Journal of Environmental Informatics Letters 10(2) 63-73 (2023)

Journal of Environmental Informatics Letters

www.iseis.org/jeil

The Importance of Intelligent Colouring for Simulation Decomposition in Environmental Analysis

A. Alam[1], M. Kozlova[2], L. T. Leifsson[3], and J. S. Yeomans[4 *]

[1] *Economics Department, Queen's University, Kingston, ON K7L 3N6, Canada*
[2] *LUT Business School, LUT University, 53850 Lappeenranta, Finland*
[3] *School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47907, USA*
[4] *Schulich School of Business, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada*

Received 23 October 2023; revised 20 November 2023; accepted 29 November 2023; published online 22 December 2023

ABSTRACT. "Real world" risk analysis in environmental contexts frequently requires the need to contrast numerous uncertain factors simultaneously and to communicate difficult-to-capture interactions. Monte Carlo simulation modelling of complex environmental systems is frequently employed to integrate uncertain inputs and to construct probability distributions of the resulting outputs. Visual analytics and data visualization can then be employed for the processing, analyzing, and communicating of the influence of any multi-variable uncertainties on the system. The simulation decomposition (SimDec) analytical technique has recently been employed in the complex assessments of environmental systems. SimDec has proved to be beneficial in revealing interdependencies in complex models, lowering computational burdens, facilitating decision-maker perceptions, and especially, making analytical components visualizable. It has been demonstrated that many analytical findings would not have been revealed without the coloured visualizations provided by SimDec. However, an ad hoc colouring scheme of the distribution output is neither sufficient nor capable of producing much of the key visualizable information requisite for an effective SimDec analysis. Instead, an approach that has recently been referred to as an intelligent colouring has been proposed. This paper outlines, highlights, and demonstrates the importance of and best-practices in an intelligent colouring scheme needed for an effective SimDec analysis of complex environmental systems.

Keywords: Monte Carlo simulation, probability distribution, SimDec, global sensitivity analysis, interaction, intelligent colouring

1. Introduction

Environmental decision-making is an inherently complicated process that frequently requires the integration of numerous socio-economic, environmental, and political uncertainties into its process (Loughlin et al., 2001; Zechman and Ranjithan, 2004; Janssen et al., 2010). While specific aspects may be clearly quantifiable, typical environmental problems also contain many components that cannot be directly incor-porated in the underlying decision models (Hipel and Ben-Haim, 1999; Mowrer, 2000; De Kok and Wind, 2003; Brugnach et al., 2007; Matthies et al., 2007; Fuerst et al., 2010; Hipel and Walker, 2011; Castelletti et al., 2012; Lund, 2012; Walker et al., 2012; Deviatkin et al., 2020). Such complexities are further compounded when stochastic uncertainties predominate (Baetz, 1990; Yeomans, 2008; Gunalay et al., 2012; Farr et al., 2016; Han et al., 2017; Kozlova and Yeomans, 2019).

Monte Carlo simulation has been employed in a variety of environmental situations in order to better capture these uncertainties (Openshaw and Whitehead, 1985; Ridlehoover, 2004; Byer and Yeomans, 2007; Byer et al., 2009; CEAA, 2011; Vithayasrichareon and Macgill, 2012; Kim et al., 2013; Farr et al., 2016; Pianosi et al., 2016; Han et al., 2017). Simulation is used to reflect various potential system impacts based on the likelihood of certain events (Byer and Yeomans, 2007; Byer et al., 2009). An effective Monte Carlo analysis should incor- porate not only the ranges of realistic possible outcome, but also the distributional nature of how the identified risks actually "behave" between identified extremes (CEAA, 2011; Kozlova and Yeomans, 2019). Although Monte Carlo simulation has been applied to wide spectrums of problems, the approach to its output analysis has remained comparatively static (Law and Kelton, 2000; Kozlova and Yeomans, 2022a, b). While simulation models enable a merger of the stochastic behaviours directly into the analysis process, they do not supply any prescriptive mechanism for determining actual system solutions (Kozlova and Yeomans, 2022b). Visual analytics involves some form of graphical representation of data to analyze, process, uncover, and communicate relationships embodied within the data. Typically, simulation outputs appear in distributional form and, therefore, visualization becomes the most important decision-support component (Byer et al., 2009).

Simulation decomposition (SimDec) has been introduced as a Monte Carlo enhancement to extend data analysis by visually revealing cause-effect relationships between combinations of input variables on the corresponding simulated outputs (Ko-

* Corresponding author. Tel.: 416 736-5074.
E-mail address: syeomans@schulich.yorku.ca (J. S. Yeomans).

ISSN: 2663-6859 print/2663-6867 online
 doi:10.3808/jeil.202300118

zlova and Yeomans, 2019; Kozlova and Yeomans, 2022b). SimDec partitions Monte Carlo outputs into sub-distributions by pre-classifying the inputs into states, grouping these states into scenarios, and then determining which outputs arise from these scenarios (Kozlova and Yeomans, 2019, 2022b). SimDec reveals previously unrecognized connections between the inputs and the outputs by visually mapping the relative contributions from each scenario onto the output distribution (Kozlova and Yeomans, 2020). This visualization permits the consequences of different combinations of the states to become recognizable to decision-makers in a straightforward manner (Kozlova and Yeomans, 2021). As the states can reflect different degrees of risks, SimDec actually generates important actionable insights to support decision-making. It is these visual analytics proficiencies that contribute substantial benefits to SimDec's practical decision-support capabilities under "real world" circumstances. SimDec is completely generalizable to any Monte Carlo model with negligible additional computational overhead and, hence, can be readily used for decision-making in situations containing considerable stochasticity (Kozlova and Yeomans, 2019). Given that SimDec is a relatively straightforward extension of Monte Carlo methods, the complexity of any underlying models is not an impediment to its adoption. In fact, the more nonlinear and/or complex the underlying models are, the more comprehensive the contributed SimDec insights have tended to be (Kozlova and Yeomans, 2022a).

The simplicity and intuitiveness of SimDec has also demonstrated how to successfully bridge the ostensible global sensitivity analysis gap existing in science and industry (Saltelli et al., 2008; Sobol and Kucherenko, 2009; Pianosi et al., 2016; Saltelli et al., 2019; Saltelli et al., 2020; Lo Piano and Benini, 2022; Kozlova et al., 2023a, b, c). One of the key strengths of SimDec in comparison to other methods is that, by design, it produces an easily interpretable visual representation of the solution without suppressing the complexities and uncertainties in the underlying problem.

Ghaffarzadegan et al. (2011) have strongly advocated for the use of intuitive system dynamics models for policy-making and Monte Carlo simulation is a well-established technique for assessing system dynamics risks (Lehar, 2005; Huang et al., 2009; Teply and Klinger, 2015). The visual analytics of SimDec have been employed in a variety of environmentally-related system dynamics settings. One of the first applications of SimDec involved a renewable energy policy analysis, in which the policy design features were analyzed in terms of the incentives they created for renewable energy investors (Kozlova et al., 2017; Ruponen et al., 2021). In these applications, SimDec was able to produce important insights into how different features of the policy interact with each other and how the investment project must be designed in order to achieve the maximum benefit from the financial support under uncertain market conditions. Another policy application of SimDec (Kozlova and Yeomans, 2019) concerned the interaction of two mechanisms for carbon capture investment support: (i) direct subsidy and (ii) carbon market. It has been shown that their mutual introduction does not have an additive effect, but that one is likely to offset the other, thereby actually diminishing the effectiveness of the policy measures. Another environmental application of SimDec considered the carbon footprint of transportation pallets (Deviatkin et al., 2021). This study focused less on uncertainty and more on the variability of pallet usage conditions within a certain country. In the case application, SimDec revealed the driving factors behind the unit emissions and their peculiar interaction resulting from emission substitution computations for incinerating pallets. This paradox was actually shown to produce ill-designed environmental incentives for the companies. Electrification of regional aviation was analyzed with SimDec through the lens of increased flying range of electric aircraft as a function of improved battery capacity and motor power (Kozlova et al., 2021). As with the previous case (Deviatkin et al., 2021), a critical interaction was revealed via the SimDec analysis. Namely, investing in electric motor power development would not prove sensible with the existing levels of battery capacity, but would produce substantial benefits should the capacity be significantly improved. SimDec has also been applied to the sustainability of urban development (Yeomans and Kozlova, 2023), the uncertainty in geological systems (Kozlova and Yeomans, 2022b), and the functioning of ecosystems (Kozlova and Yeomans, 2022a). In all of these applications, the analysis revealed critical nonlinear behaviours and interactions.

Saltelli (2023) has recently referred to SimDec as an "*intelligent colouring*" approach due to its power to visually convey the sensitivity of any output-to-input changes in conjunction with its capabilities for uncovering various complicated interactions that are present within a model. However, while not inherently "incorrect" from an analytical perspective, some applications have neglected certain requisite "*intelligent colouring*" components in the decomposition, thereby suppressing several major information benefits of the methodology (Liu et al., 2022; Raul et al., 2022). In order to illustrate the relative deficiencies of discounting such best-practice visualization recommendations, the simple model of Liu et al. (2022) is decomposed and the analytical contributions of SimDec with and without intelligent colouring are compared. Consequently, the chief goal of this particular study is to demonstrate and stress the efficacy of ensuring the actual application of an intelligent colouring approach to SimDec in order to effectively advance its numerous visual analytics contributions to environmental decision-making.

2. Simulation Decomposition

As outlined in the introduction, SimDec provides an innovative, visual analytics approach to enhance Monte Carlo analysis by detecting previously concealed relationships and interactions inherent within the studied system (Kozlova and Yeoposes the simulated outputs by clustering selected inputs into states, creating a collectively exhaustive list of the combinations of these states, and then mapping the resulting state partitions onto the output distribution (Kozlova and Yeomans, 2019, 2022b; Deviatkin et al., 2021). Upon termination of the algorithm, one can observe the overall output distribution, as in classical Monte Carlo simulation, in conjunction with the simulta-

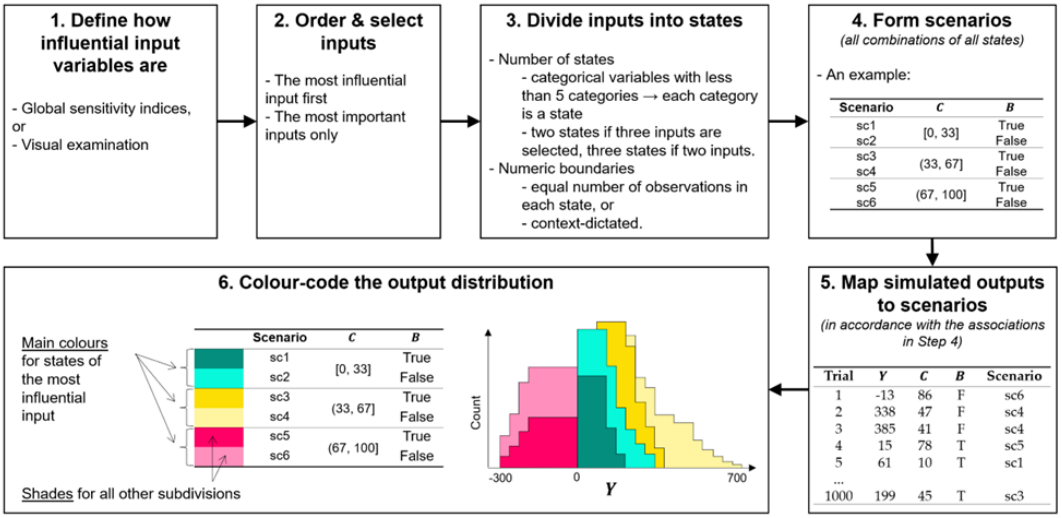


Figure 1. Schematic representation of the simulation decomposition approach adapted and modified from Kozlova et al. (2023c).

neous projections of the various input variable partitions onto this distribution (Kozlova and Yeomans, 2021, 2022b). Any need for multiple simulation runs is eliminated by noting the input variable values that generated each output during the single simulation run and by then creating the partitions based on the input variable states. Consequently, the additional computational overhead requirements of SimDec are insignificant (Kozlova and Yeomans, 2021).

In order to ensure that the colouring system implemented can be considered "intelligent", however, the states must be displayed by employing different colouring gradations of primary colours during the visual analytics stage (Saltelli, 2023). In particular, the variable partitions designated as most influential (or of most interest) need to be allocated dissimilar base colour schemes. The influential variables can be directly selected based on user preferences or determined via the calculation of Shapley values, Sobol indices, or other forms of sensitivity indices (Shapley et al., 1953; Cukier et al., 1973; Sobol, 1993; Sobol and Saltelli, 1995; Homma and Saltelli, 1996; Kozlova et al., 2023a). All remaining states in the decomposition (those considered less influential or of lesser interest) must be colour-coded according to shading gradations of this basic colour (e.g., darkest-to-lightest) to facilitate consistency in prevailing human perception and also to preserve visual acuity (Kozlova and Yeomans, 2020, 2022b). By visually displaying segmented multivariable groups as partitions of the output variable, SimDec can intuitively expose various interactions and nonlinearities previously concealed within the model variables (Kozlova et al., 2023a). This exposition frequently leads to the detection of unforeseen relationships (Kozlova and Yeomans, 2020, 2022b). Computational codes for the SimDec algorithm are available in several different programming languages on the simulation decomposition GitHub page (Kozlova et al., 2023b). The decomposition process for the SimDec algorithm is as follows:

Step 1: Define the importance of the input variables in the model. This can be done by either computing global sensitivity indices [the simple binning approach is recommended (Kozlova et al. (2023a)] or by iterative visual examination of the effect of each input variable. Appropriate visualization types would be scattered plots or single-variable SimDec.

Step 2: Select input variables for decomposition based upon how much they influence the output. The appropriate selection contains only the most important input variables in the order of their importance.

Step 3: Break up each selected variable into a set of states (for example, low-medium-high or optimistic-expected-pessimistic, etc.). The default formation of numeric boundaries ensures an equal number of datapoints in each state. Also, the state boundaries could be custom-defined to satisfy the peculiarities of a certain decision situation. The appropriate number of states for continuous numeric variables is (i) two if three or more input variables are chosen, (ii) three states for two input variables, and (iii) up to five for categorical variables where each category becomes a state.

Step 4: Form scenarios by creating an exhaustive list of the combinations of the states of the input variables. The total number of these scenarios is, thus, a combinatorial function of the number of states for each variable. For example, for C possessing three states [0, 33], (33, 67], and (67, 100] and B two states {'True', 'False'} we have six possible variable-state combinations.

Step 5: Record the values for each input variable while running the Monte Carlo Simulation. Define each scenario that matches each outcome based on the variable-state combinations. The correct sub-distribution for each iteration is achieved by mapping the randomly generated values of each individual input to their corresponding states and then mapping these states into the appropriate multivariable scenario.

Step 6: Plot a stacked histogram, where the scenarios form the series of the chart, and colour-code the figure based upon the identified states created by the partition combinations. The colour-coding should follow the principle of assigning the main distinct colour to the states of the most influential input variable

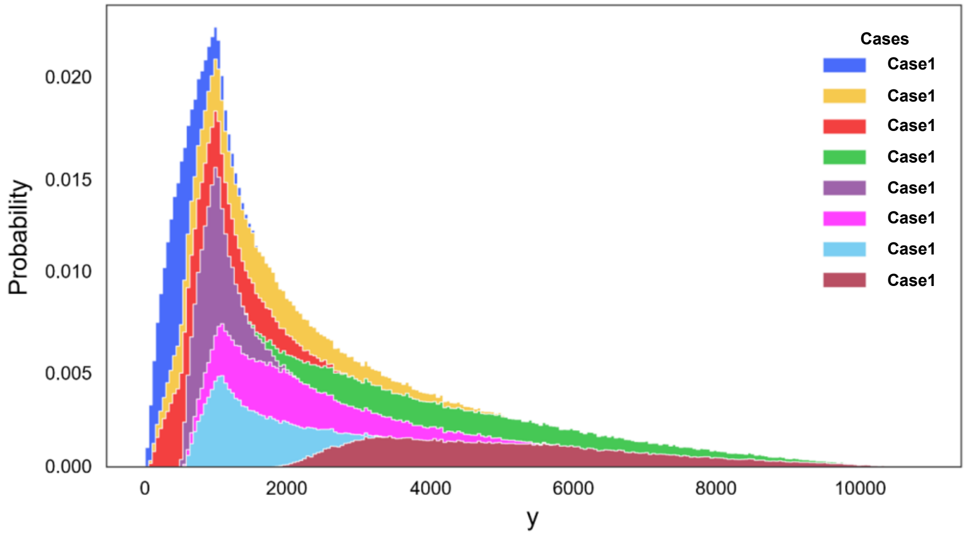


Figure 2. Simulation decomposition of the model problem (1) by Liu et al. (2022).

entering the decomposition first, and the shades of the main colours to the further subdivisions. Also, calculate the desired summary statistics for the full distribution and each group. For example, probability, weighted mean, standard deviation, minimum, and maximum. A detailed schematic representation of the SimDec algorithm appears in Figure 1. The SimDec algorithm has been implemented in several open-source code packages, including Python, R, Julia, and Matlab (Kozlova et al., 2023b).

3. Example Decomposition

Liu et al. (2022) performed a SimDec analysis of the agricultural food-water-energy system in the United States. In addition to decomposing their two large, "real world" environmental cases, the authors illustrated their approach to decomposition colouring through the example model:

$$y = x_1 + x_2 x_3^2 \tag{1}$$

in which the numeric assumptions of input variables are presented in Table 1.

Table 1. Numeric Assumptions of Input Variables

Input Variables	Distribution
x_1	U (1, 1000)
x_2	U (1, 100)
x_3	U (1, 10)

The simulated output distribution was constructed based on 106 iterations of the model. The decomposition was conducted using all three inputs, x_1, x_2, and x_3 (stated in order of their decomposition) by partitioning the numeric range of each input into two states of equal-sized sub-ranges. The corresponding decomposition partitioned the outputs into 8 (2 × 2 × 2) variable-state combinations [referred to "Cases" in Liu et al. (2022)] as shown in Figure 2.

Figure 2 had been constructed correctly from a technical decomposition perspective, and it guided the authors to the correct conclusion that x_3 was the most significant input. However, the graph itself provides very limited interpretability and readability due to the following shortcomings:

(1) The colour-coding does not follow an intelligent colouring scheme that initially separates the most influential input's states into distinctive main colours and assigns gradations of these colours to respective scenarios contained within them. Assigning a distinct base colour (together with the subsequent colour gradients) provides a means to visually project the impacts from input volatilities onto the output.

(2) The order of decomposition does not start with the most influential variable. This can lead to a misleading perception of the figure and, of much greater concern, can confound key interactions so much that they may remain essentially undetectable.

(3) The legend does not provide any attribution of input variables to specific scenarios.

Consequently, it is impossible to interpret which scenario maps onto which input states on the figure. This omission completely removes any ability to readily construe key insights from this single graphical visualization.

4. Best-Practice Decomposition and Its Variations

This section first replicates the data-generating process of Liu et al. (2022) and then visually contrasts the results from the other decompositions including (i) a decomposition using all of the input variables, (ii) a decomposition based upon a best-practice ordering of the inputs, (iii) a decomposition without the best ordering of the inputs, and (iv) a decomposition of each individual input variable.

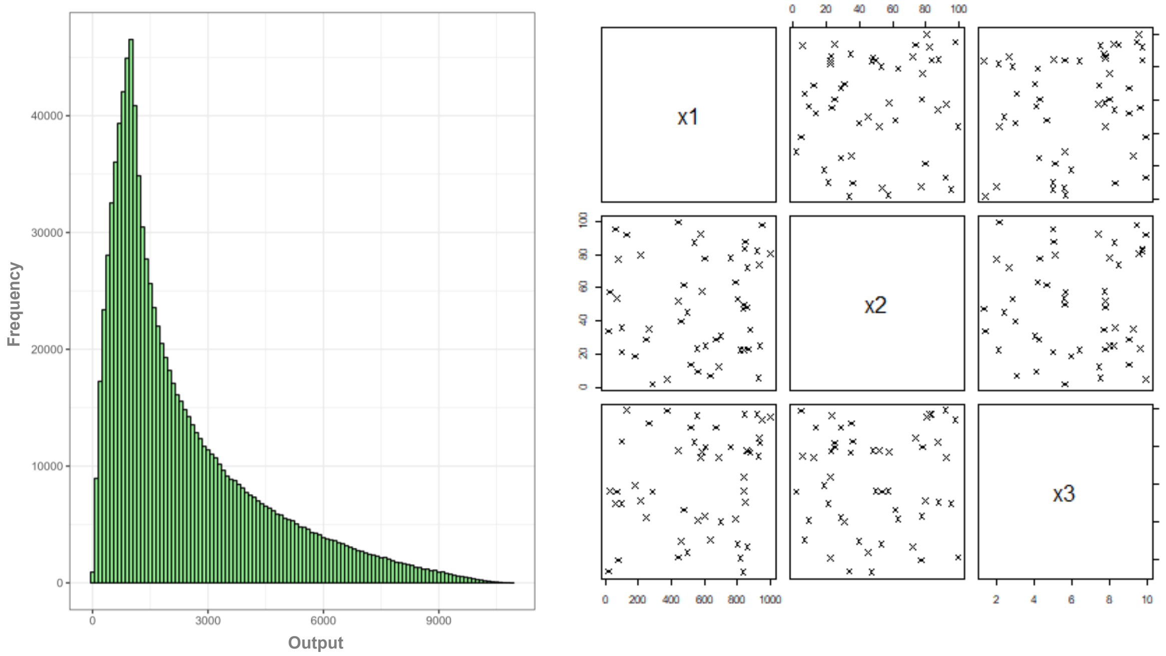
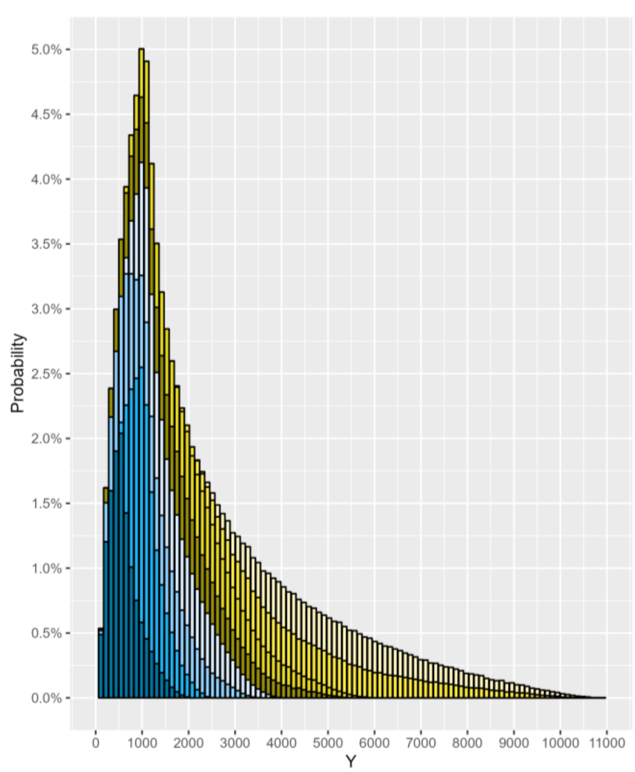


Figure 3. Reproduction of the simulation data of Liu et al. (2022).



Color	x3	x2	x1	min(Y)	mean(Y)	max(Y)	probability
1	low	low	low	3	565	1985	13%
2		low	high	504	1067	2509	12%
3		high	low	64	1172	3462	12%
4		high	high	556	1673	3999	13%
5	high	low	low	42	1845	5488	12%
6		low	high	540	2346	5972	13%
7		high	low	1579	4891	10353	12%
8		high	high	2086	5397	10921	13%

Figure 4. Decomposition on all input variables.

4.1. Replicating the Data Generating Process

This initial simulation experiment could be viewed as a general assessment of the dynamics of the model. The basic Monte Carlo simulation should produce an output distribution possessing an overall shape congruent to that of Figure 2 (but displayed using only a single colour).

To replicate the data from Liu et al. (2022), 10^6 observations of each input are randomly generated from the respective uniform distributions $x_1 \sim U$ (1,1000), $x_2 \sim U$ (1,100) and $x_3 \sim U$ (1,10), Table 1, using Latin Hypercube Sampling (McKay et al., 1979). These input variables are used to determine corresponding output values calculated using equation (1). The complete sample data with its output distribution appear in Figure 3. As expected, the shape of the output distribution in Figure 3 directly corresponds to the shape of Figure 2.

4.2. Decomposition by All Inputs (Suboptimal)

This sub-section conducts a decomposition on all inputs with the ordering based upon perceived level of influence. In any SimDec analysis, the overall distribution is broken down (or decomposed) into non-overlapping scenarios of input variable combinations that are stacked on top of each other in the figure. Altogether, all possible combinations of the state settings for the inputs form separate scenarios. The probability distribution of the model output is partitioned using these scenarios and judiciously colour-coded. The colouring logic is important to facilitate the overall interpretability of the visual perceptions. The states of the most influential input variables assume distinct main colours, and each of these main colours is then subshaded to further highlight the partitions. The visualized interpretation of the SimDec breakdown into attributes of the entire output dataset can be clarified using the histogram's legend. The legend assigns colour-shaded gradations to specific state combinations of the input variables, while also identifying each combination by assigning corresponding scenario indices to them.

Figure 4 decomposes the model using all three inputs with the variable decomposition ordering of x_3, x_2, x_1. Each variable is partitioned into low and high states. Specifically, with three variables each partitioned into two states, there are eight possible combinations of the input variable states in total. Each of these eight combinations is considered a scenario. The base colour scheme assigns blue to the low state of x_3 and yellow to the high state. Furthermore, to convey additional analytical information, the (i) minimum, (ii) mean, and (iii) maximum values of the output are computed for each scenario. Clearly, the minimum and maximum values correspond to the extreme edges of each scenario's coloured distribution on the horizontal axis within the histogram and provide a numerically convenient, ancillary interpretation of the figure.

Because of the visualization from the colouring scheme, the influence of x_3 on the output becomes immediately apparent. The darker blue colour gradations (which correspond to the low state of x_3) are all clustered on the leftmost portion of the distribution, while the yellow shades (representing the high state of x_3) are on the right with a long-tail stretching into the high output value region of the distribution. To visually determine influence from the remaining variables, one can examine the gradations of colour within each of the blue and yellow portions noting that the darker hues correspond to lower-valued states of x_1 and x_2. If x_1 and x_2 had little-to-no effect, then the distributions of the yellows and blues would appear "uniformly" distributed, with equal portions of the blue/yellow hues appearing stacked on top of each other. However, as the darker coloured (low state) hues are clustered toward the leftmost portions of both the blue and yellow regions with the lighter (high state) gradations to the right together with a relatively clear differentiation between the differently shaded colour blocks moving left-to-right the visualization further reveals that there is also an influential relationship between the output values and x_1 and/or x_2.

4.3. Best-Practice SimDec

Intuitively, by examining Equation (1), it can be seen that x_1 only acts like an intercept by shifting the output value in a type of levelling effect on the output. While the intelligent colouring shown in Figure 4 maps each scenario into its respective region, any interaction effect between x_3 and x_2 is not easily discernible due to the inclusion of x_1. Figure 5 provides a best-practices application of SimDec, in which only the most influential inputs are x_3 and x_2 are used for decomposition, with x_1 omitted from the scenario formation.

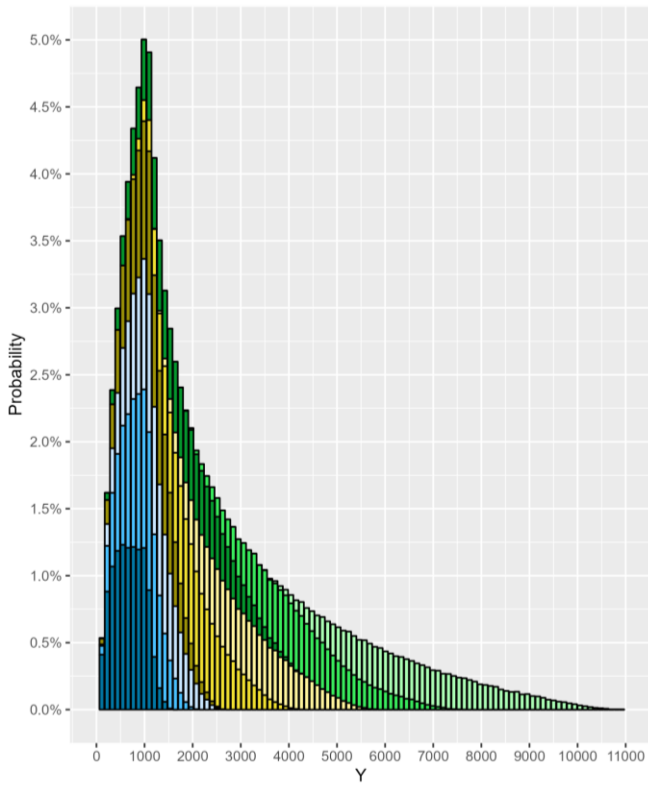
The best ordering for decomposition is x_3 followed by x_2, with each of the variables shown after being partitioned into the three states: low, medium, and high. There are 9 multivariate scenario combinations, with the base colours blue, yellow, and green assigned, respectively, to the low, medium, and high states of the primary variable x_3. The colour gradations for these base colours are adjusted to their respective shades reflecting each state of variable x_2.

In the resulting plot, the visualization advantages from using intelligent colouring become readily apparent. Firstly, the scenario data in the legend table can be clearly mapped to the stacked histogram making the plot easily interpretable even from a cursory glance. The distinct left-to-right blocks of blue /yellow/green regions clearly indicate the influence of the low /medium/high states of x_3 on the output. More importantly, however, the interaction effects between x_2 and x_3 can be visually deduced because, within each primary colour block, a distinct left-to-right pattern of darkest-to-lightest colour-gradation differentiation is clearly evident. Namely, it becomes readily apparent that the influence of x_2 increases as the value of x_3 increases [which is evident upon examination of the variable structure in Equation (1)]. Moreover, leaving out irrelevant variables from the decomposition has allowed one to more easily discern the various important underlying effects, which reinforces the importance of input variable ordering in the analysis. Visually, this is observable by the fact that the shades of green stretch further along the right tail of the distribution compared to the shades of yellow and blue.

In direct contrast to the visualizations from Figure 5, the interaction effect is not clearly discernible in the colours of Figure 2 because (i) Liu et al. (2022) have not employed intelligentc olouring, (ii) the number and order of variables used for the decomposition are not ideal, and (iii) the absence of a well-constructed scenario legend.

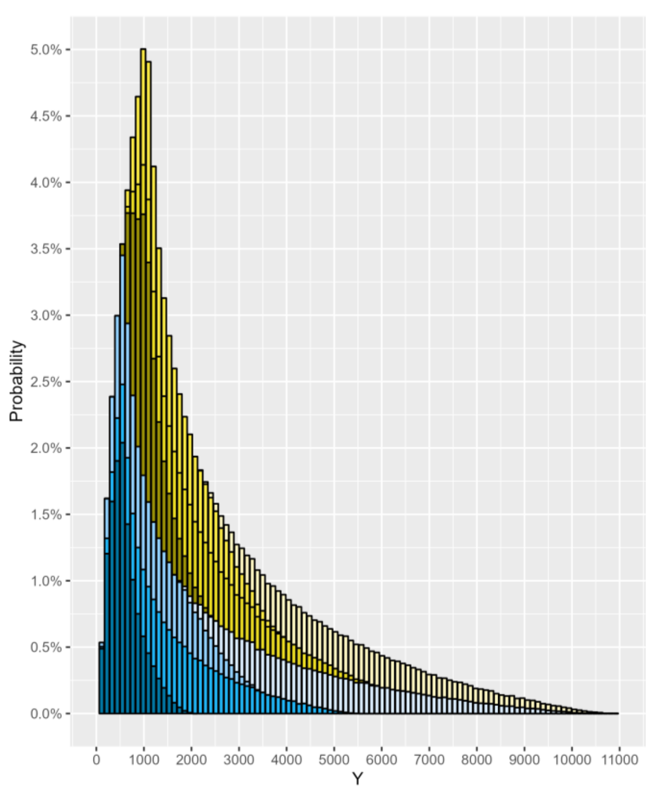
4.4. Decomposition with a Suboptimal Ordering

For illustrative purposes, the following case again contains



Color	x3	x2	min(Y)	mean(Y)	max(Y)	probability
1	**low**	low	3	621	1536	11%
2		medium	44	853	2054	11%
3		high	79	1086	2574	11%
4	**medium**	low	26	1041	2623	11%
5		medium	577	2068	4246	11%
6		high	1109	3086	5847	11%
7	**high**	low	66	1782	4297	11%
8		medium	1682	4188	7643	11%
9		high	3335	6583	10921	11%

Figure 5. The best-practice SimDec: decomposition by the most important variables in the order of their importance.



Color	x1	x2	x3	min(Y)	mean(Y)	max(Y)	probability
1	**low**	low	low	3	565	1985	13%
2		low	high	42	1845	5488	12%
3		high	low	64	1172	3462	12%
4		high	high	1579	4891	10353	12%
5	**high**	low	low	504	1067	2509	12%
6		low	high	540	2346	5972	13%
7		high	low	556	1673	3999	13%
8		high	high	2086	5397	10921	13%

Figure 6. Decomposition with a sub-optimal order of input variables that does not follow their sensitivity indices.

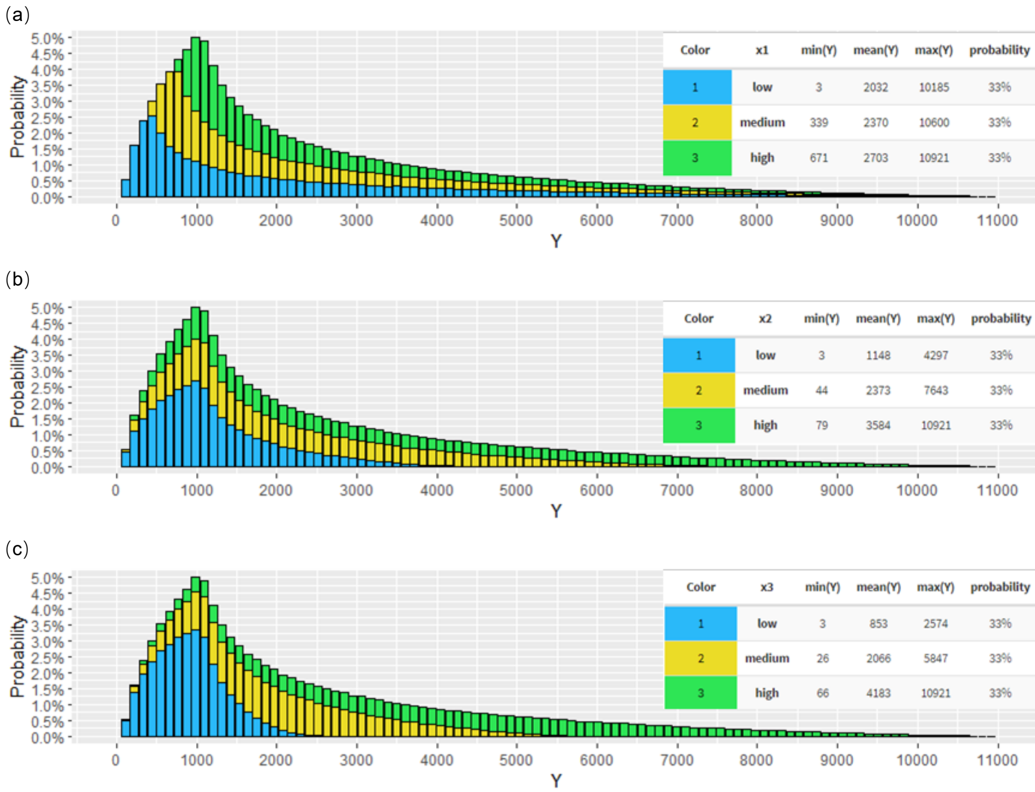


Figure 7. Decompositions by single input variable (a) SimDec plot of x_1; (b) SimDec plot of x_2; (c) SimDec plot of x_3 (note: the more the sub-distributions overlap, the less impact the input variable has on the output).

all inputs, but does not assign any importance to variable ordering in the decomposition. Each variable is decomposed into two states which creates 8 scenario combinations. Figure 6 clearly highlights a significant horizontal overlap of the colours. Although intelligent colouring maps the information from the scenario legend onto its respective region within the distribution, both the first-order effect of x_3 and the interactions between x_3 and x_2 become much harder to differentiate due to the order of decomposition employed. Consequently, in order to convey the most information, the decomposition ordering should start with the most influential variable whenever possible to prevent key interactions from becoming visually confounded within the figure.

4.5. Examining Individual Inputs with SimDec

The relative influence of each input variable on the output distribution can also be examined via decomposition. Figure 7 decomposes each variable into three states to highlight SimDec's relative abilities to assess the influence of individual inputs. The fact that x_1 only has a levelling effect is clearly evident not only from the legend table (note how the summary statistics of how the output changes with each state of x_1) but also from the layering of the colours on top of each other the low, medium, and high states cover the entire range of the distribution essentially exhibiting the same likelihoods, throughout. Namely, it is readily observable that there is not a strongly influential impact of this individual variable on the overall output. Furthermore, an equivalent lack-of-influence interpretation also applies to the individual decompositions of x_2 and x_3, respectively. The major analytical contribution of SimDec arises from the identification of the strong interaction effects detected between x_2 and x_3, as described in the earlier sections.

5. Conclusions

This paper has highlighted several differences between the best-practice approaches of the intelligent colouring analyses of SimDec to somewhat less meticulous applications of the colouring concept. Proper visual insights and interaction effects can only be achieved when the following elements of the SimDec algorithm have been incorporated into the analysis: (i) an appropriate assignment and gradation of colouring logic has been employed, (ii) an appropriate ordering of influential inputs has been adopted in the decomposition, and (iii) non-contributory input variables have been removed.

For visualization purposes, it is imperative that the scenarios of the states for the most influential variables be assigned distinct primary colours with all remaining states receiving gradations of these. SimDec has demonstrated that many analytical discoveries cannot be accomplished without appropriate colour scheme visualizations. This requires the best-practice adoption of the methods that have now become more broadly referred to as "intelligent colouring". Conversely, it is clear that an *ad hoc* colouring approach can prove neither sufficient nor capable of producing the key visualizable information requisite for an effective analysis. As demonstrated in this paper, variable "pruning" can also become a crucial factor in narrowing down and identifying important interactions.

The illustrative example used in this paper is a simple three-variable model that includes addition, multiplication, and power operations, components that are essential for any environmental system. And yet even with this simple model, it was shown how critical it is to follow the algorithmic colouring procedure to ensure the optimality of the information representation and to provide correct visual insights. This colouring is especially critical for systems that possess heterogeneous behaviour which are characteristic components inherent in most environmental models (Kozlova et al., 2023c).

Consequently, this paper has outlined, highlighted, and demonstrated why the specific application of many of the important best-practice intelligent colouring schemes would become absolutely essential in order to conduct a complete and effective SimDec analysis of more complex environmental systems. Future avenues related to SimDec research to environmental applications would include: (i) Designing an approach for using SimDec to help inform surrogate model building for otherwise too computationally heavy environmental models; (ii) Further developing the sensitivity indices computation algorithm employed behind the SimDec procedure; (iii) Utilizing SimDec's ability to incorporate dependent data into the design procedures for analyzing the model effects of aggregate variables (e.g., stages of LCA instead of individual inputs); (iv) Delving deeper into a computational world that combines nonlinear interactions and correlations within environmental models, the nature of which can already be conveniently exposed with SimDec.

Acknowledgments. This research was supported in part by grant OGP0155871 from the Natural Sciences and Engineering Research Council; and by grant #220178 from Finnish Foundation for Economic Foundation. The authors would like to thank the Editor, Associate Editor, and two anonymous reviewers for providing numerous helpful suggestions that have substantially improved the content of the paper.

References

Baetz, B.W. (1990). Optimization/simulation modeling for waste management capacity planning. *ASCE Journal of Urban Planning and Development.* 116(2), 59-79. https://doi.org/10.1061/(ASCE)0733-9488(1990)1162(59)

Brugnach, M., Tagg, A., Keil, F. and De Lange, W.J. (2007). Uncertainty matters: Computer models at the science-policy interface. Water Resources Management. 21, 1075-1090. https://doi.org/10.1007/ s11269-006-9099-y

Byer, P.H. and Yeomans, J.S. (2007). Methods for addressing climate change uncertainties in project environmental impact assessments. *Impact Assessment and Project Appraisal.* 25(2), 85-99. https://doi.org/10.3152/146155107X205841

Byer, P.H., Lalani, M.J. and Yeomans, J.S. (2009). Addressing and communicating climate change and its uncertainties in project environmental impact assessments. *Journal of Environmental Assessment Policy and Management.* 11(1), 29-50. https://doi.org/10.1142/S1464333209003245

CEAA (2011). *Decision Making under Uncertainties for Adapting to Climate Change in Project Environmental Assessments*. Canadian Environmental Assessment Agency's Research and Development Program (Accessed May 11, 2023).

Castelletti, A., Galelli, S., Restelli, M., and Soncini-Sessa, R. (2012). Data-driven dynamic emulation modelling for the optimal management of environmental systems. *Environmental Modelling and Software*. 34(3), 30-43. https://doi.org/10.1016/j.envsoft.2011.09.003

Cukier, R., Fortuin, C., Shuler, K.E., Petschek, A. and Schaibly, J.H. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *The Journal of Chemical Physics.* 59(8), 3873-3878. https://doi.org/10.1063/1.1680571

De Kok, J.L. and Wind, H.G. (2003). Design and application of decision support systems for integrated water management; lessons to be learnt. *Physics and Chemistry of the Earth.* 28(14-15), 571-578. https://doi.org/10.1016/S1474-7065(03)00103-7

Deviatkin, I., Kozlova, M. and Yeomans, J.S. (2021). Simulation decomposition for environmental sustainability: Enhanced decision-making in carbon footprint analysis. *Socio-Economic Planning Sciences*. 75, 100837. https://doi.org/10.1016/j.seps.2020.100837

Farr, J.V., Faber, I.J., Ganguly, A., Martin, W.A. and Larson, S.L. (2016). Simulation-based costing for early phase life cycle cost analysis: Example application to an environmental remediation project. *Engineering Economist.* 61(3), 207-222. https://doi.org/10.1080/0013791X.2015.1062582

Fuerst, C., Volk, M. and Makeschin, F. (2010). Squaring the circle? Combining models, indicators, experts and end-users in integrated land-use management support tools. *Environmental Management.* 46(6), 829-833. https://doi.org/10.1007/s00267-010-9574-3

Ghaffarzadegan, N., Lyneis, J. and Richardson, G.P. (2011). How small system dynamics models can help the public policy process. *System Dynamics Review*. 27(1), 22-44. https://doi.org/10.1002/sdr.442

Gunalay, Y., Yeomans, J.S. and Huang, G.H. (2012). Modelling to generate alternative policies in highly uncertain environments: An application to municipal solid waste management planning. *Journal of Environmental Informatics*. 19(2), 58-69. https://doi.org/10.3808/jei.201200209

Han, Z., Porras-Alvarado, J.D., Sun, J.R. and Zhang, Z.M. (2017). Monte Carlo simulation-based assessment of risks associated with public-private partnership investments in toll highway infrastructure. *Journal of the Transportation Research Board*. 2670(1), 59-67. https://doi.org/10.3141/2670-08

Hipel, K. and Ben-Haim, Y. (1999). Decision making in an uncertain world: information-gap modeling in water resources management. *IEEE Transactions on Systems, Man and Cybernetics - part C: Applications and Reviews.* 29(4), 506-517. https://doi.org/10.1109/5326.798765

Hipel, K.W. and Walker, S.G.B. (2011). Conflict analysis in environmental management. *Environmetrics*. 22(3), 279-293. https://doi.org/10.1002/env.1048

Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*. 52(1), 1-17. https://doi.org/10.1016/0951-8320(96)00002-6

Huang, X., Zhou, H. and Zhu, H.B. (2009). A framework for assessing the systemic risk of major financial institutions. *Journal of Banking and Finance*. 33(11), 2036-2049. https://doi.org/10.1016/j.jbankfin.2009. 05.017

Janssen, J.A.E.B., Krol, M.S., Schielen, R.M.J. and Hoekstra A.Y. (2010). The effect of modelling quantified expert knowledge and uncertainty information on model-based decision making. *Environmental Science and Policy*. 13(3), 229-238. https://doi.org/10.1016/j.envsci.2010.03.003

Kim, T.H., Choi, J.S., Park, Y.J. and Son, K. (2013). Life cycle costing: Maintenance and repair costs of hospital facilities using Monte Carlo simulation. *The Korean Institute of Building Construction*. 13(6), 541

-548. https://doi.org/10.5345/JKIBC.2013.13.6.541

Kozlova, M., Collan, M. and Luukka, P. (2017). Simulation decomposition: New approach for better simulation analysis of multi-variable investment projects. *Fuzzy Economic Review*. 21(2), 3. https://doi.org/10.25102/fer.2016.02.01

Kozlova, M. and Yeomans, J.S. (2019). Multi-Variable simulation decomposition in environmental planning: an application to carbon capture and storage. *Journal of Environmental Informatics Letters*. 1(1), 20-26. https://doi.org/10.3808/jeil.201900003

Kozlova, M., Nykänen, T. and Yeomans, J.S. (2021). Technical advances in aviation electrification: Enhancing strategic R&D investment analysis through simulation decomposition. *Sustainability*. 14 (1), 414. https://doi.org/10.3390/su14010414

Kozlova, M. and Yeomans, J.S. (2022a). Extending simulation decomposition analysis into systemic risk planning for Domino-like cascading effects in environmental systems. *Journal of Environmental Informatics Letters*. 7, 64-68. http://dx.doi.org/10.3808/jeil.202200079

Kozlova, M. and Yeomans, J.S. (2022b). Monte Carlo enhancement via simulation decomposition: A "Must-Have" inclusion for many disciplines. *INFORMS Transactions on Education*. 22(3), 147-159. https://doi.org/10.1287/ited.2019.0240

Kozlova, M., Ahola, A., Roy, P. and Yeomans, J.S. (2023a). Simple binning algorithm and simdec visualization for comprehensive sensitivity analysis of complex computational models. arXiv preprint arXiv:2310.13446. https://doi.org/10.48550/arXiv.2310.13446

Kozlova, M., Moss, R.J., Roy, P. and Alam, A. (2023b). *Simulation Decomposition, GitHub*. https://github.com/Simulation-Decomposition (Accessed May 11, 2023).

Kozlova, M., Moss, R.J., Yeomans, J.S. and Caers, J. (2023c). Uncovering heterogeneous effects in computational models for sustainable decision-making. *Environmental Modelling and Software*. 171, 105898. https://doi.org/10. 1016/j.envsoft.2023.105898

Law, A.M. and Kelton, W.D. (2000). *Simulation Modeling and Analysis (3rd edn)*. McGraw-Hill.

Lehar, A. (2005). Measuring systemic risk: A risk management approach. *Journal of Banking and Finance*. 29(10), 2577-2603. https://doi.org/10.1016/j.jbankfin.2004.09.007

Liu, Y.C., Leifsson, L., Pietrenko-Dabrowska, A. and Koziel, S. (2022). Analysis of agricultural and engineering systems using simulation decomposition. *Computational Science - ICCS 2022*, London, 435-444.

Lo Piano, S. and Benini, L. (2022). A critical perspective on uncertainty appraisal and sensitivity analysis in life cycle assessment. *Journal of Industrial Ecology*. 26(3), 763-781. https://doi.org/10.1111/jiec.13237

Loughlin, D.H., Ranjithan, S.R., Brill E.D. and Baugh, J.W. (2001). Genetic algorithm approaches for addressing unmodeled objectives in optimization problems. *Engineering Optimization*. 33(5), 549-569. https://doi.org/10.1080/03052150108940933

Lund, J.R. (2012). Provoking more productive discussion of wicked problems. *Journal of Water Resources Planning and Management*. 138(3), 193-195. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000190

Matthies, M., Giupponi, C. and Ostendorf, B. (2007). Environmental decision support systems: Current issues, methods and tools. *Environmental Modelling and Software*. 22(2), 123-127. https://doi.org/10.1016/j.envsoft.2005.09.005

McKay, M.D., Beckman, R.J. and Conover, W.J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*. 42(1), 55-61. https://doi.org/10.2307/1268522

Mowrer, H.T. (2000). Uncertainty in natural resource decision support systems: Sources, interpretation and importance. *Computers and Electronics in Agriculture*. 27(1-3), 139-154. https://doi.org/10.1016/S0168-1699(00)00113-7

Openshaw, B.W. and Whitehead, P. (1985). A Monte Carlo simulation approach to solving multicriteria optimization problems related to plan making, evaluation, and monitoring in local planning. *Environment and Planning B: Planning and Design*. 12, 321-334. https://doi.org/10.1068/b120321

Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B. and Wagener, T. (2016). Sensitivity analysis of environmental models: a systematic review with practical workflow. *Environmental Modelling and Software*. 79, 214-232. https://doi.org/10.1016/j.envsoft.2016.02.008

Raul, V., Liu, Y.C., Leifsson, L. and Kaleita, A. (2022). Effects of weather on Iowa nitrogen export estimated by simulation-based decomposition. *Sustainability*. 14(3), 1060. https://doi.org/10.3390/su14031060

Ridlehoover, J. (2004). Applying Monte Carlo simulation and risk analysis to the facility location problem. *The Engineering Economist*. 49(3), 237-252. https://doi.org/10.1080/00137910490498942

Ruponen, I., Kozlova, M. and Collan, M. (2021). Ex-Ante study of biofuel policies-analyzing policy-induced flexibility. *Sustainability*. 14(1), 147. https://doi.org/10.3390/su14010147

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D. and Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. John Wiley & Sons, pp 1-292. https://doi.org/10.1002/9780470725184

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Sushan, L. and Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling and Software*. 114, 29-39. https://doi.org/10.1016/j.envsoft.2019.01.012

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W., Kay, J., Lo Piano, S., Mayo, D., Pielke Jr, R., Portaluri, T., Porter, T., Puy, A., Rafols, I., Ravetz, J., Reinert, E., Sarewitz, D., Stark, P., Stirling, A., van der Sluijs, J. and Vineis, P. (2020). Five ways to ensure that models serve society: A manifesto. *Nature*. 582, 482-484. https://doi.org/10.1038/d41586-020-01812-9

Saltelli, A. (2023). *Sensitivity Analysis Made Easy*. http://www.andreasaltelli.eu/file/repository/Talk_SA_June_28_BIS.pdf (Accessed May 11, 2023).

Shapley, L.S. (1953). *A Value for N-Person Games*. Contributions to the Theory of Games (AM-28), Volume II. Princeton University Press, pp 307-317. https://doi.org/10.1515/9781400881970-018

Sobol, I.M. (1993). Sensitivity analysis for non-linear mathematical models, originally sensitivity estimates for non-linear mathematical models. *Mathematical Modeling and Computational Experiment*. 1, 407-414.

Sobol, I.M. and Saltelli, A. (1995). About the use of rank transformation in sensitivity analysis of model output. *Reliability Engineering and System Safety*. 50(3), 225-239. https://doi.org/10.1016/0951-8320(95)00099-2

Sobol, I.M. and Kucherenko, S. (2009). Derivative based global sensitivity measures and their link with global sensitivity indices. *Mathematics and Computers in Simulation*. 79(10), 3009-3017. https://doi.org/10.1016/j.matcom.2009.01.023

Teply, P. and Klinger, T. (2015). Monte Carlo simulations: A case study of systemic risk modelling. *The Second International Conference on Electrical, Electronics, Computer Engineering and their Applications* (EECEA 2015), Barcelona, 43.

Vithayasrichareon, P. and Macgill, I.F. (2012). A Monte Carlo based decision-support tool for assessing generation portfolios in future carbon constrained electricity industries. *Energy Policy.* 41, 374-392. https://doi.org/10.1016/j.enpol.2011.10.060

Walker, S.G.B., Hipel K.W. and Inohara, T. (2012). Attitudes and preferences: Approaches to representing decision maker desires. *Applied Mathematics and Computation*. 218(12), 6637-6647. https://doi.org/10.1016/j.amc.2011.11.102

Yeomans, J.S. (2008). Applications of simulation-optimization methods in environmental policy planning under uncertainty. *Journal of*

Environmental Informatics. 12(2), 174-186. https://doi.org/10.3808/jei.200800135

Yeomans, J.S. and Kozlova, M. (2023). Extending system dynamics modeling using simulation decomposition to improve the urban planning process. *Frontiers in Sustainable Cities*. 5, 1129316. https://doi.org/10.3389/frsc.2023.1129316

Zechman, E.M. and Ranjithan, S.R. (2004). An evolutionary algorithm to generate alternatives (EAGA) for engineering optimization problems. *Engineering Optimization*. 36(5), 539-553. https://doi.org/10.1080/03052150410001704863