samo 2022

10th International Conference on Sensitivity Analysis of Model Output

14th - 16th of March 2022

Tallahassee, Florida

FLORIDA STATE UNIVERSITY

Table of Contents

Keynotes

“Variable importance and explainable AI” – Art Owen

“Towards more general sensitivity estimates: Applications considering model structural uncertainties, grouping of parameters, and large-scale analyses” – Juliane Mai

“Global Sensitivity Analysis: a novel generation of mighty estimators based on rank statistics” – Agnès Lagnoux

“A kernel-based ANOVA decomposition: extending sensitivity indices and Shapley effects with kernels” – Sébastien Da Veiga

“Understanding the modelling process and model use” – Samuele Lo Piano

“(Non)linear dimension reduction of input parameter space using gradient information” – Clémentine Prieur

Keynote 1

Variable importance and explainable AI

Art Owen
Department of Statistics, Stanford University, Stanford, CA, U.S.A

Abstract: In order to explain what a black box algorithm does we can start by studying which variables are important for its decisions. Variable importance is studied by making hypothetical changes to predictor variables. Changing parameters one at a time can produce input combinations that are outliers or very unlikely. They can be physically impossible, or even logically impossible. It is problematic to base an explanation on outputs corresponding to impossible inputs. We introduced the cohort Shapley (CS) measure to avoid this problem, based on Shapley value from cooperative game theory. There are many tradeoffs in picking a variable importance measure, so CS is not the unique reasonable choice. One interesting property of CS is that it can detect `redlining', meaning the impact of a protected variable on an algorithm's output when that algorithm was trained without the protected variable.

This talk is based on recent joint work with Masayoshi Mase and Ben Seilert. The opinions expressed are my own, and not those of Stanford, the National Science Foundation, or Hitachi, Ltd.

Keynote 2

Towards more general sensitivity estimates: Applications considering model structural uncertainties, grouping of parameters, and large-scale analyses

Juliane Mai
Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON, Canada

Abstract: Sensitivities of model outputs are traditionally evaluated for the parameters specific to a given model of interest simulating a specific output, for example, streamflow. This presentation will focus on attempts leading to more general sensitivity estimates that hold for more than one specific model through (1) the inclusion of model structural uncertainties as parameters in the analysis, (2) grouping parameters such that sensitivities are not parameter specific but process specific, and (3) the deployment of these methods to large regions such that underlying patterns can be identified and transferred to locations that might have not been analysed before.

These approaches have been recently applied to hydrologic models across North America evaluating their sensitivity to simulated streamflow. This presentation will describe the underlying methods applied and present results derived from analysing a blended hydrologic model structure, which includes not only parametric, but also structural uncertainties over more than 3000 basins across North America. Furthermore, it will be described how the results of the 3000 basins were used to derive an approximation of sensitivities based on physiographic and climatologic data such that sensitivities can be estimated without the expensive analysis. The interactive website sharing detailed spatio-temporal inputs and results of this study will be shown.

Keynote 3

Global Sensitivity Analysis: a novel generation of mighty estimators based on rank statistics

Agnès Lagnoux
Institut de Mathématiques de Toulouse, Université Toulouse 2 Jean Jaurès, Toulouse, France

Abstract: In this talk, I present a new statistical estimation framework for a large family of global sensitivity analysis indices that we have proposed in a recent paper published in 2021. Our approach is based on rank statistics and uses an empirical correlation coefficient recently introduced by Chatterjee. We show how to apply this approach to compute not only the Cramér-von-Mises indices, directly related to Chatterjee's notion of correlation, but also first-order Sobol' indices, general metric space indices and higher-order moment indices. We establish the consistency of the resulting estimators and demonstrate their numerical efficiency, especially for small sample sizes. In addition, we prove a central limit theorem for the estimators of the first-order Sobol' indices.

Keynote 4

A kernel-based ANOVA decomposition: extending sensitivity indices and Shapley effects with kernels

Sébastien Da Veiga
Safran Tech, Paris, France

Abstract: Global sensitivity analysis is the main quantitative technique for identifying the most influential input variables in a numerical model.
In particular when the inputs are independent, Sobol' sensitivity indices attribute a portion of the output variance to each input and all possible interactions in the model, thanks to a functional ANOVA decomposition.
On the other hand, moment-independent sensitivity indices focus on the impact of inputs on the whole output distribution instead of the variance only, thus providing complementary insight on the inputs/output relationship. But they do not enjoy the nice decomposition property of Sobol' indices and are consequently harder to analyze.
In this talk, we introduce two moment-independent indices based on kernel-embeddings of probability distributions and show that the RKHS framework makes it possible to exhibit a kernel-based ANOVA decomposition.
This is the first time such a desirable property is proved for sensitivity indices apart from Sobol' ones. With dependent inputs, we also use these new sensitivity indices as building blocks to design kernel-embedding Shapley effects which generalize the traditional ones.
Several estimation procedures are discussed and illustrated on test cases with various output types such as categorical variables and probability distributions. All these examples show their potential for enhancing sensitivity analysis with a kernel viewpoint.

Keynote 5

Understanding the modelling process and model use

Samuele Lo Piano
University of Reading, Reading, UK

Abstract: Models are used to represent systems and their possible evolutions, gaining insights to be translated into decisions on the real system they aim to represent. The steps of model development and use are especially critical when these are used at the policy-making interface. It is not infrequent that the model replaces the system modelled as locus of attention and that the use of a given model may be extrapolated well beyond the function it has initially been conceived for. In these settings, uncertainty and sensitivity analysis, however useful to draw inference on model's robustness and stability, may be insufficient to acknowledge these kinds of issues leading to potentially regrettable decisions. In this contribution, I will discuss the approaches proposed for thorough scrutiny of the modelling activities and their use at the science-policy interface. I will conclude by examining practical examples in the context of recent initiatives where efforts have been put forth to mainstream these practices and approaches.

Keynote 6

(Non)linear dimension reduction of input parameter space using gradient information

Clémentine Prieur
Université Grenoble Alpes, Grenoble, France

Abstract: Many problems that arise in uncertainty quantification, e.g., integrating or approximating multivariate functions, suffer from the curse of dimensionality. The cost of computing a sufficiently accurate approximation grows indeed dramatically with the dimension of input parameter space. It thus seems important to identify and exploit some notion of low-dimensional structure as, e.g., the intrinsic dimension of the model. A function varying primarily along a a low dimensional manifold embedded in the high-dimensional input parameter space is said of low intrinsic dimension. In that setting, algorithms for quantifying uncertainty focusing on the most relevant features of input parameter space are expected to reduce the overall cost. Our presentation goes from global sensitivity analysis to (non)linear gradient-based dimension reduction, generalizing the active subspace methodology.

Speaker: Changcong Zhou

Global Sensitivity Analysis Based on Active Subspaces and Kriging

Changcong Zhou, Zhuangke Shi
Department of Engineering Mechanics, Northwestern Polytechnical University, Xi'an 710072, China, changcongzhou@nwpu.edu.cn
Sergei Kucherenko
Imperial College London, London, SWT 2AZ, UK,
s.kucherenko@imperial.ac.uk

Global sensitivity analysis (GSA) measures the effects of input variables on the model output by considering their whole ranges of uncertainty. Many GSA measures have been developed over the past decades, among which the most popular is the variance based method of Sobol' sensitivity indices [1]. Another popular among practitioners measure is the derivative-based global sensitivity measure (DGSM) [2]. Sobol' and Kucherenko [3] showed that there is a link between the DGSM and variance-based sensitivity indices in a form of inequality relationship.

Active subspaces have recently become a new valuable method among tools of GSA [4]. It is based on the observation that model output often depends only on a limited number of directions in the input space. Each direction corresponds to a linear combination of the original input variables, which is referred to as the active subspace. The active subspace can be identified by performing the eigenvalue decomposition of the covariance-like matrix of response gradients. The original input variables can be mapped into the active subspace and the input-output model can be set in the active subspaces. Once the model in the active subspace becomes a low-dimensional one, it is much easier to build an accurate surrogate model and to overcome the "curse of dimensionality". We propose a framework based on the application of active subspaces followed by Kriging [5] for building a surrogate model in the reduced space and subsequent computation of three types of GSA measures, namely DGSM, activity score and Sobol' total effect index. The proposed approach includes the following general steps: Step 1: Generate the full sample set; Step 2: Estimate the covariance-like matrix and DGSM; Step 3: Perform the eigenvalue decomposition, identify the active subspace, and compute the activity score; Step 4: Build the Kriging surrogate model by an adaptive procedure in the active subspace. Step 5: Estimate the total effect index with the Kriging model.

As an example we consider a high-dimensional quadratic model: $y=(\sum c_i x_i)^2$ (i=1,2,…,100). Here x_i follows the uniform distribution on [0,1], the values of coefficients c_i are given in Table 1. Obviously, the importance of input variables depends on the coefficients: a larger coefficient means higher importance.

Table 1 Values of the coefficients

c_5	c_{15}	c_{25}	c_{35}	c_{45}	c_{55}	c_{65}	c_{75}	c_{85}	c_{95}	The rest
5	15	25	35	45	55	65	75	85	95	1

The eigenvalues plotted on a logarithmic scale are shown in Fig. 1. The large gap between the first and the second eigenvalue implies that there exists one-dimensional active subspace. The Kriging model is built in the active subspace. The training sample points and the Kriging model in the one dimensional active subspace are shown in Fig. 2(a). For comparison, the Kriging model is also built in the two-dimensional active subspace as shown in Fig. 2(b). It can be noticed that the Kriging model built in one-dimensional active subspace is sufficient to capture the general behavior of output. The three GSA indices of all the input variables obtained by the proposed approach (in one-dimensional active subspace) as well by the straightforward Monte Carlo simulation (MCS) are presented in Fig. 3. The results obtained by the proposed approach agree well with those obtained by MCS, which proves that the proposed

approach is applicable in the high-dimensional case. Normalized values of DGSM and activity score are practically the same as the values of total effect indices. Computed GSA indices have successfully identified the importance of the input variables as follows: $x_{95}>x_{85}>x_{75}>x_{65}>x_{55}>x_{45}>x_{35}>x_{25}>x_{15}>x_5>$other inputs.

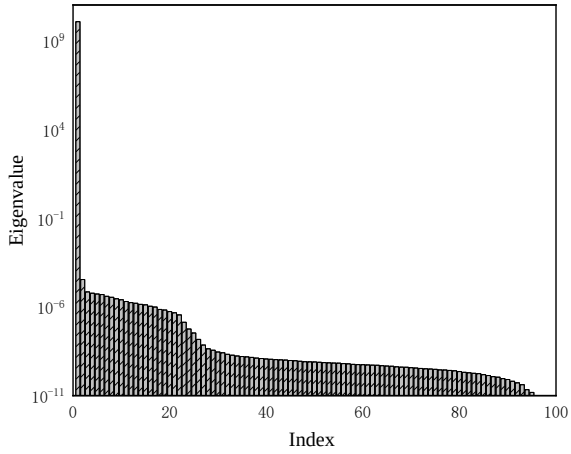
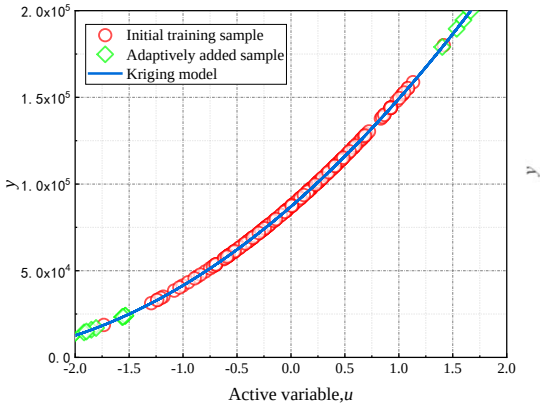
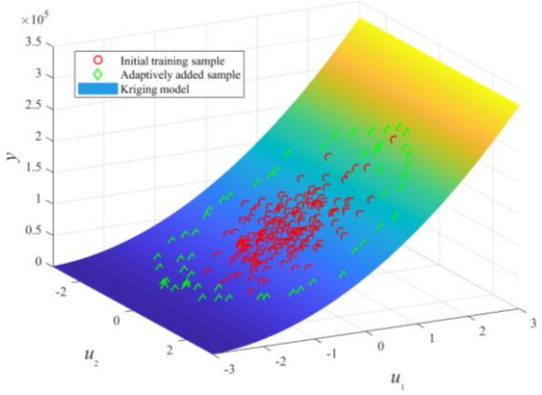


Fig. 1. The eigenvalues





(a) One-dimensional active subspace (b) Two-dimensional active subspace

Fig. 2. The training samples and Kriging model in the active subspace

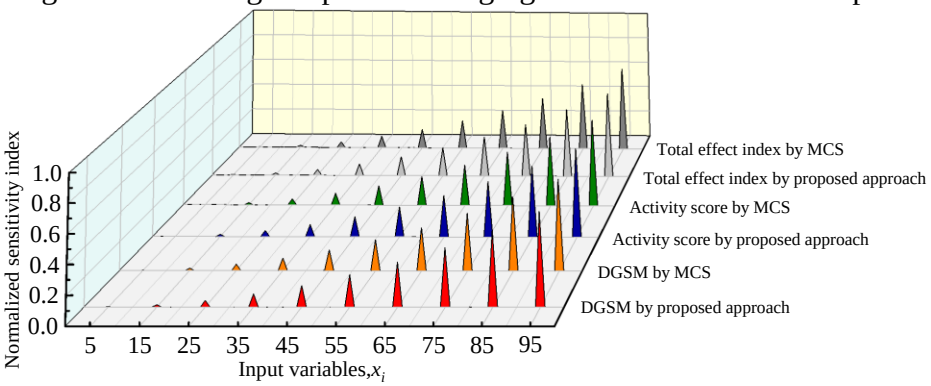


Fig. 3. Normalized GSA indices

(For a sensitivity index γ_i, the normalized value is expressed as $\bar{\gamma}_i = \gamma_i / (\sum \gamma_i^2)^{1/2}$

[1] I.M. Sobol', Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.*, 55: 271-280, 2001.

[2] S. Kucherenko, M. Rodriguez-Fernandez, C. Pantelides, N. Shah, Monte Carlo evaluation of derivative-based global sensitivity measures, *Reliab. Eng. Syst. Saf.*, 94: 1135–1148, 2009.

[3] I.M. Sobol', S. Kucherenko, Derivative based global sensitivity measures and their link with global sensitivity indices, *Math. Comput. Simul.*, 79: 3009–3017, 2009.

[4] P.G. Constantine, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, SIAM, Philadelphia, 2015.

[5] H. Liu, J. Cai, Y-S Ong, An adaptive sampling approach for Kriging metamodeling by maximizing expected prediction error, *Comput. Chem. Eng.*, 106: 171–182, 2017.

Speaker: Sergei Kucherenko

Comparison of Active Subspaces and Global Sensitivity Measures for Problems with Rotations and Dependent Variables

Sergei Kucherenko
Imperial College London, London, SWT 2AZ, UK,
s.kucherenko@imperial.ac.uk
Peter Yatsyshin
The Alan Turing Institute, UK, pyatsyshin@turing.ac.uk
Nilay Shah
Imperial College London, London, SWT 2AZ, UK,
N.Shah@imperial.ac.uk

Often the lower-dimensional subspaces that impact estimates of uncertainty are efficiently described by combinations of parameters. By moving beyond identifying directions aligned with the axes of the parameter space, significant dimension reduction can be achieved. Ideally, sensitivity analysis (SA) should identify directions that are most influential regardless of their orientation. *Active subspaces* (AS) is a method which identifies important directions in the parameter space and allows dimension reduction (Constantine **Error! Reference source not found.**).

In this work we compared the AS method with variance based and DGSM methods. Consider a nonlinear numerical model $f(\boldsymbol{x}) \in L_1$, $\boldsymbol{x} \in \mathbb{R}^n$ distributed according to pdf $p(\boldsymbol{x})$. Compute the symmetric positive semidefinite matrix $\boldsymbol{C} = \boldsymbol{E}[\nabla \boldsymbol{f} \nabla \boldsymbol{f}^T]$ and its eigenvalue decomposition: $\boldsymbol{C} = \mathbf{W}\boldsymbol{\Lambda}\mathbf{W}^{\mathbf{T}}$, where $\boldsymbol{\Lambda} = \mathbf{diag}(\boldsymbol{\lambda_1}, \ldots \boldsymbol{\lambda_n})$, $\boldsymbol{\lambda_1} \geq \cdots \geq \boldsymbol{\lambda_n}$ are eigenvalues**,** and $\mathbf{W}$ is the orthogonal matrix of the corresponding eigenvectors forming the basis of $\mathbb{R}^n$. The main idea of the AS method is to find a partition $\mathbf{W} = \mathbf{W_1} + \mathbf{W_2}$, where $\mathbf{W_1}$ is formed by the eigenvectors of the top $\boldsymbol{k}$ eigenvalues (where $\boldsymbol{k} \ll \boldsymbol{n}$), such that $\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{g}(\boldsymbol{y})$, where $\mathbf{y} = \mathbf{W_1^T x}$ and $\boldsymbol{y} \in \mathbb{R}^k$. The span of the top $\boldsymbol{k}$ eigenvectors of $\boldsymbol{C}$ is called the "active subspace. It can be shown that optimal (in a certain sense) $\boldsymbol{k}$ corresponds to the largest gap in the spectrum of $\boldsymbol{C}$ [1]. Notice that unlike the global sensitivity analysis techniques, where subsets of input parameters that can be neglected are identified, the active spaces approach seeks to find important linear combinations of all input parameters, which span the "active subspace". Perturbing the inputs along the important directions causes greater change in the prediction, on average, than perturbing along the unimportant ones. The active variables are the linear combinations of the input parameters with weights from the important eigenvectors. We note that DGSM are the diagonal elements of Matrix $\boldsymbol{C}$ (1). Variance based sensitivity indices measure the proportion of the variance attributed to each parameter while the eigenvector identifies an important direction in the parameter space, where "importance" is measured by the eigenvalue. In other words, perturbing x along $\mathbf{W_1}$ changes $f(\mathbf{x})$ the most, on average.

We consider the Ishigami function $f(x) = sin\, x_1 + a\, sin^2 x_2 + bx_3^4\, sin\, x_1$, $-\pi \leq x_i \leq \pi, i$=1, 2, 3, a=7, b=0.1. GSA gives the following values of $S_i^{tot} = \{0.55, 0.44, 0.24\}$ and $G_i = \{305,\ 967, 433\}$. Presented results for DGSM show that the second variable is the most important, with the third variable being second most important. We note, that analysis based on

Sobol' indices presents a different picture, namely ranking of inputs in the order of importance is 1, 2, 3. Applying AS we find the eigenvalues: Λ=[239.32; 121.49; 18.51]. We consider the first two eigenvalues to be dominant, hence k=2. $\mathbf{W_1^T}$ corresponding to these two eigenvalues $\mathbf{W_1^T} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$: AS lies along the second x_2 and the third x_3 direction which is in agreement with importance of inputs derived from using DGSM. In the second scenario we consider only the first eigenvalue to be dominant, k=1. It allows a 1D approximation of the original 3D function. Comparison of the fist and the second scenarios reveals that the second scenario in which dimension is reduced to 1 is sufficient to approximate the original function with good accuracy.

Consider a model in transformed coordinates $\boldsymbol{x'} = \widehat{\boldsymbol{R}}\boldsymbol{x}$, where the rotation matrix $\widehat{\boldsymbol{R}} = \begin{bmatrix} cos(\theta) & -sin(\theta) & 0 \\ sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$. This is transformation in two dimensions. It rotates the (x_1, x_2) axes clockwise into the (x_1', x_2') axesConsider the Ishigami function $\boldsymbol{f}(\widehat{\boldsymbol{R}}\boldsymbol{x}) = \boldsymbol{\varphi}(\boldsymbol{x})$ with rotation on π and assume that the user is agnostic with regards to transformation, that is function $\boldsymbol{\varphi}(\boldsymbol{x})$ is analysed in the original coordinates. From the values of $S_i = \{0.08, 0.06, 0.0\}$, $S_i^{tot} = \{0.73, 0.72, 0.22\}$ and $G_i = \{37,\ 38,\ 24\}$, no certain conclusions can be made as it seems that the first and the second inputs are equally important through interactions. Neither variance based nor DGSM based global sensitivity analysis is able to identify critical directions.

We note that $\boldsymbol{\nabla_x f}(\widehat{\boldsymbol{R}}\boldsymbol{x}) = \widehat{\boldsymbol{R}}\boldsymbol{\nabla_{x'} f}(\boldsymbol{x'})$, then $\boldsymbol{C_R} = \int \boldsymbol{\nabla_x f \nabla_x f^T}\, \boldsymbol{dx} = \widehat{\boldsymbol{R}}\boldsymbol{C}\widehat{\boldsymbol{R}}^{\boldsymbol{T}}$. Here $\boldsymbol{C} = \mathbf{W\Lambda W^T}$ is a matrix of the model without rotation. Applying the AS methodology we get $\boldsymbol{y} = \mathbf{W_1^T}\widehat{\boldsymbol{R}}^{\boldsymbol{T}}\boldsymbol{x} = (\widehat{\boldsymbol{R}}\mathbf{W_1})^{\boldsymbol{T}}\boldsymbol{x}$. Hence all the results for AS are exactly the same as in the case with no rotation, however the active directions are now given by the matrix $\widehat{\boldsymbol{R}}\mathbf{W_1}$ and a low-dimensional approximation of $\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{g}((\widehat{\boldsymbol{R}}\mathbf{W_1})^{\boldsymbol{T}}\boldsymbol{x})$. Scatter diagrams of the output versus active and non active directions (not presented here) show that the AS method was capable of finding AS in which the scatter diagrams look exactly the same as in the case of no rotation. We can conclude that traditional global sensitivity analysis (GSA) methods such as the variance based method of Sobol' sensitivity indices [2] or DGSM [3] are unable to identify orthogonal directions and rank parameters in active subspaces rather than in original directions.

We also generalised the AS methodology for the case of models with dependent variables and compared the generalised AS methodology with that of generalised Sobol' indices [4].

[1] P.G. Constantine, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, SIAM, Philadelphia, 2015.

[2] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, *Sensitivity Analysis in Practice*, Wiley (eds.), 2004.

[3] I.M. Sobol', S. Kucherenko, Derivative based Global Sensitivity Measures and their link with global sensitivity indices, *Mathematics and Computers in Simulation*, 79 (10), 3009-3017, 2009.

[4] S. Kucherenko, S. Tarantola, P. Annoni. Estimation of global sensitivity indices for models with dependent variables, *Computer Physics Communications*, 183, 937–946, 2012.

Speaker: John Barr

Kernel Methods for Global Sensitivity Analysis

John Barr[1] and *Herschel Rabitz*[2]

[1]*Department of Chemistry, Princeton University, United States of America,* jbarr@princeton.edu

[2]*Department of Chemistry, Princeton University, United States of America*, hrabitz@princeton.edu

Global Sensitivity Analysis (GSA) aims at quantifying the influence of the inputs on an output defined by an input-output system due to changes of the input variables over their entire domain. The evolution of global sensitivity analysis (GSA) in recent years has focused on developing new tools that broaden its scope of applications. These developments have included introducing new sensitivity measures for certain types of outputs, lifting restrictions on assumptions of previous approaches (such as input independence), and developing goal-oriented methods that define classes of sensitivity measures capable of analyzing different aspects of how an input can influence the output distribution [1, 2].

Kernel-based procedures have enjoyed considerable success in areas such as machine learning and pattern analysis, and naturally can be extended to GSA. The practicality of these techniques extends to solving difficult non-parametric problems by embedding data points into higher dimensional reproducing kernel Hilbert spaces (RKHSs), allowing for non-linear algorithms based upon traditional linear approaches. These concepts have been generalized to embed probability distributions into RKHSs, producing powerful methods for dealing with higher order statistics [3]. Recently, the embedding of probability distributions into RKHS has been applied to define unique GSA procedures [1,4].

In this work we propose a new class of GSA measures based on the embedding of the joint probability distribution of a system's output into a RKHS, which has several key advantages over many traditional sensitivity techniques. Specifically, the proposed methodology introduces a unified class of sensitivity measures that are well defined for an arbitrary type of output, computationally feasible for high-dimensional outputs, do not require input variable independence, and have the capacity to be goal-oriented such that the practitioner can change the sensitivity indicator to select the features of the output distribution that are deemed important.

This new kernel-based technique encapsulates measures that are both moment-independent and moment-based. This advance includes the introduction of novel GSA metrics as well as showing that certain previously proposed GSA measures, such as the variance-based indicators, are special cases of the kernel-based procedure. This new GSA procedure is tested against multiple benchmark applications to highlight the broad-range systems it can be applied to such as systems with high-dimensional or categorical outputs.

Building on this research, future studies will focus on two directions. The first concerns the kernel-selection process and how an optimal kernel-metric can be chosen. The second will focus on new algorithms to increase numerical efficiency so this methodology can be applied in regimes where data sampling can be expensive.

References

[1] Barr, J. and Rabitz, H. A Generalized Kernel Method for Global Sensitivity Analysis. *SIAM/ASA Journal on Uncertainty Quantification*, (Forthcoming, Accepted).

[2] Borgonovo, E. and Plischke, E. Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, *248* (3), pp. 869-887, 2016.

[3] Smola, A., Gretton, A., Song, L. and Schölkopf, B. A Hilbert space embedding for distributions. *International Conference on Algorithmic Learning Theory*, pp. 13-31. 2015.

[4] Da Veiga, S. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, *85*(7), pp.1283-1305. 2015.

Speaker: Franziska Henze

Dynamic Sampling Strategy for Morris' Method of Elementary Effects

Franziska Henze
Institute of Measurement and Control Systems, Karlsruhe Institute of Technology, Germany, franziska.henze@kit.edu

Markus Rußer
AUDI AG, Germany and University of Applied Sciences Kempten, Germany, markus.russer@audi.de and markus.russer@stud.hs-kempten.de

Dennis Faßbender
AUDI AG, Germany, dennis.fassbender@audi.de

Christoph Stiller
Institute of Measurement and Control Systems, Karlsruhe Institute of Technology, Germany, stiller@kit.edu

Stefan-Alexander Schneider
Faculty of Electrical Engineering, University of Applied Sciences Kempten, Germany, stefan-alexander.schneider@hs-kempten.de

Highly automated driving functions are developed to navigate a driverless vehicle through traffic. Unfortunately, understanding planning decisions and errors becomes more and more difficult even for function developers themselves [7]. Therefore, developing frameworks for transparent explanations will be crucial for the future success of this new technology [5, 10, 11]. For this, we use sensitivity analysis (SA) methods to identify the parameters that are most influential for the decision [6] and that we expect to build the basis for future explanations. Since updating the plan usually happens at a high frequency ($\sim 10\,\mathrm{Hz}$ [1, 8]), the SA is required to be computationally effortless. Morris' method of elementary effects is a one-factor-at-a-time approach [9], which is expected to require only a few sample evaluations and therefore has a manageable computational effort while providing sufficient information for our cause [4]. However, previous work showed that for analyzing planning algorithms with many inputs, a sample size of $M \in \{4, \ldots, 10\}$ may not be sufficient to identify all relevant information [3, 6]. Increasing the sample size $M \geq 100$ is one option, but it comes at the cost of evaluating unnecessarily many samples in simple situations. Therefore, we propose a dynamic stop criterion (DSC) to adapt the number of samples to the complexity of the system. We first shortly recapitulate Morris' method and define the DSC. Afterwards, we apply it to an example, before we discuss possibilities and limitations.

Morris' method of elementary effects is based on characterizing the distribution of gradient approximations of a function at random points [9]: For a function $\boldsymbol{f} : \Omega \subseteq \mathbb{R}^k \to \mathbb{R}^m$

$(k, m \in \mathbb{N})$, the directional derivative along the unit vector $\boldsymbol{e}_i \in \mathbb{R}^k$ is approximated for a set of random samples $\{\tilde{\boldsymbol{x}}_j\}_{j=1}^{M_i}$, $\Delta_i \in \mathbb{R}$ by

$$\boldsymbol{d}_i(\tilde{\boldsymbol{x}}_j) := \frac{\boldsymbol{f}(\tilde{\boldsymbol{x}}_j + \Delta_i \boldsymbol{e}_i) - \boldsymbol{f}(\tilde{\boldsymbol{x}}_j)}{\Delta_i} \quad (i \in \{1, \ldots, k\}).$$

Campolongo et al. [2] suggest using the sample mean of $\{|\boldsymbol{d}_i(\tilde{\boldsymbol{x}}_j)|\}_{j=1}^{M_i}$ defined by

$$\boldsymbol{\mu}^*_{i,M_i} := \sum_{j=1}^{M_i} \frac{|\boldsymbol{d}_i(\tilde{\boldsymbol{x}}_j)|}{M_i}, \quad \boldsymbol{\mu}^*_{\boldsymbol{M}} = \left(\boldsymbol{\mu}^*_{1,M_1}, \ldots, \boldsymbol{\mu}^*_{k,M_k}\right) \in \mathbb{R}^{m \times k}, \quad \boldsymbol{M} = (M_1, \ldots, M_k)^T \in \mathbb{N}^k,$$

with componentwise absolute value $|\cdot|$, to classify an input i as relevant ($\left\|\boldsymbol{\mu}^*_{i,M_i}\right\| \gg 0$) or irrelevant ($\left\|\boldsymbol{\mu}^*_{i,M_i}\right\| \approx 0$). Touzani et al. [12] introduce a global error criterion to determine if the sensitivity index of a quantitative derivative-based global sensitivity measure (DGSM) has converged. We modify it for the qualitative SA measure $\boldsymbol{\mu}^*_{i,M_i}$ to iteratively adapt the number of samples to the classification difficulty, but keep the main idea of averaging the relative deviations over the past 10 iterations [12]. To construct an iterative process over sample size M, we define a component-wise relative error by

$$\epsilon^2_{i,M_i} := \frac{1}{10} \sum_{l=1}^{10} \frac{\left\|\boldsymbol{\mu}^*_{i,M_i-l} - \boldsymbol{\mu}^*_{i,M_i}\right\|^2}{\left\|\boldsymbol{\mu}^*_{i,M_i}\right\|^2}$$

to evaluate the convergence of the SA measure for input i, and a global error measure

$$\epsilon^2_{\boldsymbol{M}} := \frac{1}{k} \left(\sum_{i \in \mathcal{A}} \epsilon^2_{i,M} + \sum_{i \in \{1,\ldots,k\} \setminus \mathcal{A}} \epsilon^2_{i,M_i} \right), \quad M_i = M \text{ for } i \in \mathcal{A} \subseteq \{1, \ldots, k\}, \text{ else } M_i < M$$

to check for convergence after M iterations and increase M by 1 otherwise. To separate the active inputs $\mathcal{A} \subseteq \{1, \ldots, k\}$ from the inputs $\{1, \ldots, k\} \setminus \mathcal{A}$ that we can classify after evaluating only a few samples, we define the iterative DSC for constants $\kappa_{stop} > \kappa_{act} > 0$ as

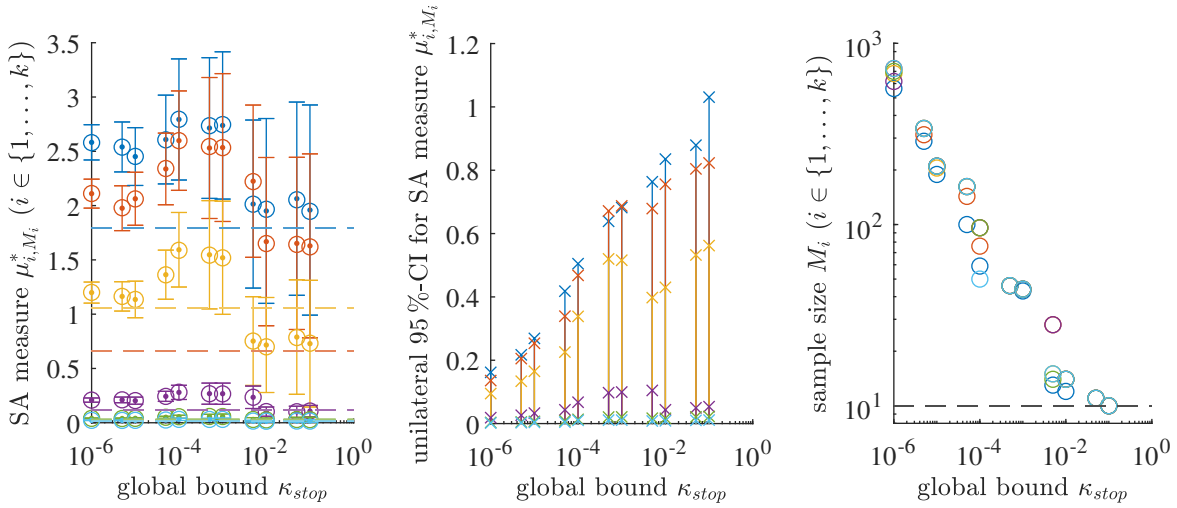
$$\epsilon^2_{i,M} \leq \kappa_{act} \quad \Rightarrow \quad \mathcal{A} := \mathcal{A} \setminus \{i\}, M_i := M, \tag{1}$$

$$\epsilon^2_{\boldsymbol{M}} \leq \kappa_{stop} \quad \Rightarrow \quad \text{stop calculation after } M \text{ iterations.} \tag{2}$$

To demonstrate the proposed criterion, we use the g-function

$$g : [0,1]^k \to \mathbb{R}, \quad g(\boldsymbol{x}) := \prod_{i=1}^{k} g_i(x_i) = \prod_{i=1}^{k} \frac{|4x_i - 2| + a_i}{1 + a_i}$$

with the same $k = 6$ parameters $[a_1, \ldots, a_6]^T = [0, 0.2, 0.9, 9, 50, 99]^T$ as in [3] ($m = 1$), where small values a_i indicate large influence. We now compare the results obtained with the DSC ($\Delta_i = 0.25$ for all $i \in \{1, \ldots, 6\}$, $M_{min} = 11$, $\kappa_{act} = 0.5\kappa_{stop}$) to the ones of a simulation similar to [3] (i.e., $M_i = 10$, $\Delta_i = 0.25$ for all $i \in \{1, \ldots, 6\}$)



(a) ⊙ SA measure μ^*_{i,M_i} after M_i samples, ⌶ Error bars give 95% CIs.

(b) ᛉ Positive half of 95% CI for SA measure μ^*_{i,M_i} (centered around 0).

(c) ○ Number of samples M_i used to calculate SA measure μ^*_{i,M_i}.

Figure 1: Analysis of global DSC bound κ_{stop} for g-function ($m = 1$, [3]). Confidence intervals (CIs) are calculated from bootstrap resampling with 10^3 bootstrap replicates as in [3]. Dashed lines - - - give results calculated according to [3] (without CIs). Colors indicate inputs: ■ $a_1 = 0$, ■ $a_2 = 0.2$, ■ $a_3 = 0.9$, ■ $a_4 = 9$, ■ $a_5 = 50$, ■ $a_6 = 99$.

and discuss parameter κ_{stop}. Fig. 1a shows the SA measures μ^*_{i,M_i} for all inputs $i \in \{1,\dots,6\}$ depending on the constant κ_{stop}: The ones calculated according to [3] (dashed lines) deviate strongly from the ones calculated with DSC (e.g., a_2) and the confidence intervals decrease with decreasing κ_{stop}, indicating a better precision for the calculations with DSC (cf. Figs. 1a, 1b). Further investigations show that this holds also for the sample mean and sample variance proposed as SA measures by Morris [9]. As expected, this comes at the expense of additional sample evaluations, cf. Fig. 1c. However, for $\kappa_{stop} \leq 10^{-2}$ the component-wise criterion (1) stops evaluating samples for some inputs earlier than for others (e.g., $i \in \{1,3,5\}$, $\mathcal{A} = \{2,4,6\}$ for $\kappa_{stop} = 5 \cdot 10^{-3}$). Compared to the straightforward approach presented in [3], which gives already reasonably good classification results, the proposed improvements for the sampling strategy seem expensive without significant improvements. However, as stated therein, choosing a sample size of $M \in \{4,\dots,10\}$ gives only bad estimates of SA measures such as $\boldsymbol{\mu}^*$, which is problematic whenever an automatic classification based on a threshold $c > 0$ is used: For $\|\boldsymbol{\mu}^*\| \approx c$, the relevance classification will naturally be unsafe for small sample sizes. With the DSC, one can improve the estimates specifically for the inputs with large fluctuations, thus gaining more reliable qualitative information, e.g. on the predominant sign of the gradient approximations (positive or negative sample mean, cf. [9]), while maintaining the overall number of samples at a low level. This is especially interesting for applications with many inputs k, where some are easy to classify without evaluating too many samples, while others need more investigation, e.g., because they are only weakly influential.

References

[1] Zlatan Ajanovic, Bakir Lacevic, Barys Shyrokau, Michael Stolz, and Martin Horn. Search-based optimal motion planning for automated driving. *IEEE/RJS International Conference on Intelligent Robots and Systems (IROS)*, 2018:4523–4530, 2018.

[2] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10):1509–1518, 2007.

[3] Francesca Campolongo and Andrea Saltelli. Sensitivity analysis of an environmental model: an application of different analysis methods. *Reliability Engineering & System Safety*, 57(1):49–69, 1997.

[4] Francesca Campolongo, Stefano Tarantola, and Andrea Saltelli. Tackling quantitatively large dimensionality problems. *Computer Physics Communications*, 117(1-2):75–85, 1999.

[5] Matthew Gadd, Daniele de Martini, Letizia Marchegiani, Paul Newman, and Lars Kunze. Sense-Assess-eXplain (SAX): Building trust in autonomous vehicles in challenging real-world driving scenarios. *IEEE Intelligent Vehicles Symposium (IV)*, pages 150–155, 2020.

[6] Franziska Henze, Dennis Faßbender, and Christoph Stiller. Sensitivity analysis of a planning algorithm considering uncertainties. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 1128–1134, 2020.

[7] Franziska Henze, Natalie Magdalena Stasinski, Dennis Faßbender, and Christoph Stiller. Developers' information needs during test drives with automated vehicles in real traffic: A focus group study. *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '21 Adjunct), September 9–14, 2021, Leeds, United Kingdom*, 2021 (accepted).

[8] Christos Katrakazas, Mohammed Quddus, Wen-Hua Chen, and Lipika Deka. Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions. *Transportation Research Part C: Emerging Technologies*, 60:416–442, 2015.

[9] Max D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.

[10] Daniel Omeiza, Helen Webb, Marina Jirotka, and Lars Kunze. Towards accountability: Providing intelligible explanations in autonomous driving. *IEEE Intelligent Vehicles Symposium (IV)*, pages 231–237, 2021.

[11] Lukas Michael Schmidt, Georgios Kontes, Axel Plinge, and Christopher Mutschler. Can you trust your autonomous car? Interpretable and verifiably safe reinforcement learning. *IEEE Intelligent Vehicles Symposium (IV)*, pages 171–178, 2021.

[12] Samir Touzani and Daniel Busby. Screening method using the derivative-based global sensitivity indices with application to reservoir simulator. *Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles*, 69(4):619–632, 2014.

Speaker: Elmar Plischke

Sensitivity Analysis with Shapley Effects: Computational Issues

Elmar Plischke
Institute of Disposal Research, Clausthal University of Technology, Germany,
epl@tu-clausthal.de
Emanuele Borgonovo
Department of Decision Science, Bocconi University, Italy,
emanuele.brogonovo@unibocconi.it

Shapley effects are attracting attention as sensitivity measures. With the value function being the conditional variance, these effects account for the individual and higher order sensitivity effects. However, one of the issues associated with their use is computational cost. We present new algorithms [1] that offer major improvements for the computation of Shapley effects, reducing computational burden by several orders of magnitude with respect to currently available implementations [3]. With these new algorithms, one may estimate also generalized (Shapley-Owen) effects for interactions.

The Shapley value is a concept from cooperative game theory. One considers a game with d players. The Shapley value is then the quantity that indicates the worth of forming coalitions and the expected payoff for each player, attributing a fair share of the grand total to each of the players. Generally, one defines the coalition worth function $\mathrm{val} : 2^{\underline{d}} \to \mathbb{R}$ with $\mathrm{val}(\emptyset) = 0$, attributing a sum of payoffs to a group of players. Here $2^{\underline{d}}$ is the powerset (set of subsets) of $\underline{d} = \{1, 2, \ldots, d\}$.

In the context of sensitivity analysis [2], one regards as players the inputs to a simulation model g: $x \mapsto y$, $\mathbb{R}^d \to \mathbb{R}$ and the variance of the conditional model output they explain as the value function, $\mathrm{val}(\alpha) = \mathbb{V}[\mathbb{E}[Y|X_\alpha]]$. Here X_α is the random vector obtained from the random input X, selecting only the coordinates indexed by α. One main reason for the interest towards Shapley effects is that they remain interpretable also in the presence of dependent inputs or domain irregularities such as holes.

The Shapley value is uniquely characterized by four axioms, Pareto-efficiency, symmetry, linearity, and null-player property. We take advantage from the following alternative definitions of the Shapley value,

$$\Phi_i(\mathrm{val}) = \frac{1}{d} \sum_{\alpha: i \notin \alpha} \binom{d-1}{|\alpha|}^{-1} \mathrm{mar}(\alpha, i), \qquad \mathrm{mar}(\alpha, i) = \mathrm{val}(\alpha \cup \{i\}) - \mathrm{val}(\alpha)$$

$$\Phi_i(\mathrm{val}) = \frac{1}{d} \sum_{\alpha: i \in \alpha} \binom{d-1}{|\alpha|-1}^{-1} \left(\mathrm{val}(\alpha) - \mathrm{val}(\sim \alpha)\right), \qquad \sim \alpha = \underline{d} \setminus \alpha,$$

$$\Phi_i(\mathrm{val}) = \sum_{\alpha: i \in \alpha} \frac{\mathrm{mob}(\alpha)}{|\alpha|}, \qquad \mathrm{mob}(\alpha) = \sum_{\beta \subset \alpha} (-1)^{|\alpha|+|\beta|} \mathrm{val}(\beta).$$

While the first one uses marginal contributions, the second one uses a balanced value function, including both the value of a coalition and the value of its anti-coalition. The third one makes use of the Möbius inverse. All formulas visit the $2^d - 1$ non-empty index combinations.

A first algorithmic refinement to compute variance-based Shapley effects via the first definition is obtained by

1. using a pick-and-freeze design instead of a brute-force double loop;
2. using a duality result for obtaining two estimators at the same computational costs;
3. estimating conditional variances via Sobol'/Saltelli and Jansen formulas;
4. Using a quasi Monte-Carlo design for improved convergence compared to a crude Monte-Carlo design.

This approach already achieves a notable speedup. However, there is room for improvements by exploiting the alternative definitions. From the second definition, the value function of any subset α is contributing to Shapley effect i either by a positive or a negative weight, depending on whether $i \in \alpha$ or not. Moreover, under independence, the computation of both $\mathrm{val}(\alpha)$ and $\mathrm{val}(\sim \alpha)$ may use the same pick-and-freeze design, halving the number of evaluations.

As the Möbius inverses used in in the third definition are formally equivalent to the functional ANOVA decomposition terms, this definition offers a viable alternative for computing the Shapley effects. However, this requires a quick way of solving for the Möbius inverse. Our results show that this can be performed efficiently.

We present deterministic algorithms for efficiently computing Shapley effects, computing all possible coalition-worth value functions. Stochastic algorithms (with a polynomial runtime) use a random subset of m permutations instead of the $d!$ ones. Some of the ideas presented here may also accelerate these stochastic algorithms.

[1] E. Plischke, G. Rabitti, E. Borgonovo. Computing Shapley Effects for Sensitivity Analysis, *SIAM/ASA Journal on Uncertainty Quantification*, in print, 2021.

[2] B. Owen, C. Prieur. On Shapley value for measuring importance of dependent inputs, *SIAM/ASA Journal on Uncertainty Quantification*, 5:986–1002, 2017.

[3] E. Song, B. L. Nelson, J. Staum Shapley effects for global sensitivity analysis: Theory and computation, *SIAM/ASA Journal on Uncertainty Quantification*, 4:1060–1083, 2016.

Speaker: Matieyendou Lamboni

Kernel-based sensitivity indices for any model behavior

Matieyendou Lamboni
[a] Department DFST, University of Guyane, Frecnh Guina, France.
[b] 228-UMR Espace-Dev, France,
matieyendou.lamboni[at]gmail.com/[at]univ-guyane.fr

We are interested in identifying the input variables that drive model output(s) in a domain of interest and/or govern specific model behaviors defined via weight functions such as outputs belonging to a given cluster from any classification approach (e.g., [3; 2; 7]). We then investigate uncertainty exploration of specific model behaviors using the multivariate weighted distribution theory, which is well-suited for altering initial distribution of inputs or output(s) (e.g., [6]). Formally, if we use ρ for the initial PDF of d variables $\mathbf{X}$, $w : \mathbb{R}^d \to \mathbb{R}_+$ for a weight function, the weighted PDF is given by

$$\rho^w(\mathbf{x}) := \frac{w(\mathbf{x})}{\mathbb{E}\left[w(\mathbf{X})\right]}\rho(\mathbf{x})\,. \tag{1}$$

Since the weighted variables $\mathbf{X}^w$ (i.e., variables having $\rho^w(\mathbf{x})$ as PDF) may be dependent, an appropriate measure of association between inputs and the target outputs related to specific model behaviors is needed. The recent works in [5; 4] make use of dependency models to provide dependent generalized sensitivity indices (dGSIs) that account for the dependency structure of inputs. For the weighted variables, a dependency model of $\mathbf{X}^w$ is given by

$$r_j : \mathbb{R}^d \to R^{d-1}, \qquad \mathbf{X}^w_{\sim j} = r_j\left(X^w_j, \mathbf{U}\right)\,, \tag{2}$$

where r_j is a known function, $\mathbf{U} \sim \mathcal{U}\left(0,\,1\right)^{d-1}$ and X^w_j is independent of $\mathbf{U}$.

Recall that Sobol' indices and dGSIs are based on the variance of sensitivity functionals (SFs), which contain the information about the contribution of inputs. Thus, Sobol' indices and dGSIs are sufficient to assess the effects of inputs when SFs are normally distributed. When a statistical test revels that SFs do not follow a Gaussian distribution, Sobol' indices and dGSIs are a second-order moment approximation of an adequate measure of association. Namely, the first-order and total dependent SFs of X^w_j are given by

$$f^{fo}_j\left(X^w_j\right) := \mathbb{E}_{\mathbf{U}}\left[f\left(X^w_j,\, r_j\left(X^w_j, \mathbf{U}\right)\right)\right] - \mathbb{E}_{X^w_j,\mathbf{U}}\left[f\left(X^w_j,\, r_j\left(X^w_j, \mathbf{U}\right)\right)\right]\,; \tag{3}$$

$$f^{tot}_j\left(\mathbf{X}^w\right) := f\left(X^w_j,\, r_j\left(X^w_j, \mathbf{U}\right)\right) - \mathbb{E}_{X^w_j}\left[f\left(X^w_j,\, r_j\left(X^w_j, \mathbf{U}\right)\right)\right]\,. \tag{4}$$

In this abstract, we propose kernel-based sensitivity indices as an adequate measure of association using the characteristic kernel corresponding to the distribution of SFs. To that end, we use $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ for a kernel defined on $\mathcal{X} := supp(f(\mathbf{X}^w))$ and corresponding to the feature map $\phi : \mathcal{X} \to \mathcal{F}$, that is, $K(\mathbf{y}, \mathbf{y}') = \langle \phi(\mathbf{y}), \phi(\mathbf{y}') \rangle_{\mathcal{F}}$ ([1]).

Definition 1 *For a symmetric, positive definite kernel with* $\mathbb{E}\left[K\left(f(\mathbf{X}^w), f(\mathbf{X}^{w'})\right)\right] \leq \infty$, *the first-order and total kernel-based SIs of* X_j^w *are defined by*

$$kS_j := \frac{\mathbb{E}\left[K\left(f_j^{fo}(X_j^w), f_j^{fo}(X_j^{w'})\right)\right]}{\mathbb{E}\left[K\left(f(\mathbf{X}^w), f(\mathbf{X}^{w'})\right)\right]}, \tag{5}$$

$$kS_{T_j} := \frac{\mathbb{E}\left[K\left(f_j^{tot}(\mathbf{X}^w), f_j^{tot}(\mathbf{X}^{w'})\right)\right]}{\mathbb{E}\left[K\left(f(\mathbf{X}^w), f(\mathbf{X}^{w'})\right)\right]}. \tag{6}$$

Proposition 1 *Let* $\mathbf{X}^{w'}$ *be an independent copy of* $\mathbf{X}^w$ *and* $K(\mathbf{y}, \mathbf{y}') = \left(\mathbf{y}^T \mathbf{y}'\right)^2$.
(i) For independent inputs and a real-valued function f, *Sobol' indices are given by*

$$S_j = \sqrt{kS_j}, \qquad S_{T_j} = \sqrt{kS_{T_j}}.$$

(ii) For any input distribution and a vector-valued function f, *the second-type dGSIs are*

$$dGSI_j^{2,M} = \sqrt{kS_j}, \qquad dGSI_{T_j}^{2,M} = \sqrt{kS_{T_j}}.$$

In conclusion, a quadratic kernel-based SIs are equivalent to i) Sobol' indices for ranking independent inputs and ii) dGSIs for ranking inputs. For non-Gaussian SFs the choice of the kernel or the feature map should be based on the distribution structure of SFs.

[1] A. Berlinet, C. Thomas, and T. A. Gnan. *Reproducing Kernel Hilbert Space in probability and statistics*. Kluwer Academic, 2004.

[2] J. C Bezdek, R. Ehrlich, and W. Full. Fcm: The fuzzy c-means clustering algorithm. *Computers and Geosciences*, 10(2-3):191–203, 1984.

[3] L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

[4] M. Lamboni. On dependent generalized sensitivity indices and asymptotic distributions. *arXiv preprint arXiv2104.12938*, 2021.

[5] M. Lamboni and S. Kucherenko. Multivariate sensitivity analysis and derivative-based global sensitivity measures with dependent variables. *RESS*, 212:107519, 2021.

[6] J. Navarro, J. M. Ruiz, Y. D. Aguila. Multivariate weighted distributions: a review and some extensions. *Statistics*, 40 (1), 51–64, 2006.

[7] S. Roux, S. Buis, F. Lafolie, and M. Lamboni. Cluster-based GSA: Global sensitivity analysis of models with temporal or spatial outputs using clustering. *EMS*,140:105046, 2021.

Speaker: Sébastien Roux

Sensitive partitioning of the model output space: principle and first results

Sébastien Roux
MISTEA, Univ Montpellier, INRAE, Montpellier SupAgro, France, sebastien.roux@inrae.fr
Patrice Loisel
MISTEA, Univ Montpellier, INRAE, Montpellier SupAgro, France, patrice.loisel@inrae.fr
Samuel Buis
EMMAH, INRAE, Univ Avignon et Pays de Vaucluse, Avignon, France, samuel.buis@inrae.fr

Context: partitioning the output space

The use of partition of the model output space is a very convenient way to define behaviors of a model which scales to any dimension of the output space and can be powerful to give interpretable characterization of model properties. Considering partitions of the output space in relation to parameter sensitivity comes from the Regional Sensitivity Analysis (RSA) approach of Spear and Hornberger [1]. This subject has lately gained interest through two research directions: i) its application in the context of reliability engineering to characterize parameter sensitivity in relation to a critical domain of the output space and using sensitivity measures compatible with rare events (Target SA, [2]), ii) its application in combination with a clustering procedure in order to characterize parameter sensitivity in relation to the dominant behaviors of the output space (Cluster-based GSA, [3]).

Optimized sensitive partitioning: principle

We introduce here a new perspective on these different approaches. Instead of trying to a priori characterize a target region of the output space, we propose to **optimize** the partitioning in order to reveal the partition of the output space **the most sensitive** to the variations of a given input, i.e. the most explained by the variations of this input. We named this approach **an optimized sensitive partitioning**. It results in associating to each input factor an optimized partition of the output space and a normalized score characterizing the influence of the parameter in driving the output from one region of the partition to the complementary one (see Fig. 1). The optimized sensitive partitioning approach thus aims to find for any model input the two most contrasted model behaviors (defined as regions of the output space) that are influenced by this input.
Two main ingredients are required to define an optimized sensitive partitioning: first a sensitivity measure relating the sensitivity of an input to a partition of the output space, second an optimization procedure that looks for a partition that maximizes the sensitivity score.

Sensitivity measures and optimization algorithms

We propose to use the cluster-based indices defined in [3]. They are defined as Sobol' indices (first order, or total) of the membership functions (MF) associated to a clustering

of the output space. MFs are typically defined in $[0, 1]$ and characterize the degree of membership of any point of the output space to a given cluster. We introduce two optimization algorithms dedicated to the case of a 2-partition problem. In this case, we look for the binarization $(C^*, \bar{C}^*)$ of the output space that maximizes a given cluster-based index.

- The first algorithm is based on a exhaustive search on a set of patches obtained using a first clustering of the output space. It can be applied with any sensitivity measure.
- The second one targets specifically first order indices and uses a property specific to the optimization with this criterion: the fact that two patches having very correlated histograms associated to the distribution of $X_i|Y \in C^*$ belong to the same region of the optimal partition $(C^*, \bar{C}^*)$. Histogram correlation can thus trigger efficiently the clustering of patches and allows to reach finer resolutions in the results.

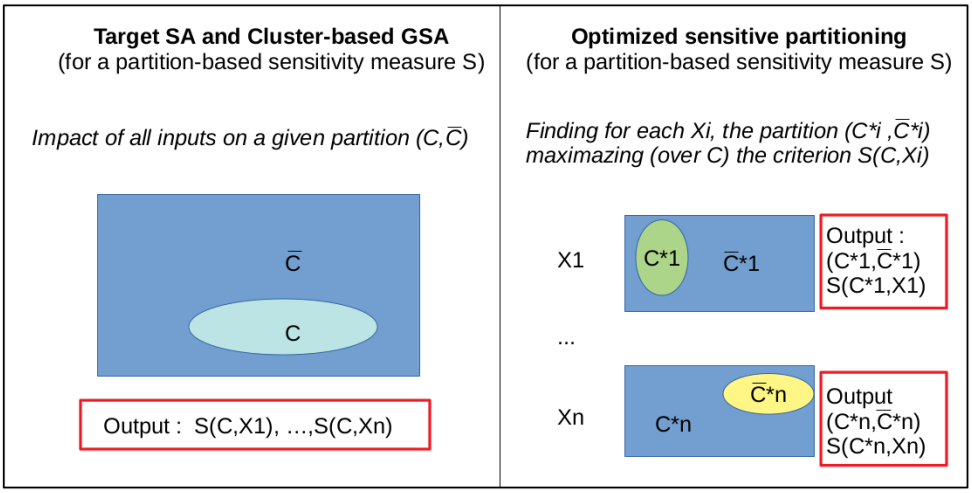


Figure 1: Principle of Optimized Sensitive Partitioning as opposed to Target SA and cluster-based GSA (here in the case of a partitioning into two regions).

First Results

We present the application of the method on different examples.

- First, a 1d example $f(X_1, X_2) = (sign(X_1) \cdot |X_2|)$, with X_1 and X_2 having uniform distributions in $[-1, 1]$ is considered. This example is used for validation purpose as the optimization problem in this case can be solved analytically.
- The two numerical algorithms are then illustrated on a 2d toy model allowing a 2d representation of the optimized partition and on a model with dynamic outputs to show the interest of the approach.

[1] Spear, R. C. and Hornberger, G. M. *Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis.* Water research, 14(1), 43-49,1980.

[2] Marrel, A. and Chabridon, V. *Statistical developments for target and conditional sensitivity analysis: application on safety studies for nuclear reactor.* Reliability Engineering & System Safety, 107711, 2021.

[3] Roux, S., Buis, S., Lafolie, F. and Lamboni, M. *Cluster-based GSA: Global sensitivity analysis of models with temporal or spatial outputs using clustering.* Environmental Modelling & Software, vol. 140, p. 105046, 2021.

Speaker: Gildas Mazo

Inference for sensitivity indices

Gildas Mazo
INRAE and Université Paris-Saclay, France, gildas.mazo@inrae.fr

Joint inference for the total sensitivity indices, the components of the Sobol-Hoeffding decomposition and the Shapley indices is performed. The inference for the Shapley indices follows directly from the inference for the components of the Sobol-Hoeffding decomposition, which is based in Möbius' inversion formula. Joint asymptotic normality, and explicit formulas for the asymptotic variance-covariance matrices, are obtained. This allows to compute exact asymptotic confidence intervals for the estimators. When the number of inputs increases, however, the algorithmic complexity for computing those confidence intervals explodes. A randomization mechanism is investigated to overcome this issue.

Speaker: Ralph Smith

Parameter Subset Selection Techniques to Determine Identifiable Parameters for a Mathematical Model of Antibody Therapies for Neurological Diseases

Kamala Dadashova
Department of Mathematics, North Carolina State University, USA, kdadash@ncsu.edu

Ralph C Smith
Department of Mathematics, North Carolina State University, USA, rsmith@ncsu.edu

Attaining sufficient delivery of antibodies across blood-brain barriers (BBB) constitutes a significant challenge in the development of drugs to treat central nervous system (CNS) disorders. Since not all antibodies can pass through blood-brain barriers, it is crucial to understand antibody exposure in the central nervous system to construct drug characteristics and identify proper dosing regimens. We focus on a minimal physiologically-based pharmacokinetic (mPBPK) model of the brain for antibody therapeutics, which was developed in [1]. This model is based on a previous multi-species platform brain PBPK model, which is reduced to 16 differential equations to improve the efficiency of simulations. The model includes 31 parameters, whose values are obtained from the original brain PBPK model. In this presentation, we will discuss the use of a sensitivity-based parameter subset selection algorithm to determine those parameters which are identifiable in the sense that they can be uniquely determined by data. We illustrate this for ascending human doses. Issues to be discussed include the computation of sensitivities using sensitivity equations and complex-step approximations and local versus quasi-global analysis. We will also discuss qualitative verification techniques in addition to quantitative techniques based on energy statistics.

[1] P. Bloomingdale, S. Bakshi, C. Maass, E. van Maanen, C. Pichardo-Almarza, D. Bumbaca Yadav, P. van der Graaf and N. Mehrotra. Minimal brain PBPK model to support the preclinical and clinical development of antibody therapeutics for CNS diseases. *J Pharmacokinet Pharmacodyn*, https://doi.org/10.1007/s10928-021-09776-7, 2021.

Speaker: Amandine Marrel

Improvements around the use HSIC-based sensitivity analysis for functional data

Amandine Marrel
CEA, DES, IRESNE, DER, Cadarache F-13108 Saint-Paul-Lez-Durance,
Institut de Mathématiques de Toulouse, 31062, Toulouse, France,
amandine.marrel@cea.fr
Reda El Amri
CEA, DES, IRESNE, DER, Cadarache F-13108 Saint-Paul-Lez-Durance,
elamri.reda@yahoo.com

As part of safety studies for nuclear reactors, numerical simulators are fundamental tools for understanding, modelling and predicting physical phenomena. These tools can take a large number of input parameters, characterizing the studied phenomenon or related to its physical and numerical modelling. The information related to some of these parameters is often limited or uncertain, this can be due to the lack or absence of data, measurement or modelling errors, or a natural variability of the parameters. In this framework, global sensitivity analysis (GSA) aims at studying the impact of the input uncertainties on the output of the model. GSA methods therefore requires to characterize the input uncertainties over their variation range, for example by assigning a probability distribution to the input vector, and are mostly based on Monte Carlo simulations of the model, i.e. on a random sampling of inputs according to their probability distributions.

We focus here on the sensitivity measures based on the Hilbert Schmidt Independence Criterion (HSIC), introduced by [5] and then [3] for GSA purpose. These dependence measures are built upon kernel-based approaches for detecting dependence, and more particularly on cross-covariance operators in reproducing kernel Hilbert spaces (RKHS) composed of mapping functions (features) and characterized by positive definite kernel function. This amounts to considering covariance between feature functions applied to two variables (here each of the d inputs $\{X_i\}_{i=1,\dots,d}$ and the output Y). This set of functions (possibly non-linear), which is defined by the space and the kernel, can be of infinite dimension and allow to capture a very broad spectrum of forms of dependency. The HSIC is then defined as the squared Hilbert-Schmidt norm of the cross-covariance operator and somehow "summarizes" the set of covariances between features. In practice, HSIC offers several other advantages, especially for time-consuming simulators. Indeed, it can be estimated from a simple Monte-Carlo sample in a very simple (direct expression with kernels) and low cost way (sample of a hundred simulations is often sufficient, estimation independent from the dimension of the inputs). Moreover, under the use of characteristic kernels, the nullity of $\mathrm{HSIC}(X_i, Y)$ is equivalent to the independence of X_i and Y. HSIC can be used quantitatively to rank the inputs by order of influence on the output, as well as qualitatively to perform a screening of the inputs. For this last purpose, statistical

independence tests built upon HSIC can be used [6, 4, 1]. These tests provides a more rigorous and accurate statistical and mathematical framework than a simple assessment and comparison of HSIC sensitivity measure values.

In this work, we focus on the extension of HSIC-based GSA to multivariate and functional outputs. The HSIC can be applied to scalar, vector or even categorical random variables as long as the notion of distance is well established in these cases. In the case of functional random variables, the kernel literature has already proposed several ways to handle curves or images for regression or classification purposes. For instance, the Principal Component Analysis(PCA)-kernel [2] can be used, as illustrated in [3]. We introduce a new kernel more suitable for functional output which relies on both functional PCA and a relevant weighted combination of (one-dimensional) kernels applied to each PCA component. We compare it with other previous definitions of kernels, all based on dimension reduction technique. The performance of the proposed kernel is illustrated through analytical examples, by showing that it led to a ranking of the inputs more consistent with the real influence of the variables. Its interest is also highlighted on several industrial applications: a compartmental epidemiological model with temporal outputs and a nuclear test case with high-dimensional multivariate output and a strongly limited budget of simulations.

[1] Mélisande Albert, Béatrice Laurent, Amandine Marrel, and Anouar Meynaoui. Adaptive test of independence based on hsic measures. *To appear in Annals of Statistics*, 2021.

[2] Hervé Cardot, Frédéric Ferraty, and Pascal Sarda. Functional linear model. *Statistics & Probability Letters*, 45(1):11–22, 1999.

[3] Sébastien Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305, 2015.

[4] Matthias De Lozzo and Amandine Marrel. New improvements in the use of dependence measures for sensitivity analysis and screening. *Journal of Statistical Computation and Simulation*, 86(15):3038–3058, 2016.

[5] Arthur Gretton, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf. Measuring statistical dependence with hilbert-schmidt norms. In *International conference on algorithmic learning theory*, pages 63–77. Springer, 2005.

[6] Arthur Gretton, Kenji Fukumizu, Choon H Teo, Le Song, Bernhard Schölkopf, and Alex J Smola. A kernel statistical test of independence. In *Advances in neural information processing systems*, pages 585–592, 2008.

Speaker: Andrea Saltelli

What can sensitivity analysis contribute to a sociology of quantification?

Monica Di Fiore, Institute for Cognitive Sciences and Technologies, Consiglio Nazionale delle Ricerche, Roma, (IT), monica.difiore@istc.cnr.it

Arnald Puy, Department of Ecology and Evolutionary Biology, M31 Guyot Hall, Princeton University, New Jersey, 08544, USA. E-Mail: apuy@princeton.edu, and Centre for the Study of the Sciences and the Humanities (SVT), University of Bergen, Parkveien 9, PB, 7805, 5020 Bergen, Norway.

Andrea Saltelli, UPF Barcelona School of Management, Carrer de Balmes, 132, 08008 Barcelona, Spain, andrea.saltelli@gmail.com

Sociology of quantification is an expanding field touching on many families where numbers are produced, from statistical and mathematical modelling to data science, algorithms, big data, quantified self and indicators of various level of aggregation [1], [2]. In principle, some of the ideas brewed in the cellars of sensitivity analysis may find application to several families of quantification, beyond that of mathematical models *stricto sensu*; for example, SA attention to issues of design of the experiment can be of broad applicability; functional and variance decomposition are also very general tools which may find use in different settings [3]. Some of the ideas put forward to 'tame' the opacity of algorithms – such as that of exploring different set of assumptions in its making [4] - resonate with SA concepts such as modelling of the modelling process. Can the wisdom of sensitivity analysis be translated into a set of norms or precepts to feed into an epistemology – or hermeneutics - of quantification? In the present contribution we extend the principles laid down in [5] for mathematical modelling toward other instances of quantification, in an attempt to feed this generalization into the broader arena of sociology of quantification.

[1] E. Popp Berman and D. Hirschman, "The Sociology of Quantification: Where Are We Now?," *Contemp. Sociol.*, vol. 47, no. 3, pp. 257–266, 2018.
[2] A. Mennicken and W. N. Espeland, "What's New with Numbers? Sociological Approaches to the Study of Quantification," *Annu. Rev. Sociol.*, vol. 45, no. 1, pp. 223–245, 2019, doi: 10.1146/annurev-soc-073117-041343.
[3] A. Saltelli *et al.*, *Global sensitivity analysis : the primer*. John Wiley, 2008.
[4] L. Amoore, *Cloud Ethics, Algorithms and the Attributes of Ourselves and Others*. Duke University Press, 2020. [Online]. Available: https://www.dukeupress.edu/cloud-ethics
[5] A. Saltelli *et al.*, "Five ways to ensure that models serve society: a manifesto," *Nature*, vol. 582, pp. 482–484, 2020.

Speaker: Robert Milton

Global Sensitivity Analysis with Multi-Output Gaussian Processes

Robert Milton
Department of Chemical and Biological Engineering, University of Sheffield, UK,
r.a.milton@sheffield.ac.uk
Aaron Yeardley
Department of Chemical and Biological Engineering, University of Sheffield, UK,
asyeardley1@sheffield.ac.uk
Solomon Brown
Department of Chemical and Biological Engineering, University of Sheffield, UK,
s.f.brown@sheffield.ac.uk

Many applications benefit from, or even rely upon, determining which inputs have the greatest effect on some output. Optimisation, control and risk management problems are usually concerned with global sensitivity of the ouput to variations over the entire input space. Often the underlying motivation is to prune design space because direct simulations/experiments are costly, and to facilitate visualisation, analysis and optimisation of the output. These reasons also favour emulation by a surrogate. Thus, surrogacy and global sensitivity analysis are apt to work hand in hand.

Much global sensitivity analysis is variance-based, focussing on the Sobol indices [1]. Usually these are calculated directly by Monte Carlo methods, but there have been some studies where a surrogate is employed [2]. The focus has, until recently, been on scalar outputs. In this work we present the full calculation of Sobol indices and their uncertainties for a significant class of multi-output Gaussian Processes (MOGP). This formulation admits L outputs as a function of M inputs, producing an $M \times L \times L$ tensor of total effect Sobol indices for the $L \times L$ output covariance matrix. This tensor may inform an assessment of input relevance in a number of ways, such as isolating which inputs are most relevant to a particular output, or the linkage (covariance) between two outputs. The Sobol index itself is developed as a random variable. Although it is not normally distributed, we calculate its variance as a measure of uncertainty, naturally as an $M \times L \times L \times L \times L$ tensor of covariances between Sobol indices.

The MOGPs used herein are limited to an ARD squared exponential kernel

$$\operatorname{cov}\left[[\mathsf{y}|x]_l, [\mathsf{y}|x']_{l'}\right] = \left[k_{\mathsf{y}}(x,x')\right]_{l\times l'} = [F]_{l\times l'} \exp\left(-\frac{(x-x')^T [\Lambda]_l^{-1} [\Lambda]_{l'}^{-1} (x-x')}{2}\right)$$

where $[\Lambda]_l$ is a diagonal $M \times M$ matrix for each $l \leq L$. Thus there is a single lengthscale associated with each input/output combination. The squared exponential form has always been widely favoured for smoothly varying outputs, and is key to analytic tractability.

The calculations have been implemented in Python and tested against known values for Ishigami and Sobol g-functions. The methods used here open up exciting possibilities for model reduction, relating explanatory power to uncertainty and the correlation between the predictions of different models. In particular, these calculations are performed in a context ready to rotate the input basis to achieve optimal dimension reduction to an active subspace of inputs. Furthermore, these calculations are formally valid even when the underlying inputs are mutually dependent. The M predictive MOGP input dimensions of x can vary independently even when the training input dimensions cannot - but the predictions will revert to the MOGP prior as there is no training data nearby. This work is significantly motivated by the hope that this circumstance may be signalled by large uncertainties in relevance measures such as the Sobol indices.

[1] Bertrand Iooss, Paul Lemaître. A review on global sensitivity analysis methods. in C. Meloni and G. Dellino. *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*. Springer, 2015.

[2] A. Marrel, B. Iooss, B. Laurent and O. Roustant. Calculations of Sobol indices for the Gaussian process metamodel. *Reliability Engineering & System Safety*, 94: 742-751 2009.

Speaker: Yaning Liu

Global Sensitivity Analysis with Surrogate Modeling using Fourier Amplitude Sensitivity Testing

Yaning Liu
Department of Mathematical and Statistical
Sciences, University of Colorado Denver, USA,
yaning.liu@ucdenver.edu
Neissrien Alhubieshi
Department of Mathematical and Statistical
Sciences, University of Colorado Denver, USA,
neissrien.alhubieshi@ucdenver.edu

Fourier Amplitude Sensitivity Testing (FAST) has long been used as a variance-based global sensitivity analysis (GSA) method. We develop a high dimensional model representation framework to use FAST as a surrogate model, called FAST-HDMR, so that not only can variance-based GSA be obtained, other analyses in uncertainty quantification, such as moment-independent GSA, forward uncertainty propagation and inverse modeling, can be investigated using the FAST surrogate as a byproduct with negligible computational cost. To enhance the accuracy of FAST surrogate models, we employ various variance reduction methods for sampling that can significantly reduce the error related to estimating the coefficients in FAST-HDMR, and sparse regression techniques that can remove the noisy components. We demonstrate the accuracy, efficiency and versatility of FAST-HDMR by a wide range of applications.

Speaker: Nora Lüthen

Poincaré chaos expansions for global sensitivity analysis and surrogate modelling

Nora Lüthen
Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich, Switzerland,
luethen@ibk.baug.ethz.ch

Olivier Roustant
INSA Toulouse and Institut de Mathématiques de Toulouse, France,
roustant@insa-toulouse.fr

Fabrice Gamboa
Institut de Mathématiques de Toulouse, France,
fabrice.gamboa@math.univ-toulouse.fr

Bertrand Iooss
EDF Lab Chatou and Institut de Mathématiques de Toulouse, France,
bertrand.iooss@edf.fr

Stefano Marelli
Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich, Switzerland,
marelli@ibk.baug.ethz.ch

Bruno Sudret
Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich, Switzerland,
sudret@ibk.baug.ethz.ch

Sobol' sensitivity analysis is a popular method for determining the global importance of independent input variables for the variance of the output of a computational model. While computing Sobol' sensitivity indices based on Monte-Carlo integration is costly, it has been found that they can be computed more efficiently by first computing a spectral expansion – i.e., a representation of the model in terms of an orthonormal basis – and then post-processing the expansion coefficients. This works whenever the spectral expansion yields an approximation to the ANOVA (a.k.a. Sobol'-Hoeffding) decomposition. One example are polynomial chaos expansions (PCE), which represent a finite-variance random variable (the model output) in terms of a basis consisting of polynomials that are orthonormal with respect to the distribution of the input random variables.

While PCE are widely used because they are easy to generate and work well in many practical applications, they are not the only possible choice for computing Sobol' indices. Any spectral expansion based on a tensor-product basis of univariate orthonormal basis functions will yield an approximation to the ANOVA decomposition, as long as each of the univariate bases contains the constant function. Recently, a new type of spectral expansion has been proposed: Poincaré chaos expansions, which have the special (and defining) property that the partial derivatives of the basis are again an orthonormal basis

with respect to the same distribution as the basis itself [1, 2, 3]. Poincaré basis functions are tensor products of univariate basis functions that are generated as eigenfunctions of the so-called Poincaré differential operator [1]. They are in general non-polynomial, with the exception of the Poincaré basis associated with the Gaussian distribution, which coincides with the Hermite polynomials.

By virtue of their special property, Poincaré expansions are ideally suited for making use of available model derivatives, as demonstrated in [2] for projection-based and in [3] for sparse regression-based computation. In this talk, we summarize the results of our recent paper [3] and present a number of further developments. We demonstrate the efficiency of sparse regression for the computation of expansion coefficients, and show that derivative-based computation results in accurate Sobol' index estimates. The special Poincaré property also allows the analytical computation of the derivative-based global sensitivity measures (DGSM), which have been shown to be upper bounds to the total Sobol' indices [4]. Furthermore, the regression formulation makes it possible to compute the coefficients by simultaneously using model evaluations and derivatives, an idea repeatedly explored in the context of PCE (gradient-enhanced PCE) and easily adaptable to Poincaré expansions.

[1] O. Roustant, F. Barthe, and B. Iooss, *Poincaré inequalities on intervals - application to sensitivity analysis*, Electronic Journal of Statistics, 2 (2017), pp. 3081–3119.

[2] O. Roustant, F. Gamboa, and B. Iooss, *Parseval inequalities and lower bounds for variance-based sensitivity indices*, Electronic Journal of Statistics, 14 (2020), pp. 386–412.

[3] N. Lüthen, O. Roustant, F. Gamboa, B. Iooss, S. Marelli, and B. Sudret, *Global sensitivity analysis using derivative-based sparse Poincaré chaos expansions*, arXiv preprint arXiv:2107.00394 (2021).

[4] M. Lamboni, B. Iooss, A.-L. Popelin, and F. Gamboa, *Derivative-based global sensitivity measures: General links with Sobol' indices and numerical tests*, Mathematics and Computers in Simulation, 87 (2013), pp. 45–54.

Speaker: Silvana M. Pesenti

Sensitivity Measures based on Scoring Functions

Tobias Fissler
Institute for Statistics and Mathematics, Vienna University of Economics and Business (WU), Austria, tobias.fissler@wu.ac.at

Silvana M. Pesenti
Department of Statistical Sciences, University of Toronto, Canada, silvana.pesenti@utoronto.ca

We propose sensitivity measures based on strictly consistent scoring functions. Scoring functions are widely used to quantify the accuracy of point forecasts or point estimates for some target functional such as the mean, some quantile, or a risk measure. Consistent scoring functions are coherent with respect to increasing information sets [4], a property our sensitivity measures inherit. Moreover, we show that our sensitivity measures adhere to the nullity property in the sense that only information relevant for modelling the target functional at hand is considered – a property we term *zero information gain*.

The mathematical framework we work with is a typical sensitivity analysis setup that consists of random input factors $\boldsymbol{X} = (X_1, \dots, X_n)$ and a corresponding output of interest $Y = g(\boldsymbol{X})$, where $g\colon \mathbb{R}^n \to \mathbb{R}$. The modelling goal is $T(Y)$, some functional of the distribution of Y. We regard the functional T as a law-determined map from a space of random variables $\mathcal{Y}$ into the reals and write $T(Y)$ for its value when $Y \in \mathcal{Y}$. The law-determined property means that $T(Y)$ depends only on the distribution of Y. For a subset of input factors $\boldsymbol{X}_{\mathcal{I}}$, $\mathcal{I} \subseteq \{1, \dots, n\}$, we denote by $T(Y|\boldsymbol{X}_{\mathcal{I}})$ the random variable corresponding to the functional evaluated at the conditional distribution of Y given $\boldsymbol{X}_{\mathcal{I}}$. Sensitivity measures assess the accuracy gain from modelling $T(Y)$ – without additional knowledge – in comparison to using the information pertaining to the input factors $\boldsymbol{X}_{\mathcal{I}}$, that is, when modelling $T(Y|\boldsymbol{X}_{\mathcal{I}})$.

A scoring function is a map $S\colon \mathbb{R}\times\mathbb{R} \to [0, \infty)$ such that a point forecast $z \in \mathbb{R}$ receives the (negatively oriented) score $S(z, y)$, if $Y = y$ materialises. A scoring function is consistent for a functional T if

$$\mathbb{E}[S(T(Y), Y)] \leq \mathbb{E}[S(z, Y)]\,, \tag{1}$$

for all $z \in \mathbb{R}$ and for all $Y \in \mathcal{Y}$. It is strictly consistent if it is consistent and if equality in (1) holds if and only if $z = T(Y)$. We assume, without loss of generality, that a fully informed forecast receives a vanishing score, i.e., $S(T(Y|\boldsymbol{X}), Y) = S(T(Y|Y), Y) = 0$. This normalisation together with (strict) consistency implies that $S(z, Y)$ is (strictly) positive if $z \neq T(Y|Y)$. The tower property of the conditional expectation together with consistency leads to the sensitivity with respect to increasing information sets [4], i.e.

$$\mathbb{E}[S(T(Y), Y)] - \mathbb{E}[S(T(Y|\boldsymbol{X}_{\mathcal{I}}), Y)] \geq 0, \tag{2}$$

which gives rise to the following definition of a sensitivity score, which subsumes the well-known Sobol indices:

Let S be a scoring function (strictly) consistent for T, then the *sensitivity score* of Y with respect to $\boldsymbol{X}_{\mathcal{I}}$, $\mathcal{I} \subseteq \{1, \dots, n\}$, is given by

$$\xi^S_{\boldsymbol{X}_{\mathcal{I}}} = \frac{\mathbb{E}[S(T(Y), Y)] - \mathbb{E}[S(T(Y|\boldsymbol{X}_{\mathcal{I}}), Y)]}{\mathbb{E}[S(T(Y), Y)]} \,. \tag{3}$$

The non-negativity of S together with (2) implies that $\xi^S_{\boldsymbol{X}_{\mathcal{I}}} \in [0, 1]$. A higher sensitivity score reflects a larger value of information, when learning the distributions of input factors. If S is strictly consistent, then $\xi^S_{\boldsymbol{X}_{\mathcal{I}}} = 0$ if and only if $\boldsymbol{X}_{\mathcal{I}}$ does not contain any information for modelling $T(Y)$ (that is, $T(Y|\boldsymbol{X}_{\mathcal{I}}) = T(Y)$). On the other hand, a sensitivity score of 1 means that the risk factor(s) is as valuable as the entire set of input factors $\boldsymbol{X}$ (that is, $T(Y|\boldsymbol{X}_{\mathcal{I}}) = T(Y|Y)$). Our sensitivity measure is closely connected to what is known in the forecasting literature as *skill scores* and in regression analysis as the *universal coefficient of determination*; see e.g., [3]. In the sensitivity analysis literature the connection between scoring functions and sensitivity measures was first discussed by [1]. Our sensitivity score can be thought of as the ratio of uncertainty, $\mathbb{E}[S(T(Y), Y)]$, which can be explained by $\boldsymbol{X}_{\mathcal{I}}$, much like the notorious R^2, which arises when T is the mean and S is the squared loss.

While there is a vast literature on sensitivity measures, only recently have sensitivity measures suitable for quantile based risk measures been introduced, e.g., [5]. In this work, we provide a holistic view on sensitivity scores suitable for models when the functional is a risk measure. Specifically, we propose sensitivity scores for the Value-at-Risk and the Expected Shortfall, relying on joint consistent scoring functions for this pair [2]. We discuss properties of sensitivity scores, non-uniqueness of scoring functions by studying Murphy diagrams, and provide an alternative to the property *nullity-implies-independence* of a sensitivity measure; for a definition see [1]. We show that, depending on the functional T, the nullity-implies-independence is too strong a requirement and introduce the notion of *zero information gain* which is consistent with the framework of sensitivity scores.

[1] E. Borgonovo, G. B. Hazen, V.R.R. Jose, and E. Plischke. Probabilistic sensitivity measures as information value. *European Journal of Operational Research*, 289(2):595–610, 2021.

[2] T. Fissler and J. F. Ziegel. Higher order elicitability and Osband's principle. *Ann. Statist.*, 44(4):1680–1707, 2016.

[3] T. Gneiting and J. Resin. Regression diagnostics meets forecast evaluation: Conditional calibration, reliability diagrams, and coefficient of determination. *arXiv:2108.03210*.

[4] H. Holzmann and M. Eulert. The role of the information set for forecasting – with applications to risk management. *Ann. Appl. Stat.*, 8:79–83, 2014.

[5] Silvana M Pesenti, Pietro Millossovich, and Andreas Tsanakas. Cascade sensitivity measures. *Risk Analysis*, 2021.

Speaker: Paul Rochet

Sensibility analysis for non-parametric variable selection

Thierry Klein
SINA, Ecole Nationale de l'Aviation Civile, France, thierry01.klein@enac.fr
Paul Rochet
SINA, Ecole Nationale de l'Aviation Civile, France, paul.rochet@enac.fr

In sensitivity analysis, Sobol indices can be used for testing non-parametrically the significance of a set of variables $X_1, ..., X_J$ to explain a real-valued square integrable variable Y. For a subset $\mathbf{u} \subset \{1, ..., J\}$, a natural notion of non-parametric significance can be defined by stating that the variables $X_j, j \notin \mathbf{u}$ are not significant to explain Y in presence of $X_j, j \in \mathbf{u}$ if the null hypothesis

$$H_0 : \mathrm{E}(Y|X_j, j \in \mathbf{u}) \overset{a.s.}{=} \mathrm{E}(Y|X_1, ..., X_J)$$

is true. Denoting by

$$S^{\mathbf{u}} = \frac{\mathrm{var}\left(\mathrm{E}(Y|X^j, j \in \mathbf{u})\right)}{\mathrm{var}(Y)}$$

the Sobol index associated to the variables $X_j, j \in \mathbf{u}$ for $\mathbf{u} \subseteq \{1, ..., J\}$, the hypothesis H_0 is actually equivalent to the equality $S^{\mathbf{u}} = S$ where S is the Sobol index associated to the whole collection $X_1, ..., X_J$. By an appropriate choice of the maximal set $X_1, ..., X_K$ and the subet $\mathbf{u}$, this approach can be used for instance to test the absolute non-parametric significance of a single variable X_j setting $H_0 : 0 = S^{\{j\}}$ (for $J = 1$ and $\mathbf{u} = \emptyset$), or of one additional variable $X_{j'}$ in the presence of $X_j, j \in \mathbf{u}$ by considering $H_0 : S^{\mathbf{u}} = S^{\mathbf{u} \cup \{j'\}}$.

This hypothesis testing framework was studied in [1] given an iid sample of the variables. However, the *Pick-Freeze* method used to derive the estimators of the Sobol indices requires a particular experimental design which forbids for instance to test different hypothesis using the same sample.

We propose a different approach to test the non-parametric significance that relies on the empirical moments of Y restrained by the values of $X_j, j = 1, ..., J$. If the null hypothesis $H_0 : S^{\mathbf{u}} = S$ holds, the process

$$\hat{\xi}(x) := \frac{1}{n}\sum_{i=1}^{n} Y_i 1\{X_i \leq x\} - \frac{1}{n}\sum_{i=1}^{n} Y_i 1\{X_i^{(\mathbf{u})} \leq x^{(\mathbf{u})}\} \times \frac{1}{n}\sum_{i=1}^{n} 1\{X_i^{(\overline{\mathbf{u}})} \leq x^{(\overline{\mathbf{u}})}\} \ , \ x \in \mathrm{R}^J$$

where $x^{(\mathbf{u})} = (x_j)_{j \in \mathbf{u}}$ and $\overline{\mathbf{u}} = \{1, ..., J\} \setminus \mathbf{u}$, is asymptotically Gaussian with zero mean. A Monte-Carlo estimation of an ℓ^2 norm of this process can then serve as a test statistics for

the non-parametric significance of $X_j, j \notin \mathbf{u}$, given that the asymptotic variance can be suitably estimated. The main advantage of this method is that it can be used to test the non-parametric significance for any given subset $\mathbf{u}$ (and any initial collection $X_1, ..., X_J$) using the same sample, where no particular design of experiment is needed.

A numerical analysis shows that this new approach performs well in terms of both significance level and power. A comparison with the Pick-Freeze based approach is carried out with convincing results.

[1] Fabrice Gamboa, Alexandre Janon, Thierry Klein, A Lagnoux, and Clémentine Prieur. Statistical inference for sobol pick-freeze monte carlo method. *Statistics*, 50(4):881–902, 2016.

Proportional marginal effects for sensitivity analysis with correlated inputs

Margot Hérin
EDF R&D and Sorbonne Université (Lip6), France, margot.herin@polytechnique.edu
Marouane Il Idrissi, Vincent Chabridon, Bertrand Iooss
EDF R&D, SINCLAIR AI Lab and Institut de Mathématiques de Toulouse, France,
marouane.il-idrissi@edf.fr, vincent.chabridon@edf.fr,
bertrand.iooss@edf.fr

Global sensitivity analysis (GSA) of a numerical model $G(.)$ aims at quantifying the inputs' influence upon its output's variability. When the inputs $X = (X_1, \dots, X_d)$ are supposed independent, the Sobol' indices [1] allow to associate each input to a percentage of the output's variance, i.e., $\mathbb{V}(G(X))$. However, when the inputs are dependent, they do not always sum up to one, and thus, do not reflect accurate variance's percentages.

In order to circumvent this issue, cooperative game theory methods, in particular *allocation rules*, have been applied to GSA, in order to allocate shares of the output's variance among the inputs. It associates to a cooperative game (D, v), a real-valued vector $(\phi_i\big((D,v)\big))_{i\in D}$, where $D = \{1, \dots, d\}$ is a set of player, and v, a real-valued positive value function defined on $\mathcal{P}_d$, the set of all subsets of D. For $i \in D$, $\phi_i\big((D,v)\big)$ represents the share of the total gain $v(D)$ allocated to the player i. One particular allocation rule is the *Shapley value*, which when applied to the game (D, S^{clos}), where $S^{clos} : A \to \mathbb{V}(\mathbb{E}[G(X)|X_A])/\mathbb{V}(G(X))$ can be computed $\forall A \in \mathcal{P}_d$, and where D represents the set of all inputs, provides an output's variance allocation among the inputs even when present a correlated probabilistic structure. These resulting allocations are known as the *Shapley effects* [2]. However, these indices present a major drawback, comonly known as the "*Shapley's joke*" [3]: an exogenous variable (i.e., not present in the model) can be granted a non-negligible share of the output's variance, as soon as it is sufficiently correlated with endogenous inputs.

To circumvent this limitation, the use of another allocation rule is proposed, namely the *proportional value*. The difference between both allocation rule can be understood by taking a two-players game $(D = \{1, 2\}, v)$. One can consider $v(\{1\})$ and $v(\{2\})$ as being the individual contributions, and $I_D = v(\{1,2\}) - v(\{1\}) - v(\{2\})$ would be the coalitional surplus. The Shapley and proportional values of the player $\{1\}$, denoted respectively $Shap_1\big((D,v)\big)$ and $PV_1\big((D,v)\big)$ are given by:

$$Shap_1\big((D,v)\big) = v\big(\{1\}\big) + \frac{I_D}{2}$$
$$PV_1\big((D,v)\big) = v\big(\{1\}\big) + \frac{v(\{1\})}{v(\{1\}) + v(\{2\})} I_D.$$

While the Shapley values redistributes the coalitional surplus equally among both players, the proportional values does redistribute them proportionally to the individual contributions.

More generally, the proportional value of a game (D, v) can be defined recursively, for positive function v, $\forall i \in D$ as:

$$PV_i\big((D,v)\big) = \frac{P(D,v)}{P(D \setminus \{i\}, v)}$$

with $P(A,v) = v(A)\left(\sum_{j\in A} P(A\setminus\{j\},v)^{-1}\right)^{-1}, \forall A \subseteq D \setminus \{\emptyset\}$, and $P(\emptyset, v) = 1$.

In order to not fall under the Shapley's joke, one needs to consider the *dual* cooperative game by taking the *marginal contribution* function $w(A) = v(D) - v(D \setminus A)$ of the game instead of the value function v itself. This means computing the allocations of the cooperative game (D, S^T) where S^T are the total Sobol' indices defined as $S^T_A = S^{clos}_D - S^{clos}_{D\setminus A}$. While the Shapley values of both initial cooperative game and its dual are equivalent, it is not the case for the proportional values. The newly proposed indices, denoted *proportional marginal effects* are defined , $\forall i \in D$ as:

$$PME_i\big((D, S^{clos}\big) = PV_i\big((D, S^T)\big).$$

Considering the proportional values of (D, S^T) leads to an interesting property: a variable with a null total Sobol' index (i.e., a variable that is not included in the model) will receive a zero share. This property is possible thanks to an extension of the originally proposed allocation to non-negative value functions, which will be presented.

The difference in behavior between the Shapley and proportional effects is outlined by means of analytical results on toy-cases. The usefulness of the proposed methodology in detecting non-influential inputs is discussed, and a data-driven algorithm for estimating the proportional marginal effects is proposed. A real-world use-case of a ultrasonic non-destructive control of a weld is also presented.

[1] I.M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1: 407—414, 1993.

[2] A.B. Owen. Sobol' indices and Shapley value. *SIAM/ASA Journal of Uncertainty Quantification*, 2: 245–251, 2014.

[3] B. Iooss, C. Prieur. Shapley effects for Sensitivity Analysis with correlated inputs : Comparisons with Sobol' Indices, Numerical Estimation and Applications. *International Journal for Uncertainty Quantification*, 9(5): 493–514, 2019.

[4] B. Feldman. The proportional value of a cooperative game, *Unpublished manuscript for a Contributed Paper at the Econometric Society World Congress 2000*, http:// netec.mcc.ac.uk., 1999

Speaker: Puy Arnald

Sensitivity analysis as a tool to probe into the relation between model complexity and uncertainty

A1 Arnald Puy
Centre for the Study of the Sciences and the Humanities, University of Bergen, Norway,
arnald.puy@uib.no
A2 Pierfrancesco Beneventano
Operation Research and Financial Engineering, Princeton University, USA,
pierb@princeton.edu
A3 Simon Levin
Dpt. of Ecology and Evolutionary Biology, Princeton University, USA,
slevin@princeton.edu
A4 Samuele Lo Piano
School of the Built Environment, University of Reading, UK,
s.lopiano@reading.ac.uk
A5 Tommaso Portaluri
Eidgenossische Technische Hochschule Zurich, ETHZ, Switzerland,
t.portaluri@gmail.com
A6 Andrea Saltelli
Open Evidence Research, Universitat Oberta de Catalunya, Spain,
andrea.saltelli@gmail.com

In this presentation we set the concept of model complexity on statistical theory and the notion of "effective dimensions" [1, 2, 4] to show that more complex mathematical models tend to produce more uncertain estimates. This is because they respectively have a higher effective dimension in the "truncation" and the "superposition" sense, i.e., in the number of influential parameters and in the order of the highest effect active in the model function. Both dimensions boost the output variance. We illustrate our case with several models of the energy, agricultural and epidemiological domain and with two meta-models that ground our findings in a very wide range of model functional forms. Our approach contributes to the debate on the so called modeling hubris [3] and may help modelers better identifying the threshold beyond which the addition of model realism no longer improves the model's fit for purpose.

[1] Russel E. Caflisch, William Morokoff, and Art Owen. Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension. *Journal of Computational Finance*, 1:27–46, 1997.

[2] S. Kucherenko, B. Feil, N. Shah, and W. Mauntz. The identification of model effective dimensions using global sensitivity analysis. *Reliability Engineering & System Safety*, 96(4):440–449, apr 2011.

[3] Andrea Saltelli, Gabriele Bammer, Isabelle Bruno, Erica Charters, Monica Di Fiore, Emmanuel Didier, Wendy Nelson Espeland, John Kay, Samuele Lo Piano, Deborah Mayo, Roger Pielke Jr, Tommaso Portaluri, Theodore M Porter, Arnald Puy, Ismael Rafols, Jerome R Ravetz, Erik Reinert, Daniel Sarewitz, Philip B Stark, Andrew Stirling, Jeroen van der Sluijs, and Paolo Vineis. Five ways to ensure that models serve society: a manifesto. *Nature*, 582(7813):482–484, jun 2020.

[4] Xiaoqun Wang and Kai Tai Fang. The effective dimension and quasi-Monte Carlo integration. *Journal of Complexity*, 19(2):101–124, 2003.

Speaker: Thierry Mara

On performing screening analysis with the Innovative Algorithm

Azzini Ivano
European Commission, Joint Research Centre, ivano.azzini@ec.europa.eu
Thierry Mara
European Commission, Joint Research Centre, thierry.mara@ec.europa.eu
Rossana Rosati
European Commission, Joint Research Centre, rossana.rosati@ec.europa.eu

Let $f(\boldsymbol{x})$ be a scalar model response with $\boldsymbol{x} = (x_1, \ldots, x_D) = (\boldsymbol{v}_1, \boldsymbol{v}_2)$ a D-dimensional independent random input vector and $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ two complementary subsets of $\boldsymbol{x}$. *Screening* analysis aims at identifying the subset of important input variables (or equivalently the subset of irrelevant variables). In this work, by important variables we mean those that have a substantial contribution to the total response variance. The Innovative Algorithm (see [1]) was recently introduced to compute the so-called first- and total-order variance-based sensitivity indices as follows,

$$\hat{S}^{IA}_{v_1} = 1 - \hat{T}^{IA}_{v_1}. \tag{1}$$

$$\hat{T}^{IA}_{v_1} = \frac{\sum_{n=1}^{N} \left(\left(f(\boldsymbol{v}^B_{1,n}, \boldsymbol{v}^B_{2,n}) - f(\boldsymbol{v}^A_{1,n}, \boldsymbol{v}^B_{2,n}) \right)^2 + \left(f(\boldsymbol{v}^A_{1,n}, \boldsymbol{v}^A_{2,n}) - f(\boldsymbol{v}^B_{1,n}, \boldsymbol{v}^A_{2,n}) \right)^2 \right)}{\sum_{n=1}^{N} \left(\left(f(\boldsymbol{v}^A_{1,n}, \boldsymbol{v}^A_{2,n}) - f(\boldsymbol{v}^B_{1,n}, \boldsymbol{v}^B_{2,n}) \right)^2 + \left(f(\boldsymbol{v}^A_{1,n}, \boldsymbol{v}^B_{2,n}) - f(\boldsymbol{v}^B_{1,n}, \boldsymbol{v}^A_{2,n}) \right)^2 \right)}. \tag{2}$$

where $\boldsymbol{x}^A = (\boldsymbol{v}^A_1, \boldsymbol{v}^A_2)$ and $\boldsymbol{x}^B = (\boldsymbol{v}^B_1, \boldsymbol{v}^B_2)$ are two independent (Quasi-) Monte Carlo samples of size N from which one creates $\boldsymbol{x}^{A_{v_1}} = (\boldsymbol{v}^A_1, \boldsymbol{v}^B_2)$ and $\boldsymbol{x}^{B_{v_1}} = (\boldsymbol{v}^B_1, \boldsymbol{v}^A_2)$. This estimator has the following desirable features, $\hat{T}^{IA}_{v_i} \geq \hat{S}^{IA}_{v_i}$ and $\hat{S}^{IA}_{v_1} + \hat{T}^{IA}_{v_2} = \hat{S}^{IA}_{v_2} + \hat{T}^{IA}_{v_1} = 1$ for any N. Hence, if $\boldsymbol{v}_i$ contains all the relevant inputs, then $\hat{T}^{IA}_{v_i} \approx \hat{S}^{IA}_{v_i} \approx 1$.

To estimate the overall individual sensitivity indices, the innovative algorithm requires $N_t = 2N(D+1)$ model runs. This might be unaffordable if D and N are high or the model run is computationally demanding. In this case, it is better to first perform a screening analysis and eventually to estimate accurately the sensitivity indices of the relevant input variables. To perform the screening analysis we propose to use the sequential bifurcation technique [3] but based on the innovative algorithm at very sample sizes N.

The algorithm will be discussed and tested on several benchmarks which involve large number of input variables. In particular, the approach will be applied to the Metafunctions recently introduced for this purpose in [2].

[1] I. Azzini and R. Rosati. Sobol' main effect index: An Innovative Algorithm (IA) using dynamic adaptive variances. *Reliability Engineering and System Safety*, 213:107647, 2021.

[2] W. Becker. Metafunctions for benchmarking in sensitivity analysis. *Reliability Engineering and System Safety*, 204:107189, 2020.

[3] B. Bettonvil and J. P. C. Kleijnen. Searching fo important factors in simulation models with many factors: Sequential bifurcation. *European Journal of Operations Research*, 96:180–194, 1996.

Speaker: Stefano Tarantola

Assessing the Performance of the Scrambled Sobol' Quasi-Number Generator: an application to Interoperability of Smart Electricity Grids

Stefano Tarantola
Directorate of Energy, Transport and Climate
Joint Research Centre, Ispra 21027, Italy
Stefano.tarantola@ec.europa.eu

Sergei Kucherenko
Imperial College London, London, SWT 2AZ, UK,
s.kucherenko@imperial.ac.uk

The aim of this work is to assess the performance of the scrambled Sobol' quasi random generator against the classic Sobol' quasi-random sampling for the estimation of global sensitivity indices.

We base our comparison on a case study in the field of smart electricity grids, which was presented at the Ninth SAMO Conference in Barcelona. In Barcelona, we used the classic Sobol' quasi-random sequences for sample generation. Here, we adopt exactly the same set-up but generate the sample using the new scrambled generator in order to guarantee a fair comparison for the estimation of the sensitivity indices.

One of the main drawbacks of using low discrepancy sequences (LDS), such as the classic Sobol' quasi-random sequences, is that there is no statistical method for computing the standard error of the estimate. The practical recipe is to compute the RMSE (the root-mean-squared relative error) using different parts of the Sobol' sequence. However, this does not have theoretical statistical justification. This means that in performing the calculations, there is no clear termination criterion for the number of points to use. On the contrary, for Scrambled LDSs, one can construct confidence intervals around the estimated value, hence providing practical error bounds. Owen [1] showed that, for a sufficiently smooth integrand f(x), the variance of the randomized net is of order (logN)^(d-1)/N^(3). Thus the integration errors are of order (logN)^((d-1)/2)/N^(3/2) in probability, which is higher than the rate (logN)^((d-1))/N attained by non-randomized (i.e., standard) LDSs. BRODA developed scrambled Sobol' Sequence generators using Owen's scrambling with additional permutations [2], which was used in this work.

The application deals with testing interoperability between a data concentrator and a set of smart meters using the methodology developed at the Smart Grids Interoperability Laboratory of the Joint Research Centre of the European Commission [3].

Data concentrators and smart meters are critical components of a smart electricity grid.

Smart meters are simple electronic devices that register real-time consumption and generation of electricity, in a household or an industry, and send the data to the electricity retailer for monitoring and billing.

A data concentrator connects smart meters from the same neighborhood providing communication capability with a large number of them. The data concentrator can be configured to request periodic electricity information (power consumption, voltage, frequency, etc.) from each smart meter.

For the correct functioning of the smart grid, components and systems must be interoperable. Interoperability is, loosely speaking, the ability of such components / systems to communicate appropriately and understand each other so that they can properly perform the function they are supposed to perform.

Design of experiments and sensitivity / uncertainty analysis are relevant components of the JRC methodology [3], which can reveal the limits of a system under test and give valuable feedback about the critical conditions that do not guarantee interoperability.

The design and analysis of experiments employed in the JRC methodology supply information about the crucial parameters that either lead to an acceptable system performance or to a failure of interoperability.

The system is stress-tested under different conditions by varying three parameters: the rate at which metering data are requested by the data concentrator (ranging from 1 second to 30 minutes), the number of smart meters connected to the data concentrator (ranging from 1 to 8) and the physical distance between the data concentrator and the smart meters (ranging from 100 to 400 meters). This latter factor could influence the success rate of the communication because of induced electromagnetic fields.

The three factors are sampled within their range using the scrambled Sobol' quasi-random number generator. For each combination of these three factors, a lab experiment is executed to measure the share of readings correctly received by the data concentrator from the smart meters. The output variable for the sensitivity analysis is the share of the successful readings.

Each experiment lasts up to 3 hours, depending on the values sampled for the first factor, plus the time to set-up the equipment. Therefore, only a limited amount of experiments can be executed, for a given time budget. As in the tests carried out for SAMO in Barcelona, we could run 16 experiments.

Tests confirm that scrambled LDS show superior performance in comparison with unscrambled Sobol' sequences.

[1] Owen, A.B., Monte Carlo variance of scrambled net quadrature. S.I.A.M. Journal of Numerical Analysis 34 (5), 1884-1910, 1997.

[2] E. Atanassov, S. Kucherenko, Description of the implementation of Owen scrambling for the Sobol' sequences, BRODA Ltd., 2018

[3] Papaioannou, I.; Tarantola, S.; Lucas, A.; Kotsakis, E.;Marinopoulos, A.; Ginocchi, M.; Olariaga-Guardiola,M.; Masera,M. Smart Grid Interoperability Testing Methodology; EUR 29416 EN; Publications Office of the European Union: Luxembourg, 2018; ISBN 978-92-79-96855-6.

Speaker: Thierry Klein

Universal sensitivity indices: application to stochastic codes and second level sensitivity analysis

Jean-Clause Fort
MAP5 Université Paris Descartes, SPC, 45 rue des Saints Pères, 75006 Paris, France.,
fort43@gmail.com

Thierry Klein
Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; ENAC - Ecole Nationale de l'Aviation Civile , Université de Toulouse, France,
thierry01.klein@enac.fr

Agnès Lagnoux
Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; CNRS. UT2J, F-31058 Toulouse, France., lagnoux@univ-tlse2.fr

The use of complex computer models for the analysis of applications from sciences, engineering and other fields is by now routine. Often, the models are expensive to run in terms of computational time. Thus it is crucial to understand the global influence of one or several inputs on the output of the system under study with a moderate number of runs afforded. When these inputs are regarded as random elements, this problem is generally called (global)sensitivity analysis.

A classical tool to perform global sensitivity analysis consists in computing the Sobol indices. They are well tailored when the output space is . Many different estimation procedures of the Sobol indices have been proposed and studied in the literature. Some are based on Monte-Carlo or quasi Monte-Carlo design of experiments (see [1, 2]). More recently a method based on nested Monte-Carlo [3] has been developed. In particular, an efficient estimation of the Sobol indices can be performed through the so-called "Pick-Freeze" method. For the description of this method and its theoretical study, we refer to [4, 5] and references therein. Some other estimation procedures are based on different designs of experiments using for example polynomial chaos expansions (see [6]).

Nowadays, the computer code output is often no longer a real-valued multidimensional variable but rather a function. Some other times, the computer code is stochastic in the sense that the same inputs can lead to different outputs. When the output of the computer code is a function (for instance, a cumulative distribution function) or when the computer code is stochastic, Sobol indices are no longer well tailored. It is then crucial to define indices adapted to the functional or random aspect of the output. When the output is vectorial or valued in an Hilbert space, some generalizations of Sobol indices are available [7, 8]. More recently, indices based on the whole distribution have been developed [9, 10]. In particular, the method relying on Cramér-von-Mises distance [11] compares the conditionnal cumulative distribution function with the unconditional

one by considering the Hoeffding decomposition of half-space indicators (rather than the Hoeffding decomposition of the output itself) and by integrating them.

In this talk we focus on two kinds of computer codes: 1) computer codes for which the output is the cumulative distribution function of a real random variable and 2) real-valued stochastic computer codes. A first step will consist in performing global sensitivity analysis for these kinds of computer codes. Further, we focus on second-level analysis that corresponds to the sensitivity analysis with respect to the input distribution. Then we will deduce how to perform second-level sensitivity analysis using the tools developed in the first step. A code with cumulative distribution function as output can be seen as a code taking values in the space of all probability measures on . This space can be endowed with a metric. This point of view allows to define indices for this kind of codes. Further, as stochastic codes will be seen as a "discrete approximation" of codes having cumulative distribution functions as values, we will define define "natural" indices for such stochastic codes. Finally, we will embedded the framework of second-level sensitivity analysis into the framework of stochastic codes. More preciselly, we consider a black-box code f defined on a product of measurable spaces $E = E_1 \times E_2 \times \ldots \times E_p$ ($p \in N^*$) taking its values in a metric space $\mathcal{X}$. The output denoted by Z is then given by

$$Z = f(X_1, \ldots, X_p). \tag{1}$$

We

The aim of this talk is to give answers to the following questions.

Question 1 How can we perform Global Sensitivity Analysis (GSA) when the output space is the space of probability distribution functions (p.d.f.) on or the space of cumulative distribution functions (c.d.f.)?

Question 2 How can we perform GSA for stochastic computer codes?

Question 3 How can we perform GSA with respect to the choice of the distributions of the input variables?

.

[1] Kucherenko, Sergei and Song, Shufang Different numerical estimators for main effect global sensitivity indices. *JReliability Engineering & System Safety*, 165: 222-238, 2017.

[2] Owen, Art B. Better Estimation of Small Sobol' Sensitivity Indices, *ACM Trans. Model. Comput. Simul.*, 23, 2 1-11,2013.

[3] Takashi Goda Computing the variance of a conditional expectation via non-nested Monte Carlo *Operations Research Letters* 45, 1, 63-67, 2017

[4] Janon, A. and Klein, T. and Lagnoux, A. and Nodet, M. and Prieur, C., Asymptotic normality and efficiency of two Sobol index estimators, *ESAIM: Probability and Statistics*, 18, 1 342-364,2014.

[5] Gamboa, F. and Janon, A. and Klein, T. and Lagnoux, A. and Prieur, C., Statistical inference for Sobol Pick-Freeze Monte Carlo method, *Statistics*, 50,4, 881-902, 2016

[6] Sudret, B., Global sensitivity analysis using polynomial chaos expansions, *Reliability Engineering & System Safety*, 93,7, 964-979,2008.

[7] Lamboni, M. and Monod, H. and Makowski, D., Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models, *Reliability Engineering & System Safety*, 96, 4, 450-459, 2011.

[8] Gamboa, F. and Janon, A. and Klein, T. and Lagnoux, A. Sensitivity analysis for multidimensional and functional outputs, *Electronic Journal of Statistics*, 8,575-603, 2014

[9] Da Veiga, S., Global sensitivity analysis with dependence measures, *J. Stat. Comput. Simul.*, 85, 7, 1283-1305, 2015.

[10] Borgonovo, E. Iooss, B.", Moment-Independent and Reliability-Based Importance Measures, *Handbook of Uncertainty Quantification*,1-23,2016.

[11] Gamboa, Fabrice and Klein, Thierry and Lagnoux, Agnès Sensitivity analysis based on Cramér von Mises distance, *SIAM/ASA Journal on Uncertainty Quantification*, 6, 2, 522-548,2018. PU janon2012asymptotic,pickfreeze

Speaker: Marouane IL IDRISSI

Shapley effects for reliability-oriented sensitivity analysis with correlated inputs

Marouane Il Idrissi
EDF R&D and Institut de Mathématiques de Toulouse (IMT), France,
marouane.il-idrissi@edf.fr

Vincent Chabridon
EDF R&D, France, vincent.chabridon@edf.fr

Bertrand Iooss
EDF R&D and IMT, France, bertrand.iooss@edf.fr

Numerical models can be of great help with critical systems for risk and reliability assessment. Tracking and understanding failures of such systems can allow to avoid potentially dramatic consequences. *Reliability-oriented sensitivity analysis* (ROSA) [1] aims at measuring how uncertainties induced by the input probabilistic model influence the variability of (potentially rare) failure events. In particular, a set of methods aim at performing what is called *target sensitivity analysis* (TSA) [2], i.e., at catching the influence of the model's inputs (considering the entire input domain) on the *occurrence* of a failure event.

Formally, for a real-valued numerical model $G(.)$, this implies to consider the ROSA variable of interest $\mathbb{1}_{\{G(X)>t\}}(X)$, where $t \in \mathbb{R}$ is a threshold above which the system is assumed to enter a failure state. The traditional TSA quantity of interest (QoI) is the *failure probability* given by $\mathbb{P}(G(X) > t)$. The main goal of TSA is to quantify the influence of the inputs on the variability of the chosen QoI, after uncertainty propagation.

Traditional global sensitivity analysis (GSA) methods (and subsequently, the common ROSA methods) require an independence assumption on the inputs. An example would be the well-known Sobol' indices, whose interpretation is dramatically altered when inputs are correlated. Recent approaches proposed GSA indices which remain interpretable even when statistical dependency is at stake: the Shapley effects [3]. They leverage the framework of *cooperative game theory*, where the central question is the redistribution of wealth among several players. By analogy with the sensitivity analysis of model output framework, it allowed to define a particular decomposition of the output's variance.

By leveraging this framework, one can obtain TSA indices that can be easily interpretable as shares of meaningful statistics. In the case of variance decomposition, the newly proposed *target Shapley effect* [4] allocated to an input variable $X_j, j = 1, \ldots, d$, can be written as:

$$\text{T-Sh}_j = \frac{1}{d} \sum_{A \subset \{-j\}} \binom{d-1}{|A|}^{-1} \left(\frac{\mathbb{V}\left(\mathbb{E}\left[\mathbb{1}_{\{G(X)>t\}}(X)|X_{A\cup\{j\}}\right]\right) - \mathbb{V}\left(\mathbb{E}\left[\mathbb{1}_{\{G(X)>t\}}(X)|X_A\right]\right)}{\mathbb{V}\left(\mathbb{E}\left[\mathbb{1}_{\{G(X)>t\}}\right]\right)} \right),$$

where $\{-j\} = \{1, \ldots, d\} \setminus \{j\}$, and verifying $\sum_{i=1}^{d} \text{T-Sh}_i = 1$. These indices can be understood as a particular aggregation of "individual", "interaction" and "dependency" effects relative to the Shapley values. They are particularly relevant in the context of TSA where high interactions and induced correlations between inputs tend to blur the global comprehension of the underlying studied failure phenomenon. Fig. 1 provides an illustrative example.

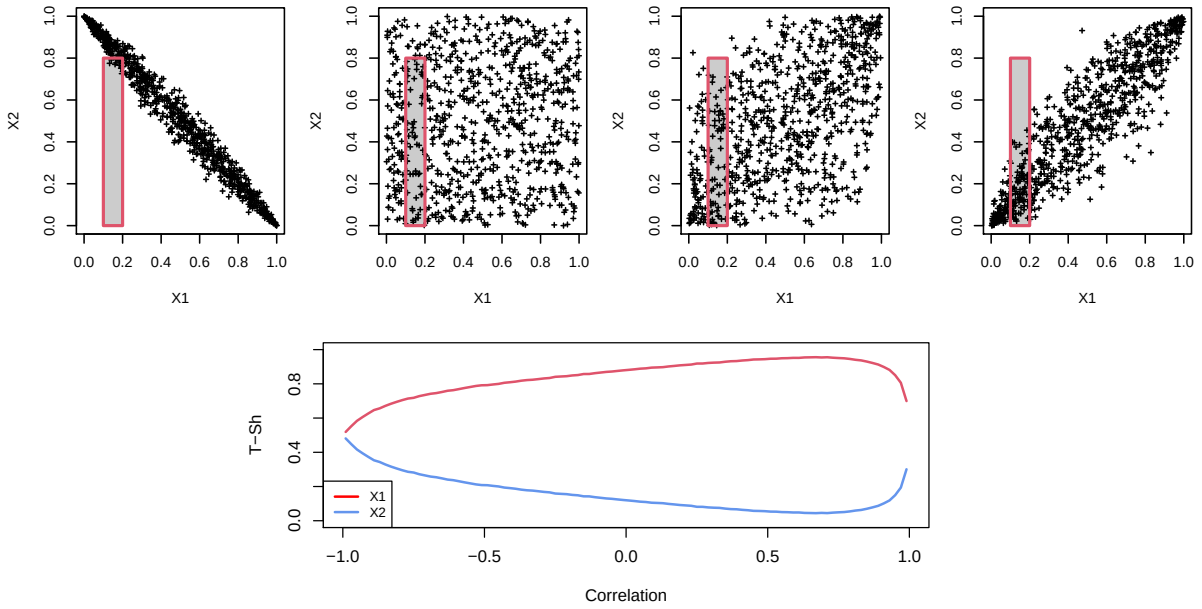


Figure 1: Target Shapley effects for a failure rectangle. The red rectangle represents the failure domain, for different correlation values between the inputs (top row) and each inputs' target Shapley effect with respect to their correlation coefficient (bottom).

Various estimation schemes will be introduced (Monte Carlo sampling as well as a given-data one applicable when a unique data sample is available), with illustrations on analytical results on toy-cases and real industrial applications (e.g., a river flood model and an ultrasonic non-destructive control of a weld).

[1] V. Chabridon. *Reliability-oriented sensitivity analysis under probabilistic model uncertainty – Application to aerospace systems*. Ph.D. thesis, Université Clermont Auvergne, 2018.

[2] A. Marrel and V. Chabridon Statistical developments for target and conditional sensitivity analysis: application on safety studies for nuclear reactor. *Reliability Engineering & System Safety*, 214: 107711, 2021.

[3] A.B. Owen. Sobol' indices and Shapley value. *SIAM/ASA Journal of Uncertainty Quantification*, 2: 245–251, 2014.

[4] M. Il Idrissi, V. Chabridon, B. Iooss. Developments and applications of Shapley effects to reliability-oriented sensitivity analysis with correlated inputs. *Environmental Modelling & Software*, 143: 105115, 2021.

Speaker: Shahroz Khan

Intra-sensitivity: An approach to regional sensitivity analysis for studying local behaviour of parametric sensitivities

Shahroz Khan
Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, Glasgow, United Kingdom, shahroz.khan@strath.ac.uk
Panagiotis Kaklis
Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, Glasgow, United Kingdom, panagiotis.kaklis@strath.ac.uk

Parametric Sensitivity Analysis (PSA) investigates the sensitivity of parameters, defining the design space of a shape-optimisation problem, for tackling the challenges of the curse of dimensionality or decreasing the uncertainty in design's performance. This is critical for complex engineering problems, especially those involving free-form shapes. Among the difficulties a robust PSA has to handle is related to the fact that a parameter can be sensitive within a certain local region of the design space but become insensitive in some other regions. Therefore, setting an applicable design space becomes a difficult and unnerving task for robust and desired results. As sensitivity analysis within a non-viable design space can be futile; either resulting in the elimination of an important parameter during dimension reduction or wastage of computational resources if uncertainty reduction is carried out with inaccurately classified sensitive parameters.

In this paper, we introduce the concept of intra-sensitivity to identify parameters whose perturbation in the range generates the largest inconsistency in the sensitivity of other parameters. For this purpose, we firstly develop and perform an ASM-based (Active Subspace Method [1]) regional sensitivity analysis, which investigates parametric sensitivity in local regions of the overall design space. The results of regional analysis are then used to conduce parameters' intra-sensitivity as they provide adequate information on how perturbing a parameter's range affects the sensitivity of itself and the remaining parameters. Once identified, intra-sensitive parameters can be tuned further to construct a viable design space, thereby avoiding uncertainty in the sensitivity results. The regional analysis has been applied in conjunction with a Dynamic Propagation Sampling approach, for tackling the computational complexity arising when high-dimensional problems are concerned. After identifying sensitive and intra-sensitive parameters, their impact on design geometry is assessed with the aid of a feature saliency map build with a Hausdorff distance-based approach [2] aiming to identify the sensitive and intra-sensitive geometrical features or regions of geometry. This feature identification also assists designers to compare different types of free-form parametrisations.

The above contributions yielded a complete pipeline for studying and analysing the regional behaviour of parametric sensitivity of free-form shape. The performance of this

pipeline is validated with an application in the area of computer-aided ship design using two parametric modellers (PM): a PD-based (Procedural Deformation) PM structured with T-splines [3, 4] and an FFD-based (Free-Form Deformation [5]) PM involving 24 and 104 geometrical parameters, respectively. The corresponding design spaces have been generated using as parent hull the KCS containership, which is extensively used for CAD and CFD experimentation in the Naval Architecture community. The volume displacement and the total resistance of the model are taken as QoI's (Quantities of Interest). Finally, the robustness and convergence performance of components of this pipeline is compared with the state-of-the-art techniques..

[1] P. G. Constantine. *Active subspaces: Emerging ideas for dimension reduction in parameter studies*. Society for Industrial and Applied Mathematics, 2015.

[2] A. Krishnamurthy., S. McMains., I. Hanniel. GPU-accelerated Hausdorff distance computation between dynamic deformable NURBS surfaces. *Computer-Aided Design*, 43(11), 1370–1379, 2011.

[3] K. V. Kostas., A. I. Ginnis., C. G. Politis., P. D. Kaklis. Ship-hull shape optimization with a T-spline based BEM–isogeometric solver. *Computer Methods in Applied Mechanics and Engineering*, 284, 611–622, 2015.

[4] T. Katsoulis., X. Wang., P. D. Kaklis. A T-splines-based parametric modeller for computer-aided ship design. *Ocean Engineering*, 191, 106433, 2019.

[5] T. W. Sederberg., S. R. Parry. Free-form deformation of solid geometric models. *In Proceedings of the 13th annual conference on Computer graphics and interactive techniques*, 151-160, 1986.

Speaker: Axelle HEGO

GSA of a green roof hydrological model with multivariate inputs and outputs

Axelle HEGO
Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France,
axelle.hego@univ-lorraine.fr
Floriane COLLIN
Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France,
floriane.collin@univ-lorraine.fr
Hugues GARNIER
Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France,
hugues.garnier@univ-lorraine.fr
Rémy CLAVERIE
Cerema Est, Team research group, F-54510 Tomblaine, France,
remy.claverie@cerema.fr

In the last decade, soil imperviousness has been one of the main urban issues in North Europe. In case of strong rain events, runoff can lead to the discharge of high volume of water and can cause water system saturation. Among all urban-water regulation systems, Green Roofs (GR) can be used to store and delay the release of rainwater to sewers [3]. GR are also considered as a sustainable solution that offers benefits such as building insulation, urban heat island cooling during summer and air pollution control.

Green roofs are composed of a substrate layer (ii) on top of which vegetation (i) is growing. Under the substrate, a geotextile (iii), a drainage layer (iv) and an insulating layer (v) protect the building (Figure 1). The hydrological performances are directly associated to the GR characteristics such as soil parameters of the different layers, dimension, type of vegetation, etc. The outflow of a GR is mainly related to the water content inside the layers. Few models exist to describe the hydrological infiltration throughout soil and they are based on the Richards' equation. This highly nonlinear partial differential equation describes the water retention capacity and the outflow in unsaturated porous media. The Richards' equation combined with the Van Genuchten - Mualem model are coded in Hydrus-1D© software to simulate hydrological behavior [4]. This software allows the set up of the green roof structure, boundary conditions, meteorological data and vegetation variables in order to reproduce the green roof real conditions.

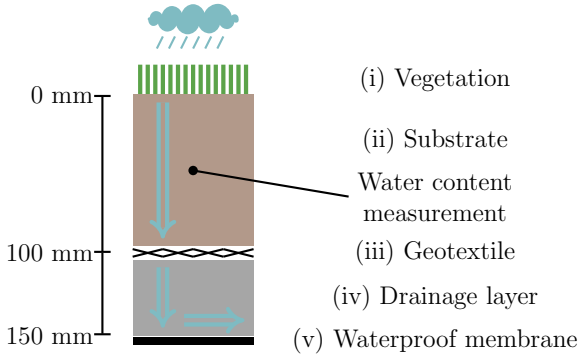


Figure 1: Profile view of a typical green roof.

However, some of the model parameters, as soil parameters or vegetation variables, are challenging to determine as they are difficult to measure accurately through experiments. All the parameter uncertainties are propagated to the water retention capacity simulated and need to be analysed with Global Sensitivity Analysis (GSA). A recent study has been dedicated to analyse the influence of soil parameters as saturated water content, porosity, etc. It has highlighted that three soil parameters of the substrate layer are more influential than the soil parameters of the other layers [2]. These results allow to reduce the number of uncertain soil parameters for the following studies. Nevertheless, it also appears that the variability of the vegetation parameters as crop height and root depth, set to a constant value in this previous study, could have effects on the water retention capacity.
In this proposed study, the GR model is dynamic and depends on static uncertain inputs, the soil parameters and on time-varying uncertain inputs, the vegetation parameters. Indeed, a one year period is analysed in order to observe different hydrological phenomenon, and the vegetation parameters as crop height or root depth vary over the seasons. Boundary conditions, meteorological data and water retention data were measured from an in-situ experimental green roof platform located in Tomblaine (France) and are used to simulate water retention with real conditions. The objective of this study is to investigate the influence of soil and vegetation parameters on the water retention capacity over time with GSA approach. The challenge is to generate samples of vegetation parameters that satisfy the random fields distribution. The random fields are assumed independent and normally distributed, defined by their means and covariance functions. One possibility is to resort to the Karhunen-Loève expansion as in [1]. Then, to compute sensitivity indices for the time-varying model output, polynomials chaos expansion is applied with a sequential [2] and PCA-based multivariate approach [2].

[1] F. Anstett-Collin, J. Goffart, T. Mara, and L. Denis-Vidal. Sensitivity analysis of complex models: coping with dynamic and static inputs. *Reliability Engineering & System Safety*, 134:268–275, 2015.

[2] A. Hégo, F. Collin, H. Garnier, and R. Claverie. Approaches for green roof dynamic model analysis using GSA. In *19th IFAC symposium on system identification: learning models for decision and control, Padova, Italy*, 2021.

[3] Y. Li and R.W. Badcock. Green roof hydrologic performance and modeling: a review. *Water Science & Technology*, 69(4):727–738, 2014.

[4] J. Simunek, M. Sejna, H. Saito, M. Sakai, and M.Th. van Genuchten. *The HYDRUS-1D software package for simulating one-dimensional movement of water, heat, and multiple solutes in variably saturated media.* Department of Environmental Sciences University of California Riverside, Riverside, California, 2008.

Speaker: Razi Sheikholeslami

Why and How Uncertainty Matters for Estimating Dissolved Phosphorus in Runoff

Razi Sheikholeslami[1,2], *Kian Golkar*[1]

[1] Civil Engineering Department, Sharif University of Technology, Iran, razi.sheikholeslami@sharif.edu

[2] Environmental Change Institute, School of Geography and the Environment, University of Oxford, UK, razi.sheikholeslami@ouce.ox.ac.uk

The transport of phosphorous (P) is one of the major factors currently deteriorating the quality of water bodies around the world, leading to eutrophication, which limits sustainable use of surface water resources [1]. The contamination of water bodies by P release from widely applied fertilizers, such as superphosphate, triple superphosphate, and mono- and diammonium phosphate, is one of the major sources of nutrients and pollutant of surface waters. Therefore, a considerable number of numerical models has been introduced to quantify the release of P from fertilizers [2]. However, current modelling frameworks developed for estimating P concentrations in runoff often fail to systematically account for input uncertainties. In this study, we employed a variance-based Global Sensitivity Analysis (GSA) technique to address this issue. We show that the apparent precision in most P modelling studies of water quality results from neglecting various important uncertainties and are therefore prone to lack of robustness.
As a case study, we used the Annual P Loss Estimator (APLE) [3] which is a popular empirical model. APLE is commonly implemented to predict release of fertilizer P on the soil surface when rain falls. This model is mainly driven by parameters representing different key physical processes related to precipitation, discharge, infiltration, and fertilizer use. Considering the severe uncertainties associated with these processes and the nonlinear nature of APLE, we investigated how several sources of uncertainty and their interaction contribute to model output variability. Our GSA results indicate that the estimated direct losses of fertilizer P in runoff dramatically turn into intervals that can span several orders of magnitude if the model input space is sufficiently explored. The results and insights gained from this study have important implications for improving fertilizer pollution management in the context of decision making.

[1] W.J. Brownlie, M. A. Sutton, D. S. Reay, K. V. Heal, L. Hermann, C. Kabbe & M. B. Spears. Global Actions for A Sustainable Phosphorus Future. *Nature Food*, 2(2): 71–74. 2021

[2] D. E. Radcliffe, D. K. Reid, K. Blombäck, C. H. Bolster, A. S. Collick, Z. M. Easton, ... & D. R. Smith. Applicability of Models to Predict Phosphorus Losses in Drained Fields: A Review. *Journal of Environmental Quality*, 44(2): 614–628. 2015.

[3] P. A. Vadas, L. W. Good, P. A. Moore & N. O. R. M. A. Widman. Estimating Phosphorus Loss in Runoff from Manure and Fertilizer for a Phosphorus Loss Quantification Tool. *Journal of Environmental Quality*, 38(4): 1645–1653. 2009

Speaker: Katarina Radišić

Sensitivity analysis of a spatio-temporal hydrological model for pesticide transfers

Katarina Radišić
INRAE, UR RIVERLY, France, katarina.radisic96@gmail.com
Emilie Rouzies
INRAE, UR RIVERLY, France, emilie.rouzies@inrae.fr
Claire Lauvernet
INRAE, UR RIVERLY, France, claire.lauvernet@inrae.fr
Arthur Vidard
Inria, AIRSEA, France, arthur.vidard@inria.fr

Pesticide use is a major issue in sustainable agriculture and water quality. Therefore, it is important to have the knowledge and the tools to best estimate the risks associated with their use and propose appropriate corrective actions. The PESHMELBA model developped by [2] aims at simulating processes involved in water and pesticide transfers at the watershed scale, in order to compare scenarios of the landscape management and its impact on the river quality. An important step in the journey to PESHMELBA's operational use is performing a thorough study on the model uncertainties. A sensitivity analysis can help trace the output uncertainties back to its input parameters, verify the model consistency with respect to the physical processes and enhance the understanding of the modeled behaviour. However, the application of global sensitivity analysis (GSA) to spatio-temporal environmental models can be very challenging and dependent on the particular case studied. No universal methodology exists for performing GSA on spatio-temporal models. Additionally, GSA methods have not yet been applied to distributed pesticide transfer models.

In the case of the PESHMELBA model, the sensitivity analysis is particularly challenging. PESHMELBA is made up of coupled code units that represent interacting physical processes simulating pesticide transfers and fate. Each code unit can be characterized by its own resolution method and its own time step. This leads to a very heterogeneous final structure, which is hard to analyse. Additionally, the PESHMELBA model is a distributed model - the resolution is performed on each spatial unit individually. Then, interactions between spatial units are integrated. The spatial aspect adds another layer of difficulty to the sensitivity analysis, as spatial interactions should be taken into account to interpret the results. Both the integration of various physical processes as distinct code units and the spatial distribution of the model contribute to drastically increasing the number of input parameters. Indeed, in the case studied, PESHMELBA counts 145 input parameters. The exploration of such a high dimensional input space is challenging as it implies time consuming model evaluations.

In this study, we propose a novel GSA methodology for the PESHMELBA model, taking into account both its high dimensionality and spatio-temporal aspect. Two solutions were taken to deal with the high dimensionality of the input space: (a) screening and (b) estimating sensitivity indices via metamodels. The proposed methodology can be separated in two conceptual steps. The first step deals with the temporal aspect of the outputs, by studying each spatial unit independently. Thus, for one spatial unit at a time, sensitivity indices are obtained for its whole output dynamics. This is done by calculating sensitivity indices on the sum of scores of the functional outputs principal components [1]. In the second step, the spatial aspect is considered. The sensitivity indices obtained for each spatial unit are aggregated to the watershed scale by means of vector projections [3]. To sum up, the proposed methodology consists in: (i) screening via the elementary effect method on the principal components of the functional outputs, (ii) calculating sensitivity indices of one spatial unit at-a-time through metamodels (polynomial chaos expansion or random forest) and (iii) aggregating the sensitivity indices to a watershed scale.

Two different PESHMELBA outputs were studied, surface moisture series and pesticide mass series. The methodology was successfully applied to surface moisture series and proved its potential for improving knowledge on model behavior. The method is promising but still needs to be improved as results were less satisfactory for pesticide mass series due to the complexity of the physical processes simulated and the non-linear interactions between them.

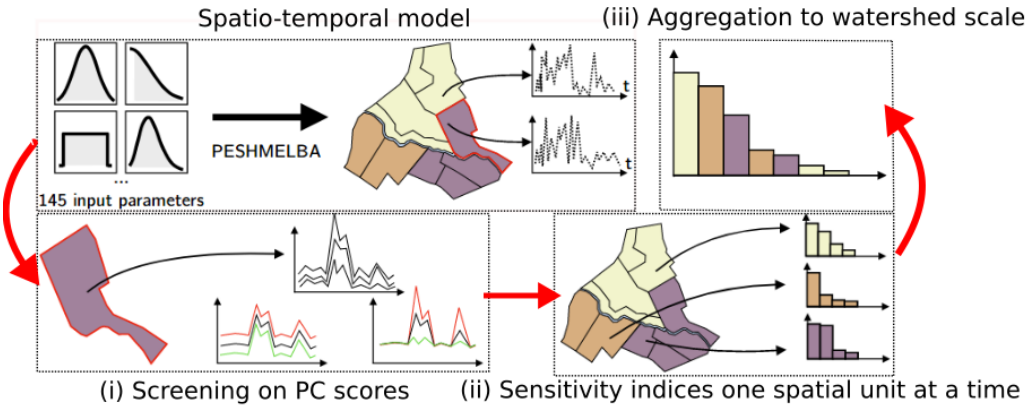


Figure 1: Schema of the proposed methodology.

[1] M. Lamboni, H. Monod, and D. Makowski. Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliability Engineering & System Safety*, 96(4):450–459, April 2011.

[2] E. Rouzies, C. Lauvernet, C. Barachet, T. Morel, F. Branger, I. Braud, and N. Carluer. From agricultural catchment to management scenarios: A modular tool to assess effects of landscape features on water and pesticide behavior. *Science of The Total Environment*, 671:1144–1160, June 2019.

[3] L. Xu, Z. Lu, and S. Xiao. Generalized sensitivity indices based on vector projection for multivariate output. *Applied Mathematical Modelling*, 66:592–610, February 2019.

Speaker: Massimo Aufiero

Surrogate-based global sensitivity analysis with confidence bounds

Massimo Aufiero
Department of Statistics, Harvard University, USA, aufiero@college.harvard.edu
Lucas Janson
Department of Statistics, Harvard University, USA, ljanson@fas.harvard.edu

Sensitivity analysis is an important practice to aid our understanding of how complex computational models are affected by their input parameters, and this topic has been studied extensively in the literature (see, e.g., [1][2][3]). However, performing a sufficient number of model evaluations to obtain accurate estimates can often be prohibitively expensive. A common solution is to use a surrogate model, or emulator, that provides reasonable approximations to the outputs of the true model but at significantly lower costs [4]. But the computational benefits of surrogate-based sensitivity analysis come with a loss in accuracy, since the difference between the surrogate and the true model will usually lead to some amount of error in the estimated sensitivity. This begs the question: **how do we know when we can trust surrogate-based sensitivity analyses?**

To answer this question, we introduce a new meta-method that can be combined with any surrogate for any computational model, augmenting the surrogate-based point estimate of the global sensitivity with a *confidence interval* so that the user can benefit from the computational speed-up of using a surrogate while still getting rigorous and accurate bounds on the *true* global sensitivity. We prove the asymptotic validity of our confidence interval under essentially no conditions on the computational model or the surrogate (except the existence of some moments), so it applies just as well to surrogates fitted via machine learning as to surrogates based on quantitative or qualitative domain knowledge, such as physically-based surrogates. While the validity of our confidence interval protects against overinterpretation of inaccurate surrogate-based sensitivity estimates, we also prove that the accuracy (width) of our confidence interval gets small as the surrogate's accuracy increases, so that when an accurate surrogate is used, the confidence interval we report will correspondingly be quite narrow, instilling appropriately high confidence in its estimate.

Letting Y denote the model output, X the set of input parameters of interest, and Z the complementary set of input parameters, we focus on the total-order sensitivity index, $\mathbb{E}[\mathrm{Var}(Y \mid Z)]/\mathrm{Var}(Y)$, which measures the total contribution of the set of inputs X both directly and through interaction effects with Z [5]. We apply our method to various computational models, including the Hymod and HBV hydrological models with Polynomial Chaos Expansion and Artificial Neural Network surrogates, demonstrating that applying our method with a high-quality surrogate provides tight bounds so that we know to trust its estimate, while also providing wide bounds when the surrogate is inaccurate to insure against overinterpretation.

[1] S. Razavi and H. Do. A new framework for comprehensive, robust, and efficient global sensitivity analysis: 1. Theory. *Water Resources Research*, 52 (1): 423–439, 2016.

[2] F. Pianosi and T. Wagener. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software*, 67: 1–11, 2015.

[3] E. Borgonovo. A new uncertainty importance measure. *Reliability Engineering and System Safety*, 92: 771–784, 2007.

[4] S. Razavi, B. A. Tolson, and D. H. Burn. Review of surrogate modeling in water resources. *Water Resources Research*, 48 (7). 2012.

[5] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis: The Primer.* John Wiley & Sons, 2008.

Speaker: Hui Duan

Bayesian PCE as a Control Variate Method for Estimating Sobol' Sensitivity Indices

Hui Duan
Department of Mathematics, Florida State University, USA, hduan@math.fsu.edu
Giray Ökten
Department of Mathematics, Florida State University, USA, okten@math.fsu.edu

Sobol' sensitivity indices are some of the most popular measures for the sensitivity of model output to input. Several Monte Carlo algorithms for the Sobol' indices have been developed in recent years. Another approach is to use a surrogate model, such as polynomial chaos expansion (PCE) [1], to estimate the Sobol' indices. An advantage of surrogate models is their reduced computational cost for computing the sensitivity indices. However, for high dimensional problems where a large number of expansion terms are required, PCE could be less efficient than Monte Carlo in estimating sensitivities.

Monte Carlo control variate algorithms using PCE expansions were introduced in Fox and Ökten [2]. In this work, we present some improved control variate algorithms using Bayesian compressive sensing (BCS) [3]. We first compare the errors of the PCE and BCS methods when they are used to estimate sensitivity indices of some test functions. The results show that BCS yields better estimation and faster convergence with fewer samples. We then compare the efficiency of the control variate algorithms with some well-known Monte Carlo algorithms [4] for estimating lower and upper Sobol' indices. The results suggest that for problems where the input function is expensive to evaluate, control variate Monte Carlo can have advantages over the other methods.

[1] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7), pp. 964-979. 2008.

[2] J. Fox, G. Ökten. Polynomial chaos as a control variate method. *SIAM Journal on Scientific Computing*, 43(3), pp. A2268-A2294, 2021.

[3] S. Babacan, R. Molina, A. Katsaggelos. Bayesian compressive sensing using Laplace priors. *IEEE Trans. Image Process.*, 19(1), pp. 53–63, 2010.

[4] A. Puy, W. Becker, S.L. Piano, A. Saltelli. A comprehensive comparison of total-order estimators for global sensitivity analysis. *International Journal for Uncertainty Quantification*, 2020.

Speaker: Gabriel Sarazin

What is hidden behind the Sobolev kernels involved in the HSIC-ANOVA decomposition?

Gabriel Sarazin
CEA, DES, IRESNE, DER, Cadarache F-13108 Saint-Paul-Lez-Durance,
gabriel.sarazin@cea.fr
Amandine Marrel
CEA, DES, IRESNE, DER, Cadarache F-13108 Saint-Paul-Lez-Durance,
Institut de Mathématiques de Toulouse, 31062, Toulouse, France,
amandine.marrel@cea.fr
Sébastien Da Veiga
Safran Tech, Modeling & Simulation Rue des Jeunes Bois, Châteaufort, 78114
Magny-Les-Hameaux, France, sebastien.daveiga@gmail.com
Vincent Chabridon
EDF R&D
6 Quai Watier, 78401, Chatou, France, vincent.chabridon@edf.fr

In many industrial fields, physical phenomena are commonly modeled by numerical simulators. These simulation codes can take a high number (dozens, if not hundreds) of uncertain parameters as input variables. Since computer-based simulation experiments are time-consuming, the permissible number of runs is most often very limited. In this context, the benefits brought by a preliminary global sensitivity analysis are twofold. On the one hand, the screening step allows to select a smaller amount of influential inputs in order to reduce dimensionality or to support the construction of a surrogate model. On the other hand, the influence of each input variable is quantified and the associated rankings provide a relevant piece of information that can be used in the sequential building process of a metamodel as in [4]. In view of these two purposes, total-order Sobol indices are often regarded as the most adapted indicators. Unfortunately, unlike their first-order counterparts, their estimation requires a too large number of simulations.

To remedy this difficulty, a new class of sensitivity measures built upon the theory of reproducing kernel Hilbert spaces (RKHS) has emerged over the last decade [1]. Each input variable is assigned a continuous positive definite kernel and is equipped with the related RKHS. The same is done for the output variable. The distance between the joint input-output bivariate distribution and the bivariate distribution obtained under independence is measured through the distance of their respective embeddings in the tensorized RKHS, which leads to the Hilbert-Schmidt Independence Criterion (HSIC) [3]. Since HSIC indices can be expressed by means of kernel-based statistical moments, their estimation can be achieved at minimal computational cost and without estimating the joint input-output distribution. However, HSIC indices are often blamed for their lack of interpretability because their sum is not equal to 1. As a consequence, they do not

fit into the advantageous framework that goes with analysis of variance (ANOVA). To bridge this gap, it has been recently proved that the use of orthogonal kernels enables an ANOVA-like decomposition for HSIC indices [2]. When the input variables are uniformly distributed, Sobolev kernels (parametrized by a smoothness parameter r) match all the required conditions. They allow to define higher order HSIC indices (especially total-order HSIC indices) and therefore to rigorously separate main effects and interactions.

The main objective of this work is to provide new insights into Sobolev kernels which have indeed been very few studied so far. First, we demonstrate that Sobolev kernels are characteristic. This ensures that the nullity of Sobolev-based HSIC indices is equivalent to independence. Then, another notable contribution of this work is the identification of one explicit feature map for Sobolev kernels, which may help the user understand the spectrum of dependency patterns that can be captured by this novel HSIC index. As with most kernels, the canonical feature map (resulting from Aronszajn's theorem) is uninformative. An additional pitfall comes from the fact that Sobolev kernels are not shift-invariant, which prevents the use of Bochner's theorem and the following characterization in terms of Fourier transform. Instead, another kind of feature map stems from Mercer's theorem which asserts that any continuous kernel on a compact set may be rewritten as an infinite sum that only implies the eigenvalues and eigenfunctions of the corresponding kernel integral operator. This key result discloses a feature map that sends the unit interval into the Hilbert space of real-valued square-summable sequences. The eigenvalue problem arising from this new feature map is tackled in two different ways. Firstly, we investigate the benefits of kernel feature analysis where eigenvalues and eigenfunctions are estimated from Gram matrix simulation. It allows to visualize eigenfunctions and sometimes to come up with closed-form candidate functions. Secondly, we demonstrate that the eigenvalue problem is equivalent to a Cauchy problem consisting of a linear homogeneous ordinary differential equation with constant coefficients and a sufficient number of boundary conditions. Then, two situations deserve a specific study. When $r = 1$, an exact analytical solution is available. On the contrary, when $r = 2$, the Cauchy problem cannot be completely solved. A solution in the form of a linear combination of analytical functions is obtained but the coefficients cannot be retrieved. An asymptotic approximation is then derived to accurately estimate small eigenvalues while robust numerical resolution is achieved by use of a semi-analytical adhoc calibration algorithm.

[1] Sébastien Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305, 2015.

[2] Sébastien Da Veiga. Kernel-based ANOVA decomposition and Shapley effects: application to global sensitivity analysis. *arXiv preprint arXiv:2101.05487*, 2021.

[3] Arthur Gretton, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf. Measuring statistical dependence with Hilbert-Schmidt norms. In *International conference on algorithmic learning theory*, pages 63–77. Springer, 2005.

[4] Amandine Marrel, Bertrand Iooss, and Vincent Chabridon. Statistical identification of penalizing configurations in high-dimensional thermalhydraulic numerical experiments: the ICSCREAM methodology. *arXiv preprint arXiv:2004.04663*, 2020.

Speaker: Matthias Van Hove

Comparison of global sensitivity analysis methods for urban scale building stock energy models

A1 Matthias Van Hove
Department of Architecture and Urban Planning, Ghent University, Belgium, matthias.vanhove@ugent.be
A2 Marc Delghust (PhD)
Department of Architecture and Urban Planning, Ghent University, Belgium, marc.delghust@ugent.be
A3 Jelle Laverge (PhD)
Department of Architecture and Urban Planning, Ghent University, Belgium, jelle.laverge@ugent.be

The aim of applying global sensitivity analysis methods (GSA) at simulation models is to characterise the impact that changes in the model input parameters have on the model output. GSA is a diagnostic tool that can guide model calibration and validation, support the prioritisation of efforts for uncertainty reduction or help with model-based decision-making (Pichery C., 2014 [1]; Song *et al*., 2015 [2]; Sarrazin *et al*., 2016 [3]).

As available processing power has increased, GSA has been steadily more utilised among building energy modellers, especially at individual building level (Tian, 2013 [4]). At building stock level, however, GSA has only scarcely been used (Fennell *et al*., 2019 [5]). Yet, several stock models are being used for decision and policy making. Even fewer studies investigated the performance of GSA at stock level (*e.g*., Cheng *et al*., 2011 [6]; Ascione *et al*., 2017 [7]), which leaves a huge gap in application at scale.

No model can be considered as a perfect representation of the world around us. All models inevitably contain uncertainty (Refsgaard *et al*., 2004 [8]). The gaps in our knowledge are bridged by assumptions, probability distributions, expert opinion, best guesses and a variety of other techniques (Gorris & Yoe, 2014 [9]). At small scale, modellers can generally rely on detailed high-quality data and less (critical) assumptions need to be made. At stock level, fewer and less detailed input parameters are generally available and the quality of the data is not always guaranteed. Furthermore, the complex physical energy equations (typically used at small scale building energy models) have to be simplified and generalised due to input parameter shortage and/or to allow for acceptable model computation times.

Hence, due to the high computation times necessary to perform a reliable GSA at scale and input data absence and/or shortage or a lack of GSA knowledge, researchers performing GSA also often limit the SA scope and/or computation time of their study by reducing the building stock size (*i.e*., the number of buildings in the stock) used for building stock model GSA as well as the number of model evaluations. The consequences of such constraints are a possibly increased uncertainty in the model output results and a reduction in GSA quality since convergence might not be reached.

Therefore, this study aims to broaden the knowledge of global sensitivity analysis application on building stock level by application of the Sobol' SA (Sobol', 1990 [10]), the Morris [11] method and DGSM [12] at an internally developed bottom-up building stock model (Delghust, 2015 [13]; Delghust *et al*., 2015 [14] [15]), based on ISO 13790 (ISO, 2007 [16]). The study will elaborate further on the required stock size, comparison of the model output of the three methods and to-be-expected uncertainty ranges for the model output and SA indices. Additionally, in order to confirm the robustness of SA results, the study will further check and elaborate on model convergence criteria.

The study is part of the IEA EBC Annex 70 on Building Energy Epidemiology (IEA EBC, 2017 [17]), where a group of research teams participated in a co-ordinated investigation to take existing GSA methods and apply them to their distinct stock models and datasets in a first attempt to quantify added value of GSA for building stock modelling. Through this process the teams aimed to examine:

- The challenges of defining input parameter uncertainties for large-scale BES models and collecting appropriate data.
- The applicability of different SA techniques in terms of robustness of results, quality assurance and computational cost.
- Key drivers of uncertainty in the models.

References

[1] Pichery, C. (2014). Sensitivity Analysis. *Encyclopedia of Toxicology (Third Edition)*. 236-237.

[2] Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., Xu, C. (2015). Global sensitivity analysis in hydrological modelling: review of concepts, methods, theoretical framework, and applications. *Journal of Hydrology*. 523, 739-757.

[3] Sarrazin, F., Pianosi, F., Wagener, T. (2016). Global Sensitivity Analysis of environmental models: Convergence and validation. *Environmental Modelling & Software*. 79, 135-152.

[4] Tian, W. (2012). A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews*. 20, 411-419.

[5] Fennell, P. J., Ruyssevelt, P. A., Mata, E., Jacob, M. (2019). A Review of the Status of Uncertainty and Sensitivity Analysis in Building-stock Energy Models. *Proceedings of the 16th International Conference of IBPSA (BS2019)*. Rome, Italy.

[6] Cheng, V., & Steemers, K. (2011). Modelling domestic energy consumption at district scale: A tool to support national and local energy policies. *Environmental Modelling & Software*. 26(10), 1186–1198.

[7] Ascione, F., Bianco, N., De Stasio, C., Mauro, G. M., & Vanoli, G. P. (2017). Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach. *Energy*, 118, 999–1017.

[8] Refsgaard J.C., Henriksen H.J. (2004) Modelling guidelines: terminology and guiding principles. *Advances in Water Resources 27*, 71-82.

[9] Gorris, L. G. M., Yoe, C. (2014). Risk Analysis: Risk Assessment: Principles, Methods, and Applications. *Encyclopedia of Food Safety*. 1, 65-72.

[10] Sobol', I. M. (1990). Sensitivity estimates for nonlinear mathematical models. *Matricheskoe Modelirovanie*. 2, 112-118 (in Russian), translated in English (1993). In: *Mathematical Modelling and Computational Experiments*. 1, 407-414.

[11] Morris M.D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33, 161-174.

[12] Sobol' I.M., Kucherenko S. (2009). Derivative based global sensitivity measures and their link with global sensitivity indices Mathematics and Computers in Simulation 79, 3009-3017.

[13] Delghust, M. (2015). Improving the predictive power of simplified residential space heating demand models : a field data and model driven study. PhD dissertation, Ghent University.

[14] Delghust, M., De Weerdt, Y., Janssens, A. (2015). Zoning and intermittency simplifications in quasi-steady state models. *Proceedings of the 6th International Building Physics Conference (IBPC2015)*. Torino, Italy.

[15] Delghust, M., Strobbe, T., De Meyer, R., Janssens, A. (2015). Enrichment of single-zone EPB-data into multi-zone models using BIM-based parametric typologies. *Proceedings of the 14th International Conference of IBPSA (BS2015)*. Hyderabad, India.

[16] ISO. (2007). ISO 13790:2007(E) Energy performance of buildings. Calculation of energy use for space heating and cooling. Geneva, Switzerland: International Organisation for Standardisation (ISO).

[17] IEA EBC (2017). Building Energy Epidemiology.

Speaker: Laura Swiler

Metamodelling sensitivity approaches versus regression and graphical methods on the basis of Geologic Cases: An International Collaboration

Laura Swiler, Dusty Brooks, Emily Stein
Sandia National Laboratories, USA
lpswile@sandia.gov, dbrooks@sandia.gov, ergiamb@sandia.gov
Klaus-Jürgen Röhlig, Elmar Plischke
Institute of Disposal Research, TU Clausthal, Germany
klaus.roehlig@tu-clausthal.de, elmar.plischke@tu-clausthal.de
Dirk-Alexander Becker, Sabine Spiessl
Repository Safety Department, GRS, Germany
Dirk-Alexander.Becker@grs.de, sabine.spiessl@grs.de
Lasse Koskinen, Pekka Kupiainen
Posiva, Finland
Lasse.Koskinen@Posiva.fi, Pekka.Kupiainen@posiva.fi
Joan Govaerts
Belgian nuclear research centre, SCK-CEN, Belgium
joan.govaerts@sckcen.be
Valentina Svitelman, Elena Saveleva
Nuclear Safety Institute of the Russian Academy of Science, Ibrae, Russia
svitelman@ibrae.ac.ru, saveleva@ibrae.ac.ru

Over the past four years, an informal working group has developed to investigate existing sensitivity analysis methods, identify best practices, and examine new sensitivity analysis methods being developed. The focus is on the use of sensitivity analysis in case studies involving geologic disposal of spent nuclear fuel or nuclear waste. We have developed multiple applicable case studies to use for testing ideas and making comparisons.[1] Four of these case studies are discussed: the GRS clay case, the SNL shale case, the Dessel case, and the IBRAE groundwater case. We present the different sensitivity analysis methods investigated by various groups, the results obtained by different groups and different implementations, and summarize our findings.

The case studies focused on repository models for underground disposal of nuclear waste. The four case studies typically included things such as a waste form (canister or steel drum encasing the waste), an engineered buffer such as bentonite or concrete, and modeling of a natural system. The processes modeled in the case studies included waste package degradation, radionuclide dissolution, radionuclide sorption and precipitation/ dissolution, radioactive decay, and radionuclide transport via advection and diffusion. The number of uncertain input parameters ranged from 6 to 20 in these case studies. The outputs included time-series data (e.g. radionuclide concentrations as a function of time and/or spatial location), dose rates, fluxes, etc.

The participating groups used a large variety of sensitivity analysis methods: scatterplots, simple correlation coefficients, rank correlation coefficients, standardized regression coefficients, main and total effects variance-based Sobol' indices estimated by methods such as EASI, RBD-FAST, distribution-based methods such as PAWN, graphical methods like CUSUNORO, and others. For each geologic case study, multiple groups presented their results using different sensitivity analysis methods and/or different implementations of the same method. The breadth and scope of the case studies as well as the number of methods used provided a rich environment to study and compare results.

We found that the first order variance-based index estimates can be easily generated from observational data (i.e. existing data which were not generated by prescribed sampling schemes) using a variety of approaches and are one of the preferred SA approaches. Linear and rank correlation coefficients and regression approaches continue to be used and are informative. More advanced methods show results mostly consistent with simpler methods but there are important differences. Graphical methods such as CUSUNORO also provide additional visualization which can show influences over the range of a variable.

We found consistency between the linear sensitivity measures (correlation and regression coefficients) calculated by the different partners but sometimes the variance-based sensitivity indices did not rank the important variables in agreement with the linear sensitivity measures. Also, there were more differences in rankings seen across the variance-based sensitivity indices from different methods, such as EASI/RBD-FAST, EFAST, PCE and RS-HDMR. Note that some of the variance-based methods make direct use of the simulation results for calculating the sensitivity indices while other methods use them to train surrogate or metamodel approximations of the simulation. All the methods used in this study relied on fixed data sets generated by the case study owners: specialized sampling of the simulations was not possible.

Parameter rankings obtained by Sobol' method are mostly consistent among different sample sizes and different surrogate models, however, there are often visible numerical issues for small sample sizes such as: negative main indices or conversely main indices slightly higher than total indices for parameters with minor or no significance, or sum of main indices more than one. This can be due to insufficient samples to accurately calculate the integrals defining the terms in the Sobol' index calculations and/or surrogate inaccuracies. We note that the choice of sampling method is of paramount importance to the resulting accuracy of both surrogate models and values of Sobol' indices.

In summary, these four cases provided a realistic set of data to study the differences in sensitivity analysis methods and implementations. We plan to continue this work with even more challenging case studies of geologic repositories, with a focus on time-dependent, highly nonlinear and/or non-monotonic behavior and inclusion of features such as spatial heterogeneity of fracture fields.

[1]: Laura P. Swiler, Dirk-Alexander Becker, Dusty Brooks, Joan Govaerts, Lasse Koskinen, Pekka Kupiainen, Elmar Plischke, Klaus-Jürgen Röhlig, Elena Saveleva, Sabine M. Spiessl, Emily Stein, Valentina Svitelman. "Sensitivity Analysis Comparisons on Geologic Case Studies: An International Collaboration." In preparation.

Acknowledgments

This sensitivity analysis group is working under the auspices of Organization for Economic Cooperation and Development (OECD)/ NEA's Integration Group for the Safety Case (IGSC).
The authors acknowledge feedback from Sergei Kucherenko with regards to metamodeling techniques and sensitivity analysis. The authors acknowledge other funding agencies, including the German Federal Ministry for Economic Affairs and Energy (BMWi).

This abstract is SAND2021-10708C. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Speaker: Helen Moore

Applying sensitivity analysis for the evaluation of quantitative systems pharmacology (QSP) models

*Sietse Braakman**
Application Engineering Department, MathWorks Inc, USA, sietsebraakman@gmail.com

Pras Pathmanathan
Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, US Food and Drug Administration, USA, Prasanna.Pathmanathan@fda.hhs.gov

Helen Moore
Laboratory for Systems Medicine, University of Florida, USA, dr.helen.moore@gmail.com

Successful drug research and development (R&D) requires tremendous time and financial resources. Achieving one approved drug in the United States (US) can take over a decade [1] and can cost more than $1.3 billion [2]. Mechanistic mathematical models can substantially reduce the time and cost of drug development. These models are known as quantitative systems pharmacology (QSP) models in the field of pharmacometrics. An early example of a QSP model for type 2 diabetes reduced an estimated 40% of the time and 66% of the cost of a Phase 1 trial [3]. QSP models also have the potential to substantially improve efficacy and safety [4]. QSP and other mechanistic systems models are thus increasingly used to support decisions in R&D, including regulatory decisions.

Along with this increase in use, assessing the predictive capability of such models is increasingly important. Several frameworks have been proposed for the development of QSP models, but few have focused on prediction assessment. In our work [5], we introduce a framework that focuses on model evaluation, which extends and builds on existing frameworks for model development. We provide recommendations for how and when various evaluation methods should be applied, including consideration of the underlying assumptions of specific analysis methods and their appropriate applications. These include sensitivity and identifiability analyses, as well as validation and uncertainty quantification. Many of these methods have been used successfully in other fields.

In this presentation, we will focus on how these methods are applied specifically to QSP model evaluation. For example, sensitivity analysis can be valuable for the planning and prioritizing of in vitro and in vivo experiments to obtain better estimates of influential parameters, in order to improve confidence in QSP model predictions. Sensitivity analysis also provides a method for reducing the number of free parameters in a model. Noninfluential parameters can be fixed (or “frozen”) to nominal values without substantially impacting model output predictions. This reduction in the number of free parameters enables more-thorough exploration of the subspace of influential parameters, providing greater confidence in model output predictions, and can provide a basis for QSP model reduction [6]. The relative influence of input parameters can also be used to prioritize potential therapeutic targets in drug discovery [7]. We also share examples in the pharmacometrics setting which demonstrate the misleading results that can be obtained when inappropriate analyses are applied.

*Current affiliation: Quantitative Translational Pharmacology, AbbVie Inc, USA.

[1] DiMasi JA, Hansen RW, Grabowski HG (2003) The price of innovation: new estimates of drug development costs. J. Health Econ. 22(2):151-85. https://doi.org/10.1016/s0167-6296(02)00126-1

[2] DiMasi JA, Grabowski HG, Hansen RW (2016) Innovation in the pharmaceutical industry: new estimates of R&D costs. J. Health Econ. 47: 20–33. https://doi.org/10.1016/j.jhealeco.2016.01.012

[3] Kansal A, Trimmer J (2005) Application of predictive biosimulation within pharmaceutical clinical development: examples of significance for translational medicine and clinical trial design. IEE Proc.-Syst. Biol. 152:214–220. https://doi.org/10.1049/ip-syb:20050043

[4] van der Graaf PH, Benson N (2018) The role of quantitative systems pharmacology in the design of first-in-human trials. Clin. Pharmacol. Ther. 104(5):797. https://doi.org/10.1002/cpt.1145

[5] Braakman S, Pathmanathan P, Moore H (2021) Evaluation Framework for QSP Models. Submitted.

[6] Hasegawa C, Duffull SB (2018) Reusing smaller versions of large models: A case example of reuse of a simplified bone model. Clin. Pharmacol. Ther. 106(6):1184-1186. https://doi.org/10.1002/cpt.1582

[7] Schoeberl B, Pace EA, Fitzgerald JB, Harms BD, Xu L, Nie L, Linggi B, Kalra A, Paragas V, Bukhalid R, Grantcharova V, Kohli N, West KA, Leszczyniecka MA, Feldhaus MJ, Kudla AJ, Nielsen UB (2009) Therapeutically Targeting ErbB3: A Key Node in Ligand-Induced Activation of the ErbB Receptor–PI3K Axis. Science Signalling 2(77) ra31 https://doi.org/10.1126/scisignal.2000352

Speaker: Sabine M. Spiessl or Dirk-A. Becker

Performance of the BSPCE Metamodeling Approach using Random and Quasi-Monto Carlo Sampling on the basis of a Time-Dependent Final Repository Model

Sabine Maria Spiessl, Dirk-Alexander Becker

Repository Safety Department, GRS, Germany, Sabine.Spiessl@grs.de, Dirk-Alexander.Becker@grs.de

Sergei Kucherenko

Imperial College London, London, UK, S.Kucherenko@imperial.ac.uk

Due to their highly nonlinear, non-monotonic or even discontinuous behaviour, final repository models may contain many parameter interactions associated with and/or triggered by the various uncertainties. Accordingly, by analysing parameter interactions, uncertainties of these models may also be better understood and may be reduced.

Polynomial Chaos Expansion (PCE) is a widely used approach in metamodeling [1]. Usually only a few terms are relevant in the PCE structure. The Bayesian Sparse PCE method (BSPCE) makes use of sparse PCE. Selection of the proposed PCE structure is based on a Bayesian approach using the Kashyap information criterion for model selection [2]. It allows efficient estimation of Sobol' sensitivity indices of the first and total orders (SI1 and SIT). We recall that SI1 represents the specific influence (or main effect) of one parameter on the model and SIT embodies the total effect of one parameter on the output in interactions of any order with all other parameters.

In this work, we tested the performance of the BSPCE method on the basis of a time-dependent model for contaminant release from a final repository in a clay formation, using random and quasi-Monto Carlo sampling. Although this model has a smooth behaviour, its analysis is still quite challenging. We analysed the time-dependent output with two approaches for sensitivity analysis, namely the pointwise and generalized approaches [3]. With the pointwise approach, the output at each time step is analysed independently. The generalized approach considers averaged output contributions at all previous time steps in the analysis of the current step. As a quasi-Monto Carlo sampling method, we utilised the LpTau sampling more commonly known as sampling using Sobol' sequences. We evaluated the model output with samples of different sizes, drawn with both sampling methods.

We compared the results for SI1 obtained with the EASI method [4]. EASI is a Fourier-based technique for the computation of the SI1 indices.

Attained SIT results were compared with the Random-Sampling High Dimensional Model Representation (RS-HDMR) metamodeling approach. In this approach, the ANOVA-HDMR expansions is truncated after the second order. The truncated terms are then approximated by orthonormal polynomials (Zuniga et al [5]). Consequently, by design, for a specific parameter, SIT in this method is a sum of SI1 plus all corresponding SI2's. RS-HDMR also belongs to the group of PCE methods. In this work we used the SobolGSA software [6,7] which contains both the RS-HDMR and BSPCE methods.

While the SI1 results of BSPCE are in good agreement with those from EASI and also with the values calculated by RS-HDMR, the SIT results differ between BSPCE and RS-HDMR. Although the SIT importance ranking of the different parameters obtained by the two metamodeling methods is the same, the SIT curves differ in shape and magnitude. This might, at least in part, be due to the different assumptions and behaviour of the numerical approaches, but it can also be a hint to the existence of relevant higher-order interactions since RS-HDMR neglects all orders higher than 2.

Previous results showed that for most direct methods as well as for the RS-HDMR metamodeling method, quasi-random LpTau sampling requires fewer runs to provide stable results than random sampling. This, however, does not seem to be the case for BSPCE. At least from the results obtained for the considered case one cannot deduce a significant superiority of LpTau above random sampling, neither for the pointwise nor for the generalised approach.

In contrast to the pointwise approach, the generalised approach takes account of the total history of sensitivity, so that typically the sensitivity curves reach nearly stable values at some time. At early times, the curves proceed similarly as those of the pointwise approach but reach medium, nearly constant values at the end. At late times, the total output variance is small, so that the relative contributions of the different parameters or parameter combinations do no longer play an essential role if evaluated in this way. For the generalised approach, the performance of the BSPCE method is the same as for the pointwise approach.

Acknowledgements

The work was financed by the German Federal Ministry for Economic Affairs and Energy (BMWi).

References

[1] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93: 964–979, 2008.
[2] Q. Shao, A. Younes, M. Fahs, T-A. Mara. Bayesian sparse polynomial chaos expansion for global sensitivity analysis. *Computer Methods in Applied Mechanics and Engineering*, 318: 474-496, 2017.
[3] A. Alexanderian, P.A. Gremaud, R.C. Smith. Variance-based sensitivity analysis for time-dependent processes. https://arxiv.org/pdf/1711.08030.pdf, 2019.
[4] E. Plischke. An effective algorithm for computing global sensitivity indices (EASI). *Reliability Engineering and System Safety*, 95: 354–360, 2010.
[5] M.M. Zuniga, S. Kucherenko, N. Shah. Metamodelling with independent and dependent inputs. *Computer Physics Communications*, 184(6): 1570–1580, 2013.
[6] S.M. Spiessl, S. Kucherenko, D.-A. Becker, O. Zaccheus. Higher-order sensitivity analysis of a final repository model with discontinuous behaviour. *Reliability Engineering and System Safety*, 187: 149-158, doi: https://doi.org/10.1016/j.ress.2018.12.004, 2019.
[7] SobolGSA software. *User manual*, https://www.imperial.ac.uk/process-systems-engineering/research/free-software/sobolgsa-software/, 2021.

Probabilistic Sensitivity via Optimal Transport

Emanuele Borgonovo, Bocconi University, Italy, emanuele.borgonovo@unibocconi.it
Alessio Figalli, ETH Zürich, Switzerland, alessio.figalli@math.ethz.ch
Elmar Plischke, Clausthal University, Germany, elmar.plischke@tu-clausthal.de
Giuseppe Savaré, Bocconi University, Italy, giuseppe.savare@unibocconi.i

May 28, 2021

Abstract

The pervasive use of scientific simulators requires approaches for providing transparent accounts of uncertainties in modeling assumptions and input parameters [5]. Such methods are collected under the name of probabilistic sensitivity analysis [4].

Methods based on linear regression and correlations are in wide use. However, they do not provide a fit-for-all-purposes solution. Variance-based methods extend the notion of a non-linear goodness-of-fit measure, providing information on the functional dependence on input assumptions. Distribution-based methods drive this idea further by measuring the distance between the present state of knowledge and the state of knowledge after information about one or more uncertain inputs has been received. As underlined in [3], the majority of state-of-the-art methods have been developed for univariate output responses and probabilistic sensitivity analysis becomes challenging when the output is a multivariate random vector, spatially or temporally distributed.

We investigate the use of the theory of optimal transport (OT) [1, 6] to address the probabilistic sensitivity analysis of multivariate output problems. We discuss theoretical aspects first. We show that, in the multivariate elliptical case, a closed form expression for OT-based sensitivity measures is found via the Wasserstein-Bures metric. The resulting sensitivity measure subsumes the variance-based sensitivity indices introduced by [2]. We address numerical quantification next and cast the estimation in a one-sample (or given data) context. We provide the explicit data driven formulation for the OT problem both in the Kantorovich and in the entropic versions. We show that given-data estimators are consistent, provided that the OT solution algorithm is consistent. From an algorithmic viewpoint, the estimation of OT-based global sensitivity measures requires the sequential solution of data-driven OT problems. This might be a computationally challenging task: We compare OT solvers based on partial reorderings, on the dual-simplex method, and on the Sinkhorn iterative approximation. We carry out a series of numerical experiments for analytical test cases and apply the findings to a well known simulator with a large output dimensionality. Overall, in the context of computer experiments, the size and sequential nature of the calculations pose a challenge to current algorithmic implementations, leading them at the edge of the most recent research efforts in the OT-solution stream. At the same time, the use of the Wasserstein-Bures metric offers an elegant as well as

numerically efficient new approach to the probabilistic sensitivity analysis of multivariate responses.

References

[1] A. Figalli. *The Monge-Ampere Equation and Its Applications.* European Mathematical Society, 2017.

[2] Fabrice Gamboa, Alexandre Janon, Thierry Klein, and Agnès Lagnoux. Sensitivity analysis for multidimensional and functional outputs. *Electronic Journal of Statistics*, 8:573–603, 2014.

[3] A Marrel, N Saint-Geours, and M De Lozzo. Sensitivity analysis of spatial and/or temporal phenomena. In *Handbook of Uncertainty Quantification*, pages 1327–1357. 2017.

[4] J Oakley and A O'Hagan. Probabilistic Sensitivity Analysis of Complex Models: a Bayesian Approach. *Journal of the Royal Statistical Society, Series B*, 66(3):751–769, 2004.

[5] A Saltelli, G Bammer, I Bruno, E Charters, M Di Fiore, E Didier, W N Espeland, J Kay, S Lo Piano, D May, R Jr. Pielke, T Portaluri, T M Porter, A Puy, I Rafols, J R Ravetz, E Reinert, D Sarewitz, P B Start, A Stirling, J P van der Sluijs, and P Vineis. Five Ways to Ensure that Models Serve Society: a Manifesto. *Nature*, 582:482–484, 2020.

[6] Cédric Villani. *Optimal Transport: Old and New.* Springer Verlag, Berlin, 2009.

Speaker: to be defined

A novel boosting algorithm for regression problems (bOOst^d) and its use for sensitivity analysis

Ivano Azzini
European Commission, Joint Research Centre (JRC), Ispra, Italy,
ivano.azzini@ec.europa.eu
Thierry Mara
European Commission, Joint Research Centre (JRC), Ispra, Italy,
thierry.mara@ec.europa.eu
Rossana Rosati
European Commission, Joint Research Centre (JRC), Ispra, Italy,
rossana.rosati@ec.europa.eu

In this work, we present a novel boosting algorithm for regression problem (called bOOst^d) and its use as meta-model for performing sensitivity analysis of computer models.

Boosting tries to build an accurate (or *strong*) predictor by combining several *weak* predictors from a given learning dataset. It was introduced in [1, 2], but it made a significant leap forward with the introduction of Adaboost for binary classification problems [3]. In [3], Adaboost was also generalized to handle multi-class (called Adaboost.M1 and M2) and regression problems (Adaboost.R). Subsequently, boosting and Adaboost have been studied intensively. However, there are still some open questions related to the Adaboost generalization error and how to extend it in a simple and optimal way in the continuum.

Our bOOst^d algorithm tries to overcome some limitations related with the use of boosting methodology in regression setting (it is very similar to Adaboost, very simple to implement and it is able to manage hypotheses with error $> \frac{1}{2}$). A full description of this algorithm in machine learning framework is in available upon request to the author of [4] (its theoretical proprieties, implementation, and first experimental results).

Some theoretical and practical characteristics of bOOst^d suggest to use it as a meta-model in sensitivity analysis using different weak learners and different estimators for Sobol' indices. In particular we explore two strategies:
1. bOOst^d with low-order PCE (Polynomial Chaos Expansion [5]) as weak learner,
2. we use in bOOst^d neural networks (or others equivalent learners) as a weak learner.

The metamodels so obtained are then coupled with the Monte Carlo estimator recently introduced in [6, 7] to compute the Sobol' indices. Comparison of the two strategies will be carried out by application on well-known GSA benchmark functions.

[1] Y. Freund, Boosting a weak learning algorithm by majority, Inform. and Comput. 121, No. 2 (September 1995), 256-285; an extended abstract appeared in ''*Proceedings of the Third Annual Workshop on Computational Learning Theory*, 1990.''

[2] R. E. Schapire, The strength of weak learnability, *Machine Learning* 5, No. 2, 197-227, 1990.

[3] Freund, Y., & Schapire, R. E. A decision-theoretic generalization of on-line learning and an

application to boosting. *Journal of computer and system sciences*, 55(1), 119-139, 1997.

[4] I. Azzini. The adaptive reductions in game theory and there applications to boosting. *Work in progress,* 2021.

[5] Shao, Q., Younes, A., Fahs, M., & Mara, T. A. Bayesian sparse polynomial chaos expansion for global sensitivity analysis. *Computer Methods in Applied Mechanics and Engineering,* 318, 474-496, 2017.

[6] Azzini, I., & Rosati, R. *Sobol'main effect index: an Innovative Algorithm (IA) using Dynamic Adaptive Variances*. Reliability Engineering & System Safety, 213, 107647, 2021.

[7] Azzini, I., Mara, T. A., & Rosati, R. *Comparison of two sets of Monte Carlo estimators of Sobol'indices*. Environmental Modelling & Software, 105167, 2021.

A case study of global sensitivity analysis using deep learning on industrially relevant large datasets – Attention to distribtion of effects in DGSM methods

Adem Rosenkvist Nielsen Aouichaoui[a], Gürkan Sin[a*]

[a] Process and Systems Engineering Center (PROSYS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Kgs. Lyngby, DK-2800, Denmark

[*]Corresponding Author's E-mail: gsi@kt.dtu.dk

Keywords: Deep learning, global sensitivity analysis, N_2O emissions, Ranking of parameter importance

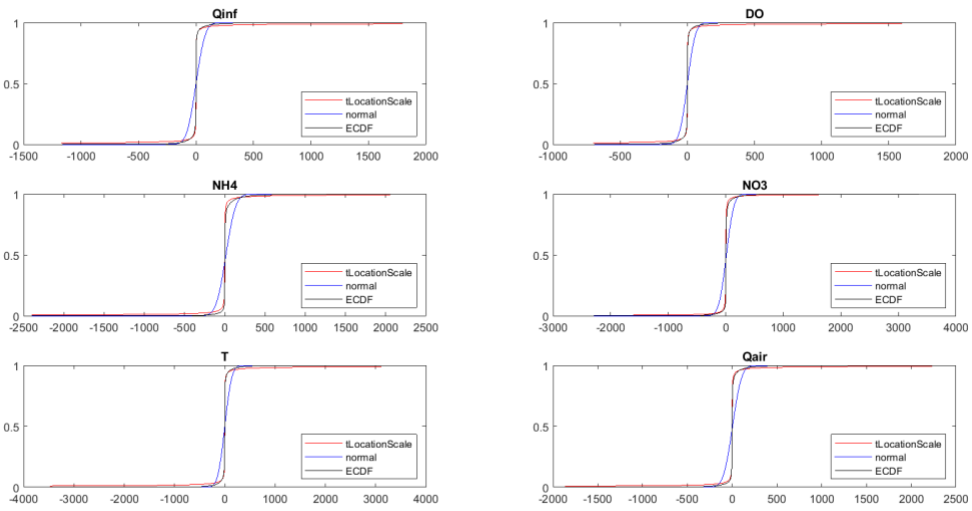
ABSTRACT

Global sensitivity analysis (GSA) is now playing an increasingly prominent role in the development and assessment of complex simulation models. Among its other uses, GSA is widely employed to generate insights into the contributions of individual model inputs, or sub-groups of inputs, to the variations in the output of a mechanistic model (Saltelli et al., 2019). However, when such models fall short of explaining a particular process phenomenon, sensitivity indices derived from these models become unreliable. One good example is nitrous-oxide emissions (N2O) from wastewater treatment plants (WWTP), for which the mechanistic understanding is still in its infancy(Daelman et al., 2015;Sin and Al., 2021). The recent breakthroughs achieved in deep learning (DL), on the other hand, offer an exciting possibility to bring new light to such poorly understood process phenomena. To this end, here we present a new software tool, named deepGSA, incorporating well-established variance-decomposition-based global sensitivity analysis methods, such as Sobol sensitivity indices, with the plant data-driven deep learning modeling techniques.

The deepGSA aims at enabling non-specialist practitioners to leverage DL-based models for GSA application purposes. To this end, the tool builds on an earlier GSA framework of the authors, easyGSA (Al et al., 2019), and is based on a recently proposed framework for DL-based and big data-driven process modeling (Hwangbo et al., 2020). By using these two frameworks, the deepGSA streamlines a number of tasks into a deep learning pipeline, such as data cleaning and preparation, model building and discrimination, model validation, Monte Carlo simulations, Sobol sensitivity analysis, Derivate-based global sensitivity analysis and effective visualizations of GSA results. The capabilities of the tool are highlighted with a case study from WWTPs concerning study of nitrogendioxide (N2O) emissions which is a potent greenhouse gas. For that purpose, a one-year-long dataset collected from four of the biological reactors of the Avedøre WWTP of Copenhagen (Denmark) (Chen et al., 2019) was used to train DL-based models. By using the tool, a number of candidate DL network topologies were evaluated, and a DL-model with satisfactory predictive performance (R^2_{test}>0.90) was obtained. By using this model in a parallelized Monte Carlo simulations procedure, the Sobol sensitivity indices were calculated so as to pinpoint underlying factors (process disturbances and conditions) related to the emissions of N2O. Several sensitivity analysis techniques were applied to understand and explain which inputs are driving the greenhouse gas emissions. As regards the variance decomposition methods, the resutls indicate that Sobol total sensitivity index (STi) is an appropriate metrics to compare input/factor importance. Furthermore accounting for dependency among inputs, only influences the main effects (Si) calculations and not the STi. Hence STi is recommended a robust measure for relative importance of the inputs for this particular study. One disadvantage of variance based decomposition methods is that they do not indicate the direction of the effects of the inputs on the outputs (namely positive or negative influence) as Si and STi both are a ratio of the conditional variance of input on the output over total variance of outputs.

Derivative based sensitivity analysis also known as local elementary effects, in reference to Morris elementary effects, can reveal the sign (negative versus positive) of the inputs on the model outputs

which provides a valuable information in engineering/scientific studies. When performed in a global setting (e.g. Sobol, I. M., & Kucherenko, S. (2010)), properties of the distribution of these effects can be used to study and rank importance of factors. In this study, the distribution of the values of the said derivative functions revealed a pattern characterized by signifcantly heavy/fat tails in both negative and positive scales of function values (see the figures below). Evaluating and benchmarking a number of distribution functions from probability for heavy tails, *t* Location-Scale distribution was found closest to describe these fat tails. The normal distribution failed to described the tails naturally and this presents an interesting challenge to the theoretical studies generalizing the sensitivity measures based on DGSM. We remark that we have used 100,000 samples (since DL was very fast computationally) for quasi monte carlo sampling for computing the sensitivity functions.



From interpretation point of view, the results indicate that all of the inputs can have a positive or negative contribution depending on the values of other inputs. The green house gas emissions is clearly a nonlinear phenomena. Looking at the sensitivity measures, the mean values of DGSM and Sobol total index STi were in agreement with each other and revealed that T (temperature), nitrate (NO3) and NH4 (ammonium) and Qair (aeration rate) are ranked as important factors. These important rankign was confirmed by Sobol total indices (STi). On the other hand, interpreting the results using the location parameter of the t Location-Scale distribution (similarly the higher the location value, the higher the importance of the factor), it revealed that nitrate, ammonia and Qair is by far the most important inputs that explains the extreme/rare events in the outputs. While temperature and DO and Qinf has location parameter as almost zero indicating that these inputs do not have much explanation to the extreme values of the outputs. These results poses an interesting question namely when the shape of the distribution of the effects (the derivative of the fucntion wrt inputs) is not normal but rather have long tails, how should the factors importance be analyzed? Pragmatically speaking the proposed sensitivity measures in the literature and the new interpretation proposed here using t Location-Scale distribution are complementary to each other. One measures looks at effects that can be explained by a normal distribtion, while the latter focuses on explaining the tails. On another note the deepGSA tool faciliates application of senstivity analysis to industrially relevant large datasets where modeling is not necessarily available.

References

Al, R., Behera, C.R., Zubov, A., Gernaey, K. V., Sin, G., 2019. Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants – An application to the BSM2 model. Comput. Chem. Eng. 127, 233–246. https://doi.org/10.1016/j.compchemeng.2019.05.015

Chen, X., Mielczarek, A.T., Habicht, K., Andersen, M.H., Thornberg, D., Sin, G., 2019. Assessment of Full-Scale N 2 O Emission Characteristics and Testing of Control Concepts in an Activated Sludge Wastewater Treatment Plant with Alternating Aerobic and Anoxic Phases. Environ. Sci.

Technol. 53, 12485–12494. https://doi.org/10.1021/acs.est.9b04889
Daelman, M.R.J., van Voorthuizen, E.M., van Dongen, U.G.J.M., Volcke, E.I.P., van Loosdrecht, M.C.M., 2015. Seasonal and diurnal variability of N 2 O emissions from a full-scale municipal wastewater treatment plant. Sci. Total Environ. 536, 1–11. https://doi.org/10.1016/j.scitotenv.2015.06.122
Hwangbo, S., Al, R., Sin, G., 2020. An integrated framework for plant data-driven process modeling using deep-learning with Monte-Carlo simulations. Comput. Chem. Eng. 143, 107071. https://doi.org/10.1016/j.compchemeng.2020.107071
Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environ. Model. Softw. 114, 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012
Sin, G., & Al, R. (2021). Activated sludge models at the crossroad of artificial intelligence—A perspective on advancing process modeling. *npj Clean Water*, *4*(1), 1-7.

Speaker: Manu Aggarwal

Sobol' indices combined with γ-indices for a more comprehensive sensitivity analysis

Manu Aggarwal
NIDDK, National Institutes of Health, USA, manu.aggarwal@nih.gov
Nick Cogan
Department of Mathematics, Florida State University, USA, cogan@math.fsu.edu

Physical, chemical, and biological processes are often viewed as closed systems consisting of controlled inputs and observed outputs. These systems can be mathematically analyzed by formulating them as a set of equations that define supposed relations between the inputs and the outputs using constants that are called parameters. We consider the scenario in which we have observed data from experimental trials of such a system. Then, a theoretically justifiable model that is validated using observed data, can be used as evidence for underlying model hypotheses and can simulate predictions for practically feasible situations that were not accounted for in the experiment setup. Then, it is to be expected that the model behavior will depend on the choice of values of at least some of the model parameters. The values of the parameters might be either measured experimentally or estimated by minimizing error between model simulations and observed data. However, both cases often have various sources of uncertainties. In the former case, there might be a lack of precision and accuracy in the measurement techniques. The latter case can be further split as follows—the model cannot account for all of the variability in the data due to a lack of complete knowledge, i.e. the degree of freedom of the model is less than that of the data; the observed data might be insufficient, i.e. the degree of freedom of the model is more than that of the data; the algebraic structure of the equations might impose impossibility of a unique solution for the parameters regardless of the nature of the observed data; and the observed data is noisy due to measurement noise. As a result of all these cases, the computed set of parameter values that match the model with the observed data can be a combination of noisy manifold(s) and noisy point(s) in the parameter space. In our work, we use the approach of viewing this noise or uncertainty in the parameter estimates as statistical distributions on parameters. Then there are two generals aims, to reduce the variance in distributions of parameter estimates, where possible, and to explain the effects of variance in the parameters on the model output. Both of these aims can be approached by analyzing the effects of variation of parameters on the output, also called sensitivity analysis. However, studying the dependence of higher moments of conditional distributions of the output on variations in parameters is not included in most senstivity analyses, possibly due to a high computational cost of estimating higher moments. In this work we claim that analysis of higher moments can possibly reveal useful information for both of the general aims, define measures to quantify this analysis, and prove a mathematical relation that can make it computationally feasible.

Sobol' indices (SIs)[1] are well-known sensitivity measures that are applicable to non-linear and non-monotonic models. Moreover, variations in all of the parameters, in a given parameter regime, are considered simultaneously. The SIs are usually reported as *total* SIs, which is further split into *first-order* and *additional* SIs. Total SIs are often used to reduce dimensionality of the parameter space and first-order SIs are used to indicate parameters that should be measured with certainty if a robust QoI is desired. These approaches can possibly reduce uncertainty in parameter estimates. However, an analysis of high additional SIs is often ignored. In this work we proved that additional SIs are related to all of the higher moments of conditional distributions of the QoI. We show that this relation can lead to inferences that make it computationally feasible to analyze higher moments of conditional distributions. We quantify this analysis by defining γ-measures pertaining to studying the dependence of expected value, variance, skewness, and kurtosis on variation in parameters. Each of these statistics is linked to a geometrical feature of the shape of the underlying statistical distribution, which can in turn have practical implications. We summarize by providing a more comprehensive and computationally feasible framework of sensitivity analysis using both Sobol' indices and γ-indices, that can provide deeper insights into the relationships between model output and variance in parameters. As an illustration, we analyze the popular non-linear non-monotonic Ishigami function using Sobol' indices and γ-indices.

[1] Sobol', Ilya M Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1 (4): 407–414, 1993.

Speaker: Clément Duhamel

A SUR adaptation of Bichon criterion for inversion

A1 Clément Duhamel
Applied Mathematics, Université Grenoble Alpes, INRIA, IFP Energies Nouvelles, France, clement.duhamel@inria.fr
A2 Céline Helbert
Univ Lyon, Ecole Centrale de Lyon, CNRS UMR 5208, Institut Camille Jordan, Ecully, France, celine.helbert@ec-lyon.fr
A3 Miguel Munoz Zuniga
Applied Mathematics Department, IFP Energies Nouvelles, France, miguel.munoz-zuniga@ifpen.fr
A4 Clémentine Prieur
Applied Mathematics, Université Grenoble Alpes, France, clementine.prieur@univ-grenoble-alpes.fr
A5 Delphine Sinoquet
Applied Mathematics Department, IFP Energies Nouvelles, France, delphine.sinoquet@ifpen.fr

Nowadays, many inversion issues are present in industry. These problems aim to find all sets of parameters such that a certain quantity of interest respects a given constraint, for example remains below a threshold. In the field of floating wind for instance, a pre-calibration step consists in estimating model parameters that fit with a given accuracy the measured data (e.g. accelerations).

An effective way to solve this problem is to use Gaussian process meta-models (Kriging) with a sequential Design of Experiment (DoE) and an inversion-adapted enrichment criterion, such as the popular Bichon (also known as Expected Feasibility Function [3]) or deviation number [5] (denoted U) criteria. It is also possible to use the more elaborate class of SUR (Stepwise Uncertainty Reduction) criteria [2]. In addition to taking into account the evaluation points and the available simulations, they quantify the uncertainty reduction which can be achieved by the addition of the new point. The goal of this work is to present a SUR version of the Bichon criterion, to find an explicit formulation and to compare its performances to some state-of-the-art criteria.

As a reminder, noting $\mathbb{X}$ the design space (compact), $\mathbf{x}_{n+1}$ the new added point to the DoE and T the fixed threshold in the inversion, the Bichon criterion is defined as:

$$\mathbf{x}_{n+1} := \underset{\mathbf{x}\in\mathbb{X}}{\operatorname{argmax}} \operatorname{EFF}(\mathbf{x}) \quad \text{with} \quad \operatorname{EFF}(\mathbf{x}) := \mathbb{E}\Big[\big(\alpha\sigma_n(\mathbf{x}) - |T - \xi(\mathbf{x})|\big)^+ \,\Big|\, \mathscr{F}_n\Big]. \tag{1}$$

ξ represents the Gaussian process meta-model of the true simulation, $\mathscr{F}_n$ the σ-algebra generated by the n first observations, σ_n the kriging standard deviation and α a fixed

positive number. Intuitively, the EFF($\mathbf{x}$) quantity is interpreted as the minimal distance of the Gaussian meta-model at point $\mathbf{x}$ with respectively the quantities $T - \alpha\sigma_n(\mathbf{x})$ and $T + \alpha\sigma_n(\mathbf{x})$, all multiplied by the indicator of $\xi(\mathbf{x}) \in [T - \alpha\sigma_n(\mathbf{x}), T + \alpha\sigma_n(\mathbf{x})]$.

From a theoretical point of view, the proposed SUR Bichon strategy is defined from a measure of uncertainty related to the Bichon criterion (integral of the Bichon criterion on the design space) as follows: $\mathbf{x}_{n+1} \in \arg\min_{\mathbf{x}\in\mathbb{X}} \mathscr{J}_n(\mathbf{x})$ with

$$\mathscr{J}_n(\mathbf{x}) := \mathbb{E}\Bigg[\underbrace{\int_{\mathbb{X}} \mathbb{E}\Big[\big(\alpha\sigma_{n+1}(\mathbf{z}) - |T - \xi(\mathbf{z})|\big)^+ \,\Big|\, \mathscr{F}_{n+1} \Big] \,\mathrm{d}\mathbb{P}_{\mathbb{X}}(\mathbf{z})}_{H_{n+1}} \,\Bigg|\, \mathbf{X}_{n+1} = \mathbf{x}, \mathscr{F}_n \Bigg]. \tag{2}$$

$\mathbb{P}_{\mathbb{X}}$ is a finite measure given on $\mathbb{X}$ (e.g. Lesbesgue measure), H_{n+1} a $\mathscr{F}_{n+1}$-measurable uncertainty measure, $\mathbf{X}_{n+1}$ the random variable corresponding to the $n+1^{\text{th}}$ evaluation of the DoE and $\mathscr{F}_{n+1}$ the σ-algebra generated by $\mathbf{X}_{n+1}$, $\xi(\mathbf{X}_{n+1})$ and $\mathscr{F}_n$. In addition, by successively using Fubini's theorem and tower property of conditional expectation, it is possible to obtain a simplified formulation of the SUR Bichon criterion dependent on $m_n(\mathbf{y})$, $\sigma_n(\mathbf{y})$, T, α and $\sigma_{n+1}(\mathbf{y})$, allowing an efficient implementation.

From a numerical point of view, performances of the SUR Bichon criterion are compared to other classic criteria, on common test functions. One of the main measures of comparison of the criteria is defined as the volume according to $\mathbb{P}_{\mathbb{X}}$ of the symmetric difference between the estimator and the true excursion set. According to this measure of comparison, the SUR Bichon criterion shows encouraging results compared to some other conventional DoE enrichment criteria, like the SUR Vorob'ev criterion [4].

The future prospects for this work are adapting this criterion to more complex data like functional uncertain input variables [1]. In this particular framework, the DoE will have to be adapted.

[1] Reda Amri, Céline Helbert, Miguel Munoz Zuniga, Clémentine Prieur, and Delphine Sinoquet. Set inversion under functional uncertainties with joint meta-models. 2020.

[2] Julien Bect, David Ginsbourger, Ling Li, Victor Picheny, and Emmanuel Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3):773–793, 2012.

[3] Barron J Bichon, Michael S Eldred, Laura Painton Swiler, Sandaran Mahadevan, and John M McFarland. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA journal*, 46(10):2459–2468, 2008.

[4] Clément Chevalier. *Fast uncertainty reduction strategies relying on Gaussian process models*. PhD thesis, Universität Bern, 2013.

[5] Benjamin Echard, Nicolas Gayton, and Maurice Lemaire. Ak-mcs: an active learning reliability method combining kriging and monte carlo simulation. *Structural Safety*, 33(2):145–154, 2011.

ABSTRACTS FOR POSTERS

Speaker: Samih Akkari

Identifiability-based Parameters Selection and Number for Calibration

Samih Akkari
Energy and processes, MINES ParisTech, samih.akkari@mines-paristech.fr
Patrick Schalbart
Energy and processes, MINES ParisTech, patrick.schalbart@mines-paristech.fr
Bruno Peuportier
Energy and processes, MINES ParisTech, bruno.peuportier@mines-paristech.fr

Bayesian calibration has been applied in several case studies in the building sector and are proved reliable in terms of enhancing the accuracy of the building energy models (BEM). The idea is that it benefits from prior distributions that express our knowledge about model parameters and it updates them to posterior distributions given measured data available.

Overparameterization due to the large number of model parameters is a common issue tackled when applying calibration. It means that different combinations of calibration parameters could result in the same model output. Moreover, calibrating a large number of parameters could be computationally intensive to such algorithms. Sensitivity analysis is normally applied to analyse the effect of all parameters on the model's output and to calibrate based on the most influential parameters. However, the selected parameters might still be collectively unidentifiable even if each parameter is confirmed to be influential.

Parameter identifiability is the concept of whether the model parameters can be uniquely inferred from the data. Correlation and interaction that might exist between the most influential parameters selected after a sensitivity analysis could make them unidentifiable. Thus, identifiablity analysis needs to be conducted before launching the calibration. Sensitivity-based identifiablity analysis that relies on the computed sensitivity indices allows to compute the collinearity of different parameter combinations and to re-rank the parameters considering their importance, correlations and interaction. This enables a further analysis concerning the acceptable number of parameters that can be considered for calibration.

The studied building corresponds to the I-BB house (Concrete construction) of the INES (National Institute of Solar Energy) "INCAS" platform, located in Le Bourget-du-Lac in France. The interior surface area is 89 m² with two floors. The house was designed to match the performance of the "PassivHaus" label, thanks to strong insulation, very low thermal bridges, and high-performance glazing. The experimental campaign has been constructed to measure the temperature profile of the building under six different scenarios with separate or combined consideration of different physical phenomena.

The building energy model COMFIE, developed by the centre for energy efficiency of Systems of MINES ParisTech was used to model the thermal behaviour of the building. The model constituted of 130 uncertain parameters on which sensitivity analysis is applied.

RBD-FAST was selected in this paper due to its computational efficiency compared to other variance-based methods. A bootstrapping technique was conducted to analyse the variability of the sensitivity indices and the ranking of the parameters. The effect of the data sample size on the variability in the parameters ranking is also analysed. Fig. plots the variability in ranking the 13 most influential parameters. It shows how the variability increases as the parameter is less

influential. It was also observed that the rank of the 6 most influential parameters stabilises with a data sample size of 2000.

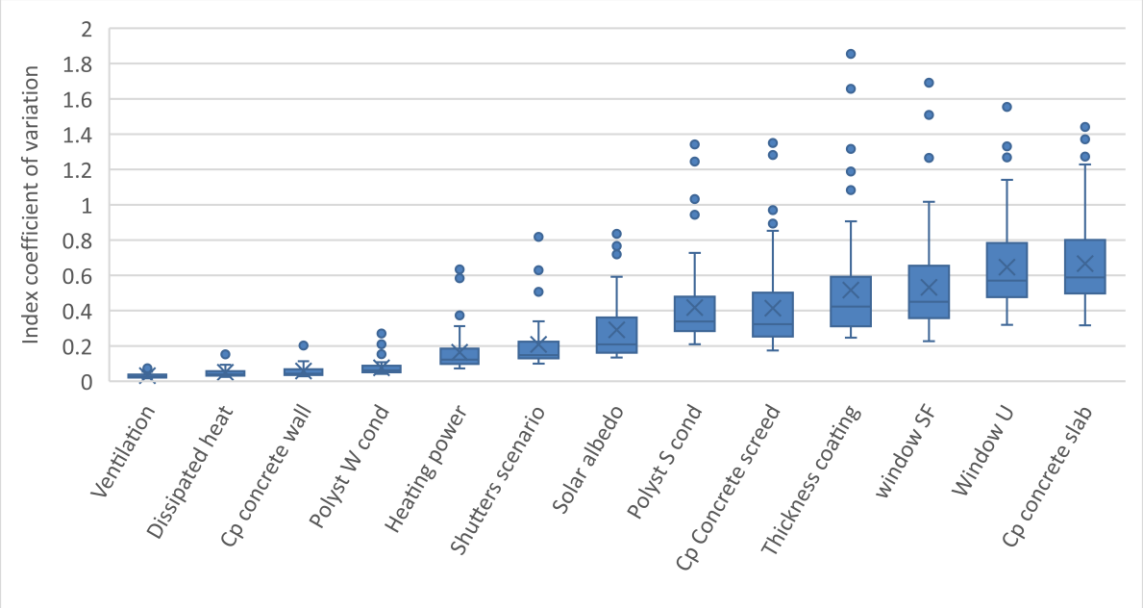


Figure 1: Variability in ranking the 13 most influential parameters

The identifiability analysis requires the selection of a collinearity index (CI) threshold below which the selected parameters are considered collectively identifiable. With a CI threshold taken 20, the maximum number of parameters that are confirmed to be identifiable is 7 as depicted in Fig.. That is that any combination of a larger size comprises of unacceptable degree of interactions and correlation.

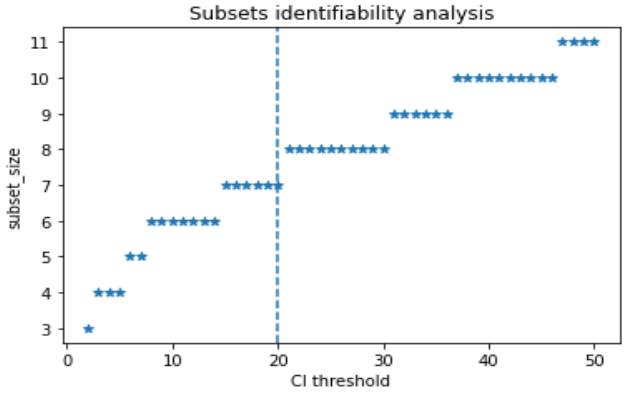


Figure 2: Collinearity index for different possible parameter combinations

The effect of this method is analysed by evaluating the performance of the whole calibration process in terms of the obtained accuracy in fitting the measurements and in well estimating the true values of the model parameters.

A related topic is the maximum number of parameters above which over-parametrisation occurs during calibration. This is related to the identifiability issue since what causes over-parametrisation is the huge interactions and correlations between the parameters. In this paper we investigate in depth how to select the CI threshold, that is that the degree of interactions permitted for undergoing inference without identifiability problems.

Speaker: Maximilian Blesch

Robust decision-making under risk and ambiguity*

Maximilian Blesch
Berlin School of Economics, Germany, maximilian.blesch@hu-berlin.de
Philipp Eisenhauer
University of Bonn, Germany, peisenha@uni-bonn.de

Decision-makers often confront uncertainties when determining their course of action. For example, individuals save to cover uncertain medical expenses in old age [9]. Firms set prices in an uncertain competitive environment [12], and policy-makers face uncertainties about future costs and benefits when voting on climate change mitigation efforts [2]. We consider the situation in which a decision-maker posits a collection of economic models to inform his decision-making process. Each model formalizes the relevant objectives and trade-offs. Within a given model, uncertainty is limited to risk, as a model induces a unique probability distribution over possible future outcomes. In addition, however, there is also ambiguity about the true model [1, 15].

In this context, we focus on the common practice in economics of estimating a subset of the model parameters outside the model and allowing the decision-makers characterized by the model to treat these point estimates as if they correspond to the true parameters. This approach ignores ambiguity about the true model resulting from the parametric uncertainty of the first-step estimation and opens the door to potential misspecification of the decision problem. As-if decision-makers, those who use the point estimates to inform decisions that would be optimal if the estimates were correct [18], face the risk of serious disappointment about their decisions. The performance of as-if decisions often turns out to be very sensitive to misspecification [22], which is particularly consequential in dynamic models where the impact of erroneous decisions accumulates over time [17]. This danger creates the need for robust decision rules that perform well over a whole range of different models instead of an as-if decision rule that performs best for one particular model. However, increasing robustness, often measured by a performance guarantee under a worst-case scenario, decreases performance in all other scenarios. Striking a balance between the two objectives thus poses a significant challenge.

We develop a framework to evaluate as-if and robust decision rules in a decision-theoretic setting by merging insights from the literature on data-driven robust optimization [6] and robust Markov decision processes [4] with statistical decision theory [5]. We set up a stochastic dynamic investment model in which the decision-maker takes ambiguity about the model's transition dynamics directly into account. Using the Kullback-Leibler divergence [16], we construct ambiguity sets for the transitions that are statistically meaningful,

computationally tractable, and anchored in empirical estimates [3]. Our work brings together and extends research in economics and operations to make econometrics useful for decision-making with models [18, 7].

As an applied example, we revisit [20]'s seminal bus replacement problem, which serves as a computational illustration in a variety of settings. In the model, the manager Harold Zurcher implements a maintenance plan for a fleet of buses. He faces uncertainty about the future mileage utilization of the buses. To make his plan, he assumes that the mileage utilization follows an exogenous distribution and uses data on past utilization for its estimation. In the standard as-if analysis, the distribution is estimated in a first step and serves as a plug-in for the true unknown distribution. Harold Zurcher makes decisions as if the estimate is correct and ignores any remaining ambiguity about future mileage utilization. We set up a robust version of the bus replacement problem to directly account for estimation uncertainty and explore the properties and relative performance of alternative decision rules.

In econometrics, there is burgeoning interest in assessing the sensitivity of findings to model or moment misspecification. Our work is most closely related to [13], who develops a measure to assess the sensitivity of results by fixing a subset of model parameters prior to estimating the remaining parameters. Our approach differs as we directly incorporate model ambiguity in the design of the decision-making process inside the model and assess the performance of a decision rule within a misspecified decision environment. As such, our focus on ambiguity faced by decision-makers inside economic models draws inspiration from the research program summarized in [11], which tackles similar concerns with a theoretical focus. We complement recent work by [21], who works in a setting similar to ours but does not use statistical decision theory to determine the optimal robust decision rule. In ongoing work, [8] use statistical decision theory to structure policy decisions in light of uncertainty about counterfactual policy predictions due to the remaining parametric uncertainty after the estimation of a model. Unlike [8], who conducts an ex-post evaluation of alternative policy proposals using decision-theoretic criteria, we perform a proper ex-ante analysis of competing decision rules. We evaluate each rule's performance under all possible parameterizations of the model and directly account for parametric uncertainty in their construction. In operations research, there are only a handful of applied examples in which data-driven robust decision-making is used in a dynamic setting including portfolio allocation [23], elective admission to hospitals [19], scheduling of liver transplantations [14], and the cost-effectiveness of colorectal cancer screening policies [10]. To the best of our knowledge, none of these applications evaluates the performance of robust decisions against the as-if alternative in a decision-theoretic framework.

[1] Kenneth J. Arrow. Alternative approaches to the theory of choice in risk-taking situations. *Econometrica*, 19(4):404–437, 1951.

[2] Michael Barnett, William Brock, and Lars Peter Hansen. Pricing uncertainty induced by climate change. *The Review of Financial Studies*, 33(3):1024–1066, 02 2020.

[3] Aharon Ben-Tal, Dick den Hertog, Anja De Waegenaere, Bertrand Melenberg, and Gijs Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357, 2013.

[4] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirowski. *Robust optimization*. Princeton University Press, Princeton, NJ, 2009.

[5] James O. Berger. *Statistical decision theory and Bayesian analysis*. Springer, New York City, NY, 2010.

[6] Dimitris Bertsimas, Vishal Gupta, and Nathan Kallus. Data-driven robust optimization. *Mathematical Programming*, 167(2):235–292, 2018.

[7] Dimitris Bertsimas and Aurlie Thiele. Robust and data-driven optimization: Modern decision making under uncertainty. In Michael P. Johnson, Bryan Norman, and Nicola Secomandi, editors, *Models, Methods, and Applications for Innovative Decision Making*, pages 95–122. INFORMS, 2006.

[8] Philipp Eisenhauer, Jano Gabler, and Lena Janys. Structural models for policy-making: Coping with parametric uncertainty. *arXiv Working Paper*, 2021.

[9] Eric French and Jae Song. The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy*, 6(2):291–337, 2014.

[10] Joel Goh, Mohsen Bayati, Stefanos A Zenios, Sundeep Singh, and David Moore. Data uncertainty in Markov chains: Application to cost-effectiveness analyses of medical innovations. *Operations Research*, 66(3):697–715, 2018.

[11] Lars Peter Hansen and Thomas J Sargent. *Robustness*. Princeton University Press, Princeton, NJ, 2016.

[12] Cosmin Ilut. Paralyzed by fear: Rigid and discrete pricing under demand uncertainty. *Econometrica*, 88(5):1899–1938, 2020.

[13] Thomas H. Jørgensen. Sensitivity to calibrated paramters. *Review of Economics and Statistics*, forthcoming, 2021.

[14] David L Kaufman, Andrew J Schaefer, and Mark S Roberts. Living-donor liver transplantation timing under ambiguous health state transition probabilities. *SSRN Working Paper*, 2017.

[15] Frank H Knight. *Risk, uncertainty and profit*. Houghton Mifflin Harcourt, Boston, MA, 1921.

[16] Solomon Kullback and Richard A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[17] Shie Mannor, Duncan Simester, Peng Sun, and John N Tsitsiklis. Bias and variance approximation in value function estimates. *Management Science*, 53(2):308–322, 2007.

[18] Charles F. Manski. Econometrics for decision making: Building foundations sketched by Haavelmo and Wald. *Econometrica*, forthcoming, 2021.

[19] Fanwen Meng, Jin Qi, Meilin Zhang, James Ang, Singfat Chu, and Melvyn Sim. A robust optimization model for managing elective admission in a public hospital. *Operations Research*, 63(6):1452–1467, 2015.

[20] John Rust. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, 55(5):999–1033, 1987.

[21] Soroush Saghafian. Ambiguous partially observable Markov decision processes: Structural results and applications. *Journal of Economic Theory*, 178:1–35, 2018.

[22] James E Smith and Robert L Winkler. The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322, 2006.

[23] Steve Zymler, Daniel Kuhn, and Berç Rustem. Worst-case value at risk of nonlinear portfolios. *Management Science*, 59(1):172–188, 2013.

Speaker: John Darges

Extreme Learning Machines for Variance-Based Global Sensitivity Analysis

John Darges
Dept. of Mathematics, North Carolina State University, USA, jedarges@ncsu.edu
Alen Alexanderian
Dept. of Mathematics, North Carolina State University, USA, alexanderian@ncsu.edu
Pierre Gremaud
Dept. of Mathematics, North Carolina State University, USA, gremaud@ncsu.edu

Variance-based global sensitivity analysis (GSA) provides useful measures, Sobol' indices, of how important individual input variables are to the output of a mathematical model. Estimation of Sobol' indices by traditional Monte Carlo methods can be infeasible for computationally intensive models. An appealing approach is to instead use a surrogate whose Sobol' indices can be computed analytically. We propose the use of extreme learning machines (ELMs), with a specific type of activation functions and a novel sparsification approach, for fast GSA. We illustrate the effectiveness of the proposed approach through application to a GSA benchmark model problem and to a system of ordinary differential equations modeling a biochemical reaction network.

Speaker: Vaibhav Dixit

Performant Global Sensitivity Analysis using GlobalSensitivity.jl

Vaibhav Dixit
Julia Computing, vaibhav.dixit@juliacomputing.com
Chris Rackauckas
Julia Computing, chris.rackauckas@juliacomputing.com
Vijay Ivaturi
Pumas-AI, vijay@pumas.ai

With increasing adoption of Global Sensitivity Analysis as part of modelling workflow in various domains with large models such as climate science or quantitative systems pharmacology, there is an urgent need for optimized and scale able GSA implementations. Such optimization can come through both algorithmic improvements, in form of new methods or modifications to existing methods, and through utilization of modern and extensible scientific computing stack. GlobalSensitivity.jl [8] is a generalized GSA package written in the julia [1] programming language which makes it capable of handling varied problems due to composability offered by julia. The built-in support for parallelism allows analysis of large models with significant simulation overhead with ease for domain scientists looking to use GSA. Currently GlobalSensitivity.jl supports the Sobol, Morris, eFAST, Regression based, DGSM, Delta Moment, EASI, Fractional Factorial and RBD-FAST GSA methods. This talk will cover a comprehensive tutorial of running various different GSA methods and analysing their results using visualizations on the Lotka-Volterra differential equation model using SciML's [7] DifferentialEquantions.jl [6] package. Further, there will be a focus on demonstrating use of GSA in Pharmacometrics by analyzing some example PK/PD, PBPK and QsP models. For this purpose the talk will include tutorials [3][4] focused on using Pumas [5] for running GSA and post processing results to derive insights on a PK/PD model of Hepatitis-C Virus (HCV) [2] and a PBPK model for Voriconazole [9].

[1] Jeff Bezanson, Stefan Karpinski, Viral B. Shah, and Alan Edelman. Julia: A fast dynamic language for technical computing. *CoRR*, abs/1209.5145, 2012.

[2] Jeremie Guedj, Caroline Bazzoli, Avidan U Neumann, and France Mentré. Design evaluation and optimization for models of hepatitis c viral dynamics. *Statistics in medicine*, 30(10):1045–1056, 2011.

[3] PumasAI. Global Sensitivity Analysis on HCV model - https://tutorials.pumas.ai/html/pkpd/hcvgsa.html, 2021.

[4] PumasAI. Global Sensitivity Analysis on Voriconazole model - https://tutorials.pumas.ai/html/pbpk/vorigsa.html, 2021.

[5] Chris Rackauckas, Yingbo Ma, Andreas Noack, Vaibhav Dixit, Patrick Kofod Mogensen, Simon Byrne, Shubham Maddhashiya, José Bayoán Santiago Calderón, Joakim Nyberg, Jogarao V.S. Gobburu, and Vijay Ivaturi. Accelerated predictive healthcare analytics with pumas, a high performance pharmaceutical modeling and simulation platform. *bioRxiv*, 2020.

[6] Christopher Rackauckas and Qing Nie. Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software*, 5(1), 2017.

[7] SciML. `https://sciml.ai`, 2019.

[8] SciML. GlobalSensitivity.jl - `https://gsa.sciml.ai/stable/`, 2020.

[9] Nicole R Zane and Dhiren R Thakker. A physiologically based pharmacokinetic model for voriconazole disposition predicts intestinal first-pass metabolism in children. *Clinical pharmacokinetics*, 53(12):1171–1182, 2014.

Speaker: Max Ehre

Global reliability sensitivity analysis of models with dependent inputs using failure samples

Max Ehre
Engineering Risk Analysis Group, Technical University of Munich, Germany,
max.ehre@tum.de
Iason Papaioannou
Engineering Risk Analysis Group, Technical University of Munich, Germany,
iason.papaioannou@tum.de
Daniel Straub
Engineering Risk Analysis Group, Technical University of Munich, Germany,
straub@tum.de

Reliability analysis is concerned with determining the probability of failure of a system subject to uncertainties. Often, knowing how sensitive a failure probability is to each of the system's uncertain inputs is as important as an accurate probability estimate. By definition, failure events are rare events and thus reliability sensitivity is conceptually different from and computationally more demanding than sensitivity analysis of model output (SAMO). Statistically dependent input uncertainties present an additional difficulty in this framework. Such dependencies require discerning between variable interactions produced by the probabilistic model describing the system inputs and the computational model describing the system itself.

In the context of SAMO, [1] proposed a set of variance-based sensitivity measures that discern between the total contribution of an input to the output variance (generated by interactions in both probabilistic and computational model) and independent contributions (generated by interactions in the computational model only). The idea is based on considering d (d is the number of inputs to the system) different isoprobabilistic transformations from the space of dependent inputs to d spaces of corresponding independent d-dimensional random vectors. These transformations are structured such that in each transformation, one of the d variables is independent of all others and another depends on all others (e.g., the Rosenblatt transform or the Nataf transform using a Cholesky decomposition have this property). By performing d variance-based sensitivity analyses on the d sets of transformed independent random vectors, it is possible to discriminate between output variance contributions stemming from the computational and the probabilistic model.

In our work, we extend this idea to reliability sensitivity analysis, i.e., we compute the independent and total contributions of all d inputs to the variance of the indicator function of the rare event. [2] described how to compute variance-based sensitivity indices of the rare event indicator using a set of failure samples only with independent inputs. We draw

on this approach and show that by considering d different isoprobabilistic transformations of a set of failure samples in the original space of dependent input variables to the d corresponding spaces of independent random vectors, we are able to recover the indices of [1] for reliability sensitivity analysis. Our approach facilitates computing these indices with a set of failure samples obtained as the byproduct of a single run of a suitable rare event simulation method such as crude Monte Carlo importance sampling or subset simulation. As opposed to [1], it is not necessary to repeat the sensitivity/reliability analysis d times. We demonstrate the approach on various artificial test functions as well as engineering problems.

[1] Thierry A. Mara and Stefano Tarantola Variance-based sensitivity indices for models with dependent inputs *Reliability Engineering & System Safety* 107: 1–115-121, 2012.

[2] Luyi Li and Iason Papaioannou and Daniel Straub. Global reliability sensitivity estimation based on failure samples. *Structural Safety* 81: 101871, 2019

Speaker: Matthias Van Hove

Types of uncertainty in simulation models: Categorisation for better identification, accounting and assessment

A1 Matthias Van Hove
Department of Architecture and Urban Planning, Ghent University, Belgium, matthias.vanhove@ugent.be
A2 Marc Delghust (PhD)
Department of Architecture and Urban Planning, Ghent University, Belgium, marc.delghust@ugent.be
A3 Jelle Laverge (PhD)
Department of Architecture and Urban Planning, Ghent University, Belgium, jelle.laverge@ugent.be

A simulation model is a rich and complex structure that maps between the input(s) and the output(s) [1]. It aims to replicate the workings and logic of a real system by using physical and/or statistical descriptions of the activities involved [3]. Consequently, no simulation model can be a perfect representation of the system it aims to emulate [25]. All simulation models inevitably contain uncertainty, which should be addressed and quantified as part of the quality assurance process of the model and as part of inferences.

Uncertainty in modelling can be defined as "any departure from the unachievable ideal of complete deterministic knowledge of the system" (Walker *et al.*, 2003). As the systems, being modelled, increase in scale and complexity, it is expected that the uncertainty in the model also increases (Langevin J., 2020). Though, a fair amount of simulation model outputs are expressed as a single value (Cerezo, 2017), which may yield misleading impressions about the certainty of model insights when used for inferences and/or policy making (Langevin J., 2020).

In the literature, several different authors have addressed sources of uncertainty in simulation models in wording and/or schemes (Booth et al., 2012 [6]; Walker et al., 2003 [8]; Coakley et al., 2014 [9]; Oberkampf et al., 2002 [23]), however a general consensus in terms of uncertainty classification and related terminology does not appear to exist (Refsgaard et al., 2007 [28]). A review of 25 existing uncertainty classification schemes ([5]-[29]) highlighted a broad pattern with types of uncertainty being grouped according to where it occurs in the modelling chain: in the model inputs, the simulation model itself or the model outputs.

In *Figure 1*, the different types of uncertainty in simulation modelling are categorised. In *Table 1*, a concise definition is given.

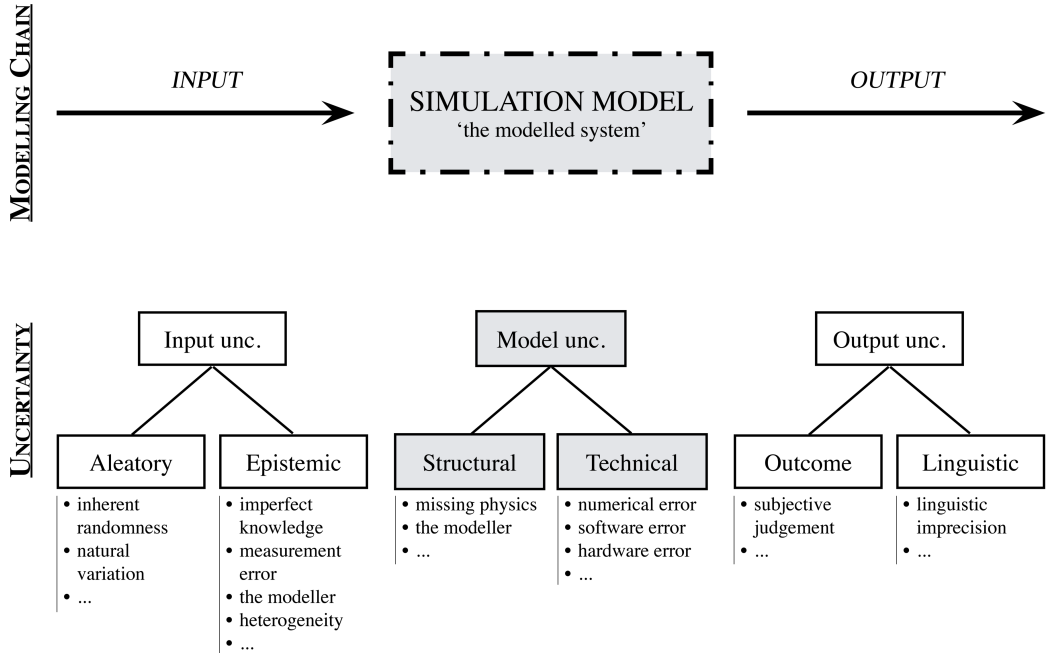


***Figure 1** - Types of uncertainty identified in existing uncertainty classification schemes. Types of uncertainty may be grouped by whether they relate to model inputs, the model itself, or model outputs.*

Definitions

Aleatory uncertainty: Uncertainty due to inherent or natural variation of the system under investigation.

Epistemic uncertainty: Uncertainty resulting from imperfect knowledge; can be quantified and reduced.

Model structural uncertainty: Uncertainty that arises from a lack of sufficient understanding of the system (past, present or future), that is the subject of the policy analysis, including the behaviour of the system and the interrelationships among its elements.

Model technical uncertainty: The uncertainty generated by software or hardware errors.

Model outcome uncertainty: Total uncertainty on the model simulation (so endogenous rather than exogenous as the other categories).

Linguistic uncertainty: Uncertainty arising from language issues; can be quantified and reduced.

Table 1 - *Definition of the types of uncertainty categorised in Figure 1.*

References

[1] Durán J.M. (2020). What is a Simulation Model? *Minds and Machines 30*, 301-323.

[2] Briggs W. (2016). *Uncertainty: The Soul of Modeling, Probability & Statistics*. Springer Nature.

[3] Snow D.A. (2001). Plant Engineer's Reference Book. Butterworth-Heinemann.

[4] Cerezo D.C. (2017). Building archetype calibration for effective urban building energy modeling. Ph.D. thesis, Massachusetts Institute of Technology.

[5] Langevin, J., Reyna, J.L., Ebrahimigharebhaghi, S., Holck Sandberg, N., Fennell, P., Nägeli, C., Laverge, J., Delghust, M., Van Hove, M., Webster, J., et al. (2020). Developing a common approach for classifying building stock energy models. *Renewable & Sustainable Energy Reviews. 133*.

[6] Booth A.T., Choudhary R., Spiegelhalter D.J. (2012). Handling uncertainty in housing stock models. *Building and Environment 48*, 35-47.

[7] Funtowicz S.O., Ravetz J.R. (1993). Science for the post-normal age. *Futures 25*, 739-755.

[8] Walker W.E., Harremoës P., Rotmans J., van der Sluijs J.P., van Asselt M.B.A., Janssen P., Krayer von Krauss M.P. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integer Assess 4(1)*, 5–17.

[9] Coakley D., Raftery P., Keane M. (2014). A review of methods to match building energy simulation models to measured data. *Reneable and Sustainable Energy Reviews 37*, 123-141.

[10] Uusitalo L, Lehikoinen A., Helle I, Myrberg K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software 63*, 24-31.

[11] Tian W., Li Z., de Wilde P., Yan D. (2018). A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews 93*, 285-301.

[12] Regan H.M., Colyvan M., Burgman M.A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications 12*, 618-628.

[13] Morgan M.G., Henrion M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge: Cambridge University Press.

[14] Andrews C.J., Hassenzahl D.M., Johnson B.B. (2004). Accommodating Uncertainty in Comparative Risk. *Risk Analysis 24*: 1323-1335.

[15] Stirling A. (2010, December 30). Keep it complex. *Nature 468*, pp. 1029-1031.

[16] Finkel A.M. (1990). *Confronting Uncertainty in Risk Management: A guide for decision-makers*. Washington, D.C.: Resources for the Future, The Center for Risk Management.

[17] Chatfield C. (1995). Model Uncertainty, Data Mining and Statistical Inference. *Journal of the Royal Statistical Society Series A (Statistics in Society) 158*, 419-466.

[18] Hopfe C.J., Hensen J.L.M. (2011). Uncertainty analysis in building performance simulation for design support. *Energy and Buildings 43*, 2798–2805.

[19] Micovic Z., Schaefer M.G., Barker B.L. (2017). Sensitivity and Uncertainty Analyses for Stochastic Flood Hazard Simulation. *Sensitivity Analysis in Earth Observation Modelling*, 213-234.

[20] Petersen A., Janssen P. (2013). *Guidance for Uncertainty Assessment and Communication (Second Edition)*. The Hague: PBL Netherlands Environmental Assessment Agency.

[21] Ding Y., Shen Y., Wang J., Shi X. (2015) Uncertainty Sources and Calculation Approaches for Building Energy Simulation Models. *Energy Procedia 78*, 2566 – 2571.

[22] Titikpina F., Caucheteux A., Charki A., Bigaud D. (2015) Uncertainty assessment in building energy performance with a simplified model. *International Journal of Metrology and Quality Engineering 6*, 308-317.

[23] Oberkampf W.L., DeLand S.M., Rutherford B.M., Diegert K.V., Alvin K.F. (2002) Error and uncertainty in modeling and simulation. *Reliability Engineering and System Safety 75*, 333-357.

[24] Mcheick H., Mohammad A.F. (2014) The Evident Use of Evidence Theory in Big Data Analytics using Cloud Computing, *presented at the Canadian Conference on Electrical and Computer Engineering*, Toronto, Ontario, Canada, May 4-7, 2014.

[25] Refsgaard J.C., Henriksen H.J. (2004) Modelling guidelines: terminology and guiding principles. *Advances in Water Resources 27*, 71-82.

[26] Eisenhower B., O'Neill Z., Narayananb S., Fonoberovc V.A., Mezic ´ I. (2012) A methodology for meta-model based optimization in building energy models. *Energy and Buildings 47*, 292–301.

[27] Rodríguez G.C., Andrés A.C., Muñoz F.D., López J.M.C., Zhang Y. (2012) Uncertainties and sensitivity analysis in building energy simulation using macroparameters. *Energy and Buildings 67*, 79–87.

[28] Refsgaard J.C., van der Sluijs J.P., Hojberg A.L., Vanrolleghem P.A. (2007) Uncertainty in the environmental modelling process: A framework and guidance. *Environmental Modelling & Software 22*, 1543-1556.

[29] Fennell, P.J., Ebrahimigharehbaghi, S., Mata, E., Kokogiannakis, G., Amrith, S., Ignatiadou, S. (2021). A Review of the Status of Uncertainty and Sensitivity Analysis in Building-Stock Energy Models. *Renewable and Sustainable Energy Reviews* (in review).

Speaker: Sergei Kucherenko

Global Sensitivity Analysis and Global Optimization of Membrane Enhanced Peptide Synthesis model

Sergei Kucherenko
Imperial College London, London, SWT 2AZ, UK,
s.kucherenko@imperial.ac.uk
Wen Chen
Imperial College London, London, SWT 2AZ, UK,
wen.chen@nus.edu.sg
Nilay Shah
Imperial College London, London, SWT 2AZ, UK,
N.Shah@imperial.ac.uk
Nicholas Lau
Imperial College London, London, SWT 2AZ, UK,
wai.lau19@imperial.ac.uk

Uncertainty and sensitivity analysis has been recognized as an essential part of model applications. Global sensitivity analysis (GSA) is used to identify key parameters whose uncertainty most affects the output. This information can be used to rank variables, fix or eliminate unessential variables and thus decrease problem dimensionality. The variance-based method of Sobol' sensitivity indices (SI) became very popular among practitioners due to its easiness of interpretation. Computation of Sobol' indices using direct Monte Carlo method generally requires a large number of function evaluations to achieve reasonable convergence. Metamodel based methods have proved to be much more efficient especially for complex practical problems. In majority of cases metamodel are used only for computing GSA measures. In this work we used the Bayesian Sparse Polynomial Chaos Expansion method (BSPCE) implemented in SobolGSA was used for GSA and GO. BSPCE makes use of sparse PCE by selecting different PCE structures using a Bayesian approach with the Kashyap information criterion for model selection [1].

In many areas practitioners are interested in minimizing/maximizing quantities of interests. We developed a unified approach in which metamodels are used for both GSA and Global Optimization (GO). It was implemented in the software tool SobolGSA [2]. We note, that there are fundamental differences between these two problem. They are summarized in the table below using a model $Y = f(\vec{X}),\ \ \vec{X} = (X_1, X_2, ..., X_n) \in \Omega$ as an example:

Global Sensitivity Analysis	Global Optimization
Interpretation: random setting - X is a vector of random variables	Interpretation: deterministic setting- X is a vector of inputs given in the deterministic domain
Objective: study the effect of input uncertainty on the variability in the model output	Objective: Search the global minima of *f* and the set of global minimizers X*

In order to illustrate the developed approach we considered Membrane Enhanced Peptide Synthesis model (MEPS) [3],[4]. Peptides are biopolymers that control, direct, and coordinate inter- and intracellular communications and cellular functions in many living systems. We cross-verified by experiment and sensitivity analysis, that the generation of truncated error sequences diminishes as extent of reaction increases, ascertaining that a vast majority of error sequences can be neglected in simulation. In addition, we showed that a double membrane system is

advantageous over single membrane due to higher retention of product, resulting in higher yield. Other process variables such as recycle ratio and diavolume are crucial in further optimizing the yield.
Considered model has 9 uncertain parameters. It is known that the model structure and parameter distributions affect the results of GSA. We compared the results of GSA using two different distributions of uncertain paraments: 1) a uniform distribution assuming +/- 20% variation around the parameter mean value μ; 2) a normal distribution assuming the standard deviation $\sigma=\mu/2$. In both case we found that only two parameters, namely the rejection rate of Piperidine and the rejection rate of the final product are important, with the rejection rate of Piperidine being the most important in the first case and the rejection rate of the final product in the second case. There are some not very significant interactions between parameters in the first case and strong interactions between parameters in the second case: the main effect Sobol' index for the rejection rate of Piperidine is equal to 0, although its total effect index is close to 0.8. On the second stage we performed global optimization maximizing total peptide yield and verified the results by running the full model in gPROMS. Although there is some discrepancy between predictions based on metamodel and validation results obtained using the full model with the maximizer values, proposed approach gives a good guidance with regards to further directions of modelling and experimental work.

Acknowledgements. We acknowledge the financial support of Eli Lilly and Company.

References

[1] Q. Shao, A. Younes, M. Fahs, T-A. Mara. Bayesian sparse polynomial chaos expansion for global sensitivity analysis. *Computer Methods in Applied Mechanics and Engineering*, 318: 474-496, 2017.

[2] SobolGSA software. *User manual*, https://www.imperial.ac.uk/process-systems-engineering/research/free-software/sobolgsa-software/, 2021

[3] W. Chen, M. Sharifzadeh, N. Shah, AG. Livingston. Implication of Side Reactions in Iterative Biopolymer Synthesis: The Case of Membrane Enhanced Peptide Synthesis. *Industrial & Engineering Chemistry Research*, 56(23):6796-804, 2017.

[4] W. Chen, M. Sharifzadeh, N. Shah, AG. Livingston. Iterative peptide synthesis in membrane cascades: Untangling operational decisions. *Computers & Chemical Engineering*, 115:275-85, 2018

Speaker: Claire Lauvernet

Metamodeling methods that incorporate qualitative variables for improved design of vegetative filter strips.

Claire Lauvernet
RIVERLY, INRAE Lyon-Villeurbanne, FR, claire.lauvernet@inrae.fr
Céline Helbert
UMR CNRS 5208, Ecole Centrale de Lyon, FR, celine.helbert@ec-lyon.fr
Zhu Xujia
ETH Zürich, Chair of Risk, Safety & Uncertainty Quantification, CH, zhu@ibk.baug.ethz.ch
Bruno Sudret
ETH Zürich, Chair of Risk, Safety & Uncertainty Quantification, CH, sudret@ethz.ch

Significant amounts of pollutant are measured in surface water, their presence due in part to the use of pesticides in agriculture. One solution to limit pesticide transfer by surface runoff is to implement vegetative filter strips (VFS) along rivers. These buffer zones are identified as the best management practices of choice for runoff mitigation to prevent and limit the transfer of pollutants from agricultural fields to water resources. They are mandatory or highly advised depending on the country and conditions. Since their location is part of the farmer's field, the sizing of these strips is a major issue. However, to be efficient, they need to be properly designed, depending on the specific context in which they are implanted (climate, soil, water table, etc.).

The BUVARD modeling toolkit was developped to design VFSs throughout France according to all these local influencing factors [1]. Processes that drive the pesticide fate are complex and interact : infiltration, surface runoff, sediment trapping, pesticide transfer, etc., and are summarized through nonlinear equations and/or conceptual and/or stochastic modeling. To represent most of them, BUVARD is composed of several models centered around the numerical model VFSMOD [4], which quantifies dynamic effects of VFS site-specific pesticide mitigation efficiency (see figure 1).

The way BUVARD is built (i.e., a chain of several models) implies a large set of parameters that are difficult to measure (for the physical modeling) or to calibrate (for the conceptual modeling). Furthermore, inputs and outputs are dynamic (for example, rainfall, surface runoff, etc.), and inputs are either quantitative or qualitative variables. For all these reasons, we get an expensive tool to use, and a high uncertainty, which has to be quantified, particularly in the case of an operationnal tool. Metamodeling BUVARD is *a priori* a relevant solution to decrease the cost and the complexity of the model, to help users design VFSs that are adapted to specific contexts.

Added to the mixed qualitative and quantitative variables, that is not often taken into account in surrogate methods, we have to deal with a huge number of zero values of

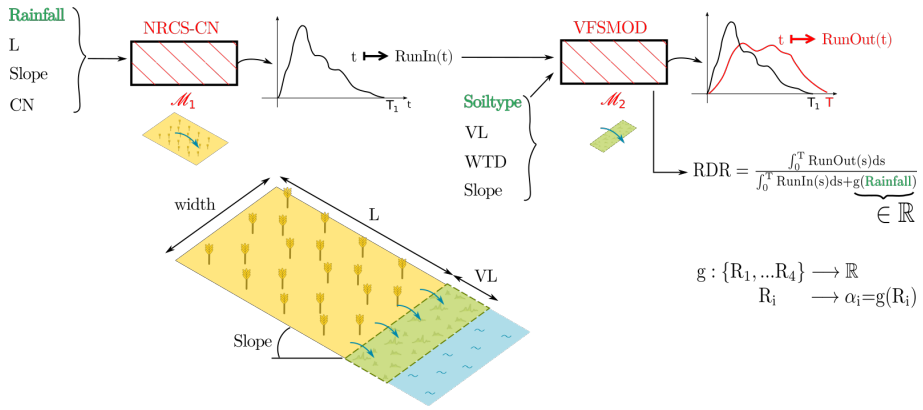


Figure 1: BUVARD toolchain, with inputs and outputs used to build the surrogate.

inputs and outputs, and to boundaries in which the main output has to range. In this study, different methods are tested: (i) Mixed-Kriging, a Kriging method that was implemented with a covariance kernel for a mixture of qualitative and quantitative inputs [3] (ii) PCE, that was also adapted to qualitative variables, encoding categorical inputs as quantitative dummy variables, thus allowing for transforming the problem into the standard regression setup [5] (iii) DeepGP (Deep Gaussian Processes), that is particularly suited for non-stationnary models [2]. We show that categorical variables are properly taken into account by the Kriging and by the PCE adaptations, and that mixed variables methods outperform the same methods applied per category, and even more with smaller samplings. DeepGP, that was not adapted to qualitative variables, does not need any classification or boundaries, and reaches the performance of the adapted methods. However, it needs repetitions for the most complex soils, with a much higher numerical cost, that is multiplied by the number of categorical variables. Finally, we perform a global sensitivity analysis with the help of the two surrogate models with the best accuracy. The results show that they give the same ranking of the importance of the input parameters.

[1] N. Carluer, C. Lauvernet, D. Noll, and R. Muñoz-Carpena. Defining context-specific scenarios to design vegetated buffer zones that limit pesticide transfer via surface runoff. *Science of The Total Environment*, 575:701–712, 2017.

[2] A. Damianou and N. D. Lawrence. Deep Gaussian processes. volume 31 of *Proceedings of Machine Learning Research*, pages 207–215, 2013. PMLR.

[3] C. Lauvernet and C. Helbert. Metamodeling methods that incorporate qualitative variables for improved design of vegetative filter strips. *Reliability Engineering & System Safety*, 204:107083, 2020.

[4] R. Muñoz-Carpena, J.E. Parsons, J.W. Gilliam. Modeling hydrology and sediment transport in vegetative filter strips. *Journal of Hydrology*, 214(1-4):111–129, 1999.

[5] X. Zhu and B. Sudret. Sparse polynomial chaos expansions for mixed categorical / continuous inputs. (in preparation).

Speaker: Nabir Mamnun

Global sensitivity analysis of a one-dimensional ocean biogeochemical model

Nabir Mamnun
Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Germany
Nabir.Mamnun@awi.de

Marvin Thelen
Research Group for Theory of Hybrid Systems, RWTH Aachen University, Germany
marvin.thelen@rwth-aachen.de

Christoph Völker
Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Germany
Christoph.Voelker@awi.de

Pascal Richter
Steinbuch Centre for Computing, Karlsruhe Institute of Technology, Germany
pascal.richter@rwth-aachen.de

Sebastian Krumscheid
Chair of Mathematics for Uncertainty Quantification, RWTH Aachen University, Germany
krumscheid@uq.rwth-aachen.de

Lars Nerger
Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Germany
Lars.Nerger@awi.de

We perform a global sensitivity analysis of an ocean biogeochemical (OBGC) model with regard to the sensitivity of its input parameters. OBGC models involve a wide variety of processes (e.g., carbon cycling, air-sea interactions and gas transfer, and components of the marine ecosystem) controlled by numerous input parameters which are not precisely known. OBGC model simulations are highly uncertain in model parameterization [1]. Global sensitivity analysis (GSA) aims to identify the parameters whose uncertainty has the largest impact on the variability of a simulated quantity of interest – for instance net primary production in the ocean. In this study we consider the biogeochemical model Regulated Ecosystem Model 2 (REcoM2) [2], coupled with the MIT General Circulation Model (MITgcm) [3] in a 1D configuration at two ocean biogeochemical time series stations with different environmental conditions – the Bermuda Atlantic Time-series

Study (BATS) the North Atlantic and the Dynamique des Flux Atmosphériquesen Méditerranée (DYFAMED) in the Mediterranean Sea. We analyze the variance-based Sobol indices [4] for GSA of ten selected biogeochemical parameters of REcoM2, by applying a Monto Carlo method for sampling from parameter space of the selected parameters. Sobol indices quantify the respective influence of uncertain parameters either through their direct effect (first order sensitivity index), through their mutual interaction (higher order sensitivity index), or collective influence of all perturbed parameters (total order sensitivity index).

In this study, we set the focus on the first and total order sensitivity indices. We found out that the chlorophyll degradation and grazing parameters are most influential for chlorophyll-a concentration simulation. For net primary production, parameters related to photosynthesis and grazing are particularly important. The total order indices indicate that some parameters have mutual dependencies.

[1] Schartau, M., et al., Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling. *Biogeosciences,* 14(6):1647-1701, 2017.

[2] Hauck, J., et al., Seasonally different carbon flux changes in the Southern Ocean in response to the southern annular mode. *Global Biogeochem Cycles,* 27(4):1236-1245, 2013.

[3] Marshall, J., et al., A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research-Oceans,* 102(C3):5753-5766, 1997.

[4] Sobol, I.M., Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation,* 55(1-3)271-280, 2001.

Speaker: to be defined

SIML@B: A European Commission Tool for Sensitivity Analysis

Thierry Mara
European Commission, Joint Research Centre (JRC), Ispra, Italy,
thierry.mara@ec.europa.eu
Ivano Azzini
European Commission, Joint Research Centre (JRC), Ispra, Italy,
ivano.azzini@ec.europa.eu
Rossana Rosati
European Commission, Joint Research Centre (JRC), Ispra, Italy,
rossana.rosati@ec.europa.eu

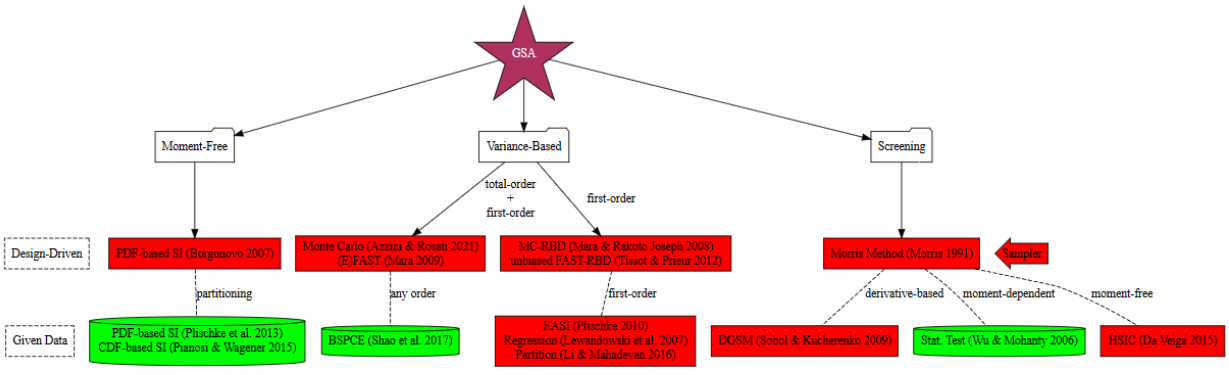
The Competence Centre on Modelling of the European Commission (EC) has been developing an online tool called **Siml@b** in order to allow modellers performing Global Sensitivity Analysis (GSA) of their models. The tool is developed with R Shiny Environment.

A possible classification of the main existing GSA methods can be the following:

Moment-free methods: which do not rely on any specific moment of y (like the variance). They quantitatively assess how different is the unconditional random variable y w.r.t. the conditional random variable $y|x_i$. One such quantitative importance measure is the one proposed by Borgonovo (2007).

Variance-based methods: which owe their groundings to the work of Ilya M. Sobol' (1993). They assess the so-called Sobol' indices which are the quantitative measure of importance of interest. These indices stem from the analysis of variance (ANOVA decomposition) of y.

Screening methods: Screening methods are used for factor fixing setting. Screening methods are also named qualitative methods because they do not allow the ranking of the input variables by order of importance. Morris method (1991) is an instance of screening method.



Some methods require specific sampling designs (Design-Driven), some others not (Given-Data). Currently, in Siml@b, only Given-Data approaches in green are implemented but its

extension to further methods is ongoing.

Requirements for using `GIVEN DATA` methods
The dataset must be an array [X,y] of size N•(d+1). The first d columns must contain the Monte Carlo sample of x (i.e. input vector), the last column the corresponding values of the model response y (i.e. scalar output).

Requirements for using `DESIGN-DATA` methods
A specific design is required in this case and one can generate the latter with the tool called **Sampler**. Then, after running the model and collecting the dataset one can create an array [X,y], the size of which depends on the method used.

Datasets must be stored in a csv or txt file on the user's computer, and may contain the variables' name in the first row. Note that Excel files are not handled.

Sim@b appears as a really versatile and user-friendly tool. Uncertainty and sensitivity analyses, Sobol' sensitivity indices, and corresponding graphics, can be obtained with a few steps. The execution of the program usually takes only a short time and can be carried out by people without any experience in programming simply by uploading the data file. A comprehensive guide is also available.

At the SAMO 2022 Conference, a demonstration will be performed and specific cases will be discussed like the case of dependent input variables (see Rosenblatt 1952 and Mara & Becker 2021).

[1] M. Rosenblatt, *Remarks on a multivariate transformation*, Ann. of Math. Stat., Vol. **23**(3), 470-472, 1952.

[2] M.D. Morris., Factorial sampling plans for preliminary computational experiments, Technometrics, Vol. **33**(2), 161-174, 1991.

[3] I.M. Sobol', Sensitivity estimates for nonlinear mathematical models, Math. Mod. and Comput. Exp., Vol **1**, 407-414, 1993.

[4] E. Borgonovo, A new uncertainty importance measure, Reliab. Eng. & Syst. Saf., Vol. 92, 474-496, 2007.

[5] T.A. Mara & W.E Becker., *Polynomial chaos expansion for sensitivity analysis of model output with dependent inputs,* Reliab. Eng. & Syst. Saf., SAMO Special Issue, 2021

Speaker: Robert Milton

Locating an Active Subspace using Gaussian Processes

Robert Milton
Department of Chemical and Biological Engineering, University of Sheffield, UK,
r.a.milton@sheffield.ac.uk
Aaron Yeardley
Department of Chemical and Biological Engineering, University of Sheffield, UK,
asyeardley1@sheffield.ac.uk
Solomon Brown
Department of Chemical and Biological Engineering, University of Sheffield, UK,
s.f.brown@sheffield.ac.uk

Analysis of experimental or simulation data is often hampered by the sheer number of input variables. Predictions become unreliable due to the notorious "curse of dimensionality", and visualisation and optimisation become virtually impossible in spaces of even moderate dimensionality. To mitigate all these difficulties, screening inputs for relevance has long been a basic tool in data analysis. Latterly it has been shown that the effectiveness of screening is enhanced dramatically by combining inputs to find the most amenable basis for screening [1]. One seeks a rotation Θ

$$z = \Theta x$$

which concentrates the relevant inputs in the first few dimensions of z. The rotation is optimised according to some measure of relevance, then z is truncated to those dimensions whose relevance exceeds a chosen threshold. The reduced input space is called an active subspace, and is often of remarkably small dimension.

Active subspace approaches [1] use local sensitivity measures directly on experimental and simulation data, which may be noisy. This invites issues regarding computational expense and exorbitant data requirements to cover the entire input space and obtain accurate estimates of relevance. Aside from which, it is surely desirable in many applications to subtract noise from the underlying data before searching for an active subspace. Clearly these issues may be tackled, if not entirely removed, by emulating the underlying data with a smoothed surrogate, and seeking an active subspace for that. An appropriate surrogate eliminates much of the noise and facilitates efficient, even analytic, computation of relevance from limited data.

This work uses a Gaussian Process (GP) surrogate for which relevance is measured analytically. Relevance measures are global, the well-known Sobol indices, although other variance-based measures may be considered. The GP uses a smooth ARD squared exponential kernel

$$\operatorname{cov}\left[\left[\mathsf{y}|x\right],\left[\mathsf{y}|x'\right]\right] = k_{\mathsf{y}}(x,x') = F\exp\left(-\frac{(x-x')^T\left[\Lambda\right]^{-2}(x-x')}{2}\right)$$

where $[\Lambda]$ is a diagonal matrix of ARD lengthscales.

An efficient algorithm is formulated to find Θ which optimally drains relevance from dimensions to be eliminated, without wasting resources distributing relevance amongst retained dimensions. This algorithm is applied to standard test functions taking 5 input dimensions, to locate their active subspaces.

It is well known that GPs are very vulnerable to the curse of dimensionality, performing poorly for input dimensions greater than 6 or so. To mitigate this, we investigate building the active subspace from the bottom up. All but 5 input dimensions are initially ignored (effectively regarded as sources of noise). The active subspace technique is used to reduce this input space to 3 dimensions. This active subspace is then enhanced with two of the inputs previously ignored to generate a new dataset with 5 inputs. Iterating through GP regression and subsequent active subspace location and enhancement enables one to add many dimensions gradually, without ever falling prey to the curse of dimensionality. This technique is applied to a suite of test functions, this time taking 10 input dimensions. The bottom-up approach is crucially enabled by the novel combination of surrogates with active subspace construction, exploiting the ability to ignore variation of inputs we wish to defer by relegating it to noise in the surrogate.

[1] P. G. Constantine. *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*. Society for Industrial and Applied Mathematics, 2015.

Speaker: Magdalena Olczyk, Marta Kuc-Czarnecka

Digital transformation and economic growth – DESI improvement and implementation

Prof. Magdalena Olczyk
Department of Management and Economics,
Gdansk University of Technology, Poland,
Magdalena.Olczyk@zie.pg.gda.pl
Marta Kuc-Czarnecka, PhD
Department of Management and Economics,
Gdansk University of Technology, Poland
marta.kuc@pg.edu.pl

The paper concentrates on digital transformation measurement and the relationship between digital transformation and economic growth. The digital transformation is underway and accelerates. The established metrics and assessment instruments cannot keep up with the rapid digital transformation pace[1]. The [2]identifies many gaps in the current framework of measuring digital transformation and recommends improving the international comparability of the indicators currently in use. Also, a better adaptation of current statistical systems to rapid changes brought about by the digital revolution is strongly recommended. The literature covers several indices used to assess the development of the digital economy, starting from the Information Society Index (1997), E-Readiness Index (2000), Technology Achievement Index (2001), E-Government Development Index (2002), ICT Development Index (2002), Networked Readiness Index (2002), Digital Access Index (2003), Knowledge Economy Index (2005), Digital Opportunity Index (2005), ICT Opportunity Index (2005), ICT Diffusion Index (2006), and ending up with the newest one – the Digital Economy and Society Index (2014). These measures propose a holistic framework for assessing the digital revolution's multi-faceted impact on society and economies

Our paper analyses the latest index, the Digital Economic and Society Index (DESI), proposed by the European Union [3]. It is based on 37 individual indicators and evaluates the digital transformation of EU countries from the point of view of e-business, e-society and e-administration. So far, the DESI has been used to assess the degree of digital economy development in particular countries. To the best of our knowledge, the methodology proposed in the DESI has not been verified or attempted to be improved yet.

We wish to fill this gap, aiming to improve the DESI and investigating its relationship with economic growth. The paper examines, using the sensitivity-based analysis [4], whether methodological modifications to the DESI structure boost its ability to capture the digitalisation digitalization of society and economies. We consider whether the selection of weights of variables included in the DESI is optimal or could be improved. Additionally, we use panel data models to check if changes in the DESI influence EU economies' growth. The research questions that will be answered in this paper are: (1) can we improve the DESI as a composite indicator of the digital transformation of the EU-28? and (2) could the DESI be used as a GDP per capita forecast measure for the EU-28? To answer those research questions, we use data from the 2015-2020 DESI reports.

Results

Our analysis of the Digital Economic and Society shows that besides reducing the set of variables, it would also be necessary to drastically change the value of the weights assigned to each pillar, subpillar, and individual variables. The analysis pointed to significant discrepancies between the original weights and the optimized ones. The issues

related to connectivity, the importance of which should be close to 0.60, have the most substantial effect on the final ranking. Due to the strong correlation, it is challenging to develop a well-balanced index without assigning zero weights to some elements. Moreover, composite indicators are inconsistent as linear aggregation is a poor way of summarising the information that, perhaps, should not be summarised in the first place. Combining certain pillars or modifying their content to provide conceptual coherence should perhaps be considered.

Our results are of significant importance to policymakers regarding the measurement, support, and deepening of digital transformation. Existing metrics and assessment instruments fail to keep up with the rapid pace of digital transformation. To measure digital transformation, countries use existing indicators drawn from various areas, including education, innovation, trade, economic and social issues, or a composite indicator such as the DESI for EU countries. Our results reveal that for current and quick analyses of digital transformation development or some international comparisons in this area, it is appropriate to consider only several indicators such as the coverage of broadband (fixed, fast, 4G), level of software skills, and the percentage of enterprises analyzing big data and selling products and services online. These indicators are most crucial from the digital transformation level point of view.

As digital technologies continue to reshape society and the economy dramatically, many countries are pursuing large-scale supporting initiatives in this area. Our study indicates that EU countries should develop fast broadband plus 4G technologies as well as invest in all education programmes aimed to create a new generation capable of adapting and working with ICTs. Citizens and employees with high digital skills are of common interest to both the state and employers, so building partnerships between the state and private sector to make people more familiar with ICTs is strongly recommended.

Additionally, our study confirms that the DESI is useful for explaining changes in GDP per capita. This is good news for the poorest EU countries because the gap between rich and poor countries in the European Union can be closed or eliminated by fast and intensive digital transformation. It is crucial for each country that uses its natural resources (such as oil or minerals) or export trade to ensure GDP growth and is still unable to take off to reach a high development stage.

[1] OECD (2019a). *OECD Skills Outlook 2019: Thriving in a Digital World*. Paris. Retrieved from https://doi.org/10.1787/df80bc12-en.

[2] OECD (2019b). *Going Digital: Shaping Policies, Improving Lives*. *Going Digital: Shaping Policies, Improving Lives*. https://doi.org/10.1787/9789264312012-en

[3] European Commission. (2020). *Digital Economy and Society Index 2020: Methodological note*. Retrieved from http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=67082

[4] Becker, W.; Saisana, M.; Paruolo, P.; Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological Indicators*, 80, 12–22, https://doi.org/10.1016/j.ecolind.2017.03.056

Speaker: Rony Parra

Understanding The Future Metabolism of Ecuador's Energy System Using MuSIASEM

A1Rony Parra
Instituto de Investigaciones Hidrocarburíferas,
Universidad Central del Ecuador, Ecuador,
Research Group on Integrated Assessment: Sociology,
Technology and the Environment (IASTE)/Spain
rmparra@uce.edu.ec
A2 Génesis Yánez
Instituto de Investigaciones Hidrocarburíferas,
Universidad Central del Ecuador, Ecuador,
gbyanez@uce.edu.ec

The rapid and unavoidable change of the countries' energy systems composition is a matter of global interest in various science fields and the public policymakers. Several institutions dedicated to conducting behavior analysis in the systems of energy forecast different future scenarios regarding the composition of the final energy consumption on a global scale in the short term. These differences in the results are due to the criteria that organizations and governments take into consideration to forecast the scenarios. The lack of an established methodology to conduct these types of assessments creates several forecasts and different understandings of the energy system. In Ecuador, the energy system has not been completely evaluated and the results of the energy forecast of some undertaken studies differ from each other in their final scenario making it difficult to know the health of the system and anticipate the preparation of public policies. Therefore, this work aims to develop a methodology based on MuSIASEM grammar to build future energy scenarios based on the understanding of the scarcity of primary energy sources and the dynamics of the end-use energy in the different compartments of socioeconomic systems, applied in the Ecuador system by 2035

The construction of the Ecuadorian energy system scenario is based on the structuration of the MuSIASEM grammar, in which the energy system elements and their relationship with the socioeconomic sectors are shown at different levels and scales.
Energy systems join functional (example: electricity production) and structural (example: thermal/hydraulic energy) subcategories within a metabolic route, so these make it possible to join two non-equivalent points of view of the metabolic pattern of a given society. For the generation of the accounting of flows (example: electricity/fuels) and funds (example: power capacity/ human activity). Moreover, within the energy grammar, it is necessary to take in mind the concepts of primary energy sources (PES), energy carriers (EC), energy systems (ES), and energy end uses (EU)

The 2035 energy demand for Ecuador indicates that it will exist an increase in the fuel and electricity metabolic rate. The increase will be from 2,6 to 4 MJ/h of fuel, and from 0 to 0,5 MJ/h of electricity for the Agriculture sector, from 7,4 and 11,3 MJ/h to 15 and 30 MJ/h of electricity and fuel respectively for the Building and Manufacturing sector, from 0 to 2 MJ/h of electricity

and from 309 to 300 MJ/h of fuel for the Transport sector, from 4,5 to 5 MJ/h of electricity and from 4 to 6 MJ/h of fuel for the Service and Government sector, from 0,2 to 0,3 MJ/h in electricity and from 0,3 to 0,5 MJ/h of fuels for the Households sector. Finally, the sector that concentrates the greater consumption by work hour will be the Energy sector, in which case the energetic intensity increases from 114,7 to 115 MJ/h of electricity and from 750,4 to 810 MJ/h of fuel. The electricity generation system shows a gross energy increase of 28.083 GWH produced in 2017 to 45.463 GWH by 2035.

The scenario considers the production mix extension to hydro-power and renewable energies while the thermal generation is reduced from 26% to 17% from the total generation of electricity. This allows maintaining the CO2 emissions at the year base levels of 6.477 KTon CO2.

[1] Castro Verdezoto, P. L., Vidoza, J. A., & Gallo, W. L. R., 2019. Analysis and projection of energy consumption in Ecuador: Energy efficiency policies in the transportation sector. Energy Policy, 134(November 2018). https://doi.org/10.1016/j.enpol.2019.110948

[2] Equinor, 2019. Energy Perspectives 2019. Long-term macro and market outlook

[3] Giampietro, M., Mayumi, K. & Sorman, A.H., 2012. The Metabolic pattern of societies : where economists fall short. London: Routledge

[4] Giampietro, M., Diaz-Maurin, F., 2014b. The Energy Grammar. In: Giampietro, M., Aspinall, R.J., Ramos-Martin, J., Bukkens, S.G.F. (Eds.), Resource Accounting for Sustainability: The Nexus between Energy, Food, Water and Land Use. Routledge, Abingdon, pp. 90–115.

[5] Giampietro M, Sorman AH., 2012. Are energy statistics useful for making energy scenarios? Energy;37:5e17.https://doi.org/10.1016/j.energy. 2011.08.038

[6] Greenpeace, 2019. https://www.greenpeace.org/international/

[7] International Energy Agency IEA, 2019. Energy Technology Perspectives. https://www.iea.org/topics/energy-technology-perspectives

[8] Odum, H.T., 1996. Environmental Accounting: Emergy and Environmental Decision Making. John Wiley, New York.

[9] Parra Narváez, R., 2015. Factor de emisión de CO2 debido a la generación de electricidad en el Ecuador durante el periodo 2001-2014. ACI Avances En Ciencias e Ingenierías, 7(2). https://doi.org/10.18272/aci.v7i2.269

[10] Parra, R., Di Felice, L. J., Giampietro, M., & Ramos-Martin, J., 2018. The metabolism of oil extraction: A bottom-up approach applied to the case of Ecuador. Energy Policy, 122(July), 63–74. https://doi.org/10.1016/j.enpol.2018.07.017

[11] Pinzón, K., 2018. Dynamics between energy consumption and economic growth in Ecuador: A granger causality analysis. Economic Analysis and Policy, 57, 88–101. https://doi.org/10.1016/j.eap.2017.09.004

[12] Velasco-Fernández, R. (2017). The pattern of Socio-Ecological Systems. Thesis Phd

[13] Velasco-Fernández, R., Giampietro, M., & Bukkens, S. G. F. (2018). Analyzing the energy performance of manufacturing across levels using the end-use matrix. Energy, 161, 559–572. https://doi.org/10.1016/j.energy.2018.07.122

[14] World Energy Council-WEC, 2019. https://www.worldenergy.org/g

Speaker: Rik Rutjens

Elementary Effects for models with dimensional inputs of arbitrary type

R.J.L. Rutjens
School of Mathematical Sciences, University of Nottingham, UK,
pmxrr3@nottingham.ac.uk
M.R. Owen
School of Mathematical Sciences, University of Nottingham, UK
L.R. Band
School of Mathematical Sciences & Plant Sciences, University of Nottingham, UK
M.D. Jones
School of Geography & Future Food Beacon, University of Nottingham, UK

Models in the biological and environmental sciences typically have many free parameters (factors) [1]. Calibration of these parameters often requires empirical data, sometimes costly or simply impossible to obtain, which introduces uncertainty into the model. However, according to the 'sparsity of factors' principle, very often only a small subset of factors in a system has a significant influence on the system output [2]. As such, it can be extremely beneficial for model development to identify unimportant parameters, so that they may be set to a fixed value, and the most important factors, so that every effort can be made to accurately estimate this group. This can greatly decrease model complexity, while increasing trust in the model.

The Elementary Effects method (EE) [3] has received increased attention in recent years as a simple and computationally efficient global screening approach. EE is designed to identify (non-)influential parameters and produces a qualitative ranking of factor importance. However, the current body of literature about EE almost exclusively treats the case where all input factors are dimensionless and take values in the unit interval. Applying EE in its current form to real-life models might lead to erroneous ranking results.

Here, we therefore discuss the application of EE to dimensional models and models where input factors take values on arbitrary intervals or contain inputs of integer or Boolean type. We show that scaling of the effects in the input direction by a function of the input range is necessary in such cases to prevent erroneous ranking results while obtaining results consistent with the general notion of sensitivity. Furthermore, we propose an alternative normalized dimensionless sensitivity index based on recent advances in the field. We show that this index provides a standardized way to identify (non)-influential factors, while avoiding the issue of scaling the effects in the output direction. Finally, we consider whether spread (as introduced by Campolongo et al. [4]) and discrepancy of the sampled parameter sets can be used as proxies for the ability to accurately rank factors.

[1] G. Qian and A. Mahdi. Sensitivity analysis methods in the biomedical sciences, *Math. Biosci.,* 323:108306, 2020.

[2] G. E. P. Box and R. D. Meyer. An analysis for unreplicated fractional factorials, *Technometrics*, 28(1):11–18, 1986.

[3] M. D. Morris. Factorial sampling plans for preliminary computational experiments, *Technometrics*, 33(2):161–174, 1991.

[4] F. Campolongo et al. An effective screening design for sensitivity analysis of large models, *Environmental Modelling and Software*, 22(10):1509–1518, 2007.

Speaker: Aleksei Sorokin

Efficiently Computing Sensitivity Indices in QMCPy

Aleksei Sorokin
Applied Mathematics, Illinois Institute of Technology, USA, asorokin@hawk.iit.edu
Jagadeeswaran Rathinavel
Applied Mathematics, Illinois Institute of Technology, USA, jrathin1@hawk.iit.edu

Sobol' indices quantify the importance of a function's inputs to explaining the output's variance [3, Appendix A]. Normalized Sobol' indices, or sensitivity indices, have been used in a variety of applications for global sensitivity analysis. Monte Carlo methods present an efficient approach for approximating these importance scores. In this talk, we will describe our implementation of such techniques into QMCPy [1], an open source Quasi-Monte Carlo library in Python. QMCPy utilizes algorithms from [2] to adaptively select an appropriate number of samples so the approximation is guaranteed to be within an desired tolerance of the true sensitivity indices.

[1] Sou-Cheng T. Choi, Fred J. Hickernell, R. Jagadeeswaran, Michael J. McCourt, and Aleksei G. Sorokin. QMCPy: A Quasi-Monte Carlo Python library, 2020+.

[2] Lluís Antoni Jiménez Rugama and Laurent Gilquin. Reliable error estimation for Sobol' indices. *Statistics and Computing*, 28(4):725–738, July 2018.

[3] Art B. Owen. *Monte Carlo theory, methods and examples.* 2018.