

A Review of the Status of Uncertainty and Sensitivity Analysis in Building-stock Energy Models

Pamela J Fennell¹, Paul A Ruyssevelt, Érika Mata², Martin Jakob³

¹Energy Institute, University College London, London, UK

²IVL Swedish Environmental Research Institute, Stockholm, Sweden

³TEP Energy GmbH, Zürich, Switzerland

Abstract

Building-Stock Energy Models (BSEMs) are emerging as a powerful tool for cities and regions seeking to reduce greenhouse gas emissions and mitigate the effects of changing climates for their populations. The potential influence of such model results coupled with the scale and complexity of the environments they aim to represent means it is essential to understand their limitations. This study undertakes a systematic review of the literature relating to such models and finds that in only a very small proportion of studies are model uncertainties even considered. This fundamental flaw is due to the computational demands of exploring the output space of such complex models. A more detailed assessment was then undertaken of the identified studies in which uncertainty analysis (UA) and sensitivity analysis (SA) had been applied to BSEMs. The adequacy of the applied methods is discussed, and recommendations proposed for the application of best practice techniques based on the underlying form of the model.

Introduction

As cities world-wide seek to reduce emissions of greenhouse gases to meet international obligations and to mitigate the effects of changing climates to address local concerns about health and well-being, tools are needed to compare the effects of different interventions and assess the impacts of different development scenarios. BSEMs, incorporate representations of large numbers of individual buildings in order to create a model of a neighbourhood, entire city (Kavgic et al., 2010) or region. In contrast to data-driven approaches which use statistical and machine learning techniques to relate overall energy consumption to characteristics of the building stock, this study focusses on BSEMs which are based on the aggregation of results for individual buildings, “bottom-up, building physics based models” (Swan & Ugursal, 2009).

As availability of processing power has increased, BSEMs 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 is the 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 constructing models 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 & 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 BSEMs will contain a considerable number of uncertainties and this should be cause for neither criticism nor alarm. While some applications of BSEMs, such as in early design, actively seek a range of possible options, it is common to see output in BSEMs for existing cities expressed as a single value (Cerezo Davila, 2017). 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.

Sources of uncertainty

Modelling energy consumption (and related greenhouse gas emissions) at scale inherently involves a wide range of uncertain inputs (Eisenhower, O’Neill, Fonoberov, & Mezić, 2012) and it is helpful to begin by considering the types of uncertainty which can be found in these models.

A number of different classification schemes have been proposed for the different sources of uncertainty (Stirling, 2001; Walker et al., 2003). For the purposes of this study, the categorisation proposed by Booth et al. (2012) is used since this was developed expressly for the case of BSEMs and makes explicit the concept of heterogeneity:

1. Aleatory uncertainty – the random variation inherent in a parameter, for example in the thermal conductance of an individual building material.
2. Epistemic uncertainty – uncertainty arising from a lack of knowledge, for example, the year of construction of a property.

3. Heterogeneity – building stock models are frequently simplified by the use of archetypes, in which a single set of parameter values is assigned to all buildings within a particular class. Heterogeneity represents the variation between the buildings assigned to a particular archetype.¹
4. Model uncertainty – uncertainty about how to model the true processes, for example, the choice of sub-model to calculate long-wave radiative transfer between buildings, or whether to exclude it entirely from calculations.

Existing discussions of uncertainty and sensitivity analysis in building-stock energy 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. Naber et al. (2017) extend the review to encompass models at an urban or national scale, and note a limited uptake of UA and SA methods. However, Naber et al. do not perform a systematic review nor offer any detailed analysis of the methods used and their appropriateness. A more detailed review is presented by Lim and Zhai (2017) who discuss the approaches to UA and SA in the reviewed models. However, rather than a systematic review, Lim and Zhai aimed to synthesise examples of UA and SA practice. As a result, it is clear that there is a need for a systematic review of the application of UA and SA to BSEMs, together with a need to determine how the form of the model in question might affect the application of UA and SA techniques.

Aim of this paper

The aim of our present paper then, is to systematically review the existing literature on the application of UA and SA to BSEMs, consider the forms of model encountered and the UA and SA methods used. Based on this, recommendations are then provided for the application of UA and SA to BSEMs.

Method

A review was undertaken of journal and conference publications together with book chapters to explore UA and SA approaches employed to date. Given the relatively recent emergence of BSEMs, the period of review was limited to publications from 2010 onwards. A title, keyword and abstract search was undertaken in the Scopus database for the following terms: (energy AND building AND model) AND (uncertainty OR sensitivity OR probabilistic OR stochastic) AND (city OR building stock). A variety of search terms were assessed and results screened to ensure that known publications on the subject were identified prior to selection of the final search string.

¹ Following the definition by Swan and Ugursal (2009), in this article, *representative buildings* is used to refer to a specific set of parameters which applied to each building in a particular class (typically determined by age, function

A total of 570 publications were identified, since this exceeded the resources available for review, the top 100 publications by citation count were selected for review. It is acknowledged that this selection procedure skews the choice of papers towards older publications (Davis & Cochran, 2015). Alternative approaches to filtering the search results were possible, such as ranking by occurrence of search terms. However, we follow the reasoning advanced by Saltelli et al. (2019) and note that the most highly cited papers are likely to be used as a benchmark to guide methodology and thus are an important indicator of the state of practice in a given field.

The selected abstracts were then reviewed manually to determine whether they met the criteria for inclusion in this study:

- A model based on aggregation of the simulated energy consumption of individual buildings in the stock, i.e. bottom-up, physics-based simulation.
- Aggregation at a neighbourhood or larger scale. The relative lack of research at scale led to a fairly low threshold being set for this criterion: 25 premises. This ensured that methods which could be demonstrated to work at a scale beyond that of the individual building would still be included.

This initial screening led to 39 publications being retained for detailed screening. The challenge of inferring methods from short abstract meant that where there was a lack of clarity about whether or not a publication met the criteria, it was retained. The full content of each of the 39 publications was then re-evaluated against the inclusion criteria. This resulted in the exclusion of a further 15 publications, either because the model described was not a bottom-up, physics-based model, or because, although the abstract discussed application at scale, the case detailed in the publication referred only to a single building. The high exclusion rate at this stage of the screening process highlights a tendency for abstracts to discuss the potential extension of methods and models to a larger scale than that which is actually addressed in the research.

The application of UA and SA in the remaining 24 publications was then reviewed. This resulted in discarding a further 10 publications which contained only narrative references to the search terms between varying inputs and outputs, or did not detail any analysis of their relationships.

Evidence from the selected literature

The details of the reviewed publications are summarised in

Table 1. We observe the following types of **uncertainty analyses**:

UA1 Calculated confidence intervals (CI)

etc.). Where the representative buildings are simulated directly and results scaled for other buildings in the class, these are referred to as sample buildings

Found in 2 refs. This approach to propagating uncertainty benefits from simplicity and requires only a small number of transformations. However, as employed in Firth et al. (2010), it entails the assumption that all input parameters are normally distributed and independent. For systems with physical limits this will not be true. In addition, it is assumed that there are no interactive effects between parameters which might result in more extreme results. Finally, for practical reasons, this approach is only possible for simple, quasi-steady state models.

UA2 Stochastic uncertainty propagation

Found in 2 refs. This approach typically applies a quasi-random sampling strategy. Ascione et al. (2017) employ a Latin Hypercube Sampling strategy (LHS) which allows probability distributions to be defined independently for each of the input parameters, avoiding the necessity to assume a normal distribution for all parameters. These methods can be extended to consider the correlation structure of the input parameters.

UA3 Bayesian calibration

Found in 2 refs. Although Bayesian calibration is not an uncertainty analysis method, it is included here since it involves the refinement of a quantified uncertainty distribution by calibration against measured data. It is important to note that with any calibration process, the uncertainty in any parameters which were deemed to be fixed will be incorporated within the uncertainty of the variable parameters. Therefore, the resulting posterior distributions for input parameters must be carefully interpreted.

These three approaches are presented in order of increasing robustness, stochastic uncertainty propagation is preferred to calculating confidence intervals since it does not rely on the assumption of normally distributed variables. Where measured data is available for comparison with the model output, Bayesian calibration offers a more precise definition of the output range of the model.

As for the **sensitivity analysis** methods used, we observe:

SA1 Scenario analysis

Found in 3 refs. A scenario is “based on a coherent and internally consistent set of assumptions about key relationships and driving forces (technology changes, prices, etc.)” and is used to explore the output space of a given set of input parameters (Walker et al. (2003). Since

the change in outputs cannot be attributed to an individual input parameter without further analysis this is not strictly a sensitivity analysis method but is included here since it is often conflated with sensitivity analysis. While scenario analyses can be useful for demonstrating the potential impact of a particular combination of input parameters, they do not systematically explore the input space of a model. As a result, the model output for a specific scenario cannot be used to infer potential outputs due to alternative combinations of input parameters. A true sensitivity analysis maps the changes in output values to changes in input parameters allowing inferences to be made about the relationships between the two. Quantification of uncertainty can provide valuable information on the range of outcomes in any given scenario.

SA2 One-at-a-Time (OAT)

Found in 6 refs. The effect of variable input parameters on the output is assessed by varying each input parameter in turn. All other input parameters are held fixed while the parameter in question is varied. One-at-a Time sensitivity analysis belongs to the class of local sensitivity analysis. By varying model inputs sequentially, the relationship between model input and output is explored at certain key points in the input space. However, the principal failure of OAT analyses is the lack of coverage of the whole input space as demonstrated by Saltelli and Annoni (2010). The limitations of the OAT analysis proposed by Firth et al. (2010) were explored by Cheng and Steemers (2011) who demonstrated that the results were only valid locally. As a result, the sensitivity analysis was of limited predictive value since it did not apply to the full range of likely or valid values for each input parameter.

SA3 Global sensitivity analysis

Found in 2 refs. These techniques aim to provide a full coverage of the model input space. However, this comes at considerable computational cost, since models must be evaluated at many points within the input space in order to produce results. For models covering large building stocks, this can be prohibitive, although approaches based on design of experiments can be used to reduce the burden as discussed later in this article. A full discussion of global sensitivity analysis methods is beyond the scope of this review, but Saltelli and Annoni (2010) provide a useful introduction to the topic.

Table 1: Applications of UA and SA in building-stock models (ordered by number of citations)

Authors	UA Methods	SA Methods	Form of model	Uncertain parameters	Size of stock	Representative buildings
Firth et al. (2010)	Calculated CI	OAT	Quasi-steady state	27	England	47
Baetens et al. (2012)	uncertainty propagation (unclear stochasticity)	n/a	Scaled dynamic	not reported	33	4
Cheng and Steemers (2011)	Calculated CI	OAT	Quasi-steady state	27	16,194	50
Wang and Chen (2014)	n/a	scenario analysis	Scaled dynamic	1	USA	9

Authors	UA Methods	SA Methods	Form of model	Uncertain parameters	Size of stock	Representative buildings
Papaefthymiou et al. (2012)	n/a	scenario analysis	Scaled dynamic	1	Germany	7
Tian and Choudhary (2012)	Bayesian calibration	Global SA	Surrogate model	16	379	1
Nouvel et al. (2013)	n/a	OAT	Quasi-steady state	2	Neighbourhood	6
Collins et al. (2010)	n/a	OAT	Scaled dynamic	4	UK	6
Booth et al. (2013)	Bayesian calibration	n/a	Surrogate model	6	City	21
Ascione et al. (2017)	Stochastic uncertainty propagation	Global SA	Surrogate model	48	8,800	1
Fonseca et al. (2016)	n/a	scenario analysis	unclear	16	Neighbourhood	?
Turhan et al. (2014)	n/a	OAT	Surrogate model	5	2,136	n/a
Kazas et al. (2017)	Stochastic uncertainty propagation	n/a	Scaled dynamic	1	144	1
Nouvel et al. (2017)	n/a	OAT graphical	Quasi-steady state	4	150	150

The results presented in Table 1 demonstrate the low penetration of UA and SA methods in the selected literature. Of the 24 publications identified as relating to the application of a bottom-up simulation of energy consumption at scale and purporting to address UA and SA, only 7 undertook some form of uncertainty analysis. This means that even for the 24 papers which a full text review had confirmed met the requirements of both model form and coverage and which included the search terms “uncertainty”, “sensitivity” or “stochastic”, over two-thirds of studies do not include even a basic assessment of the bounds of inference of the model presented. Our findings are therefore in agreement with the more general results of the systematic review undertaken by Saltelli et al. (2019) which found three quarters of publications selected through a search focussed on sensitivity analysis did not contain an uncertainty analysis.

Our corresponding findings for sensitivity analysis are similar to those by Shin et al. (2013) in their review of the application of sensitivity analysis in hydrological modelling who found 36% of the total papers mentioned some form of SA (46% in this review) but only 19% of those discussing SA actually undertook some form of analysis and only half used robust methods (i.e Global SA for a non-linear model). In our review 33% of the papers which met the modelling criteria applied some form of SA of which only a quarter were robust techniques.

Only two of the reviewed publications (Ascione et al., 2017; Tian & Choudhary, 2012) applied global sensitivity analysis methods. Both use regression techniques to map changes in input parameters to changes in output values. In both cases, SA is undertaken at a sample building level rather than at a stock level. The results of SA are then used to generate a surrogate model which is used to generate the building stock. Tian and Choudhary (2012) use the SA results directly to create a regression model while Ascione et al. (2017) use the SA to exclude non-

influential parameters and then build a neural network model based on the reduced parameter set. Tian and Choudhary create a sample of the whole stock by sampling from the results of the calibrated surrogate model and scaling for each individual instance. This process is repeated multiple times to generate a distribution for the whole stock. The scaling process employed by Ascione et al. is not detailed in the reviewed publication.

Types of uncertainty considered

Figure 1 shows the categories of uncertain parameter used in the reviewed publications and which were most common. The approaches taken to determining uncertain parameters and the overall form and purpose of model seem to have a strong influence:

- Location is more likely to be a variable parameter in studies undertaken for national stocks (Cheng & Steemers, 2011; Firth et al., 2010). This is however not related to sensitivity or uncertainty analysis in itself but used as means of upscaling, in which a very limited amount of representative buildings are modelled in the different climate zones of the geographic region studied.
- Inputs which some studies treat as a categorical variable defining an archetype are treated as an uncertain parameter in others (e.g. Collins et al., 2010 take "dwelling type" as an uncertain parameter).
- The most common category of uncertain parameter considered is building fabric. This may owe more to the ease of modelling and conceptual simplicity than likelihood of being influential.

A clear gap concerns Booth et al.'s (2012) fourth category of uncertainty as discussed earlier: the uncertainty in how to model the true processes. This category of uncertainty was not addressed in any of the reviewed publications.

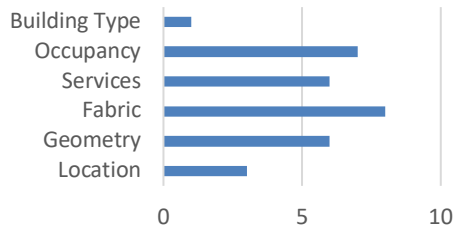


Figure 1: Occurrences of categories of uncertain parameter (count by use of category in each publication)

Potential applicability of UA and SA for other building-stock modelling approaches

The stochastic uncertainty propagation and global sensitivity analysis techniques which represent best practice in UA and SA methods are well established and extensively applied at the individual building scale (e.g. De Wit, 1997; Eisenhower et al., 2012). However, as demonstrated by this review, these techniques are infrequently applied in larger scale models. The challenges of dimensionality which underlie scale simulation are exacerbated by techniques which increase the number of simulations required by one or more orders of magnitude.

The approach to scaling from individual buildings to multiple buildings is dependent on the form of the model. Models were categorised following the schema used by Gaetani et al. (2016). While robust uncertainty propagation techniques were not applied in all the reviewed cases, they would have been applicable to each form of model. A suggested approach to this and the potential implications are discussed in the following sections.

Quasi-steady state models

The processing power required to compute detailed outputs for a complex building in granular time-steps can be considerable. As a consequence, many researchers have sought to reduce the complexity of calculations. In some cases this is done by using quasi-steady state models which average variables over long time periods and keep all building parameters fixed (Raslan & Davies, 2010). Quasi-steady state models extend this approach by incorporating some transient parameters, typically weather by using degree days data for example Cheng & Steemers' model of UK housing (2011). These models can be evaluated on a scaled sample basis (by modelling sample buildings and scaling results to be representative of the stock) (e.g. Firth et al., 2010). The proposed 3 step approach to uncertainty propagation follows that suggested by Tian and Choudhary (2012):

1. Quantify range for each uncertain parameter, paying particular attention to heterogeneity and ensuring that the distribution and ranges of each parameter are carefully considered (Shin et al., 2013). Paying attention to heterogeneity means considering all the ways in which the real buildings might vary from the archetype, orientation is a key example of this. The correlation structure of the inputs should also be

considered (Zhao, 2012) since correlations between inputs can result in both over and under estimation of sensitivity depending on the direction of correlation.

2. For each reference building, create a distribution of output values by repeated sampling from the input parameter distributions. Saltelli et al. (2008) suggest 500 to 1000 evaluation of each archetype. A low discrepancy sampling method such as Latin Hypercube Sampling or Sobol' sequences should be used. The resulting distributions for each archetype are then divided by floor area to create a metric which can be applied to all buildings in the stock.
3. A single sample of the whole stock is generated by creating a sample for each building in the stock through random sampling from the appropriate reference building distribution and scaling by the appropriate parameter (e.g. floor area). This process is repeated until the desired number of samples of the stock have been obtained. A sample of the order of 1000 is recommended, in order to have a sufficiently large sample to apply regression based sensitivity analysis (e.g. the approach used in Tian & Choudhary, 2012) as a subsequent step.

Some quasi-steady state models are evaluated on a building by building basis. In such case, the archetype input data defines the input data for a particular class of buildings but as each building is evaluated separately, more of the details of the individual building are retained with archetype data typically being used to provide building fabric and systems information combined with details of the specific geometry and orientation of the target building. Following construction of the input distribution for each sample building as described in point 1 above, an output estimate for each real building is evaluated by random selection from the relevant sample building input distribution. The real building model is evaluated, and the process repeated the desired number of times. As the initial number of model evaluations is given by the number of buildings in the stock, rather than the number of samples as in the previous case, this is a significantly more onerous undertaking. The simplicity of the underlying building energy model means that this could potentially be applied at an urban scale, but larger scales are likely to remain beyond the scope of current computational results.

Nodal-network models

An alternative approach to reducing the complexity of calculations which underpin the model is the use of resistor-capacitor models, which model the energy flows within each building using an analogy with an electronic circuit (Fonseca et al., 2016). This approach allows the urban scene to be simulated as a single entity. Consequently, as detailed for quasi-steady state models in which all buildings are directly evaluated, following quantification of input distributions, each model evaluation requires sampling of all inputs from the defined distributions. A random sample of each of the inputs is generated and the model evaluated until sufficient samples have been obtained to produce a stable

output distribution. Even if a relatively small number of archetypes is used, with only a few uncertain parameters for each, the total number of uncertain parameters quickly reaches the hundreds. If a sample sufficient for regression analysis is not feasible, a design of experiments method may be possible. These approaches are less robust than the regression method described above but offer the advantage of requiring less computational resource. Fractional factorial (Hajas, 1998) and elementary effects methods (Campolongo, Saltelli, & Cariboni, 2011) use a carefully constructed exploration of the input space to minimise the number of model evaluations required.

Scaled dynamic models

Scaled dynamic models are based on detailed characterisation of sample buildings to calculate the transfer of energy to, from and within the building. This calculation is dynamic with the results of one time-step forming the initial conditions for calculations in the next, often performed for sub-hourly timesteps. The normalised outputs for reference buildings are used develop metrics, e.g. energy use intensity per square metre which can then be scaled according to the proportion of the overall stock represented by that building type. Examples include Heiple and Sailor's model of Houston (2008) and Shimoda et al.'s model of residential building stock in Osaka, Japan (2004).

The approach detailed for quasi-steady state models, in which only the archetype buildings are directly evaluated is appropriate here, but the same concerns apply regarding the importance of accounting for all significant differences between sample and real buildings. The increased number of model inputs may make this more difficult.

Surrogate models

As with scaled dynamic models, detailed simulation is used to generate models for a set of sample buildings. Following this machine learning techniques are used to generate a meta-model or response surface which links inputs and outputs. This surrogate model can then be applied to inputs for the full building stock. (Ascione et al., 2017; Tian & Choudhary, 2012) One approach to this is to develop individual surrogate models for each reference building. SA may be applied at the level of the sample buildings in order to develop the surrogate models. Uncertainty in the whole stock can be generated through stochastic sampling as described for quasi-steady state models above. As before, if sufficient samples are drawn, these can be reused to undertake a regression-based SA of the whole stock.

Full dynamic simulation models

An additional form of BSEM which was not encountered in the publications reviewed in this study is dynamic simulation models, in which each building in the stock is directly evaluated in a full dynamic simulation. Examples of dynamic simulation models include: ECCABS initially developed to evaluate the energy demand of the Swedish residential stock (Mata, Kalagasidis, & Johnsson, 2013) and MIT's UBEM (Cerezo Davila, 2017).

Full dynamic simulation models are the most computationally demanding form of stock model and recent developments aimed at quantifying interactive exchanges between buildings and the environment (Hong & Luo, 2018) are likely to increase these demands. Nonetheless, these models represent an important area of emerging work and merit some discussion here (Reinhart & Cerezo Davila, 2016). For these models, as with individually evaluated quasi-steady state models, dimensionality is a significant concern, compounded by the complexity of the underlying simulation model. While new approaches are emerging to deal with such models (e.g. Becker, Tarantola, & Deman, 2018), uncertainty quantification remains challenging. For these models, evaluating sensitivity using the techniques discussed previously but with factors considered in groups rather than individually is proposed as a strategy for qualitative SA (Saltelli et al., 2008, p. 121).

Dynamic forecasting models

Although no publications about dynamic forecasting models were identified in the review, these are an important category of BSEM, which would be strengthened by quantification of uncertainty and from which insights could be obtained through use of SA (Sandberg, Sartori, & Brattebø, 2014). These models combine changes in the stock of buildings and energy mix over time with bottom-up evaluation of heat balances and energy services to generate long range forecasts of energy consumption. In addition to the uncertain parameters considered above for models of an unchanging building stock, the uncertainty in parameters such as rates of construction, demolition and renewal of the stock and the energy consumption of new and refurbished buildings needs to be considered. The uncertainty in these parameters is driven by a range of complex political and economic factors and may dominate the uncertainty in the baseline parameters. For these models the range of sample buildings can be extended to incorporate representations of new and updated stock and the uncertainty quantification approach set out above for quasi-steady state models could be applied with the weighting of each reference building in the stock treated as a stochastic variable with uncertain parameters.

Conclusions

We conclude that Uncertainty Analysis (UA) and Sensitivity Analysis (SA) are not common practice in building-stock energy modelling and that if UA and SA are performed, only a few parameters are assessed and that methodologies are not standardized. To foster UA (preferably) and SA (subsidiary) we recommend to (i) launch a dedicated research programme and to (ii) to make such analysis a pre-requisite for scientific journal papers. The research programme should (i) highlight the benefits and the costs of different UA and SA approaches, (ii) generate indications about which types and sets of variables, input data and modelling approaches are most appropriate for different types of model output (e.g. energy use and consumption, emissions, costs and benefits), differentiating by use case of the model and

spatial and temporal resolution, and (iii) provide guidelines about good and best practice on how to perform UA and SA in different cases. This should raise awareness about the relevance of uncertainties, give modellers and their clients a guideline on how to prioritize UA and SA depending on the use case and enable them to state the impact of uncertainties at least in a rough semi-quantitative manner. The programme could be based on the outcome of the IEA EBC Annex 70, for instance.

References

- Ascione, F., Bianco, N., De Stasio, C., Mauro, G. M., & Vanoli, G. P. (2017). Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach. *Energy*, *118*, 999–1017.
- Baetens, R., De Coninck, R., Van Roy, J., Verbruggen, B., Driesen, J., Helsen, L., & Saelens, D. (2012). Assessing electrical bottlenecks at feeder level for residential net zero-energy buildings by integrated system simulation. *Applied Energy*, *96*, 74–83.
- 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. <https://doi.org/10.1080/00949655.2018.1450876>
- Booth, A. T., Choudhary, R., & Spiegelhalter, D. J. (2012). Handling uncertainty in housing stock models. *Building and Environment*, *48*, 35–47.
- Booth, A. T., Choudhary, R., & Spiegelhalter, D. J. (2013). A hierarchical Bayesian framework for calibrating micro-level models with macro-level data. *Journal of Building Performance Simulation*, *6*(4), 293–318.
- 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., Saltelli, A., & Cariboni, J. (2011). From screening to quantitative sensitivity analysis. A unified approach. *Computer Physics Communications*, *182*(4), 978–988.
- Cerezo Davila, C. (2017). *Building Archetype Calibration for Effective Urban Building Energy Modeling*. Massachusetts Institute of Technology, Cambridge, Massachusetts.
- 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.
- Collins, L., Natarajan, S., & Levermore, G. (2010). Climate change and future energy consumption in UK housing stock. *Building Services Engineering Research and Technology*, *31*(1), 75–90.
- Davis, P. M., & Cochran, A. (2015). Cited Half-Life of the Journal Literature. *CoRR*, *abs/1504.07479*.
- De Wit, M. S. (1997). Identification of the important parameters in thermal building simulation models. *Journal of Statistical Computation and Simulation*, *57*(1–4), 305–320.
- 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. (2017). 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.
- Firth, S. K., Lomas, K. J., & Wright, A. J. (2010). Targeting household energy-efficiency measures using sensitivity analysis. *Building Research & Information*, *38*(1), 25–41.
- Fonseca, J. A., Nguyen, T.-A., Schlueter, A., & Marechal, F. (2016). City Energy Analyst (CEA): Integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts. *Energy and Buildings*, *113*, 202–226.
- Gaetani, I., Hoes, P.-J., & Hensen, J. L. M. (2016). Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. *Energy and Buildings*, *121*, 188–204.
- Garcia Sanchez, D., Lacarrière, B., Musy, M., & Bourges, B. (2014). Application of sensitivity analysis in building energy simulations: Combining first- and second-order elementary effects methods. *Energy and Buildings*, *68*, 741–750.
- Hajas, W. (1998). *Introducing Confidence Bounds and Confidence Levels into Iterated Fractional Factorial Design Analysis*. University of Manitoba, Winnipeg, Canada.
- Heiple, S., & Sailor, D. J. (2008). Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles. *Energy and Buildings*, *40*(8), 1426–1436.
- Heo, Y., Augenbroe, G., & Choudhary, R. (2011). *Analysis Methodology for Large Organizations' Investments in Energy Retrofit of Buildings*. Retrieved
- Hong, T., & Luo, X. (2018, September 26). *MODELING BUILDING ENERGY PERFORMANCE IN URBAN CONTEXT*. 8. Chicago, Illinois, USA.
- Kavgic, M., Mavrogianni, A., Mumovic, D., Summerfield, A., Stevanovic, Z., & Djurovic-Petrovic, M. (2010). A review of bottom-up building stock models for energy consumption in the residential sector. *Building and Environment*, *45*(7), 1683–1697.

- Kazas, G., Fabrizio, E., & Perino, M. (2017). Energy demand profile generation with detailed time resolution at an urban district scale: A reference building approach and case study. *Applied Energy*, *193*, 243–262.
- Lim, H., & Zhai, Z. J. (2017). Review on stochastic modeling methods for building stock energy prediction. *Building Simulation*, *10*(5), 607–624. <https://doi.org/10.1007/s12273-017-0383-y>
- Mata, É., Kalagasidis, A. S., & Johnsson, F. (2013). A modelling strategy for energy, carbon, and cost assessments of building stocks. *Energy and Buildings*, *56*, 100–108.
- Naber, E., Volk, R., & Schultmann, F. (2017). From the Building Level Energy Performance Assessment to the National Level: How are Uncertainties Handled in Building Stock Models. *Procedia Engineering*, *180*, 1443–1452.
- Nouvel, R., Schulte, C., Eicker, U., Pietruschka, D., & Coors, V. (2013). CITYGML-BASED 3D CITY MODEL FOR ENERGY DIAGNOSTICS AND URBAN ENERGY POLICY SUPPORT. *Proceedings of BS2013*, 218–225. Chambéry, France.
- Nouvel, R., Zirak, M., Coors, V., & Eicker, U. (2017). The influence of data quality on urban heating demand modeling using 3D city models. *Computers, Environment and Urban Systems*, *64*, 68–80.
- Papaefthymiou, G., Hasche, B., & Nabe, C. (2012). Potential of Heat Pumps for Demand Side Management and Wind Power Integration in the German Electricity Market. *IEEE Transactions on Sustainable Energy*, *3*(4), 636–642.
- Raslan, R., & Davies, M. (2010). Results variability in accredited building energy performance compliance demonstration software in the UK: an inter-model comparative study. *Journal of Building Performance Simulation*, *3*(1), 63–85.
- Refsgaard, J. C., & Henriksen, H. J. (2004). Modelling guidelines - terminology and guiding principles. *Advances in Water Resources*, *27*(1), 71–82.
- Reinhart, C. F., & Cerezo Davila, C. (2016). Urban building energy modeling – A review of a nascent field. *Building and Environment*, *97*, 196–202.
- Robinson, D., Haldi, F., Kämpf, J., Leroux, P., Perez, D., Rasheed, A., & Wilke, U. (2009). *CITYSIM: COMPREHENSIVE MICRO-SIMULATION OF RESOURCE FLOWS FOR SUSTAINABLE URBAN PLANNING*. 8.
- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., ... Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*.
- Saltelli, A., & Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, *25*(12), 1508–1517.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., ... Tarantola, S. (2008). *Global Sensitivity Analysis. The Primer*. Hoboken: John Wiley & Sons Ltd.
- Sandberg, N. H., Sartori, I., & Brattebø, H. (2014). Sensitivity analysis in long-term dynamic building stock modeling—Exploring the importance of uncertainty of input parameters in Norwegian segmented dwelling stock model. *Energy and Buildings*, *85*, 136–144.
- Shimoda, Y., Fujii, T., Morikawa, T., & Mizuno, M. (2004). Residential end-use energy simulation at city scale. *Building and Environment*, *39*(8), 959–967.
- Shin, M.-J., Guillaume, J. H. A., Croke, B. F. W., & Jakeman, A. J. (2013). Addressing ten questions about conceptual rainfall-runoff models with global sensitivity analyses in R. *Journal of Hydrology*, *503*, 135–152.
- Stirling, A. (2001). *On Science and Precaution in the Management of Technological Risk - Volume II* (No. Report EUR 19056/EN/2). Retrieved from ESTO-IPTS website:
- Swan, L. G., & Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, *13*, 1819–1835.
- Tian, W. (2013). A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews*, *20*, 411–419.
- Tian, W., & Choudhary, R. (2012). A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London. *Energy and Buildings*, *54*, 1–11.
- Turhan, C., Kazanasmaz, T., Uygun, I. E., Ekmen, K. E., & Akkurt, G. G. (2014). Comparative study of a building energy performance software (KEP-IYTE-ESS) and ANN-based building heat load estimation. *Energy and Buildings*, *85*, 115–125.
- Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & Kreyer von Krauss, M. P. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment*, *4*(1), 5–17.
- Wang, H., & Chen, Q. (2014). Impact of climate change heating and cooling energy use in buildings in the United States. *Energy and Buildings*, *82*, 428–436.
- Zhao, J. (2012). *Development of a Novel Statistical Method and Procedure for Material Characterization and a Probabilistic Approach to Assessing the Hygrothermal Performance of Building Enclosure Assemblies* (PhD Thesis, Syracuse University).

