

Sensitivity analysis for sets : application to pollutant concentration maps

N. Fellmann^{1,2}, M. Pasquier², C. Blanchet-Scalliet¹, C. Helbert¹, A. Spagnol², and D. Sinoquet²

¹École Centrale de Lyon, CNRS UMR 5208, Institut Camille Jordan, 36 Avenue Guy de Collongue, 69134 Écully, France,

²IFP Energies Nouvelles

{noe.fellmann, christophette.blanchet, celine.helbert}@ec-lyon.fr

{noe.fellmann, mathis.pasquier, delphine.sinoquet, adrien.spagnol}@ifpen.fr

Abstract

In the context of air quality control, our objective is to quantify the impact of uncertain inputs such as meteorological conditions and traffic parameters on pollutant dispersion maps. It is worth noting that the majority of sensitivity analysis methods are designed to deal with scalar or vector outputs and are ill suited to a map-valued output space. To address this, we propose two classes of methods. The first technique focuses on pointwise indices. Sobol indices are calculated for each position on the map to obtain Sobol index maps. Additionally, aggregated Sobol indices are calculated. Another approach treats the maps as sets and proposes a sensitivity analysis of a set-valued output with three different types of sensitivity indices. The first ones are inspired by Sobol indices but are adapted to sets based on the theory of random sets. The second ones adapt *universal* indices defined for a general metric output space. The last set indices use kernel-based sensitivity indices adapted to sets. The proposed methodologies are implemented to carry out an uncertainty analysis for time-averaged concentration maps of pollutants in an urban environment in the Greater Paris area. This entails taking into account uncertain meteorological aspects, such as incoming wind speed and direction, and uncertain traffic factors, such as injected traffic volume, percentage of diesel vehicles, and speed limits on the road network.

1 Introduction

Air quality and dispersion of atmospheric pollutants are key aspects tackled by the emerging discipline of urban physics. Using computational fluid dynamics (CFD) as a tool, this branch of physics simulates pollutant dispersion to aid in urban planning and emergency decision-making. In Pasquier et al. 2023, a complete modelling chain has been developed to simulate traffic-related pollutant dispersion at the local urban scale. The method combines a microscopic traffic simulator with a physical engine model to estimate realistic road emissions, which are used as input of a CFD code to model unsteady atmospheric dispersion and compute two-dimensional time-averaged concentration maps at ground level. A presentation of this application is displayed in figure 1.

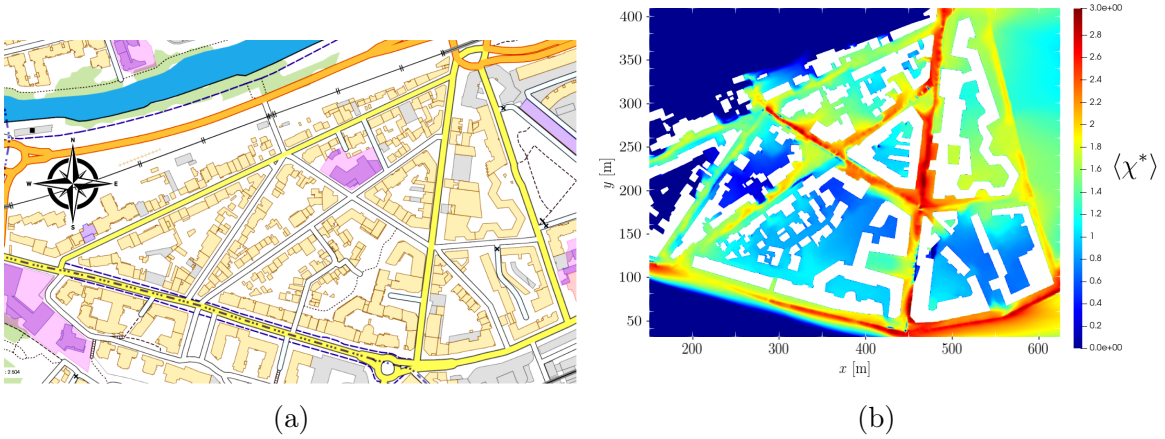


Figure 1 – Concentration maps from CFD numerical simulations of pollutant dispersion at the urban scale (Pasquier et al. 2023).

(a) Domain of interest, captured from OpenStreetMaps.

(b) Time-averaged map of the logarithm of the concentration, denoted $\langle \chi^* \rangle$, obtained from a run of the CFD code with realistic road emissions. The wind is coming from the left in this case.

The modelling chain outputs two-dimensional concentration maps from a set of inputs associated with meteorological and traffic-related uncertainties. These five uncertain inputs are the incoming wind direction θ [radians] and intensity U_∞ [metres/second], together with the volume of traffic injected into the road network in the urban geometry, q [vehicles/hour], the proportion of diesel and petrol engines in the fleet β [-], and the speed limit over the network, ν_{max} [kilometres/hour]. The modelling chain with the corresponding uncertain variables is displayed in figure 2 and the probability distributions associated with each variable are given in table 1 and plotted in figure 3.

To efficiently perform a global sensitivity analysis (SA) to compare the relative influence of these uncertain inputs on the output maps, we construct a metamodel of the entire modelling chain using a combination of principal component analysis and Gaussian process regression (GPR), known in the literature as POD-GPR (Guo and Hesthaven 2018; Marrel et al. 2014; Nony et al. 2023). The metamodel is trained on a Latin Hypercube Sampling (LHS) design of experiments of 150 expensive CFD simulations – each simulation requires about 7 hours of computational time on 960 cores on the high performance computing facility of IFP Energies Nouvelles – and validated on a set of 85 additional simulations. A global coefficient of determination, defined as in Nony et al. 2023, is computed

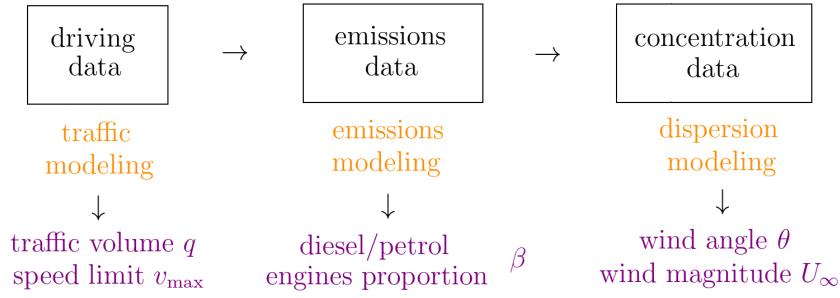


Figure 2 – Whole modelling chain designed in Pasquier et al. 2023 for the calculation of time-averaged concentration maps. The five uncertain parameters considered are indicated under each block.

variable	distribution	bounds	parameters
θ	truncated normal	[5; 10] [rad]	$\mu = 0.7, \sigma = 0.5$
U_∞	truncated normal	[-0.35; 1.75] [m/s]	$\mu = 8, \sigma = 2$
q	truncated skewed normal	[100; 500] [vehicle/h]	$\xi = 450, \omega = 100, \alpha = -3$
β	uniform	[0; 1] [-]	-
ν_{max}	uniform	[30; 50] [km/h]	-

Table 1 – Probability distributions and associated parameters for the five uncertain pollutant dispersion variables

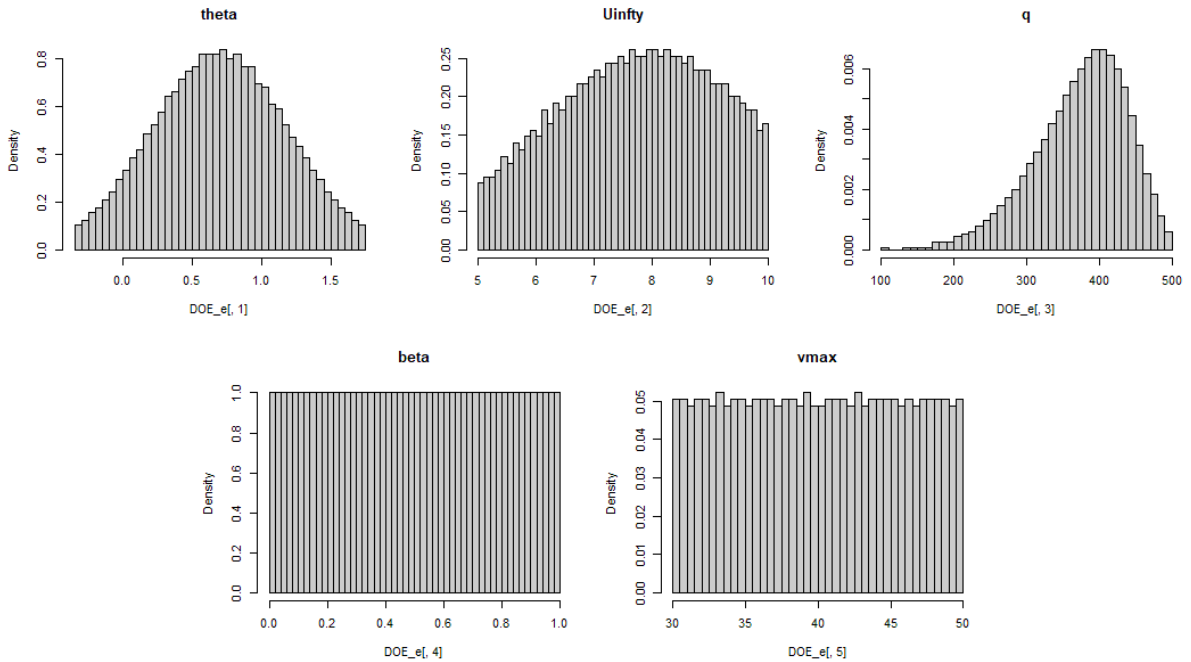


Figure 3 – Latin hypercube samples of size 10^4 of the probability distributions chosen for the five uncertain variables considered in the pollutant dispersion case.

to quantitatively evaluate the predictivity of the POD-GPR metamodel, which reaches about 86%. The lack of predictivity is related to residual noise in the numerical simulation results due to the difficult statistical convergence of the time-averaged concentration

fields in CFD simulations.

Based on this cheap metamodel, we want to perform a global sensitivity analysis for concentration maps, which can be seen as instances of a functional output from $\mathcal{D} \subset \mathbb{R}^2$ to \mathbb{R} . Initially, sensitivity analysis focused primarily on models with scalar inputs and outputs. Standard methods such as Morris analysis (Morris 1991) and the Sobol' index (Sobol 2001) were widely used to evaluate the variation of outputs in response to variations in input parameters. Morris analysis, for example, relies on sampling scenarios in which each parameter is modified one at a time to assess its effect on the model's output. However, as modelling and complex systems have evolved, the applications of sensitivity analysis have expanded to include more complex inputs and outputs. Inputs can be vectors, categorical data, time series and more. Outputs can be multi-dimensional, non-scalar or even probability distributions. To meet these needs, new advanced sensitivity analysis techniques have emerged. For example, Sobol' indices have been generalised for vectorial and functional outputs (Gamboa et al. 2013). However, the latter uses a truncated expansion of the functional output to then reuse the proposed index on vector outputs, which implies a loss of information. In Perrin et al. 2021, the authors also propose a method to conduct a global sensitivity analysis in the case of spatial output through functional principal component analysis. Saint-Geours et al. 2014 also perform sensitivity analysis for spatial output by computing maps of sensitivity indices. In Fort et al. 2021, the authors propose universal sensitivity indices that can be computed in any metric output space. The more recent indices based on the Hilbert-Schmidt Independence Criterion (HSIC) were also initially defined for scalar and vector inputs and outputs (Gretton et al. 2005 and Da Veiga 2013), but are much more permissive and can be extended to arbitrary outputs as long as a kernel is defined on the output space. For example, HSIC-based indices have been used for probability distribution outputs in Da Veiga 2021.

The aim of this paper is to propose and compare different methods for carrying out a sensitivity analysis of a model that produces two-dimensional maps. We first adopt a pointwise approach, specifically a sensitivity analysis technique using Sobol' indices computed at each point of the map. The resulting Sobol' index maps can be interpreted directly or used to compute *generalised* Sobol' indices, similar to those proposed in Gamboa et al. 2013, where the discretised map is considered as a vector. The other proposed approach analyses the spatial output without discretisation and instead aims to establish an index that quantifies the influence of each input on the overall spatial output. To achieve this, three new SA indices based on sets are introduced and adapted to maps. Each index is tested on the POD-GPR metamodel.

The paper is structured as follows. Section 2 proposes a pointwise method for performing sensitivity analysis on maps, where section 2.1 establishes essential notations and section 2.2 defines Sobol' indices at each point of the maps and proposes *generalised* indices that aggregate all pointwise indices. Section 3 is devoted to three new set-based SA indices adapted to maps. Section 3.1 provides the necessary notations for dealing with sets, while section 3.2 deals with sensitivity analysis based on random set theory. Section 3.3 adapts universal indices from Fort et al. 2021 to set-valued outputs, and in section 3.4 we use indices based on kernels and the Hilbert-Schmidt independence criterion in the context of spatial outputs, as done in Fellmann et al. 2023. Section 4 is devoted to a comparative study of the various sensitivity indices introduced. Finally, our conclusion (section 5) ties these sections together and summarises our main findings.

2 Pointwise sensitivity analysis for maps

2.1 Notations

Let us define a map-valued model Φ by

$$\begin{aligned} \mathcal{U} &\longrightarrow \mathcal{F}(\mathcal{D}, \mathbb{R}) \\ \Phi : \mathbf{u} &\mapsto \Phi_{\mathbf{u}} : x_1, x_2 \mapsto \Phi_{\mathbf{u}}(x_1, x_2) \end{aligned}$$

with $\mathcal{D} \subset \mathbb{R}^2$, $\mathcal{U} = \mathcal{U}_1 \times \dots \times \mathcal{U}_p \subset \mathbb{R}^p$ and $\mathcal{F}(\mathcal{D}, \mathbb{R})$ the space of functions from \mathcal{D} to \mathbb{R} . As in many sensitivity analysis framework, the inputs are assumed to be a random vector $\mathbf{U} = (U_1, \dots, U_p) \in \mathcal{U}$ with independent components. It is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with known distributions $\mathbb{P}_{\mathbf{U}} = \bigotimes_{i=1}^p \mathbb{P}_{U_i}$.

In the case of the POD-GPR metamodel of pollutant concentration maps, we have

- $\mathcal{U} = [5; 10] \times [-0.35; 1.75] \times [100; 500] \times [0; 1] \times [30; 50]$ and the \mathbb{P}_{U_i} are given in Table 1.
- $\mathcal{D} = [0; 550] \times [0; 400]$
- $\Phi_{\mathbf{u}}(x_1, x_2)$ is the pollutant concentration at position $x_1, x_2 \in \mathcal{D}$ in the physical domain when the uncertain parameters are equal to $\mathbf{u} \in \mathcal{U}$.

2.2 Sobol' indices maps and generalized Sobol' indices

First, we propose to analyse the sensitivity of the concentration fields to the uncertain inputs by calculating the Sobol' indices of the spatial maps. In other words, for each $x_1, x_2 \in \mathcal{D}$, we compute the Sobol' indices of the output concentration $\Phi_{\mathbf{U}}(x_1, x_2)$.

At each point x_1, x_2 , from the ANOVA decomposition of $\mathbf{u} \mapsto \Phi_{\mathbf{u}}(x_1, x_2)$, a variance decomposition of $\Phi_{\mathbf{U}}(x_1, x_2)$ can be obtained. The first-order Sobol' indices are defined from the first terms of the ANOVA decomposition, each of which depends on a single uncertain variable. The first-order index corresponding to the i -th uncertain variable is defined as

$$S_i(x_1, x_2) = \frac{\text{Var}[\mathbb{E}[\Phi_{\mathbf{U}}(x_1, x_2)|U_i]]}{\text{Var}[\Phi_{\mathbf{U}}(x_1, x_2)]}.$$

These indices indicate the proportion of the total variance that can be attributed to the individual contribution of each uncertain variable U_i .

Such first-order indices are computed using the standard partial variance estimator recalled in Saltelli et al. 2010 and defined as

$$\text{Var}[\mathbb{E}[\Phi_{\mathbf{U}}(x_1, x_2)|U_i]] \approx \frac{1}{N} \sum_{j=1}^N \Phi_{\mathbf{U}'^{(j)}}(x_1, x_2) (\Phi_{\tilde{\mathbf{U}}^{i(j)}}(x_1, x_2) - \Phi_{\mathbf{U}^{(j)}}(x_1, x_2)), \quad (1)$$

where $(\mathbf{U}^{(j)}, \mathbf{U}'^{(j)})_{j=1}^N$ are independent and identically distributed (iid) samples of $(\mathbf{U}, \mathbf{U}')$, with \mathbf{U}' being an independent copy of \mathbf{U} . The matrix $\tilde{\mathbf{U}}^i$ is the matrix \mathbf{U} with the i -th column replaced by that of \mathbf{U}' . The denominator $\text{Var}[\Phi_{\mathbf{U}}(x_1, x_2)]$ is estimated using the classical unbiased sample variance estimator. The indices are estimated for each location $\{x_1^{(j)}, x_2^{(j)}\}_{j=1}^m$ of a discretised computational grid composed of m nodes to obtain two-dimensional first-order sensitivity maps, as shown in Figure 4.

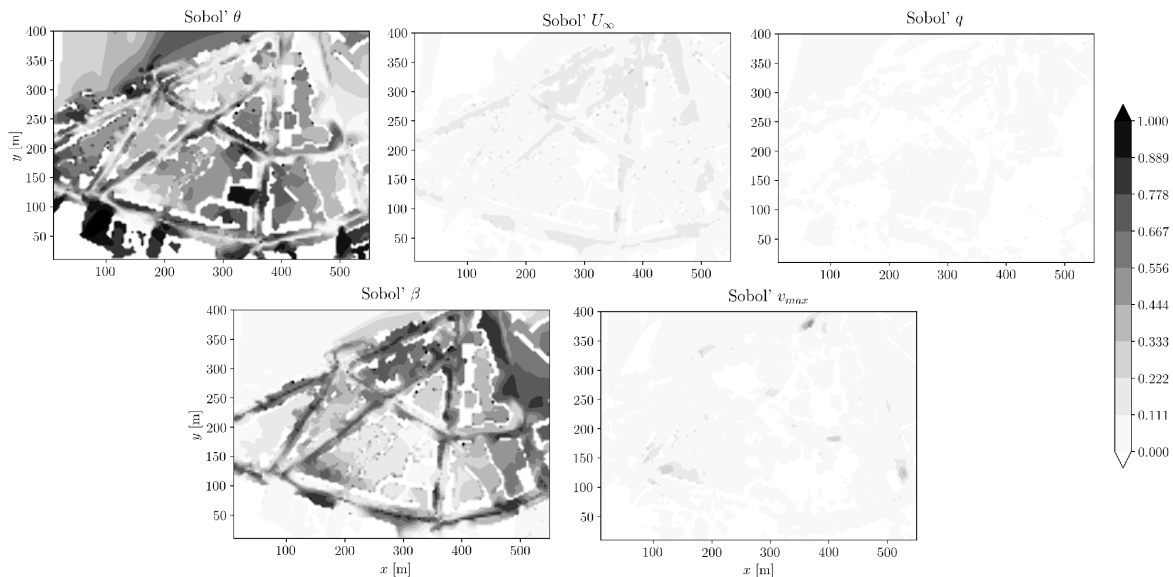


Figure 4 – Two-dimensional maps of pointwise first-order Sobol' indices for time-averaged pollutant concentration, estimated by quasi-Monte-Carlo using 2^{10} sample points and Halton sequences.

The information provided by such two-dimensional maps can be aggregated into scalar values by weighting the grid values of the first-order Sobol' indices with the corresponding estimated variances. They correspond to the vectorial Sobol' indices defined in Gamboa et al. 2013. Such *generalised* Sobol' indices are defined as follows:

$$S_i^{\text{gen}} := \sum_{j=1}^m w_j S_i(x_1^{(j)}, x_2^{(j)}) \quad \text{with} \quad w_j = \frac{\text{Var}[\Phi_{\mathbf{U}}(x_1^{(j)}, x_2^{(j)})]}{\sum_{k=1}^m \text{Var}[\Phi_{\mathbf{U}}(x_1^{(k)}, x_2^{(k)})]}.$$

S_{θ}^{gen}	$S_{U_{\infty}}^{\text{gen}}$	S_q^{gen}	S_{β}^{gen}	$S_{v_{max}}^{\text{gen}}$
5.48×10^{-1}	4.47×10^{-2}	7.56×10^{-4}	3.22×10^{-1}	1.73×10^{-2}

Table 2 – Estimated generalised first-order Sobol' indices for each uncertain variable from two-dimensional sensitivity maps, using 2^{12} sample points from Halton sequences.

From the results shown in Figure 4 and 5 it can be seen that the wind angle and the rate of diesel engines in the fleet contribute to the major part of the variance of the output concentration field. The generalised indices reach approximately 55% and 32% for these two variables respectively, while the other indices fall below 10%, and in particular the value of S_q^{gen} is very close to zero, showing that the traffic volume q is not influential in this study. The two-dimensional Sobol' maps highlight the fact that the sensitivity to the two dominant variables varies depending on the location in the urban neighbourhood: thus, if we choose to focus our analysis on a specific part of the domain - e.g. a specific building such as a hospital or a primary school - we may end up with a different hierarchy among the variables. We also carried out a second analysis (the results of which are not shown here) by fixing θ and β to their mean values and evaluating the Sobol' maps for the three remaining variables. This showed that the wind intensity dominates over the speed limit (with a generalised index of about 70% for the former and 30% for the latter)

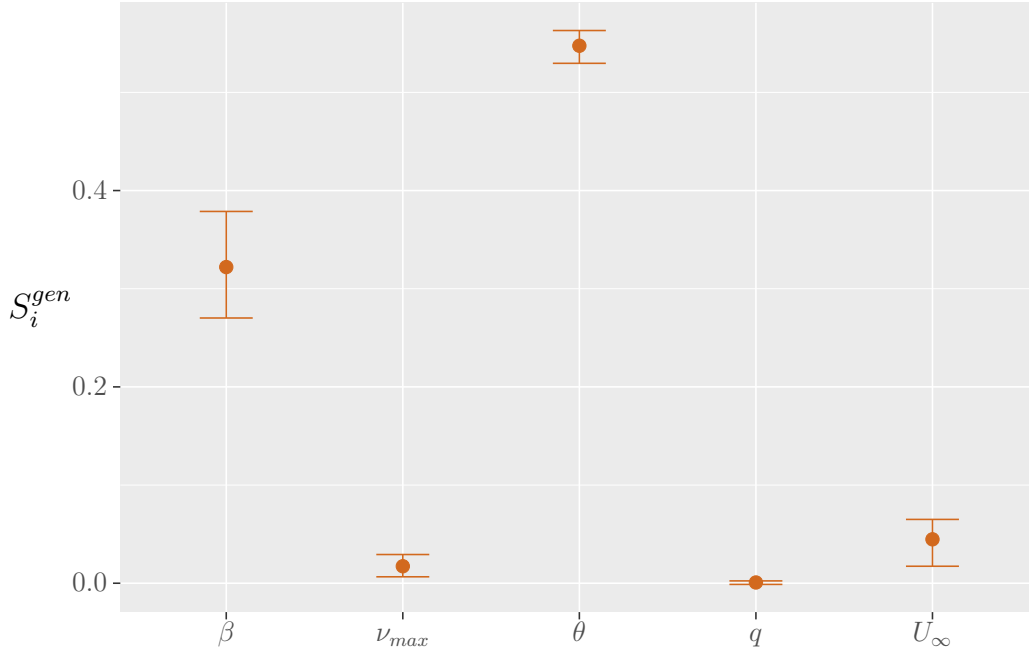


Figure 5 – Estimated generalised first-order Sobol’ indices for each uncertain variable from two-dimensional sensitivity maps, using 2^{12} sample points from Halton sequences and 100 bootstrap resamples to construct 95% confidence intervals.

and the traffic volume remains negligible. We can also observe that the indices sum to about 95%, which means that there are almost no interactions between the inputs.

Although this sensitivity analysis is performed point by point of the pollutant concentration map, it allows a first screening of the influencing variables and, in particular, generalised indices allow a quantitative ranking of the inputs. The next section presents another approach that considers the output as a subset of \mathbb{R}^3 and performs a sensitivity analysis for such outputs.

3 New sensitivity analysis indices for sets

In this section we propose three types of indices designed to perform sensitivity analysis for sets, which we use on map-valued outputs. The first section 3.1 introduces the notations to see map-valued models as examples of set-valued models. Then, three sections are dedicated to introduce the three indices. In section 3.2, we propose new indices based on Vorob’ev’s theory of random sets (Molchanov 2005). In section 3.3 we adapt the universal indices of Fort et al. 2021 to sets. Finally, in Section 3.4, we present indices based on the Hilbert-Schmidt Independence Criterion (HSIC) in the context of set-valued output, as done in Fellmann et al. 2023.

3.1 Notations

We want to perform a sensitivity analysis of a non-discretised map-valued model Φ . To do this, we propose to consider Φ as a set-valued model Ψ defined by:

$$\Psi : \mathbf{u} \mapsto \Gamma_{\mathbf{u}} = \{(x_1, x_2, c) \in \mathcal{D} \times \mathcal{C}, c - \Phi_{\mathbf{u}}(x_1, x_2) \leq 0\}$$

where $\mathcal{C} = [c_{min}, c_{max}]$ and $\mathcal{L}(\mathcal{X})$ is the space of Lebesgue measurable subsets of $\mathcal{X} = \mathcal{D} \times \mathcal{C}$. $\Gamma = \Psi(U)$ is then a random set that depends directly on the inputs U_i whose effect is to be measured.

For estimation purposes, \mathcal{X} is discretised with a grid of m points denoted $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)})$. This allows the volume $\lambda(A)$ of $A \subset \mathcal{X}$ to be estimated as

$$\hat{\lambda}_m(A) = \frac{1}{m} \sum_{k=1}^m \mathbb{1}_A(\mathbf{x}^{(k)}).$$

Let also $(\mathbf{U}^{(i)}, \Gamma^{(i)}), i = 1, \dots, n$ be an independent and identically distributed (i.i.d.) sample of (\mathbf{U}, Γ) with $\Gamma = \Psi(\mathbf{U})$.

3.2 Sensitivity analysis based on random set theory

Working with sets instead of scalars or vectors makes the study of random elements much more difficult. For example, expectations and variance are not easy to define. In Molchanov 2005, a complete theory of random sets is developed with definitions of expectations, medians and deviations of random sets. In particular, based on the Vorob'ev median and deviation, we propose to define indices inspired by Sobol' indices but adapted to random sets.

When dealing with a scalar output Y , first-order Sobol' indices S_i can be defined by :

$$S_i = 1 - \frac{\mathbb{E}[\text{Var}[Y|U_i]]}{\text{Var} Y}.$$

Here we want to define similar indices, but for set-valued outputs. To do this, we propose to replace the variance with the Vorob'ev median deviation, which is a possible adaptation of the median deviation to random sets. This requires defining the Vorob'ev median and the Vorob'ev median deviation, but also defining a Vorob'ev conditional median deviation.

Definition 3.1 (Vorob'ev median). *The Vorob'ev median $Q_{0.5}(\Gamma)$ of a random set Γ is defined as the set of points whose probability of being in Γ is at least 0.5, i.e.*

$$Q_{0.5}(\Gamma) = \{x \in \mathcal{X}, \mathbb{P}(x \in \Gamma) \geq 0.5\}$$

Using the Vorob'ev median, the Vorob'ev Median Deviation is defined as the Vorob'ev deviation between the random set and its median. We also propose conditional version of the Vorob'ev Median Deviation.

Definition 3.2. *The Vorob'ev Median Deviation of a random set Γ , denoted $\text{VMD}(\Gamma)$, is defined by:*

$$\text{VMD}(\Gamma) = \mathbb{E}[\lambda(\Gamma \Delta Q_{0.5})].$$

where Δ is the symmetric difference. The Vorob'ev conditional median deviation is the random variable $\text{VMD}(\Gamma|U_i)$ defined by:

$$\text{VMD}(\Gamma|U_i) = \mathbb{E}[\lambda(\Gamma\Delta Q_{0.5}^i)|U_i]$$

where $Q_{0.5}^i$ is the Vorob'ev conditional median defined by:

$$Q_{0.5}^i = \{\mathbf{x} \in \mathcal{X}, \mathbb{P}(\mathbf{x} \in \Gamma|U_i) \geq 0.5\}.$$

We can now define indices that describe the part of the Vorob'ev median deviation of a random set that is due to an input U_i .

Definition 3.3. Using the previous definition, we define a sensitivity index S_V^i by:

$$\begin{aligned} S_V^i &:= 1 - \frac{\mathbb{E}[\text{VMD}(\Gamma|U_i)]}{\text{VMD}(\Gamma)} \\ &= 1 - \frac{\mathbb{E}[\lambda(\Gamma\Delta Q_{0.5}^i)]}{\mathbb{E}[\lambda(\Gamma\Delta Q_{0.5})]} \end{aligned}$$

This index quantifies the effect of U_i on the output Γ , since it's zero if Γ and U_i are independent (simply as $Q_{0.5}^i = Q_{0.5}$ if Γ and U_i are independent). However, as first-order Sobol' indices, the converse doesn't hold. We show in Proposition 3.1 that it is between zero and one, as desired for sensitivity indices. The interpretation is close to a part of the total deviation explained by an input. However, there is no decomposition of the indices and they could sum to more than 1.

Proposition 3.1. $0 \leq S_V^i \leq 1$

Proof. It is clear that $S_V^i \leq 1$. To show that it's non-negative, we use the tower property of the conditional expectation on the numerator and the denominator :

$$\frac{\mathbb{E}[\lambda(\Gamma\Delta Q_{0.5}^i)]}{\mathbb{E}[\lambda(\Gamma\Delta Q_{0.5})]} = \frac{\mathbb{E}\mathbb{E}[\lambda(\Gamma\Delta Q_{0.5}^i)|U_i]}{\mathbb{E}\mathbb{E}[\lambda(\Gamma\Delta Q_{0.5})|U_i]}.$$

Then we show that

$$\mathbb{E}[\lambda(\Gamma\Delta Q_{0.5}^i)|U_i] \leq \mathbb{E}[\lambda(\Gamma\Delta Q_{0.5})|U_i]$$

almost surely, using the fact that the Vorob'ev median minimises the deviation for any measurable sets (see Proposition 2.4 of Molchanov 2005). Finally by taking the expectation we conclude that S_V^i is non-negative. \square

The estimation of S_V^i is done by a double loop, since we are not aware of any cheaper estimator. Let $U_i^{(j)}$, $j \in (1, n)$ be n independent copies of a given input U_i , the conditional median $Q_{0.5}^i(U_i^{(j)}) = \{x \in \mathcal{X}, \mathbb{P}(x \in \Gamma|U_i = U_i^{(j)}) \geq 0.5\}$ is estimated by independently drawing another sample $U_{-i}^{(j,l)}$, $l \in (1, n)$ of U_{-i} and evaluating the output sets $\Gamma_l(U_i^{(j)}) = \Psi(U_i^{(j)}, U_{-i}^{(j,l)})$. The estimator of the conditional Vorob'ev median can then be derived:

$$\hat{Q}_{0.5}^i(U_i^{(j)}) := \{x \in \mathcal{X}, \frac{1}{n} \sum_{l=1}^n \mathbb{1}_{\Gamma_l(U_i^{(j)})}(x) \geq 0.5\}$$

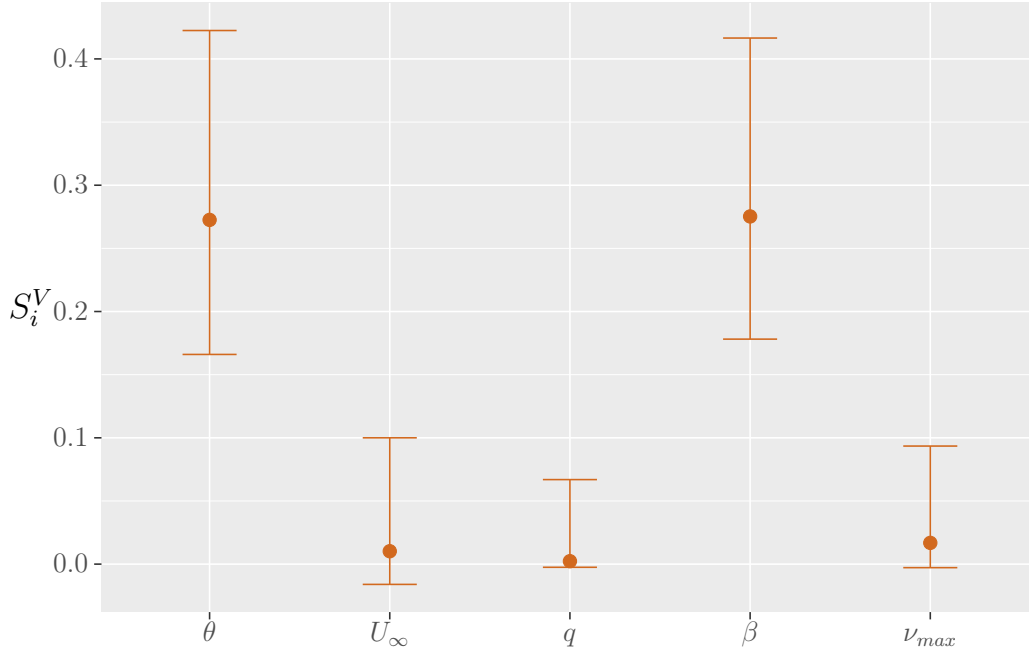


Figure 6 – Estimation of S_V^i for each input with 1024 model evaluations and a confidence interval estimated with 100 bootstrap resamples.

Then S_i^V is estimated by:

$$\tilde{S}_i^V = 1 - \frac{\sum_{j=1}^n \hat{\lambda}_m(\Gamma^{(j)} \Delta \hat{Q}_{0.5}^i(U_i^{(j)}))}{\sum_{j=1}^n \hat{\lambda}_m(\Gamma^{(j)} \Delta \hat{Q}_{0.5})}. \quad (2)$$

where $\Gamma^{(j)} = \Psi(U_i^{(j)}, U_{-i}^{(j)})$.

The index is estimated by $n = 32$ model evaluations per loop, resulting in a total of $n^2 = 1024$ model evaluations. Confidence intervals are obtained by bootstrapping with 100 resamples. The results are shown in Figure 6.

As these indices do not have a decomposition, it is difficult to justify a comparison between the value of each index. The only possible interpretation is that an index far from zero means that the input has an effect on the output. Here, the intervals of β and θ do not include zero, which means that they have an effect. The other three, however, are close to 0, implying that they may not be influential. The confidence intervals are also wide, which means that in another application we might fail to detect influential inputs. To become significant, additional model evaluations or a different estimation method is required. However, methods such as pick and freeze or rank-based estimation do not appear to be applicable in this scenario, as the quantity being estimated is not strictly a conditional variance.

3.3 Universal sensitivity indices

In this second part, we propose to use *universal* indices from Fort et al. 2021. In the latter paper, the authors define indices that can be applied to any metric output space. Using them with set-valued outputs, we define the following indices:

Definition 3.4.

$$S_i^{Univ}(\gamma, \mathbb{Q}) = \frac{\int_0^1 \text{Var} \mathbb{E} [\lambda(\Gamma \Delta \gamma_a) \mid U_i] d\mathbb{Q}(a)}{\int_0^1 \text{Var} (\lambda(\Gamma \Delta \gamma_a)) d\mathbb{Q}(a)}.$$

where $(\gamma_a)_{a \in [0,1]} \subset \mathcal{X}$ is a collection of test sets parameterised by $a \in [0,1]$ and \mathbb{Q} is a probability distribution on $[0,1]$ to choose.

The idea behind these indices is to quantify the effect of inputs on multiple scalar transformations of the set-valued output, rather than on the set-valued output itself. Specifically, we measure this variability by the volume of the symmetric difference between the output set and a collection of test sets.

Similar to Sobol' indices, the contribution of each input and their interactions can be decomposed to rank the inputs. Indeed, the decomposition can be obtained by taking the Sobol-Hoeffding decomposition of $\text{Var} (\lambda(\gamma \Delta \gamma_a))$ for each a and then summing for 0 to 1 with respect to the measure \mathbb{Q} . An independence test is not accessible, mainly because the family of all transformations does not necessarily characterise the whole distribution of the set-valued output. However, screening can still be done, either by keeping only the inputs above a certain threshold, or by keeping the first inputs in the ranking.

After selecting the test sets and the law \mathbb{Q} , these indices can be estimated in a similar way to Sobol's indices. We use a rank-based estimator, which is recalled in Fort et al. 2021. Plugging in the volume estimation, the estimator \widehat{S}_i^{Univ} of S_i^{Univ} is given by the ratio between

$$\begin{aligned} \widehat{S}_{num} = & \frac{1}{N_a} \sum_{l=1}^{N_a} \left[\frac{1}{n} \sum_{j=1}^n \left(\widehat{\lambda}_m(\Gamma^{(j)} \Delta \gamma_{a_l}) \right) \left(\widehat{\lambda}_m(\Gamma^{(N_i(j))} \Delta \gamma_{a_l}) \right) \right] \\ & - \frac{1}{N_a} \sum_{l=1}^{N_a} \left[\frac{1}{n} \sum_{j=1}^n \left(\widehat{\lambda}_m(\Gamma^{(j)} \Delta \gamma_{a_l}) \right) \right]^2 \end{aligned} \quad (3)$$

and,

$$\widehat{S}_{den} = \frac{1}{N_a} \sum_{l=1}^{N_a} \left[\frac{1}{n} \sum_{j=1}^n \left(\widehat{\lambda}_m(\Gamma^{(j)} \Delta \gamma_{a_l}) \right)^2 \right] - \frac{1}{N_a} \sum_{l=1}^{N_a} \left[\frac{1}{n} \sum_{j=1}^n \left(\widehat{\lambda}_m(\Gamma^{(j)} \Delta \gamma_{a_l}) \right) \right]^2, \quad (4)$$

where (a_l) is an iid sample of the law \mathbb{Q} and $N_i(j)$ is the index in the sample (U_i^l) that comes after U_i^j when (U_i^l) is sorted in ascending order (see Gamboa et al. 2022 for details).

We compute the indices with $n = 1000$ model evaluations and $N_a = 100$ test sets. To remain general, \mathcal{X} is rescaled to $[0,1]^3$ which is centered on $x^0 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. We test four families of test sets defined on $[0,1]^3$:

- Centered balls: $\gamma_a = B(x^0, a)$ with $\mathbb{Q} \sim \mathcal{U}([0, \frac{1}{2}])$
- Centered squares: $\gamma_a = \{x \in [0,1]^3, \|x - x^0\|_\infty \leq a\}$ with $\mathbb{Q} \sim \mathcal{U}([0, \frac{1}{2}])$
- Slides along the i -th dimension: $\gamma_a^i = \{x \in [0,1]^3, x^i \leq a\}$ with $\mathbb{Q} \sim \mathcal{U}([0,1])$
- Vorob'ev quantiles: $\gamma_a = \{x \in [0,1]^3, \mathbb{P}(x \in \Gamma) \geq a\}$ with $\mathbb{Q} \sim \mathcal{N}(\frac{1}{2}, \sigma^2)$

The slides are used with $i = 3$ to get levels of concentration test sets. Confidence intervals are obtained by bootstrapping with $B = 100$ resamples. We use bootstrap samples of size $0.8n = 800$ without replacement because replacement introduces high bias in rank-based estimation. The intervals are adjusted by a correction factor. The results are displayed in Figure 7.

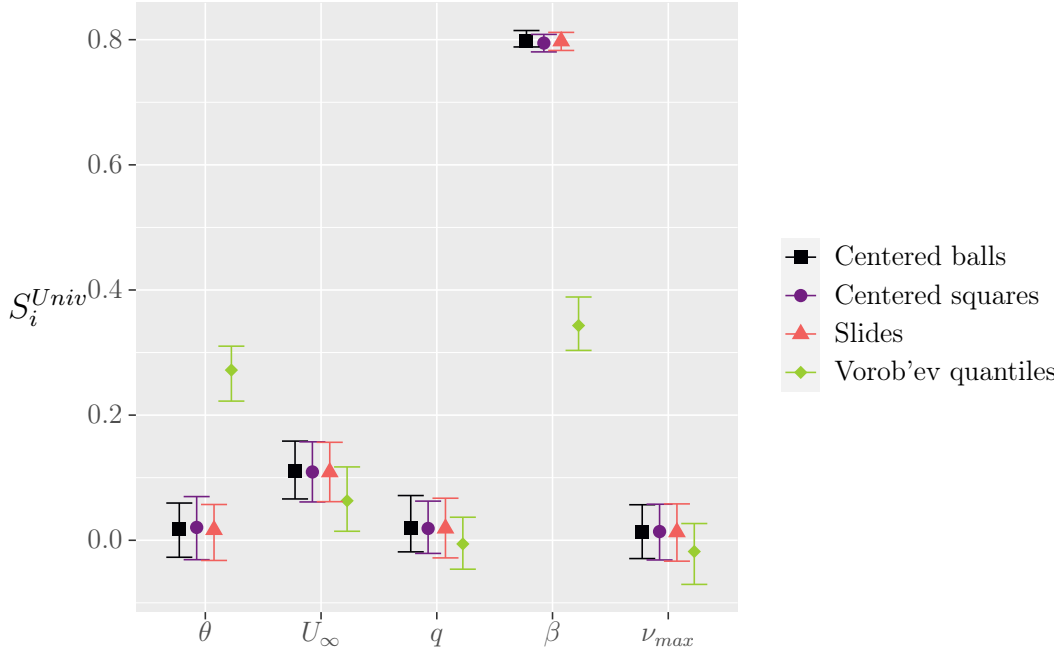


Figure 7 – Estimation of the universal indices S_i^{Univ} for each input and for four different test sets and $N_a = 100$, with 1000 model evaluations. Confidence intervals are obtained with 100 bootstrap samples

The results are comparable when using centered balls, centered squares, and slices. β is the most dominant input with an index value of about 80%. The value of U_∞ is approximately 10%, and all other inputs have minimal impact. In contrast, when Vorob'ev quantiles are the test sets, β and θ are the most important inputs, and U_∞ also has some influence to a lesser extent. The output is clearly not influenced by either q or ν_{max} . This difference shows that the choice of test sets is crucial, as it has a strong influence on the values of the indices. When the law \mathbb{Q} of Vorob'ev quantiles is approximately 0.5, their geometry is similar to Γ 's realisations. Therefore, changes in Γ can be accurately detected by measuring the volume of the symmetric difference. Conversely, it may not be fully possible to detect all variations of the random set by comparing realisations of Γ with simple sets as squares, cubes or slices.

3.4 HSIC-based sensitivity analysis

In this section, we propose to define kernel-based sensitivity indices with set-valued outputs, as the use of kernels makes such indices very permissive in the type of considered outputs (or inputs).

Sensitivity analysis, based on measures of dependence, consists in examining the influence of an input X_i on an output Y by measuring their mutual dependence. To do this, a distance is calculated between the joint distribution $\mathbb{P}_{X_i, Y}$ and the product of their marginal distributions $\mathbb{P}_{X_i} \otimes \mathbb{P}_Y$. If this distance is zero, then X_i and Y are independent, meaning X_i has no impact on Y . A frequently used measure for dependence is the Hilbert-Schmidt Independence Criterion (HSIC). It operates by selecting appropriate input and output kernels, and subsequently measuring the squared distance between the mean embeddings of $\mathbb{P}_{X_i, Y}$ and $\mathbb{P}_{X_i} \otimes \mathbb{P}_Y$ in the corresponding Reproducing Kernel Hilbert

Space (RKHS), see Gretton et al. 2005 for a proper definition of the tools needed to define the HSIC. This calculation can be performed using just one sample, which is among the reasons why HSICs have gained popularity.

If a *characteristic* kernel is available on any space, meaning a kernel whose mean embedding in the RKHS is injective, then HSIC-based indices can be used. Consequently, their application is appropriate for sensitivity analysis of set-valued models. Fellmann et al. 2023 introduced a kernel k_{set} defined on sets, which is proved to be characteristic. It is then used to define HSIC-based indices on sets. The latter definition requires to have

- a characteristic kernel on $\mathcal{L}(\mathcal{X})$. The proposed kernel in Fellmann et al. 2023 is

$$k_{set}(\gamma_1, \gamma_2) = \exp\left(-\frac{\lambda(\gamma_1 \Delta \gamma_2)}{2\sigma^2}\right) \quad \forall \gamma_1, \gamma_2 \subset \mathcal{X},$$

where σ^2 is a normalization constant.

- p characteristic kernels K_i on the input space \mathcal{U}_i . As shown in Da Veiga 2021, an ANOVA-like decomposition of HSIC exists if the input kernels are ANOVA, i.e. they can be written as $1 + k_i$, where k_i is a kernel whose induced RKHS consists of zero-mean functions. Sobolev kernels are the most commonly used kernels that fulfil these requirements. Furthermore, classic characteristic kernels can be modified to become ANOVA, as described in Ginsbourger et al. 2016.

Definition 3.5. *With the previous kernels, HSIC-based indices on sets are defined by*

$$\forall i \in \{1, \dots, p\}, S_i^{H_{set}} := \frac{H_{set}(U_i, \Gamma)}{H_{set}(\mathbf{U}, \Gamma)},$$

where

$$\begin{aligned} H_{set}(U_i, \Gamma) &= \mathbb{E}[(K_i(U_i, U_i') - 1)k_{set}(\Gamma, \Gamma')], \\ H_{set}(\mathbf{U}, \Gamma) &= \mathbb{E}[(K(\mathbf{U}, \mathbf{U}') - 1)k_{set}(\Gamma, \Gamma')], \end{aligned}$$

with

$$K = \bigotimes_{j=1}^p K_j,$$

and (\mathbf{U}', Γ') an independent copy of (\mathbf{U}, Γ) .

Since these indices are zero if and only if the considered input U_i and the output Γ are independent, they are suitable for screening. Using this result, p-values associated with the independent tests can be calculated to screen the inputs. Besides, we choose the input kernels K_i as ANOVA kernels so that the indices $S_i^{H_{set}}$ have an ANOVA decomposition as done in Fellmann et al. 2023. Subsequently, the inputs can be ranked by their impact based on each index's value.

The indices can be estimated using:

$$\widehat{H}_{set}(U_i, \Gamma) = \frac{2}{n(n-1)} \sum_{j < l}^n \left(K_i(U_i^{(j)}, U_i^{(l)}) - 1 \right) \exp\left(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \hat{\lambda}_m(\Gamma^{(j)} \Delta \Gamma^{(l)})\right). \quad (5)$$

We compute the indices for each of the five inputs using a sample size of $n = 1000$. The results for five different input ANOVA kernels are compared using the following kernels:

- the Sobolev kernel of order 1, $k_{sob}(x, y) = 1 + (x - \frac{1}{2})(y - \frac{1}{2}) + \frac{1}{2}[(x - y)^2 - |x - y| + \frac{1}{6}]$

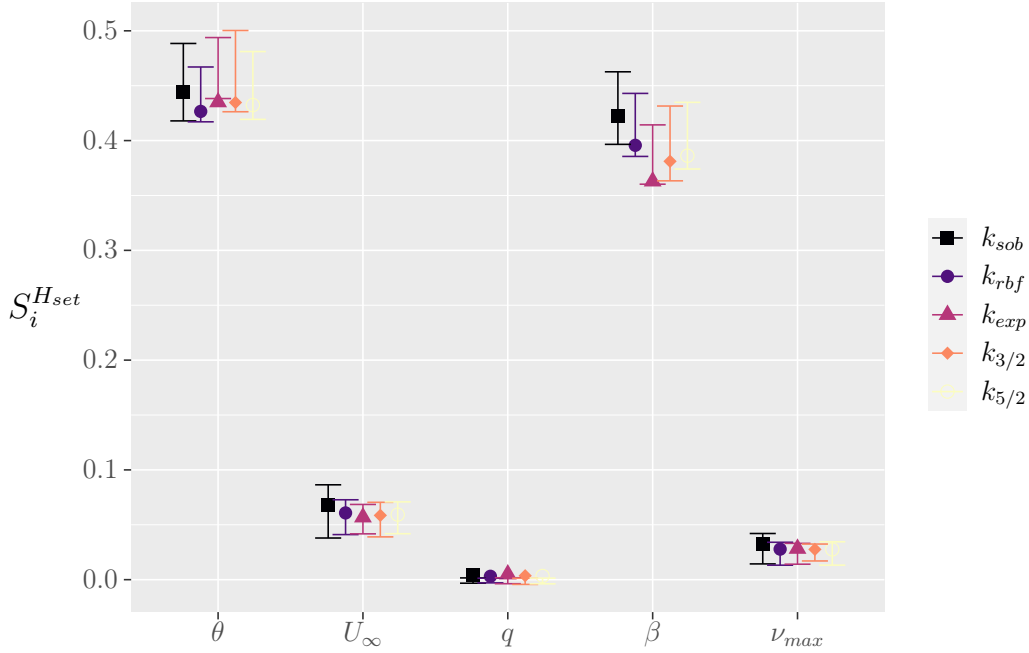


Figure 8 – Estimation of $S_i^{H_{set}}$ for five input kernels, 1000 model evaluations. Confidence intervals are obtained by bootstrap with 100 resamples

- the Gaussian kernel, $k_{rbf}(x, y) = e^{-\frac{1}{2}\left(\frac{x-y}{\sigma}\right)^2}$ with $\sigma > 0$,
- the Laplace kernel, $k_{exp}(x, y) = e^{-\frac{|x-y|}{h}}$ with $h > 0$,
- the Matérn 3/2, $k_{3/2}(x, y) = \left(1 + \sqrt{3}\frac{|x-y|}{h}\right) e^{-\sqrt{3}\frac{|x-y|}{h}}$ with $h > 0$,
- the Matérn 5/2, $k_{5/2}(x, y) = \left(1 + \sqrt{5}\frac{|x-y|}{h} + \frac{5}{3}\frac{|x-y|^2}{h^2}\right) e^{-\sqrt{5}\frac{|x-y|}{h}}$ with $h > 0$.

The Sobolev kernel is already ANOVA, whereas four additional kernels need adaptations to become ANOVA (see Ginsbourger et al. 2016 for the detailed transformation). Confidence intervals are estimated using bootstrap methods with $B = 100$ resamples. The value of σ is taken as the empirical median of $\hat{\lambda}_m(\Gamma^{(j)}\Delta\Gamma^{(l)})$ for $j < l$. Figure 8 shows our results. We observe that the choice of the input kernel does not significantly affect the values of the indices: regardless of the kernel chosen, β and θ remain the most influential variables (with about 40% each) and the other three inputs the least influential (with about 10%). The input ranking can be used for screening purposes. For instance, one may decide to keep only the top three most influential inputs. However, using HSIC-based indices guarantees trustworthy screening by computing the p-values of independence tests between an input and an output. This is done in Table 3, where the p-values for each input are computed by asymptotic estimation (Song et al. 2007) for each input kernel. If the p-value exceeds 0.05, then the input has no influence on the output. Specifically, in this case, we observe that only q , for each kernel, has a p-value above 0.05. Additionally, a standard deviation is estimated for q through bootstrap analysis, which is given in the second column. We confirm that despite this deviation, the p-values of q remain above 0.05. For the remaining four inputs, p-values below 10^{-16} were obtained, and standard deviations of the same scale indicate that these four inputs are influential. The chosen input kernel slightly modifies the p-values in this case, but it does not affect the screening results. However, in other test cases, the choice of kernel could impact the results, which

	β	q	θ	U_∞	ν_{max}
k_{sob}	$< 10^{-16}$	0.74 ± 0.6	$< 10^{-16}$	$< 10^{-16}$	$< 10^{-16}$
k_{rbf}	$< 10^{-16}$	0.76 ± 0.6	$< 10^{-16}$	$< 10^{-16}$	$< 10^{-16}$
k_{exp}	$< 10^{-16}$	0.89 ± 0.5	$< 10^{-16}$	$< 10^{-16}$	$< 10^{-16}$
$k_{3/2}$	$< 10^{-16}$	0.82 ± 0.6	$< 10^{-16}$	$< 10^{-16}$	$< 10^{-16}$
$k_{5/2}$	$< 10^{-16}$	0.80 ± 0.6	$< 10^{-16}$	$< 10^{-16}$	$< 10^{-16}$

Table 3 – Asymptotic estimation of the p-values of HSIC-ANOVA-based independence test with five input kernel and 1000 model evaluations and 100 bootstrap samples.

is the primary limitation of the HSIC-based indices. Some research has also been conducted to aid in kernel selection. For instance, in El Amri and Marrel 2021, the authors suggest selecting the kernel that maximises test power. Without further investigation into the impact of kernels, the Sobolev kernel is preferable as it does not require any choice of hyperparameters.

4 Comparison of the indices

In this section we compare the four indices recalled here :

- the generalised first-order Sobol’ indices S_i^{gen} ;
- the Vorob’ev based indices S_i^V ;
- the universal indices S_i^{Univ} computed using Vorob’ev quantiles as test sets;
- the HSIC-based indices S_i^{Hset} using the Sobolev kernel as ANOVA input kernel.

Each index is computed with a budget of about 1000 model evaluations. S_i^{gen} are estimated with the pick and freeze estimator given in (1). S_i^V are estimated using a double loop method as described in (2), with 32 model evaluations in each loop, resulting in 1024 evaluations. The universal indices are evaluated using the rank-based estimation approach described in (3) and (4), with $N_a = 100$ test sets. Lastly, the estimate for S_i^{Hset} is given by (5).

The results are shown in Figure 9. First we check that the four indices give similar results. The most influential inputs are θ and β , and the least influential are U_∞ , q and ν_{max} . However, there are differences in the obtained results between the indices regarding screening and ranking. The HSIC is the only index that allows access to a reliable independence test, which categorizes q as non-influential in screening (according to the p-values given in Table 3). For the other three indices, a possible screening method is to select a threshold under which inputs will be deemed non-influential. For instance, if a 10% threshold is chosen, then U_∞ , q , and ν_{max} would be considered negligible. However, with a smaller threshold, conclusions would be challenging to draw as some confidence intervals overlap. The four significant inputs, identified by HSIC-based tests, can be confidently ranked since their confidence intervals do not overlap. In contrast, the ranking of the remaining inputs varies for each index. Similar to S_i^{Hset} , S_i^V places θ and β at the same level. However, θ is much more influential than β according to the generalised Sobol’ indices, and the opposite is true for the universal indices. The ranking is justified in theory for all indices except for S_i^V , which lacks any decomposition.

Examining the confidence intervals, all four indices remain satisfactory with 1000 model evaluations, but the HSIC-based indices, and to a lesser extent the universal indices,

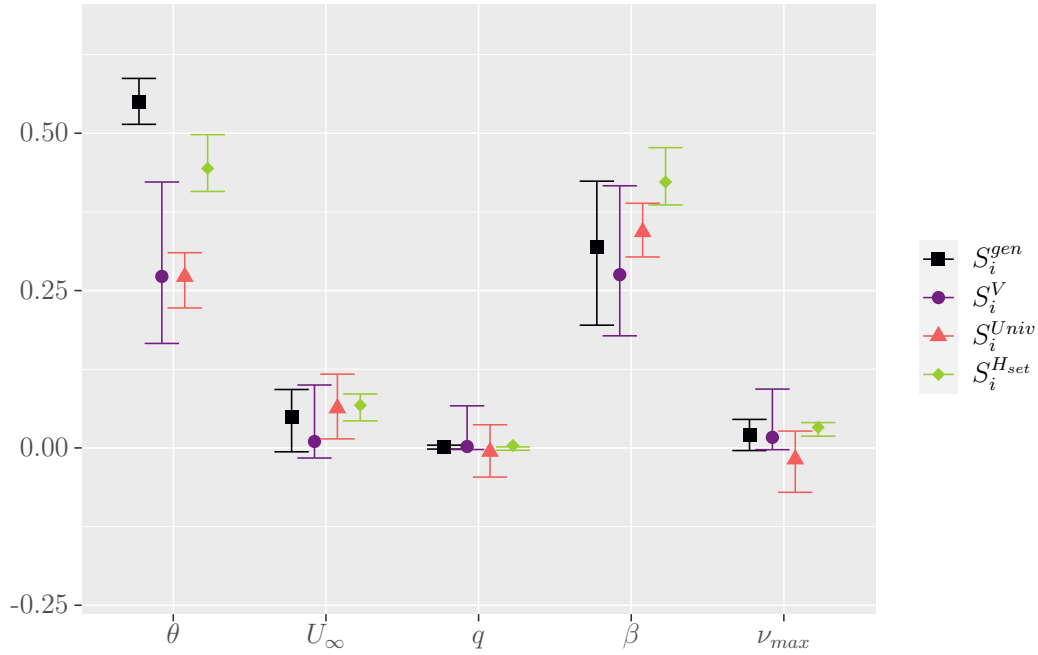


Figure 9 – Comparison of the four indices with a total budget of $n = 1000$ model evaluations. 100 bootstrap sample are used to estimate confidence intervals

have much less variability. To emphasise this disparity, we create an identical plot on Figure 10 with only 100 model evaluations (10 in each loop for S_i^V). This time the HSIC-based indices can be used as they have relatively small confidence intervals. The same conclusion in terms of screening and ranking can be reached with a single 100-sample. The universal indices can also still be used in this case, but their variability makes it possible to have non-positive indices. As for the other two indices, it is important to note that they have not yet converged, rendering them unusable with only 100 model evaluations.

The strengths and limitations of each index are summarised in Table 4.

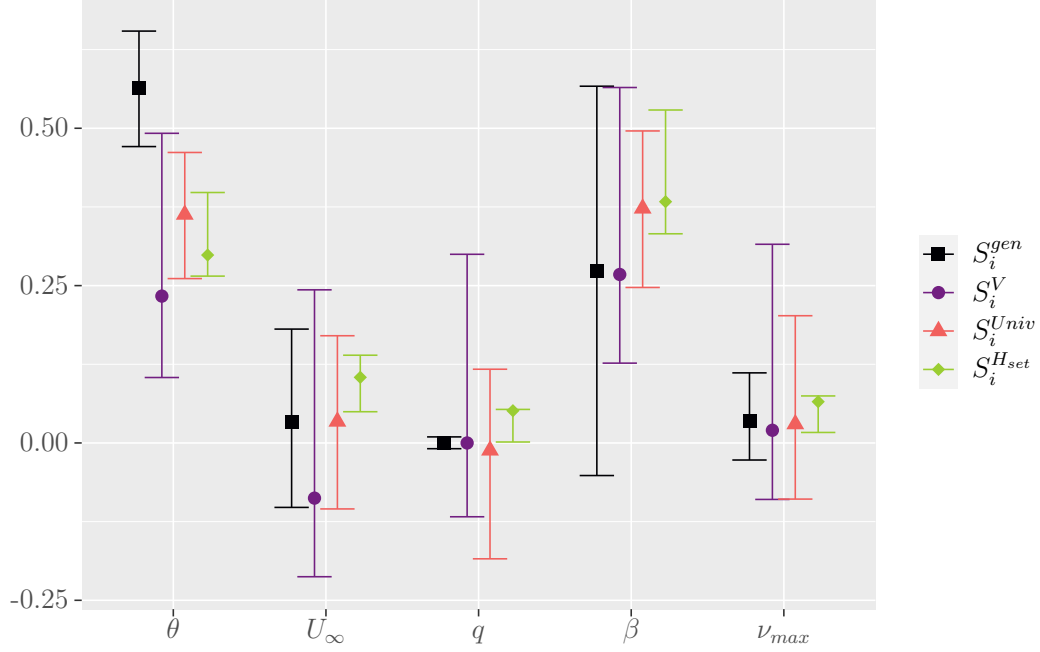


Figure 10 – Comparison of the four indices with a total budget of $n = 100$ model evaluations. 100 bootstrap sample are used to estimate confidence intervals

	Ranking	Screening	Evaluations	Limitations
S_i^{gen}	✓ ANOVA	~ threshold	~ $(p + 1)n$	✗ pointwise influence
S_i^V	✗ no decomposition	✗ no screening method	✗ n^2 (double loop)	✓ no choice to be made
S_i^{Univ}	✓ ANOVA	~ threshold	~ n but big confidence intervals	✗ choice of the test sets
S_i^{Hset}	✓ ANOVA	✓ independence test	✓ n and small confidence intervals	~ choice of kernel ✗ Interactions interpretations

Table 4 – Indices comparison

5 Conclusion

In this paper, we aim to quantify the impact of inputs on a map-valued model, specifically a concentration map model for a realistic urban setting in the greater Paris region (France). To achieve this goal, we propose a sensitivity analysis of the model. As a starting point, we perform a pointwise sensitivity analysis on each point of the map. This produces maps of Sobol’ indices and allows us to observe the influence of each input in each zone of the map. Generalised indices are also derived to synthesise the overall impact of each input. A second approach deals with the model as a whole. Three types of indices, defined for set-valued outputs, are introduced and used with the map-valued outputs. The first is based on quantities from random set theory. The second is based on universal indices that can be used with set outputs. Finally, the last relies on kernel-based sensitivity analysis and, in particular, on HSIC-based indices. A comparison of each index reveals that the HSIC-based index is the most efficient in terms of the number of model evaluations and can be used for both screening and ranking. However, it may not be the most suitable option if we are interested in examining the interactions between inputs. In such situations, universal indices and generalized Sobol’ may be more appropriate.

The introduction of sensitivity indices for sets was initially devised for their application to maps. However, their utility extends beyond maps and can be effectively employed for any model with set-valued outputs, like for example sets of feasible points in mechanical engineering. While we specifically addressed the two-dimensional case to directly handle maps, it is also worth noting that these indices can be used with \mathcal{X} of any finite dimension.

6 Acknowledgment

This research was conducted with the support of the consortium in Applied Mathematics CIROQUO, gathering partners in technological research and academia in the development of advanced methods for Computer Experiments.

References

- Da Veiga, Sébastien (2013). *Global Sensitivity Analysis with Dependence Measures*. arXiv: 1311.2483 [math.ST].
- (Jan. 2021). “Kernel-based ANOVA decomposition and Shapley effects - Application to global sensitivity analysis”. working paper or preprint.
- El Amri, Reda and Amandine Marrel (Oct. 2021). “More powerful HSIC-based independence tests, extension to space-filling designs and functional data”. working paper or preprint.
- Fellmann, Noé, Christophe Blanchet-Scalliet, Céline Helbert, Adrien Spagnol, and Delphine Sinoquet (2023). *Kernel-based sensitivity analysis for (excursion) sets*. arXiv: 2305.09268 [math.ST].
- Fort, Jean-Claude, Thierry Klein, and Agnès Lagnoux (2021). “Global Sensitivity Analysis and Wasserstein Spaces”. In: *SIAM/ASA Journal on Uncertainty Quantification* 9.2, pp. 880–921. DOI: 10.1137/20M1354957.

- Gamboa, Fabrice, Pierre Gremaud, Thierry Klein, and Agnès Lagnoux (2022). “Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics”. In: *Bernoulli* 28.4, pp. 2345–2374. DOI: 10.3150/21-BEJ1421.
- Gamboa, Fabrice, Alexandre Janon, Thierry Klein, and Agnès Lagnoux (2013). *Sensitivity analysis for multidimensional and functional outputs*. arXiv: 1311.1797 [stat.AP].
- Ginsbourger, David, Olivier Roustant, Dominic Schuhmacher, Nicolas Durrande, and Nicolas Lenz (Jan. 2016). “On ANOVA Decompositions of Kernels and Gaussian Random Field Paths”. In: pp. 315–330. ISBN: 978-3-319-33505-6. DOI: 10.1007/978-3-319-33507-0_15.
- Gretton, Arthur, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf (2005). “Measuring Statistical Dependence with Hilbert-Schmidt Norms”. In: *Algorithmic Learning Theory*. Springer, pp. 63–77. ISBN: 978-3-540-31696-1.
- Guo, Mengwu and Jan S. Hesthaven (2018). “Reduced order modeling for nonlinear structural analysis using Gaussian process regression”. In: *Computer Methods in Applied Mechanics and Engineering* 341, pp. 807–826. DOI: <https://doi.org/10.1016/j.cma.2018.07.017>.
- Marrel, Amandine, Nadia Pérot, and Clementine Mottet (July 2014). “Development of a surrogate model and sensitivity analysis for spatio-temporal numerical simulators”. In: *Stochastic Environmental Research and Risk Assessment* 29. DOI: 10.1007/s00477-014-0927-y.
- Molchanov, Ilya (2005). *Theory of random sets*. Vol. 19. 2. Springer. DOI: 10.1007/978-1-4471-7349-6.
- Morris, Max D. (1991). “Factorial Sampling Plans for Preliminary Computational Experiments”. In: *Technometrics* 33.2, pp. 161–174. ISSN: 00401706.
- Nony, B., Mélanie Rochoux, Thomas Jaravel, and Didier Lucor (Feb. 2023). “Reduced-order modeling for parameterized large-eddy simulations of atmospheric pollutant dispersion”. In: *Stochastic Environmental Research and Risk Assessment* 37, pp. 1–28. DOI: 10.1007/s00477-023-02383-7.
- Pasquier, Mathis, Stéphane Jay, Jérôme Jacob, and Pierre Sagaut (2023). “A Lattice-Boltzmann-based modelling chain for traffic-related atmospheric pollutant dispersion at the local urban scale”. In: *Building and Environment* 242, p. 110562. ISSN: 0360-1323. DOI: <https://doi.org/10.1016/j.buildenv.2023.110562>.
- Perrin, T.V.E., O. Roustant, J. Rohmer, O. Alata, J.P. Naulin, D. Idier, R. Pedreros, D. Moncoulon, and P. Tinarid (2021). “Functional principal component analysis for global sensitivity analysis of model with spatial output”. In: *Reliability Engineering and System Safety* 211, p. 107522. ISSN: 0951-8320. DOI: <https://doi.org/10.1016/j.ress.2021.107522>.
- Saint-Geours, Nathalie, Jean-Stéphane Bailly, Frédéric Grelot, and Christian Lavergne (2014). “Multi-scale spatial sensitivity analysis of a model for economic appraisal of flood risk management policies”. In: *Environmental Modelling & Software* 60, pp. 153–166. ISSN: 1364-8152. DOI: <https://doi.org/10.1016/j.envsoft.2014.06.012>.
- Saltelli, Andrea, Paola Annoni, Ivano Azzini, Francesca Campolongo, M. Ratto, and Stefano Tarantola (2010). “Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index”. In: *Computer Physics Communications* 181, pp. 259–270. DOI: 10.1016/j.cpc.2009.09.018.

Sobol, I. M (Feb. 15, 2001). “Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates”. In: *Mathematics and Computers in Simulation*. The Second IMACS Seminar on Monte Carlo Methods 55.1, pp. 271–280. ISSN: 0378-4754. DOI: 10.1016/S0378-4754(00)00270-6.

Song, Le, Alex Smola, Arthur Gretton, Karsten Borgwardt, and Justin Bedo (Apr. 20, 2007). *Supervised Feature Selection via Dependence Estimation*. DOI: 10.48550/arXiv.0704.2668. arXiv: 0704.2668[cs].