A comparison of performance of three variance-based sensitivity analysis methods on an urban-scale building energy model.

Pamela Jane Fennell[1], Ivan Korolija[1], Paul Ruyssevelt[1]
[1]Bartlett School of Environment, Energy and Resources, University College London, London, UK

Abstract

The vast number of different inputs required to model a complex urban environment makes it impossible to precisely quantify all inputs and complex energy flows within models must be simplified to achieve tractable solutions. As a result, the outputs of these models inevitably have a significant range of variation. Without understanding these limits of inference resulting policy advice is inherently defective. Uncertainty Analysis (UA) and Sensitivity Analysis (SA) offer essential tools to determine the limits of inference of a model and explore the factors which have the most effect on the model outputs. Despite a well-established body of work applying UA and SA to models of individual buildings, very limited work has been done to apply these tools to urban scale models.

This study presents a systematic comparison of a range of three different variance-based SA methods to a high resolution, dynamic thermal simulation of a mixed-use neighbourhood in North London. Accuracy, processing time and complexity of application of each SA method is evaluated to provide guidance which can inform the application of these methods to other urban and large-scale building energy models.

Key Innovations

There is an urgent need for guidance on the relative performance of SA methods for large scale models with fine resolution and large numbers of uncertain parameters. To address this need, this study evaluates **3** different SA variance based-methods (Sobol', Derivative-based Global Sensitivity Measure and Elementary Effects) applied to a dynamic thermal simulation of a neighbourhood model.

Practical Implications

This study suggests that the elementary effects method (an enhanced version of Morris Method) is the most appropriate Sensitivity Analysis method for Building Stock Energy Models due to its ability to the stepped nature of responses to temperature variations and the discrimination between scales of input parameters.

Introduction

Large scale building energy models capitalise on the increasing accessibility of large-scale urban data sets and allow the rapid evaluation of competing policy options. The urgency of the need to reduce carbon emissions means there is little time left for trial and error and the has importance of targeting policy at the interventions which will have the greatest impact in reducing emissions is critical. Building Stock Energy Models (BSEMs) are therefore a vital tool for urban responses to the climate emergency, enabling the translation of declarations and targets into meaningful action and effective public policy.

In recent years the scale and complexity of these models has progressed rapidly with a trend away from bespoke, standalone models to stock models "designed for wider applicability, allowing core modelling structures to be transferred to other cities or countries by varying model input data" (Langevin et al., 2020). Existing quality assurance approaches are increasingly inadequate when models are used in critical policy decision-making settings and applied to new contexts: model validation is typically applied to the aggregate annual output of the whole model, giving little insight into the ability of the model to capture the changes in emissions resulting from changes in different parts of the city building stock (Cerezo Davila, 2017). Further, such approaches cannot identify the most significant drivers for emissions. The vast number of different inputs required to model a complex urban environment makes it impossible to precisely quantify all inputs and the complex energy flows within models must be simplified to achieve tractable solutions, as a result, the outputs of these models inevitably have a significant range of variation. Without understanding these limits of inference, resulting policy advice is inherently defective and the potential for assumptions suitable for the original context to be erroneously carried through to the new context is high.

Uncertainty Analysis and Sensitivity Analysis

Uncertainty Analysis (UA) and Sensitivity Analysis (SA) offer essential tools to determine the limits of inference of a model and explore the factors which have the most effect on the model outputs. UA does this through propagation of input uncertainties through the model to understand the resulting model output distribution and thus the limits of inference. SA is used to explore the relationship between input and output uncertainties. By understanding which input factors have the greatest impact on outputs, data collection efforts can be focussed where they will have most impact and non-influential factors can be fixed. However, despite a well-established body of work applying UA and SA to models of individual buildings, a review of the literature relating to energy models for larger groups of buildings, undertaken by (Fennell et al., 2019) highlighted very limited application.

Challenges of applying UA and SA to UBEMs

As simulation models are always a simplification of real physical processes, all models inevitably contain uncertainty (Refsgaard & Henriksen, 2004). Uncertainty can be defined as 'any departure from the unachievable ideal of complete deterministic knowledge of the system' (Walker et al., 2003) and as the systems being modelled increase in scale and complexity, the uncertainty in the model will also increase. Simulation models on individual building level as well as at scale involve a broad spectrum of uncertain inputs (Calleja Rodríguez et al., 2013; Eisenhower et al., 2012) and model uncertainties (i.e.,model structural and model technical; (Refsgaard et al., 2007)).

Existing conceptualisations of the application of sensitivity analysis (e.g. Saltelli et al., 2019) view the model as a simple black box, with clearly defined inputs and outputs and a simple workflow which consists of sampling from input distributions to create sets of model inputs, running the model and calculating the appropriate SA indicators from the resulting outputs.

The limited application of SA to UBEMs is at least partly due to the inadequacy of this simple input-output process to describe the complexity of the UBEM work flow:

- Models are typically amalgams

  UBEMs are a class of model defined by their outputs rather than their structure or inputs, consequently a very large variety of approachs exists. In some UBEMs the unit of simulation is the neighbourhood, but in many, individual buildings are simulated and results aggregated with these aggregate results often being used as inputs for larger scale models. Determining the level of model to which SA is to be applied and how these choices impact on final model sensitivities is important.

- Model inputs are ill-defined

  For a city building stock comprising millions of premises it is not possible to specify the parameters of each premises individually and aggregation techniques must be used, typically this means defining clusters or groups with similar characteristics within the stock and assigning identical inputs for some or all aspects of the inputs to each member of the cluster. These clusters are often referred to as archetypes. Booth et al. (2012) introduce the concept of heterogeneity uncertainty which considered the variation between the value for a specific building and that which is assigned for the archetype. For example, while the epistemic uncertainty around the u-value of a particular wall might be small, if the wall is part of archetype specification then the uncertainty which should be modelled is range of u-values for walls in all buildings assigned to that archetype. It should be clear that this uncertainty is much greater since the number of archetypes is smaller than the potential variations in the stock and also that if the choice of archetypes is a subjective matter, there is unlikely to be precise data available for that range

- Model inputs are highly diverse

  Characterising uncertainties in model inputs is challenging even for simple cases such as material properties due to limited data availability but the types of input data used in UBEMs are highly diverse, often including semantic data obtained from public records to determine the use of premises, national survey data to determine occupancy and usage profiles and LiDAR or similar data to determine geometric inputs. Each of these forms of data has different input uncertainties which need to be characterised in different ways. Different inputs may be dealt with differently in the model with some, such as geometric inputs being deeply embedded in the model and challenging to access as a result.

The practical consequence of this picture of complex and varied models which use data in different ways is that the application of UA and SA is necessarily highly tailored to the specific model, generally with the aim of answering a model specific question and that little research exists evaluating the suitability of different SA methods for use with UBEMs.

Aims of this study

The computational challenges of applying SA to UBEMs mean that while significant limitations exist in the characterisation of uncertainties as described above, the scale of the exercise is overwhelming. As a result, it is necessary to proceed incrementally:

1. Determine which SA methods offer the best trade-off between precision and computational burden.
2. Apply the resulting SA methods to determine which input factors can be fixed.
3. Broaden the scope of SA to encompass the missing types of uncertainty in an interative process in which assumptions around factor fixing are sytematically retested.

This study aims to address the first step in this process. Three different SA methods are applied to a high resolution dynamic thermal simulation of a neighbourhood to determine (i) number of model evaluations required to ensure robust results and (ii) the relative performance of the different methods.

The following sections set out the SA methods which are applied and the framework which has been establised for assessing them. The model and the study case are described and results are presented, followed by a discussion of the implications and limiatations of the results and the planned extensions of this work.

Methods

SimStock modelling platform

SimStock is a modelling platform which combines data from multiple sources to automatically generate dynamic building energy simulation models ready to be executed by EnergyPlus, an open-source whole-building energy modelling (BEM) engine.

High Performance Computing (HPC) or cloud computing is used to allow a large number of models to be simulated in parallel. Simulation outputs are collected and post-processed automatically.

SimStock allows the automatic creation of dynamic thermal simulation models of all buildings within an area of analysis; allowing a wide range of scenario analyses to be performed. A key feature of the SimStock modelling platform is its ability to accommodate mixed-use buildings, and combined addresses through the use of the Self-Contained Unit (SCU) as the smallest division of the building stock. Evans et al. (2017b) define a SCU as the smallest unit which into which the stock can be disaggregated without splitting either premises or building polygons. In the simplest case, a SCU might be a single building such as a single-family home but in dense urban centres, much more complex mixes of ownership and use need to be modelled. The SimStock modelling platform automatically generates Energyplus input files from collections of SCUs bounded by roads or other natural breakpoints, these are referred to as built islands. A single thermal zone is created per floor of each SCU. This approach results in a single EnergyPlus model for each built island. Thermal exchanges are not considered between built islands but all buildings within a 25m radius are included as shading elements.

Model case

A mixed commercial and residential neighbourhood in North London was selected as the case for this study. Geometric and activity data is extracted from the 3DStock model (Evans et al., 2017a).

The study model comprises 41 built islands within which 4 use-types are defined: Office, Retail Sales, Restaurant and Dwelling. SCUs are assigned to a use-type based on the dominant use type identified in the UK property tax records. Restricting the model to 4 use types requires some gross simplification, with less common uses such as education and religious facilities being included within the office use type and all food service premises being included within the single Restaurant use-type. Such simplifications were considered necessary to reduce model complexity since the number of use-types drives the number of parameters and consequently the number of model evaluations required.

In total the study area comprises 648 SCUs in 1779 thermal zones (84 Office, 119 Retail sales, 42 Restaurant and 1534 Dwelling). Occupancy and equipment usage schedules are derived from National Calculation Method (NCM) (Department for Communities and Local Government, 2008) profiles. Variant high and low profiles are developed based on the NCM profiles and assigned stochastically to introduce an elementt of diversity in the levels of usage across the stock.

The simulation models are configured to calculate heating and cooling demand based on ideal loads with equipment efficiency calculations added in post-processing if required.

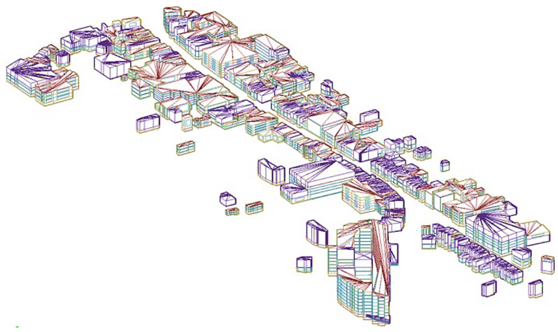

Figure 1: Wireframe model of study area

Figure 2: Satellite view of study area (Google Imagery, 2020)

Input uncertainties

The choice of uncertainties is fundamental to any SA since parameters treated as fixed are, by definition, excluded from the analysis and the choice of which parameters to evaluate must reflect the aims of the study. Although geometric and model uncertainties are of considerable interest, and have received little attention in the literature, they have been deferred to a later stage of this study, to be explored once the most appropriate methods have been selected. The uncertainties selected for evaluation in this study are limited to readily accessible input parameters which typically represent either choices made by occupants or building parameters which might be impacted by retrofits. Parameters which are variable across the stock are treated stochastically as described earlier.

After initial testing with a range of material parameters demonstrated that stable results were not possible within the available computational resources due to the large number of uncertain parameters (103), only a single parameter for each material was retained. In total 50 uncertain parameters were evaluated. Parameters were characterised with triangular distributions to avoid introducing technically infeasible values in the tails of distributions and to ensure that extreme values were not over-emphasised. Space constraints preclude a full listing of input distributions, a summary is provided in table 1 with the full listing available on request from the authors.

Table 1: characterisation of input uncertainties

Parameter type	Instances	Uncertainties considered
Set point temperatures (heating, cooling, natural ventilation	Defined separately for the 4 use-types (12 total)	Heterogeneity across the use-type
Occupant density	Defined separately for the 4 use-types (4 total)	Heterogeneity across the use-type
Ventilation rates	Defined separately for the 4 use-types (4 total)	Heterogeneity across the use-type
Power densities (equipment & lighting)	Defined separately for the 4 use-types (8 total)	Heterogeneity across the use-type
Material conductivity	Defined per material (14 total)	Aleatory uncertainty
Glazing transmittance & emissivity	Defined per material (4 total)	Aleatory uncertainty
Infiltration rates	Defined separately for the 4 use-types (4 total)	Heterogeneity across the use-type

SA methods

SA methods can be either local (focussed at a single point in the input space) or global (assessing sensitivity across the full range of the input space). While local methods, which involve varying a single parameter at a time to assess the effect on the output, are appropriate for linear, additive models, they do not account for the interaction between parameters making them generally unsuitable for non-linear models such as UBEMs. Three methods are evaluated in this study:

- Sobol' analysis (Saltelli et al., 2010) – global method in which all parameters are varied simulataneous and the output variance is decomposed into first and higher-order effects, thus accounting for interactions between parameters. This method is implemented using the SALib library (Herman & Usher, 2017) The Saltelli implementation is similar in form to the EER method and requires (k+2)*N model evaluations, where k is the number of parameters and N the desired number of estimates.

$$\hat{V}_{Ti} = \frac{1}{2N}\sum_{j=1}^{N}\left|f\left(\boldsymbol{x}_j^{(i')}\right) - f(\boldsymbol{x}_j)\right|^2 \quad (1)$$

- Elementary effects (EER) (Campolongo et al., 2007) – a repeated One At a Time design which averages estimates calculated at different points in the input space and thus accounts for parameter interactions.

$$\hat{\mu}_i^* = \frac{1}{N}\sum_{j=1}^{N}\frac{\left|f\left(x_j^{(i')}\right)-f(x_j)\right|}{\left|x_{ji}^{(i')}-x_{ji}\right|} \quad (2)$$

- Derivative based (DGSM) (Becker et al., 2018) – Similar to EER, this method uses a smaller increment.

$$\hat{\upsilon}_i = \frac{1}{N}\sum_{j=1}^{N}\frac{\left|f\left(x_j^{i''}\right)-f(x_j)\right|^2}{\left|x_{ji}^{i''} - x_{ji}\right|} \quad (3)$$

Where:

N is the total number of estimates required

X is a matrix in which each column represents a vector containing the model inputs

x_j is the jth set of model inputs

$x_j^{(i')}$ is a point which differs from x_j only in value of $x_j^{(i)}$

$x_j^{(i'')}$ is a point which differs from x_j only in value of $x_j^{(i)}$ by only a small increment (1x10^{-5} when sampling with respect to the unit hypercube). In equations (2) and (3), the difference between inputs is normalised to take account of the very different scales of different parameters.

Sobol sequences are used for sampling to ensure good coverage of the input space and SA measures were applied by use-type and for the overall model.

Evaluating the different methods and cases

The SA literature relies of the use of test functions for which analytical solutions are availble to evaluate the performance of different SA methods. Since there is no analytical solution for the case considered here, in line with Saltelli et al. (2008), the baseline performance was set by the Sobol' analysis with the highest number of model evaluations. For this study the highest practical number of evaluations was 40, which represents a total of 86,920 individual simulations requiring a total of 6,793 hours of CPU processing time. Two metrics are used to compare SA results for different number of model evaluations and different methods with this baseline:

- Ranking performance - Kendall's rank correlation (τ) (Kendall, 1938) is a non parametric measure of the correspondence between two rankings. Values close to one indicate strong agreement between the two rankings.
- Screening performance - Becker et al. (2018) propose the number of parameters wrongly identified as influential as a fraction of the number of influential parameters as a test of the accuracy of screening. In this study, wrongly excluding influential parameters is considered less desireable than wrongly including non-influential parameters and so the fraction of false negatives is also calculated. The set of influential parameters is defined as the minimum set of parameters which accounts for 95% of the total sensitivity.

The following cases were considered in this study:

- Sobol' method with 40 evaluations (the baseline)
- Sobol' method with 20 and 10 evaluations
- Elementary Effects method with 20 and 10 evaluations
- DGSM with 10 evaluations.

Results

Validation of the baseline result

Bootstrapping with replacement was used to resample the results to construct the 95% confidence interval for sensitivity results. The number of model evaluations would be considered sufficient when all influential parameters are captured at the 95% confidence level. **Error! Reference source not found.** shows the Sobol sensitivity indices with confidence intervals ranked and normalised by the most sensitive parameter for the whole stock for heating and cooling energy respectively. It can be seen that while the indices for cooling energy for the whole stock are relatively stable, those for heating are not. This partial validation of the screening results suggests that more model runs should have been undertaken. However, additional runs were considered to be beyond the limit of availble resources. It can also be seen that the ranking of parameters is affected by the make-up of the stock with parameters relating to the least common use- type, restaurants (4% of thermal zones) having significantly less influence in the whole stock.

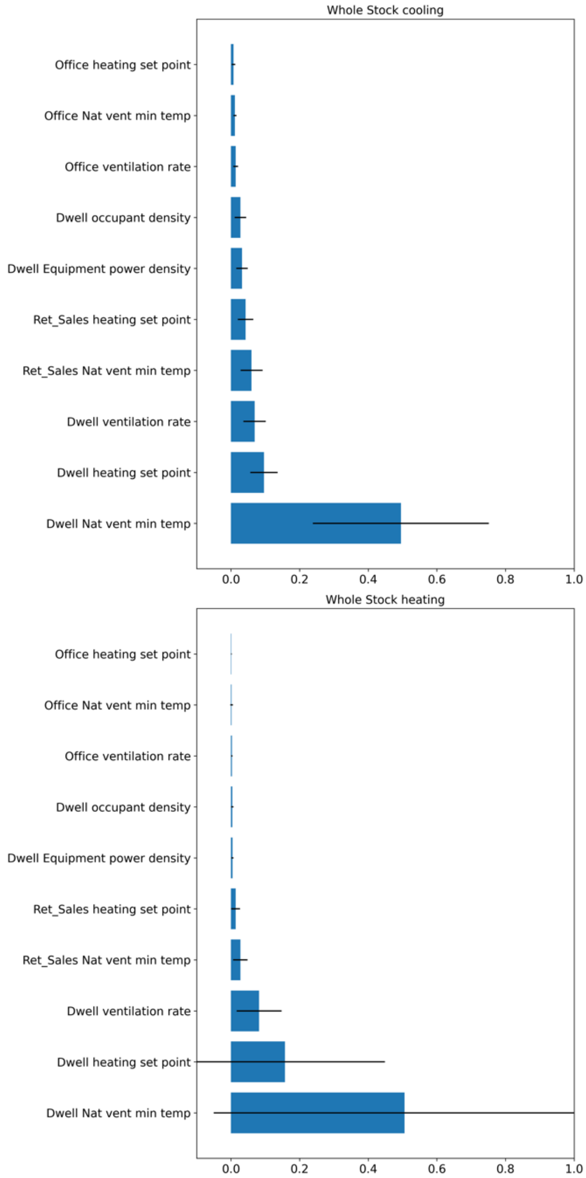

Figure 3: Most influential parameters for heating and cooling energy across the whole stock

Ranking performance

The rank correlation between the baseline results for the whole stock (Sobol' 40 evaluations) and each subsequent set was used to assess the relative performance of the different methods and numbers of evaluations. Figure 4 shows results for both heating and cooling energy with relative computational burden (measured as CPU hours) shown on the x-axis. The results for heating and cooling are identical. Reducing the number of evaluations for Sobol' method has only a small effect on the ranking performance, with a rank correlation of EER results offer marginal time savings but perform less well. Results for DGSM show poor performance.

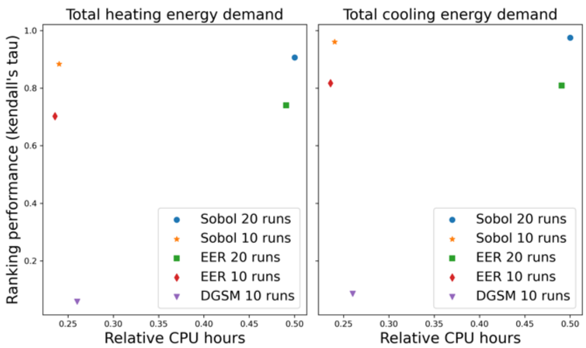

Figure 4: Ranking performance

Screening performance

Screening the input parameters to identify the most influential is a key application of SA. Errors in screening can be either type I: incorrectly identifying a non-influential parameter as influential (false positives) or type II: missing influential parameters (false negatives) The picture of screening performance shown in Figure 5 is similar, to the results for the ranking performance with both EER results and the Sobol' 20 results showing no false negatives. It should be noted that while the relative performance of the methods is the same for both heating and cooling outputs, the fraction of false negatives is lower.

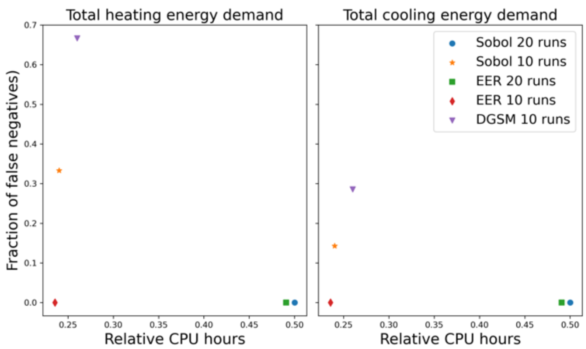

Figure 5: screening performance (fraction of false negatives)

Figure 6 shows the other side of this picture, with EER methods identifying significantly more false positives than Sobol' and DGSM methods.

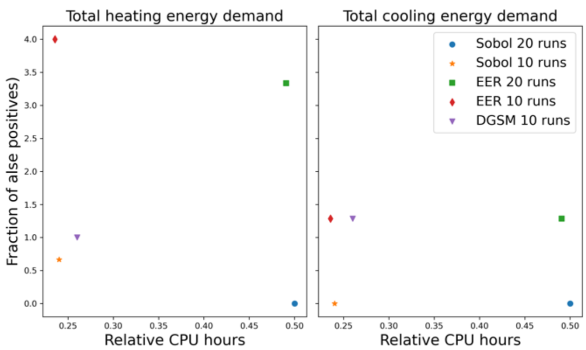

Figure 6: Screening performance (fraction of false positives)

The DGSM method performs least well of all those considered. This seems to be due to the underlying nature of the model - Becker et al. (2018) demonstrated that allow the DGSM method performs well on smooth functions, the small increment results in poor performanc in step-functions. It is likely that the setpoint temperature regime for heating and cooling demand represents a significant enough step to make this method unsuitable for use with building energy simulation.

The results presented in Figure 4, Figure 5 and Figure 6 indicate that the EER method output performs Sobol' at lower numbers of evaluations. Although the methods are very similar in approach, a key difference is the inclusion of the input difference in calculating the μ* index as shown in equation (2). Including this input difference means that parameters are highlighted as influential when a small perturbation in the input parameters results in a relatively large change in the output even if the overall change in the output is not enough to indicate influence on its own.

These results suggest that the widely differing scales of input parameters in UBEMs are more effectively explored using the EER method.

Influential parameters by use-type

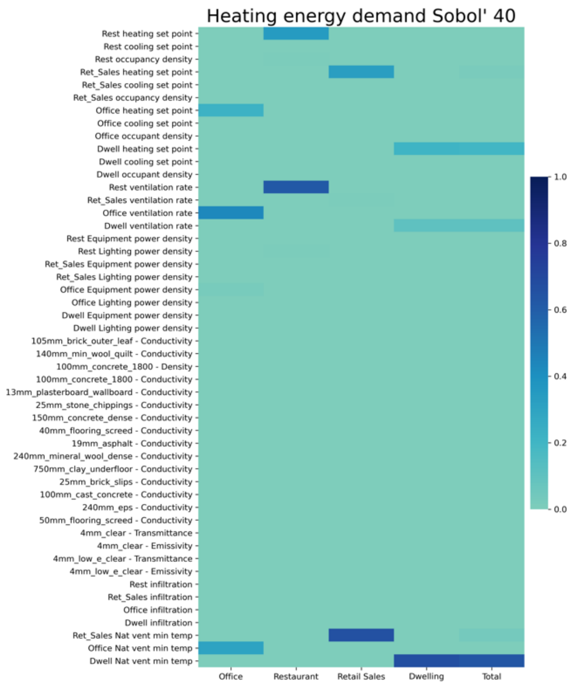

Figure 7: parameter influence by use-type and whole stock

Figure 7 contrasts the influential parameters for the whole stock, shown in the final column in each plot with those for the individual use-types. This highlights the dominance of the Dwelling use-type and the importance of assessing parameter sensitivity at a more granular scale as well as at the whole stock level to ensure that weak signals from smaller segments are not lost.

Care has been taken in this document to avoid focussing on the parameters which have been identified as influential. The aim of this exercise has been to evaluate the performance of different methods rather than to undertake a comprehensive sensitivity analysis for this particular building stock. Nonetheless, some notes about which parameters are identified as influential are important: Firstly, only those parameters which were included in the study can be shown to be influential, this does not mean that other parameters are non-nfluential, only that they were not included. Factors such as glazing ratio and storey height are assumptions within the model but embedded within the model code. This lack of accessibility led to their exclusion from this study although it might be expected that they would have been shown to be influential if included. Secondly, care also needs to be taken with the specification of input parameters, glazing emissivity shows more influence than other building parameters but this is likely to be related to how materials are specified across the stock – the glazing is common to much of the stock while other materials are not included in as many buildings and thus show little influence at the stock level.

Suitability of OAT methods

Fennell et al. (2019) highlighted the predominance of OAT methods within the literature for BSEMs despite the warnings in the broader SA literature that OAT methods are unlikely to be suitable for these projects, this was supported by Cheng and Steemers' (2011) finding that similar models (DECM) were only linear in small ranges about the baseline values. Since EER is a repeated OAT method, the evolution of parameter rankings as successive estimates are added gives an insight into the validity of OAT methods for the model in question. Figure 8 shows how ranks for each parameter vary in each estimate for heating energy. It is clear that the stability of the ranking varies considerably depending on the parameter, with some varying little and others moving considerably. This confirms the caution with which OAT analysis of BSEMs should be treated.

Conclusion and further work

The results of this study highlight the importance of choosing sensitivity measures which are well-suited to the underlying model. For a BSEM based on bottom-up dynamical thermal simulation, two important considerations emerged: (i) the need to incorporate widely differening scales of input parameter and (ii) the stepped nature of responses to changes in temperature. Together these considerations suggest that EER is most appropriate method.

It is also clear that the outputs for which sensitivity indices are calculated need to be carefully considered, in this case, much information is lost if results are considered only at the whole stock level and not at the level of use types.

The performance of the EER method has been shown to be acceptable at fairly low numbers of evaluations (10 evaluations for each index). However, this since each evaluation requires (k+1) simulations, where k is the number of input parameters this still requires a total of 1600 of CPU time meaning that access to high performance computing resources is required.

The aim of this study has been to evaluate the relative performance of a subset of SA methods for a fine spatial grain of UBEM based on dynamic thermal simulation. The methods evaluated here belong to the same class of variance-based methods and further work remains to be done to compare the performance of other classes of SA method including regression methods.

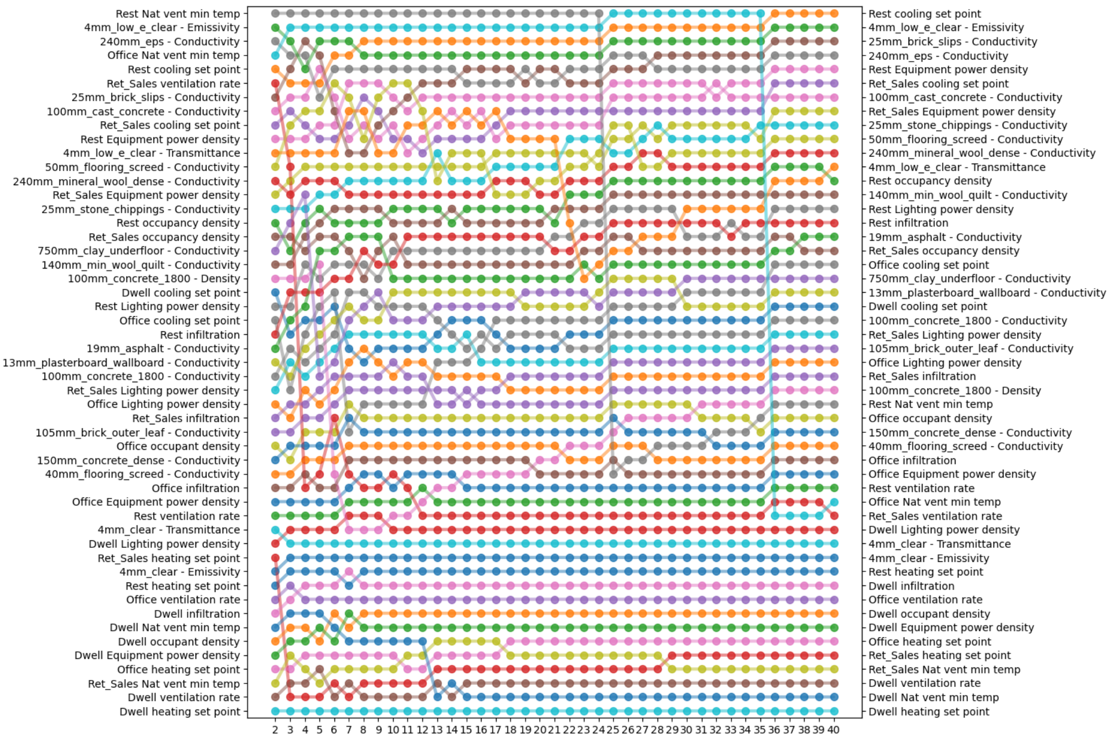

Figure 8: Evolution of parameter ranks over successive iterations

Acknowledgments

The author's work has been supported by the iNUMBER project [grant number EP/R008620/1.

References

Becker, W. E., Tarantola, S., & Deman, G. (2018). Sensitivity analysis approaches to high-dimensional screening problems at low sample size. Journal of Statistical Computation and Simulation, 88(11), 2089–2110.

Booth, A. T., Choudhary, R., & Spiegelhalter, D. J. (2012). Handling uncertainty in housing stock models. Building and Environment, 48, 35–47.

Calleja Rodríguez, G., Carrillo Andrés, A., Domínguez Muñoz, F., Cejudo López, J. M., & Zhang, Y. (2013). Uncertainties and sensitivity analysis in building energy simulation using macroparameters. Energy and Buildings, 67, 79–87.

Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software, 22(10), 1509–1518.

Cerezo Davila, C. (2017). Building Archetype Calibration for Effective Urban Building Energy Modeling. Massachusetts Institute of Technology.

Cheng, V., & Steemers, K. (2011). Modelling domestic energy consumption at district scale: A tool to support national and local energy policies. Environmental Modelling & Software, 26(10), 1186–1198.

Department for Communities and Local Government. (2008). National calculation methodology (NCM) modelling guide (for buildings other than dwellings in England and Wales). Communities and Local Government Publications.

Eisenhower, B., O'Neill, Z., Fonoberov, V. A., & Mezić, I. (2012). Uncertainty and sensitivity decomposition of building energy models. Journal of Building Performance Simulation, 5(3), 171–184.

Evans, S., Liddiard, R., & Steadman, P. (2017a). 3DStock: A new kind of three-dimensional model of the building stock of England and Wales, for use in energy analysis. Environment and Planning B: Urban Analytics and City Science, 44(2), 227–255.

Evans, S., Liddiard, R., & Steadman, P. (2017b). Modelling a whole building stock: Domestic, non-domestic and mixed use. Building Research & Information, 1–17.

Fennell, P. J., Ruyssevelt, P. A., Mata, É., & Jakob, M. (2019). A Review of the Status of Uncertainty and Sensitivity Analysis in Building-stock Energy Models. In V. Corrado & A. Gasparella (Eds.), Proceedings of Building Simulation 2019 (p. 9).

Google Imagery. (2020). Google maps [Online]. https://www.google.co.uk/maps/place/Mornington+Cres,+London/@51.5352529,-0.1411737,720m/data=!3m1!1e3!4m5!3m4!1s0x48761b20619d3675:0x2ae69acf3476c243!8m2!3d51.5333916!4d-0.140483

Herman, J., & Usher, W. (2017). SALib: An open-source Python library for Sensitivity Analysis. Journal of Open Source Software, 2(9), 97.

Kendall, M. G. (1938). A NEW MEASURE OF RANK CORRELATION. 13.

Langevin, J., Reyna, J. L., Ebrahimigharebhaghi, S., Holck Sandberg, N., Fennell, P., Nägeli, C., Laverge, J., Delghust, M., Van Hove, M., Webster, J., Federico, F., Jakob, M., & Camarasa, C. (2020). Developing a common approach for classifying building stock energy models. Renewable and Sustainable Energy Reviews, 133.

Refsgaard, J. C., & Henriksen, H. J. (2004). Modelling guidelines—Terminology and guiding principles. Advances in Water Resources, 27(1), 71–82. https://doi.org/10.1016/j.advwatres.2003.08.006

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process – A framework and guidance. Environmental Modelling & Software, 22(11), 1543–1556.

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

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications, 181(2), 259–270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). Global Sensitivity Analysis. The Primer. John Wiley & Sons Ltd.

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & Krayer von Krauss, M. P. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment, 4(1), 5–17.