LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis

Martino Lacirignola, Philippe Blanc, Robin Girard, Paula Perez-Lopez, Isabelle Blanc

Science of the Total Environment xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis

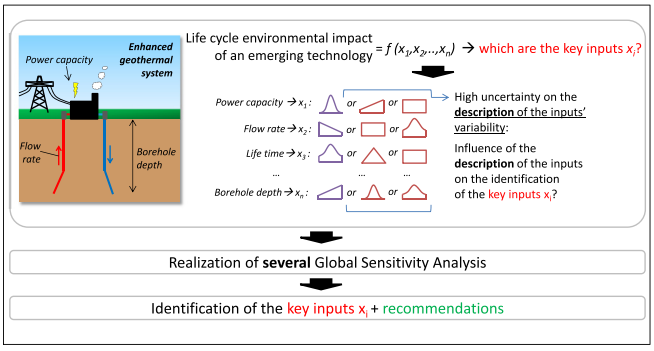
Martino Lacirignola [a,*], Philippe Blanc [b], Robin Girard [c], Paula Pérez-López [b], Isabelle Blanc [b]

[a] French Environment and Energy Management Agency (ADEME), Energy Networks and Renewable Energy Department, 27 rue Louis Vicat, 75737 Paris Cedex 15, France
[b] MINES ParisTech, PSL Research University, Centre Observation, Impacts, Energy (O.I.E.), CS 10207, F-06904 Sophia-Antipolis, France
[c] MINES ParisTech, PSL Research University, Centre for Processes, Renewable Energies and Energy Systems (PERSEE), CS 10207, F-06904 Sophia-Antipolis, France

HIGHLIGHTS

- Global sensitivity analysis (GSA) is sensitive to the description of LCA inputs.
- The robustness of GSA results in the LCA of innovative products needs to be assessed.
- A strategy to analyze the influence of the description of the inputs variability
- Several GSA are reiterated and recommendations retrieved for the key inputs selection.
- Case study: identification of key parameters of enhanced geothermal systems LCA

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 7 August 2016
Received in revised form 8 October 2016
Accepted 9 October 2016
Available online xxxx

Editor: D. Barcelo

Keywords:
Global sensitivity analysis
Life cycle assessment
Enhanced geothermal systems
Environmental model
Sobol indices

ABSTRACT

In the life cycle assessment (LCA) context, global sensitivity analysis (GSA) has been identified by several authors as a relevant practice to enhance the understanding of the model's structure and ensure reliability and credibility of the LCA results. GSA allows establishing a ranking among the input parameters, according to their influence on the variability of the output. Such feature is of high interest in particular when aiming at defining parameterized LCA models.

When performing a GSA, the description of the variability of each input parameter may affect the results. This aspect is critical when studying new products or emerging technologies, where data regarding the model inputs are very uncertain and may cause misleading GSA outcomes, such as inappropriate input rankings. A systematic assessment of this sensitivity issue is now proposed.

We develop a methodology to analyze the sensitivity of the GSA results (i.e. the stability of the ranking of the inputs) with respect to the description of such inputs of the model (i.e. the definition of their inherent variability). With this research, we aim at enriching the debate on the application of GSA to LCAs affected by high uncertainties.

We illustrate its application with a case study, aiming at the elaboration of a simple model expressing the life cycle greenhouse gas emissions of enhanced geothermal systems (EGS) as a function of few key parameters. Our methodology allows identifying the key inputs of the LCA model, taking into account the uncertainty related to their description.

* Corresponding author.
E-mail address: martino.lacirignola@ademe.fr (M. Lacirignola).

http://dx.doi.org/10.1016/j.scitotenv.2016.10.066
0048-9697/

1. Introduction

Life cycle assessment (LCA) is widely considered as the most relevant methodology to assess the environmental performances of products and processes over their life cycle and is currently applied to different industrial sectors (Jacquemin et al., 2012; Moomaw et al., 2011). Due to the inherent variability of the input parameters, the large number of assumptions and sometimes the incomplete knowledge of modeled process, the importance of assessing uncertainties through sensitivity analysis (SA) has been stressed since the early development of the LCA methodology (Heijungs, 1996; Huijbregts, 1998; Lloyd and Ries, 2007). The ISO standard for LCA (ISO 14040, 2006; ISO 14044, 2006) also indicates SA as a fundamental part of the analysis, without however recommending a particular calculation technique.

In the LCA context, global sensitivity analysis (GSA) has been recently identified by several authors as a relevant practice to address several issues: (i) to study the combined influence of the different input parameters (Padey et al., 2013), (ii) to assess the robustness of the results (Wei et al., 2015), (iii) to enhance the understanding of the structure of the model (Cucurachi et al., 2016) (iv) to ensure transparency, reliability and credibility of LCA practices (Bisinella et al., 2016) and (v) to contribute to the decision-making process (Andrianandraina et al., 2015). GSA allows establishing a ranking among the input parameters and identifying the most influential on the variability of the output of the model. The identification of such key parameters is fundamental when aiming at the simplification of the uncertainty quantification: in fact, based on the GSA results, the efforts to minimize the uncertainty can be focused only on few key input variables while the others can be fixed to average values without influencing the results (Bisinella et al., 2016; Wei et al., 2015). Identifying the most influent variables also allows developing simplified parameterized LCA models (Padey et al., 2013). In general, GSA techniques support the execution of LCAs and facilitate its interpretation, promoting an enhanced decision making process (Cucurachi et al., 2016).

The question of how to perform GSA in a LCA context has been addressed by few studies (Cucurachi et al., 2016; Andrianandraina et al., 2015; Wei et al., 2015; Bisinella et al., 2016). In particular, the recent work of Cucurachi et al. (2016) proposes a comprehensive multi-step protocol for the integration of sensitivity and uncertainty analysis in the impact assessment phase of LCAs. Examples of application of GSA techniques to LCAs are also to be found in Lacirignola et al. (2014), Marini and Blanc (2014), Azadi et al. (2015) and Cucurachi and Heijungs (2014).

When conducting a GSA, the description of the variability of each input parameter is one of the most important steps, because it could significantly affect the GSA results (Wei et al., 2015). This step is called in this paper "*description of the inputs*" and consists in defining (i) the minimum and maximum values they can assume and (ii) if some values are more probable than others within those boundaries. Such description is done by the LCA modeler, who allocates a probability distribution (Gaussian, uniform, or any other) over the defined range of variability of each input. This process is based on expert opinions, literature survey or even better on field data. While mentioning the importance of the description of the inputs, the above-mentioned studies however do not propose a systematic assessment of its influence on the GSA results.

In an ideal case, the modeler has a high confidence on the performed description. This can be found for instance in the GSA performed by Padey et al. (2013) on the LCA of wind turbines: in this study, the descriptions of two of the input variables (the load factor and the nominal power of the machines) are based on data collected from hundreds of turbines currently installed in France. In general, when the system analyzed is well known, the application of GSA protocols available in literature (see for instance Cucurachi et al., 2016) is adequate to clearly identify the main drivers of the model.

On the contrary in other cases, especially when studying new products or emerging technologies, the level of confidence of the inputs' description is significantly low due to the limited amount of available information. For instance such issue is to be found in Lacirignola et al. (2014), where GSA is applied to the LCA of an innovative energy technology: the enhanced geothermal systems (EGS), of which less than ten installations exist today in Europe (Van Wees et al., 2013; Baujard et al., 2015). In such LCA the description of the input parameters is quite uncertain because it is only based on the few data available from the industry, discussion with experts and literature survey (for instance Genter et al., 2010 and Huenges, 2010). In this context the robustness of the GSA results needs to be further investigated.

The goal of the present study is to set up a methodology to perform such investigation, overcoming the critical issue of handling very uncertain information regarding the input parameters. It also aims at enriching the debate on the application of GSA to LCAs, by focusing on this specific critical issue. Our research focuses on the sensitivity of the GSA itself (i.e. the variability of the GSA results and the identification of the key inputs) with respect to the description of such inputs (i.e. the definition of their variability). This issue has never been addressed before despite being of paramount importance, especially when studying innovative products.

Starting from the GSA protocol presented by Cucurachi et al. (2016), we propose a methodology that relies on the reiteration of several GSA calculations under different hypothesis regarding the description of the input parameters. This allows assessing the stability of the parameters' ranking, while considering the level of confidence of their description. We also analyze whether the contribution of one input to the output's variance appears to vary significantly from one calculation to another. We then retrieve relevant recommendations for the selection of the key parameters of the model. Our methodology is presented in Section 2. We illustrate its application with a case study (Section 3) focused on the environmental analysis of enhanced geothermal systems (EGS) (based on Lacirignola et al., 2014). The latter aims at the elaboration of a simplified LCA model for the estimation of the greenhouse gases (GHG) emissions of this innovative energy technology. Starting from a reference parameterized model, we take advantage of GSA to obtain a reduced model only based on the key parameters. The case study is divided in two parts. We first apply a "baseline" GSA approach based on Sobol Indices, then we highlight the sensitivity of the GSA results with respect to the inputs' description with an example. In the second part, we illustrate how our methodology is addressing this sensitivity issue and allows identifying the most influent input parameters accounting for the uncertainty related to their description. Section 4 proposes a critical discussion of our approach, presenting the possible variants and axes of improvement, in particular aiming at reducing the computational costs. Section 5 presents the concluding remarks, underlining the importance of a wise use of GSA techniques for a correct understanding of the LCA model.

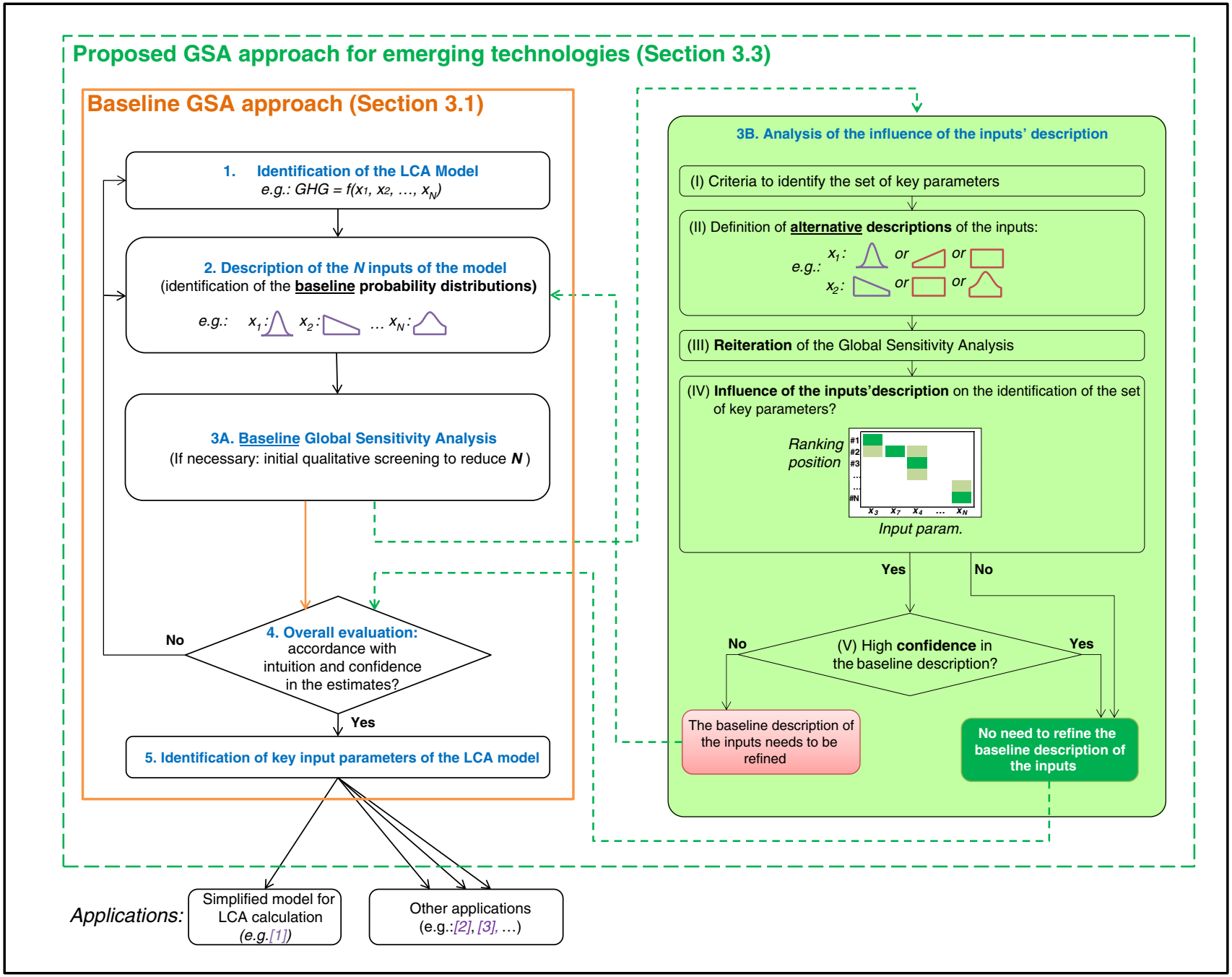
2. Methodology

The methodology we propose to perform GSA in the LCA context handling very uncertain assumptions regarding the inputs' description is presented in Fig. 1. It starts from the protocol presented by Cucurachi et al. (2016), which is further extended by setting a strategy to analyze the influence of the inputs' description on their ranking obtained through the GSA. Such ranking, i.e. the ordered list of the parameters from the more to the less influent on the model output, allows identifying the key inputs of the LCA model. As stated in the Introduction, other authors have also presented frameworks of application of the GSA in the LCA context (Andrianandraina et al., 2015; Wei et al., 2015; Bisinella et al., 2016): in Fig. 1, we use the term "baseline" GSA approach to refer to a GSA

Fig. 1. Methodology for the identification of key input parameters of a LCA model accounting for the influence of the inputs' description, starting from the GSA protocol proposed by Cucurachi et al. (2016). ([1] = Lacirignola et al. (2014); [2] = Andrianandraina et al. (2015); [3] = Bisinella et al. (2016)).

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066



procedure that does not take into account our additional analysis (proposed in Step 3B) of the influence of the inputs' description.

2.1. Step 1: Identification of the LCA model

In the initial step the modeler defines the calculation model, namely the computational structure used to estimate the life cycle impacts for the studied impact category according to a set of model parameters. According to the goal and scope of the model, the inputs are identified (such as the type and the amount of materials involved over the life cycle of the product and the energy required) and the input-output relation is formalized (for instance, the model allowing to calculate the life cycle emissions of greenhouse gases).

2.2. Step 2: Description of the inputs of the model

The modeler identifies for each of the N input parameters (a) the boundaries of its range of variability and (b) one probability distribution applied to such variability range (we call it "baseline" distribution to distinguish it from the "alternative" ones identified in Step 3B, see Section 2.4.2). For a given parameter, the baseline probability distribution (Gaussian, uniform, lognormal, or other) reflects the best current knowledge regarding the variability of the input, according to the goal and scope of the model. This is based on available data (from literature or other sources) and expert knowledge. Once the distributions are established, a random sample of each of the inputs is generated.

Based on the information provided in this step, it is also possible to analyze the propagation of the inputs' uncertainty by generating a Monte Carlo sample of the output (or also using other propagation methods, such as those dedicated to high dimensional input spaces and faster convergence).

2.3. Step 3A: Baseline global sensitivity analysis

In this step, the global sensitivity analysis is performed. Since this is based on the baseline probability distributions established in the previous phase, we call it the "baseline" GSA, which allows distinguishing it from the other GSAs performed further in Step 3B (see Section 2.4.3).

If the number of model inputs is high, the modeler may here performs an initial screening in order to identify the non-influential parameters and fix them to average values, as proposed by Andrianandraina et al. (2015). This can be done by applying different screening methods available in the literature, such as the qualitative approach illustrated by Morris (1991).

The methodology that we propose relies on the hypothesis that the input parameters are independent. In this case it is possible to describe, through a probability distribution, the variability of each parameter independently. Several GSA methods can be found in the literature, see for instance Groen et al. (2016), Wolf et al. (2016) and Padey et al. (2013). For comprehensive reviews of the available options we suggest referring to Saltelli et al. (2008) or Iooss (2011). In our framework we suggest to use the methodology proposed by Sobol' (2001), which is based on the decomposition of the variance and estimates sensitivity indicators called Sobol indices. They are appropriate for our analysis since they provide a quantitative measure, thus they allow computing easily the ranking among the variables, and they have a convenient interpretation in terms of explained variance of output. For instance, for a given model $z = f(x_1, x_2, \ldots, x_N)$, the first order Sobol index, denoted S_i^{First}, indicates the contribution of the variance of the input x_i to the overall variance of the output z (Eq. 1).

$$S_i^{First} = \frac{V[\mathbb{E}[z|x_i]]}{V[z]} \quad (1)$$

Thus if for example the input x_i has a $S_i^{First} = 0.2$ he contributes to 20% of the overall variance of the output z.

Since the GSA is based on random samples of data (generated in Step 2), we recommend performing bootstrapping to assess the confidence of the GSA results. The modeler must also assess if the size of such samples is large enough to reduce the effects of numerical instability. For instance, if overlapping is observed among the uncertainty ranges of the Sobol indices, the modeler must reduce it to the minimum by enlarging the sample size.

The final output of Step 3A provides the ranking among the input parameters of the model and a quantitative measure of their relative importance. This allows identifying the ones that are responsible for most of the variability of the output, i.e. those displaying the largest Sobol Indices. If the indices of two or more parameters are very close, different bootstraps of the baseline GSA may produce different rankings. The modeler must take note of this residual ranking instability, because the latter will also appear in the results of Step 3B.

If the modeler is sufficiently confident on all the probability distributions established to describe the inputs, after Step 3A he can proceed directly to an overall evaluation of the results (Step 4) and to the selection of the key parameters (Step 5). On the contrary, if the description of one or more inputs is particularly uncertain (e.g. when studying new products or emerging technologies), we propose to proceed from Step 3A to the additional Step 3B before the conclusive Steps 4 and 5.

2.4. Step 3B: Analysis of the influence of the inputs' description

This step aims at studying if and how the identification of the set of key parameters is influenced by the description of the inputs. To achieve this, the GSA will be reiterated several times under different input conditions: this will lead to the production of a number of possibly different rankings of the inputs.

2.4.1. 3B – (I) Criteria to identify the set of key parameters

In this step, the modeler must define the criteria to detect the inputs' description that are eventually influent on the identification of the set of key parameters. To do that, the modeler must first clarify what is the condition for being identified within the set of key parameters, by establishing a targeted threshold for their "aggregated contribution", for example 60% (or more). In this case that the key parameters (showing the highest S_i^{First}) must be together responsible of at least 60% of the overall variability of the output: namely the sum of their S_i^{First} must be higher than 0.6. Indeed, the number of selected key parameters depends on this threshold: for instance two key parameters may be sufficient in the baseline scenario, nevertheless a deeper analysis may show that - under different hypothesis - three or even four parameters may be necessary to achieve the targeted 60%. Therefore, the modeler will be interested in observing whether the set of key parameters remains the same or not after different GSA calculations. If such ambiguity is found, then the description of the inputs has a significant influence.

An alternative approach for the selection of the key parameters consists in focusing on their single contribution rather than their aggregated one (i.e. observing if each S_i^{First} is above a certain threshold). However, such approach alone may not be sufficient to identify a set able to cover a given share of the output variance: in the case study, we'll use it only for a complementary analysis (for more details see also to the Supporting information S8).

2.4.2. 3B – (II) Definition of alternative descriptions of the inputs

This phase conceptually corresponds to the Step 2 previously described (Section 2.2). For each of the N input parameters, the modeler identifies a number of possible alternative descriptions (i.e. other possible probability distributions applied to its interval of variability, different from the baseline one established in Step 2). The number of k_i alternative descriptions (including the baseline one) is set by the modeler and may be different for each *i-th* parameter.

For instance, if the baseline distribution of one parameter is Gaussian with $\mu = 10$ and $\sigma = 1$, an alternative distribution may have a different shape (e.g. triangular), or a different mean (e.g. $\mu = 12$), or a different standard deviation (e.g. $\sigma = 2$), or a combination of all these changes.

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066

In any case, at this step it is necessary to consider realistic constraints: this can mean for example avoiding negative values for a parameter that is meant to be only positive. The general condition is that each of the alternative distribution must be plausible according to the current knowledge and to the goal and scope of the model. A random sample is then generated from the established descriptions. The difference between the alternative and the baseline distributions of one variable can be quantified by measuring the different means and standard deviations: this will enhance the understanding of how the space of distributions is explored (see the Supporting information S7).

2.4.3. 3B – (III) Reiteration of the GSA

This step conceptually corresponds to Step 3A (performing the GSA, Section 2.3) but the GSA is here reiterated several times, each time considering a different set of descriptions of the input parameters. The same sensitivity indices chosen in Step 3A are used here. Concretely, a *one factor at a time* (OAT) approach is used for reiteration of the GSAs, as detailed hereafter. The analysis starts with the observation of the first of the N input parameters. Here, k_1 GSAs are performed, each time considering one of the several k_1-*th* distributions for the first parameter, while the baseline distribution is set for all the other parameters. Bootstrapping can also be used at every stage of calculation (in the Eq.2, B is the amount of bootstraps per GSA). The same process is repeated for the other *i-th* parameters. At the end of each bootstrap, the global sensitivity indices and the obtained ranking (i.e. the sorted order of the inputs according to their influence on the output's variability) is stored. Without repeating the case where all distributions are the baseline ones, the total number of GSAs performed in this study (i.e. the total number of rankings recorded) is given in Eq.2:

$$\text{Total number of rankings} = \left[\sum_{i}^{N} k_i - 1 + 1\right] \cdot B \tag{2}$$

Reiterating the GSA with such an OAT approach regarding the input conditions does not account for all the possible combinations of descriptions of the input parameters. However, it still allows formulating relevant observations for the scope of the study, while keeping the process relatively simple. A more global approach, which would consider all the possible combinations of descriptions of the inputs (requiring a much higher computational cost), is discussed in Section 4 (Discussion).

2.4.4. 3B – (IV) Analysis of the influence of the inputs' description

Based on the calculations performed in the previous phase, the modeler here analyzes the sensitivity of the GSA outputs (the global sensitivity indices and the collection of rankings of the input parameters) with respect to the description of the inputs. This is done according to the criteria established at the beginning of Step 3B (Section 2.4.1).

In this phase, the modeler first identifies *how many* key parameters need to be selected to achieve the targeted threshold for their aggregated contribution (defined in Step 3B – (I)). Then he identifies which description of the inputs has a significant influence on this selection process. In other words, the modeler here finds out if describing one input with e.g. a Gaussian instead of a uniform distribution leads him to identify different sets of key parameters. Such analysis is performed one factor at the time (observing the $k_i \cdot B$ rankings related to each single parameter) and also by examining the aggregated results of all the GSAs (observing the whole set of obtained rankings). The modeler will identify *which are* the key parameters later, in Step 5: it is worth to remind that if one parameter is selected as "key", it doesn't necessarily mean that the description of its variability is influent on the selection process.

2.4.5. 3B – (V) Consideration of the level of confidence of the inputs' description

If the description of a parameter is found to be influent on the identification of the key parameters and its level of confidence is low (i.e. it is based on numerous assumptions), such description should be refined. In other words, the modeler should try to collect further data to validate or improve the baseline probability distribution used. However, especially when studying new products or emerging technologies, it may happen that a more detailed analysis is not possible because of lack of existing additional data: in this case, an alert must appear when exploiting the GSA results. This may affect the identification of the key parameters: with a conservative approach, a larger number of key parameters may eventually be selected (i.e. including those affected by the alerts).

Conversely, the modeler may find out that the description of a parameter is particularly influent on the GSA output, while also being confident about the baseline probability distribution chosen in Step 2 (obtained for example from a sufficiently large statistical sample of data). In this case the input's description doesn't need to be refined.

As discussed in Section 4 (Discussion), this Step 3B - (V) can also be performed earlier, inquiring about the level of confidence before the beginning of Step 3B in order to simplify the calculation process. It is also important to remind that the aim of this analysis is not to identify which is the best description for a parameter among several alternatives. The goal is to enhance the understanding of the model, and to formulate appropriate alerts and recommendations for the use of the GSA results.

2.5. Step 4: Overall evaluation

At the end of every step of the proposed methodology, the modeler should verify if the obtained results are in accordance with intuition, check for misleading interpretations and eventually reiterate partially or totally the calculation process if needed. For instance, the modeler must be alerted by results of the output of the LCA model (generated in Step 2) that are too far from those available in literature, or by the observation of drastic changes in the ranking position (e.g. from the first to the last) of one parameter during Step 3B. Such process of consistency check should be continuous, but for simplicity we represent in our methodology (Fig. 1) just one step of overall evaluation of the results, labeled as Step 4: this constitute the minimum requirement in terms of consistency check.

2.6. Step 5: Identification of key input parameters of the LCA model

Based on the outcome of the previous steps, the modeler formulates conclusions and recommendations for the selection of the key inputs of the model. Namely, he identifies *which are* the key parameters able to cover a sufficient share of the variability of the output (for instance 60%, as in the example provided in Section 2.4.1). If the description of one or more inputs provoked an alert in Step 3B – (V), the modeler must take it into account in the process of selection of the key parameters, as discussed in Section 2.4.5.

2.7. Application of the GSA results

As stated in the Introduction (Section 1), one possible application of the GSA in the LCA context is the elaboration of simplified calculation models, where the life cycle impacts are expressed as a function of few key parameters identified through the GSA. Other possible uses of the GSA results can be found in Andrianandraina et al. (2015), where eco-designed scenarios are established using the lower values of the most influential drivers, and in Bisinella et al. (2016), where the authors propose to recalculate the uncertainty propagation considering only the key parameters.

3. Case study

The object of the case study is a renewable energy based emerging technology: the enhanced geothermal systems. The methodology described in Section 2 is applied to identify the key input parameters of an EGS LCA model. The latter is based on the one recently published

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066

by Lacirignola et al. (2014) and is presented here in an updated form that accounts for the most recent characterization factors corresponding to the IPCC, 2013 method (IPCC, 2013).

Section 3 is structured in two parts. In Section 3.1 we apply a baseline GSA approach, without further questioning the description of the inputs of the model (i.e. we perform Steps 1, 2, 3A, 4 and 5, boxes within

Table 1
Description of the nine parameters of the EGS Reference parameterized model.

Parameter	Value range [unit]	Baseline distribution (type 1)	Alternative distributions			
			Type 2	Type 3	Type 4	Type 5
Borehole depth (z)	2,000 – 6,000 [Meters]	Uniform μ: 4,000 σ: 1,155	Δμ: 0% Δσ: −43%	Δμ: −11% Δσ: −8%	Δμ: 11% Δσ: −8%	Δμ: 0% Δσ: −14%
Scaling factor enhan. (Sfe)	0.5 – 10 [Ad.]	Trunc–lognormal μ: 2.10 σ: 1.53	Δμ: 99% Δσ: 65%	Δμ: 150% Δσ: 79%	Δμ: 150% Δσ: 2%	Δμ: 150% Δσ: 53%
Flow rate (f)	25 – 100 [kg/s]	Uniform μ: 62.5 σ: 21.7	Δμ: 0% Δσ: −43%	Δμ: −13% Δσ: −8%	Δμ: 13% Δσ: −8%	Δμ: 0% Δσ: −14%
Fuel for drilling (d)	3,000 – 7,000 [MJ/m]	Uniform μ:5,000 σ:1,155	Δμ: 0% Δσ: −43%	Δμ: −9% Δσ: −8%	Δμ: 9% Δσ: −7%	Δμ: 0% Δσ: −14%
Lifetime (LT)	20 – 40 [Years]	Trunc–gaussian μ: 30 σ: 3.3	Δμ: 0% Δσ: 76%	Δμ: −7% Δσ: 62%	Δμ: 7% Δσ: 62%	Δμ: 0% Δσ: 51%
Load factor (LF)	0.85 – 0.95 [Ad.]	Uniform μ: 0.90 σ: 0.03	Δμ: 0% Δσ: −43%	Δμ: −1% Δσ: −8%	Δμ: 1% Δσ: −8%	Δμ: 0% Δσ: −14%
Pumps specific power (P_p)	3.6 – 8.6 [kW/(kg/s)]	Uniform μ: 6.1 σ: 1.4	Δμ: 0% Δσ: −43%	Δμ: −9% Δσ: −8%	Δμ: 9% Δσ: −7%	Δμ: 0% Δσ: −14%
Number of wells (Nw)	2 or 3 [Ad.]	50% 50% μ: 2.5 σ: 0.5	70% 30% Δμ: −8% Δσ: −8%	30% 70% Δμ: 8% Δσ: −8%	–	–
Installed capacity ORC (P_{ORC})	1,250 – 3,500 [kW]	Uniform μ: 2,375 σ: 650	Δμ: 0% Δσ: −43%	Δμ: −11% Δσ: −8%	Δμ: 11% Δσ: −8%	Δμ: 0% Δσ: −14%

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066

the orange square in Fig. 1). After showing the interest of assessing the sensitivity of the GSA results (Section 3.2), in Section 3.3 we perform such investigation, by handling the same five steps mentioned above with the addition of Step 3B (boxes within the green dotted square in Fig. 1).

The final purpose of identifying the key inputs in this case study is the setting of a simplified model, aimed for decision makers, able to estimate life-cycle GHG emissions of EGS as a function of a very limited number of parameters. The development of simplified LCA models based on the key parameters, initiated by Padey et al. (2013), is one the possible applications of GSA in the LCA context and is emerging as a useful technique aimed at decision makers. Simplified LCA models are designed to address two of the commonly agreed drawbacks of LCAs: (i) the complexity of the LCA process, which is time consuming and requires expert knowledge (ii) the applicability of the results, which often correspond only to specific configurations of the system analyzed. These drawbacks may lead policy-makers to consider LCA as a pretty inconclusive method, especially when aiming at comparing different technological options (Heath et al., 2010). Simplified LCA models aim at overcoming these issues.

3.1. Application of the baseline GSA approach

3.1.1. Step 1: Identification of the LCA model

The LCA model analyzed in this case study, called "*Reference model*" (Eq. 3 below), is designed for the analysis of the GHG performances of EGS installed in central Europe and takes into account current technologies for all the equipment (Genter et al., 2010; Hettkamp et al., 2011; Baujard et al., 2015; Bestec, 2012). The boundaries of the system analyzed include both sub-surface elements (i.e. the geothermal wells) and surface equipment, like for instance the pumps, the heat exchanger and the elements of an organic Rankine cycle (ORC) for the electricity production (no cogeneration). Materials and energy flows related to the hydraulic and chemical stimulation of the geothermal reservoir are also accounted for.

The *Reference model* is a function of nine parameters (first column of Table 1): the borehole depth (***z***), the produced flow rate (***f***), the number of wells (***Nw***), the amount of fuel consumed during the drilling phase (***d***), the load factor expressing the amount of equivalent operating hours at nominal power in one year (***LF***), the lifetime (***LT***), a dimensionless factor expressing the intensity of the stimulation of the reservoir (***SFe***),[1] the specific power of the pumps of the geothermal loop ($\boldsymbol{P_p}$) and the installed capacity of the ORC ($\boldsymbol{P_{ORC}}$). These nine parameters allow the calculation of the GHG performances as they are sufficient to determine the size of the plant, the amount of material and energy flows involved over its lifecycle and the total amount of the electricity produced. Data regarding the background processes (e.g. raw material extraction or steel production) necessary to elaborate the EGS life cycle inventory are retrieved from Ecoinvent v2.2 (Ecoinvent Centre, 2010). The selected functional unit is the kWh of net electricity produced over the lifetime and delivered to the grid. Therefore, the output results are expressed in terms of grams of CO_2 equivalents per kWh.

The parameterized *Reference model* is presented in Eq. 3. Further details on the development of this formula can be found in Lacirignola et al. (2014) and in the Supporting information S1.

$$\mathrm{GHG}_{\mathrm{EGS\_Ref}}\left[\frac{\mathrm{gCO_2eq}}{\mathrm{kWh}}\right] = \frac{z \cdot Nw \cdot \alpha_1 + \alpha_2 \cdot d + LT \cdot f \cdot \alpha_3 + P_{ORC} \cdot LT \cdot \alpha_4 + Nw \cdot SFe \cdot \alpha_5}{LT \cdot LF \cdot P_{ORC} - f \cdot P_p \cdot 8760} \quad (3)$$

$\alpha_1 = 498,\ 761.36\ [\mathrm{gCO_2eq/m}];\ \alpha_2 = 90.56\ [\mathrm{gCO_2eq/(MJ \cdot \mathit{m})}];$
$\alpha_3 = 487,\ 363.03\ [gCO_2eq \cdot \mathrm{s}/(kg \cdot \mathrm{y})];$
$\alpha_4 = 50,\ 603.13 [gCO_2eq/(kW \cdot \mathrm{y})];\ \alpha_5 = 25,\ 757,\ 089.05\ [gCO_2eq];$

Starting from this model, GSA is applied to reduce the number of parameters of Eq. 3, thus obtaining a more simple calculation tool (called "*Reduced model*"), which can be easily used by decision makers. Indeed, once identified the key parameters, we express the GHG performances as a function of only those few variables, while fixing the others to their median value.

3.1.2. Step 2: Description of the inputs of the model

In order to consider a large panel of possible EGS configurations, each of the nine parameters is characterized by a variability range that reflects the possible characteristics (i.e. interval of realistic values for each parameter) of the EGS in accordance with the scope of the model (see second column of Table 1). Moreover, a baseline probability distribution is associated to each variability range (third column of Table 1). This is established based on technical survey, literature review and discussion with experts. The boundaries of the intervals and the baseline distributions reflect the best current knowledge on EGS within the scope of the model, considering the current installations and the future potential power plants. In this paper, we use the same boundaries and distributions proposed by Lacirignola et al. (2014). A random sample of 500,000 values for each of the nine input parameters is then generated.

Based on these settings, a Monte Carlo sample of the output is also obtained. That is, the GHG performances of 500,000 possible randomly generated EGS configurations. We use the simple Monte Carlo sampling scheme because the dimension of the input space is relatively limited and it allows the computation of confidence intervals with bootstrap. In the Supporting information S2, we present a comparison between the output of the *Reference model* and the results of different case studies available in literature (Huenges, 2010; Frick et al., 2010; Bauer et al., 2008; Platt et al., 2012; Sullivan et al., 2013; Lacirignola and Blanc, 2013). Globally we observe that our results are coherent with literature and that most of the GHG results lay within the 23–40 $\mathrm{gCO_2eq/kWh}$ range.

3.1.3. Step 3A: Baseline global sensitivity analysis

In this step, the "baseline" GSA is performed: each input parameter is characterized by its baseline distribution within its variability range (second and third column in Table 1). The nine parameters are statistically independent (see Lacirignola et al. (2014) for a more detailed

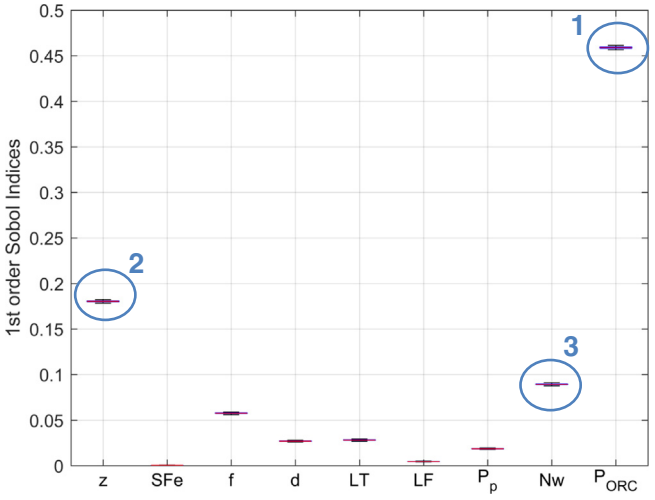


Fig. 2. Results of the baseline GSA: first order Sobol indices of the nine input parameters: the circles indicate the three highest parameters of the ranking.

[1] ***SFe*** is a multiplying factor used to scale up the values of the life cycle inventory of a "base case" stimulation of the geothermal reservoir. High values of ***SFe*** correspond to a strong enhancement campaign (i.e. high consumption of water, fuel and acid).

discussion on the independency). Under this assumption, we estimate the global sensitivity through the Sobol Indices.

The results of the baseline GSA are shown in Fig. 2. We observe that the parameter responsible for most of the variability of the GHG results is the installed capacity $\boldsymbol{P_{ORC}}$, with a S_i^{First} of 0.46. The second within the ranking is the borehole depth $\boldsymbol{z}$ ($S_i^{First} = 0.18$), followed by the number of wells $\boldsymbol{Nw}$ ($S_i^{First} = 0.09$). These three key-parameters are together responsible for about 73% of the variability of the output (sum of their first order Sobol Indices). The other six variables have a S_i^{First} lower than 0.06. Moreover, by observing the total order Sobol Indices (presented in the Supporting information S3), we see that no major interaction effects occur: hence the S_i^{First} are sufficient to identify the most influent parameters.

100 bootstraps of 500,000 random samples are performed, showing no major fluctuations of the S_i^{First} (cf. narrow boxplots in Fig. 2). Therefore, the size of the input samples is satisfactory. In the Supporting information S4, we propose an illustration of the numerical instability of the GSA results that occurs when a too small sample is used.

We also observe that the S_i^{First} of $\boldsymbol{d}$ and $\boldsymbol{LT}$ are very close, with their interquartile ranges overlapping: this means that their ranking positions may swap when running two different bootstraps. Such behavior may be avoided but with a prohibitory increase of the sample size: thus, to reduce the computational costs, we keep the size of 500,000 random samples and we take note of this residual numerical instability (presented in the Supporting information S5).

3.1.4. Steps 4 and 5: Overall evaluation and identification of the key parameters

An evaluation of the results was realized all along the calculations of the previous steps, without finding inconsistencies: results are also coherent with those of Lacirignola et al. (2014), which constitutes the basis of the case study.

For the selection of the key parameters, we set the threshold for their aggregated contribution to 66%, i.e. they must cover at least two thirds of the variability of the output. This is set to ensure a sufficient representativeness of the *Reduced model*. Within this framework, we conclude that $\boldsymbol{P_{ORC}}$, $\boldsymbol{z}$ and $\boldsymbol{Nw}$ can be identified as key parameters: they show the three highest Sobol Indices and the sum of their S_i^{First} is indeed higher than 0.66. The obtained results are also in accordance with intuition: the energy produced over the lifetime is directly related with the installed capacity ($\boldsymbol{P_{ORC}}$), and the amount of drilled meter (which depend on $\boldsymbol{z}$ and $\boldsymbol{Nw}$) is widely considered the main source of GHG emissions (Menberg et al., 2016).

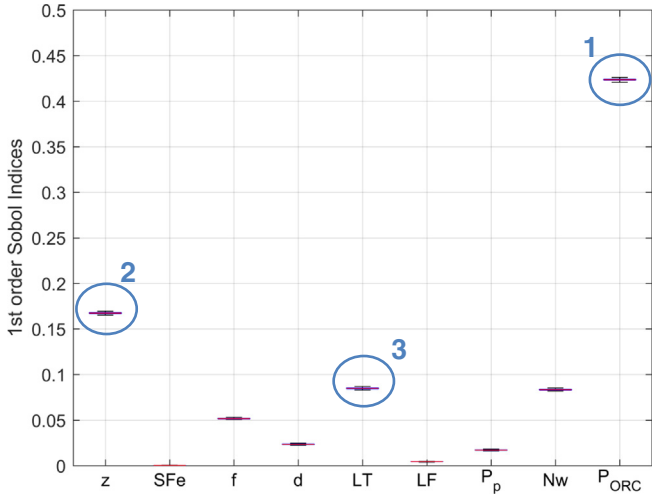


Fig. 3. Results of a GSA under a different hypothesis regarding the description of the variability of $\boldsymbol{LT}$ (for all the other parameters, the baseline distributions are used): the circles indicate the three highest parameters of the ranking.

3.2. Sensitivity of the GSA results: an example based on the case study

Now let's see what happen if we had made different choices during Step 2. Let's imagine that for some reason (e.g. lack of data or choice to have a more conservative approach) we had set, for the variable $\boldsymbol{LT}$, a uniform instead of a Gaussian distribution for the baseline scenario. The results are shown in Fig. 3 (in this simulation, for the other eight variables, the baseline distribution are unchanged i.e. are the ones shown in the third column of Table 1). We observe a doubling of the $S_{T_i}^{First}$ of $\boldsymbol{LT}$, which now appears to be the third more relevant variable (slightly overtaking $\boldsymbol{Nw}$) while it was only the 5th of the ranking in Fig. 2. Therefore, in this case we would rather identify $\boldsymbol{P_{ORC}}$, $\boldsymbol{z}$ and $\boldsymbol{LT}$ (instead of $\boldsymbol{Nw}$) as key parameters together responsible for more than 66% of the variability of the output. Other conclusions may also be drawn when changing the description of other variables (another example is provided in the Supporting information S6, where an alternative description of $\boldsymbol{P_{ORC}}$ is tested).

This simple example shows that the description of the variability of the inputs is essential and has a high influence on the identification of the key parameters. Such description is usually based on quite uncertain assumptions, especially when studying an emerging technology like the EGS. In these cases, the analysis of the robustness of the GSA results is essential to investigate their sensitivity. Such analysis can be executed with the approach showed in the next section, namely by performing the additional Step 3B.

3.3. Proposed GSA approach to address high uncertainty regarding the inputs' description

3.3.1. Steps 1, 2 and 3A

As discussed in Section 2 (Methodology), the proposed strategy to address the lack of confidence in the inputs' description represents an extension of the baseline GSA approach presented in Section 3.1 (see Fig. 1). Steps 1, 2 and 3A are to be performed exactly in the same way as illustrated in Sections 3.1.1, 3.1.2 and 3.1.3, therefore there is no need to repeat their content here. Once completed the baseline GSA (Step 3A), we proceed to the analysis of the influence of the distributions (Step 3B) in order to investigate the robustness of the GSA results.

3.3.2. Step 3B: Analysis of the influence of the inputs' description

3.3.2.1. 3B – (I) Criteria to identify the set of key parameters. The threshold for the aggregated contribution of the key parameters is set to 66% (same as in Section 3.1): this will allow a comparison between the outcome of the baseline and the proposed GSA approaches.

3.3.2.2. 3B – (II) Definition of alternative descriptions of the inputs. For each of the nine inputs, several possible alternative distributions are identified. As discussed in the Methodology (Section 2), the number and the characteristics of the alternatives is set by the modeler, provided that all of them are realistic according to the current knowledge and the goal and scope of the analysis. In this study, we consider in total five types of continuous distribution (including the baseline one) per parameter, except for the number of wells $\boldsymbol{Nw}$ (three discrete distributions) as shown in Table 1 (third to seventh column). It is important to note that, for each single parameter, the boundaries of its variability interval (where to apply the alternative distributions) remain unchanged. In fact, such boundaries represent the minimum and maximum values for the *i-th* parameter according to the goal and scope of this case study.

In the baseline case, most of the parameters are characterized by a uniform distribution: this is essentially due to lack of data (Lacirignola et al., 2014), resulting in a conservative assessment. With the alternative distributions, we account for the realistic possibility that the values at the boundaries (either closer to the minimum or the maximum of the range) are the most probable: this is done by establishing trapezoid

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066

Table 2
Representation of the calculation process for the GSA reiteration. We observe one parameter at a time: each *§i-th* GSA takes into account one of the several k_i-*th* non-baseline distributions for the *i-th* parameter, while the baseline distribution is set for all the others. 100 bootstraps per GSA are performed.

§ GSA	z	SFe	f	d	LF	LT	P_p	Nw	P_{ORC}
§1 (Baseline) ×100 bootstraps	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 2 (×100)	TYPE 2	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 3 (×100)	TYPE 3	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 4 (×100)	TYPE 4	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 5 (×100)	TYPE 5	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 6 (×100)	Baseline	TYPE 2	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 7 (×100)	Baseline	TYPE 3	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 8 (×100)	Baseline	TYPE 4	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
§ 9 (×100)	Baseline	TYPE 5	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline
[…]	[…]	[…]	[…]	[…]	[…]	[…]	[…]	[…]	[…]
§ 32 (×100)	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	TYPE 2
§ 33 (×100)	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	TYPE 3
§ 34 (×100)	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	TYPE 4
§ 35 (×100)	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	Baseline	TYPE 5

distributions (or step functions for ***Nw***). We also explore the possibility of a Gaussian description of the variability of the inputs, considering two possible standard deviations. According to the knowledge of the model, other methods for generating alternative distribution can be also considered (e.g. displacement of the mean or the mode of the distribution, homothetic transformation, etc.)

We define the following settings for our alternative distributions: (i) uniform: all values equiprobable; (ii) truncated Gaussian with standard deviation set to 1/6 of the interval's width; (iii) truncated centered Gaussian with standard deviation set to 1/3 of the interval's width; (iv) trapezoid, with the probability associated to the left boundary of the interval five times higher to the one associated to the right boundary ; (v) trapezoid, with the probability associated to the left boundary of the interval five times lower to the one associated to the right boundary. The parameter ***Nw*** is characterized by a discrete variability range with two values (i.e. two or three wells): the probability distributions are hence step functions allocating to those two values respectively a probability of: 50% and 50% (equiprobability), 70% and 30% or 30% and 70%.

Table 1 also displays the difference between the alternative and the baseline distributions in terms of relative variation of the mean μ and standard deviation σ (see also the Supporting Information S7 for a more detailed discussion of the differences).

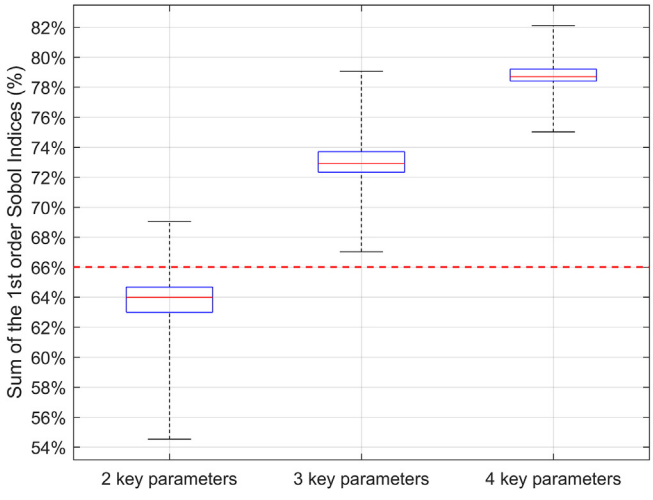


Fig. 4. Sum of the S_i^{First} of the top two, three and four parameters in the ranking. The boxplots are based the results of 3,500 GSAs as defined in Table 2.

3.3.2.3. 3B – (III) Reiteration of the GSA. In this step, several GSA are performed, according to the strategy represented in Table 2. As defined in the previous steps, the *Reference model* is based on 9 parameters ($N = 9$) and 5 probability distributions per parameter are considered except for ***Nw***. Therefore, the calculation strategy is based on 35 different combinations of the distributions (indicated as *§i* in the first column of Table 2). Bootstrapping is also performed: each of the 35 GSA is repeated 100 times. Therefore, according to Eq. 2, 3,500 potentially different rankings are obtained in this study (indeed, each bootstrap corresponds to a new GSA calculation).

3.3.2.4. 3B – (IV) Analysis of the influence of the inputs' description. Fig. 4 shows that, as a result of the several GSAs performed, the sum of the S_i^{First} of the first two parameters of the ranking can range from 0.55 to

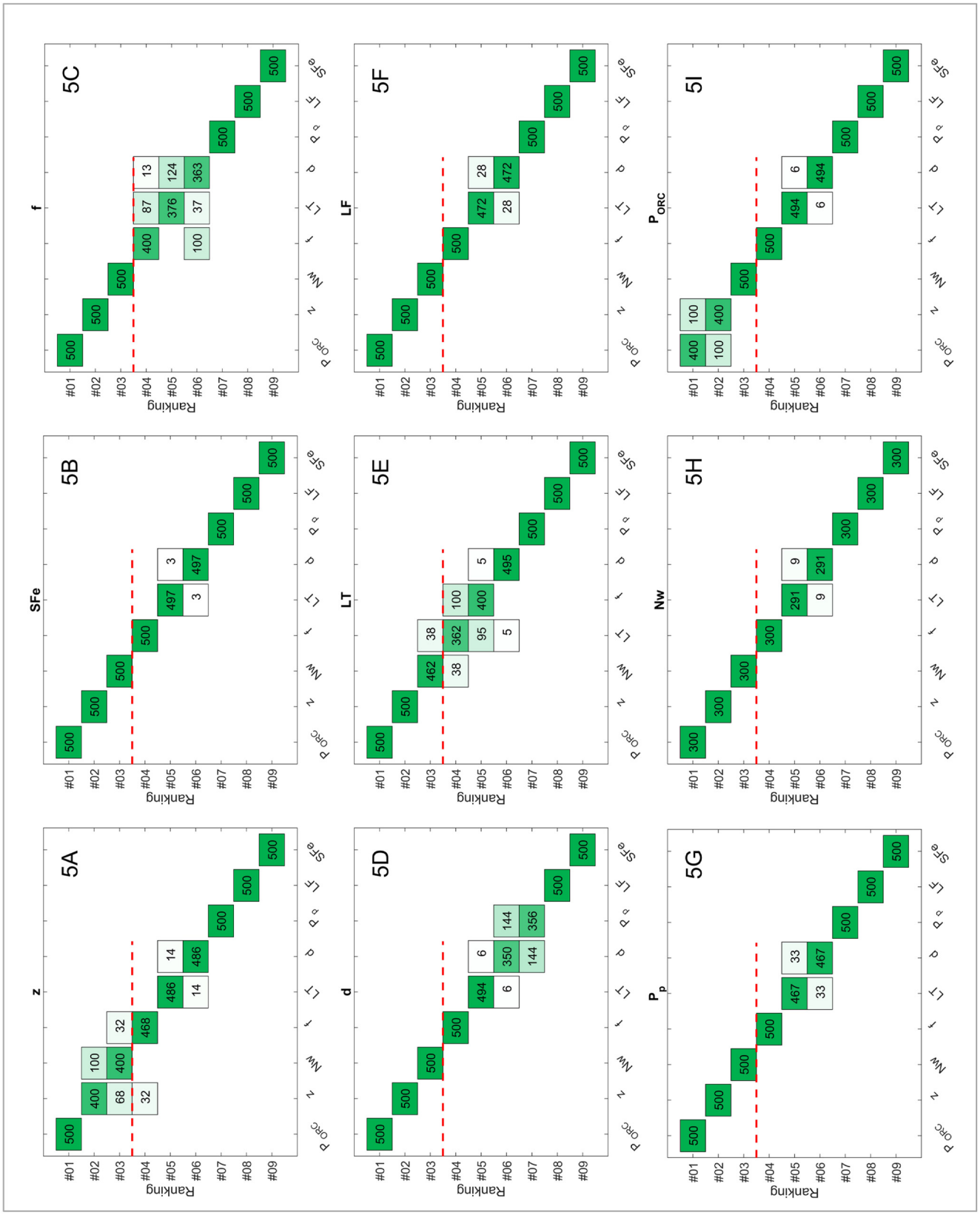


Fig. 5. Sensitivity of the ranking with respect to the description of the input parameters: analysis *one parameter at a time*.

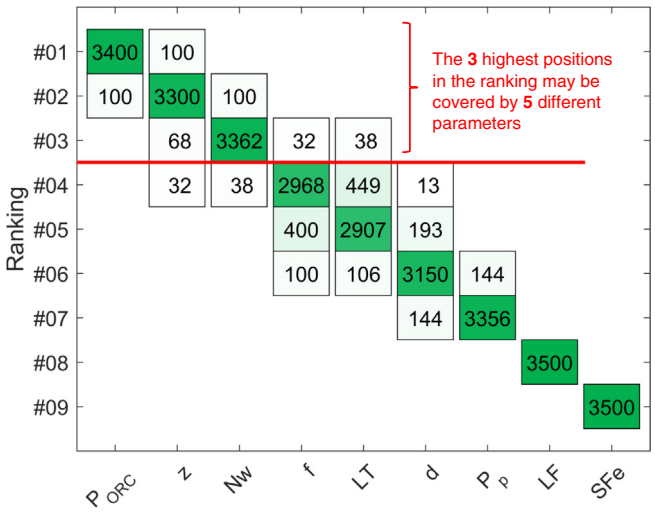


Fig. 6. Sensitivity of the ranking with respect to the description of the input parameters: aggregated results of the analysis *one parameter at a time*.

0.69 (the boxplots are based on the S_i^{First} issued from 3,500 GSAs). This means that, in most cases, two key inputs are responsible for less than 66% of the overall variability of the output (in the remaining cases, their aggregated contribution can achieve 69% at best). Conversely, the sum of the S_i^{First} of the first three parameters of the ranking is always higher than 0.66. We conclude that in order to be sure to cover at least 66% of the variability of the output (as established at the beginning of Step 3B), it is necessary to constitute a set of at least three key parameters.

We now analyze the results of the ranking among the 9 inputs of the model, considering that those in the first three positions should be selected as key parameters. We start by observing the results of the five GSA calculations in which we modify the description of the borehole depth (***z***): these GSAs are noted §1, §2, §3, §4 and §5 in Table 2 and each of them is repeated 100 times (bootstrapping). Therefore, we obtain 500 ranking results. The outcome is presented in Fig. 5: the numbers in the boxes indicate the amount of times the parameters is found in that given ranking position. For instance, the parameter ***z*** (borehole depth, Fig. 5A) results 2nd in the ranking after 400 GSAs, 3rd after 68 GSAs and 4th after the remaining 32 GSAs. Then, we observe the results of the 500 GSAs in which we modify the description of the next input parameter, ***SFe*** (GSAs noted as §1, §6, §7, §8, §9 in Table 2, each of those repeated 100 times, Fig. 5B), and so on. The aggregated result of all 3,500 GSA calculations is presented in Fig. 6.

Based on these figures, the following observations can be formulated:

a) When modifying the description of the variability of the depth (***z***) (Fig. 5A), we see that the flow rate (***f***) and ***z*** itself may pass from being "key" (above the red dotted line) to "non-key" (below the red dotted line). For instance, the flow rate is in the 4th position (non-key) for most of the calculations, but in 32 GSAs it results 3rd in the ranking (key). The same thing is observed when modifying the description of the variability of the lifetime (***LT***) (Fig. 5E). Here the number of wells (***Nw***) and the ***LT*** itself oscillate between key and non-key positions. Therefore, the description of the variability of ***z*** and ***LT*** has a significant influence on the selection of the key parameters. The importance of the description of ***z*** is also shown in the Supporting Information S8, where we analyze the changes observed on the single contribution of each parameter.
b) The description of the variability of $\boldsymbol{P_{ORC}}$ (installed capacity, Fig. 5I) may cause a swap among the first two positions of the ranking. However, this influence is not significant because in these calculations the top three positions are always covered by the same three variables ($\boldsymbol{P_{ORC}}$, ***z*** and ***Nw***)
c) Also the descriptions of ***SFe***, ***f***, ***d***, ***LF***, ***Pp*** and ***Nw*** (Figs. 5B, 5C, 5D, 5F, 5G, 5H) do not have an influence on the selection of the key parameters, because they cause changes in the ranking only between the 4th and the 7th positions.
d) As expected, we observe that the lifetime (***LT***) and the fuel for drilling (***d***) often swap their ranking position (respectively the 5th and 6th), even when no change is made on their distributions. This is due to the fact that their S_i^{First} are very close (cf. Fig. 2) and even small perturbations in the random samples used for the GSA can easily "generate" the swaps. This phenomenon of residual numerical instability, already discussed in Section 3.1.3, is observed in all of the results of Fig. 5.
e) When observing the aggregate results of the 3,500 rankings (Fig. 6), we see that the three highest positions, due to the sensitivity of the parameters to their description, may be covered by five different variables: $\boldsymbol{P_{ORC}}$, ***z***, ***Nw***, ***f*** and ***LT***. Those are all potential key-variables to be selected for the *Reduced model*.

Based on these results, we conclude that the descriptions that have an influence on the identification of the key parameters are the ones of ***z*** and ***LT***.

3.3.2.5. 3B – (V) Consideration of the level of confidence of the inputs' description. Once identified the importance of the description of the variability of ***z*** and ***LT*** according to the scope of the study, we must inquire about their level of confidence. As stated before, only few EGS power plants currently exist in Europe. Thus, the baseline distributions chosen are based on few data and expert knowledge (Lacirignola et al., 2014). Therefore on one hand their level of confidence is low (it may change in the future when new data will be available), but on the other hand it can't be further improved since it reflects the best current knowledge. In conclusion, an alert regarding the description of these two parameters must appear when the conclusions of the analysis are formulated (Step 5, see Section 3.3.4).

3.3.3. Step 4: Overall evaluation

A consistency check was performed all along the application of the methodology and no counterintuitive results were spotted. The aggregated results of Step 3B (Figs. 5 and 6) are coherent with the one of the baseline GSA (Step 3A, Fig. 2). They provide sufficient information to enhance the understanding of the model and to formulate alerts for the final phase of the study, i.e. the identification of the key parameters and the generation the *Reduced model*. Such *Reduced model* will be obtained by fixing the non-key parameters to their median value.

3.3.4. Step 5: Identification of key input parameters of the LCA model

If we were relying only on the baseline GSA (Step 3A), we would have selected without hesitation only three key parameters: $\boldsymbol{P_{ORC}}$, ***z*** and ***Nw*** (the results of the baseline GSA, Section 3.1, shows that $\boldsymbol{P_{ORC}}$, ***z*** and ***Nw*** are together responsible for more than 66% of the variability of the output). However, the results of Step 3B provide useful additional information. Indeed, we found (Fig. 4) that three key parameters are sufficient to cover 66% of the output's variability (as requested at the beginning of Step 3B) no matter the type of distribution used to describe the inputs. However, Fig. 6 shows that the top three positions of the ranking may be covered by five different variables ($\boldsymbol{P_{ORC}}$, ***z***, ***Nw***, ***f*** and ***LT***) depending on the description of the inputs (especially ***z*** and ***LT*** as discussed in Section 3.3.2.4). We are also alerted on the uncertainty carried by the descriptions of ***z*** and ***LT***, since their level of confidence is low and no improvement is to be foreseen with the current knowledge. In conclusion, given that the descriptions of some inputs are influent on the GSA result and are also uncertain, it is preferred to select as key parameters all the five that could possibly cover the three highest positions of the ranking.

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066

3.3.5. Application of the GSA results: generation of the Reduced model

Based on this enhanced knowledge of the *Reference model*, we conclude that the *Reduced model* must be a function of these five variables: ***P*ORC**, ***z*, *Nw*, *f*** and ***LT***. We will let the user of the *Reduced model* define their values, instead of fixing them to their median value beforehand.

Hence, the resulting formula for the estimation of the GHG performances of EGS is given by Eq. 4:

$$\mathrm{GHG}_{\mathrm{EGS\_Reduced}}\left[\frac{\mathrm{gCO_2eq}}{\mathrm{kWh}}\right] = \frac{Nw\cdot\omega_1\cdot z+\omega_2+LT\cdot\omega_3\cdot f+\omega_4\cdot P_{ORC}}{LT\cdot P_{ORC}-f\cdot\omega_5} \quad (4)$$

With:

$$\omega_1 = 120.70\ [\mathrm{gCO_2/m\cdot h/y}];\ \omega_2 = 5,161.87\ [\mathrm{gCO_2/h/y}];$$
$$\omega_3 = 61.82\ [\mathrm{gCO_2eq}\cdot s/\mathrm{kg}\cdot h];\ \omega_4 = 6.42\ [\mathrm{gCO_2eq}/kWh];$$
$$\omega_5 = 6.10\ [kW\cdot s/kg];$$

A comparison of the results of some LCAs of EGS proposed by different authors from literature (Frick et al., 2010; Bauer et al., 2008; Huenges, 2010) with the results obtained through the formula of the *Reduced model* (Eq. 4) is proposed in the Supporting Information S9. Globally, we observe that the results of the *Reduced model* are coherent with those from literature, attesting the robustness of this simplified tool based on these 5 key parameters that allow a rapid calculation of the life cycle impacts without undertaking the long and complex LCA procedure.

4. Discussion

With the case study, we show that the description of the inputs may have a significant influence on the GSA results (for instance, the ranking of the ***LT*** fluctuates over four positions, depending on the description of its variability) and on their exploitation (in our example, we finally decide to include ***f*** and ***LT*** in the set of key variables). However, this aspect is not addressed by the literature nor by the available calculation softwares. Indeed, tools to perform local and global sensitivity analysis are available in both commercial and open-source softwares, however these tools do not consider the question of the "sensitivity of the sensitivity analysis". We propose here a first methodological approach to understand the magnitude of the influence of the inputs' description on the identification of the key parameters and to take appropriate actions if needed.

Of course, when comparing to a baseline GSA approach, our strategy entails an additional task (i.e. Step 3B of the methodology), however the increase in complexity and calculation cost is reasonable, considering the added value it provides. As stated in the Introduction, our strategy is principally aimed for modeler analyzing emerging products or technologies. In fact, when the object of the LCA is a new or innovative item, the entry data of the GSA are usually lacking or highly uncertain and can possibly lead to completely misleading assessments. Obviously, unreliable results are of little interest: our methodology allows increasing their robustness and pointing sources of unreliability that need to be discussed. Moreover, considering that the GSA itself has a non-negligible computational cost, investigating its robustness is essential to consolidate the effort made by the modeler (which otherwise may be vain or questionable).

As discussed in Section 2, the proposed calculation strategy (Table 2) is based on the observation of the ranking's variations that occur when we modify the description of just one parameter at a time (while, for the others, the baseline description is set). A more advanced approach would consider all the possible random combinations of descriptions of the inputs. Referring to our case study this means that, for one single GSA, each of the nine parameters would be randomly characterized by one of the alternative distributions: concretely such strategy would correspond to the realization of a sort of "GSA of the GSA". However, in this paper we keep an OAT approach for the reiteration of the GSAs for several reasons. To start, our first objective is to highlight the interest of investigating the sensitivity of the GSA output and to present, with a concrete application, the importance of the initial hypothesis on the final results. Then, given the seminal character of this publication, we aimed at presenting a first, explorative approach to address the sensitivity problem: indeed the interpretation of the OAT results is easier, given that the changes in the output can be ascribed only to one input at the time. Other than being conceptually challenging, the GSA of the GSA would also entail a dramatic increase of the computational resources needed. In fact, in order to explore the entire space of combinations, the number of ranking to be established would be the following:

$$\textit{Total number of rankings (GSA of the GSA)} = \left[\prod_i^N k_i\right]\cdot B \quad (5)$$

For our case study, this corresponds to about 120 millions of rankings. Even with lower values for k_i or B, the total amount of repeated GSAs is in the order of millions, namely three orders of magnitude higher than the one used in this paper. In conclusion, we believe that our OAT reiteration of the GSA allows drawing useful recommendations while keeping the computational strategy relatively simple. It also constitutes a first brick in the investigation of the sensitivity of the GSA results. The exploration of a "GSA of the GSA" strategy (and the set-up of appropriate indicators to exploit its results) will be object of further studies.

On the other hand, some remarks can be also formulated with the aim of simplifying the calculation strategy of Step 3B. Indeed, the computational cost of the analysis may also be a problem when the number of input parameters N is high (e.g. several dozens). In this case, two solutions are suggested:

- Perform an initial screening to reduce the number of uncertain parameter. This can be done with a qualitative sensitivity method like the one proposed by Morris (1991). Such approach is for instance presented by Andrianandraina et al. (2015) and Wei et al. (2015).
- Exclude from the analysis the parameters with a high level of confidence regarding their description. Indeed, in the methodology that we described, we analyze the sensitivity of all the N parameters (Step 3B – (III)) and we inquire a posteriori about the level of confidence of their description (Step 3B - (V)). This is proposed in order to acquire a global understanding of the model. However, since the highly reliable descriptions will not be challenged further (even if they results being very influent on the ranking), the modeler may exclude them from the analysis performed in Step 3B (in this case, Step 3B – (V) will be moved before Step 3B – (I)).

Another relevant remark is that the conclusions of the analysis of course depend on the choices and hypothesis made by the modeler (e.g. the features of the alternative distributions). However, this is an intrinsic characteristic of any sensitivity analysis. We believe that the proposed methodology, while relying on some reasonable assumptions, still allows enhancing the understanding of the model.

5. Conclusions

Global sensitivity analysis is a powerful tool to study the influence of the different parameters of complex models and to establish a ranking among them, in order to identify the ones that are most influent on the variability of the output. However, the application of GSA has to be handled with care, since its results can be heavily influenced by the initial assumptions: this aspect is particularly critical when studying new products or emerging technologies. With the EGS case study we

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066

provided a clear illustration of how the description of the variability of one input can affect its position in the ranking and its contribution to the output's variance. This research proposes a strategy for a wise use of GSA in the LCA context investigating the stability of the parameters' ranking while considering their level of confidence. We develop a methodology that allows estimating the sensitivity of the GSA results with respect to the description of the variability of the several inputs and to take appropriate action for a relevant identification of key the parameters of the model.

While our methodology implies an augmentation of the computational cost (when compared to the execution of one single GSA), the increase in complexity is still reasonable and the results are sufficiently readable. Some options to simplify or to enhance the calculation strategy are also discussed.

In conclusion, our methodology allows an enhanced understanding of the LCA model and in particular of the relevancy of the inputs' description. This is fundamental to assess the robustness of the GSA results and for the development of future experiments. The strategy illustrated here has not been proposed so far in literature, hence this study provides a relevant contribution to the debate on the application of GSA to LCAs. Moreover, since GSA is also applied in many other different fields of science and engineering, the analysis proposed in this article may also contribute to the enrichment of the sensitivity studies outside the LCA context.

Acknowledgements

We would like to acknowledge Dr. Thierry Ranchin and all the members of the O.I.E. centre of MINES ParisTech for their useful comments and warm support during the realization of this study. We also acknowledge the three anonymous reviewers for their valuable comments that helped us to improve significantly the quality of this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.scitotenv.2016.10.066.

References

Andrianandraina, Ventura, A., Senga Kiessé, T., Cazacliu, B., Idir, R., Van der Werf, H.M.G., 2015. Sensitivity analysis of environmental process modeling in a life cycle context: a case study of hemp crop production, J. Ind. Ecol. 19, 978–993. (doi:10.1111/jiec.12228).

Azadi, P., Brownbridge, G., Mosbach, S., Inderwildib, O., Kraft, M., 2015. Simulation and life cycle assessment of algae gasification process in dual fluidized bed gasifiers. Green Chem. 17:1793–1801. http://dx.doi.org/10.1039/C4GC01698J.

Bauer, C., Dones, R., Heck, T., Hirschberg, S., 2008. Environmental assessment of current and future Swiss electricity supply options. Proceedings of the International Conference on the Physics of Reactors "Nuclear Power: A Sustainable Resource", 14–19 September 2008. Switzerland, Interlaken.

Baujard, C., Genter, A., Graff, J.J., Maurer, V., Dalmais, E., 2015. ECOGI, a New Deep EGS Project in Alsace, Rhine Graben, France. Proceedings World Geothermal Congress 2015, 19–25 April 2015 (Melbourne, Australia).

Bestec, 2012. The Insheim Geothermal Project (http://www.bestec-for-nature.com/j2510m/index.php/en/projects-en/insheim-en, Accessed July 2016).

Bisinella, V., Conradsen, K., Højlund Christensen, T., Fruergaard Astrup, T., 2016. A global approach for sparse representation of uncertainty in life cycle assessments of waste management systems. Int. J. Life Cycle Assess. 21 (3):378–394. http://dx.doi.org/10.1007/s11367-015-1014-4.

Cucurachi, S., Heijungs, R., 2014. Characterisation factors for life cycle impact assessment of sound emissions. Sci. Total Environ. 468-469:280–291. http://dx.doi.org/10.1016/j.scitotenv.2013.07.080.

Cucurachi, S., Borgonovo, E., Heijungs, R., 2016. A protocol for the global sensitivity analysis of impact assessment models in life cycle assessment. Risk Anal. 36:357–377. http://dx.doi.org/10.1111/risa.12443.

Ecoinvent Centre, 2010. Ecoinvent data v2.2. Ecoinvent reports No.1–25 (http://www.ecoinvent.org. Accessed July 2016).

Frick, S., Kaltschmitt, M., Schröder, G., 2010. Life cycle assessment of geothermal binary power plants using enhanced low-temperature reservoirs. Energy 35, 2281–2294.

Genter, A., Baumgartner, J., Cuenot, N., Graff, J.J., Kolbel, T., Sanjuan, B., 2010. The EGS Soultz Case Study: Lessons Learnt after Two Decades of Geothermal Researches. Second European Geothermal Review, Mainz.

Groen, E.A., van Zanten, H.H.E., Heijungs, R., Bokkers, E.A.M., de Boer, I.J.M., 2016. Sensitivity analysis of greenhouse gas emissions from a pork production chain. J. Clean. Prod. 129:202–211. http://dx.doi.org/10.1016/j.jclepro.2016.04.081.

Heath, G., Plevin, R., Kim, H.-C., Sovacool, B., 2010. special session on meta-analysis of energy LCAs. Presented at Life Cycle Assessment X: Bridging Science, Policy and the Public, 2–4 November 2010 (Portland, USA).

Heijungs, R., 1996. Identification of key issues for further investigation in improving the reliability of life-cycle assessments. J. Clean. Prod. 4 (3–4), 159–166.

Hettkamp, T., Baumgärtner, J., Teza, D., Gandy, T., Frost, J., 2011. Production pump technology in the Landau and Insheim geothermal projects. Proceedings of the Soultz Geothermal Conference, 5–6 Oct 2011 (Soultz-sous-Forêts, France).

Huenges, E., 2010. Geothermal Energy Systems: Exploration, Development and Utilization. Wiley-VCH.

Huijbregts, M.A.J., 1998. Part I: a general framework for the analysis of uncertainty and variability in life cycle assessment. Int. J. Life Cycle Assess. 3 (5), 273–280.

Iooss, B., 2011. Revue sur l'analyse de sensibilité globale de modèles numériques. Journal de la Société Française de Statistique 152, 3–25.

IPCC, 2013. The IPCC Fifth Assessment Report - Climate Change 2013: The Physical Science Basis. Working Group I. IPCC Secretariat, Geneva, Switzerland.

ISO 14040, 2006. Environmental management - Life cycle assessment - Principles and framework. International Standard Organization.

ISO 14044, 2006. Environmental management - Life Cycle Assessment - Requirements and guidelines. International Standard Organization.

Jacquemin, L., Pontalier, P.-Y., Sablayrolles, C., 2012. Life cycle assessment (LCA) applied to the process industry: a review. Int. J. Life Cycle Assess. 17:1028–1041. http://dx.doi.org/10.1007/s11367-012-0432-9.

Lacirignola, M., Blanc, I., 2013. Environmental analysis of practical design options for enhanced geothermal systems (EGS) through life-cycle assessment. Renew. Energy 50: 901–914. http://dx.doi.org/10.1016/j.renene.2012.08.005.

Lacirignola, M., Hage Meany, B., Padey, P., Blanc, I., 2014. A simplified model for the estimation of life-cycle greenhouse gas emissions of enhanced geothermal systems. Geothermal Energy 2 (8). http://dx.doi.org/10.1186/s40517-014-0008-y.

Lloyd, S.M., Ries, R., 2007. Characterizing, propagating, and analyzing uncertainty in life-cycle assessment a survey of quantitative approaches. J. Ind. Ecol. 11, 161–179.

Marini, C., Blanc, I., 2014. Towards prospective life cycle assessment: how to identify key parameters inducing most uncertainties in the future? Application to photovoltaic systems installed in Spain. Proceedings of the International Conference on Computational Science and Its Applications, Guimarães, Portugal, June 30 – July 3, 2014, Part III:pp. 691–706 http://dx.doi.org/10.1007/978-3-319-09150-1_51.

Menberg, K., Pfister, S., Blum, P., Bayer, P., 2016. A matter of meters: state of the art in the life cycle assessment of enhanced geothermal systems. Energy Environ. Sci. 9: 2720–2743. http://dx.doi.org/10.1039/C6EE01043A.

Moomaw, W., Burgherr, P., Heath, G., Lenzen, M., Nyboer, J., Verbruggen, A., 2011. Annex II: Methodology, IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2), 161–174.

Padey, P., Girard, R., Le Boulch, D., Banc, I., 2013. From LCAs to simplified models: a generic methodology applied to wind power electricity. Environmental Science & Technology 47, 1231–1238.

Platt, M., Balon, B., Appelhans, K., Bracke, R., 2012. Ökobilanzdaten der geothermischen Strom- und Wärmeerzeugung in Deutschland imBereichTreibhausgase. Proceedings of the German Geothermal Congress, 13–16 November 2012. Karlsruhe, Germany.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis—The Primer (Chichester).

Sobol', I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math. Comput. Simul. 55, 271–280.

Sullivan, J.L., Clark, C., Han, J., Harto, C., Wang, M.Q., 2013. Cumulative energy, emissions and water consumption for Geothermal Power Production. J. Renew. Sustain. Energy 5 (023127. doi: 10.1063/1.4798315).

Van Wees, J.-D., Boxem, T., Angelino, L., Dumas, P., 2013. A prospective study on the geothermal potential in the EU, GEOELEC Deliverable n. 2.5. (http://www.geoelec.eu/concep/library/, Accessed July 2016).

Wei, W., Larrey-Lassalle, P., Faure, T., Dumoulin, N., Roux, P., Mathias, J.D., 2015. How to conduct a proper sensitivity analysis in life cycle assessment: taking into account correlations within LCI data and interactions within the LCA calculation model. Environmental Science & Technology 49 (1):377–385. http://dx.doi.org/10.1021/es502128k.

Wolf, P., Groen, E.A., Berg, W., Prochnow, A., Bokkers, E.A.M., Heijungs, R., de Boer, I.J.M., 2016. Assessing greenhouse gas emissions of milk production: which parameters are essential? Int. J. Life Cycle Assess. http://dx.doi.org/10.1007/s11367-016-1165-y.

Please cite this article as: Lacirignola, M., et al., LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis, Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.10.066