statistik.tu-dortmund.de` (J. Fruth), `roustant@emse.fr` (O. Roustant), `sonja.kuhnt@fh-dortmund.de` (S. Kuhnt)

the functional domain. For illustration purposes, let us consider the particular application in sheet metal forming that originally motivated this research. Here a scalar output, springback, depends on two functional inputs: friction and blankholder force. The aim is to analyze the sensitivity of springback as a measure of geometrical accuracy to the input variables not only as a whole but as functions over time corresponding to the different stages of the forming process. Let us look at a few preview runs, where only the friction between tool and metal is varied (Tab. 1). In the first two runs, the friction is constant over the 15 seconds of the punch travel. Then different functional input settings are used, indicating that varying friction can dramatically reduce springback. Although the four last runs have equal mean friction over time, we get very clear differences in the springback results. This clearly motivates the exploration of the functional behavior of the process, and more generally the sensitivity analysis of functional input parameters.

The most common method to study data with functional input is functional linear regression (see Ramsay and Silverman [1]), a framework to approximate, analyze and predict data with functional input. However, a design of experiments is not considered – data are assumed as already provided – and the interpretation of the influence of different time regions is not easy (James et al. [2]). Furthermore, it is restricted to the assumption of linearity. Morris [3] introduced a Gaussian process model for the analysis of computer models with functional input and output, taking also the design of the functional experiments into account. However, he did not consider sensitivity analysis. Three ideas for the sensitivity analysis of computer experiments with one functional input are presented in Iooss and Ribatet [4]. They result in one uncertainty index for the input, and thus summarize the functional domain without giving an insight into the sensitivity with respect to changes at specific time intervals during the process.

In this paper, sensitivity analysis of functional inputs is obtained by first transforming the functional problem into a scalar one by considering the basis of piecewise constant functions, often suggested in functional data analysis (see e.g. Ramsay and Silverman [1]). That framework also allows to consider both scalar and functional inputs. Then we adapt the ideas of sequential bifurcation (Bettonvil [5]) in order to analyze more and more locally the effect of one specific part of the functional domain. For that purpose, we introduce a normalized sensitivity index, which allows comparison between two different steps, and investigate at a theoretical level its properties.

The paper is structured as follows. Sections 2 and 3 present the method-

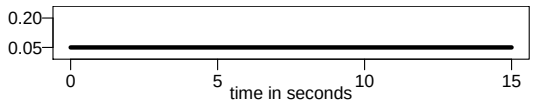
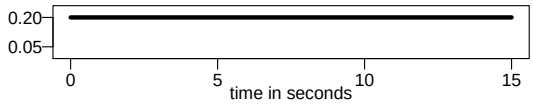
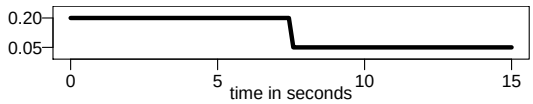
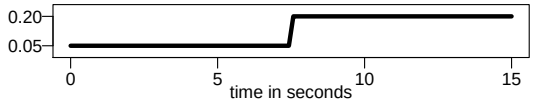
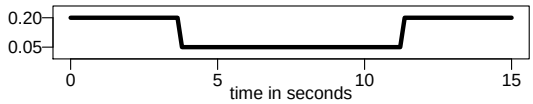
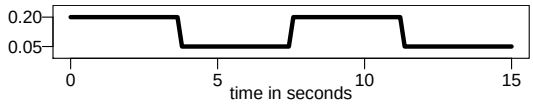
functional input: friction (unitless)	scalar output: springback (mm)
	6.1
	2.1
	14.7
	1.4
	0.8
	6.0

Table 1: Preview runs to compare different functional settings of the parameter friction.

ology, together with properties and a way for graphical presentation. Section 4 describes the sequential design of the functional input. An application to sheet metal forming is given in Section 5, exploring the sensitivity of the time intervals of two functional inputs on springback. A discussion concludes the paper.

2. Framework

Our aim is to analyze an experiment involving d_s scalar input variables $x_i \in [-1, 1]$, $i = 1, \ldots, d_s$ and d_f functional input variables $g_j : D_j \mapsto [-1, 1]$, $j = 1, \ldots, d_f$ with a scalar output $Y \in \mathbb{R}$. For convenience the input parameters, scalar as well as functional, are bounded to fall into $[-1, 1]$, and the input functions are defined on $D_j = [0, 1]$ for each $j = 1, \ldots, d_f$. The functions depend on a single argument, which in practical applications

might often be the time, but the approach is not limited to it. The argument is not necessarily the same for all functional inputs. Beside the boundedness, there are no further conditions on the shape of the functions. Input and output parameters are connected by an unknown black-box function $f : [-1,1]^{d_s} \times \mathcal{F}_{[0,1]}^{d_f} \mapsto \mathbb{R}$,

$$Y = f\left(x_1, \ldots, x_{d_s}, g_1, \ldots, g_{d_f}\right),$$

where $\mathcal{F}_{[0,1]}$ denotes the space of all functions on $[0,1]$. We assume this black-box function to be very expensive or time-consuming to execute.

Due to functional inputs, the input domain is infinite dimensional. To cope with this issue, we propose to explore subintervals of the functional domains as a whole instead of looking at each point. The subintervals are then chosen sequentially to increasingly finer divisions. This approach is explained later in Section 4. For now, say we have a decomposition of the domain of each functional input g_j into $p_j \in \mathbb{N}^+$ subintervals at splitting points $\mathbf{a}_j = (a_j^0, \ldots, a_j^{p_j})$ with $0 = a_j^0 < a_j^1 < \cdots < a_j^{p_j-1} < a_j^{p_j} = 1$,

$$D_j = [a_j^0, a_j^1[\ \cup \ [a_j^1, a_j^2[\ \cup \ \ldots \ \cup \ [a_j^{p_j-1}, a_j^{p_j}].$$

We restrict each g_j to belong to $V_{\mathbf{a}_j}$, the space of piecewise constant functions over the subintervals defined by $\mathbf{a}_j$,

$$V_{\mathbf{a}_j} = \left\{ Z_j^{(1)} \mathbb{1}_{[0,a_j^1[}(t) + \cdots + Z_j^{(p_j)} \mathbb{1}_{[a_j^{p_j-1},1]}(t), \text{ with } Z_j^{(k)} \in [-1,1], 1 \leq k \leq p_j \right\}. \tag{1}$$

For a given set of splitting points $\mathbf{a}_j$ an element of $V_{\mathbf{a}_j}$ is then defined by the vector $(Z_j^{(1)}, \ldots, Z_j^{(p_j)})$. Thus, we obtain a transformation of the input space of the black-box function from functional to scalar,

$$\begin{aligned} Y &= f\left(x_1, \ldots, x_{d_s}, g_1, \ldots, g_{d_f}\right) \\ &= \tilde{f}_{a^1,\ldots,a^{p_{d_f}}}\left(x_1, \ldots, x_{d_s}, Z_1^{(1)}, \ldots, Z_1^{(p_1)}, \ldots, Z_{d_f}^{(1)}, \ldots, Z_{d_f}^{(p_{d_f})}\right). \end{aligned} \tag{2}$$

The described way to design the functional input can also be regarded in the context of function representation by basis functions (see e.g. Ramsay and Silverman [1], Chapter 3). They can be viewed as B-splines of order

one: linear combinations of piecewise constant functions being constant over one interval and zero otherwise (Ramsay and Silverman [1], de Boor [6]). They are also connected to wavelet theory, especially to Haar-Wavelets (see e.g. Walker [7]), where the basic functions are piecewise constant functions, being scaled and shifted to the desired functional form. In this work we restrict ourselves to the basis of piecewise constant functions which allows a direct and easy interpretation. Nevertheless, we will borrow from wavelet theory the idea of splitting the domain sequentially. Hence, we will consider a sequence of embedded spaces $V_{\mathbf{a}_j} \subseteq V_{\mathbf{a}'_j} \subseteq V_{\mathbf{a}''_j} \subseteq \dots$ corresponding to refined time intervals.

A unified framework for scalar and functional inputs. Let us remark that the space of piecewise constant functions also includes scalar inputs, through the one-to-one correspondence:

$$\begin{aligned} [-1,1] &\rightarrow V_{(0,1)} \\ x &\rightarrow x\mathbb{1}_{[0,1]}(t) \end{aligned}$$

In other words any scalar input with value equal to x can be considered as a functional input with a constant value over $[0,1]$ equal to x. This allows to consider only functional inputs, having in mind that scalar inputs correspond to functional inputs that are kept constant over $[0,1]$, and thus never split. We adopt this point of view from now on and denote by $d = d_s + d_f$ the total number of inputs.

3. Sensitivity Indices

With the framework of Section 2, unifying the case of scalar and functional inputs, it is possible to perform sensitivity analysis on

$$Y = \tilde{f}_{a^1,\dots,a^{p_d}}\left(Z_1^{(1)},\dots,Z_1^{(p_1)},\dots,Z_d^{(1)},\dots,Z_d^{(p_d)}\right)$$

A specific sensitivity analysis method can be chosen according to the supposed behavior and evaluation costs of the experiment among methods such as regression analysis, Morris screening, Sobol indices or others (see e.g. Saltelli et al. [8] or Confalonieri et al. [9] for an overview). Since the black-box function is assumed to be expensive, we have a very low budget. Thus, we suggest the method of regression analysis and use it in the further explorations. Nevertheless, the sequential design strategy proposed in Section 4

can be adapted to other methods.

Design of experiments for regression analysis usually only cover few levels of the input values, for instance two extreme values, encoded by -1 and $+1$. An equal scaling of the input factors is important to preserve comparability. Saltelli et al. [8] even suggest scaling input and output to have a mean of zero and variance of one, that is to use standardized regression coefficients. A linear model is considered

$$Y = \alpha + \sum_{j=1}^{d} \sum_{k=1}^{p_j} \beta_j^{(k)} Z_j^{(k)} + \delta. \tag{3}$$

The coefficients α and β_j^k are commonly computed by the method of Least Squares using a design matrix $\mathbf{Z}$ on the values of $Z_j^{(k)}$ sufficient for estimation, i.e. such that $\mathbf{Z}'\mathbf{Z}$ is invertible (see e.g. Saltelli et al. [8], p. 124). The estimates $\widehat{\alpha}$ and $\widehat{\beta}_j^{(k)}$ are obtained by the well-known Least Squares formula $(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$, with $\mathbf{y}$ being the vector of the scalar outputs. As the model is deterministic no assumptions are made for δ, the difference between response and regression model and the Least Squares approach is interpreted as simple curve fitting. The estimates $\widehat{\beta}_j^{(k)}$ can then be used as sensitivity indices as they indicate the linear influence of the corresponding scalar input or functional interval on the output. Additionally, depending on the budget, higher order effects of interest can be considered, especially the interaction effect between intervals of the same functional input, leading to the model

$$Y = \alpha + \sum_{j=1}^{d} \sum_{k=1}^{p_j} \beta_j^{(k)} Z_j^{(k)} + \sum_{j=1}^{d} \sum_{1 \le k < k' \le p_j} \beta_j^{(kk')} Z_j^{(k)} Z_j^{(k')} + \delta. \tag{4}$$

3.1. Normalization

The drawback of $\widehat{\beta}$ as a sensitivity estimate is that it depends on the chosen decomposition of the functional input domain. If for instance two adjoint intervals have a common effect, we expect their combined interval to have the same effect. This can be achieved by normalizing by the interval size.

Definition 1. *As in Section 2, let us consider a set of splitting points* $\mathbf{a}_1, \ldots, \mathbf{a}_{p_d}$ *and assume that* $g_j \in V_{\mathbf{a}_j}, j = 1, \ldots, d : g_j(t) = \sum Z_j^{(k)} \mathbb{1}_{[a_j^{k-1}, a_j^k[}(t)$.

Denote by $\widehat{\beta}_j^k$ and $\widehat{\beta}_j^{(kk')}$ the estimated first-order and second-order regression coefficients, then we define by

$$\widehat{H}_j^k = \frac{\widehat{\beta}_j^{(k)}}{a_j^k - a_j^{k-1}} \quad \textit{and} \quad \widehat{H}_j^{kk'} = \frac{\widehat{\beta}_j^{(kk')}}{(a_j^k - a_j^{k-1})(a_j^{k'} - a_j^{k'-1})},$$

the so-called normalized regression index of $Z_j^{(k)}$ and the normalized interaction regression index of $Z_j^{(k)}$ and $Z_j^{(k')}$ resp. for $j \in \{1, \ldots, d\}, 1 \leq k < k' \leq p_j$.

Thus $\widehat{H}$ can be interpreted as the change in the output when increasing the functional input over the corresponding interval by one unit of time.

Let us consider some particular shapes for the underlying unknown model f, depending on only one functional input g. The shapes are ideal in the sense that a pure linear behaviour is assumed for f. It can be shown that in these cases the returned functional indices have desirable properties. The corresponding proofs are given in the Appendix.

Proposition 1. *Let f be the integral over the functional input g weighted by an integrable function $w : [0, 1] \mapsto \mathbb{R}$*

$$f(g) = \alpha + \int_0^1 w(t)g(t)\, dt,$$

with $\alpha \in \mathbb{R}$. Then for any splitting of the functional domain of g, $\mathbf{a} = (a^0, \ldots, a^p)$ the normalized regression indices return the mean values of the weight function over the interval and each interaction index is zero.

$$\widehat{H}^k = \frac{\int_{a^{k-1}}^{a^k} w(t)\, dt}{a^k - a^{k-1}} \quad \textit{and} \quad \widehat{H}^{kk'} = 0, \quad k, k' = 1, \ldots, p.$$

Asymptotically, if the interval size goes to zero the normalized regression indices return the exact value of the weight function

$$\lim_{\delta \to 0} \frac{\int_t^{t+\delta} w(u)\, du}{\delta} = w(t).$$

In the same way interactions between time regions are recaptured.

Proposition 2. *Let f be the product of two integrals of g over two different time regions $[a^{i^*-1}, a^{i^*}]$ and $[a^{j^*-1}, a^{j^*}]$ with $c, \alpha \in \mathbb{R}$*

$$f(g) = \alpha + c \int_{a^{i^*-1}}^{a^{i^*}} g(t)\, dt \times \int_{a^{j^*-1}}^{a^{j^*}} g(t)\, dt.$$

Then for any splitting $\mathbf{a}$ of the functional domain of g with $\left\{a^{i^*-1}, a^{i^*}, a^{j^*-1}, a^{j^*}\right\} \in \mathbf{a}$, the estimates of the indices and the interaction indices hold

$$\widehat{H}^k = 0, k = 1, \ldots, p \quad \textit{and} \quad \widehat{H}^{kk'} = \begin{cases} c, & k = i^*, k' = j^*, \\ 0, & \textit{otherwise.} \end{cases}$$

Furthermore we can show that the definition is robust against non-linear transformations of the input g in the computer model.

Proposition 3. *If g is transformed by a not necessarily linear, but strictly monotonic increasing function $\zeta : [0, 1] \mapsto \mathbb{R}$*

$$f(g) = \alpha + \int_0^1 w(t)\zeta(g(t))\, dt$$

then for a two-level design the values $\widehat{H}^k$ are linear in the mean values of the weight function, and the weighting function still determines the importance of the intervals compared to each other. Thus, there exists $\lambda > 0$ such that

$$\widehat{H}^k = \lambda \frac{\int_{a^{k-1}}^{a^k} w(t)\, dt}{a^k - a^{k-1}}, \quad k = 1, \ldots, p.$$

3.2. Interpretation and graphical representation

The indices can be represented by plotting them on the time scale of each functional input. An artificial example of such a visualization is depicted in Fig. 1. A positive bar size indicates that a raise of the input at this time interval causes an increase in the mean output, a negative size causes a decrease. Here, one can directly see the increasing influence with the time. Yet, care has to be taken in the interpretation, as the indices represent mean values, as mentioned above in Prop. 1. In an extreme case, the joint interval of two intervals with influences of different signs could show zero influence when the intervals cancel each other out. So any interpretation of the indices can only relate to the mean influence of the interval.

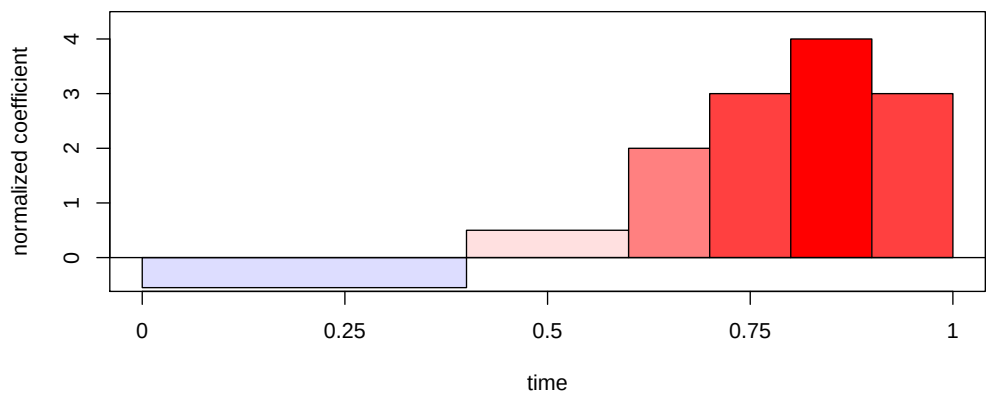


Figure 1: Graphical representation of functional sensitivity indices.

4. Splitting and design

4.1. Sequential splitting approach

To keep the number of runs as low as possible, we propose a sequential design approach, where runs are saved by omitting regions of 'low interest' in the analysis. The approach is based on group factor screening, a method to identify important factors with very small designs by grouping factors and then sequentially dividing active groups (Watson, G. S. [10] or for an overview Morris [11]).

For easier readability, we consider only one functional input $g : D \mapsto [-1, 1]$ as sole input, i.e. $d = 1$. The method is easily extended to more inputs by considering them as addditional groups. For a specific iteration step r we denote the splitting points

$$\mathbf{a}^r = (a^{0,r}, \ldots, a^{p^r,r}),\ 0 = a^{0,r} < a^{1,r} < \cdots < a^{p^r-1,r} < a^{p^r,r} = 1$$

and the space of piecewise constant functions

$$V_{\mathbf{a}^r} = \left\{ Z^{(1,r)} 1_{[0,a^{1,r}[}(t) + \cdots + Z^{(p^r,r)} 1_{[a^{p^r-1,r},1]}(t), Z^{(k,r)} \in [0,1] \right\}.$$

The idea of group factor screening is transferred to functional input screening by interpreting the intervals as groups of factors. Interesting intervals

are then split into smaller intervals corresponding to dividing active groups of factors. The algorithm runs as follows: Start with a low number p^1 of intervals, $\mathbf{a}^1 = (a^{0,1}, \ldots, a^{p^1,1})$ and choose the suitable design for the corresponding variables $Z^{(1,1)}, Z^{(2,1)}, \ldots, Z^{(p^1,1)}$. Then perform the experiments and compute the indices $\widehat{H}^{1,1}, \ldots, \widehat{H}^{p^1,1}$. From the results decide which intervals shall be explored further in the second step and which can be ignored. We recommend to take the decision in close cooperation with experts. Generally, an interval should be seen as important, if its index is bigger than an assumed approximation error. However, it has to be kept in mind, that the indices are only mean values as explained in Section 3.2. Consequently, intervals where a change in sign can be suspected should rather be explored further, even if the index is small. All intervals of interest are then split in the middle, that is in the second step, the vector of splitting points $\mathbf{a}^2$ will include all points of $\mathbf{a}^1$ and for each interval $\left[a^{1,k-1}, a^{1,k}\right]$ to be split at the point $\frac{a^{k-1,1}+a^{k,1}}{2}$. Then a next design allowing for the analysis of these new intervals is set up. The procedure of estimating and splitting is repeated until the functional domain is satisfyingly explored.

4.2. Sequential designs

In the literature on group factor screening, there are various ways on how to design the groups at the different steps, which differ e.g. in their assumptions, orthogonality, the treatment of interactions, or the way of reusing. A very economical way to reuse the points is sequential bifurcation (Bettonvil [5]). Let us first consider the case of a linear model without interactions. Adapted to our functional framework the design is set in the following way. The factors are designed on two extreme levels, encoded to -1 and 1. It starts with only one interval, corresponding to a constant curve, one set to -1 and one to 1. Using the result of the two runs, the normalized regression index of the whole function can be estimated. Then the interval is split in the middle, resulting in two intervals in the second step. Only one additional run is required to estimate the coefficients of both intervals, where we set the first interval to $+1$ and the second interval to -1. Values for $(+1, +1)$, $(-1, -1)$ are already known from the previous step. The result for $(-1, +1)$ would just be a linear combination of the responses from the other 3 runs and is not needed for the calculation of the indices. Generally, for each split interval, only one run is required to estimate the influence of both new intervals, the one with $+1$ up to the cut point and -1 from there is used. The approach is depicted in Fig. 2. It shows three sequential steps requiring

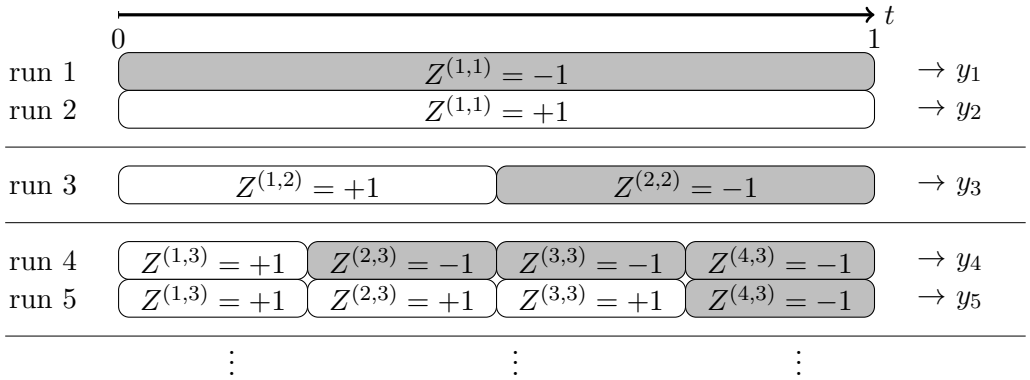


Figure 2: Scheme for the design based on sequential bifurcation. White and grey shading indicate the setting to $+1$ and -1.

a total of 5 runs. In the first step, the mean influence of the whole interval can be estimated by $\widehat{H}^{1,1} = \frac{0.5(y_2-y_1)}{1}$. In the second step the result of run 3 can be used to estimate $\widehat{H}^{1,2} = \frac{0.5(y_3-y_1)}{0.5}$ and $\widehat{H}^{2,2} = \frac{0.5(y_2-y_3)}{0.5}$. Splitting both intervals, from run 4 and run 5 we get $\widehat{H}^{1,3} = \frac{0.5(y_4-y_1)}{0.25}$, $\widehat{H}^{2,3} = \frac{0.5(y_3-y_4)}{0.25}$, and so on.

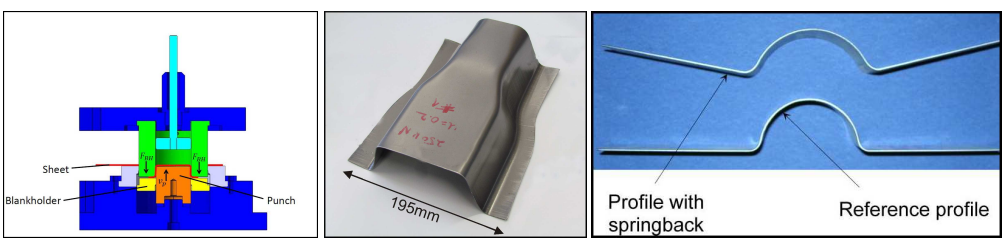
In presence of interactions, Bettonvil [5] suggest to add mirror runs to the design, that means adding to each run another run with opposite signs in order to obtain unbiased estimates of the main effect coefficients.

If orthogonality and/or the possibility to estimate interactions are required, factorial or fractional factorial designs are further design options. A new such design is then constructed in each step r on all factors of interest $Z^{(1,r)}, \ldots, Z^{(p^r,r)}$, all non-interesting intervals are set to a constant value, e.g. 0, in the design and thus are not regarded in the sensitivity analysis any more. Here again, runs from former steps can be reused. It is easy to show that for full factorial designs in a step where all intervals have been split, half of the required runs can be reused. For fractional factorial designs, the reuse possibilities depend strongly on which effects are confounded in the sequential designs, as different confounding requires different runs.

5. Application

Deep drawing is an important sheet metal forming process, especially in the automotive industry. The main aim of deep drawing is to form a flat sheet metal into the desired shape by pressing it with a punch into the die. An example for such a press is shown in Fig. 3 (a) and the formed workpiece can be seen in Fig. 3 (b). In the deep drawing, special care has to be taken in terms of shape accuracy, especially springback, which leads to serious problems during the assembly of different parts. Such an inaccuracy in flange is depicted in Fig. 3 (c) at the top. The increased use of advanced high strength steels, due to their advantages of light weight and higher strength, has additionally enhanced the requirements to the shape accuracy, as they tend to have higher springback.

The forming process can be adjusted by several parameters, concerning for instance metal properties, the geometry of the die or process parameters. Here we consider two of the process parameters, blankholder force and friction. Blankholder force is the pressure of the blankholder, which keeps the metal flange on its place during the forming (see Fig. 3, left). Friction measures the friction between tool and metal at various points. The pro-



(a) (b) (c)

Figure 3: Sheet metal forming press (a), example of a resulting formed metal (b), formed parts to demonstrate springback (c).

cess is usually studied by complex Finite Element simulation. In previous studies the parameters could only be kept on a fix value during the forming process (Gösling et al. [12]). Adjustments to the machines and to the simulation tool give us now the possibility to temporally change two of the inputs, blankholder force and friction, that is to vary their values during the punch travel.

Hence, there are no scalar inputs and 2 functional inputs, friction and blankholder force. We performed three steps of functional sensitivity analysis

following the design based on sequential bifurcation of Section 4.2, taking 24 runs in total. We refer to the Appendix for the complete design. Figure 4 shows the results of the three steps. In the first step both functional inputs are split into two intervals. A stronger influence of friction than blankholder force is visible as well as a general tendency towards a positive influence in the first half and negative influence in the second half for both. As all intervals are at this point interesting to the engineers, all are split in the second step. There it can be seen that for both factors, the last interval has the greatest impact on the springback. On the other hand, the first and third intervals have relatively small influence. Therefore, in the third step of the sequential analysis, these intervals are not explored further and each second and fourth interval is split again. Overall a similar behavior of both factors can be clearly seen. A clear break in about half of the time can be observed. It is also noted that the magnitude of the influence increases towards the end of both halves. These intervals are most important for the springback behavior.

6. Conclusion

This paper has proposed a solution to perform sensitivity analysis for time-consuming computer experiments with both scalar and functional inputs, focusing not only on the sensitivity of the functional inputs as a whole but on the sensitivity behavior over the functional domain as well. Our motivation came from an industrial sheet metal forming simulation where the behavior of two process parameters over the forming process should be explored.

We proposed a framework based on piecewise constant functions, applicable for scalar and functional inputs. The method's idea is to reduce the dimension by dividing the functional domain into subintervals. The subintervals are then sequentially screened following methods from group factor screening, resulting in an increasingly fine decomposition of the interesting parts of the functional domains. Normalized regression indices were proposed as sensitivity indices. They meet desirable requirements to represent the behavior over the functional domain.

The method shows some limitations, due to the strongly limited number of runs. The indices capture only linear behavior, and thus possibly ignore non-linear influences. Though we showed that under some circumstances the method is robust against a direct non-linear transformation, global sensitiv-

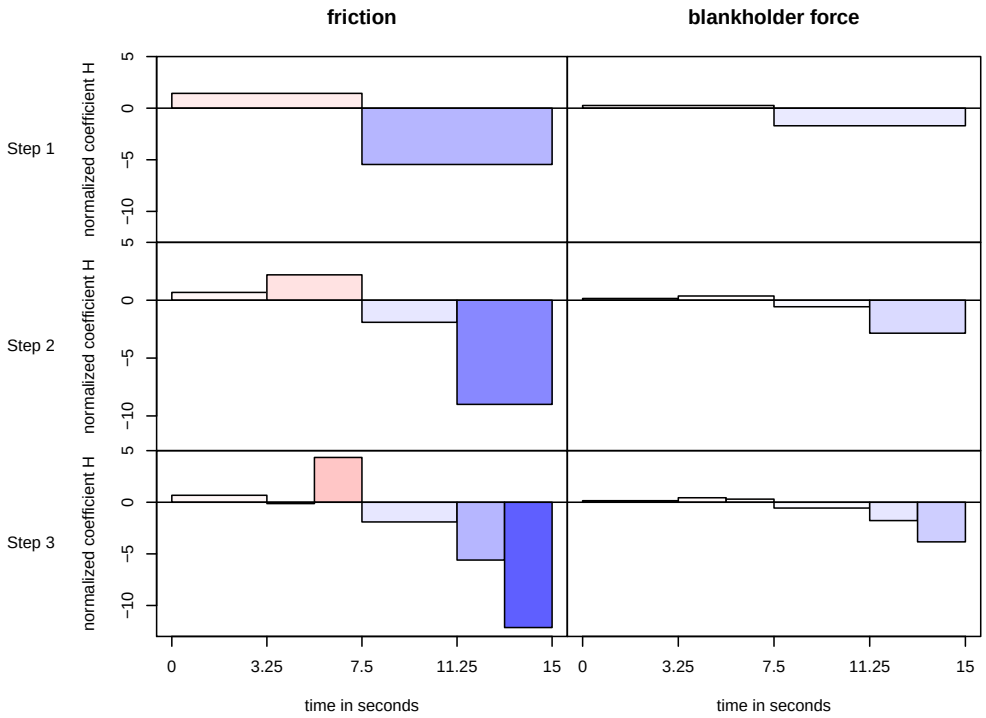


Figure 4: Three steps of the sheet metal application. First step (top). Second step (middle). Third step (bottom). The bar colour references sign and amplitude of the bar. Recall that these amplitudes are comparable to each other at each step.

ity indices like Sobol indices can as well be employed in the method and are a good, but more expensive alternative in strong non-linear cases. Another limitation concerns the interpretation of the indices, a well-known issue in group factor screening. As they can only be interpreted as mean values over the subinterval, theoretically strong sensitivities of different sign inside one subinterval could get canceled out. In our experiences, this problem could be handled well by user knowledge.

Acknowledgements

We gratefully thank Hamad ul Hassan for the conduction of the sheet metal forming simulations and his engineering insights, and David Steinberg for helpful hints about group factor screening. This paper is based on investigations of the collaborative research centre SFB 708, project C3. The

authors would like to thank the Deutsche Forschungsgemeinschaft (DFG) for funding. All the statistical parts of this work have been performed within the R environment.

Appendix A. Proofs

In all three propositions, recall that g is a piecewise constant function

$$g(t) = \sum_{k=1}^{p} Z^{(k)} \mathbb{1}_{[a^{k-1},a^k[}(t).$$

Proof of Prop. 1. As g is bounded and w is integrable, wg is integrable.

$$\begin{aligned} f(g) =& \alpha + \int_0^1 w(t)g(t)\,dt, \\ =& \alpha + Z^{(1)} \int_{a^0}^{a^1} w(t)\,dt + \cdots + Z^{(p)} \int_{a^{p-1}}^{a^p} w(t)\,dt \\ =& (\mathbf{1}, Z^{(1)}, \ldots, Z^{(p)})\eta \end{aligned}$$

with $\eta = \left(1, \int_{a^0}^{a^1} w(t)\,dt, \ldots, \int_{a^{p-1}}^{a^p} w(t)\,dt\right)'$. Now using a design matrix $\mathbf{Z}$ on the values of $Z^{(1)}, \ldots, Z^{(p)}$ such that $(\mathbf{Z}'\mathbf{Z})$ is invertible, and denoting by $\mathbf{y}$ the output, we have $y = \mathbf{Z}\eta$. The Least Squares coefficients are given by

$$\widehat{\beta} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y} = \eta,$$

which leads to

$$\widehat{H}^k = \frac{\int_{a^{k-1}}^{a^k} w(t)}{a^k - a^{k-1}}.$$

Finally, as there are no interaction terms in $\mathbf{Z}$, we have $\widehat{H}^{kk'} = 0, \quad k, k' = 1, \ldots, p.$. □

Proof of Prop. 2.

$$\begin{aligned} f(g) =& \alpha + c \int_{a^{i^*-1}}^{a^{i^*}} g(t)\,dt \times \int_{a^{j^*-1}}^{a^{j^*}} g(t)\,dt \\ =& \alpha + c \times Z^{(i^*)}(a^{i^*} - a^{i^*-1}) \times Z^{(j^*)}(a^{j^*} - a^{j^*-1}) \end{aligned}$$

Then similarly to Prop. 1 we obtain $\widehat{\beta}^k = 0$, $k = 1, \ldots, p$ and

$$\widehat{\beta}^{kk'} = \begin{cases} c(a^{i^*} - a^{i^*-1})(a^{j^*} - a^{j^*-1}), & k = i^*, k' = j^*, \\ 0, & \text{otherwise.} \end{cases}$$

leading to the result. □

Proof of Prop. 3.

$$\begin{aligned} f(g) =& \alpha + \int_0^1 w(t)\zeta(g(t))\, dt, \\ =& \alpha + \zeta(Z^{(1)}) \int_{a^0}^{a^1} w(t)\, dt + \cdots + \zeta(Z^{(p)}) \int_{a^{p-1}}^{a^p} w(t)\, dt. \end{aligned}$$

Now for a design on two levels {-1,1}, we can define λ and $\kappa \in \mathbb{R}$, such that

$$\zeta(z) = \lambda z + \kappa \quad \text{for the two values } z = -1, z = 1.$$

Since ζ is increasing, we have $\lambda > 0$. It follows for all $Z^{(k)} \in \{-1, 1\}, k = 1, \ldots, p$ that

$$\begin{aligned} f(g) =& \alpha + (\lambda Z^{(1)} + \kappa) \int_{a^0}^{a^1} w(t)\, dt + \cdots + (\lambda Z^{(p)} + \kappa) \int_{a^{p-1}}^{a^p} w(t)\, dt \\ =& \left[\alpha + \kappa \sum_{k=1}^{p} \int_{a^{k-1}}^{a^k} w(t) dt \right] + \sum_{k=1}^{p} \left[\lambda \int_{a^{k-1}}^{a^k} w(t) dt \right] Z^{(k)} \end{aligned}$$

Then for the computations of the coefficients, we obtain

$$\begin{aligned} \widehat{\beta}^k =& \lambda \int_{a^{k-1}}^{a^k} w(t), \quad k = 1, \ldots, p \\ \Rightarrow \widehat{H}^k =& \lambda \frac{\int_{a^{k-1}}^{a^k} w(t)}{a^k - a^{k-1}}, \quad k = 1, \ldots, p. \end{aligned}$$

□

Appendix B. Application data

Sequential bifurcation design on 2 factors: for friction 0.05 (–) and 0.2 (+), for blankholder force 67.5 (–) and 125 (+). Mirror runs are shaded.

step	run	time intervals: friction								time intervals: blankholder force								y
		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	
1	1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	6.2
	2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.7
	3	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	2.1
	4	–	–	–	–	–	–	–	–	+	+	+	+	+	+	+	+	4.7
	5	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	9.1
	6	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	0.7
	7	+	+	+	+	+	+	+	+	+	+	+	+	–	–	–	–	2.3
	8	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	4.3
2	9	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	6.8
	10	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.7
	11	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–	–	6.9
	12	–	–	–	–	–	–	+	+	+	+	+	+	+	+	+	+	0.5
	13	+	+	+	+	+	+	+	+	+	+	–	–	–	–	–	–	2.2
	14	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	4.7
	15	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	1.9
	16	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	4.5
3	17	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	6.8
	18	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	0.7
	19	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–	4.2
	20	–	–	–	–	–	–	–	+	+	+	+	+	+	+	+	+	0.5
	21	+	+	+	+	+	+	+	+	+	+	+	–	–	–	–	–	2.2
	22	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	4.4
	23	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	0.7
	24	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	4.2

References

[1] J. O. Ramsay, B. W. Silverman, Functional data analysis, Springer series in statistics, Springer, New York, 1997.

[2] G. M. James, J. Wang, J. Zhu, Functional linear regression that's interpretable, The Annals of Statistics 37 (2009) 2083–2108.

[3] M. D. Morris, Gaussian surrogates for computer models with time-varying inputs and outputs, Technometrics 54 (2012) 42–50.

[4] B. Iooss, M. Ribatet, Global sensitivity analysis of computer models with functional inputs, Reliability Engineering & System Safety 94 (2009) 1194–1204.

[5] B. Bettonvil, Factor screening by sequential bifurcation, Communications in Statistics-Simulation and Computation 24 (1995) 165–185.

[6] C. de Boor, A practical guide to splines, volume 27 of *Applied mathematical sciences*, Springer, New York, 2001.

[7] J. S. Walker, A primer on wavelets and their scientific applications, Chapman & Hall/CRC, London, 1999.

[8] A. Saltelli, K. Chan, E. M. Scott (Eds.), Sensitivity analysis, Wiley series in probability and statistics, Wiley, Chichester, 2000.

[9] R. Confalonieri, G. Bellocchi, S. Bregaglio, M. Donatelli, M. Acutis, Comparison of sensitivity analysis techniques: A case study with the rice model warm, Ecological Modelling 221 (2010) 1897–1906.

[10] Watson, G. S., A study of the group screening method, Technometrics 3 (1961) 371–388.

[11] M. D. Morris, An overview of group factor screening, in: A. Dean, S. Lewis (Eds.), Screening methods for experimentation in industry, drug discovery, and genetics, Springer, New York, 2006, pp. 191–206.

[12] M. Gösling, H. Kracker, A. Brosius, U. Gather, A. E. Tekkaya, Study of the influence of input parameters on a springback prediction by FEA, in: Proceedings of the IDDRG 2007, pp. 397–404.