Challenges and Lessons Learned in Applying Sensitivity Analysis to Building Stock Energy Models

Citation for the original published paper (version of record):
Fennell, P., Van Hove, M., Weinberg, L. et al (2022). Challenges and Lessons Learned in Applying Sensitivity Analysis to Building Stock Energy Models. Building Simulation Conference Proceedings: 2203-2210. http://dx.doi.org/10.26868/25222708.2021.30960

Challenges and Lessons Learned in Applying Sensitivity Analysis to Building Stock Energy Models

Pamela Jane Fennell[1], Matthias Van Hove[2], Lia Weinberg[3], George J Bennett[1], Marc Delghust[2], Sebastian Forthuber[4], Martin Jakob[3], Érika Mata[5], Claudio Nägeli[6], Janet L Reyna[7], Giacomo Catenazzi[3]
[1]Bartlett School of Environment, Energy and Resources, University College London, London, UK
[2]Ghent University, Ghent, Belgium
[3]TEP Energy GmbH, Zurich, Switzerland
[4]Vienna University of Technology, Vienna, Austria
[5]IVL Swedish Environmental Research Institute, Gothenburg, Sweden
[6]Chalmers University of Technology, Gothenburg, Sweden
[7]National Renewable Energy Laboratory, Golden, CO, USA

Abstract

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. 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 at larger scales. This contribution describes the efforts undertaken by a group of research teams in the context of IEA-EBC Annex 70 working with a diverse set of Building Stock Models (BSMs) to apply global sensitivity analysis methods and compare their results. Since BSMs are a class of model defined by their output and coverage rather than their structure and inputs, they represent a diverse set of modelling approaches. Key challenges for the application of SA are identified and explored, including the influence of model form, input data types and model outputs. This study combines results from 7 different modelling teams, each using different models across a range of urban areas to explore these challenges and begin the process of developing standardised workflows for SA of BSMs.

Key Innovations

A co-ordinated sensitivity analysis exercise was undertaken for 7 diverse BSMs.

Practical Implications

A framework for understanding how input uncertainties are applied in BSMs is established and different SA methods are evaluated, suggesting that Morris Method/EE is a good option for this type of model.

Introduction

As availability of processing power has increased, building stock energy models are emerging as powerful tools in urban planning, offering detailed insights into diagnosing energy consumption across a building stock, allowing energy efficiency interventions to be targeted at areas of greatest need. The impact of potential intervention strategies across the stock, including the application of renewable energy technologies, can be assessed, allowing competing strategies to be ranked. A key benefit of these models is their potential to explore future scenarios, such as changing climate and different development pathways for the stock.

The wide range of potential applications coupled with the complexity of a modern city makes it essential to understand the limitations of the predictive power of such models. The process of model building is, by definition, one of simplification. No model can be a perfect representation of the system it aims to emulate and all models inevitably contain uncertainty (Refsgaard and Henriksen, 2004). Uncertainty can be defined as "any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system(Walker et al., 2003). It is to be expected that as the systems being modelled increase in scale and complexity, the uncertainty in the model will also increase. Consequently, it is inevitable that building stock energy models will contain a considerable number of uncertainties. This should be cause for neither criticism nor alarm, as scale and complexity could also be utilised to hedge uncertainties. Such a discussion is taken in several fields, for instance, Leamer (2010) stated in relation to econometrics "If the range of inferences that can reasonably be supported by the data we have is too wide to point to one and only one decision, we need to admit that the data leave us confused." Saltelli et al. (2019) warn that science in general faces a potential loss of credibility if these weaknesses in methodology are allowed to persist. Therefore, model uncertainties should be explored and reported to guarantee reliance on the model for decision support. The quantification of the level of uncertainty in the model output, is essential. Sensitivity analysis, while less fundamental, offers scope for significant insights by apportioning the uncertainty in the model output to individual input factors.

Definitions

Uncertainty Analysis (UA) characterises the uncertainty in model output while Sensitivity Analysis (SA) explores how that uncertainty can be apportioned to the different sources of uncertainty in the model inputs (Saltelli et al., 2019). Both UA and SA are concerned with understanding the full range of model outputs. In contrast,

https://doi.org/10.26868/25222708.2021.30960

Scenario Analysis focuses on the subset of model outputs which can be attributed to variations in a defined sub-set of model inputs.

Applying uncertainty and sensitivity analysis in large scale models

Uncertainty and sensitivity analysis of models for individual buildings is a well explored topic. In particular, Tian (2013) presents a detailed review of the subject, together with recommendations for appropriate methods for different problem settings. Mavromatidis (2017) presents a comprehensive review of approaches to input uncertainties in large scale models. However, as Naber et al. (2017) note, there has generally been limited application to urban and larger scale models. Traditional sensitivity analysis methods view the model as a clearly bounded "black-box" shown on the left in *Figure 1*. However, for urban scale models the mapping of inputs to outputs is often more complex, with sequential models and aggregation of sub-model results being common features as shown on the right.

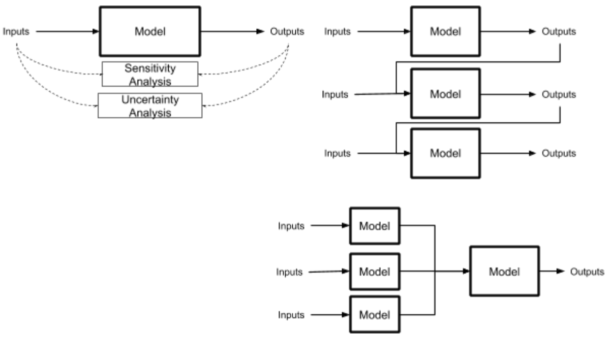

Figure 1: Application of UA and SA to large scale models

Input data is typically gathered from a variety of data sources such as taxation data, LiDAR surveys and utility company records. Since each data source has been compiled for different reasons often using bespoke definitions, considerable data processing is required to align the different elements of input data (See Evans et al., 2017 for a detailed account of the case of London, UK). In this context, the boundary between data processing and modelling is defined subjectively, based on the perspective and objective of the modeller. One result of this complexity is a lack of clarity about how UA and SA can be applied to BSMs.

Aims of this study

To begin the process of addressing the gap in the application of SA to models of whole building stocks, an exercise was established as part of IEA-EBC Annex 70 to undertake a co-ordinated investigation to apply existing global sensitivity analysis methods to a diverse set of BSMs and datasets. Through this process the modelling teams aimed to explore:

1. The challenges of defining input uncertainties for large scale models and collecting appropriate data
2. The appropriateness of different SA techniques in terms of robustness of results and computing burden
3. Key drivers of uncertainty in the models.

Due to limitations of space, this contribution focusses on the high-level comparison of the intra and inter model results. We begin by explaining the practical rationale for this approach before providing brief details of the models included in the study. A new framework is introduced for the definition of uncertainties and the challenges encountered in their definition and application. The sensitivity analysis methods used are then discussed and their performance compared, followed by a brief overview of the comparative results. We conclude with a summary of the lessons learned through the study.

Methods

Since BSMs are a class of model defined by their output and scope rather than their structure and inputs, they represent a diverse set of modelling approaches (Langevin et al., 2020). This diversity, coupled with the diversity in structure and coverage of input data sets in the possession and the computation expense of model evaluation led to the development of a comparison exercise driven by the research aims of each research team.

Following collective review and appraisal of the literature and available toolboxes, each team selected the inputs and analysis methods which were most appropriate to their models and datasets with a common set of model outputs being agreed between the teams. This approach allowed each team to undertake analysis which fitted with their underlying research objectives while allowing key challenges for the application of SA to be identified and explored, including the influence of model form, input data types and model outputs. Results from 7 models covering European and US housing stocks and some non-residential stocks are included in this paper. We begin with a brief description of each of the models and the input data sets used before detailing the sensitivity analysis methods and results and discussing their implications. We conclude with insights for other modelling teams and recommendations for further work in this field.

The models

Each of the models in this study is based on a disaggregated bottom-up building physics-based calculation of building stock energy consumption, three of the models (ECCABS, Invert and BSM) are dynamic stock simulations which are designed to evaluate competing policy options over the long-term. With the exception of SimStock which is based on direct representation of individual building geometry, the models take an archetype-based approach to the stock with a stock of archetypes aggregated up to form the representation of the whole stock. ECCABS, BSM and SimStock cover residential and non-residential building stocks while the other models focus on residential buildings. Models were classified in line with Langevin et al. (2020).

(A) ECCABS (Physics simulation and agent-based market share of technologies - Q4). Previous applications include assessment of the transformation of the residential building stock and non-residential

https://doi.org/10.26868/25222708.2021.30960

building stock of Sweden and various EU-MS (Mata et al., 2013, 2018)

(B) Delghust Model (Physics simulation - Q4). Previous applications include assessment of the effect of energy saving measures in terms of reducing energy consumption in relation to costs in the residential sector, and the development, impact assessment and comparison of different calculation methods and performance criteria in the context of regulatory energy performance calculations and policy (Bracke et al., 2019).

(C) Simstock (Physics simulation - Q4). Previous applications include assessment of retrofit options for a medieval city centre in France (Claude et al., 2019) and schools in the UK (Grassie et al., 2018).

(D) Invert/EE-Lab (Techno-socio-economic simulation - Q4). Previous applications include assessment of the effects of different framework conditions (of economic and regulatory incentives) on the total energy demand, energy carrier mix, CO2 reductions and costs for space heating, cooling and hot water preparations in buildings (Müller, 2015). The model has been applied in more than 40 projects for the EU-27 (+UK and selected neighbouring countries), covering residential and non-residential buildings (Kranzl et al., 2019).

(E) BSM (Physics simulation and agent-based to model building stock dynamics - Q4). Previous applications include state and transformation of the residential building stock of Switzerland (Nägeli et al., 2018, 2020) as well as several European countries (Jakob et al., 2020b, 2020a; Ostermeyer et al., 2019a, 2019b).

(F) ResStock (Physics simulation - Q4). Previous applications include state-level energy efficiency fact sheets (https://resstock.nrel.gov/factsheets/), Single-Family Housing EE Potential (Wilson et al., 2017) Low-income efficiency potential (Wilson et al., 2019), and residential loads for LA100 - Los Angeles 100% Renewable Energy Goals. Other applications: https://resstock.nrel.gov/page/publications.

(G) NHM (Physics simulation - Q4). UK Government Department of Business Energy and Industrial Strategy internal impact assessment model of stock implications.

Input Uncertainties

A variety of different classification schemes for uncertainty (not only in the field of building simulations) exist, for example: (Booth et al., 2013; Coakley et al., 2014; Oberkampf et al., 2000; Walker et al., 2003). The principle aim of these classification schemes has been to map the diverse range of possible sources of uncertainty. These can be broadly grouped into input, model and output uncertainties as shown in Figure 2:

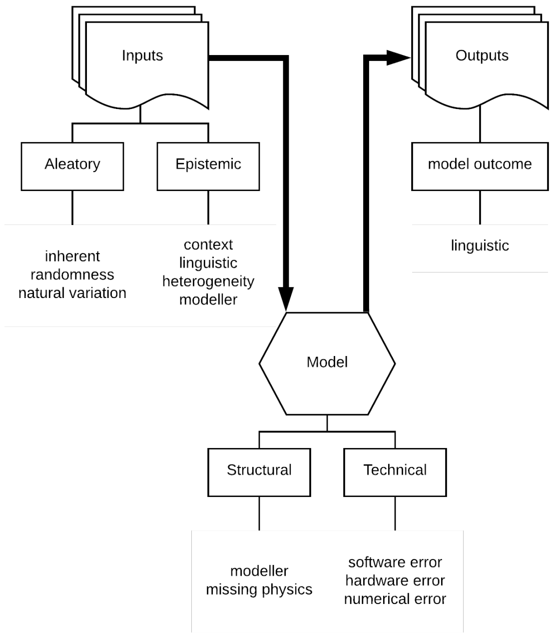

Figure 2: Sources of uncertainty

Fennell et al.'s review (2019) found an overwhelming focus on input uncertainties. All BSMs used in this study use bottom-up building physics-based calculations for energy consumptions. For these types of model consistent definitions of input uncertainties are complicated by the use of subgroups or archetypes as a means of reducing the complexity of model inputs (it should be noted that the probabilistic building characterisation method recently proposed by De Jaeger et al. (2021) allows a better representation of variation within the overall stock). The degree of information which is shared within an archetype sub-group varies between models:

- Complete archetypes - members of an archetype sub-group share all characteristics, e.g. ECCABS, NHM, ResStock
- Partial archetypes - members of an archetype sub-group share some characteristics (e.g. building fabric), but combine this with individual information for other characteristics (e.g. geometric data), e.g. SimStock.
- Distribution archetypes - members of an archetype sub-group are defined based on a distribution of values for that sub-group e.g. Delghust Model, BSM.

In each case, a large number of input parameters need to be defined at different (dis)aggregation levels: ranging from aggregated stock level (e.g., discount rates, climatic data) to the level of the individually calculated buildings (e.g., average insulation levels) or even further (e.g., insulation levels of walls, floors etc.). Those input values will be drawn from different sources but, more importantly from a methodological point of view, they will be assigned across the units of the bottom-up BSM through different approaches. Three main approaches are identified based on how input parameter values are assigned across the units in the BSM, of which the second

https://doi.org/10.26868/25222708.2021.30960

and third will result in a larger variation of input values within stock (and thus of spread of sub-results):

a) a fixed value assignment, with a single value being assigned for that parameter to all members of the group (e.g., a fixed insulation level for all houses of a certain age)
b) distribution based assignment of values across the members of the group (e.g., considering a normal distribution, mean and standard deviation)
c) individual, discrete case-based assignment of different values across the members of the group based on a direct sampling from a (reference) stock dataset (e.g., selecting a reference sample of N cases and their discrete parameter values from a governmental energy performance database, assigning those to N calculated buildings within the BSM).

The assignment of physical parameters to different archetypes or variations of the archetype within the BSM is an obvious illustration of this categorisation of approaches, but this categorisation also applies to other aspects that can be disaggregated within a building stock (e.g., the heating habit or income of households). It is important to note that the three approaches can be used together within a single BSM, having e.g., a fixed insulation level for all houses of a certain age (approach (a)), a distribution-based assignment of air tightness values based on field studies in literature (approach (b)) and a discrete, sampling-based definition of geometrical properties of individual houses taken directly from the governmental cadastre database or GIS (approach (c)).

Approaches (a) and (b) are strongly related. In one way, the deterministic approach (a) could be defined as a specific, simplified case of approach (b): considering a distribution with a variance of zero. The other way around, most often, the distributions in approach (b) could be defined by means of individual input parameters fixed at group level (e.g., a mean and a standard deviation), requiring only a random generator within the modelling engine to generate the matching discrete values for each individual group member. This latter translation of distributions into characterizing, discrete distribution parameters at stock level allows for them to be included in common methods of uncertainty and sensitivity analysis, by varying those discrete values across stocks when generating the required BSMs for applying e.g., the Sobol or Morris Method.

Sensitivity Analysis Methods

As noted earlier, after collective review and debate, each research team selected the methods which they felt best fitted their models. The teams focussed on variance-based GSA methods as these are the most prevalent in the literature with one team applying a machine-learning based approach. In each case the output of interest is calculated at the stock level rather than at the archetype level (although one team used preliminary SA of archetypes to identify inputs for the stock level model which was ultimately the subject of a stock level SA). The methods selected and application details are summarised in the following sections.

One-at-a-Time (OAT) methods or differential sensitivity methods investigate local sensitivity about the base case values. The computational cost of the analysis is low but the fraction of the input space assessed is small. OAT methods were explored for ECCABS, Delghust model and NHM. The other methods used can be classed as global sensitivity analysis measures meaning they explore the full input space of the model with all parameters free to vary simultaneously. In principle, this should mean more reliable results.

Sobol' method is based on decomposing the variance in the model output into the fractions which can be attributed to the different input parameters. First-order effects are those attributable to variance in each input on its own. Higher order effects are attributable to interactions between inputs. Total effects encompass all the interaction terms. Sobol' method was employed on ECCABS, Delghust, SimStock and NHM (Saltelli's implementation in each case).

Morris Method/EE uses a design of experiments approach to maximise the coverage of the input space for as small a computational cost as possible. The approach is effectively an OAT design which is repeated at different points in the input space and averages the results from each point. Campolongo et al. (2007) improved and extended the method, referring to it as Elementary Effects (EE).

Derivative-based Global Sensitivity Measure (DGSM) is the integral of the squared partial derivatives which is very similar in form to Saltelli's implementation of the Sobol total effects measure but is based on an increment between samples of the order of $1x10^{-5}$.

FAST (Fourier Amplitude Sensitivity Test) – uses multiple Fourier series expansions of the output function to decompose the output variance into the conditional variances.

RDP-FAST (Random Balance Design – FAST) uses a more computationally efficient extension to FAST by exploring a subset of the input space and scrambling inputs.

Sobol', Morris/EE. FAST and RBD-FAST all belong to the variance-based class of sensitivity measures. As such they focus on the first moment of the output distribution (i.e., the variance) with the inherent assumption that the output distribution is Gaussian in nature. The indices calculated as a result may not be accurate when model outputs are non-Gaussian (Razavi et al., 2020), for example non-symmetric distributions or showing higher kurtosis.

DMIM (Delta Moment Independent Measure) seeks to address the focus on variance as the sole descriptor of the output distribution by using a moment independent measure.

The final method used in this analysis was the application of *Random Forest* feature importance method to explore pre-simulated data by studying the impact of adding features to a machine learning model. This approach was used on the largest model due to its computational

https://doi.org/10.26868/25222708.2021.30960

efficiency but has the potential for bias due to focus on continuous or high-cardinality categorical variables.

The methods and use cases applied to each model are summarised below:

ECCABS

(A1a) OAT (Firth et al., 2010) energy need module, 1% variation; residential and non-residential stock of France, Germany, Spain and the UK (Mata et al., 2014), (A1b) OAT, energy need module, +/-10% variation; residential stock of France (72 building archetypes), (A2a) Sobol' method (Saltelli et al., 2010) energy need module: 11 variable inputs, (A2b) Sobol' method, delivered energy module: 3 variable inputs.

Delghust Model

(B1) OAT (Firth et al., 2010), (B2) Sobol (Saltelli et al., 2010), (B3) FAST (Saltelli et al., 1999), (B4) RDP-FAST (Tarantola et al., 2006), (B5) DMIM (Borgonovo, 2007). Sample sizes of 20, 100 &1000 buildings for each method.

SimStock

(C1) Sobol' method, (C2) Derivative-based Global Sensitivity Measure (DGSM) (Becker et al., 2018), (C3) Elementary Effects (Morris Method) (Campolongo et al., 2007), 51 variable parameters in an urban district in North London. Each method evaluated at 10, 20 and 50 estimates.

Invert/EE-Lab

Elementary Effects (Morris Method) applied to four country cases (FRA, ESP, CZE, SWE) using two kinds of variation: (D1) a general variation: sample within a -/+ 30% margin and (D2) a variation within a range based on expert guess for each of the parameters. 11 variable technical, economical and behavioural input parameters for a reference scenario simulation of national residential and service sector building stock development from 2012 to 2030, three evaluated output variables (installed power of heat pumps in the year 2030, share of final energy use in heat pumps in 2030, final energy consumption for gas for space heating and hot water in 2030)

BSM

Morris Method applied to two cases using grouped input parameters (E1) Status quo of the residential building stock (2017) and non-residential building stock. (E2) Model dynamics until 2040 of the non-residential stock. 13 groups of parameters for the national residential building stock of Switzerland. 13 groups of parameters for the national non-residential building stock of Switzerland. 13 groups of parameters for the national non-residential building stock of Switzerland. Outputs are aggregate useful and final energy demand as well as total resulting gas emissions.

ResStock

(F1) Morris Method applied to individual archetypes used to determine most influential parameters for electricity and gas demand. 100 buildings in Chicago, Miami and Los Angeles ~30 parameters per building. 116 parameters selected and included in stock model. (F2) Random forest trained for each output quantity of interest and input parameters rerun for each of 116 parameters of interest. Applied to all US homes by region.

NHM

(G1) Sobol' (Saltelli et al., 2010), (G2) Morris Method. 7 Parameters applied in (1) and (2) using 2 methods (a, b) of manipulating underlying model inputs. (G3) OAT (Firth et al., 2010)

Relative computational cost

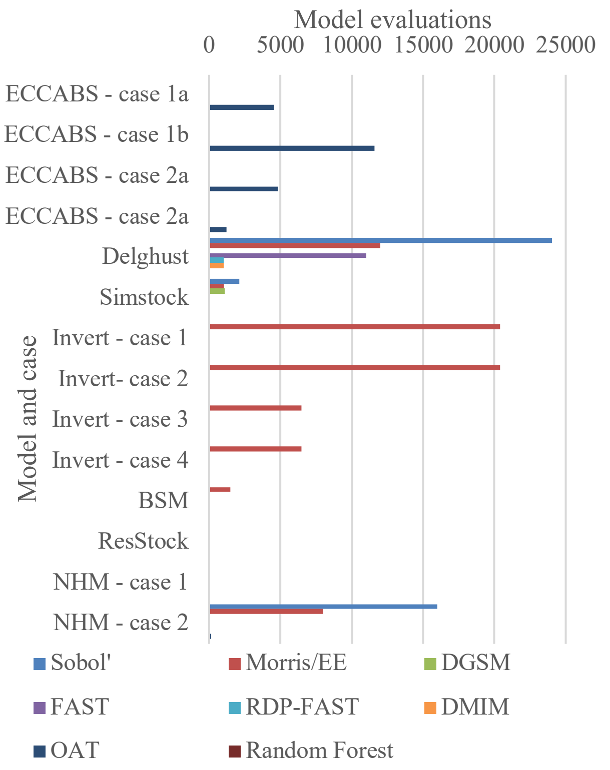

Figure 3: Model evaluations and variable parameters

The relative computational cost of each SA method is assessed by comparing the number of model evaluations required for the analysis in relation to the number required for a single iteration of the BSM (in some instances a single iteration of the BSM comprises multiple aggregated sub-model runs). For all of the methods considered, the number of evaluations is closely linked to the number of variable parameters. This results in a trade-off between complexity of SA (defined by number of variable parameters) and the resources required for the results to reach an appropriate validation metric. This trade-off was most evident for the most complex models, those based on dynamic rather than quasi-steady state energy balances (c.f. SimStock (C) vs the Delghust Model (B)). For model E, this trade-off was managed by grouping parameters. The Sobol' method was the most computationally expensive method employed, followed by Morris/EE and DGSM. FAST and RDP-FAST offered significant savings in computation time.

For model F, sensitivity analysis was done at the individual model level to identify potentially important EnergyPlus inputs that weren't specified by the larger

https://doi.org/10.26868/25222708.2021.30960

stock model. Additionally, model F trained a random forest on a set of model inputs and outputs to identify the relative importance of stock-level distributions (e.g. heating fuel, equipment efficiency, wall material, etc.) to key quantities of interest in the output (e.g. annual total, average daily peak timing and magnitude by season, etc.). These identified variables informed an ongoing calibration process of model F.

Validation of results

Approaches to validation of results varied across the teams, with the three most common approaches being (1) inclusion of dummy variables (BSM, Invert/EELab), (2) use of boot-strapped confidence intervals to establish discrete sets of influential and non-influential parameters (screening threshold 0.05) (SimStock) and (3) convergence of sensitivity indices in line with Sarrazin et al. (2016) (BSM, Delghust Model). In some cases, e.g., SimStock, complete convergence was not possible within the limit of available processing power.

Intra-method comparisons

Intra-method comparisons were undertaken on four models: ECCABS (A), Delghust model (B), SimStock (C) and NHM (G). Results from the Delghust model highlighted the value of nested sampling schemes (e.g., Sobol' sequences) allowing the addition of extra evaluations until convergence is reached when carrying out initial analysis compared with space-filling designs which preclude this (e.g., Latin Hypercube Sampling). Intra-method comparisons on the Delghust model (Sobol', Morris, FAST, RBD-FAST, DMIM) suggested that the parameter ranking of sensitive parameters have similar results for all tested methods. Delghust model results show that DMIM produced slightly different results from the variance-based methods, but needs much fewer model evaluations to reach convergence of indices. Overall DMIM seems to have more potential as a screening method. SimStock results suggested that DGSM performed poorly due to the stepped nature of responses to changes in temperature, while the risk of false negatives in screening (incorrectly assigning influential parameters to the non-influential set) was reduced when using the Morris/EE methods compared with the Sobol' method. Analysis of NHM suggested that OAT analysis had some value for initial screening of large numbers of parameters.

Sensitivity Analysis Results

The primary focus of this exercise has been on the practical challenges of applying SA to BSMs. Early analysis of model structures and input definitions quickly highlighted the impossibility of rigorous inter-model comparisons and drove a focus on intra-model methodological comparisons and practical considerations. As a result, much of the analysis referenced in this study is exploratory in nature with only small numbers of variable inputs considered and wide variation in the robustness of the underlying data sources between models. While ResStock and NHM draw their inputs from large datasets, others have relied more heavily on modeller experience to determine ranges for parameters for which a large evidence base does not currently exist. Nonetheless, it is clear from these preliminary analyses that temperature set-points are a key driver in most models.

In the Delghust model and BSM, the internal temperature or heating set point results to be the most sensitive parameter on the total building energy use and final energy demand, respectively. In the Delghust model, the effect seems much larger compared to the BSM. In the NHM model, on the other hand, the wall U-value has the largest sensitivity on the total energy demand, followed by the set point temperature. The SimStock model investigates the sensitivity of the output parameters heating and cooling of different building types or whole building stocks. Still, the minimum temperature and cooling setpoint show the largest sensitivity on the whole stock heating and cooling, respectively.

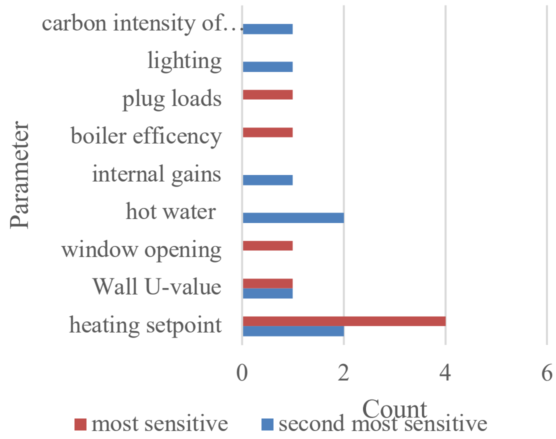

Figure 4: Most influential parameters

Lessons learned

A key objective of the exercise was a practical exploration of the challenges and benefits of undertaking SA of BSMs. For all the participating teams, initial application required detailed consideration of the model structure to determine the most efficient application strategy. In the case of SimStock and the Delghust model, this highlighted changes to the model input structures which would be necessary to allow all parameters to be accessed easily for analysis. In the case of NHM, SA allowed the identification of a number of legacy model input variables which were in fact overwritten by default values during model evaluation.

The importance of the scale of the building stock being assessed depended on the model form, models which were based on complete archetypes were much less sensitive to the scale of the stock than those which were based on partial archetypes in which each unit within the stock was evaluated independently (albeit based on some shared information).

Most of the teams undertook SA using modified versions of existing toolboxes (e.g.,(Herman and Usher, 2017)). A common theme raised was the need for careful scrutiny of

https://doi.org/10.26868/25222708.2021.30960

sampling algorithms to ensure that the most appropriate strategy was being applied.

The methods of UA and SA discussed in this study are exemplary of sampling-based forward uncertainty analysis approaches, as systematised by Tian et al.(2013) Their application raises the confidence that the built stock energy models developed in a bottom-up manner correctly represent the essential aspects of the phenomenon. The synthesis of this knowledge with the empirical findings made with inverse methods, e.g. exemplified by Zhuravchak et al. (2021) may be of significant further interest for large-scale building energy research.

Conclusions

In the context of IEA-EBC annex 70, seven modelling teams undertook a co-ordinated Sensitivity Analysis exercise in order to explore the practical challenges of SA of BSMs. The exercise undertaken was a preliminary exploration, focussed on methodological issues rather than inter-model comparison. The majority of the methods applied were variance-based although a machine learning approach and a moment-based method were also applied. The machine learning based random forest method was the only tractable solution to SA of the US residential building stock.

A key benefit of SA was the systematic interrogation of the model and its results with the teams developing much better understanding of the underlying models and potential model improvements being identified. It should, however, be emphasised that the conclusions made through SA explain the mechanism of the model, not the phenomenon itself.

The methodological comparisons undertaken suggest that Morris Method/EE offers a good trade-off between computational cost and accuracy of results. OAT analysis may have a small role to play in the initial screening of large numbers of parameters.

Computational cost is strongly driven by the number of parameters and the size of the underlying stock when partial archetypes are used. When complete archetypes are used the influence of stock size is reduced.

The variation in models and data used, coupled with the limited numbers of variables included in some studies, make it inappropriate to draw firm conclusions about the most sensitive parameters across the stock. However, the dominance of set-point temperatures and the importance of hot-water operation in models which included it as a variable are notable.

Although much work remains to be done to explore additional forms of uncertainty (e.g., model uncertainty and geometrical uncertainty), the challenges of non-Gaussian outputs and correlated parameters, the benefits of even exploratory analysis were clear and it is hoped that further work will enable robust and efficient workflows to be established for the application of SA to BSMs.

References

Becker, W.E., Tarantola, S., and Deman, G. (2018). Sensitivity analysis approaches to high-dimensional screening problems at low sample size. J. Stat. Comput. Simul. *88*, 2089–2110.

Booth, A.T., Choudhary, R., and Spiegelhalter, D.J. (2013). A hierarchical Bayesian framework for calibrating micro-level models with macro-level data. J. Build. Perform. Simul. *6*, 293–318.

Borgonovo, E. (2007). A new uncertainty importance measure. Reliab. Eng. Syst. Saf. *92*, 771–784.

Bracke, W., Delghust, M., Laverge, J., and Janssens, A. (2019). Building energy performance: sphere area as a fair normalization concept. Build. Res. Inf. *47*, 549–566.

Campolongo, F., Cariboni, J., and Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. Environ. Model. Softw. *22*, 1509–1518.

Claude, S., Ginestet, S., Bonhomme, M., Escadeillas, G., Taylor, J., Marincioni, V., Korolija, I., and Altamirano, H. (2019). Evaluating retrofit options in a historical city center: Relevance of bio-based insulation and the need to consider complex urban form in decision-making. Energy Build. *182*, 196–204.

Coakley, D., Raftery, P., and Keane, M. (2014). A review of methods to match building energy simulation models to measured data. Renew. Sustain. Energy Rev. *37*, 123–141.

De Jaeger, I., Lago, J., and Saelens, D. (2021). A probabilistic building characterization method for district energy simulations. Energy Build. *230*, 110566.

Evans, S., Liddiard, R., and Steadman, P. (2017). Modelling a whole building stock: domestic, non-domestic and mixed use. Build. Res. Inf. 1–17.

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

Firth, S.K., Lomas, K.J., and Wright, A.J. (2010). Targeting household energy-efficiency measures using sensitivity analysis. Build. Res. Inf. *38*, 25–41.

Grassie, D., Korolija, I., Mumovic, D., and Ruyssevelt, P.A. (2018). Feedback and feedforward mechanisms for generating occupant datasets for UK school stock simulation modelling. In Proceedings of Building Simulation and Optimization 2018, (Cambridge, UK: International Building Simulation Association, England), p. 8.

Herman, J., and Usher, W. (2017). SALib: An open-source Python library for Sensitivity Analysis. J. Open Source Softw. *2*, 97.

Jakob, M., Ostermeyer, Y., Nägeli, C., Palacios, A., Camarasa, C., Saraf, S., Catenazzi, G., and Taylor, N. (2020a). Building Market Brief: Slovenia 2020.

Jakob, M., Ostermeyer, Y., Nägeli, C., Palacios, A., Camarasa, C., Saraf, S., Catenazzi, G., Mahmoud, R.,

https://doi.org/10.26868/25222708.2021.30960

Himpe, E., and Laverge, J. (2020b). Building Market Brief: Belgium 2020.

Kranzl, L., Aichinger, E., Büchele, R., Forthuber, S., Hartner, M., Müller, A., and Toleikyte, A. (2019). Are scenarios of energy demand in the building stock in line with Paris targets? Energy Effic. *12*, 225–243.

Langevin, J., Reyna, J.L., Ebrahimigharebhaghi, S., Holck Sandberg, N., Fennell, P., Nägeli, C., Laverge, J., Delghust, M., Van Hove, M., Webster, J., et al. (2020). Developing a common approach for classifying building stock energy models. Renew. Sustain. Energy Rev. *133*.

Leamer, E.E. (2010). Tantalus on the Road to Asymptopia. J. Econ. Perspect. *24*, 31–46.

Mata, É., Sasic Kalagasidis, A., and Johnsson, F. (2013). Energy usage and technical potential for energy saving measures in the Swedish residential building stock. Energy Policy *55*, 404–414.

Mata, É., Sasic Kalagasidis, A., and Johnsson, F. (2014). Building-stock aggregation through archetype buildings: France, Germany, Spain and the UK. Build. Environ. *81*, 270–282.

Mata, É., Kalagasidis, A.S., and Johnsson, F. (2018). Contributions of building retrofitting in five member states to EU targets for energy savings. Renew. Sustain. Energy Rev. *93*, 759–774.

Mavromatidis, G. (2017). Model-based Design of Distributed Urban Energy Systems under Uncertainty. PhD Thesis. ETH Zurich.

Müller, A. (2015). Energy Demand Assessment for Space Conditioning and Domestic Hot Water: A Case Study for the Austrian Building Stock. PhD Thesis. Technische Universität Wien.

Naber, E., Volk, R., and Schultmann, F. (2017). From the Building Level Energy Performance Assessment to the National Level: How are Uncertainties Handled in Building Stock Models. Procedia Eng. *180*, 1443–1452.

Nägeli, C., Camarasa, C., Jakob, M., Catenazzi, G., and Ostermeyer, Y. (2018). Synthetic Building Stocks as a Way to Assess the Energy Demand and Greenhouse Gas Emissions of National Building Stocks. Energy Build.

Nägeli, C., Jakob, M., Catenazzi, G., and Ostermeyer, Y. (2020). Policies to decarbonize the Swiss residential building stock: An agent-based building stock modeling assessment. Energy Policy *146*, 111814.

Oberkampf, W.L., DeLand, S.M., Rutherford, B.M., Diegert, K.V., and Alvin, K.F. (2000). Estimation of Total Uncertainty in Modeling and Simulation (Albuquerque, New Mexico, USA: Sandia National Laboratories).

Ostermeyer, Y., Camarasa, C., Nägeli, C., Saraf, S., Jakob, M., Palacios, A., Visscher, H., Meijer, A., van den Brom, P., Catenazzi, G., et al. (2019a). Building Market Brief The Netherlands 2019.

Ostermeyer, Y., Camarasa, C., Nägeli, C., Saraf, S., Jakob, M., Palacios, A., Hamilton, I., Catenazzi, G., de Baranda, E.S., and Goatman, D. (2019b). Building Market Brief: UK 2019.

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., et al. (2020). The Future of Sensitivity Analysis: An Essential Discipline for Systems Modeling and Policy Support. Environ. Model. Softw. 104954.

Refsgaard, J.C., and Henriksen, H.J. (2004). Modelling guidelines - terminology and guiding principles. Adv. Water Resour. *27*, 71–82.

Saltelli, A., Tarantola, S., and Chan, K.P.-S. (1999). A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. Technometrics *41*, 39–56.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Commun. *181*, 259–270.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., and Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environ. Model. Softw.

Sarrazin, F., Pianosi, F., and Wagener, T. (2016). Global Sensitivity Analysis of environmental models: Convergence and validation. Environ. Model. Softw. *79*, 135–152.

Tarantola, S., Gatelli, D., and Mara, T.A. (2006). Random balance designs for the estimation of first order global sensitivity indices. Reliab. Eng. Syst. Saf. *91*, 717–727.

Tian, W. (2013). A review of sensitivity analysis methods in building energy analysis. Renew. Sustain. Energy Rev. *20*, 411–419.

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

Wilson, E.J., Christensen, C.B., Horowitz, S.G., Robertson, J.J., and Maguire, J.B. (2017). Energy Efficiency Potential in the U.S. Single-Family Housing Stock.

Wilson, E.J.H., Harris, C.B., Robertson, J.J., and Agan, J. (2019). Evaluating energy efficiency potential in low-income households: A flexible and granular approach. Energy Policy *129*, 710–737.

Zhuravchak, R., Pedrero, R.A., del Granado, P.C., Nord, N., and Brattebø, H. (2021). Top-down spatially-explicit probabilistic estimation of building energy performance at a scale. Energy Build. 110786.

https://doi.org/10.26868/25222708.2021.30960