Modelling of the Modelling Process for Sustainable Development

Andrea Saltelli[1] and Alessio Lachi[2](✉)

[1] UPF Barcelona School of Management, Barcelona, Spain
[2] Saint Camillus University, Rome, Italy
alessio.lachi@unicamillus.org

Abstract. Much discussions surrounds the issue of "analytic flexibility" seen in multi-analyst studies. In these, researchers crunch data following diverse plausible model structures as to produce some form of useful inference, only to find a "universe of uncertainty" when comparing their findings. Uncertainty in models, be those mathematical, statistical, or other, can arise from various sources, including parametric and structural factors. We discuss the issue of analytic flexibility in general, and how global sensitivity analysis can be usefully applied to tackle the problem. We illustrate the main concepts via a very simple example taken from sustainability science. Specifically we show with our example how uncertainty increases when transitioning from purely parametric variations to structural modifications in a simple three-factor model inspired by the Human Development Index. Global sensitivity analysis, performed using variance-based methods, reveals that structural uncertainty increases interaction effects, leading to a higher effective dimension and model complexity. These findings emphasize the importance of considering structural uncertainty in model assessments, including in applications where aggregation rules influence outcomes, such as for the case of many composite indicators for sustainability.

Keywords: Garden of forking paths · Global sensitivity analysis · Mathematical models · Human Development Index

1 Introduction

A considerable number of authors [1,2,7,21] have engaged in discussing the "analytic flexibility" seen in recent multi-analyst studies, whereby different research teams collectively engage to check whether they obtain the same result when they venture in a so-called "garden of forking paths" [4], a metaphor for the many choices modellers do in the construction of an analysis, from curating the data to choosing the analytic architecture of their study. Here, apparently inconsequential modelling choices ended up producing a large variation in the results. In a sense these dangers were clearly identified by Edward Leamer whose work [8] advocated global sensitivity analysis (GSA) to test the robustness of inference. As shown in

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
E. di Bella et al. (Eds.): SIS 2025, ISSSAS, pp. 115–120, 2025.
https://doi.org/10.1007/978-3-031-96736-8_20

[18] the problem of analytic flexibility can be properly tackled using global sensitivity analysis, that permit to explore the dimension of the garden *ex ante*, i.e. before a multi-analyst study is undertaken. GSA offers efficient methods for exploring multidimensional spaces (for example by proper experimental design and use of quasi-random sequences) that permit charting the "universe of uncertainty" hidden in multi-analyst studies. We use the term "modelling of the modelling process" ($MOMP$) [9,15,16] to designate the application of GSA to a garden of the forking path. A more captivating term proposed in the literature is "multiverse analysis" [2]. According to Pierre Duhem inferences from mathematical models of phenomena to real physical applications must also be demonstrated to be approximately correct when the assumptions of the model are only approximately true (see [3]). This position is known as Duhem's principle of stability, where instability can arise from different sources: 1) Uncertainty cascade: adding model parameters by growing the size of the model increases the uncertainty in the output [5,22]; 2) Butterfly effect [20]: a model is very sensitive to initial or boundary condition; 3) Hawkmoth effect [23]: A model is very sensitive to structural assumptions. In environmental modelling the concept of uncertainty cascade is also known as the O'Neil conjecture [11,13]. We wish here show with a simple example what happens when different types of uncertainty are involved. For example, what happens when one moves from exploring pure parametric uncertainty, whereby one sample input variables from given distribution, to structural uncertainty, when one samples structural aspects of a model. As we shall see, the more one moves away from straight parametric uncertainty and toward model uncertainty, the more the model uncertainty increase. When using GSA, complexity can be associated to the concept of effective dimension, intended as the true dimensionality of the problem: it grows with the number of parameters that are truly influential on the input and with the number and strength of their interactions. A model with k uncertain inputs may have an effective dimension as low as one and as large as k, with the latter case corresponding to higher complexity and larger uncertainty in the output [6]. Interestingly, in the multiverse analysis studied in [18], it was discovered that the variation in the results could not be explained by any factor in isolation, but only by interaction effects involving several factors. Hence no much variance could be seen by varying the input uncertain factors one at a time. Since this may be counter-intuitive to non-statisticians, we illustrate here how this can be taught with a didactic model.

2 Methods

We shall play here with a simple three-factor model where the factors can either be added or multiplied with one another. A practical occurrence of this kind of model is the Human Development Index [10], a composite indicator where three dimensions (wealth, health and education) are combined; before 2010 the HDI was built by adding the dimensions, and in subsequent editions the dimensions were instead multiplied, see [12] for a discussion of the relative sensitivity analyses. Let's start with a quasi-additive function (Case 1a, the original HDI) defined in Eq. (1):

$$y = (x_1 + x_2 + x_3)^{\frac{1}{3}}. \tag{1}$$

We allow x_1, x_2, and x_3 to vary according to predefined distributions, i.e., $f_i(x_i) \sim \mathscr{U}(0,1)$, and we use three quasi-random numbers θ_i defined in $[0,1]$, for $i \in \{1,2,3\}$, to sample them (Case 1). Similarly, let us consider the case of a multiplicative function (Case 1b, as the later version of HDI) defined in Eq. (2):

$$y = (x_1 \times x_2 \times x_3)^{\frac{1}{3}}. \tag{2}$$

As a second step (Case 2), see Eq. (3), we modify the model so that, instead of being a sum of the three variables, y results from different combinations of sums and products:

$$y = (x_1 \circledast x_2 \circledast x_3)^{\frac{1}{3}}, \tag{3}$$

where $\circledast$ represents either addition (+) or multiplication (×). We introduce a trigger ζ, which can take eight possible values, to determine the model type. This allows us to test all possible model combinations ($\mathbb{I}_{0,..,7}$) defined in Eq. (4):

$$y = \begin{cases} (x_1 + x_2 + x_3)^{\frac{1}{3}} & \text{if } \zeta = 0 \\ (x_1 + x_2 \times x_3)^{\frac{1}{3}} & \text{if } \zeta = 1 \\ (x_1 \times x_2 + x_3)^{\frac{1}{3}} & \text{if } \zeta = 2 \\ (x_1 \times x_3 + x_2)^{\frac{1}{3}} & \text{if } \zeta = 3 \\ [x_1 \times (x_2 + x_3)]^{\frac{1}{3}} & \text{if } \zeta = 4 \\ [x_2 \times (x_1 + x_3)]^{\frac{1}{3}} & \text{if } \zeta = 5 \\ [x_3 \times (x_1 + x_2)]^{\frac{1}{3}} & \text{if } \zeta = 6 \\ (x_1 \times x_2 \times x_3)^{\frac{1}{3}} & \text{if } \zeta = 7. \end{cases} \tag{4}$$

As a third step (Case 3), we use a binary trigger ξ to switch each variable x_i between two distributions that share the same mean and variance ($\mathbb{I}_{0,1}$) defined in Eq. (5):

$$f_i(x_i) \sim \begin{cases} \mathscr{U}(a = 0, b = 1) & \text{if } \xi = 0 \\ \mathscr{N}_{[0,1]}(\mu = \frac{1}{2}, \sigma^2 = \frac{1}{12}) & \text{if } \xi = 1. \end{cases} \tag{5}$$

3 Results

Let us now calculate Sobol' indices [17] on the model output y for each test case. The total order sensitivity index (T_i) gives the total strength of a factor, by itself plus via all its interactions with the other factors. Instead, the first order sensitivity index (S_i) gives the strength of a factor by itself, and is equal to the expected reduction in the variance of the output that would be achieved on average if a factor X_i could be fixed. In this work, we consider as model output the actual value obtained from y.

If we consider Case 1a (see first row, first column of Fig. 1), we observe that there are almost no interaction terms since the model is quasi-additive (Eq. (1)). In fact, the total indices are equal to the first-order indices. Instead, if we consider Case 1b (see first row, second column of Fig. 1), we observe that some

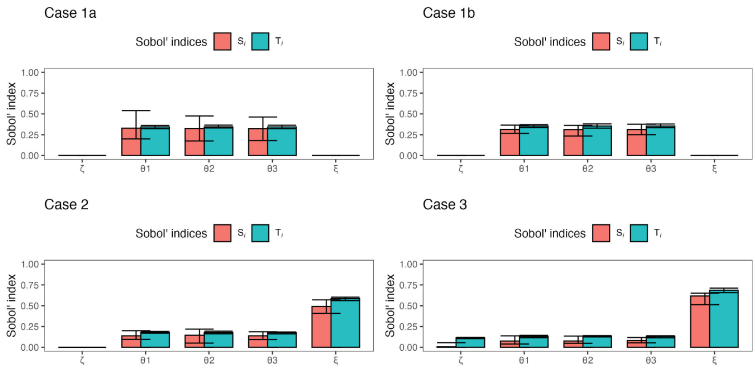

Fig. 1. Sobol indexes calculated on y for each test case.

interaction terms are present, since the model is multiplicative (Eq. (2)). If we consider Case 2 (see second row, first column of Fig. 1), where, with respect to Cases 1a and 1b we introduce the choice of the model function (trigger ξ), we observe that the structural uncertainty driven by ξ dominates the variance. If we consider Case 3 (see second row, second column of Fig. 1), where, with respect to Case 2, we introduce the choice of distribution (trigger ζ), we observe a slight increase in interaction terms flagged by the increase in effective dimension, see Table 1, while the variance remains dominated by the structural parameter ξ. With Case 3 the total variance decreases (Table 1) as a result of the introduction of normal distributions, that make the input more concentrated (less uncertain) than is the case of uniform distributions.

Table 1. Variance and coefficient of variation of Y. Mean dimension (MD) represents the index proposed by Hoyt and Owen (2021) [6] that quantify the force of the interaction among model inputs. MD is defined as $\sum_{i=1}^{k} T_i$, where T_i is the total order sensitivity index and k the number of model input considered in the analysis. Note that MD varies in $[1, k]$.

	Case 1a	Case 1b	Case 2	Case 3
Variance	0.019	0.038	0.072	0.052
Coefficient of Variation	0.122	0.462	0.336	0.278
Mean Dimension	1.032	1.067	1.117	1.193
K	3	3	4	5

4 Discussion

The results presented highlight how the nature of uncertainty influences the structure and complexity of model outputs. In particular, we observe that the introduction of structural uncertainty moving from Case 1 to Case 2 doubles the variance of the output and increases the effective dimension of the problem [6], suggesting that models incorporating structural variability inherently exhibit greater complexity. This didactic example underscores the importance of explicitly accounting for structural uncertainty when conducting sensitivity analysis: approaches limited to variations in input parameters may underestimate the full extent of uncertainty in real-world applications [14,19]. Finally, in practical applications such as composite indices, e.g., the Human Development Index [10], understanding the implications of different aggregation rules is critical. As discussed in [12], the choice between quasi-additive and multiplicative aggregation is not neutral and can significantly impact both the robustness and the policy implications of the measure. These results call for attention toward best practices for model selection, particularly in policy-relevant contexts where uncertainty quantification is crucial for decision-making.

References

1. Botvinik-Nezer, R., et al.: Variability in the analysis of a single neuroimaging dataset by many teams. Nature **582**(7810), 84–88 (2020). ISSN 1476-4687
2. Breznau, N., Rinke, E.M., Wuttke, A.: Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. PNAS **119**(44), e2203150119 (2022)
3. Fletcher, S.C.: The principle of stability. In: Philosopher's Imprint 20.3 (2020). Company: Michigan Publishing, University of Michigan LibraryDistributor: Michigan Publishing, University of Michigan LibraryInstitution: Michigan Publishing, University of Michigan LibraryLabel: Michigan Publishing, University of Michigan Librarypublisher-place: Ann Arbor, MIpublisher: University of Michigan. ISSN 1533-628X
4. Gelman, A., Loken, E.: The garden of forking paths. In: Working Paper Department of Statistics, Columbia University (2013)
5. Hillerbrand, R.: Climate simulations: uncertain projections for an uncertain world. J. General Philos. Sci./Zeitschrift für allgemeine Wissenschaftstheorie **45**, 17–32 (2014). ISSN 0925-4560
6. Hoyt, C., Owen, A.B.: Efficient estimation of the ANOVA mean dimension, with an application to neural net classification. SIAM/ASA J. Uncertainty Quantification **9**(2), 708–730 (2021)
7. Huntington-Klein, N., et al.: The influence of hidden researcher decisions in applied microeconomics. Econ. Inquiry **59**(3), 944–960 (2021). ISSN 1465-7295
8. Leamer, E.E.: Sensitivity analyses would help. Am. Econ. Rev. **75**(3), 308–313 (1985)
9. Piano, S.L., et al.: Unpacking the modeling process for energy policy making. Risk Anal. **44**(6), 1273–1284 (2023). ISSN 1539-6924
10. United Nations. Human Development Index. UN (2024)
11. O'Neill, R.V.: Error analysis of ecological models. In: Nelson, D.J. (ed.) Radionuclides in Ecosystems, Proceedings of the Third National Symposium in Radioecology, Oak Ridge, Tennessee, May 1971, pp. 898–907 (1971)
12. Paruolo, P., Saisana, M., Saltelli, A.: Ratings and rankings: voodoo or science? J. R. Stat. Soc. Series A (Stat. Soc.) **176**(3), 609–634 (2013). ISSN 09641998

13. Saltelli, A.: Statistical versus mathematical modelling: a short comment. Nat. Commun. **10**, 1–3 (2019)
14. Saltelli, A., Annoni, P.: How to avoid a perfunctory sensitivity analysis. Environ. Model. Softw. **25**(12), 1508–1517 (2010). ISSN 13648152
15. Saltelli, A., Puy, A.: What can mathematical modelling contribute to a sociology of quantification? Humanities Soc. Sci. Commun. **10**(213) (2023)
16. Saltelli, A., Puy, A., Di Fiore, M.: Mathematical models: a state of exception. Int. Rev. Appl. Econ. (2024)
17. Saltelli, A., et al.: Global Sensitivity Analysis: The Primer. Wiley, Hoboken (2008)
18. Saltelli, A., et al.: Leamer reloaded: a long-neglected suggestion is rediscovered to apply global sensitivity analysis to a quantification before publishing it. MetaArXiv (2024)
19. Saltelli, A., et al.: Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. Environ. Model. Softw. **114**, 29–39 (2019). ISSN 1364-8152
20. Schuster, H.G.: Deterministic Chaos: An Introduction. English, 2nd Rev edition. Weinheim: Vch Pub. (1998). ISBN 978-0-89573-611-6
21. Silberzahn, R., et al.: Many analysts, one data set: making transparent how variations in analytic choices affect results. Adv. Methods Pract. Psychol. Sci. **1**(3), 337–356 (2018). ISSN 2515-2459
22. Wilby, R.L., Dessai, S.: Robust adaptation to climate change. Weather **65**(7), 180–185 (2010). ISSN 1477-8696
23. Winsberg, E.: Appendix: structural stability and the "hawkmoth effect". In: Philosophy and Climate Science, pp. 232–246. Cambridge University Press, Cambridge (2018). ISBN 978-1-107-19569-1