



**HAL**  
open science

# Computationally efficient sensitivity analysis for building ecodeign with many-level categorical input factors

Marie-Lise Pannier, Patrick Schalbart, Bruno Peuportier

## ► To cite this version:

Marie-Lise Pannier, Patrick Schalbart, Bruno Peuportier. Computationally efficient sensitivity analysis for building ecodeign with many-level categorical input factors. *International Journal of Metrology and Quality Engineering*, 2023, 14, pp.15. 10.1051/ijmqe/2023016 . hal-04324021

**HAL Id: hal-04324021**

**<https://hal.science/hal-04324021v1>**

Submitted on 5 Dec 2023

**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.



Distributed under a Creative Commons Attribution 4.0 International License

# Computationally efficient sensitivity analysis for building ecodesign with many-level categorical input factors

Marie-Lise Pannier<sup>1,2</sup>, Patrick Schalbart<sup>1</sup>, Bruno Peuportier<sup>1</sup>

<sup>1</sup>Mines Paris, PSL University, Centre for energy efficiency of systems (CES), 75006 Paris, France

<sup>2</sup>Univ Angers, LARIS, SFR MATHSTIC, F-49000 Angers, France

**ABSTRACT:** Dynamic building energy simulation and life cycle assessment (LCA) are useful ecodesign tools to mitigate the energy and environmental impacts of buildings. Various uncertain factors can affect the building energy and environmental modelling, including continuous and categorical factors (i.e. discrete factors without logical ranking). Sensitivity analysis (SA) is applied to identify the most influential factors on which additional research efforts are needed to increase the robustness of results. The Sobol method (Sobol) is the reference SA method, but it requires a significant amount of computation. Less time-consuming methods, such as an adaptation of the Morris screening (Morris), have shown a good ability to quantify the influence of factors, but their performance has not been investigated for categorical factors having many (more than two) levels. Two strategies (2LA-Morris and MA-Morris) based on the adaptation of Morris are proposed to handle many-level factors. Their performance is compared to that of Sobol based on four criteria: computation time, factor's relative influence, factor's ranking, and ability to detect interactions. For the LCA of a house including 24 uncertain factors, MA-Morris was able to quantify the influence of factors in the same way as Sobol, while reducing the computation time by a factor of 12.

**Keywords:** Dynamic building energy simulation, Life cycle assessment, Sensitivity analysis, Morris screening, Sobol indices.

**Corresponding author:** Marie-Lise Pannier, e-mail: marie-lise.pannier@univ-angers.fr

## 1. INTRODUCTION

The building sector is responsible for high energy consumptions and environmental impacts [1–3]. Ecodesign tools are increasingly used to mitigate these impacts. Among them dynamic building energy simulation (DBES) precisely assesses the temporal evolution of heating and cooling loads, and comfort. In addition, the life cycle assessment (LCA) methodology [4,5] applied to buildings, allows to evaluate the environmental impacts of a construction over its long life cycle, as well as compare the performance of building alternatives. Associating DBES and LCA is essential for a precise assessment of the impacts of the building sector [6].

Because of the variability of the use of buildings and their long lifetime, many input factors of DBES and LCA are uncertain. These uncertainties could question the reliability of decisions based on building LCA. Uncertainties have been discussed in building LCA since the mid-1990s. However, they are still rarely addressed and most studies present deterministic results [7,8]. Uncertain factors in building energy and environmental simulations have been classified according to their origin (i.e. lack of knowledge of the true value of a quantity, stochastic variations, mistakes...), or to the life cycle stage at which they occur (i.e. construction,

use, renovation, end-of-life) [8–10]. Beside these general classifications, factors can also be sorted according to their type, which ranges from continuous factors (e.g. uncertainties on materials and building lifetimes) to categorical inputs which are discrete factors with no logical ranking (e.g. scenarios describing climate or occupancy variability).

Uncertainty analysis and sensitivity analysis (SA) methods are used to better understand the effect of uncertain factors. They allow uncertainties to be quantified and they can help improving both the reliability of results and the quality of the subsequent decision-making. For instance, SA methods are useful to identify the most uncertain factors that could change the decision and that should be further investigated.

Many SA methods are available to identify the most uncertain factors [11]. Variance-based global SA methods have proven to be the most relevant [12]; among them, the Sobol method [13] has often been used in DBES and LCA [14]. It has the advantage of handling all types of uncertain factors, such as continuous, discrete and categorical factors. Despite being accurate, this method has the disadvantage of requiring a significant computation effort to perform all DBES and LCA, as well as to calculate Sobol indices. Other methods, such as local SA or screening, require less effort but not cover the entire variation space of uncertain factors,

and therefore quantify the influence of uncertain factors less precisely.

Previous studies compared the performances of different SA methods in DBES and LCA [14–19] and found that linear regression-based SA, Morris and SA applied to metamodels yielded a good compromise between accuracy and computation time. However, they did not focus on the ability of low-computationally intensive methods to deal with all types of factors, such as categorical inputs with more than two levels. In this article, two less computationally intensive methods based on the Morris method [20] are proposed to deal with many-level inputs. They are compared on their ability to quickly rank influential factors in the same way as the Sobol method.

## 2. METHODOLOGY

The two methods proposed to deal with many-level inputs are based on an adaptation of the Morris screening method. This adaptation, noted A-Morris hereafter, has been shown to precisely quantify the influence of uncertain factors [14]. The two adaptations and the comparison metrics are presented in the next paragraphs.

### 2.1. SA methods

**Sobol method.** The variance-based Sobol method [13] is used as a reference to investigate the performance of the proposed methods. As a global SA method, the continuous uncertain factors are assessed over their entire variation range, following their probability distribution. For discrete and categorical inputs, a possible value is randomly sampled. Total Sobol indices  $TS_i$  [21,22] are computed for the comparison as they include factors' non-linear and interactions effects. They are calculated, as in equation (1), for each uncertain factor  $X_i$ , by studying the variation of the model output  $Y = f(X)$  when the value of the  $i^{th}$  factor changes while all others are kept constant.

$$TS_i = \frac{E[Var(Y|X_{-i})]}{Var(Y)} \quad (1)$$

$Var(Y)$  and  $E[Y]$  are the variance and the expectation.  $N(K + 2)$  model evaluations are required,  $K$  being the number of uncertain factors and  $N$  the sample size.

The total Sobol Jansen indices [21] are used as they are the best estimate of Sobol indices according to Saltelli [22]. For further investigations, the first

order indices were assessed following Saltelli [22] for consistency reasons [23,24].

**Morris method.** In the Morris method [20], the uncertain factors variation ranges are discretised into levels. The exploration space is then a grid of nodes on which different trajectories are performed. A trajectory corresponds to a list of  $K + 1$  simulations. Between two consecutive simulations of the trajectory, the value of only one factor is modified. A One-factor-At-Time design of experiment (DoE) is thus followed. Starting from one randomly chosen point  $X$ , values of each coordinates  $X_i$  are changed one after the other, according to the discretised levels available, considering a jump of length  $\delta$  on the grid of nodes.  $r$  trajectories are repeated, so that  $r \cdot (K + 1)$  simulations are required. Elementary effects (EE) can then be computed for each factor  $i$  and each trajectory  $j$ , as in (2).

$$EE_i^j = \frac{f(X_1^j, \dots, X_i^j + \delta x_i^j, \dots, X_K^j) - f(X_1^j, \dots, X_i^j, \dots, X_K^j)}{\delta} \quad (2)$$

In the original Morris method, the influence of a factor is quantified by the mean of the absolute value of EE over the trajectories  $\mu_i^*$ . In addition, the standard deviation of EE  $\sigma_i$  informs on the non-linearity and on the presence of interactions between the  $i^{th}$  factor and the other factors.

**A-Morris method.** A-Morris is an improvement of the Morris methods proposed in [14]. It allows the influence of uncertain factors to be quantified more precisely than with the original Morris method but with far fewer simulations than for the Sobol method. In A-Morris, the DoE, i.e. the list of simulations to perform, remains the same as in the original method. However, for each uncertain factor  $i$  and repetition  $j$ , an elementary variance (EV) is calculated, instead of an EE, as in (3).

$$EV_i^j = [f(X_1^j, \dots, X_i^j + \delta x_i^j, \dots, X_K^j) - f(X_1^j, \dots, X_i^j, \dots, X_K^j)]^2 \quad (3)$$

Then, a sensitivity index  $S_i$ , similar to the expectancy of the variance computed for the Sobol method, is obtained by averaging the  $EV_j^i$  over  $r$  repetitions, as in (4).

$$S_i = \frac{1}{r} \sum_{j=1}^r EV_j^i \quad (4)$$

The performances of A-Morris were investigated for continuous factors and for categorical factors with only two levels by Pannier et al. [14]. In each trajectory, the value of the categorical input was

changed from one level to the other one. As categorical factors may have more than two levels, two strategies are proposed and assessed in this article to deal with many-level inputs with A-Morris.

**2LA-Morris method.** In the first strategy, only two levels that give extreme results for all outputs are selected and used for all repetitions. At each repetition, a jump is performed from the first to the second level. This strategy is called 2LA-Morris (for two Level Adapted Morris).

**MA-Morris method.** In this second strategy, at each repetition, two different levels are randomly sampled and the jump is performed from one level to the other. The levels sampled may be different at each repetition. In this strategy,  $2 * r$  levels are explored. This strategy is called MA-Morris (for Many-level Adapted Morris).

## 2.2. Comparison metrics

The results of the two strategies (2LA-Morris and MA-Morris) are compared with the results of the Sobol method based on four criteria.

Firstly, the **computation time** is studied. It is defined as the time required to run all the simulations and to compute the sensitivity indices.

Secondly, the **relative influence** (RI) is computed as in [14] to get the precision of each method. It is defined as the share of the sensitivity index of one factor relatively to the sum of the sensitivity indices of all factors, as in (5).

$$RI_i = \frac{S_i}{\sum_i S_i} \quad (5)$$

RI is useful to identify a relevant set of uncertain factors. After sorting the  $RI_i$  by increasing order, the number of uncertain factors that cover some share of variance can be identified for all methods.

Thirdly, the **factors' ranking** is compared using the method proposed by Akkari [19]. Factors are plotted according to their rank. In addition, the Kendall rank correlation  $\tau$  is calculated as in (6) to quantify the ranking similarities:

$$\tau = \frac{2(n_c - n_d)}{K(K - 1)} \quad (6)$$

with  $n_c$  and  $n_d$ , the number of concordant and discordant pairs. Concordant pairs  $(a_1, b_1)$  and  $(a_2, b_2)$  are so that  $sgn(a_2 - a_1) = sgn(b_2 - b_1)$ , while discordant pairs have  $sgn(a_2 - a_1) = -sgn(b_2 - b_1)$ ,  $sgn$  being the signum function defined as below:

$$sgn(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} \quad (7)$$

A  $\tau$  value close to -1 (resp. +1) indicates a strong positive (resp. negative) correlation, while a  $\tau$  value close to 0 indicates an absence of correlation.

Finally, the ability of 2LA-Morris and MA-Morris to **detect interaction** effects is assessed.  $\sqrt{\mu_j^*}$  is therefore plotted as a function of  $\sqrt{\sigma_j}$ . Non-linear or interaction effects are identified when  $\sqrt{\sigma_j}$  is of the same order of magnitude or greater than  $\sqrt{\mu_j^*}$ . For the reference Sobol method, interactions effects can be detected by comparing first  $S_i$  and total  $TS_i$  order sensitivity indices. If the total order indices are significantly higher than the first order ones, the factor has high interaction effects. Non-linear effects are quantified both by  $S_i$  and  $TS_i$  as Sobol is a global SA method.

## 3. CASE STUDY

After presenting the DBES and LCA calculation methods, the building used as a case study and the 24 uncertain factors are described hereafter.

### 3.1. DBES and LCA framework

The COMFIE [25] model of the software Pleiades [26] is used to run the DBES. In this model, whose reliability has been validated [27,28], the building is divided into thermal zones of homogeneous operative temperature before being meshed. A thermal balance is applied to each mesh to obtain the zone temperatures, as well as the heating and cooling loads. This provides a precise assessment of the energy consumptions during the building use phase [6], which is the life cycle phase having the highest share of impacts [2].

Since climate and user behaviour have a strong influence on the energy performance [29], two additional models linked to COMFIE are used to generate a set of realistic meteorological years [30] and occupancy scenarios [31].

The model EQUER [32,33] of Pleiades [26] is used to run the LCA simulations, based on the DBES results. All building life cycle phase are considered and the impacts are calculated for the 14 indicators listed in Table 1. The results reliability of EQUER has been verified by comparison with other building LCA models [34–36].

Table 1. Environmental indicators and fluxes.

Indicator / flux	Legend	Unit	Method
Climate change GWP100	Clim. Change	kg CO <sub>2</sub> eq	IPCC 2013 [37]
Cumulative Energy Demand	CED	MJ	CED [38]
Abiotic Depletion of Minerals	Resources	kg Sb	CML [39]
Mass of Waste	Waste	kg	Flux from ecoinvent [40]
Water use	Water	litre	Flux from ecoinvent [40]
Fine particulate matter formation	Particulate M.	kg PM10 <sub>eq</sub>	ReCiPe [41]
Photochemical ozone formation	P. Ozone	kg NMVOC <sub>eq</sub>	ReCiPe [41]
Volume occupied by radioactive waste	Rad. Waste	m <sup>3</sup>	Flux from ecoinvent [40]
Ionising radiation	Ioni. Rad.	kg 235U <sub>eq</sub>	ReCiPe [41]
Human toxicity	Hum. Tox.	CTUh	USEtox [42]
Eutrophication	Eutrop.	kg PO43- eq	CML [39]
Acidification	Acid.	kg SO <sub>2</sub> eq	CML [39]
Ecotoxicity	Ecotox.	CTUe	USEtox
Land use	Land use	points	ReCiPe [41]

### 3.2. Building

The studied building is a single-family house of 90 m<sup>2</sup>. It has a shuttered concrete structure, an external insulation, a double-flow ventilation system and is electrically heated. The performance of this building corresponds to the passive house standards. A complete description of the DBES and LCA assumptions is provided in [10,43].

### 3.3. List of uncertain factors

In a first step, 153 uncertain factors were considered. They are related to the building site, materials, components, construction processes, systems or use. In addition, some uncertainties on the background environmental data and the life cycle impact assessment methods were also included. The complete list can be found in [10,43].

The computation time of SA methods depends on the number of uncertain factors and it can be very long for the Sobol method. Therefore, a screening was performed to select the factors whose cumulative relative influence exceeded 99 % for both 2LA-Morris and MA-Morris. The three methods are applied on this set of 24 uncertain factors, given in Table 2. Uniform distributions were chosen for all continuous factors in this study. Reference values, variation units and variation ranges (e.g. lower (LB) and upper (UP) bounds) are given in Table 2 along with the data source for the uncertainty characterisation.

20 uncertain factors are continuous. The remaining four factors are categorical. One of them have two possible levels: the choice of the blowing agent for the polystyrene extrusion process. The three other categorical factors may have many-level. For the variability of climate, a set of 2,000 realistic climate files was generated with the model of Ligier et al. [30]. Similarly, a set 2,000 realistic occupancy scenarios was generated with the model of Vorger et al. [31]. Finally, a set of 2,000 background environmental data was generated using the uncertainty distributions provided by the ecoinvent environmental database [46].

For the Sobol method, a Latin hypercube sampling of size  $N = 1,000$  was applied, allowing the 2,000 levels available to be used for the categorical inputs with many-level.

Regarding 2LA-Morris, the two extreme scenarios for the three many-level categorical factors are selected as follow:

- Variability of climate: the available scenarios are ranked by increasing mean outside temperature during the year and during the heating period. Two scenarios close to the 2.5% and 97.5% bounds of the two distributions are selected.
- Variability of occupancy and environmental background data: DBES followed by LCA simulations are performed for all levels. The results are ranked by increasing order. Two scenarios for occupancy and two scenarios for background data, close to the 2.5% and 97.5% bounds of the 14 distributions of the environmental indicators, are selected.

Regarding the MA-Morris, the two different scenarios are used for each repetition. For a previously performed uncertainty analysis, it has been shown that after using 200 scenarios, convergence was reached for the variability of climate and of occupancy, and almost reached for the environmental background data [10]. 200 scenarios are then used for this method, corresponding to 100 repetitions. For consistency reason, 100 repetitions are also performed for 2LA-Morris.

Table 2. List of uncertain factors.

Factor	Source	Ref. value	LB	UB	Unit
Characterisation factor HFC134a	[44]	1301.3	1190.5	1412.0	kg CO <sub>2</sub> -eq
Construction material waste	[45]	5	0	10	%
Environmental background data	[46]	<b>Categorical:</b> Sampling in the available levels			-
Fan energy consumption	[47]	0.475	0.105	0.845	Wh/m <sup>3</sup>
Lifetime: building	[2,3]	80	40	200	yr.
Lifetime: covering	[45]	40	15	100	yr.
Lifetime: doors	[45]	37	10	80	yr.
Lifetime: hot water tank	[45]	15	8	20	yr.
Lifetime: insulation	[45]	38	20	60	yr.
Lifetime: painting	[45]	16	5	50	yr.
Lifetime: roof tile	[45]	50	15	100	yr.
Lifetime: tiled floor	[45]	50	15	100	yr.
Lifetime: windows	[45]	38	15	60	yr.
Polystyrene: conductivity	[48,49]	0.030	0.027	0.033	W/m/K
Polystyrene extrusion process	-	<b>Categorical:</b> CO <sub>2</sub> or HFC134a			-
Steel rate in concrete	[50]	3	0.2	4.2	-
Thermal bridge: living room	[51]	RT2012	-0.52 * V <sub>ref</sub>	0 * V <sub>ref</sub>	W/K
Transp. distance of new materials to building site	[52,53]	75	0	200	km
Transp. distance of waste materials from building to end-of-life site	[52,53]	25	0	100	km
Variability of Climate	[30]	<b>Categorical:</b> Sampling in the available levels			-
Variability of Occupancy	[31]	<b>Categorical:</b> Sampling in the available levels			-
Ventilation double flow heat exchanger efficiency	[49]	90.0	85.5	92	%
Ventilation rate in the crawl space	[28]	1	0.5	1.5	vol/h
Water network efficiency	[54]	74	37	100	%

## 4. RESULTS

### 4.1. Computation time

Given the number of uncertain factors and the sampling size for Sobol, 26,000 model simulations were parallelised on a six-core desktop computer. This corresponds to 36 h of computation to pre-process data, run simulations and post-process results. For 2LA-Morris and MA-Morris, the 2 500 model simulations were parallelised on the same computer, leading to 3 h of computation.

### 4.2. Relative influence (RI)

The comparison results in terms of RI are given in Figure 1 and in selected to reach RI 95, highlighting a ranking difference for this factor between the methods.

For a given share of RI (80 to 99 %), factors identified by Sobol are always identified with MA-Morris. However, this method may select additional factors that are not selected by Sobol. As an example, in RI 80, MA-Morris selected two additional factors, the insulation lifetime and the covering lifetime, which are later selected in RI 85 and in RI 90 by Sobol. On the contrary, there is a risk of missing influential factors with the 2LA-Morris: some factors selected with Sobol may not be selected (e.g. the water network efficiency for RI 80, or the insulation lifetime in RI 85). MA-Morris is thus preferred in order to avoid not selecting highly influential factors.

Table 3. In Figure 1, each group of three bars is related to one indicator (the names of indicators are shown above the groups of bars), and each bar stands for one SA method (the names of the methods are indicated below the bars). In each bar, each colour corresponds to the relative influence  $RI_j$  of an uncertain factor. The larger a colour segment in a bar, the more influential the corresponding factor is. The most influential factors identified across all indicators and methods are: the variability of occupancy (green), the environmental background data (red), and the building lifetime (blue). Visually, MA-Morris performs better than 2LA-Morris for all indicators. The relative influences of factors given with MA-Morris are close to those of Sobol for most indicators (the different colour are equally represented), except for the eutrophication and the human toxicity. However, visually, the ranking of the factors appears identical.

The conclusion of the visual interpretation are confirmed in selected to reach RI 95, highlighting a ranking difference for this factor between the methods.

For a given share of RI (80 to 99 %), factors identified by Sobol are always identified with MA-Morris. However, this method may select additional factors that are not selected by Sobol. As an example, in RI 80, MA-Morris selected two additional factors, the insulation lifetime and the covering lifetime, which are later selected in RI 85 and in RI 90 by Sobol. On the contrary, there is a risk of missing influential factors with the 2LA-Morris: some factors selected with Sobol may not be selected (e.g. the water network efficiency for RI 80, or the insulation lifetime in RI 85). MA-Morris is thus preferred in order to avoid not selecting highly influential factors.

Table 3. The upper part of this table summarises, which factors (in rows) are selected for a certain share of RI, for each SA method (in column). In addition, the lower part of the table explains the number of factors to select in order to reach 80 to 99 % of the RI for all 14 environmental indicators. The corresponding sets are called RI 80 to RI 99. For instance, for the Sobol method, RI 80 consists of the six factors on the dark green background (environmental data, building and tiled floor lifetime, polystyrene extrusion process, variability of occupancy, and water network efficiency). RI 85 is reached by adding the insulation lifetime. In addition, the water network efficiency should be selected to reach RI 80 for Sobol and MA-Morris. However, this factor is selected for RI 80 with 2LA-Morris. It appears to be

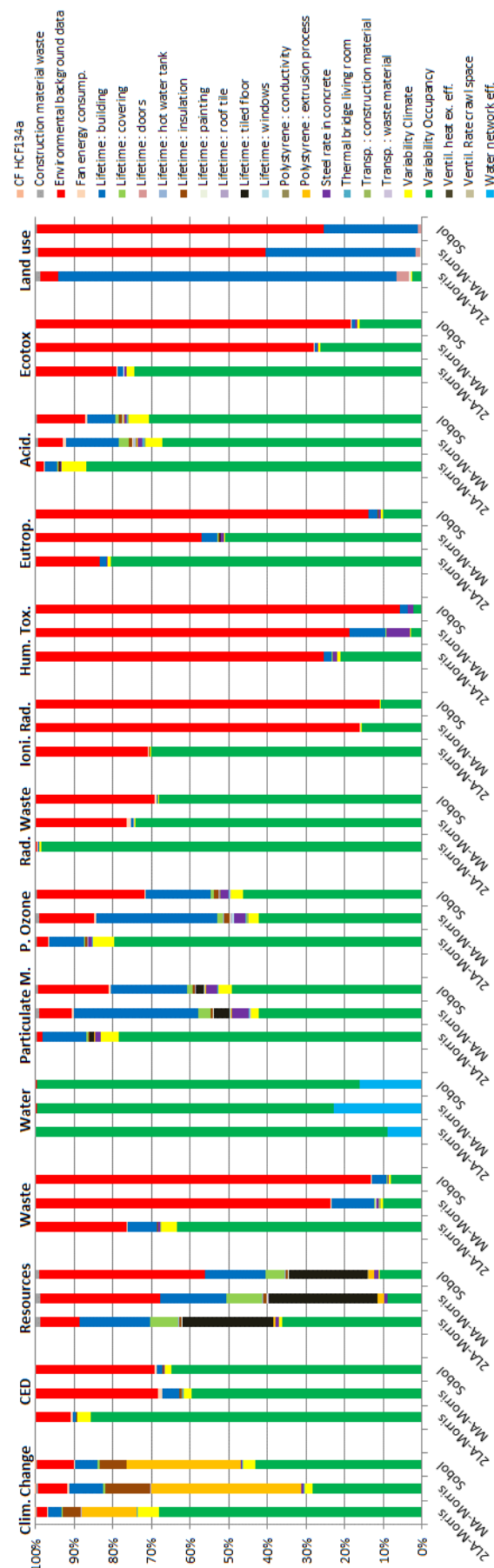


Figure 1: RI of the three SA methods.

selected to reach RI 95, highlighting a ranking difference for this factor between the methods.

For a given share of RI (80 to 99 %), factors identified by Sobol are always identified with MA-Morris. However, this method may select additional factors that are not selected by Sobol. As an example, in RI 80, MA-Morris selected two additional factors, the insulation lifetime and the covering lifetime, which are later selected in RI 85 and in RI 90 by Sobol. On the contrary, there is a risk of missing influential factors with the 2LA-Morris: some factors selected with Sobol may not be selected (e.g. the water network efficiency for RI 80, or the insulation lifetime in RI 85). MA-Morris is thus preferred in order to avoid not selecting highly influential factors.

Table 3. RI 80 to RI 99 factors for the two A-Morris strategies and for Sobol for all indicators.

Factor	Sobol	2LA-Morris	MA-Morris
Const. material waste	97.5	97.5	97.5
Envi. background data	80	80	80
Fan energy consump.	99	99	97.5
Lifetime: building	80	80	80
Lifetime: covering	90	90	80
Lifetime: doors	99	95	97.5
Lifetime: insulation	85	90	80
Lifetime: painting	99		99
Lifetime: roof tile			99
Lifetime: tiled floor	80	80	80
Lifetime: windows			99
Polystyrene: extrusion process	80	80	80
Steel rate in concrete	95	95	85
Thermal bridge living room	99		97.5
Transp. of const. material	99	99	99
Variability Occupancy	80	80	80
Variability Climate	90	85	90
Water network eff.	80	95	80
<b>Nb. of factors for 80 %</b>	6	5	8
<b>Nb. of factors for 85 %</b>	7	6	9
<b>Nb. of factors for 90 %</b>	9	8	10
<b>Nb. of factors for 95 %</b>	10	11	10
<b>Nb. of factors for 97.5 %</b>	11	12	14
<b>Nb. of factors for 99 %</b>	16	14	18

### 4.3. Factors' ranking

The comparison of the methods in terms of ranking of factors is given in Figure 2 (for the climate change indicator) and Table 4 (for all indicators).

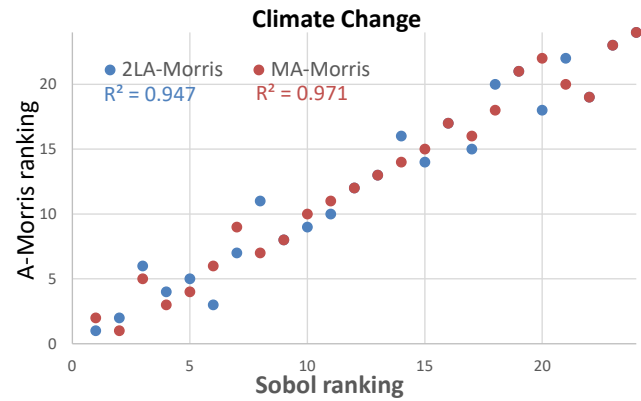


Figure 2: Correlations between Sobol and A-Morris rankings.

In Figure 2, factors are ranked by increasing order of influence for the Sobol method (x-axis). The ranking for the A-Morris methods can be read on the y-axis. The blue series represents the correlations between Sobol and 2LA-Morris, while the red series represents the correlations between Sobol and MA-Morris. As the first blue point has coordinates (1,1), the most influential factor is the same for Sobol and 2LA-Morris. The first red point has coordinates (1,2), meaning that the most influential factor for Sobol is the second most influential factor for MA-Morris. The factors ranking is similar in this case study as an almost linear relation is found between the results of Sobol and A-Morris for the climate change indicator. However, visually the ranking of MA-Morris is a little closer to the one of Sobol. This is confirmed by the results of the linear regressions performed: the  $R^2$  is higher for MA-Morris (0.971) than for 2LA-Morris (0.947).

The ranking similarities are quantified using the Kendall rank correlation  $\tau$  in Table 4.  $\tau$  is more adapted than the  $R^2$  to quantify the extent to which the order of factors is respected between two datasets. A value of +1 for  $\tau$  indicates that the ranking is the same for all factors, while a value of -1 indicates a perfect ranking inversion. Correlations are weak when the  $\tau$  value is close to zero. In this study,  $\tau$  being generally close to +1, there are overall strong positive correlations between Sobol and A-Morris. For almost all indicators,  $\tau$  is higher for the correlations between MA-Morris and Sobol, indicating that the MA-Morris ranking is closer to the Sobol ranking, than the 2LA-Morris ranking. 2LA-Morris gives a better ranking only for the ionising radiation and human toxicity (red font). Both A-Morris methods perform



equally well, in terms of uncertain factors ranking, for the cumulative energy demand (yellow font).

Table 4. Kendall rank correlation  $\tau$  for all indicators.

	$\tau$ Sobol vs. 2LA-Morris	$\tau$ Sobol vs. MA-Morris
Clim. Change	0.884	0.920
CED	0.957	0.957
Resources	0.870	0.920
Waste	0.928	0.935
Water	0.710	0.754
Particulate M.	0.862	0.949
P. Ozone	0.891	0.913
Rad. Waste	0.841	0.920
Ioni. Rad.	0.877	0.862
Hum. Tox.	0.855	0.826
Eutrop.	0.833	0.891
Acid.	0.688	0.899
Ecotox.	0.819	0.877
Land use	0.862	0.935

#### 4.4. Interaction effects

In Figure 3 and Figure 4, the interaction effects found in A-Morris and Sobol are shown for the six most influential factors of two environmental indicators. The upper parts of Figure 3 and Figure 4 give the first order Sobol indices (in red) and the total order indices (in blue).  $TS_i$  values significantly higher than  $S_i$  indicate interactions effects between factors. The 2LA-Morris (respectively the MA-Morris results) are shown in the middle part (respectively in the lower part) of Figure 3 and Figure 4. In these four graphs, the x-axis corresponds to the square of the mean of the absolute value of the EE over the trajectories  $\sqrt{\mu_i^*}$ ; a large value for  $\sqrt{\mu_i^*}$  representing a significant influence of the uncertain factor due to linear effect. The y-axis value corresponds to the square of the standard deviation of the EE  $\sqrt{\sigma_i}$ ; a factor with a large value for  $\sqrt{\sigma_i}$  is influential because it has strong non-linear or interaction effects. Factors with a large  $\sqrt{\mu_i^*}$ , a large  $\sqrt{\sigma_i}$ , or both are therefore particularly influential.

In these results, the convergence of the first order Sobol indices is not completely reached due to the low sampling size chosen: the uncertainty bars obtained by bootstrapping remain large for the first order indices. However, for climate change,  $TS_i$  is

only slightly larger than  $S_i$  for the variability of climate, the background environmental data, the lifetime of insulation, and the polystyrene extrusion process. This indicates weak interaction effects between these factors. For the three factors mentioned previously, interaction or non-linear effects are more important for MA-Morris, than for 2LA-Morris, as  $\sqrt{\mu_i^*}$  is closer to  $\sqrt{\sigma_i}$ . In addition, the lifetime of insulation has a  $\sqrt{\sigma_i}$  larger than its  $\sqrt{\mu_i^*}$ , suggesting the presence of non-linear or interaction effect in the two A-Morris strategies. In fact, interaction effects are expected in this case, as the number of replacements of the insulation material during the building lifetime depends on both the insulation and the building lifetimes.

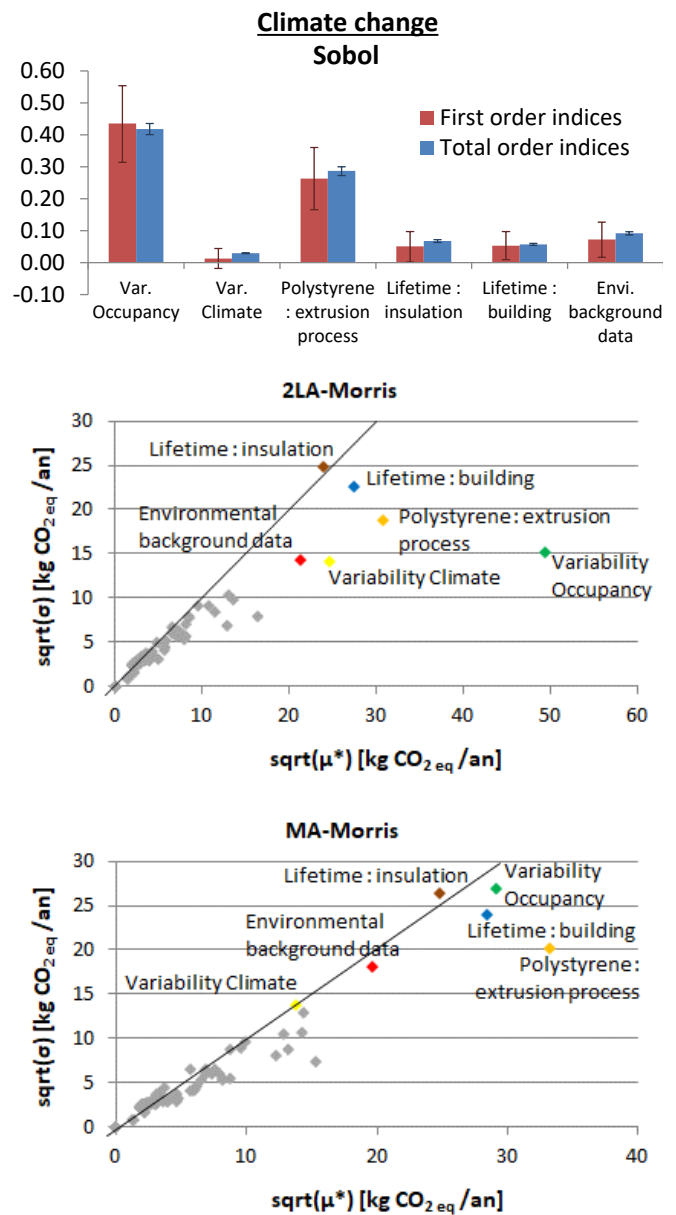


Figure 3: Detection of non-linear and interaction effects for the climate change indicator.

The same phenomenon is observed with the CED indicator in Figure 4. Small interactions effects are pointed out by the Sobol method for the variability of occupancy and the environmental background data. For these factors,  $\sqrt{\sigma_j}$  is of the same order of magnitude as  $\sqrt{\mu_j^*}$  for MA-Morris, but not for 2LA-Morris. In addition, interaction effects are found on the insulation and building lifetime for both 2LA-Morris and MA-Morris.

a linear effect. In comparison, MA-Morris includes as well non-linearities.

## 5. DISCUSSION

Overall, in this building LCA case study, the two strategies based on A-Morris required twelve times less computation effort. This is a lower bound since Sobol usually requires a larger sample size. This makes them more suitable for application in current practice by building designers.

In A-Morris, the computation time depends on the number of uncertain factors included in the SA, as well as on the number of repetitions. 100 repetitions were performed in this study in order to run simulations in MA-Morris, with 200 different levels for the many-level factors (variability of climate, variability of occupancy and environmental background data). It was decided to integrate 200 levels because a previous uncertainty analysis had shown that convergence was reached on the environmental indicators after this number of levels. However, the number of levels to include (and thus the number of repetitions) needs to be further investigated, in order to get good performances with as few simulations as possible. There is a potential for a reduction of the number of simulations with A-Morris.

Overall, MA-Morris strategy is more adapted than 2LA-Morris to handle many-level categorical uncertain factors. In addition, with MA-Morris, the preliminary step of selection of scenarios giving extreme results is not necessary since all available scenarios are investigated. The use of MA-Morris is thus easier and recommended.

The results given by MA-Morris are closer to those given by the Sobol method in terms of relative influence quantification, and ranking of uncertain factors. The discrepancy between MA-Morris and Sobol between two of the environmental indicators must be further investigated. The PDF of these indicators display a long tail due to the environmental background data (see annex N of [10]). There is higher risk that Sobol samples in this tail and therefore overestimates the influence of background data. This risk could be minimised by somehow truncating the tails of the distributions.

Three of the four categorical factors (variability of climate, environmental background data and polystyrene extrusion process) are among the most influential factors. They are selected from RI 80 for all SA methods. Categorical factors tend to have

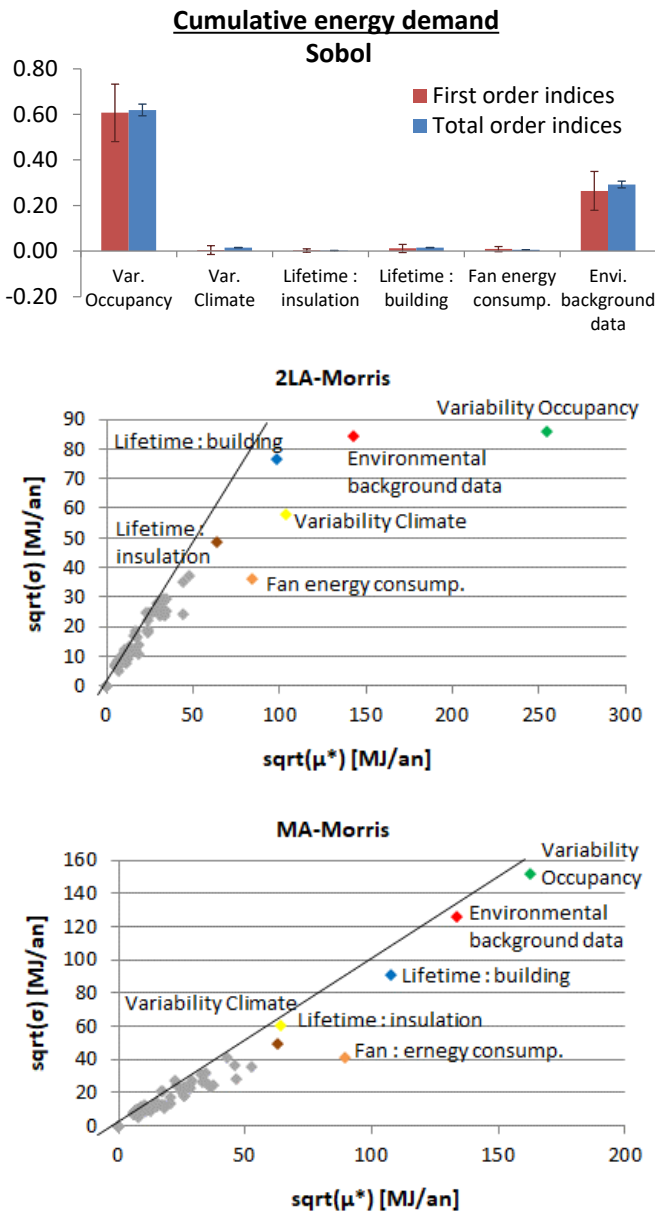


Figure 4: Detection of non-linear and interaction effects for the cumulative energy demand indicator.

The difference between the two A-Morris strategies is linked to the number of levels. In 2LA-Morris, because only two levels are included, the effect of the factor is more likely to be interpreted as

more influence by construction because changing the level for each repetition creates variability in the results and induces non-linearities. In addition, there are some known interaction effects between the three most influential categorical factors and other factors, especially regarding the variability of occupancy (linked to many factors affecting the building use stage) and the environmental data (linked to all life cycle uncertain factors). Even though these interaction effects remain small based on the comparison of first order and total indices for Sobol, they are better identified with MA-Morris than with 2LA-Morris.

The performance of MA-Morris should be verified for other case studies, for instance for cases with more non-linear or interaction effects, or for cases for which continuous factors have a larger influence than categorical ones. These conditions could be met assuming larger variation spaces for continuous factors: a larger variation space for some factors would indeed potentially increase the influence of these factors, which could have more interaction effects, and thus non-linear properties could emerge. An increase in these values corresponds to the case of the early design phase of a building project when the values of the factors are not yet well known. In addition, discrete factors were not taken into account in this case study. The behaviour of these factors is expected to be similar to the one of continuous factors in A-Morris, as the possible values have a logical order. However, this needs to be verified for case studies including continuous, discrete and categorical factors.

## 6. CONCLUSION

SA methods have been applied in DBES and building LCA to quantify the influence of uncertain factors. The reliable and recognised Sobol SA method is applicable for all kinds of uncertain factors, but requires large computational efforts. Its long computation time makes it hardly manageable for practical application in building projects. Previous research identified other SA methods as good compromises between accuracy and computation time. However, the ability of the low-computationally intensive methods to deal with categorical inputs with more than two levels has not yet been investigated. That was the aim of this work, in which two strategies based on A-Morris are proposed to deal with many-level inputs.

The results given by the two strategies were compared to those of the Sobol method, in order to investigate drivers among a set of 24 uncertain factors for the LCA of a single family passive house. The four categorical factors (variability of occupancy, variability of climate, uncertainties on the environmental background data, and choice of the blowing agent for the polystyrene extrusion) were found to be among the most influential factors. Both strategies have shown a good ability to quantify the influence of uncertain factors in twelve times fewer simulations than the Sobol method. The MA-Morris strategy, for which two different levels are randomly sampled for the categorical factors at each repetition, has shown results closer to those of Sobol for the investigated performance criteria. This strategy quantifies the relative influence in the same way as Sobol and does not miss any influential factors for a given share of relative influence. The factor ranking given by MA-Morris is closer to the one given by Sobol for 11 of the 14 environmental indicators.

Further work is needed to investigate the performance of the MA-Morris method for other case studies with strong non-linear or interaction effects, or for which other types of factors (such as continuous or discrete factors) are more influential than the categorical factors. The optimal number of repetitions to perform needs to be examined as well.

## ACKNOWLEDGEMENT

This work was supported by the research Chair ParisTech VINCI “Eco-design of buildings and infrastructure”. We would like to thank the organisers of the MQDS2023 conference for offering to publish our article in the IJMQE journal.

## REFERENCES

- [1] IEA, Global Status Report for Buildings and Construction 2019 – Towards a zero-emissions, efficient and resilient buildings and construction sector, 2019. <https://www.iea.org/reports/global-status-report-for-buildings-and-construction-2019>.
- [2] A. Sharma, A. Saxena, M. Sethi, V. Shree, Varun, Life cycle assessment of buildings: A review, *Renew. Sustain. Energy Rev.* 15 (2011) 871–875. <https://doi.org/10.1016/j.rser.2010.09.008>.
- [3] L. Cabeza, L. Rincon, V. Vilarino, G. Perez, A. Castell, Life cycle assessment (LCA) and life cycle energy analysis (LCEA) of buildings and the buildings sector: A review, *Renew. Sustain. Energy Rev.* 29 (2014) 394–416. <https://doi.org/10.1016/j.rser.2013.08.037>.
- [4] ISO 14040, Environmental management - Life cycle assessment - Principles and framework, 2006.

- [5] ISO 14044, Environmental management - Life cycle assessment - Requirements and guidelines, 2006.
- [6] C. Roux, P. Schalbart, B. Peuportier, Accounting for temporal variation of electricity production and consumption in the LCA of an energy-efficient house, *J. Clean. Prod.* 113 (2016) 532–540. <https://doi.org/10.1016/j.jclepro.2015.11.052>.
- [7] H. Feng, J. Zhao, H. Zhang, S. Zhu, D. Li, N. Thurairajah, Uncertainties in whole-building life cycle assessment: A systematic review, *J. Build. Eng.* (2022) 104191. <https://doi.org/10.1016/j.jobe.2022.104191>.
- [8] E. Marsh, S. Allen, L. Hattam, Tackling uncertainty in life cycle assessments for the built environment: A review, *Build. Environ.* (2022) 109941. <https://doi.org/10.1016/j.buildenv.2022.109941>.
- [9] A.E. Björklund, Survey of approaches to improve reliability in LCA, *Int. J. Life Cycle Assess.* 7 (2002) 64–72. <https://doi.org/10.1007/BF02978849>.
- [10] M.-L. Pannier, Etude de la quantification des incertitudes en analyse de cycle de vie des bâtiments, PhD thesis, MINES ParisTech PSL, 2017.
- [11] B. Iooss, P. Lemaître, A review on Global Sensitivity Analysis Methods, in: *Uncertain. Manag. Simul.-Optim. Complex Syst.*, Springer, 2014: pp. 101–122. [http://link.springer.com/chapter/10.1007/978-1-4899-7547-8\\_5](http://link.springer.com/chapter/10.1007/978-1-4899-7547-8_5).
- [12] R. Faivre, B. Iooss, S. Mahévas, D. Makowski, H. Monod, *Analyse de sensibilité et exploration de modèles - Application aux sciences de la nature et de l'environnement*, Quae, 2013.
- [13] I. Sobol, Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. Exp.* 1 (1993) 407–414. [https://doi.org/1061-7590/93/04407-008\\$9.00](https://doi.org/1061-7590/93/04407-008$9.00).
- [14] M.-L. Pannier, P. Schalbart, B. Peuportier, Comprehensive assessment of sensitivity analysis methods for the identification of influential factors in building life cycle assessment, *J. Clean. Prod.* 199 (2018) 466–480. <https://doi.org/10.1016/j.jclepro.2018.07.070>.
- [15] A.-T. Nguyen, S. Reiter, A performance comparison of sensitivity analysis methods for building energy models, *Build. Simul.* 8 (2015) 651–664. <https://doi.org/10.1007/s12273-015-0245-4>.
- [16] E.A. Groen, E.A.M. Bokkers, R. Heijungs, I.J.M. de Boer, Methods for global sensitivity analysis in life cycle assessment, *Int. J. Life Cycle Assess.* (2016) 1–13. <https://doi.org/10.1007/s11367-016-1217-3>.
- [17] M.H. Kristensen, S. Petersen, Choosing the appropriate sensitivity analysis method for building energy model-based investigations, *Energy Build.* 130 (2016) 166–176. <https://doi.org/10.1016/j.enbuild.2016.08.038>.
- [18] S. Yang, W. Tian, E. Cubi, Q. Meng, Y. Liu, L. Wei, Comparison of Sensitivity Analysis Methods in Building Energy Assessment, *Procedia Eng.* 146 (2016) 174–181. <https://doi.org/10.1016/j.proeng.2016.06.369>.
- [19] S. Akkari, Etude et amélioration de l'application de Bayesian calibration dans modèle énergétique des bâtiment, PhD thesis, Université Paris sciences et lettres, 2022. <https://www.theses.fr/s233656>.
- [20] M.D. Morris, Factorial Sampling Plans for Preliminary Computational Experiments, *Technometrics.* 33 (1991) 161–174. <https://doi.org/10.1080/00401706.1991.10484804>.
- [21] M.J.W. Jansen, Analysis of variance designs for model output, *Comput. Phys. Commun.* 117 (1999) 35–43. [https://doi.org/10.1016/S0010-4655\(98\)00154-4](https://doi.org/10.1016/S0010-4655(98)00154-4).
- [22] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, *Comput. Phys. Commun.* 181 (2010) 259–270. <https://doi.org/10.1016/j.cpc.2009.09.018>.
- [23] H. Monod, C. Naud, D. Makowski, Uncertainty and sensitivity analysis for crop models, in: *Work. Dyn. Crop Models Eval. Anal. Parameterization Appl.*, 2006: pp. 55–100.
- [24] A. Janon, T. Klein, A. Lagnoux-Renaudie, M. Nodet, C. Prieur, Asymptotic normality and efficiency of two Sobol index estimators, *ESAIM Probab. Stat. EDP Sci.* (2014) 342–364. <https://doi.org/10.1051/ps/2013040>.
- [25] B. Peuportier, I. Blanc-Sommereux, Simulation tool with its expert interface for the thermal design of multizone buildings, *Int. J. Sol. Energy.* 8 (1990) 109–120. <https://doi.org/10.1080/01425919008909714>.
- [26] IZUBA Energies, Software Pleiades, IZUBA Énerg. <https://www.izuba.fr/logiciels/outils-logiciels/>.
- [27] A. Brun, C. Spitz, E. Wurtz, L. Mora, Behavioural comparison of some predictive tools used in a low-energy building, in: *Elev. Int. IBPSA Conf.*, Glasgow, Scotland, 27–30 July 2009, 2009: pp. 27–30.
- [28] F. Munaretto, T. Recht, P. Schalbart, B. Peuportier, Empirical validation of different internal superficial heat transfer models on a full-scale passive house, *J. Build. Perform. Simul.* (2017) 1–22. <https://doi.org/10.1080/19401493.2017.1331376>.
- [29] W. Tian, A review of sensitivity analysis methods in buildings energy analysis, *Renew. Sustain. Energy Rev.* 20 (2013) 411–419. <https://doi.org/10.1016/j.rser.2012.12.014>.
- [30] S. Ligier, P. Schalbart, B. Peuportier, Development of a Methodology to Guaranteed Energy Performance, in: *Eco-Des. Build. Infrastruct.*, CRC Press, 2020.
- [31] P. Schalbart, E. Vorger, B. Peuportier, Stochastic Prediction of Residents' Activities and Related Energy Management, in: S. Ploix, M. Amayri, N. Bouguila (Eds.), *Energy Smart Homes Algorithms Technol. Appl.*, Springer International Publishing, Cham, 2021: pp. 543–604. [https://doi.org/10.1007/978-3-030-76477-7\\_17](https://doi.org/10.1007/978-3-030-76477-7_17).
- [32] B. Polster, Contribution à l'étude de l'impact environnemental des bâtiments par analyse du cycle de vie, PhD thesis, École nationale supérieure des mines de Paris, 1995.
- [33] E. Popovici, Contribution to the life cycle assessment of settlements, PhD thesis, École Nationale Supérieure des Mines de Paris, 2005.
- [34] B. Peuportier, D. Kellenberger, D. Anink, H. Mötzl, J. Anderson, S. Vares, J. Chevalier, H. König, Inter-comparison and benchmarking of LCA-based environmental assessment and design tool, in: *Varsovie*, 2004: p. 74.
- [35] N. Salmon, L. Duclos, F. Filit, B. Peuportier, G. Herfray, J. Chevalier, N. Schiopu, S. Lasvaux, A. Lebert, J.-L. SÉNÉGAS, R. Mikolase, O. Sidler, T. Riester, *Connaissance de l'impact environnemental des bâtiments // COIMBA 2011, Développement des outils d'évaluation de la qualité environnementale des bâtiments par analyse de cycle de vie*, Nobatek, ARMINES, CSTB, IZUBA Energies,

- ENERTECH, 2011. [http://www.nobatek.com/downloads/Etudes%20publiques/Coimba\\_Synthese\\_NOBATEK.pdf](http://www.nobatek.com/downloads/Etudes%20publiques/Coimba_Synthese_NOBATEK.pdf).
- [36] A. Lebert, ARMINES - CES, BYCN, CSTB, Fédération maisons de qualité, IZUBA Energies, Projet ANR BENEFIS : Bilan ENergétique et Environnemental Fiable Simple et reproductible des bâtiments, Tâche 5 : diffusion et communication, Rapport final, ANR 2011 VILD 001 01, DEE/EICV - 14.100, 2014. [extranet.cstb.fr/sites/anr/benefis/.../Rapports%20du%20projet/BENEFIS\\_5\\_Com.pdf](http://extranet.cstb.fr/sites/anr/benefis/.../Rapports%20du%20projet/BENEFIS_5_Com.pdf).
- [37] IPCC Working Group I, Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2013.
- [38] R. Frischknecht, F. Wyss, S.B. Knöpfel, T. Lützkendorf, M. Balouktsi, Cumulative energy demand in LCA: the energy harvested approach, *Int. J. Life Cycle Assess.* 20 (2015) 957–969. <https://doi.org/10.1007/s11367-015-0897-4>.
- [39] J.B. Guinée, Handbook on life cycle assessment operational guide to the ISO standards, *Int. J. Life Cycle Assess.* 7 (2002) 311–313. <https://doi.org/10.1007/BF02978897>.
- [40] B.P. Weidema, Bauer, R. Hischer, C. Mutel, T. Nemecek, J. Reinhard, C. Vadenbo, G. Wernet, Overview and methodology, Data quality guideline for the Ecoinvent database version3, Ecoinvent Report 1(v3), St. Gallen: The ecoinvent Centre, 2013.
- [41] M. Goedkoop, R. Heijungs, M. Huijbregts, A. de Schryver, J. Struijs, R. van Zelm, ReCiPe 2008 - A life cycle impact assessment method wich comprises harmonised category indicators at the midpoint and the endpoint level, First edition, Report I : Characterisation, 2013.
- [42] R.K. Rosenbaum, T.M. Bachmann, L.S. Gold, M.A.J. Huijbregts, O. Jolliet, R. Juraske, A. Koehler, H.F. Larsen, M. MacLeod, M. Margni, T.E. McKone, J. Payet, M. Schuhmacher, D. van de Meent, M.Z. Hauschild, USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment, *Int. J. Life Cycle Assess.* 13 (2008) 532–546. <https://doi.org/10.1007/s11367-008-0038-4>.
- [43] Pannier, Marie-Lise, Patrick Schalbart, et Bruno Peuportier. 2023. « Dealing with Uncertainties in Comparative Building Life Cycle Assessment ». *Building and Environment* 242 (août): 110543. <https://doi.org/10.1016/j.buildenv.2023.110543>.
- [44] G. Myhre, S. Drew, Anthropogenic and Natural Radiative Forcing - Supplementary Material of Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], 2013. <http://www.ipcc.ch/report/ar5/wg1/>.
- [45] INIES, French environmental database for construction product : Base de donnée INIES, (2015). <http://www.base-inies.fr/inies/>.
- [46] ecoinvent, Ecoinvent database, <https://ecoinvent.org/>.
- [47] A. Merzkirch, S. Maas, F. Scholzen, D. Waldmann, Field tests of centralized and decentralized ventilation units in residential buildings – Specific fan power, heat recovery efficiency, shortcuts and volume flow unbalances, *Energy Build.* 116 (2016) 376–383. <https://doi.org/10.1016/j.enbuild.2015.12.008>.
- [48] I.A. Macdonald, Quantifying the effects of uncertainty in building simulation, University of Strathclyde. Dept. of Mechanical Engineering, 2002.
- [49] C. Spitz, L. Mora, E. Wurtz, A. Jay, Practical application of uncertainty analysis and sensitivity analysis on an experimental house, *Energy Build.* 55 (2012) 459–470. <https://doi.org/10.1016/j.enbuild.2012.08.013>.
- [50] E. Hoxha, G. Habert, J. Chevalier, M. Bazzana, R. Le Roy, Method to analyse the contribution of material’s sensitivity in buildings’ environmental impact, *J. Clean. Prod.* 66 (2014) 54–64. <https://doi.org/10.1016/j.jclepro.2013.10.056>.
- [51] M.-L. Pannier, Analyse de cycle de vie des maisons passives de la plateforme INCAS, INSA de Strasbourg, 2014. [http://eprints2.insa-strasbourg.fr/1743/1/Pannier\\_Marie-Lise\\_GCE\\_-\\_PFE\\_-\\_Synth%C3%A8se.pdf](http://eprints2.insa-strasbourg.fr/1743/1/Pannier_Marie-Lise_GCE_-_PFE_-_Synth%C3%A8se.pdf).
- [52] CETE Sud-Ouest, Transports de matériaux de carrières et de construction en Poitou-Charentes, 2002. <http://www.ort-poitou-charentes.asso.fr/wp-content/uploads/2015/06/materiauxdecar.pdf>.
- [53] CETE Lyon, Prévention et gestion des déchets issus de chantier du bâtiment et des travaux publics. Guide méthodologique à l’attention des Conseils Généraux et de la région Ile-de-France, 2013. [http://www.lagazettedescommunes.com/wp-content/uploads/ct69\\_detc\\_dspes\\_guide\\_planification\\_BTP.pdf](http://www.lagazettedescommunes.com/wp-content/uploads/ct69_detc_dspes_guide_planification_BTP.pdf).
- [54] Sispea, Observatoire national des services d’eau et d’assainissement, (2015). <https://www.services.eaufrance.fr/donnees/telechargement>.