Best practice reporting guideline for building stock energy models

Nägeli, C.[a,*], Camarasa, C.[b], Delghust, M.[c], Fennell, P.[d], Hamilton, I.[d], Jakob, M.[e], Langevin, J.[f], Laverge, J.[c], Reyna, J.L.[g], Sandberg, N.H.[h], Webster, J.[i]

[a] *Chalmers University of Technology, Gothenburg, Sweden*
[b] *UNEP-DTU Partnership, Copenhagen, Denmark*
[c] *Ghent University, Department of Architecture and urban planning, Research group building physics, Gent, Belgium*
[d] *University College London, London, UK*
[e] *TEP Energy GmbH, Zurich, Switzerland*
[f] *Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
[g] *National Renewable Energy Laboratory, Golden, CO, USA*
[h] *SINTEF Community, Department of Architecture, Materials and Structures, Trondheim, Norway*
[i] *Natural Resources Canada, CanmetENERGY-Ottawa, Ottawa, CA*

Abstract

Buildings are responsible for 38% of global greenhouse gas (GHG) emissions and, therefore, pathways to reduce their impact are crucial to achieve climate targets. Building stock energy models (BSEMs) have long been used as a tool to assess the current and future energy demand and environmental impact of building stocks. BSEMs have become more and more complex and are often tailored to case-specific datasets, which results in a high degree of heterogeneity among models. This heterogeneity, together with a lack of consistency in the reporting hinders the understanding of these models and, thereby, an accurate interpretation and comparison of results. In this paper we present a reporting guideline in order to improve reporting practices of BSEMs. The guideline was developed by experts as part of the IEA's Annex 70 and builds upon reporting guidelines from other fields. It consists of five topics (Overview, Model Components, Input and Output, Quality Assurance and Additional Information), which are further subdivided into subtopics. We explain which model aspects should be described in each subtopic, and provide illustrative examples on how to apply the guideline. The reporting guideline is consistent with the model classification framework and online model registry also developed in the Annex.

Keywords: Building Stock Energy Models, Urban Building Energy Modelling, Model Reporting, Energy Epidemiology, IEA Annex 70

1. Introduction

Buildings will play a critical role in meeting climate change goals. The global buildings and construction sector are together responsible for over a third of the global final energy consumption and 38% of total greenhouse gas (GHG) emissions [1]. Furthermore, annual emissions from the building sector continue to increase, driven by improved access to energy in developing countries, larger ownership and use of energy-operating appliances, and rapid growth in global buildings floor area [1]. To achieve the Paris Agreement goals, the global building stock needs to transition to being net zero carbon, highly efficient and resilient by 2050 [2].

Building stock energy models (BSEMs) offer a tool to assess the energy demand and environmental impact of building stocks, and can demonstrate and evaluate pathways for reducing their

*corresponding author details, claudio.naegeli@chalmers.se

energy demand and respective GHG emissions [3]. Here, we define BSEMs as models that (a) represent multiple buildings that are often - though not always - geographically co-located; (b) produce energy use metrics as an output; and (c) generate out-of-sample predictions. BSEMs can be used to evaluate policy scenarios [4, 5, 6, 7]; energy planning in an urban context [8, 9, 10]; assess life cycle performance of building stocks [11]; develop refurbishment strategies [12, 13, 14], and assess health impacts [15, 16]. Historically, BSEM have encompassed a range of approaches, and recent developments in the BSEM field have focused on bottom-up methodologies [17, 8]. This development has been driven by widening access to building-specific data, in the form of building registries, 3D city models, and energy performance certificates, and increasing computational capabilities. As such BSEMs have become more and more complex. Moreover, due to the case-specific nature of the data availability, BSEMs are usually tailored to the specific datasets and vary widely in terms of their structure and outputs resulting in a high degree of heterogeneity and structural uniqueness exhibited among BSEM [17]. Langevin et al. [17] gives a more detailed overview and classification of the state of the art of different BSEM approaches.

The heterogeneity of BSEMs, together with a lack of consistency in the description and reporting of the models often hinders understanding of the model, impeding accurate interpretation and/or comparison of results. This is a known issue in other domains such as agent-based and medical modelling [18, 19]. In these fields, the issue has been partially addressed through the development of dedicated reporting guidelines and protocols that aim to bring transparency and consistency to reporting on models and their results. The overall aim being to increase robustness of and confidence in reported results. [18, 19, 20, 21].

In order to be relevant, reporting guidelines need to be tailored to the domain in question, as model aspects and methodologies differ from one field to another. Thus far, a dedicated reporting guideline for building stock energy modelling has been lacking and documentation of BSEMs is often poor and inconsistent. In this paper, we present a novel reporting guideline tailored to BSEMs in order to improve reporting practices in the field of building stock energy modelling. The guideline is intended to provide a simple and structured tools for when preparing reports and manuscripts and with a minimum list of information needed to ensure transparency and clarity. The guideline was constructed by modelling experts as part of the International Energy Agency's Annex 70 on Building Energy Epidemiology; the guideline builds upon and adapts existing reporting guidelines from other fields [18, 19, 20, 21] and is informed by a recent review of modelling approaches, which was also carried out within the Annex 70 [17]. The reporting guideline together with the model classification framework presented in Langevin et al. [17] will be complemented with an online registry of BSEMs as part of work in Annex 70 to increase the transparency within the field of building stock energy modelling.

2. Building Stock Energy Model Reporting Guideline

The aim of the reporting guideline is to provide a consistent structure for the description of BSEM and their results (see Table 1). Doing so will help readers, reviewers and other interested parties find relevant information about a model, thereby facilitating interpretation of model results and creating the potential for replicability including scaling of BSEM-based research and services. The guideline can be used to generate stand-alone reports describing a model, to guide substantive or supplementary descriptions in journal publications, and for internal or promotional business documentation purposes. For instance, planners or program managers in governments and utilities procuring BSEM-based services can benefit from improved consistency in model documentation. In a recent review study, a preliminary version of the guideline has even been used as a guideline to collect information on existing models [22].

The guideline is structured into five topics (Overview, Model Components, Input and Output, Quality Assurance and Additional Information), which in turn are broken down into individual subtopics each of them linked to guiding questions to help researchers to interpret the content of different sections. The Overview section gives general information and context of the model and its aim and scope, drawing from the quadrant-based model classification scheme of Langevin et al. [17] to describe the model's high-level structure. Based on that, the Model Components section delves deeper into different aspects of the model and their underlying methodologies. In the Input and Output section, the main model inputs and outputs are described, giving readers a detailed view on what it takes to apply the model and what results might be expected. Next, Quality Assurance summarises results from sensitivity and uncertainty assessments of the model as well as documenting the model's main limitations. In the Additional Information section, further details about the model such as implementation, access and funding are documented.

The set-up of the guideline provided in Table 1 is meant as a way to structure and group different relevant aspects about a given model. In practice, some subtopics can be combined or even left out in order to adjust the guideline to a given model. Another way to use the guideline would be to simply use it as a checklist to see if all necessary information is included in an already existing model documentation.

To help researchers and anyone else seeking to apply the guideline for the first time, a template is provided in the Online Appendix along with example applications to different BSEMs. In the following sections different topics and subtopics are explained further including through guiding questions and extracts from examples provided in the Appendix. This exemplifies the required content per topic and sub-topic.

Table 1: Structure and guiding questions of different aspects of the reporting guideline.

Topic	**Subtopic**	**Guiding questions**
Overview	Aim and scope	What is the overall aim and scope of the model? What are the main use cases addressed?
	Modelling approach	What is the general modelling approach and how is it structured? What are the main model parts and components and how do they relate to each other? What are the key steps in the modelling workflow?
	System boundary	What are the system boundaries (temporal, geographical, building types, energy services, economic sectors, etc.) of the model?
	Spatio-temporal resolution	What is the spatio-temporal resolution on which energy demand calculations are performed? What are common spatial and temporal aggregations on which outputs are reported in relation to use cases?
Model components	Building stock	How are buildings and the building stock represented and characterized in the model? What building attributes are used to characterize buildings on either an individual or archetype basis?
	People	How are people (e.g. occupant behavior) represented in the model?
	Environment	How is the environment (e.g., climatic, policy, economic, context) represented in the model? How does the model account for spatial differences in these environmental aspects?

Topic	Subtopic	Guiding questions
	Energy	How are energy demand (useful, final, primary) and related performance indicators (e.g., GHG emissions) assessed?
	Costs	How are costs (capital and/or operational) assessed?
	Dynamics	How are building stock dynamics (i.e., changes of the stock over time) modeled? Which of the above aspects are modeled dynamically? Which of these dynamics are endogenously defined, what is modeled endogenously?
	Other aspects	Are there other relevant aspects of the model not covered by the above?
Input and output	Data sources	What are the primary data sources used for the model and how they are structured?
	Data processing	How has the data been cleaned, matched or otherwise processed to become input into the model?
	Key assumptions	What are the main input assumptions made to address any information gaps in the data and/or model system?
	Scenario	What model inputs are introduced or modified to describe a scenario?
	Output parameters	What are the main model outputs? At what levels of aggregation and in which formats are they available?
Quality assurance	Calibration	With what method(s) and sources of information has the model been calibrated? What was the outcome?
	Validation	With what method(s) and sources of information has the model been validated? What was the outcome?
	Limitations	What are the (current) limitations of the model and its results? How do modelling assumptions or data limitations affect the model application and/or interpretation of the model results?
	Uncertainty	What are key sources of uncertainty in the model? With what method(s) or sources of information has the uncertainty of the model and results been assessed? What was the outcome?
	Sensitivity	With what method(s) and sources of information has the sensitivity of the model and results been assessed? What was the outcome?
Additional information	Implementation	What software, programming language, packages, libraries or other models are used in or necessary for the model to be used?
	Access	Who owns the model? To whom and under what licence/condition is the model and the necessary data available?
	Funding and contributors	How has the model development and/or underlying research been funded?
	Areas of application	In which geographical areas and for what use cases has it been applied?
	Key references	What are key references showing previous applications and documenting the model and its parts?

2.1. Overview

The overview section of the reporting guideline is ‘abstract’ of the model which includes a high-level description of the model. It is structured into four subtopics: aim and scope, modelling approach, system boundary, and spatio-temporal resolution. This topic is important to quickly communicate a model’s overall structure and focus as well as the relevant system boundaries.

2.1.1. Aim and scope

This section covers the overarching aim and scope of the model, which is critical for understanding the design choices made by the modelers. This section should give a brief description of the aim (i.e., the overall purpose of the model) and scope of the model (i.e., the extent of the area or subject matter the model covers). It should also include a brief description of how the model is used. Answers might include aims such as policy evaluation, urban energy planning, building energy retrofit strategy development, energy poverty identification and alleviation, utility conservation and demand management program planning, etc.

Guiding questions: What is the overall aim and scope of the model? What are the main use cases addressed?

> ***Example (SimStock):***
> SimStock is a Python-based modelling platform which combines data from multiple sources to automatically generate dynamic building energy simulation models ready to be executed by EnergyPlus [1], an open-source whole-building energy modelling (BEM) engine. Typical applications of SimStock include analysis of competing retrofit strategies, exploring potential for integration of renewable technologies or energy storage systems, overheating analysis and identification of indoor air quality risks.

2.1.2. Modelling approach

This section summarizes the model’s overall design (top-down vs. bottom-up) and degree of transparency (black-box vs. white box) using the quadrant-based classification scheme of Langevin et al. [17] as a recommended guide. This allows a reader to identify the model design at a high-level. Additionally, this section provides an overview of key model components, how they relate to each other, and the main steps in the modelling workflow, i.e. the steps in how the model moves information from input through to output when executing it. High-level information in this section is expanded upon in the following sections (e.g., section 2.2 Model Components). Answers report the model’s classification quadrant(s) and high-level modeling technique(s) (e.g., “Q1 (system dynamics)”; “Hybrid Q1/Q4 (techno-econometric/end-use distribution)”, etc.) per the guidance in [17]) as well as describe the high level workflow of the model.

Guiding questions: What is the general modelling approach and how is it structured? What are the main model parts and components included in the model and how do they relate to each other? What are the key steps in the modelling workflow?

> ***Example (ABBSM):***
> The ABBSM applies a bottom-up modeling approach that falls into the hybrid models of classification quadrant Q4 (Bottom-up white box models) according to the classification of Langevin et al. (2020) as it applies a physics-simulation to simulate energy demand in combination with an agent-based modelling approach to model building stock dynamics.

2.1.3. System boundary

The system boundary section encompasses the main delimitations of the model in terms of boundaries of systems covered. This information is necessary for a reader to interpret the boundaries and use-case applicability of modelled results and allows for direct comparison between different models. Such boundaries may include its temporal boundaries (i.e. the time span covered by the model), geographical boundaries (i.e. the geographical area covered), the covered typologies (i.e. residential, commercial, healthcare, etc.), the covered sectors (i.e. what political, economic, social, technological, environmental, legal, activity or end-uses, etc. areas are covered) or the covered energy services (i.e. does it only cover heating demand or also other energy services such as ventilation, appliances, cooling, etc.). Possible system boundaries for different dimensions are:

- temporal: Current state, future scenarios until 2050
- geographical: National, regional, city, district, portfolio
- economic sectors: Residential, non-residential, industrial
- building types: Single-dwelling buildings, multi-dwelling buildings, office buildings, etc.
- energy services: Space heating, hot water, space cooling, ventilation, appliances, lighting

Guiding questions: What are the system boundaries (temporal, geographical, economic sectors, building types, energy services, etc.) of the model?

> ***Example (SimStock):***
> SimStock covers both domestic and non-domestic building stocks and accommodates mixed-use buildings. The geographic boundary is determined by the input data;temporal boundaries are determined by the weather data assigned to the EnergyPlus simulation. Models are limited to thermal and electrical energy consumption within the buildings.

2.1.4. Spatio-temporal resolution

This section covers model resolution in terms of its spatial and temporal aspects. Temporal resolution consists of time step (e.g. sub-hourly, hourly, daily, monthly, annual) and time frame or period over which data inputs were obtained and outputs produced, including baseline and future scenarios. Ultimately, the use case will guide the temporal resolution on which outputs are required. It is good practice to identify the main temporal resolution on which energy demand calculations are performed and common time steps to which outputs are aggregated according to use case. Time frames will vary and may be reported using specific implementations as examples. Spatial resolution describes the spatial scale and scope or geographical extent covered by a models inputs and outputs (e.g. County, climate zones, city, district, individual building). For example, buildings data may be obtained at the dwelling unit level for the extent of an entire municipality. Spatially aggregated outputs are also generated in service of use cases. It is good practice to identify the main spatial resolution on which energy demand calculations are performed as well as any other common geographic aggregations for which outputs are produced, according to use case. For models that are explicitly spatial, it is also good practice to describe the common georeferencing attributes to which other variables are linked.

Guiding questions: What is the spatio-temporal resolution on which energy demand calculations are performed? What are common spatial and temporal aggregations on which outputs are reported in relation to use cases?

> ***Example (Scout):***
> Results are assessed and reported down to the regional level, where region options are selected by the user and include American Institute of Architects (AIA) climate zones [2], U.S. Energy Information Administration (EIA) Electricity Market Module (EMM) regions [3], or U.S. states. The model typically uses a time step of one year, though for measures that affect electricity demand under assessment of EMM regions, users may re-attribute annual results for individual EMM regions down to an hourly time step and report results using this finer temporal resolution.

2.2. Model components

This section goes deeper into the design of the different components (parts/modules) that make up the complete BSEM. In particular, it has sub-sections describing the modelling design and assumptions for the Building stock, People, Environment, Energy and Cost, components of the BSEM, followed by a discussion on the implementation of Dynamics in the stock. This information is crucial for users trying to understand at a deeper level how a model works and the nuances of outputs of the model. The general structure of the BSEM in which these parts are organised and connected should already have been discussed in the Overview section under sub-section Modelling approach (2.1.2) and should therefore not be repeated here. This section however, should contain a short introductory paragraph summarizing the major design choices of each of the components and any pertinent specifics, in order to succinctly communicate more detailed design information.

2.2.1. Building stock

The modelling of buildings is at the core of any BSEM, but many different approaches exist. This section covers the modelling choices for the building characteristics (technical, geometrical and other physical building attributes) and as well as general representation of buildings (e.g., on an individual, archetype, or other aggregation basis), and the more detailed attributes of these representations. This includes the level of detail (LOD) of the geometry that is included, both indoors (e.g. is there room-level information or not, what appliances are in the building?) and on the exterior (e.g. size and orientation of each window or general window to wall ratio). Including this information is important to understand how closely tied a model is to its source data, or if alternate (e.g., newer) data could be used instead. This section also includes how these attributes are selected. E.g., for an archetype approach, the selection and number of archetypes and their attributes based on input data should be discussed under this sub-topic. Examples for different building representations are individual building, building archetypes and statistical sample of buildings. Examples of building attributes are building type, age, floor area, building geometry, characteristics of envelope components (walls, windows, doors), mechanical systems such as heating, ventilation and air conditioning (HVAC), building usage, etc.

Guiding questions: How are buildings and the building stock represented and characterized in the model? What building attributes are used to characterize buildings on either an individual or archetype basis?

> ***Example (RE-BUILDS):***
> The model applies multiple archetypes. For the residential building stock, the archetypes are defined by dwelling type, cohort, and renovation state. For the non-residential building stock, the archetypes are defined by building type and energy performance level. Renovation of buildings may or may not lead make the buildings change energy performance level. Each archetype has the attributes average energy intensity, energy mix and use of local renewable energy sources.

2.2.2. People

This section covers modelling of the people using, occupying and interacting with the buildings. Human behavior has a significant impact on building energy use that interacts strongly with the technical components of the building stock, so it is important to capture these aspects in the reporting protocol. Again, this builds upon the general design approach and goes into details about specific attributes. The general modelling assumptions are the representation of people (e.g. as fixed occupancy schedules, as autonomous individual agents or as interacting families). The discussion of specific attributes is focused on their impact on the behaviour of the people in the model (e.g. socio-economic status, age, gender, preferences, etc.) and how behaviour is further taken into account in the modelling.

Guiding questions: How are people (e.g. in terms of occupant behavior) represented in the model?

> ***Example (SimStock):***
> Occupants are represented by fixed schedules for occupancy, ventilation behaviour, equipment usage and setpoint temperatures. A range of schedules are developed per use-type and assigned randomly to introduce an element of diversity.

2.2.3. *Environment*

The description of the environment and the interaction between the environment and buildings and people are to be described in this section. The environment includes all external influences on the behaviour of the buildings and people in the model (e.g. climate, policy, economy, physical context). This includes both the macro level (e.g. economic cycle), the meso level (e.g. urban heat island effect) and the micro level (e.g. shading by local structures and greenery). Each of these elements can be included in the BSEM directly as inputs or be contained in sub-models. These elements all heavily influence resulting building stock energy use, so it is important to capture if and how they're accounted for in the model.

Guiding questions: How is the environment (e.g., climatic, policy, economic, physical context) represented in the model? How does the model account for spatial differences in these environmental aspects?

> ***Example (SimStock):***
> Policy and economic context are not included in the model. Climate is defined by the weather file, typically a single file is used per city. Solar shading within a user defined radius is incorporated during geometry processing.

2.2.4. *Energy*

In this section, the modelling of energy flows in the BSEM is discussed. This includes type of modelling (e.g. first principles based dynamic multi-zone heat loss simulation, energy intensity estimation, etc.). It can also include further details about the internal resolution of energy calculations for specific energy uses (e.g. higher frequency modelling of a specific energy use in the background as opposed to lower, aggregated resolution of outputs). Examples may include energy demand of building based on measured utility data, through hourly building energy simulation based on EnergyPlus, through monthly building energy simulation based on standard ISO EN 52016-1, and annual energy demand assessment based on heating degree days, etc. Each of these different energy flow methodologies has different implications for the accuracy of the simulated results, so it's important to acknowledge the selected approach.

Guiding questions: How are the energy demand (useful, final, primary) and related performance indicators (e.g., GHG emissions) assessed?

> ***Example (RE-BUILDS):***
> Yearly average energy intensities per archetype is applied to the simulated stock and distribution to archetypes. Energy need intensities for heating and domestic hot water are applied together with assumptions on energy mix, system efficiencies per energy carrier and extent of local renewable energy production. Electric load is added in addition.

2.2.5. *Costs*

This section presents a discussion of the calculation of costs and how they interact with the other components of the model. Examples of such interactions include decisions to invest taken by the people influenced by the expected return on investment and non-energy benefits, such as aesthetic upgrades, etc. In the real-world, cost is an important driver of building energy use, so it is often accounted for in BSEMs.

Guiding questions: How are costs (capital and/or operational) assessed?

> ***Example (ABBSM):***
> The model assesses the life cycle costs of building measures (reinstatement, retrofit or replacement) for both building envelope components and heating systems. The model calculates investment, maintenance and operation as well as energy costs of these technologies and thereby covers the entire life cycle costs.

2.2.6. Dynamics

The dynamics of the building stock are all processes that change the composition of the stock in terms of buildings, people, environment, energy and costs over space and time. This section includes a discussion on how and with what methodologies these processes are modeled in the BSEM. Building stock dynamics are highly complex as well as highly influential on building stock energy use, so the inclusion of these in the BSEM should be well-described. Examples of different methodologies to model building stock dynamics are based on fixed renovation and construction rates, using agent-based methods, or on system dynamics. An example of dynamics in the Environment aspect of BSEMs is the representation of climate change: considered following a predefined scenario or including a feedback-loop with the results of the BSEM (e.g. calculated building thermal balances impacting the heat islands being considered).

Guiding questions: How are building stock dynamics (i.e., changes of the stock over time) modeled? Which of the above aspects are modeled dynamically? Which of these dynamics are endogenously defined, what is modeled endogenously?

Example (RE-BUILDS):
The core of the model is the long-term development of the building stock resulting from the population's demand for floor area in buildings of various types. For all years, the demand for floor area A in buildings of type t is simulated by multiplying the size of the population which is using the given building type Pt with the corresponding average demand for floor area per person AP in building type t [...]. This gives the total demand for floor area and hence the building stock size and its distribution to various building types for each year, At.

2.3. Input and output

This section describes the necessary inputs and generated outputs of the BSEM. In particular, it outlines necessary data sources, required data processing and assumptions made therein and structure and resolution of model outputs. If relevant for a given BSEM, inputs associated with different scenarios can also be described. These sections are critical for communicating the outputs of a model and the necessary data (as well as any introduced data bias) within the BSEM.

2.3.1. Data sources

BSEMs are typically highly data intensive models and often tailor-made using specific datasets. As such, a description of the data used for the case presented is essential in order to understand how the model works, what level of detail it is designed for and the use cases to which it might plausibly be applied. In this section the (main) data sources for the model are outlined and described. Where possible for each source, the relevant information on type of data (e.g. energy consumption, building characteristics, stock forecast, etc.), data publication/collection year, sample size (if appropriate), author name / institution, website and DOI (if available) should be included in the description. Examples of different sources might be building registries, energy performance certificates, 3D city models, census data, building typology descriptions, survey data, etc. Where the model is dependent on particular data sets which cannot be substituted, this should be clearly noted.

Guiding questions: What are the primary data sources used for the model and how they are structured?

Example (SimStock):
Data sources depend on the study area. Models of UK cities are based on Ordnance Survey data for building footprints. LiDAR data is used for building heights and to identify different footprints on different floors, UK Buildings Survey (building age), Valuation Office Agency (property tax data for use classes). Building fabric and occupancy details are based on national standards.

2.3.2. Data processing

Many BSEMs process input data to some degree before the main modelling tasks are performed, especially if they built on microdata such as building registries, energy performance certificates or similar. Typically, many data cleaning steps such as removal of outliers and duplicates, filling in missing values, and matching and merging of data sets are needed in order to prepare input data for use in the model. These steps should be described in this section, citing the reason for each step and implications for modelling results. This is necessary to communicate any uncertainty or bias introduced into the BSEM by these procedures.

Guiding questions: How has the data been processed to become model inputs?

> ***Example (ABBSM):***
> Data from different sources outlined under data sources [...] is processed and merged in order to generate a representative synthetic building stock according to the methodology described in (Nägeli et al., 2018). The method is structured into three steps: [...]

2.3.3. Key assumptions

In order to address information gaps in input data and/or model system, BSEMs often rely on several assumptions to fill these gaps. These assumptions are not always made explicit when documenting model results, which hinders transparency of results, so it is critical to include these in the reporting protocol. In this section key assumptions made regarding input data and in model design should be documented.

Guiding questions: What are the main input assumptions made to address any information gaps in the data and/or model system?

> ***Example (SimStock):***
> Key assumptions are:
> - Floor to ceiling height is taken from VOA data where available or assumed to be 3m.
> - Building fabric data is inferred from age and function data.
> - Building services are based on ideal-loads with efficiencies added in post-processing.
> - Occupancy, equipment use and set-point temperature schedules are assigned stochastically based on use type.

2.3.4. Scenario

A common use of BSEMs is to model building stock scenarios over time or investigate the effect of large scale implementation of measures across the stock [23]. Knowing which model parameters are considered in scenario development is important for understanding how scenario results are generated, interpreting the results and guiding alternative scenario generation using the same model. This section should therefore outline the main parameters comprising a typical scenario and describe how they affect model outcomes. This section is intended for models with predictive capabilities and can be omitted for informational models whose aim is to simply describe the status quo of the building stock. Example variables that may be modified include: number of residents and employees, number and type of new construction, number and type of retrofit(s) to existing dwellings, equipment penetration rate, utility rate increases etc..

Guiding questions: What model inputs are modified to describe a scenario?

Example (RE-BUILDS):
The Baseline scenario assumes a development that would have been likely if zero emission technologies as part of the ZEB concept are not introduced. Here, new construction is according to the TEK 17 standard in the period 2020-2024, which is the current conventional building code for new built in Norway. From 2025 onwards, new construction is assumed to be according to the Norwegian passive-house standards (Standards Norway, 2013, 2012). Furthermore, the Baseline scenario assumes that buildings that are renovated are energy upgraded corresponding to current practice until 2035. From 2035 onwards, advanced renovation with higher energy savings is assumed. [...]

2.3.5. Output parameters

BSEMs can often deliver a multitude of different results at different levels of aggregation, not all of which might be shown in a given report or publication. In order to understand a models capabilities, it is therefore important to clearly describe what kind of outputs or results it is able to produce and at what level of detail.

Guiding questions: What are the main model outputs? At what levels of aggregation and in which formats are they available?

Example (RE-BUILDS):
Building stock development: Stock development during the period under study, and the stock's distribution to types and archetypes. Delivered energy: Aggregated delivered energy per year for the total stock, per building type, per construction period or per energy carrier. Scenario specific. Average energy intensity: Average energy intensity across the entire system, per square meter or per person. Scenario specific. GHG emissions: GHG emissions per year for the total stock, per building type, per construction period or per energy carrier. Scenario specific and according to the five variants of the emission intensities for electricity.

2.4. Quality assurance

Quality assurance of simulation models is most often discussed in the context of the software itself, for example in the development of the BESTEST procedures which define standardised testing procedures against which simulation tools are validated [24]. However, these procedures relate to the simulation of individual buildings and focus on "testing a software package's ability to model thermal processes associated with the building envelope". For building stock energy models which couple the energy demands (and potentially interactions) of large numbers of individual buildings and frequently integrated assessment of the impacts of specific policy drivers, verification of the underlying simulation tool is a necessary but far from complete element of model quality assurance. As a result, the model reporting guidelines focus on the following critical aspects of quality assurance: calibration, validation, limitations and uncertainty.

2.4.1. Calibration

The fine-tuning of model parameters within carefully specified ranges to allow the model output to be matched to a specified "ground-truth" value, typically the measured energy consumption but sometimes an output from an existing verified model or statistics. It should be recognised that calibration is almost always an under-determined problem and a degree of heuristic insight is typically required. It is unlikely that any one calibrated solution is the true value and therefore a better approach is to search for a small set of plausible solutions [25] . Therefore, it is important to note that a model is only calibrated for the temporal and spatial resolution for which the calibration exercise was undertaken. For example, a model calibrated using annual energy consumption data cannot be considered to be calibrated if used to produce monthly outputs [26]. In order for readers to judge the validity of a model in a given application it is therefore important to know how it has been calibrated, which should be described in this section.

Guiding questions: With what method(s) and sources of information has the model been calibrated? What was the outcome?

> ***Example (ABBSM):***
> The model decision model and the resulting retrofit and heating system adoption behaviour has been calibrated based on the historic observed adoption behaviour in the Swiss residential building stock. The results of which are documented in (Nägeli et al., 2020a).

2.4.2. Validation

In simulation modelling, validation is the comparison of the model output against the "ground-truth" value. The acceptable deviation from the ground-truth value depends on the specific context of the model. For BSEMs there is typically significant variation between validation results at the aggregated and individual building scale [27], therefore validation results should always be presented at multiple levels whenever possible and feasible to increase transparency on the modeled accuracy.

Guiding questions: With what method(s) and sources of information has the model been validated? What was the outcome?

> ***Example (ResStock):***
> For the End-use Load Profiles project, ResStock was validated against multiple years of electric utility AMI/LRD for regions that weren't the specific focus of calibration. This is detailed in a forthcoming project report (expected early 2022). More information can be found at: https://www.nrel.gov/buildings/end-use-load-profiles.html.

2.4.3. Limitations

BSEMs as any kind of model inherently have different limitations in terms of functionalities and the results they can produce due to simplifications and assumptions made in the modelling process and/or the specificities and limitations of the training data set in case of calibrated or fully data-driven models model. In this section, these model limitations and its results should be clearly described in order to aid in interpretation of the results.

Guiding questions: What are the (current) limitations of the model and its results? How do modelling assumptions or data limitations affect the model application and/or interpretation of the model results?

> ***Example (ABBSM):***
> Known limitations of the model are documented and discussed in detail in Nägeli et al. (2020a, 2020b, 2018), as summary of which is given here: The use of synthetically generated building agents makes it possible to represent the heterogeneity in the building stock, however, as the different input distributions are assumed to be independent from each other the method may lead to unrealistic combinations of attributes in some cases. [...]

2.4.4. Uncertainty

Due to the complexity in the models as well as the often incomplete data used in BSEMs, model results are inherently uncertain. In this section, the sources of uncertainty in the model and its results should be documented and described. Ideally this is done in combination with a quantitative assessment of the uncertainty using established uncertainty analysis methods. At the time of this publication, uncertainty modelling is not common practice for BSEMs, but it is critical for improving the quality, confidence, and transparency of the BSEM field.

Guiding questions: What are key sources of uncertainty in the model? With what method(s) has the uncertainty of the model and results been assessed? What was the outcome?

> ***Example (Scout):***
> Quantitative assessment of model uncertainty has not yet been undertaken; principal uncertainties likely arise from the following modeled elements: Measure installed cost and performance inputs and existing equipment stock turnover rates may vary widely across locations, individual building configurations, and consumer segments, and disruptive changes to technology characteristics may occur that are unforeseen in the building technology policy roadmaps that Scout relies upon to inform prospective efficiency, flexibility, and fuel switching measure definitions. [...]

2.4.5. Sensitivity

The complexity of BSEMs means that relationships between inputs and outputs may not be intuitive for readers. With large numbers of input parameters, understanding those which are key drivers of model output is important for understanding. While model sensitivity is driven in part by the uncertainties inherent in the input parameters used to describe a specific case or location, for complex models it is also a function of model structure. The relative importance of different sub-models or ranges within which different operating regimes are often unknown without sensitivity analysis. Consequently, while sensitivity analysis results cannot be assumed to be consistent across all use cases, they provide essential information to aid the interpretation of model results. Scenario analysis, while useful for exploring variations due to particular narratives of change, should be used as an adjunct to and not a substitute for Sensitivity Analysis. The non-linear nature of BSEMs means a Global Sensitivity Analysis approach such as the Sobol method [28] or Elementary Effects method [29] is strongly preferred.

Guiding questions: With what method(s) has the sensitivity of the model and results been assessed? What was the outcome?

> ***Example (SimStock):***
> An initial sensitivity analysis has been undertaken indicating that set-point temperatures are the most influential parameter. This was stable across a range of different Global Sensitivity Analysis methods [3]. The impact of floor to ceiling height and assignment of building fabric types was not considered in the analysis.

2.5. Additional information

This section covers additional information on the model for any interested parties to understand the context of the model as well as for potential new users to consider whether the model is accessible and suitable for use in support of policy, programs or research. This topic should include how the model is implemented, in terms of software requirements for use of the model, the accessibility for new users, transparency in terms of listing funding resources and intellectual property, geographical areas where the model has already been applied as well as the key references. This information is necessary to ensure that the model can be located, accessed, and sources of funding or bias are acknowledged.

2.5.1. Implementation

The implementation section should describe software packages, libraries or other models that are used in or necessary for the model to run (either embedded or coupled). It should also specify which programming languages have been used to develop the model. This is necessary for identifying any prerequisites as well as any software-specific intricacies of the BSEM. Examples of the software and programming include: Excel, Access, HOT2000, Open Studio, Sketchup, EnergyPLUS, ArcGIS, Geomedia, MapInfo, FME, SAS, SPSS, Python, Ruby, C++, SQL, etc.

Guiding questions: What software, programming language, packages, libraries or other models are used and necessary for the model to be run?

Example (RE-BUILDS):
The residential building stock model is a combined Excel/MATLAB model, where inputs and outputs are given in Excel whereas the model is run in MATLAB. The non-residential building stock model is implemented in Excel.

2.5.2. Access

This section explains how to access the model and ancillary information to make use of the model and generate results. Aspects related to intellectual property include model ownership, who can typically access the model and on what basis, and licensing approach and type. Data sharing agreements required to access key data sets should be mentioned. If it is an open source model, a link to the code repository could be included. Contact information for a developer or other proponent should be provided. This information is necesesary for communicating with any potential users of the BSEM.

Guiding questions: Who owns the model? To whom and under what type of licence is the model and the necessary data available?

Example (SimStock):
UCL Energy Institute are the owners and developers of the SimStock modelling platform, for access see: https://www.ucl.ac.uk/energy-models/models/simstock

2.5.3. Funding and contributors

The transparency section seeks to provide information on funding. It should also include the list of authors and acknowledgments for the various parts and contributions to the model. This is important both to acknowledge the organisations funding the development of the BSEM as well as to reveal any funding-related bias in the BSEM.

Guiding questions: How has the model development and/or underlying research been funded?

Example (ResStock):
Most of the core development of ResStock has been funded by the U.S. Department of Energy Building Technologies Office. Enhancements and user-specific features have been developed by a range of organizations, key contributors include the Bonneville Power Administration and the Los Angeles Department of Water and Power.

2.5.4. Areas of application

The areas of applications refers to both the geographical (i.e. region, country, city, district, etc.), as well as subject areas of application (i.e. assessing and monitoring, forecasting, strategy development, etc.). This is important to understand how existing users have deployed the BSEM. Examples of geographical areas are: EU27, Germany, Berlin, Wedding. Examples of areas of application are: spatial analysis of renewable energy sources in Copenhagen.

Guiding questions: In which geographical areas and for what use cases has the model been applied?

Example (ABBSM):
The model has been originally developed and apply to the building stock of Switzerland. However, versions of the model have also been applied to model the building stock of France, Germany, England, the Netherlands, Italy, Spain, Poland, Belgium and Slovenia.

2.5.5. Key references

The purpose of this section is to provide the reader to find further information and examples of use. To this end, a list of key references should be provided, referring to sources where the model has already been described and/or applied. For example, reference to a more detailed technical description of the model and model parts (e.g. in pseudo code, UML, etc.) or an online code documentation can be given (e.g. HTML based code documentation using the open source tool Sphinx for Python-based models [30]).

Guiding questions: What are key references showing previous applications and documenting of the model or parts of the model?

> ***Example (RE-BUILDS):***
> Sandberg, 2017: Dynamic modelling of national dwelling stocks. Understanding phenomena of historical observed energy demand and future estimated energy savings in the Norwegian dwelling stock. PhD thesis. Sartori et al., 2016: Dynamic dwelling stock model – model description and exemplification for Norway Sandberg et al., 2014b: Segmented dynamic dwelling stock model Sandberg et al., 2014a: Sensitivity analysis for the segmented dynamic dwelling stock model [...]

3. Discussion and Conclusion

This model reporting guideline should be used as a tool by authors, reviewers, and journal editors, in order to promote best practices in reporting building stock models and their results. It is our hope that through the application of the guideline the transparency and understanding of BSEMs, their results and their reliability, are improved. In addition, that making use of the guideline also offers benefits to modellers in terms of providing a clear framework for how to describe and report their models, which should make it easier to write as well as read model documentation as it will have a consistent form. Moreover, using the guideline as a checklist seeks to ensure that important information is not omitted in reporting. Lastly, using a standardised format for model documentation will make reporting modelling results in future publications more straightforward as (parts of) the documentation can be reused. For readers, the guideline will make it easier to find relevant information about a model, which will also make the comparison of different models more straightforward as highlighted by the recent review study based on a preliminary version of the guideline [22]. A limitation identified in this review is that because models are continuously being improved, any model description should be viewed as a snapshot in time with more recent improvements not captured. Secondly, while the guideline is comprehensive, certain details may be intentionally or unintentionally omitted by developers. For example, private sector developers will describe their models in less detail to protect their intellectual property.

Further to the description in this article, to assist users in the application of this guideline we also provide a template document which can be used by modelers to describe their models. Moreover, we also provide example applications of the guideline in the description of models developed by the authors in the Online Appendix. These examples highlight different ways of applying the guideline to different types of models and at different levels of detail. These examples show that despite the guideline giving some boundaries in what and how models should be reported on, there is still plenty of freedom to adapt and tailor the guideline to a given model or set of models for reporting purposes as well as adjusting the length, make-up and level of detail of the model description to the intended purpose.

The field of building stock modelling and its modelling methods are still very much evolving and with it are its reporting requirements. The reporting guidelines presented here are therefore designed to be flexible enough to be applied to different model types and applications. Therefore, as the guideline is applied more and more and with it the knowledge around reporting on BSEMs accumulates, there will be need for an adaptation of the guideline in the future as has already been done in other fields [31, 18].

4. Acknowledgments

The authors gratefully acknowledge the strong support of Annex 70 from the International Energy Agency Energy in Buildings and Communities Programme (IEA-EBC). This work was authored in part by the Regents of the University of California for the U.S. Department of Energy (DOE) under contract no. DE-AC02-05CH11231 and by the National Renewable Energy Laboratory, operated by Alliance for Sustain-able Energy, LLC, under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Building Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. This work was partly funded by the Swiss Federal Office of Energy (SFOE) through its research program "Buildings and Cities" as part of the Swiss contribution to IEA-EBC Annex 70 and partly by Enova SF through their funding of the Norwegian contribution to IEA-EBC Annex 70. This work was partly funded by the Bitten and Mads Clausen Fundation and the Copenhagen Centre on Energy Efficiency (C2E2). The C2E2 is institutionally part of the UNEP DTU Partnership – operating under a tripartite agreement between the Danish Ministry of Foreign Affairs, UN Environment Programme and the Technical University of Denmark (DTU). This work was partly funded through the UK EPSRC project (EP/R035288/10) and the iNumber project (EP/R008620/1). This work was partly funded through the Program of Energy Research and Development of Natural Resources Canada.

Nomenclature

BEM Building energy modelling

$BSEM$ Building stock energy modelling

DOI Digital Object Identifier

GHG Greenhouse gases

$HVAC$ Heating, ventilation and air conditioning

IEA International Energy Agency

LOD Level of detail

ZEB Zero-emission building

References

[1] UNEP, 2020 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector, Technical Report, United Nations Environment Programme, Nairobi, 2020. URL: http://www.ren21.net/resources/publications/.

[2] J. Laski, V. Burrows, From Thousands to billions: Coordinated Action towards 100% Net Zero Carbon Buildings By 2050, Technical Report, World Green Building Council, London, United Kingdom, 2017. URL: http://www.worldgbc.org/sites/default/files/FromThousandsToBillionsWorldGBCreport_FINALissue310517.compressed.pdf.

[3] M. Kavgic, D. Mumovic, A. Summerfield, Z. Stevanovic, O. Ecim-djuric, Uncertainty and modeling energy consumption : Sensitivity analysis for a city-scale domestic energy model, Energy & Buildings 60 (2013) 1–11. URL: http://dx.doi.org/10.1016/j.enbuild.2013.01.005. doi:10.1016/j.enbuild.2013.01.005.

[4] R. McKenna, E. Merkel, D. Fehrenbach, S. Mehne, W. Fichtner, Energy efficiency in the German residential sector: A bottom-up building-stock-model-based analysis in the context of energy-political targets, Building and Environment 62 (2013) 77–88. URL: http://dx.doi.org/10.1016/j.buildenv.2013.01.002. doi:10.1016/j.buildenv.2013.01.002.

[5] J. W. Bleyl, M. A. Casas, A. Hulshoff, M. Robertson, M. Bareit, B. D. Bruyn, S. Mitchell, Building deep energy retrofit : Using dynamic cash flow analysis and multiple benefits to convince investors, ECEEE Summer Study Proceedings (2017).

[6] N. H. Sandberg, J. S. Næss, H. Brattebø, I. Andresen, A. Gustavsen, Large potentials for energy saving and greenhouse gas emission reductions from large-scale deployment of zero emission building technologies in a national building stock, Energy Policy 152 (2021) 112114. doi:https://doi.org/10.1016/j.enpol.2020.112114.

[7] C. Nägeli, M. Jakob, G. Catenazzi, Y. Ostermeyer, Policies to decarbonize the Swiss residential building stock: An agent-based building stock modeling assessment, Energy Policy 146 (2020). doi:10.1016/j.enpol.2020.111814.

[8] C. F. Reinhart, C. Cerezo Davila, Urban building energy modeling - A review of a nascent field, Building and Environment 97 (2016) 196–202. URL: http://dx.doi.org/10.1016/j.buildenv.2015.12.001. doi:10.1016/j.buildenv.2015.12.001.

[9] N. Kohler, P. Steadman, U. Hassler, Research on the building stock and its applications, 2009. doi:10.1080/09613210903189384.

[10] J. Webster, B. Korteling, R. Boulter, K. Cooper, A. Mohareb, L. Saikali, R. Tooke, Evaluating Residential Energy, Emissions and Cost Scenarios for Prince George's Official Community Plan: ICEM Approach, Methods and SCEC 3 Model Results, Technical Report March, Natural Resources Canada, 2013.

[11] A. Mastrucci, A. Marvuglia, U. Leopold, E. Benetto, Life Cycle Assessment of building stocks from urban to transnational scales: A review, Renewable and Sustainable Energy Reviews 74 (2017) 316–332. doi:10.1016/j.rser.2017.02.060.

[12] M. Österbring, C. Nägeli, C. Camarasa, L. Thuvander, H. Wallbaum, Prioritizing deep renovation for housing portfolios, Energy & Buildings 202 (2019). doi:10.1016/j.enbuild.2019.109361.

[13] J. L. Reyna, M. V. Chester, The Growth of Urban Building Stock: Unintended Lock-in and Embedded Environmental Effects, Journal of Industrial Ecology (2015). doi:10.1111/jiec.12211.

[14] C. Nägeli, A. Farahani, M. Österbring, J. O. Dalenbäck, H. Wallbaum, A service-life cycle approach to maintenance and energy retrofit planning for building portfolios, Building and Environment (2019). doi:10.1016/j.buildenv.2019.106212.

[15] I. Hamilton, J. Milner, Z. Chalabi, P. Das, B. Jones, C. Shrubsole, M. Davies, P. Wilkinson, Health effects of home energy efficiency interventions in England: A modelling study, BMJ Open (2015). doi:10.1136/bmjopen-2014-007298.

[16] M. Ferreira, M. Almeida, Benefits from energy related building renovation beyond costs, energy and emissions, in: Energy Procedia, 2015. doi:10.1016/j.egypro.2015.11.199.

[17] J. Langevin, J. L. Reyna, S. Ebrahimigharehbaghi, N. Sandberg, P. Fennell, C. Nägeli, J. Laverge, M. Delghust, Mata, M. Van Hove, J. Webster, F. Federico, M. Jakob, C. Camarasa, Developing a common approach for classifying building stock energy models, Renewable and Sustainable Energy Reviews (2020). doi:10.1016/j.rser.2020.110276.

[18] V. Grimm, U. Berger, D. L. Deangelis, J. G. Polhill, J. Giske, S. F. Railsback, The ODD protocol : A review and first update, Ecological Modelling 221 (2010) 2760–2768. URL: http://dx.doi.org/10.1016/j.ecolmodel.2010.08.019. doi:10.1016/j.ecolmodel.2010.08.019.

[19] C. Bennett, D. G. Manuel, Reporting guidelines for modelling studies (2012).

[20] C. Hogan, A. Gretchen, M. Ezzat, J. T. Gr, C. J. L. Murray, P. J. Sim, T. Vos, W. The, The GATHER Statement : Explanation and Elaboration (2016) 1–41.

[21] G. S. Collins, J. B. Reitsma, D. G. Altman, K. G. Moons, Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement, European Urology (2015). doi:10.1016/j.eururo.2014.11.025.

[22] P. Claudino, J. Webster, Scan of Building Energy Mapping Applications, Technical Report, Posterity Group, 2021.

[23] A. Mastrucci, A. Marvuglia, U. Leopold, E. Benetto, Life Cycle Assessment of building stocks from urban to transnational scales : A review, Renewable and Sustainable Energy Reviews 74 (2017) 316–332. URL: http://dx.doi.org/10.1016/j.rser.2017.02.060. doi:10.1016/j.rser.2017.02.060.

[24] R. D. Judkoff, J. Neymark, The BESTEST Method for Evaluating and Diagnosing Building Energy Software, ACEE Summer Study of Energy Efficiency in Building (1998).

[25] T. A. Reddy, I. Maor, C. Panjapornpon, Calibrating detailed building energy simulation programs with measured data—part i: General methodology (rp-1051), HVAC&R Research 13 (2007) 221–241. URL: https://www.tandfonline.com/doi/abs/10.1080/10789669.2007.10390952. doi:10.1080/10789669.2007.10390952. arXiv:https://www.tandfonline.com/doi/pdf/10.1080/10789669.2007.103

[26] F. Khayatian, A. Bollinger, P. Heer, Temporal resolution of measurements and the effects on calibrating building energy models, 2020. arXiv:2011.08974.

[27] C. Cerezo, J. Sokol, S. AlKhaled, C. Reinhart, A. Al-Mumin, A. Hajiah, Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): A residential case study in Kuwait City, Energy and Buildings 154 (2017) 321–334. URL: https://doi.org/10.1016/j.enbuild.2017.08.029. doi:10.1016/j.enbuild.2017.08.029.

[28] I. M. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computers in Simulation (2001). doi:10.1016/S0378-4754(00)00270-6.

[29] F. Campolongo, J. Cariboni, A. Saltelli, An effective screening design for sensitivity analysis of large models, Environmental Modelling Software 22 (2007) 1509–1518. URL: https://www.sciencedirect.com/science/article/pii/S1364815206002805. doi:https://doi.org/10.1016/j.envsoft.2006.10.004, modelling, computer-assisted simulations, and mapping of dangerous phenomena for hazard assessment.

[30] Sphinx, Sphinx 4.0.0+ documentation, 2020. URL: https://www.sphinx-doc.org/en/master/.

[31] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jørgensen, W. M. Mooij, B. Müller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R. A. Stillman, R. Vabø, U. Visser, D. L. DeAngelis, A standard protocol for describing individual-based and agent-based models, Ecological Modelling (2006). doi:10.1016/j.ecolmodel.2006.04.023.