Energy and Buildings 130 (2016) 166–176

Contents lists available at ScienceDirect

Energy and Buildings

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

Choosing the appropriate sensitivity analysis method for building energy model-based investigations

Martin Heine Kristensen*, Steffen Petersen

Department of Engineering, Aarhus University, Inge Lehmanns Gade 10, 8000 Aarhus, Denmark

A R T I C L E I N F O

Article history:
Received 17 June 2016
Received in revised form 27 July 2016
Accepted 12 August 2016
Available online 18 August 2016

Keywords:
Sensitivity analysis
Morris method
Sobol method
Building energy modelling
Input information
Probability density functions

A B S T R A C T

Current literature holds various examples of investigations that make use of a building energy model (BEM) combined with a sensitivity analysis (SA) technique to identify and rank the BEM input parameters that the model output is most sensitive to. However, a sound argumentation that vouches for the reliability, validity and necessary complexity of the chosen SA method for the specific purpose of the BEM-based analysis, is rare. This paper reports on an investigation of how two different levels of a-priori information about input parameters, applied to three different SA methods (Local, Morris and Sobol), influenced the identification and ranking of the input parameters that the annual energy need output of a quasi-steady-state BEM using monthly time steps, and a simple dynamic BEM using hourly time steps, is most sensitive to. It was found that the three SA methods, to a great extent, were able to identify the same cluster of most sensitive input parameters, independent of the level of a-priori input parameter information and BEM. However, the ranking of most sensitive input parameters varied with the applied SA method, BEM, and level of a-priori input parameter information. From a practical point of view, the choice of appropriate SA method is concluded to depend on the purpose of the SA analysis.

1. Introduction

In the context of using building energy modelling (BEM) for performance predictions, it is often valuable to conduct a sensitivity analysis (SA) to explore the model behaviour and to identify which input parameters that drive the majority of the model output variation. Sensitivity analysis is thus a generic description of different techniques for quantification of how variability in model output can be apportioned to the variability and uncertainty of the model input parameters. SA methods are often categorised as either local sensitivity analysis (LSA) or global sensitivity analysis (GSA) [1].

1.1. Local sensitivity analysis

LSA methods rely on an OAT-methodology (one-parameter-at-a-time) where the effect of the variation of a single input parameter to a BEM tool is valuated at discrete points of the input space while all other input parameters are held constant at their reference value. The nature and behaviour of the input parameters are not taken into account, i.e. all values have an equal probability of occurrence without considering the effect of range and shape of the probability density function (PDF). Consequently, LSA methods do not consider any effects from correlated input parameters or any non-linear and non-additive model behaviour.

LSA methods have previously been used for various types of BEM-based analysis [2–5]. Petersen and Svendsen [2] used LSA together with a simple hourly dynamic BEM tool to provide building designers with an overview of the consequences of adjusting a performance-decisive parameter, in terms of energy performance and indoor environment, prior to any actual design decision. In a study of macro-parameters for net zero energy building design by Sun [3], LSA was used to quantify the impact of design parameters on the outcome of BEM tool TRNSYS. Besides this direct application, LSA has also been used to reduce large sets of parameters to smaller manageable sets before applying more complicated and computationally demanding GSA methods [4,5].

1.2. Global sensitivity analysis

GSA is a generic description of methods that evaluates the effect of an input parameter on the output by varying not only the parameter in question, but all other input parameters chosen for analysis as well. GSA uses a probabilistic framework where the effect of range and shape of input PDFs are incorporated. The

* Corresponding author.
E-mail address: mhk@eng.au.dk (M.H. Kristensen).

http://dx.doi.org/10.1016/j.enbuild.2016.08.038
0378-7788/

assignment of an individual PDF for each input parameter is an important and often difficult task; however, in most cases one can narrow down the range of variation and chose an appropriate PDF describing the variation [4]. The probabilistic framework of GSA requires that the model output be evaluated multiple times on randomly selected input samples from the entire input space. This requires a large number of Monte Carlo-based (MC) evaluations of the model. The accuracy of a MC analysis strongly depends on the sampling technique that must ensure good coverage of the input space [6]. Several different sampling techniques are available, but specific GSA methods often require a specific type of sampling. The prevailing sampling techniques includes simple random sampling (e.g. Monte Carlo sampling), quasi-random low-discrepancy sampling (e.g. Sobol' sequences [7,8]) and stratified sampling (e.g. Latin hypercube sampling (LHS) [9]). Quasi-random sampling using Sobol' sequences and LHS are known to outperform crude Monte Carlo random sampling [7,9]. Moreover, Helton et al. [10] argues that LHS is a good choice for computationally demanding models, because its efficient stratification properties allow a broad coverage of the entire probability distribution, at a relatively low computational cost.

A commonly used group of GSA methods for BEM-based analysis are the so-called screening-based methods. These methods seek to identify the least important input parameters that can be fixed at any given value without considerably reducing output variance. In this way, these methods are capable of ranking input variables by their importance in descending order using only a relatively small number of model evaluations [11]. Screening-based methods are efficient for computational heavy models, and/or models with a large number of input parameters, as e.g. building energy models [4]. They are, however, most reliable when the number of important input parameters is low, but as this is often the case, they are widely applicable [11]. For a detailed review on screening-based methods, see Saltelli et al. [12]. The screening-based method by Morris [13] has been widely used for BEM-based analysis, e.g. in an identification of important design parameters in sustainable buildings [4], in an uncertainty study of retrofits for residential buildings [14], in an investigation of important parameters for the performance of different active cooling systems [15], in an evaluation of how geometry effects building energy use [16], and for reducing the number of uncertain parameters in early building design [17,18].

Another group of GSA methods is the so-called variance-based methods. They rely on a decomposition of the model output variance; thus, they are known as ANOVA (analysis of variance) methods. They are regarded as model independent black-box methods making them suitable for complex non-linear and non-additive models. These advantages come, however, at the price of a high computational cost compared to the screening-based methods. A popular variance-based GSA method is the method of Sobol' [19], which, based on a decomposition of the output variance, is capable of estimating sensitivity indices describing the contribution of first-order effects for each input parameter alone, second-order interaction effects between two parameters, third-order effects and so on. The original method has since evolved as described in Borgonovo and Plischke [1] and Saltelli et al. [7]. Another popular variance-based method is the FAST (Fourier Amplitude Sensitivity Test) [20], and the extended FAST [21], which both uses a Fourier decomposition of the model output to estimate first-order and total-order effects. The main difference between Sobol' and FAST lies in the numerical estimation of the multidimensional integrals of the model necessary for the computation of the variances. The method by Sobol' applies Monte Carlo integration loops while FAST applies a sinusoidal function for pattern search. Variance-based methods are not used as widely for BEM-based analysis as screening-based methods; the few examples, to the knowledge of the authors, are the studies by Mechri et al. [22] who used the FAST method to assess important parameters in office energy design, Shen and Tzempelikos [23] who applied a FAST sensitivity analysis of daylighting and energy performance of offices, Mara and Tarantola [6] who used the Sobol' method in an investigation of the thermal behaviour of an experimental test cell, and Spitz et al. [5] who used the Sobol' method in an experimental setup where measured parameter uncertainty was used as input data for the sensitivity analysis.

1.3. Information level in input parameters

The information level of input parameters to a BEM is important to the outcome of the sensitivity analysis; therefore, one has to specify an input parameter space of interest to be examined. The input parameter space for the SA can be defined by a range and a distribution of the likelihood for each value within this range. If the purpose is to explore the effect of equally possible parameter values, for example in a design situation, this space can be assumed uniformly distributed across a defined range meaning that the likelihood of different input parameter values are given equal weight. This uniform distribution is called a *non-informative distribution* as no information can be extracted from it besides the range of variation. One can also explore the output variability of a non-uniform distribution. A non-uniform distribution is called an *informative distribution* because it is based on a-priori information about the variation of the parameters, which could be available from e.g. expert judgements, historical data, or sample measurements. All SA methods are able to handle non-informative distributions, but not all SA methods are able to handle informative distributions.

1.4. Aim of this paper

The previous sections mention several examples where SA has been applied for BEM-based analysis. However, a sound argumentation that vouches for the reliability, validity and necessary complexity of the chosen SA method in the context it is being used, namely for BEM-based analysis, is rare.

The aim of this paper is to provide future BEM-based research with an argumentation for choosing an appropriate SA method, when using SA for identifying and ranking the most important model parameters, given a certain a-priori information level about the parameter input space. Three different but commonly used SA methods were therefore applied to a building test zone, representing an existing residential building stock, using two different BEM models and two different degrees of a-priori input parameter information (a uniform and a non-uniform distribution within the same range). The outputs from these analyses were then compared to identify whether the choice of SA method, BEM, and/or information level in input parameters, affected the identification and ranking of the input parameters most sensitive to the BEM output.

2. Methods

The methodology used to investigate the research statement is illustrated in Fig. 1. First, PDFs were assigned to each input parameter describing the a-priori beliefs of their shape and range of variation (see section 2.2 for details). For input scenario A, the information about shape of the individual input parameters were ignored by only considering the range of the PDFs using the 0.01 and 0.99 quantiles as boundaries. Hereby parameter variation is treated as uniformly distributed between the boundaries, which would correspond to the level of information available in the early building design process where the likelihood of the actual parameter value is uniformly distributed. For input scenario B, all information about the shape of the PDFs was taken into account, which could be the situation in the case of uncertainty analysis where the likelihood of