

Target and Conditional Sensitivity Analysis with Emphasis on Dependence Measures

H. Raguet, A. Marrel

► **To cite this version:**

H. Raguet, A. Marrel. Target and Conditional Sensitivity Analysis with Emphasis on Dependence Measures. 2018. cea-02339856

HAL Id: cea-02339856

<https://hal-cea.archives-ouvertes.fr/cea-02339856>

Submitted on 5 Nov 2019

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

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

NEW STATISTICAL DEVELOPMENTS FOR TARGET AND CONDITIONAL SENSITIVITY ANALYSIS

HUGO RAGUET* AND AMANDINE MARREL*

Abstract. In the context of sensitivity analysis of complex phenomena in presence of uncertainty, we motivate and precise the idea of orienting the analysis towards a critical domain of the studied phenomenon. For this, target and conditional sensitivity analyses are defined. We make a brief history of related approaches in the literature, and propose a more general and systematic approach. Nonparametric measures of dependence being well-suited to this approach, we also make a review of available methods and of their use for sensitivity analysis, and clarify some of their properties. Then, we focus our attention on sensitivity indices based on correlation ratio, namely Sobol' indices, and on two dependence measures: the kernel quadratic dependence measure also called Hilbert–Schmidt independence criterion and the Csiszár divergence dependence measure. We propose adapted versions of these tools for target and conditional analysis, by considering transformation of the output using hard or smooth weight functions. Finally, we show on synthetic numerical experiments both the interest of target and conditional sensitivity analysis, and the efficiency of the dependence measures. We also illustrate the relevance of the proposed smooth versions for conditional estimators.

Key words. Sensitivity analysis, computer experiments, target and conditional sensitivity analysis, dependence measure, correlation ratio

AMS subject classifications. 62G05, 62G99

1. Introduction. Nowadays, many phenomena are modeled by mathematical equations which are implemented and solved using complex computer programs. These numerical simulators are used to model and predict the underlying physical phenomena and their results can guide decisions, which can involve important financial, societal and safety stakes. However, they often take as inputs a high number of numerical, physical or even conception parameters. Because of a lack of phenomenon knowledge and characterization or a need to investigate various configurations, many of these input parameters are uncertain (or considered as such) and it is important to assess how these uncertainties can affect the model output. In a probabilistic framework, the uncertain parameters, also called factors, are modeled by random variables characterized by probabilistic distributions. Sensitivity analysis methods are performed to evaluate how input uncertainties contribute, qualitatively or quantitatively, to the variation of the output. The variety of approaches and applications of sensitivity analysis brings forth a diversity of objects and terms. Before presenting the goals of this document and the extent of our work, we introduce a few notations.

In the classical framework, we assume the modeling of a phenomenon Y depending on a set of *factors* $(X_i)_{1 \leq i \leq d}$ following a deterministic relation $Y \stackrel{\text{def}}{=} f(X_1, \dots, X_d)$, f being the numerical simulator in our industrial applications and Y the output(s) of interest. Uncertainties are taken into account by modeling the factors as random variables, defined over the same implicit probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ with Ω denoting the sample space, \mathfrak{F} the set of events and \mathbb{P} the associated probability measure. It will also be convenient to consider a generic random variable X , usually standing for a group of one or several factors. For such a variable, we note its range $\mathcal{X} \stackrel{\text{def}}{=} \text{ran}(X)$ and its law $\mathbb{P}_X \stackrel{\text{def}}{=} \mathbb{P} \circ X^{-1}$ is the induced probability measure over \mathcal{X} . We also particularize the range of the output $\mathcal{Y} \stackrel{\text{def}}{=} \text{ran}(Y)$.

*CEA, DEN, DER, SESI, LEMS, F-13108, Saint-Paul-lez-Durance, France; hugo.raguet@gmail.com, amandine.marrel@cea.fr

45 **1.1. Global, target and conditional sensitivity analysis.** *Global sensitivity*
 46 *analysis* aims at measuring how the variations of one or several factors contribute to
 47 the variation of the studied phenomenon, over the whole domain of possible values.
 48 Many authors agree with Saltelli et al. (2008) to distinguish several use of sensi-
 49 tivity analysis. First, the *ranking* of factors by importance is the starting point of
 50 any application. Identifying the factors which are most influential in a phenomenon
 51 might help understanding it or guide resource investment for controlling it. Then,
 52 *screening* the factors for insignificant ones is considered, for instance for the purpose of
 53 model simplification. This sometimes calls on statistical tests, which might be practical
 54 but cannot be satisfying for all applications. In our experience, screening is often in
 55 practice interpretation of ranking, either through the expertise of the practitioner or
 56 with some cross-validation process. Finally, *factor mapping* is often described as a
 57 finer identification of functional relationship between the specific domains of values
 58 of the factors and of the phenomenon. This last use of sensitivity analysis consists
 59 in determining *which values* of these factors are responsible of the occurrence of the
 60 phenomenon in a given domain.

61

62 In our work, we also focus on specific domains of values of the phenomenon
 63 but we want to determine *which factors* contribute most in the occurrence of the
 64 phenomenon in a given domain. For this, we first define the *target sensitivity analysis*
 65 which aims at measuring the influence of the factors over a restricted domain of the
 66 studied phenomenon, and in particular over the *occurrence* of the phenomenon in this
 67 restricted domain. Such domain of interest would usually be extreme and relatively
 68 rare, constituting a risk or an opportunity; we call it *critical domain*, noted $\mathcal{C} \subset \mathcal{Y}$ and
 69 associated to a *critical probability* $P(Y \in \mathcal{C}) = P_Y(\mathcal{C})$. Alternatively, we also define the
 70 *conditional sensitivity analysis* which evaluates the influence of the factors *within* the
 71 critical domain only, ignoring what happens outside. Let us underline that those two
 72 notions can widely differ; this point will be illustrated by the numerical applications
 73 proposed in this paper.

74

1.2. Goals and Structure of the Paper. In this paper, we aims at proposing
 75 news methods and tools for target and conditional sensitivity analysis. It seems to
 76 us that there are numerous, direct applications, especially for, but not restricted to,
 77 industrial safety. Still, while global sensitivity analysis has been an active research field
 78 for several decades, it seems that target sensitivity analysis is less understood, and
 79 until recently has not been studied systematically as such. This is why we sometimes
 80 introduce our own terminology, which we discuss along with the description of similar
 81 concepts that we identify in the literature. Finally, let us point out that we are mostly
 82 interested in phenomena influenced by many factors and of which only limited under-
 83 standing is available. Typical situations include complex systems observed through
 84 heavy computer simulations or costly physical measures. These applicative constraints
 85 should be taken into account when selecting and proposing dedicated tools.

86

87 In section 2, we propose a review on the existing approaches and tools for target
 88 sensitivity analysis before introducing our contributions in this framework. Then, the
 89 actual sensitivity analysis tools on which our work relies are more precisely described in
 90 section 3. In section 4, we get back to our initial problematic of target and conditional
 91 sensitivity analysis: we propose a simple dedicated framework and describe some
 92 resulting tools. Finally, we give numerical evaluations of our methods along section 5,
 93 on various synthetic data.

94 **2. Review on Existing Approaches.** We propose here a coarse classification
 95 of methods relating to target sensitivity analysis, according to both chronological and
 96 methodological criteria.

97 **2.1. Regional Sensitivity Analysis.** The very notion of target sensitivity anal-
 98 ysis dates back at least to Spear and Hornberger (1980), motivated by environmental
 99 science applications. The proposed methodology compares the distribution of the
 100 factors within the critical domain against their distribution outside. The authors
 101 choose to use the *Kolmogorov distance*, almost systematically reused ever since:

$$102 \quad \sup_{x \in \mathcal{X}} |F_{X|Y \in \mathcal{C}}(x) - F_{X|Y \in \mathcal{Y} \setminus \mathcal{C}}(x)|,$$

103 where $F_{X|A}$ is the *cumulative distribution function* of a real random variable X (i.e.
 104 $\mathcal{X} \subseteq \mathbb{R}$) conditioned by an event $A \in \mathfrak{F}$ of nonzero probability (see section 4.1.2 for
 105 details).

106 They call it *regional sensitivity analysis*, or sometimes generalized sensitivity
 107 analysis. The former name could fit our purpose, if it was not for two inconveniences.
 108 First, it evokes more of a sensitivity analysis within the critical domain (what we call
 109 conditional sensitivity analysis) rather than its occurrence; second, for the past three
 110 decades in the literature it referred exclusively to the above methodology. It appears
 111 to be the generalization of no other method, explaining why the alternative name is
 112 not used anymore. Finally, one may encounter the term *Monte Carlo Filtering*, which
 113 might be vague and restrictive.

114 Comparing distributions conditionally to the critical domain seems a good choice
 115 for target sensitivity analysis. It involves only two conditionings, which facilitates
 116 its estimation, for instance with Monte Carlo method. One difficulty, mentioned by
 117 the authors and common to all target sensitivity methods, arises when the critical
 118 probability is low. Another deficiency pointed out by the authors is the difficulty to
 119 study factors in interaction. From this viewpoint, observe that a metric comparing
 120 cumulative distribution functions can be extended to multidimensional settings, which
 121 would allow to regroup several factors. However, the particular metric used here, namely
 122 the supremum norm over the differences, is sensitive to outliers. Both aspects make it
 123 particularly unsuitable for categorical factors.

124 Strangely enough, regional sensitivity is mostly used in the literature as a mean of
 125 global sensitivity analysis, the partition of the domain of values of the phenomenon
 126 into several regions losing its original sense and becoming more or less arbitrary.

127 **2.2. Reliability Sensitivity Analysis.** Another field dealing with target sensi-
 128 tivity analysis is motivated by applications in structural reliability, where the term
 129 *reliability sensitivity analysis* is commonly used. In this context, critical domains are
 130 failure domains, and the developed methods are influenced by two typical features:
 131 failure probabilities are small in comparison to the number of available observations,
 132 and the probability distributions of the factors are assumed to be known.

133 The first methods developed seek, in a suitable transformation of the factors space,
 134 to determine a “most probable failure point”, and to estimate the critical probability
 135 from linear or quadratic approximations of the boundary of the critical domain around
 136 that point. This yields the *first- and second-order reliability methods*, reviewed by
 137 Rackwitz (2001). It is possible to give to each factor an importance measure based on
 138 the position of the most probable failure point. The geometrical assumption about the
 139 failure domain seems however restrictive, implying in particular that the factors have
 140 a monotonous effect on the phenomenon.

141 The sensitivity measures which later prevail in the field are based on derivatives of
 142 the critical probability, with respect to the parameters defining the probability laws of
 143 the factors or of their transformation. This framework seems once again restrictive for
 144 our purpose. Nonetheless, the approaches developed in parallel for dealing with low
 145 critical probabilities deserves to be incidentally noted, because they could be adapted to
 146 other sensitivity measures. Let us mention the methods based on *importance sampling*
 147 (see for instance the adaptation of Wu, 1994), as well as the approach of *sequential*
 148 *Monte Carlo* as proposed by Au and Beck (2001), who call it *subset simulation*. As
 149 further developed by Song et al. (2009) and Cérou et al. (2012), the latter is based on
 150 Markov chain Monte Carlo with the Metropolis–Hasting algorithm.

151 Still in the reliability context, the Ph.D. dissertation of Lemaître (2014) is the first
 152 systematic study of target sensitivity analysis. With this purpose in mind, the author
 153 compares more general methods of global sensitivity analysis, of which we give a brief
 154 overview below. We can already mention that he identifies the need to transform the
 155 variable modeling the phenomenon into a binary variable encoding the occurrence in
 156 the critical domain; that is $1_{\mathcal{C}}(Y)$, where $1_{\mathcal{C}}: y \mapsto 1$ if $y \in \mathcal{C}$, 0 otherwise. This is one
 157 of the approach on which we focus in this work (see subsection 4.1).

158 A first sensitivity analysis method considered is the estimation of (square) *correla-*
 159 *tion ratio* between the factors and the phenomenon (real, with finite variance),

$$160 \quad (2.1) \quad \eta^2(X, Y) \stackrel{\text{def}}{=} \frac{V(\mathbb{E}[Y | X])}{V(Y)} .$$

161 Resulting quantities are often called Sobol’ (1990, 1993) indices. These are nowadays
 162 standard for global sensitivity analysis, notably because they can be interpreted in
 163 terms of decomposition of the variance of the studied phenomenon. Lemaître shows
 164 how these indices applied to the binary transformation of the observed phenomenon,
 165 $\eta^2(X, 1_{\mathcal{C}}(Y))$, are relevant at least for cases that are simple and where the number of
 166 available observations is high enough.

167 A second method is based on the *total variation* which we develop later (see
 168 section 3.2.2), once again applied to the binary transformation. Unfortunately, the
 169 proposed estimation methods might be inadequate and the analysis is too brief; the
 170 author mentions a “positive bias” without further explanations.

171 Another set of methods is based on *binary classification trees*. The author lists
 172 many ways of defining classification trees, and even more ways of deducing sensitivity
 173 indices. This indicates a lack of generality and robustness, actually revealed by some
 174 numerical experiments. For the sake of brevity, we do not elaborate here and invite
 175 the interested reader to refer to the dissertation for more details.

176 Then, Lemaître takes over regional sensitivity described above with some modifi-
 177 cations. He compares the probability laws conditionally to the critical domain against
 178 the (known) marginal probability laws. In addition to Kolmogorov distance, he tries
 179 other discrepancy measures between cumulative distribution functions classically used
 180 in statistical tests, namely Cramér–Von Mises and Anderson–Darling, and shows that
 181 this choice can influence the importance ranking of factors. More importantly, he
 182 suggests that sequential Monte Carlo approaches are well adapted to methods based
 183 on comparisons of factor distributions conditionally to the critical domain.

184 Finally, closer to the classical sensitivity measures for reliability mentioned above,
 185 the author proposes its own measures, quantifying how modifications of the factors
 186 probability laws impact on the critical probability. Although it has specific advan-
 187 tages, such as quantification of uncertainties due to estimation errors on the model’s

188 parameters, this framework seems somewhat artificial and restrictive.

189 Altogether, this Ph.D. dissertation is an interesting entry point to target sensitivity
 190 analysis. However, more numerical experiments seem necessary in order to conclude
 191 about the advantages and drawbacks of the different considered approaches, and those
 192 which should be retained for further improvements and comparisons are not clearly
 193 identified.

194 **2.3. Sensitivity Analysis of a Specific Statistic.** Another recent approach
 195 for target sensitivity analysis is due to Fort et al. (2013). Their formulation is more
 196 precise than ours: they are interested in the sensitivity of *an estimator of a statistical*
 197 *quantity* of the studied phenomenon. For this, they introduce the term of *goal-oriented*
 198 *sensitivity analysis*.¹

199 From the relations $V(E[Y | X]) = E\left(\left(E[Y | X] - E[Y]\right)^2\right) = V(Y) - E(V[Y | X])$,
 200 the authors show how the correlation ratio, (2.1), is, in their sense, a measure adapted
 201 to the sensitivity of the *expectation* of the phenomenon: indeed, it measures a distance
 202 between expectations, quantified by a difference of variances. Now for a generic real
 203 random variable Y , expectation and variance can be defined through an optimization
 204 problem, $E[Y] = \arg \min_{\theta \in \mathbb{R}} E\left((Y - \theta)^2\right)$ and $V[Y] = \min_{\theta \in \mathbb{R}} E\left((Y - \theta)^2\right)$, where
 205 the functional $(y, \theta) \mapsto (y - \theta)^2$ plays the role of a *contrast function*.

206 The generalization of the correlation ratio to a statistic defined by another con-
 207 trast function ψ becomes² $\min_{\theta \in \mathbb{R}} E(\psi(Y, \theta)) - E(\min_{\theta \in \mathbb{R}} E[\psi(Y, \theta) | X])$. In practice,
 208 in order to study extreme values, they focus on the *quantiles* of the phenomenon,
 209 considering for a level $\alpha \in]0, 1[$, the contrast function $(y, \theta) \mapsto (y - \theta)(1_{\{y \leq \theta\}} - \alpha)$.
 210 However, resulting indices turn out to be difficult to estimate, as shown by the recent
 211 developments of Browne et al. (2017) and Maume-Deschamps and Niang (2017).

212 Les us mention that Kucherenko and Song (2016) propose another adaptation of
 213 the correlation ratio to analysis of sensitivity of quantiles, more direct: expectations
 214 are simply replaced by quantiles³ of level $\alpha \in]0, 1[$, $E\left(\left(F_{Y|X}^{-1}(\alpha) - F_Y^{-1}(\alpha)\right)^2\right)$, where
 215 F^{-1} is the generalized inverse of a cumulative distribution function. As its estimation
 216 is also difficult, the authors propose to approximate the quantiles conditionally to
 217 factors values by a rough form of kernel method.

218 At last, let us add that the quantile is a peculiar notion and in our opinion, its use for
 219 sensitivity analysis raises some troubles. Beyond difficulty of definition and estimation,
 220 these tools are adapted only to phenomena which are unidimensional and continuous.
 221 Moreover, the “sensitivity of a quantile” has a less straightforward interpretation than
 222 the sensitivity of the occurrence of a phenomenon, or of the variation of a phenomenon,
 223 in a critical domain.

224 **2.4. Contributions.** The majority of the methods previously described are
 225 originally developed for particular applications; we would like to make abstraction
 226 of the problem to get more general methods. To this end, rather than defining or
 227 enhancing specific methods, we seek modifications or generalizations of global sensitivity
 228 analysis tools, which would be adapted to target or conditional sensitivity analysis.

229 Such modifications boil down, for a given analysis tool considered, to weighting
 230 the observations according to the critical domain. The weights can operate following

¹Beware that this term already exists in the literature referring to tools of different nature.

²Provided that the random variable $\min_{\theta \in \mathbb{R}} E[\psi(Y, \theta) | X]$ is well defined.

³The random variable $F_{Y|X}^{-1}$ is now defined through conditional distribution.

231 two principles: either as a *transformation* of the phenomenon prior to the application
 232 of the tool, or as a *modification* of the parameters and objects which define the tool
 233 itself. This includes, but is not restricted to, the natural notion of conditioning. Several
 234 variations around these principles are presented along [section 4](#); before that, the actual
 235 sensitivity analysis tools must be introduced.

236 **3. Correlation Ratio and Dependence Measures for Sensitivity Analy-**
 237 **sis.** We present here the measures of sensitivity analysis upon which we construct our
 238 tools. First, sensitivity indices based on correlation ratio, the popular *Sobol' indices*,
 239 are introduced. Now, it appears that sensitivity analysis based on *nonparametric*
 240 *dependence measures*, recently advocated by [Da Veiga \(2015\)](#), is particularly adapted
 241 to our framework. This will retain most our attention in the following, starting from
 242 [subsection 3.2](#) where we review available methods for measuring statistical dependence
 243 and detail the use of some of them in the context of sensitivity analysis.

244 **3.1. Correlation Ratio yielding Sobol' Indices.** Given a group of factors
 245 $I \subset \{1, \dots, d\}$, we write $X_I \stackrel{\text{def}}{=} (X_i)_{i \in I}$ for the corresponding random tuple, and
 246 $\complement I \stackrel{\text{def}}{=} \{1, \dots, d\} \setminus I$ for the complementary group of factors. Moreover, we abusively
 247 note the concatenation $(X_I, X_{\complement I}) \stackrel{\text{def}}{=} (X_i)_{1 \leq i \leq d}$.

248 The use of correlation ratio for sensitivity analysis has been proposed by [Iman](#)
 249 [and Hora \(1990\)](#) and [Ishigami and Homma \(1990\)](#), and independently by [Sobol' \(1990,](#)
 250 [1993\)](#). The latter was the most popularized, introducing modifications of correlation
 251 ratios of groups of factors to achieve a convenient decomposition of the total variance
 252 of the phenomenon, provided that the factors are independent; these are the Sobol'
 253 indices. While they are theoretically interesting for studying specific interactions of
 254 factors, in practice the most useful sensitivity indices are the *first-order indices* and the
 255 *total-order indices*. The former tends to evaluate the influence of a group of factor I
 256 on its own and is simply $\eta^2(X_I, Y)$, and the latter incorporate all possible interactions
 257 with other factors, defined as $1 - \eta^2(X_{\complement I}, Y)$.

258 Estimation of correlation ratio can be expensive because it involves the term
 259 $E(E[Y | X_I]^2)$. Most common efficient estimators develop the square conditional ex-
 260 pectation as the product $E[f(X_I, X_{\complement I}) | X_I] E[f(X_I, X'_{\complement I}) | X_I]$ where $(X_I, X'_{\complement I})$ is
 261 distributed identically to $(X_I, X_{\complement I})$, which in turn is $E[f(X_I, X_{\complement I})f(X_I, X'_{\complement I}) | X_I]$,
 262 provided that $X_{\complement I}$ and $X'_{\complement I}$ are independent conditionally to X_I . In practice, this is
 263 ensured when the *input factors are independent*. The expectation of the last expression
 264 is nothing but $E(f(X_I, X_{\complement I})f(X_I, X'_{\complement I}))$, which is now easier to handle. Typical esti-
 265 mator consists in drawing $2n$ independent observations $(X_I^{(j)}, X_{\complement I}^{(j)})_{1 \leq j \leq 2n}$ distributed
 266 as $(X_I, X_{\complement I})$, and evaluating the model at specifically chosen factors combinations,
 267 typically

$$268 \quad (3.1) \quad E(E[Y | X_I]^2)_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{j=1}^n f(X_I^{(j)}, X_{\complement I}^{(j)}) f(X_I^{(j)}, X_{\complement I}^{(n+j)}) .$$

269 Let us mention that this approach, usually referred to as *pick-and-freeze*, has two
 270 drawbacks: first, this constrains the *experience design* (the set of points at which the
 271 model must be observed or computed), and second, the required number of model
 272 evaluations grows with the number of factors to be investigated.

273 **3.2. Sensitivity Analysis with Dependence Measures.** Sensitivity analysis
 274 based on correlation ratio as described above is fairly general and can be readily adapted

275 for target and conditional analysis, as we propose later in sections 4.2.1 and 4.2.2.
 276 However, several weaknesses can be pointed out.

277 First, accurate estimation is known for requiring many observations. In addition,
 278 although statistical independence implies zero correlation ratio, some variables can
 279 be significantly related and yet their correlation ratio be zero as well; in such case,
 280 one must resort to total-order indices to identify a relationship. More generally, the
 281 statistical variance of the phenomenon might not be the most representative mode of
 282 variation. Finally, the above extension for multidimensional phenomenon might not be
 283 satisfying. In the classical probabilistic framework, we believe with Da Veiga (2015)
 284 that a more general and more versatile notion of sensitivity of a phenomenon to a group
 285 of factors can be captured by the notion of statistical dependence. Moreover, De Lozzo
 286 and Marrel (2016) recently investigated the use of dependence measures for sensitivity
 287 analysis of costly model and illustrated the efficiency of associated significance tests
 288 for screening purpose.

289 We propose somewhere else (Raguet and Marrel, 2018, §§ 3.1 and 3.2), a classifi-
 290 cation of dependence measures in general, and an extensive discussion on their use for
 291 sensitivity analysis in particular. We refer the interested reader to the above article for
 292 details; for now, let us focus on two important classes. Note that our choice is mainly
 293 guided by ease of implementation (notably the possibility of writing estimators as
 294 empirical expectations), aim for generality (factors and phenomenon of any nature
 295 and dimension), good invariance properties, and ease of adaptation for target and
 296 conditional sensitivity analysis. They both rely on the same principle: measuring
 297 the statistical dependence between two variables X and Y by comparing their joint
 298 distribution $P_{X,Y}$ to their product $P_X \otimes P_Y$; the two being equal if, and only if, X
 299 and Y are independent.

300 **3.2.1. Kernel Quadratic Dependence Measure also called Hilbert–Schmidt**

301 **Independence Criterion.** The first class of dependence measures which we consider
 302 arises in the literature from the comparison of the distributions according to their
 303 *probability density* or *characteristic functions*, with help of weighted L_2 norms. However,
 304 more recent interpretations in terms of *kernel embeddings* of probability distributions
 305 yield the *kernel quadratic dependence measure*, following the terminology of Achard
 306 et al. (2003) and Diks and Panchenko (2007), also called *Hilbert–Schmidt independence*
 307 *criterion* by Gretton et al. (2005).

308 In brief, if P is a probability distribution over a generic space \mathcal{Z} , and if $k: \mathcal{Z}^2 \rightarrow \mathbb{R}$
 309 is a suitable *positive definite kernel*, then the mapping $z \mapsto \int k(z, z') dP(z')$ is an
 310 element of the *reproducing kernel Hilbert space* induced by k (see the introduction
 311 of Berlinet and Thomas-Agnan, 2003, Chapter 4). The norm between such *kernel*
 312 *embeddings* of two different probability distributions is called their *kernel distance*.

313 A measure of the dependence between X and Y is thus defined by the ker-
 314 nel distance between $P_{X,Y}$ and $P_X \otimes P_Y$. These are probability distributions over
 315 the space $\mathcal{X} \times \mathcal{Y}$; a useful particular case arises when the kernel k is separable as
 316 $((x, y), (x', y')) \mapsto k_{\mathcal{X}}(x, x')k_{\mathcal{Y}}(y, y')$, where $k_{\mathcal{X}}$ and $k_{\mathcal{Y}}$ are positive definite kernels
 317 over \mathcal{X} and \mathcal{Y} , respectively. The square of the resulting kernel distance is the kernel
 318 quadratic dependence measure, and can be expressed as
 319

320 (3.2) $\text{QDM}_{k_{\mathcal{X}}, k_{\mathcal{Y}}}(X, Y) \stackrel{\text{def}}{=} \text{E}(k_{\mathcal{X}}(X, X')k_{\mathcal{Y}}(Y, Y')) + \text{E}(k_{\mathcal{X}}(X, X')) \text{E}(k_{\mathcal{Y}}(Y, Y')) - 2 \text{E}(k_{\mathcal{X}}(X, X')k_{\mathcal{Y}}(Y, Y'')) ,$
 321

322 provided that (X', Y') is independent of, and distributed identically to, (X, Y) , and
 323 Y'' is independent of X, Y, X', Y' and distributed identically to Y .
 324

325 A straightforward estimator, given $(X^{(i)}, Y^{(i)})_{1 \leq i \leq n}$ independent observations
 326 distributed identically to (X, Y) , is

$$\begin{aligned} \text{QDM}_{k_{\mathcal{X}}, k_{\mathcal{Y}}}(X, Y)_n &\stackrel{\text{def}}{=} \frac{1}{n^2} \sum_{i,j=1}^n k_{\mathcal{X}}(X^{(i)}, X^{(j)}) k_{\mathcal{Y}}(Y^{(i)}, Y^{(j)}) \\ &+ \frac{1}{n^2} \left(\sum_{i,j=1}^n k_{\mathcal{X}}(X^{(i)}, X^{(j)}) \right) \frac{1}{n^2} \left(\sum_{i,j=1}^n k_{\mathcal{Y}}(Y^{(i)}, Y^{(j)}) \right) \\ &- \frac{2}{n} \sum_{i=1}^n \left(\frac{1}{n} \sum_{j=1}^n k_{\mathcal{X}}(X^{(i)}, X^{(j)}) \right) \left(\frac{1}{n} \sum_{j=1}^n k_{\mathcal{Y}}(Y^{(i)}, Y^{(j)}) \right), \end{aligned}$$

328 which should be put under the following handier form for practical implementation,

$$329 \frac{1}{n^2} \sum_{i,j=1}^n \left(k_{\mathcal{X}}(X^{(i)}, X^{(j)}) - \frac{1}{n} \sum_{\ell=1}^n k_{\mathcal{X}}(X^{(i)}, X^{(\ell)}) \right) \left(k_{\mathcal{Y}}(Y^{(i)}, Y^{(j)}) - \frac{1}{n} \sum_{\ell=1}^n k_{\mathcal{Y}}(Y^{(i)}, Y^{(\ell)}) \right).$$

330 Note that some authors prefer normalizing with factors $n - 1$ and $n - 2$, or add some
 331 other debiasing modifications, which are of little interest here. The required number
 332 of computations grows as $O(n^2)$, which is acceptable in situations where the cost for
 333 obtaining each observation is large.

334 When \mathcal{X} is a finite set, the *categorical kernel* $(x, x') \mapsto 1$ if $x = x'$, 0 otherwise, is
 335 most typically used. When \mathcal{X} is a normed vector space, the *Gaussian kernel* $(x, x') \mapsto$
 336 $\exp(-\|x - x'\|^2 / 2\sigma^2)$ for some parameter $\sigma^2 \in \mathbb{R}$, dependent in practice on the data,
 337 is also typically used, but others variants are popular; [Sejdinovic et al. \(2013\)](#) show
 338 that the *distance covariance* of [Székely et al. \(2007\)](#) is a particular case.

339 These last authors propose to normalize their dependence measure in the spirit
 340 of the linear correlation coefficient, yielding the *distance correlation*, which is scale-
 341 invariant. When [Da Veiga \(2015\)](#) highlights the potential use of the quadratic de-
 342 pendence measure for sensitivity analysis, he also advocates for such normalization,
 343 leading to the sensitivity index

$$344 \overline{\text{QDM}}_{k_{\mathcal{X}}, k_{\mathcal{Y}}}(X, Y) \stackrel{\text{def}}{=} \frac{\text{QDM}_{k_{\mathcal{X}}, k_{\mathcal{Y}}}(X, Y)}{\sqrt{\text{QDM}_{k_{\mathcal{X}}, k_{\mathcal{X}}}(X, X)} \sqrt{\text{QDM}_{k_{\mathcal{Y}}, k_{\mathcal{Y}}}(Y, Y)}},$$

345 and similarly for its *plug-in* estimator $\overline{\text{QDM}}_{k_{\mathcal{X}}, k_{\mathcal{Y}}}(X, Y)_n$.

346 **3.2.2. Csiszár Divergence Dependence Measure.** The second class of de-
 347 pendence measures which we consider compares the distributions through *Csiszár*
 348 [\(1972\) divergences](#). Considering again P, Q two probability distributions over a generic
 349 space \mathcal{Z} , a *Csiszár divergence* between P and Q can be defined as

$$350 \text{div}_{\phi}(P, Q) \stackrel{\text{def}}{=} \int \phi \left(\frac{dP}{dQ} \right) dQ,$$

351 where $\phi: \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{+\infty\}$ is a convex function vanishing at unity, and $\frac{dP}{dQ}$ is the
 352 Radon–Nikodym derivative of P with respect to Q ; note that this can be conveniently
 353 extended to cases where P is not dominated by Q . Notable examples include the
 354 *(reverse) Kullback–Leibler divergence* with $\phi: t \mapsto -\log(t)$ and the *total variation*
 355 *distance* with $\phi: t \mapsto |t - 1|$.

356 Again, one can measure dependence between X and Y by measuring discrepancy
 357 between $P_{X,Y}$ and $P_X \otimes P_Y$ with help of this tool, yielding *Csiszár divergence depen-*
 358 *dence measure*, $\text{CDM}_\phi(X, Y) \stackrel{\text{def}}{=} \text{div}_\phi(P_X \otimes P_Y, P_{X,Y})$. The famous special case of the
 359 *mutual information*, stemming from the work of [Shannon \(1948\)](#), is obtained with the
 360 Kullback–Leibler divergence, $\text{div}_{-\log}(P_X \otimes P_Y, P_{X,Y})$. In the context of sensitivity
 361 analysis, [Park and Ahn \(1994\)](#) use a form of mutual information, and later [Borgonovo](#)
 362 [\(2007\)](#) uses the *total variation* dependence measure. In both cases, estimating the
 363 Csiszár divergence is problematic, and the authors must call on *ad hoc* parametric
 364 density fits.

365 In fact, estimations of Csiszár divergences have been studied in many contexts, often
 366 focused on specific versions defined by a given function ϕ or on specific knowledge about
 367 the involved distributions. In order to devise a general tool, we choose in the current
 368 work to rely on nonparametric estimations of the Radon–Nikodym derivatives. Densities
 369 at specific points can be estimated through *kernel* or *nearest-neighbors* methods, see for
 370 instance the monograph of [Silverman \(1986\)](#). Probabilities are estimated by empirical
 371 frequencies. Both can be combined if necessary.

372 We justify somewhere else ([Raguét and Marrel, 2018](#), § 3.4.2) that a “support”
 373 version, $\text{sCDM}_\phi(X, Y)$, where the integration is performed over the range of the
 374 joint variable (X, Y) rather than the whole product space $\mathcal{X} \times \mathcal{Y}$, is convenient. The
 375 corresponding estimator is

(3.3)

$$376 \quad \text{sCDM}_\phi(X, Y)_{k_{\mathcal{X}}, k_{\mathcal{Y}}, n} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n \phi \left(\frac{\left(\frac{1}{n} \sum_{j=1}^n k_{\mathcal{X}}(X^{(i)}, X^{(j)}) \right) \left(\frac{1}{n} \sum_{j=1}^n k_{\mathcal{Y}}(Y^{(i)}, Y^{(j)}) \right)}{\frac{1}{n} \sum_{j=1}^n k_{\mathcal{X}, \mathcal{Y}}((X^{(i)}, Y^{(i)}), (X^{(j)}, Y^{(j)}))} \right)$$

377 where $k_{\mathcal{X}}$, $k_{\mathcal{Y}}$ and $k_{\mathcal{X}, \mathcal{Y}}$ are the kernels used for estimating densities or probabilities;
 378 typically (normalized) Gaussian and categorical, respectively. It has a computational
 379 cost of $O(n^2)$, just as for the kernel quadratic dependence measure.

380 Unfortunately, normalization is not as natural as for the kernel quadratic depen-
 381 dence measure which derives from a square norm. Consider moreover that, for instance,
 382 $\text{sCDM}_\phi(X, X)$ might be infinite. We propose to normalize *the estimators* as

$$383 \quad \overline{\text{sCDM}}_\phi(X, Y)_{k_{\mathcal{X}}, k_{\mathcal{Y}}, n} \stackrel{\text{def}}{=} \frac{\text{sCDM}_\phi(X, Y)_{k_{\mathcal{X}}, k_{\mathcal{Y}}, n}}{\text{sCDM}_\phi(X, X)_{k_{\mathcal{X}}, k_{\mathcal{X}}, n}}.$$

384 Let us mention to the interested reader that this can be seen as a rough generalization
 385 of the normalization proposed by [Joe \(1989\)](#) for mutual information of categorical
 386 variables, because $\text{sCDM}_{-\log}(X, X)$ is in that case the *Shannon entropy* of X .

387 **4. Some Tools for Conditional and Target Sensitivity Analysis.** All of
 388 the sensitivity measures detailed above can be easily adapted to target and conditional
 389 sensitivity analysis. We describe first general approaches which can be applied to any
 390 sensitivity measure. Further details are then given for each tool that we consider.

391 **4.1. Transformations and Weights.** Our general approaches are based on
 392 transformations of the variable quantifying the phenomenon and on conditioning;
 393 specific notions and notations are introduced here.

394 **4.1.1. Targeting with Transformations.** In order to study the *occurrences* of
 395 the phenomenon Y within the critical domain $\mathcal{C} \subset \mathcal{Y}$, the natural transformation which
 396 comes to mind is a binary random variable encoding directly the actual phenomenon of

397 interest and suppressing uninformative fluctuations. This leads to consider the weight
 398 function $1_{\mathcal{C}}: \mathcal{Y} \rightarrow \{0, 1\}: y \mapsto 1$ if $y \in \mathcal{C}$, 0 otherwise.

399 Now, recall that a limited number of observations is usually assumed, so that
 400 estimation considerations cannot be ignored. The binary transformation above might
 401 result in a significant loss of the information conveyed by the relative values of Y .
 402 Indeed, when the critical probability $P_Y(\mathcal{C})$ is low, most data is summed up to a bunch
 403 of zeroes.

404 Fortunately, a sensible relaxation of the binary assumption can be given as soon as
 405 one can evaluate some sort of distance $d_{\mathcal{C}}: \mathcal{Y} \rightarrow \mathbb{R}_+$ between each point in \mathcal{Y} and the
 406 critical domain \mathcal{C} . One can compose it by a decreasing real function $\mathbb{R} \rightarrow [0,1]$, with
 407 the rationals that the closer is an observation to the critical domain, the more likely it
 408 is to convey similar information. This of course assumes some kind of regularity of the
 409 phenomenon's statistical properties. When \mathcal{Y} lies in an Euclidean space, we typically
 410 consider the weight function $y \mapsto \exp(-d_{\mathcal{C}}(y)/s)$, where $d_{\mathcal{C}}(y) \stackrel{\text{def}}{=} \inf_{y' \in \mathcal{C}} \|y - y'\|$. Here,
 411 the exponential function encodes multiplicative contributions, and s is a smoothing
 412 parameter depending typically on a measure of dispersion of the values of Y .

413 In all the following, $w: \mathcal{Y} \rightarrow [0,1]$ is any kind of the above weight functions, either
 414 used deterministically, or as a transformation yielding a random variable through the
 415 composition $w(Y)$. Any sensitivity measure between a group of factors X and $w(Y)$
 416 yields a target sensitivity measure.

417 **4.1.2. Conditioning with Weighted Probabilities.** Alternatively, in order
 418 to study the *behavior* of the phenomenon within the critical domain, a natural idea
 419 is *conditioning by the event* $\{Y \in \mathcal{C}\}$. Given an initial probability space $(\Omega, \mathfrak{F}, P)$,
 420 if $A \in \mathfrak{F}$ is an event of nonzero probability, then conditioning by A simply means
 421 *endowing the measurable space* (Ω, \mathfrak{F}) *with the probability measure* $P_{|A}$, defined as
 422 $P_{|A}(B) \stackrel{\text{def}}{=} P(B \cap A)/P(A)$ for all $B \in \mathfrak{F}$. If X is a random variable over $(\Omega, \mathfrak{F}, P)$, then
 423 its law conditionally to A is the law of the mapping X over the conditioned probability
 424 space $(\Omega, \mathfrak{F}, P_{|A})$, that is $P_{X|A} \stackrel{\text{def}}{=} P_{|A} \circ X^{-1}$.

425 Just as we introduced smooth relaxation of the binary transformation above, it
 426 might be useful to consider extensions of conditioning allowing to take into account
 427 some of the information outside the critical domain. This can be easily done by
 428 observing that $P_{|A}(B)$ can be expressed as $\int_B 1_A dP / \int_{\Omega} 1_A dP$. If W is a positive
 429 nonzero random variable over $(\Omega, \mathfrak{F}, P)$ with finite expectation, we define the *probability*
 430 *P weighted by W*, noted P^W , with for all $B \in \mathfrak{F}$, $P^W(B) \stackrel{\text{def}}{=} \int_B W dP / \int_{\Omega} W dP$. In
 431 other words, P^W is the probability distribution absolutely continuous with respect to P
 432 whose density is proportional to W . In addition, if X is a generic random variable, we
 433 clarify that the notation P_X^W stands for the image measure $(P^W)_X$; although strictly
 434 speaking, it cannot be confused with a weighted image measure $(P_X)^W$ since W is
 435 defined over Ω and not over the range of X . Let us also exemplify the particular cases
 436 of weighted probabilities which are actual conditional probabilities, $P_{|A} = P^{1_A}$, and
 437 $P_{X|A} = P_X^{1_A}$.

438 In a probabilistic framework, any sensitivity measure is defined depending on a
 439 (usually implicit) probability space. When conditioning by weight W , we change the
 440 underlying probability measure, but the mappings defining the random variables are
 441 left unchanged; in such case, the notations are prefixed by $[P^W]$. Let us underline
 442 here that, provided that the expectations exist, $[P^W] E(X) = E(WX)/E(W)$.

443 For conditional sensitivity analysis, we typically use conditioning by weights
 444 $W \stackrel{\text{set}}{=} w(Y)$ as defined above.

445 **4.2. Correlation Ratio.** As presented in [subsection 3.1](#), sensitivity indices based
 446 on correlation ratio (widely known as Sobol' indices) all consists in (possibly weighted
 447 sums of) correlation ratios of the phenomenon Y with well chosen groups of factors,
 448 noted generically X .

449 **4.2.1. Target Correlation Ratio.** Correlation ratios can be directly applied
 450 to the transformation $w(Y)$, yielding target sensitivity analysis indices based on
 451 $\eta^2(X, w(Y))$. Observe that even for multidimensional Y , the transformation $w(Y)$
 452 takes values in $[0,1]$, thus sparing us the trouble of interpreting multidimensional
 453 extensions of correlation ratio.

454 **4.2.2. Proposition of Hybrid Conditional Correlation Ratio.** Following
 455 [section 4.1.2](#), the correlation ratio conditioned by the critical domain is the quantity
 456 $[P^{w(Y)}] \eta(X, Y)$. It is important to note that even if the factors are independent
 457 under P , they usually are not under $P^{w(Y)}$. The covariance estimator in [\(2.1\)](#) cannot
 458 be used anymore, hindering the estimation of the correlation ratio as explained in
 459 [subsection 3.1](#).

460 Alternatively, it is possible to define a conditional correlation ratio by another
 461 transformation of Y . We have seen that $w(Y)$, keeping no memory of the actual values
 462 of Y , is more adapted to target sensitivity; for conditional sensitivity, it is preferable
 463 to weight multiplicatively the values, as $w(Y)Y$. However, the fact that w vanishes on
 464 regions away from the critical domain seems arbitrary: the value zero might not be
 465 meaningful for the phenomenon at hand. Since the correlation ratio is a measure of
 466 variance, it still seems relevant to set a constant value over these regions, but equal to
 467 the expectation of the resulting transformation; they would then not contribute to the
 468 variance of the phenomenon. We thus define the transformation

$$469 Y_w \stackrel{\text{def}}{=} w(Y)Y + (1-w(Y))y_0 \quad \text{such that} \quad y_0 \stackrel{\text{def}}{=} E(Y_w); \quad \text{yielding} \quad y_0 = \frac{E(w(Y)Y)}{E(w(Y))}.$$

470 Observe that with $w \stackrel{\text{set}}{=} 1_{\mathcal{C}}$, $E(w(Y)) = P(Y \in \mathcal{C})$ and $y_0 = E[Y | Y \in \mathcal{C}]$; more
 471 generally, we have $y_0 = [P^{w(Y)}] E(Y)$. In any case, it is easy to estimate with
 472 $\sum_{i=1}^n w(Y^{(i)})Y^{(i)} / \sum_{i=1}^n w(Y^{(i)})$, and $\eta^2(X, Y_w)$ can be estimated in turn with any
 473 usual method, with the advantage over the conditional correlation ratio that even
 474 observations associated to null weight are somehow taken into account.

475 **4.3. Kernel Quadratic Dependence Measure.** We recall that this depen-
 476 dence measure is also known as Hilbert–Schmidt independence criterion, and is detailed
 477 in [section 3.2.1](#).

478 **4.3.1. Target Kernel Quadratic Dependence Measure.** Just as with the
 479 correlation ratio, target sensitivity measure of a group of factors can be obtained
 480 through the weight transformations $w(Y)$, that is to say $\text{QDM}_{k_X, k_{w(Y)}}(X, w(Y))$. Our
 481 notation reminds that the kernels depend on the underlying spaces; in the particular
 482 case of the binary transformation $w \stackrel{\text{set}}{=} 1_{\mathcal{C}}$, it seems natural to use a categorical kernel for
 483 $k_{\{0,1\}}$. Let us mention that this last case was already suggested and briefly illustrated
 484 by [Da Veiga \(2015\)](#).

485 **4.3.2. Conditional Kernel Quadratic Dependence Measure.** The condi-
 486 tional version $[P^{w(Y)}] \text{QDM}_{k_X, k_Y}(X, Y)$ is defined through kernel distance and can

487 be again expressed as expectations of kernels analogously to (3.2)

$$488 \quad \begin{aligned} & \mathbb{E}(k_{\mathcal{X}}(X, X')k_{\mathcal{Y}}(Y, Y')\bar{w}(Y)\bar{w}(Y')) \\ & + \mathbb{E}(k_{\mathcal{X}}(X, X')\bar{w}(Y)\bar{w}(Y')) \mathbb{E}(k_{\mathcal{Y}}(Y, Y')\bar{w}(Y)\bar{w}(Y')) \\ & - 2 \mathbb{E}(k_{\mathcal{X}}(X, X')k_{\mathcal{Y}}(Y, Y'')\bar{w}(Y)\bar{w}(Y')\bar{w}(Y'')) , \end{aligned}$$

489 having taken care of normalizing the weights $\bar{w} \stackrel{\text{def}}{=} \mathbb{E}(w(Y))^{-1}w$. This can also be
490 analogously estimated, by replacing empirical averages by weighted averages

$$491 \quad \begin{aligned} & \sum_{i,j=1}^n \left(k_{\mathcal{X}}(X^{(i)}, X^{(j)}) - \sum_{\ell=1}^n k_{\mathcal{X}}(X^{(i)}, X^{(\ell)})\hat{w}(Y^{(\ell)}) \right) \\ & \times \left(k_{\mathcal{Y}}(Y^{(i)}, Y^{(j)}) - \sum_{\ell=1}^n k_{\mathcal{Y}}(Y^{(\ell)}, Y^{(j)})\hat{w}(Y^{(\ell)}) \right) \times \hat{w}(Y^{(i)})\hat{w}(Y^{(j)}) , \end{aligned}$$

492 with empirical normalized weights $\hat{w} \stackrel{\text{def}}{=} (\sum_{i=1}^n w(Y^{(i)}))^{-1}w$.

493 **4.4. Csiszár Divergence Dependence Measure.** We refer to section 3.2.2
494 for the definitions of the “support” Csiszár divergence dependence measure.

495 **4.4.1. Target Csiszár Divergence Dependence Measure.** As previously,
496 target sensitivity measure of a group of factors can be obtained through Csiszár diver-
497 gence dependence measures of the transformations $w(Y)$, that is to say $\text{sCDM}_{\phi}(X, w(Y))$ ■
498 Let us emphasize that, in the case of the binary transformation $w \stackrel{\text{set}}{=} 1_{\mathcal{C}}$, Radon-
499 Nikodym derivatives should be estimated with normalized categorical kernel.

500 **4.4.2. Conditional Csiszár Divergence Dependence Measure.** The condi-
501 tional versions are respectively $[\mathbb{P}^{w(Y)}] \text{CDM}_{\phi}(X, Y) = \text{div}_{\phi}(\mathbb{P}_{X,Y}^{w(Y)}, \mathbb{P}_X^{w(Y)} \otimes \mathbb{P}_Y^{w(Y)})$
502 and $[\mathbb{P}^{w(Y)}] \text{sCDM}_{\phi}(X, Y) = \text{sdiv}_{\phi}(\mathbb{P}_X^{w(Y)} \otimes \mathbb{P}_Y^{w(Y)}, \mathbb{P}_{X,Y}^{w(Y)})$. In the estimator in (3.3),
503 the weights are influencing the expectations in each density estimation and each integral,
504 yielding with empirical normalized weights $\hat{w} \stackrel{\text{def}}{=} (\sum_{i=1}^n w(Y^{(i)}))^{-1}w$,

$$505 \quad \sum_{i=1}^n \hat{\phi} \left(\frac{\left(\sum_{j=1}^n k_{\mathcal{X}}(X^{(i)}, X^{(j)})\hat{w}(Y^{(j)}) \right) \left(\sum_{j=1}^n k_{\mathcal{Y}}(Y^{(i)}, Y^{(j)})\hat{w}(Y^{(j)}) \right)}{\sum_{j=1}^n k_{\mathcal{X},\mathcal{Y}}((X^{(i)}, Y^{(i)}), (X^{(j)}, Y^{(j)}))\hat{w}(Y^{(j)})} \right) \hat{w}(Y^{(i)}) .$$

506 Versions with nearest-neighbors density estimation can also be easily adapted. For
507 instance, the k -th nearest-neighbor distance of the point $(x, y) \in \mathcal{X} \times \mathcal{Y}$ is the smallest
508 distance d_k such that the cumulative sum of the weights of the points within d_k
509 distance to (x, y) reaches k . If copula transforms are used, recall that they are also
510 modified by weighted probabilities.

511 **5. Numerical Illustrations.** We conduct here numerical illustrations and com-
512 parisons of the different adapted tools that we propose for target and conditional
513 sensitivity analysis. These concise examples also demonstrate that target and condi-
514 tional sensitivity analysis explore aspects of a model which are both different from
515 global sensitivity analysis and valuable for practitioners.

516 Note that all the above tools are implemented in the language R, interfaced with
517 C++ for some routines; we intend to integrate them to the *Sensitivity* package of R.

518 **5.1. Presentation of Test Case Functions.** To illustrate target and condi-
 519 tional sensitivity analysis, we first propose a model with a simple but strong nonlin-
 520 earity, which we call *minimum-normal-uniform*. It is defined in dimension $d \stackrel{\text{set}}{=} 2$, with
 521 $f: x \mapsto \min(x_1, x_2)$, with independent factors conveniently noted $X_1 \stackrel{\text{set}}{=} N$ and $X_2 \stackrel{\text{set}}{=} U$,
 522 following respectively a standard normal distribution, and a uniform distribution over
 523 $[0,1]$.

524

525 We then explore the more complicated Ishigami–Homma model which is well-
 526 known from the sensitivity analysis community. This model is defined in dimension
 527 $d \stackrel{\text{set}}{=} 3$ by

$$528 \quad f: x \mapsto \sin(x_1) + a \sin^2(x_2) + bx_3^4 \sin(x_1),$$

529 where $a, b \in \mathbb{R}_+$; all factors (X_1, X_2, X_3) are independent and uniformly distributed
 530 over $[-\pi, \pi]$. The influence of the factor X_2 is purely additive, its importance being
 531 modulated by the parameter a . The influence of the factor X_1 includes an additive
 532 part and an interaction with the factor X_3 , the balance being tuned by parameter
 533 b . We set here the parameters $a \stackrel{\text{set}}{=} 5$ and $b \stackrel{\text{set}}{=} 0.1$, so that first-order Sobol’ indices
 534 are $\eta^2(X_1, Y) = 0.40$, $\eta^2(X_2, Y) = 0.29$ and $\eta^2(X_3, Y) = 0$, while total-order ones are
 535 $1 - \eta^2(X_{c_{\{1\}}}, Y) = 0.71$, $1 - \eta^2(X_{c_{\{2\}}}, Y) = 0.29$, and $1 - \eta^2(X_{c_{\{3\}}}, Y) = 0.31$.

536

537 In both models, we suppose that the critical domain \mathcal{C} is defined by Y exceeding
 538 a given critical value: $\mathcal{C} \stackrel{\text{set}}{=} \{y \in \mathcal{Y} \mid y \geq c\}$, chosen as the *ninth decile* of Y computed
 539 empirically, $c \stackrel{\text{set}}{=} F_{Y,n}^{-1}(0.9)$. Recall that target and conditional sensitivity measures are
 540 defined via weight functions $w: \mathcal{Y} \rightarrow [0,1]$ which depends on \mathcal{C} . In both models, we
 541 use the indicator function $1_{\mathcal{C}}$, and a smooth relaxation in accordance with the notion
 542 of distance over the reals,

$$543 \quad (5.1) \quad w_{\mathcal{C}}: y \mapsto \exp\left(-\frac{\max(c - y, 0)}{s \sigma_Y}\right);$$

544 where σ_Y is an estimation of the standard deviation of Y , and $s \stackrel{\text{set}}{=} 1/5$ is a factor
 545 tuning the smoothness, chosen so that $w_{\mathcal{C}}$ almost vanishes one standard deviation
 546 away from \mathcal{C} .

547 **5.2. Tested Target and Conditional Sensitivity Tools.** Among the large
 548 choice of interesting sensitivity measures, we consider those in [Table 5.1](#). Correlation
 549 ratios estimated with pick-and-freeze factors combinations are included because they
 550 are currently the most popular for global sensitivity analysis. Recall however from
 551 [section 4.2.2](#) that they do not allow for proper conditional versions, because conditioning
 552 introduces dependence between factors. Consequently, we use what we call the “hybrid”
 553 version. We report here results only for the first-order indices, but we can mention that
 554 the total-order indices behave similarly for target and conditional sensitivity analysis
 555 of both analytical models.

556 Then, we include the quadratic dependence measure with Gaussian kernel, and
 557 the mutual information dependence measure with truncated nearest-neighbors copula
 558 density estimation [Blumentritt and Schmid \(2012\)](#). For the hard target versions, recall
 559 that $1_{\mathcal{C}}(Y)$ is a discrete random variable over $\{0, 1\}$. For the mutual information,
 560 its law is estimated by empirical frequencies and the law of the joint $(X_i, 1_{\mathcal{C}}(Y))$
 561 is estimated by conditioning. For the quadratic dependence measure, we use a categorical
 562 kernel for $k_{\{0,1\}}$.

Table 5.1: Sensitivity measures used for target and conditional analysis experiments. The generic weight function w is either $1_{\mathcal{C}}$, or the smooth relaxation $w_{\mathcal{C}}$ defined in (5.1).

Notation	Definition	Expression for factor i
$S_{\text{PF}}^{(1,\text{tgt},w)}$	First-order correlation ratio target sensitivity measure	$\eta^2(X_i, w(Y))_n$
$S_{\text{PF}}^{(1,\text{hbd},w)}$	First-order correlation ratio hybrid sensitivity measure	$\eta^2(X_i, Y_w)_n$
$\text{QDM}_{\text{G}}^{(\text{tgt},w)}$	Normalized target kernel quadratic dependence measure	$\overline{\text{QDM}}_{k_{\mathcal{X}},k_{w(Y)}}(X_i, w(Y))_n$
$\text{QDM}_{\text{G}}^{(\text{cnd},w)}$	Normalized conditional kernel quadratic dependence measure	$[\text{P}^{w(Y)}] \overline{\text{QDM}}_{k_{\mathcal{X}},k_Y}(X_i, Y)_n$
$\text{MI}_{\text{c,nn}}^{(\text{tgt},w)}$	Normalized target mutual information	$\overline{\text{sCDM}}_{-\log}(X_i, w(Y))_{k_{\text{nn}},n}$
$\text{MI}_{\text{c,nn}}^{(\text{cnd},w)}$	Normalized conditional mutual information	$[\text{P}^{w(Y)}] \overline{\text{sCDM}}_{-\log}(X_i, Y)_{k_{\text{nn}},n}$

563 **5.3. Numerical Experiments and Results.** For each model, we draw hun-
564 dred different samples of size $n \stackrel{\text{set}}{=} 1000$ and schematize the resulting distribution of
565 each conditional or target sensitivity measure, together with their global sensitivity
566 counterpart, with Tuckey box plots on Figures 5.1 and 5.2.

567 On the minimum-normal-uniform model, the critical value is $c = 0.62$. The global
568 analysis, on Figures 5.1(a) to 5.1(c), is unanimous: the factor N is much more important
569 than the factor U . This is not surprising, since N presents more variability and takes
570 values far below the minimum of U .

571 The target analysis indicates that the ordering of the factors is the same, although
572 the relative importance difference is less drastic. This is again not surprising because N
573 has a higher probability to be below the critical value than U , hence still determining
574 again the outcome of interest here, but in the same time, the variability of N below the
575 threshold has no influence anymore. The correlation ratio on Figure 5.1(d), is much
576 less precise than the dependence measures (e) and (g). The target mutual information
577 shows an important bias, but this does not impact the ordering of the factors. It
578 can be noted that the smoothed versions present less variability while still ordering
579 correctly the factors. However, it is unclear if this is thanks to better behavior of the
580 smooth estimator, or simply because the estimated smoothed quantity is some kind of
581 interpolation between target and global measures. In the latter case, this effect would
582 turn out unfavorable if the ordering of the factors were different in both analysis. The
583 smoothed target mutual information on Figure 5.1(i) is clearly problematic, as it yields
584 the same importance measures as the global version (c). This can be explained by the
585 fact that the density estimation is based on copula transforms, and that Y and $w_{\mathcal{C}}(Y)$
586 have very similar copula transforms with the level of smoothing that we used; in this
587 case and for this particular estimator, smoothing is not judicious.

588 The conditional analysis tells a whole different story: now U is more important
589 than N . Indeed, conditionally to both U and N being no less than c , U varies in $[c,1]$
590 while N varies in $[c,+\infty[$, in such a way that the former has more chance to determine

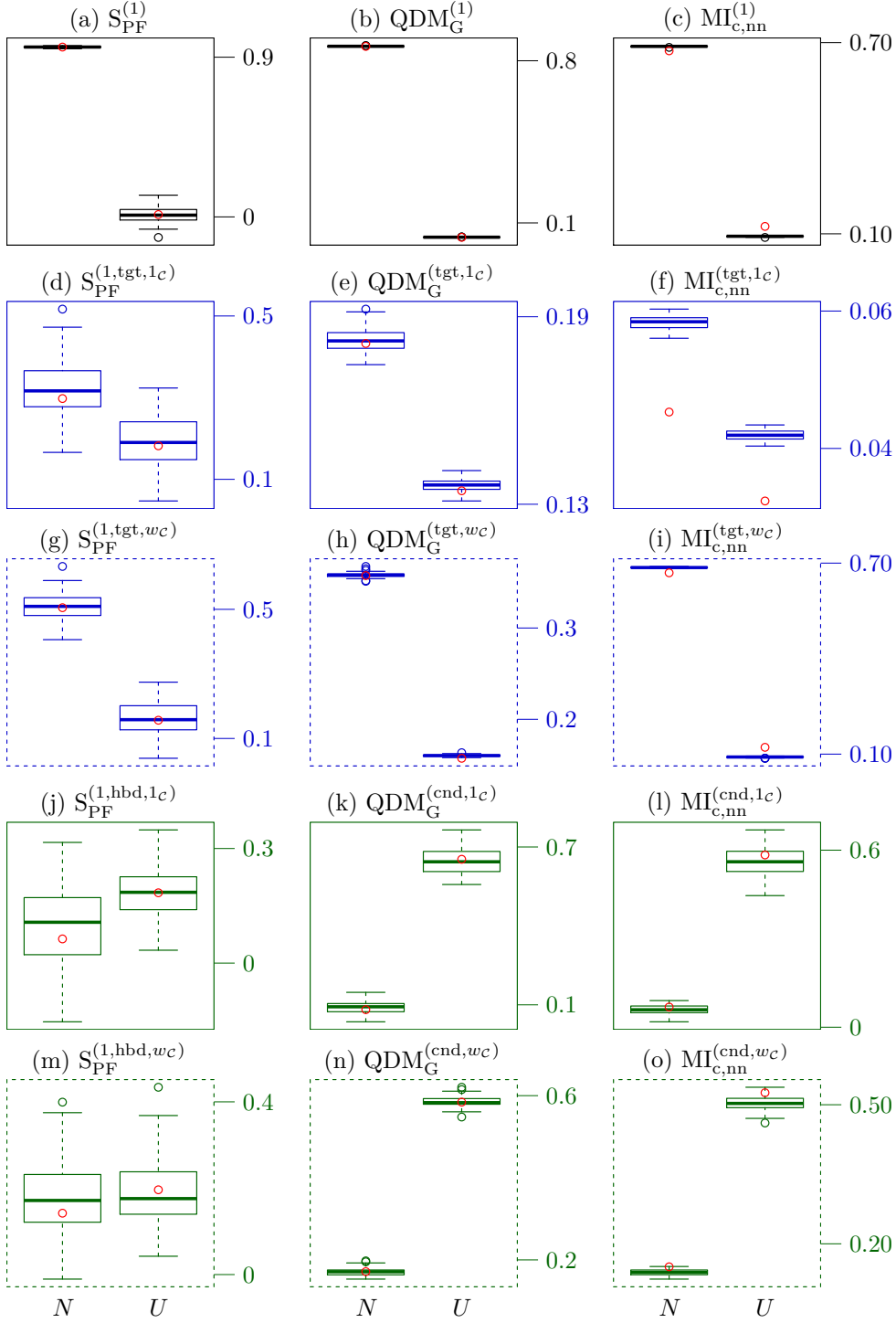


Figure 5.1: Global (black), target (blue) and conditional (green) sensitivity analysis of minimum-normal-uniform model on samples of size $n^{\text{set}} = 1000$. Red circles are asymptotic values estimated on samples of size $n = 10000$.

591 the value of their minimum. This is clearly captured by both dependence measures
 592 considered, on Figures 5.1(k) and (l), and moreover their smoothed conditional versions
 593 improve perceptibly their precision, as indicated by the relative height of the box plots
 594 in Figures 5.1(n) and (o), Hybrid correlation ratio adapted to pick-and-freeze estimator
 595 follows the same trend on Figures 5.1(j) and (m), but precision is not satisfying at all.

596 For the Ishigami–Homma model, the critical value is $c = 6.31$. Here, the relative
 597 importances of the factors are different in each analysis case. In the global analysis, the
 598 factor X_1 is the most important, and the factors X_2 and X_3 have lower importance,
 599 being ranked differently according to different sensibility measures (Figures 5.2(a), (b)
 600 and (f)).

601 In the target analysis, X_3 has now similar importance to X_1 , while X_2 has much
 602 less. Indeed, the combined effect of X_1 and X_3 easily exceeds the critical value, while the
 603 isolated action of X_2 can merely approach the critical value (recall that the parameter
 604 $a \stackrel{\text{set}}{=} 5$ is significantly less than c). As previously, the dependence measures offer more
 605 precision than the correlation ratio with pick-and-freeze estimator. It can be noted that
 606 they do not agree exactly on the relative importance of X_1 and X_3 on Figures 5.2(e)
 607 and (f), and that the target kernel quadratic dependence measure does not differ much
 608 from its global version in (b). Once again, the smoothed versions for target analysis
 609 are not particularly relevant: even if they seem to slightly reduce the variability of the
 610 estimators of kernel quadratic dependence measures, they completely fail to improve
 611 the estimators of the mutual information computed through copula density.

612 In the conditional analysis, X_3 becomes the dominant factor: being raised to the
 613 fourth power, the corresponding term presents steep derivatives in the regime of high
 614 values. The mutual information on Figure 5.2(l) seems the most suitable method
 615 for putting this into evidence. Once again for conditional analysis, the smoothing
 616 techniques do improve the quality of both dependence measures considered, even
 617 enabling kernel quadratic dependence measure to capture the dominance of X_3 .

618 **6. Conclusion.** In the context of sensitivity analysis of complex phenomena
 619 in presence of uncertainty, this work motivates and precises the idea of orienting
 620 the analysis towards a critical domain of the studied phenomenon. This gives rise
 621 to the notions of target and conditional sensitivity analysis. We show that many
 622 concepts in the literature relate to them, although usually in more specific frameworks
 623 depending on considered applications. Building up on modern statistical tools, we define
 624 mathematically a broad range of sensitivity measures which make as few assumptions
 625 as possible on the model at hand, while remaining flexible enough to be adapted to
 626 many particular situations.

627 To provide dedicated tools for target and conditional sensitivity analysis, we focus
 628 our attention on the popular sensitivity indices based on correlation ratio, namely
 629 Sobol’ indices, and on dependence measures which seem to us particularly well-adapted
 630 to our problematic. More particularly, we consider two dependence measures: the kernel
 631 quadratic dependence measure also called Hilbert–Schmidt independence criterion and
 632 the Csiszár divergence dependence measure, the mutual information being a particular
 633 case of the latter. For these different selected sensitivity measures, we propose adapted
 634 versions for target and conditional analysis, by considering transformation of the
 635 output using hard or smooth weight functions. We also propose an hybrid version for
 636 correlation ratio.

637 The proposed tools are illustrated and compared on analytical test cases. These
 638 experiments on synthetic data clearly illustrate the interest of target and conditional
 639 sensitivity analysis which can differ from global one. They also show that dependence

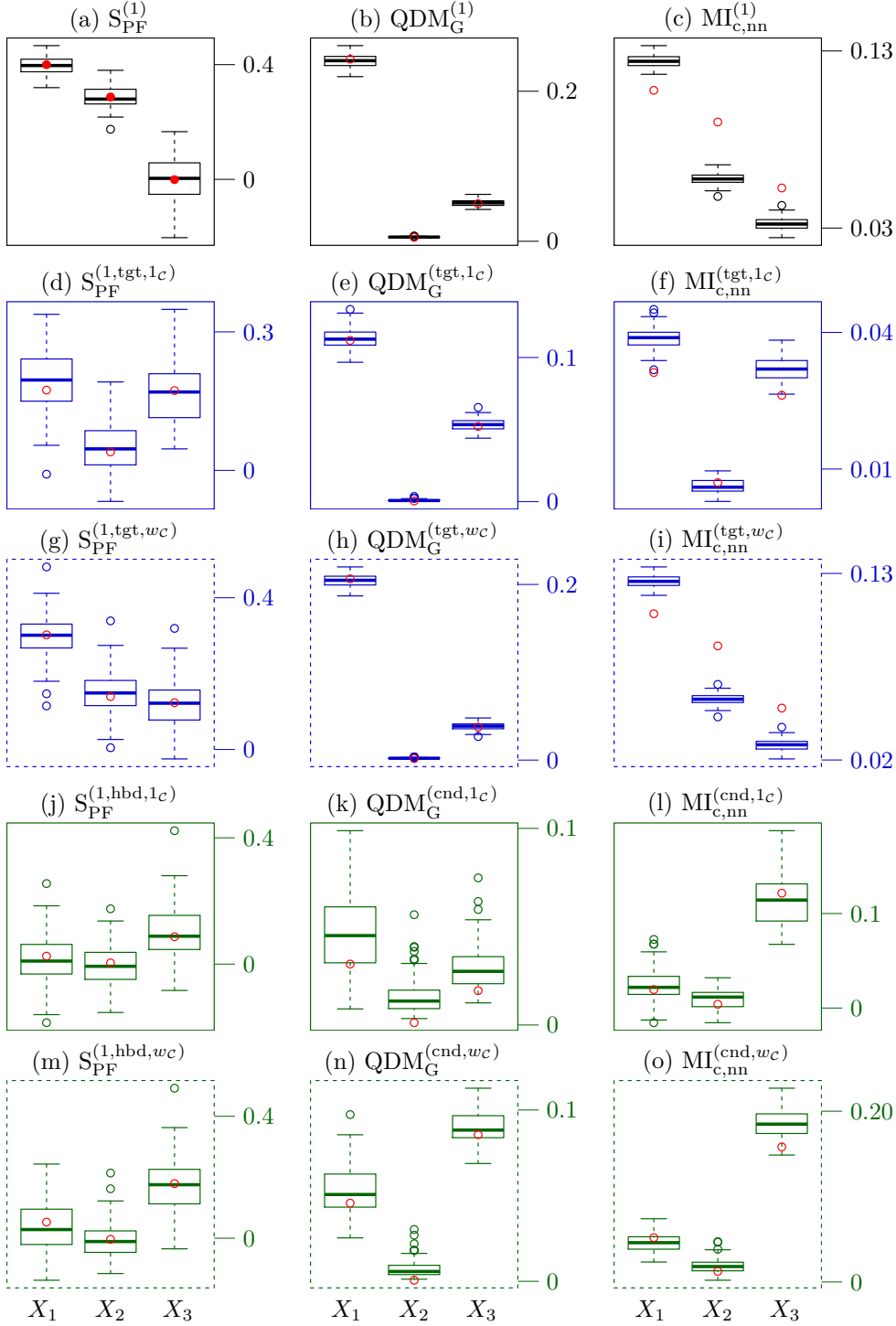


Figure 5.2: Global (black), target (blue) and conditional (green) sensitivity analysis of Ishigami–Homma model on samples of size $n^{\text{set}} = 1000$. Filled red dots are analytical values, hollow red circles are asymptotic values estimated on samples of size $n^{\text{set}} = 10000$.

640 measures are well suited for this task and are more precise than the correlation ratio.
641 Our preliminary results favor the use of kernel quadratic dependence measures rather
642 than correlation ratio. The mutual information with truncated nearest-neighbors cop-
643 ulla density estimation is also relevant (low variability and good capacity to capture
644 influence), but more adjustments should be required to reduce its bias. Furthermore,
645 even if more numerical explorations are necessary before drawing further conclusions,
646 the proposed smooth versions of estimators seem clearly suited for conditional estima-
647 tors, especially when the number of available observations in the critical domain is
648 low. However, their use for target sensitivity analysis remains questionable yet.

649

650 Altogether, this work is a good starting point towards sensitivity measures which
651 are more powerful and more adapted to questions raised by experimenters. There is still
652 much to do before actually establishing good practice. Naturally, we do not pretend to
653 exhaustiveness, since we cannot evaluate in this work all existing dependence measures.
654 Other popular approaches of global sensitivity analysis could be adapted to target or
655 conditional sensitivity analysis. We voluntarily set those aside for brevity, but other
656 approaches such as the regional sensitivity analysis ought to be more deeply studied;
657 e.g. by considering other measures of discrepancy between probability distributions
658 rather than Kolmogorov distance.

659

660 Then, it is important to test the target and conditional sensitivity measures in
661 more challenging situations, in particular where the critical probability is low, or to put
662 it otherwise, where less critical observations are available. In that respect, we believe
663 that the smoothing technique is promising, if correctly tuned. Last but not least, all
664 these sensitivity measures can only be completely assessed through confrontation to
665 real data.

666

667 **Acknowledgement.** The authors are grateful to Sébastien Da Veiga for sug-
668 gesting the use of dependence measures and for fruitful discussions about target
669 sensitivity analysis, in the framework of the MASCOT-NUM Research Group. We also
668 thank Bertrand Iooss for his helpful advice, within the project "Uncertainty and model
669 validation" of the CEA-EDF-AREVA Institute.

References.

- 670
671 Sophie Achard, Dinh Tuan Pham, and Christian Jutten. Quadratic dependence measure
672 for nonlinear blind sources separation. In *International Conference on Independent*
673 *Component Analysis and Blind Source Separation*, pages 263–268, 2003.
- 674 Siu-Kui Au and James L. Beck. Estimation of small failure probabilities in high
675 dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–
676 277, 2001.
- 677 A. Berlinet and C. Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability*
678 *and Statistics*. Springer US, 2003.
- 679 Thomas Blumentritt and Friedrich Schmid. Mutual information as a measure of
680 multivariate association: analytical properties and statistical estimation. *Journal of*
681 *Statistical Computation and Simulation*, 82(9):1257–1274, 2012.
- 682 E. Borgonovo. A new uncertainty importance measure. *Reliability Engineering &*
683 *System Safety*, 92(6):771–784, 2007.
- 684 Thomas Browne, Jean-Claude Fort, Bertrand Iooss, and Loïc Le Gratiet. Estimate of
685 quantile-oriented sensitivity indices. *HAL preprint*, 2017.
- 686 F. Cérou, P. Del Moral, T. Furon, and A. Guyader. Sequential Monte Carlo for rare
687 event estimation. *Statistics and Computing*, 22(3):795–808, 2012.
- 688 I. Csiszár. A class of measures of informativity of observation channels. *Periodica*
689 *Mathematica Hungarica*, 2(1):191–213, 1972.
- 690 Sebastien Da Veiga. Global sensitivity analysis with dependence measures. *Journal of*
691 *Statistical Computation and Simulation*, 85(7):1283–1305, 2015.
- 692 Matthias De Lozzo and Amandine Marrel. New improvements in the use of dependence
693 measures for sensitivity analysis and screening. *Journal of Statistical Computation*
694 *and Simulation*, 86(15):3038–3058, 2016.
- 695 Cees Diks and Valentyn Panchenko. Nonparametric tests for serial independence based
696 on quadratic forms. *Statistica Sinica*, 17(1):81–98, 2007.
- 697 Jean-Claude Fort, Thierry Klein, and Nabil Rachdi. New sensitivity analysis subordi-
698 nated to a contrast. *HAL preprint*, 2013.
- 699 Arthur Gretton, Olivier Bousquet, Alex Smola, and Bernhard Schölkopf. Measuring
700 statistical dependence with hilbert-schmidt norms. In Sanjay Jain, Hans Ulrich
701 Simon, and Etsuji Tomita, editors, *Algorithmic Learning Theory*, pages 63–77, Berlin,
702 Heidelberg, 2005. Springer Berlin Heidelberg.
- 703 Ronald L. Iman and Stephen C. Hora. A robust measure of uncertainty importance
704 for use in fault tree system analysis. *Risk Analysis*, 10(3):401–406, 1990.
- 705 T. Ishigami and T. Homma. An importance quantification technique in uncertainty
706 analysis for computer models. In *International Symposium on Uncertainty Modeling*
707 *and Analysis*, pages 398–403, 1990.
- 708 Harry Joe. Relative entropy measures of multivariate dependence. *Journal of the*
709 *American Statistical Association*, 84(405):157–164, 1989.
- 710 Sergei Kucherenko and Shufang Song. Quantile based global sensitivity measures.
711 *arXiv preprint*, 2016.
- 712 Paul Lemaître. *Analyse de sensibilité en fiabilité des structures*. PhD thesis, University
713 of Bordeaux, 2014.
- 714 Véronique Maume-Deschamps and Ibrahima Niang. Estimation of quantile oriented
715 sensitivity indices. *HAL preprint*, 2017.
- 716 Chang K. Park and Kwang-Il Ahn. A new approach for measuring uncertainty impor-
717 tance and distributional sensitivity in probabilistic safety assessment. *Reliability*
718 *Engineering & System Safety*, 46(3):253–261, 1994.
- 719 Rüdiger Rackwitz. Reliability analysis—a review and some perspectives. *Structural*

- 720 *Safety*, 23(4):365–395, 2001.
- 721 Hugo Raguet and Amandine Marrel. Target and conditional sensitivity analysis with
722 emphasis on dependence measures. *preprint*, 2018.
- 723 Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni,
724 Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global Sensitivity Analysis:
725 The Primer*. Wiley-Interscience, 2008.
- 726 Dino Sejdinovic, Bharath Sriperumbudur, Arthur Gretton, and Kenji Fukumizu.
727 Equivalence of distance-based and rkhs-based statistics in hypothesis testing. *The
728 Annals of Statistics*, 41(5):2263–2291, 2013.
- 729 C. E. Shannon. A mathematical theory of communication. *The Bell System Technical
730 Journal*, 27:379–423, 1948.
- 731 B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Monographs
732 on Statistics and Applied Probability. Springer US, 1986.
- 733 I. M. Sobol'. Sensitivity estimates for nonlinear mathematical models (in russian).
734 *Matematicheskoe Modelirovanie*, 2:112–118, 1990.
- 735 I. M. Sobol'. Sensitivity estimates for nonlinear mathematical models. *Mathematical
736 Modeling and Computational Experiment*, 1:407–414, 1993.
- 737 Shufang Song, Zhenzhou Lu, and Hongwei Qiao. Subset simulation for structural
738 reliability sensitivity analysis. *Reliability Engineering & System Safety*, 94(2):658–
739 665, 2009.
- 740 R.C. Spear and G.M. Hornberger. Eutrophication in peel inlet II. identification of
741 critical uncertainties via generalized sensitivity analysis. *Water Research*, 14(1):
742 43–49, 1980.
- 743 Gábor J. Székely, Maria L. Rizzo, and Nail K. Bakirov. Measuring and testing
744 dependence by correlation of distances. *The Annals of Statistics*, 35(6):2769–2794,
745 2007.
- 746 Y.-T. Wu. Computational methods for efficient structural reliability and reliability
747 sensitivity analysis. *AIAA Journal*, 32(8):1717–1723, 1994.