

Evaluating the impact of data quality on the accuracy of the predicted energy performance for a fixed building design using probabilistic energy performance simulations and uncertainty analysis



Tomas Ekström^{a,b,*}, Stephen Burke^{a,b}, Magnus Wiktorsson^c, Samer Hassanie^d, Lars-Erik Harderup^a, Jesper Arfvidsson^a

^a Division of Building Physics at the Department of Building and Environmental Technology, Lund University, Lund, Sweden

^b NCC AB, 205 47 Malmö, Sweden

^c Centre for Mathematical Sciences, Lund University, Lund, Sweden

^d EQUA Simulations AB, Sweden

ARTICLE INFO

Article history:

Received 8 April 2021

Revised 16 June 2021

Accepted 22 June 2021

Available online 24 June 2021

Keywords:

Probabilistic method

Dynamic energy simulations

Field measurements

Data quality

Monte Carlo method

Multi-family house

Case study

Design known uncertainties

ABSTRACT

Probabilistic building performance simulations enables a method for evaluating how the data quality used to create input models affects the accuracy of the predicted energy performance. The method described in this paper explores an uncertainty analysis of a fixed building design and system by including the discrepancy between the declared property of a product, material, or behaviour quantified under specific conditions, and the actual performance in real-world use. The method was tested using a case study, a multi-family building, using two datasets based on different data quality to quantify the discrepancies. The models' outcome was validated against field measurements from 28 buildings built using the fixed building design. The main findings were that data quality significantly shifted the probability density curves and consequently impacted the predictions and accuracy of the predictive models, showing the importance of high-quality input models and a validation process based on a probability distribution. The study also indicated that higher quality data does not equal narrower distributions in the input models. In this case, the case study showed that, despite well-known building properties, narrowing the performance gap can only occur through larger variation in the input data models, resulting in a larger predicted energy performance interval.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many countries strive for continued reductions in energy demand from buildings by implementing more stringent building regulations to mitigate climate change [1]. Consequently, the importance of accurate and efficient building performance simulations (BPS) to predict the influence of different building designs on energy performance (EP) is increasing [2]. As observed in earlier studies [3–8], there is a performance gap between the predicted EP – using deterministic methods for BPS – and the building's actual EP. Overcoming the performance gap is one of the main challenges that remain with BPSs, according to Hong et al. [9]. One technique of identifying and quantifying the performance gap is to use uncertainty analysis [10] based on probabilistic meth-

ods to model and include the relevant uncertainties of the real world in the input data of the predictions. The development to include uncertainty quantification and stochastic parameters in BPS appears to have started in the 1990 s and has developed since then [11–13].

A state of the art review regarding uncertainty analysis in building energy assessment is presented by Tian et al. [10], showcasing input data models, sampling methods and software available for uncertainty analysis. Tian et al. identified the need for more research regarding optimisation under uncertainty, focusing on uncertain factors during the design phase. Using uncertainty analysis for decision-making by quantifying the risk of a building design was proposed by Pettersen already in 1997 [13], and based on the same idea, Sun et al. [14] developed a framework for risk analysis using probabilistic simulations. However, their model was never validated against a population of building with the same design.

* Corresponding author at: Division of Building Physics at the Department of Building and Environmental Technology, Lund University, 221 00 Lund, Sweden.

E-mail address: tomas.ekstrom@ncc.se (T. Ekström).

One method that applies this concept was developed and tested by Burke et al. [15]; later, Ekström et al. [16] applied the method and expanded the method's scope to include risk analysis. The concept of a risk analysis is to quantify the probability and consequence of an event occurring as a basis for quantifying a risk level of a proposed building design. In this case, the unwanted event, or failure, is defined as the building's predicted or actual EP not attaining the design criteria for the EP. However, the predicted risk level is dependent on an accurate prediction of both the probability and the consequence. In the study by Ekström et al. [15], the accuracy between the predicted probability of failure based on simulations and the actual EP based on field measurements (FM) was less than expected. Thus, investigating how to accurately predict the probability of failure using a predictive model and the model accuracy's dependence on data quality is necessary.

This study investigates how the data's quality used to create models for stochastic parameters affects the accuracy of the predictive model, created using uncertainty analysis, and how to improve the accuracy of the predictive model. The prediction accuracy is evaluated regarding the trueness and precision of the simulated results compared to the reference values, i.e. actual EP based on FM from buildings, illustrated in Fig. 1, as described in ISO 5725 [16]. The trueness refers to the proximity of measured results to the true value, while precision refers to the measurement's repeatability or reproducibility [16]. The aim was to validate the predictive model against the actual variation in EP in buildings built with the same design quantified by FM. The aim was also to investigate the influence of the decision-making process on the probability of failure by using different options for the design criteria as a sensitivity analysis. The options evaluated were identified from the original contracts from the built objects used in the FM. The options showcase the influence and consequence of the decision-making process and why the probabilistic process of producing data enables a data-informed process for decision-making. The process and results provide new knowledge about predicting the performance gap using a probabilistic approach and how to apply this knowledge to predict the probability of failure for a building design during the design phase of the building process. This also provides insight into how to approach the generation of data sets that are more suitable for different objects or uses while showing how incorrect data can impact the BPS results.

Although the method's application in this paper is within a Swedish context and based on Swedish conditions, it is not limited to this specific use case. The climate's influencing factor was included deterministically, excluding uncertainties about this

parameter from the simulations in this study. The quantified uncertainties are limited to those that could arise in the design phase and during the first two years of building use. The quality of the input data depends on the various sources that could be obtained. In these cases, the input data is not explicitly related to the measured objects since collecting data from all the individual inhabitants or apartments would be too time-consuming and not within this project's scope. In these cases, the data quality was evaluated qualitatively by the authors related to the studied objects.

1.1. Types of uncertainties

Uncertainties in BPS originate from various causes. Examples include; uncertain design choices; component properties; limited data; the operator's experience, knowledge and their interpretation of information; time-constraints; limitations in the simulation software; and the deviation between designed and real-world operation and occupant behaviour in the finished building [7,12]. The uncertainties can be divided into different classes, depending on the source of the uncertainty. The classification of uncertainties presented by Rezaee, R. et al. [17] clarifies and defines the different types of uncertainties, as shown in Table 1.

1.2. Energy performance of buildings

EP was defined based on the Energy performance of buildings directive [18] as delivered energy used in the building for space heating, domestic hot water (DHW), and electricity for building services. The process for calculating the EP was based on the Swedish building regulation [19]. The building's EP ($\text{kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$) was defined as delivered energy used per year ($\text{kWh}\cdot\text{a}^{-1}$) in the building for space heating (E_{SH}), DHW (E_{DHW}), and electricity for building services (E_{EBS}) divided by the floor area (A_{temp}) (m^2) and calculated according to eq. (1). The floor area referenced in this study is the heated floor area if not otherwise stated. Active cooling is uncommon in Swedish residential buildings and was not available in the case study buildings. Thus, cooling is excluded from the method described.

$$EP = \frac{E_{\text{SH}} + E_{\text{DHW}} + E_{\text{EBS}}}{A_{\text{temp}}} \quad (1)$$

1.3. Influencing factors and sources of uncertainty

Six main factors influence the EP of buildings according to Annex 53 [20]; 1. climate, 2. building envelope characteristics, 3. building services and energy systems characteristics, 4. building operation and maintenance, 5. occupant activities and behaviour, and 6. indoor environmental quality provided. Within each influencing factor, there are various sources of uncertainty affecting the actual performance of a factor. Based on the categories of influencing factors and the types of uncertainties, examples of sources of uncertainty found in the literature are described below.

Table 1
Classification of uncertainties in BPS [17].

Type	Description
Numerical	Computational and numerical imperfections
Modelling	Modellers interpretation and abstraction of reality
Scenario	Boundary conditions, a scenario for building use
Design	Parameters with a specific design decision, physical uncertainty quantifying variability in material properties
Design	Uncertainty in the evolution of the design
Undecided	

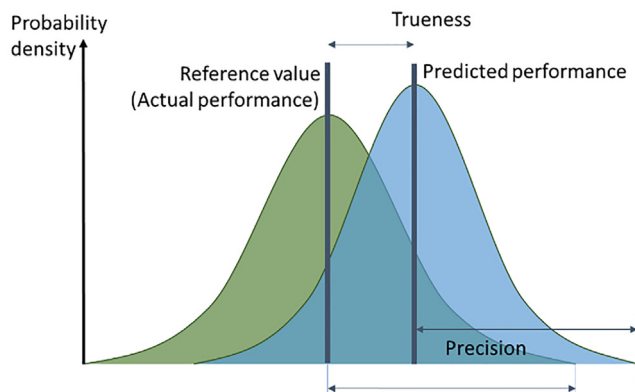


Fig. 1. Illustration exemplifying the trueness and precision of the predicted outcome compared to the reference value (actual outcome) based on terminology from ISO 5725.

1.3.1. The building envelope

The building envelope includes various building elements that consist of several dissimilar materials and layers that affect the thermal transmittance of the building elements. These variations in the individual elements influence the overall thermal performance of the building envelope. In studies investigating and optimising a building's design space, the *design undecided* uncertainty is often included for the building envelope, i.e., by including alternative designs, the external wall's specific design is included as an uncertainty. However, even if the component's design is known, eliminating the *design undecided* uncertainty, the *design decided* uncertainty remains. An example of quantifying the *design decided* uncertainty for mineral wool's material property and a window element's thermal transmittance can be seen in Burke et al. [21]. Using data from laboratory tests - commissioned by the Swedish National Board of Housing, Building and Planning - observations show a discrepancy between the real-world performance of mineral wool batts [22] and thermal transmittance of windows [23] and their declared values. The discrepancy highlights the relative deviation and stochasticity of parameters commonly assumed to be deterministic. Another source of uncertainty is the airtightness of the building envelope. Even when constructed according to specification, the performance depends on the workmanship during the construction process. When measured, significant deviations can be found regarding the airtightness at both a local level, for example, separate apartments of a building, and a building level - which can be seen in Wisth [24]. There are even differences at a building stock level, as shown by Fahlén et al. [25] when evaluating a neighbourhood of 26 identical buildings where the building envelopes were manufactured on an on-site factory. Another often ignored uncertainty is the built-in moisture losses. Two common energy losses can occur due to moisture in building materials. The first, which is considered in this project, is when moisture from the building materials, for example, concrete, evaporates into the indoor air. This process requires energy [26,27]. The second example, which is not relevant to the case study and thus excluded from this study, is when moisture is trapped between two impervious layers, and the moisture evaporates and condenses between the warm and cold sides. One of the earliest energy calculation programs to incorporate this was WUFI Plus, 2003 [28].

1.3.2. The HVAC system

The heating, ventilation, and air conditioning system consists of several systems and products to ensure the indoor climate and EP. This system has several sources of uncertainty. An example of this is the air handling unit (AHU) used for mechanical ventilation in this study's case building, a central unit servicing the whole building with a constant air volume system and a rotating heat exchanger. Connected to the AHU is the forced kitchen hood ventilation function, used when cooking food to increase the airflow and bypass the heat exchanger to avoid contamination. This system introduces a *scenario* uncertainty on how often and for how long each occupant uses this feature. The supply and return airflows are designed and built to specific flows. However, the allowed variations in the measurement and calibration standards [29] introduce *design decided* uncertainties for the actual airflow. Other examples of uncertainties affecting the AHU's performance are the heat exchanger's efficiency, supply and exhaust fan efficiencies, and the temperature setpoint controller for the supply air. The specific fan power depends on the total resistance in the ventilation system, which is dependent on, i.e. the length, number and angle of bends, and smoothness of the ducts and the filter's resistance.

1.3.3. Occupant activities and behaviour

Research has shown that the occupant activities and behaviour significantly impact a buildings' EP. The influencing factors are differences in occupant's culture, lifestyle, occupation, and preferences [30]. Dividing the occupants' impact on the EP of a building into factors introduces the following parameters. The number of occupants in a space, the age and size of the occupants, when the occupants are present in the space, how they move and act in that space, the intensity and type of activity, the heat gain from the occupant, the household electricity, the DHW use, the preferred indoor temperature setpoint, the opening and closing of windows or usage of blinds, and usage of the kitchen hood fan. In a building stock, all these parameters are stochastic and thus challenging to model and predict.

1.3.4. Other factors

The design phase introduces other sources of uncertainty. An overarching and significant factor is the heated floor area, A_{temp} (defined by BBR [19] as the floor area in m² heated to above ten °C), used in the EP normalisation process. The floor area is commonly quantified either by measurements from drawings or as an output from a building's computer model. The output from both methods are dependent on the interpretation and accuracy of the person performing the measurement. The uncertainty regarding a building's actual heated floor area was one parameter that emerged during a BPS competition held in Sweden [31]. During the competition to evaluate which contestant could best predict the actual EP, an unexpected variation in the evaluated building's heated floor area emerged when evaluating the answers [32]. Pasichnyi et al. [33] quantified a significant deviation - ranging from 0 to 120 per cent - in a study where the declared heated floor area for the same objects (15 068 metering points) from two data sources - the database for EP certificate and the district heating suppliers - were compared for inconsistencies. These examples highlight the impact of the modellers' interpretation of the problem and introduce a *modelling* uncertainty source. The deviation significantly impacts the quantification of a building's EP.

Thermal bridges are another example of *modelling* uncertainty. The thermal bridges are sometimes included in BPS using a simplified method, for example, quantified using a standardised value - in per cent - added to the building envelope's overall heat transmission coefficient. Studies by Berggren et al. [33] have shown how imprecise the simplified method of including thermal bridges is compared to calculated values. While using the standard value for thermal bridges is a flawed method, it is advantageous in early phases when details of the building envelope are undecided (introducing the *design undecided* uncertainty). The more precise method of modelling and simulating thermal bridges' geometry requires that the design (or design options) and materials are known while introducing a *modelling* uncertainty regarding how the modeller interprets and approaches the problem, as shown in Berggren et al. [34]. In this study, an experiment was performed to quantify the *modelling* uncertainty for thermal bridges; see section *uncertainty quantification*. These parameters and more were quantified and modelled in this project.

2. Methodology

This study further develops and demonstrates a practical application of using a probabilistic approach to quantify a buildings' EP previously developed by Burke et al. [35] and include the probabilistic approach in risk analysis as evaluated by Ekström et al. [15]. The selected research strategy to evaluate the impact of data quality on the prediction model's accuracy when using probabilistic methods to quantify a buildings' EP was the case study. In the

case study, two datasets with different data quality levels were applied to the same building design and using the method to quantify the effect of data quality, see section 4.1 *Scenarios*. The scope of the current version of the method does not include how to perform design space exploration, although future versions of the method could likely implement this aspect. Methods for exploring the design space has been developed by others, from the manual process by Ekström et al. [36] to more automated processes as by Østergård et al. [37].

The study does not include all uncertainties that impact a buildings' EP. Of the uncertainties described in section 1.1 *Types of uncertainties*, the *numerical* uncertainty was ignored in this study. The *design undecided* uncertainty was eliminated by using a specific building design as an example in the case study presented below. An expanded interpretation of the *design decided* uncertainty was used in this study to include factors from the construction process that could affect the parameter's value, i.e., the airtightness's performance based on workmanship; thus, the expectation was that the results do not predict the measured performance's variation and shape with perfect precision. The sources of uncertainty included for this study's stochastic parameters were identified based on the building contractor's perspective and a project using the Swedish regulatory benchmarking system's design criteria for the EP. According to this system [19], the building's real-world performance must be verified within the first two years of use. Thus, the parameter uncertainties included in this study's case study were limited to those that could occur during the design phase until the first two years of use.

The case study building was selected based on defined criteria while focusing on two key aspects: testing the hypothesis and validating the probabilistic method. For this, a building design used for multiple existing buildings was needed, and where the EP was measured to enable the comparison between predicted and actual EP. Several perspectives, such as the study's aim and other limiting factors, i.e. time, cost, and quality, were considered within these two aspects. The case selection criteria reduced the list to one viable option, with high-quality data regarding the buildings' actual EP, built multiple times based on the same building design, with measured data available for analysis. The selected option was a multi-family building based on a product design built all over Sweden where the EP is measured and already compiled in a database for extraction and analysis. There were no options where multiple buildings with the same design had been built in the exact same location in enough numbers. Instead, a normalisation process was decided on and applied to the buildings' FM since the same design was used in various locations.

The outcome of the case study – the output from the simulations using the two datasets – were validated against observations from FM of the actual EP in buildings built according to the evaluated building design. The evaluation was performed in a double-blind setting. The compilation of data and creation of the input data sets, simulations, verification of the results, and creation of predictive models were performed before the compilation of FM to quantify the actual EP, not to influence the uncertainty quantification process. The FM were used to evaluate the impact of data quality on the predictive model's accuracy. The data for actual EP were compiled from NCC's internal database, containing data from buildings constructed by NCC and otherwise used for the verification process according to the Swedish building regulation (see section 4.3.1 *Field measurements*).

A *normalisation process* was needed to validate and compare the predicted and actual EP by eliminating the effect of unwanted factors from the data. The normalisation process for buildings' EP enables benchmarking various buildings and building designs against standard design criteria. Using a building stock introduces several unwanted effects in the collected FM alleviated by a nor-

malisation process. For the FM, the primary factor to be removed was the objects being in various geographical locations, thus having different climates. Another effect is that the measured data are from different years, and the weather differs from year to year. Thus, the normalisation process was applied to space heating, both simulated and measured; see description in section 4.3.2 *Normalisation process*.

3. Uncertainty quantification and analysis framework

The developed method for uncertainty quantification of input data and uncertainty analysis using probabilistic BPS methods is described below, starting with an overview. The uncertainty quantification process was divided into several steps, illustrated in Fig. 2. The first step was to identify influencing factors for a buildings' EP and, for these factors, define and classify types and sources of uncertainties. The next step was to identify data sources for the uncertainty quantification of the identified types and sources of uncertainty and classify the data sources using the definition of data quality. Next was the uncertainty quantification, using compiled data and models from various sources to define and create the probability density function (PDF) used for sampling input data. The last step was to perform the BPSs and visualise and analyse the results to quantify the probability of failure. More detailed descriptions of the steps are in the sections below, followed by applying the method in section 4 *Case study*.

3.1. Parameter uncertainty - Identification, classification, quantification, and modelling

Identifying parameter uncertainties and dependencies influencing a building's EP throughout the building's life cycle began with a diverging phase. Using a top-down approach where all parts of the building impacting the energy balance were separated and structured based on the project group's prior knowledge and experience and knowledge gathering through literature search and review. The focus then switched to how uncertainties impact the parameters. The uncertainties associated with BPS and modelling parameters can be divided into different classes depending on the source of uncertainty. The classification of uncertainties presented by Rezaee, R. et al. [17] clarifies and defines the different types of uncertainties, as shown in Table 1. The definition of the types of uncertainties was used to classify the sources of uncertainty identified in this study. How the sources of uncertainty were either excluded or included and simplified in the building model is also described. There are sources of uncertainty regarding the actual performance within the influencing factors described above, as exemplified in section 3 *Influencing factors and sources of uncertainty*.

A qualitative method was used to evaluate factors by describing the identified parameters in a master list containing; classification of the parameter divided into influencing factors (1), the phase of the building process in which it occurs (2), and types of uncertainties (3) impacting the parameter. When the master list had been created, the converging phase began, reducing the number of parameters included in the study based on the study's scope, prioritisation, and limitations. The sources of uncertainties for the stochastic parameters were categorised as either main- or sub-parameters. For parameters with sub-parameters, the sub-parameters were combined to quantify the main parameter variation. The parameters' dependency determined how this combination was performed, either through multiplication or addition.

The parameters are included in the building model either internally in the dynamic simulation or externally as a fixed addition. The numerical value for each relevant parameter of the simulations

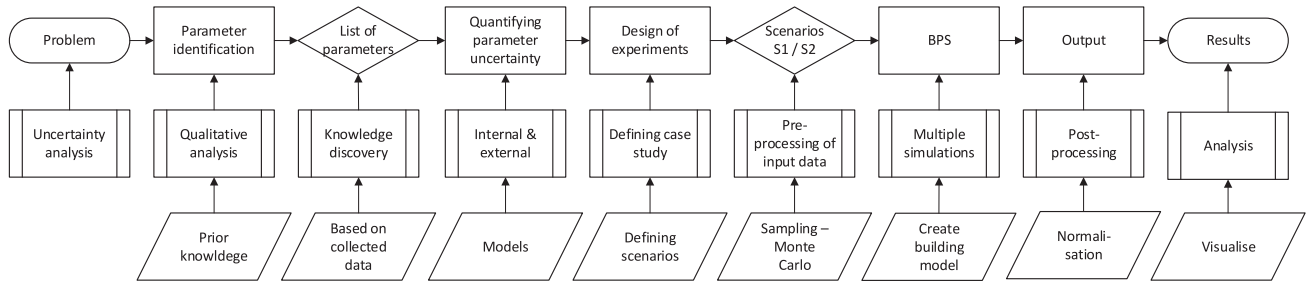


Fig. 2. Overview showing the steps of the method.

were quantified either deterministically, using the same value in all simulations, or probabilistically, using a sampled numerical value from the PDF for each simulation, as shown in eq. (2). Probabilistic parameters (x) will be subscripted by “ i ” (x_i) to indicate changes in values for each iteration. The parameter models that could not be included internally as a dynamic parameter in the simulation model are included as external parameters. These models represent the total energy losses for the specific parameter and are instead added as extra energy losses to the simulated results, as a numerical fixed value sampled from an assigned probability distribution in each iteration simulated. Thus, no share of the heating losses are included as internal heat gains.

$$y = f(x_{m,det,int.} + x_{n,prob,int.,i}) + x_{k,det,ext.} + x_{l,prob,ext.,i} \quad (2)$$

$$m = 1, 2, \dots, n - 1$$

3.2. Uncertainty propagation and sampling method

The probabilistic method requires a large set of outcomes to enable statistical analyses. A process of multiple simulations - using randomised input data for the stochastic parameters (X) and constant values for the deterministic parameters for the iterations (i) - was used to produce the dataset. A sampling method that can handle various input data types - with different units and variations - was needed to use the probability distributions for generating input data for simulations. This study chose the Monte Carlo method because of the focus on quantifying the building design's aggregated uncertainty. The random sampling method was used to maintain the uncertainty propagation throughout the process. Using the random sampling Monte-Carlo technique, a numerical value for each parameter and iteration was sampled from the stochastic parameter's cumulative distribution function (CDF) to create a dataset, represented by the n -by- m sampling matrix X (see eq. (3)).

$$\bar{X} = [X_1, X_2, \dots, X_m] = \begin{bmatrix} x_{1,1} & \dots & x_{1,m} \\ \vdots & \ddots & \vdots \\ x_{N,1} & \dots & x_{N,m} \end{bmatrix} \quad (3)$$

The last step was to use the sampled dataset to evaluate the building model N times, once for each row of the sample (3), creating an input-output map within the input space defined by the uncertainties and for each dataset and compile the obtained output, N possible results, see eq. (4):

$$\bar{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} f(x_{1,1}, x_{1,2}, \dots, x_{1,m}) \\ f(x_{2,1}, x_{2,2}, \dots, x_{2,m}) \\ \vdots \\ f(x_{N,1}, x_{N,2}, \dots, x_{N,m}) \end{bmatrix} \quad (4)$$

3.3. Uncertainty quantification

The next step was to quantify the uncertainties using the list of parameter uncertainties identified in the previous step. The focus was on quantifying uncertainties from all types of uncertainties to showcase how this could be performed, including several sources of uncertainty for the same parameter. The primary method used for uncertainty quantification was the compilation of data from various literature sources. Data triangulation was used when possible when creating this study's models to enhance the credibility of the researcher's data, models, and results. The variation could be estimated with increased confidence by finding multiple sources that quantified the same parameter uncertainty.

Based on the types of uncertainties presented in the section *introduction*, examples of these uncertainties were identified, and models for quantifying the uncertainties were compiled or created. Where no prior model was available, or the dataset was limited, expert estimation was used to determine the variation and shape of the PDF for the parameter, using the design value or standardised data as a basis.

Regarding the *modelling* uncertainty, few examples were found. Instead, this was quantified using an experiment where multiple modellers tried to solve the same problem, using the same software, method, and source material. The difference in output could be attributed to the modeller and their prior knowledge and interpretation of the problem. The aspects influencing the modellers are dependent on the group or company's culture and routines. Thus, this quantification will be case dependent but could give a general indication of the uncertainty.

3.3.1. Data collection, quality, and sources

The data and models were obtained from various sources. The data handling process used in this study to compile, analyse, and model using data is as follows:

1. Define questions and objectives - How and for what will the data be used.
2. Data gathering - performed through a literature search using the databases. The compiled data was prioritised by empirical evidence (1), theoretical calculations or modelling and simulations (2), estimations based on experience (3).
3. Data quality - When compiled and structured, the data's quality needs to be assessed before it is used for further exploration, analysis and modelling used for decisions.

The data validation and verification process in this study for assessing and ensuring the data quality used the following two steps:

- I. Data validation, ensuring that the data is within an acceptable range.

II. Data verification performs a check of the data to ensure that it is accurate, consistent, and reflects the intended purpose.

The data that did not pass the validation and verification steps were excluded from further analysis, i.e. an object's attribute was missing and this could not be solved, then the object was excluded from further analysis. The data was then visualised using different plots, i.e. scatter plots and density plots, to identify any outliers that needed further investigation before being either included or excluded. However, the most common limitation was the lack of data. Passing this step, the data was then used to create a model and implemented in the simulation.

3.3.2. Input data – Stochastic parameters and probability density functions

Performance data (observations) or estimations are compiled and analysed to determine each stochastic parameter's distribution. If performance data is available, the distributions are converted to a probability distribution. If only limited data were available, the data were used in combination with an estimation based on experience to determine the variation and the most relevant distribution. The stochastic parameters included in the study after the identification and information gathering phase was described by; what real-world event or effect the stochasticity model aims to imitate, how the parameter uncertainty was identified, and how it was quantified and modelled. The scaling and shape of the distribution are described and summarised in section 5 Results. If a model for a stochastic parameter was found during the literature search, and the model was deemed applicable to this study's aim, it was implemented and used. If no model was available, a shape was assigned based on the data available. When estimated, the distributions' shapes are simplified into two distributions: uniform and triangular, see Fig. 3. The triangular could be either symmetric or skewed. The triangular distribution was used for most naturally occurring phenomena to simplify the normal distribution. This simplification was used to resolve the issue with unrealistic numerical values for input data, e.g. a negative number of people. The uniform distribution was used when there is a lack of data indicating a more suitable distribution. The skewed triangular distributions exemplify an expected performance that could not be either below or above this level; however, because of various causes, it could be either reduced or increased. It is important to note that the input data were based on a known design (Case Study) and the distributions were for specific material and component properties. The exception was with user-behaviour related parameters since it is complicated to measure these in Sweden due to privacy laws. See 5.1 Parameter uncertainty and distributions for parameter values.

3.4. Building performance simulation

Performing the BPS begins with creating an abstracted version of the building as a model in the BPS tool using a three-

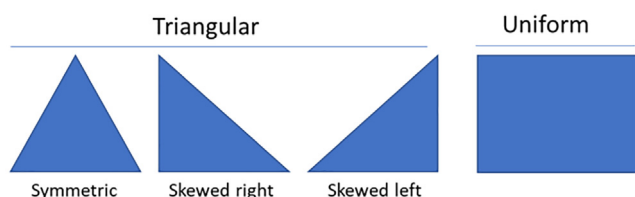


Fig. 3. Illustrates the probability distributions assigned to the input data in the energy simulations of the case study, symmetrical triangular distribution (T_S), skewed triangular distribution, right (T_{SR}) or left (T_{SL}), and uniform distribution (U).

dimensional model-builder and the building characteristics, with a numerical value for all relevant parameters, as for any deterministic BPS. This project used a customised version of the commercial dynamic energy simulation software IDA-ICE, developed by EQUA [38], to enable the probabilistic approach for BPS. The customised version of the software has the option of performing the process of sampling input data from probability distributions, running the multiple simulations, and logging the results required for the probabilistic energy calculations compared to the previously manual labour of performing this process, often being a limiting factor for using the uncertainty approach. The probabilistic approach results in N outcomes – N being the chosen number of simulations – based on variations in input and consequently a variation in outcomes, used as the basis for a probability distribution from which the predictive model is created. The process of including uncertainty is different from the deterministic approach of using a single value for each parameter resulting in a single output value. The performed simulations were dynamic and had a simulation period of one year with a time step of one hour to match the measured data's fidelity and enable the model's validation compared to the FM.

3.4.1. Building model resolution - level of detail

Before creating a building model and implementing quantified uncertainties, an important aspect to decide on is the level of detail or granularity to use. A building model is divided into zones. These zones could be on the whole building level (single-zone model) or divided into smaller zones (multi-zone model), with alternatives such as the whole floor plan, whole apartments, or rooms, and is often based on how parts of the building will be used. The decision about the level of detail for the building model determines the level of detail when deciding the source and model used to quantify the included parameters' input data. In this study, the stochastic parameters were either assigned on a *global level* - with one sampled value for the whole building model - or on a *local level*, with one sampled value for each building model zone. The chosen level of detail and model type determines how to define the input data. An example of how the used level of detail impacts the model is the parameter household electricity, where the standardised value in Sweden for household electricity is $30 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$, according to BEN2 [39]. However, in recent years the use of energy-efficient household appliances has increased, possibly reducing the household electricity compared to the data used as the basis for the standardised value. Instead of using an average value for the whole building, a study by Westin et al. [40] presented a more detailed model for household electricity at the apartment (local) level. This model was based on measured energy use for household electricity from over 3000 apartments in newly built Swedish multi-family buildings. An example of using this model with sampled values at the local level for the case building is presented in Fig. 4. When applying this local level model for household electricity to the case building's local level, the resulting variation and distribution of the building's aggregated household electricity are shown in Fig. 5. The difference between these distributions shows the importance of using assigned distributions and variations based on data and models for the correct level.

3.5. Probability of failure

The use of a probabilistic approach to quantify and predict the EP of buildings enables the quantification of the probability of failure. A method used to quantify the probability of failure, based on probability theory, is the Bernoulli's Theorem [41] using Bernoulli trials, where the experiment is defined so as an event – a set of outcomes – has only two outcomes; these outcomes are either success or failure. Success (p) includes the observations that the event

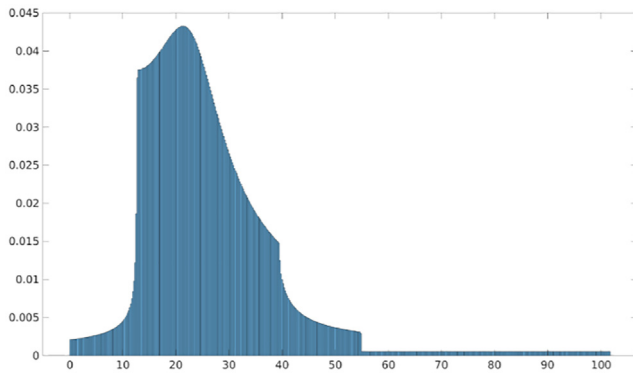


Fig. 4. An example of PDF for the household electricity at a local level was found in Westin [40] and implemented in this study's case building.

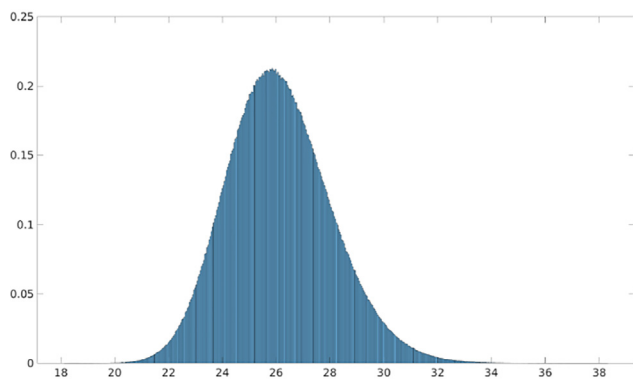


Fig. 5. The aggregated household electricity for the case study building when sampled using the PDF in Fig. 4.

occurred, and the specified conditions were met, and failure (q) includes the observations that a complimentary – unwanted – event has occurred. The probability of these options can then be presented using eq. (5):

$$1 = p + q \quad (5)$$

The probability of either success or failure can be calculated using eq. (5). The unwanted event in this study is the EP not fulfilling the design criteria. Based on this critical condition, the probability of failure needs to be quantified using the aggregated distribution for design options' EP, as described in the study by Ekström et al. [15]; an example of how to determine the probability of failure using the CDF is shown in Fig. 6. A vertical line from the horizontal axis represents the level required to fulfil the design criteria – starting from the design criteria for EP – up to the CDF curve and then a horizontal line out to the vertical axis to quantify the probability of failure (q).

4. Case study

This section describes the case study building, the options for design criteria influenced by the decision-making process, the two scenarios used to quantify the impact of data quality, the quantification of the uncertain parameters included in the scenarios, and the validation process for compiling FM and comparing predicted and actual EP. The building used for the case study was based on a standardised product of a multi-family tower building. This product was built with minor changes all over Sweden, with measured data accessible to the research project. The original val-

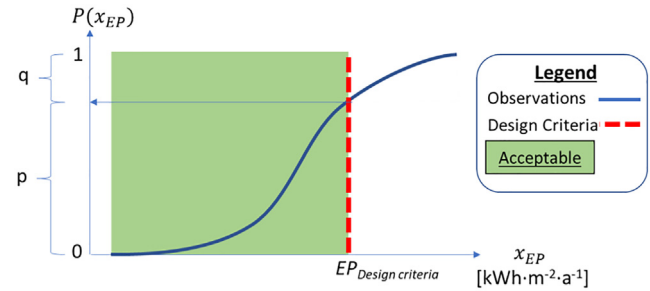


Fig. 6. Illustration of a CDF (blue S-shaped curve) based on observations of EP for a building – either predicted or actual. The vertical axis shows the two outcomes of the Bernoulli trial – success (p) or failure (q) – used to quantify the probability of failure. The level is decided based on the design criteria for EP (red dashed line) on the horizontal axis and the CDF curve. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ues for parameters used in the BPS of the product are shown in Table 2. Electricity for building services include the electricity used by the technical systems for heating, ventilation and DHW, including fans, pumps, electronics, electrical energy input to ventilation systems and heat recovery systems, lighting in common areas, and the elevator. The household electricity is excluded from this comparison, and active cooling is not available in the building.

The design criteria for the building product's EP was initially $80 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$. However, when compiling data from the objects, the data for contracted EP showed objects deviating from the national building requirement and general project requirement with a more stringent level for the design criteria. The identified levels were used as a basis for the evaluation regarding how the choice of design criteria impacts the probability of the building design fulfilling the requirements: National building regulation (dependent on geographical location): $90\text{--}130 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$; General project requirement: $80 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$; Specific project requirement: $65, 70, \text{ and } 75 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$.

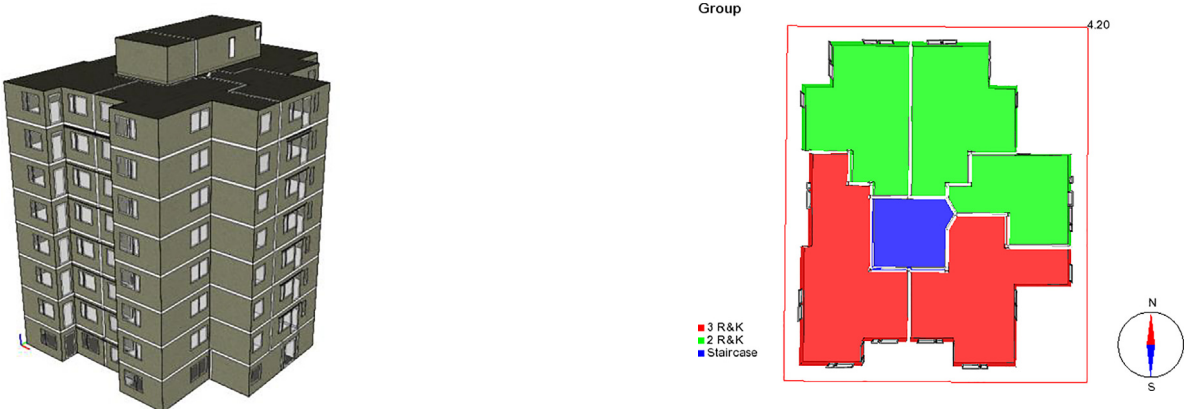
4.1. Scenarios

The two scenarios – S1 and S2 – are used as examples to show the impact on accuracy by using various sources to compile data with differing quality and granularity and by using different models to quantify uncertainties. Thus, whenever possible, the datasets were compiled from different data sources. When a source of uncertainty was only possible to quantify using data from a single data source, this was assigned to both scenarios. It is explicitly stated if both scenarios used the same model and data for parameters. For parameters with the building envelope and building services characteristic decided by the building design, both scenarios primarily used the design value while adding quantified sources of uncertainties to the parameters, thus changing the distribution's scaling and shape.

The first dataset's – scenario 1 (S1) – was defined to showcase a minimum data quality level when simulating a known building design (within acceptable standards). The basis was the standardised input data from the Swedish building regulation's benchmarking system, where the deterministic models for parameters were assigned probability distributions using standardised input data and expert elicitation for the quantification process. Thus, S1 uses a combination of (1) specific design values for the building design and standardised values for parameters with unknown design value – i.e. occupant dependent parameters – from the data sources Sveby [42] and BEN [39], and (2) expert elicitation to estimate the variation and distribution of the parameter based on the value from step 1 and the experience within the project group.

Table 2

Deterministic parameters and design values from the original deterministic model for the object simulated, values in brackets () indicates that the parameters are included probabilistic in this study.

			
Parameter		Deterministic value	
General	Location	Malmö (for simulations, not measured objects)	
	Geographical coordinates	Lat.: 55.592 N, Long.: 13.025 E	
	Floor area	2570 m ²	
	Number of floors	8 + room for central air handling unit	
	Infiltration model	Wind-driven, semi-exposed	
	Building model - Zones	Total: 52 zones, of which: 37 residential zones, 15 non-residential zones	
	Building model - Zone types	Residential: 2 R&K = 21 zones, 3 R&K = 16 zones, Non-residential: Mechanical room, staircase (9 zones), storage room (2 zones), bicycle storage room, laundry room, room for central air handling unit	
Building Envelope components	-	Surface area	Thermal transmittance
	External walls - Heavy	453 m ²	(0.18) W·m ⁻² ·K ⁻¹
	External walls - Light	926 m ²	(0.14) W·m ⁻² ·K ⁻¹
	Roof	301 m ²	(0.11) W·m ⁻² ·K ⁻¹
	Floors towards ground	307 m ²	0.21 W·m ⁻² ·K ⁻¹
	Windows - Main Entrance	3 m ²	1.2 W·m ⁻² ·K ⁻¹
	Windows - Apartments	241 m ²	(1.0) W·m ⁻² ·K ⁻¹
	Doors - glazed	79 m ²	(1.1) W·m ⁻² ·K ⁻¹
	Doors	3 m ²	2.1 W·m ⁻² ·K ⁻¹
Schedules and intensity	AHU schedule	Always on, 100%	
	Occupant activity & schedule	Activity: 0.8 MET, Schedule: Home 14 h / day (00–08, 18–24), away 08–18	
Electricity for building services:	Household electricity schedule	Always on (00–24), at 70% (30% not internal included as internal heat gain)	
	Elevator electricity	Usage: 50 kWh per apartment per year. Lighting: 350 kWh per year.	
	Pump electricity	It is quantified as an addition of 1 per cent of the space heating (dynamic).	
	External lighting	200 kWh per year.	

The second dataset, scenario 2 (S2), aimed to improve the data quality and models for all identified uncertainties, based on specific data gathered from public sources - from studies focusing on quantifying specific parameters - and either implementing existing or creating new data models with probability distributions. The level of detail was also increased, including more parameters on a local level and based on more detailed data.

4.2. Uncertainty quantification and modelling of stochastic parameters - data and models

The stochastic parameters of this study are described below using the above-described process of uncertainty quantification and modelling. Following the process, the description includes what type and source of uncertainty were modelled and what method, data, and data source was used to create the models. Also described was the model's level of detail, the probability distributions, how it differs between the two scenarios and any other lim-

itations applied to the parameter. The quantified uncertainties are summarised in Table 4.

4.2.1. Internal stochastic parameters

Heated floor area, A_{temp} - The source of uncertainty included as a modelling uncertainty is the inaccuracy of measuring the building's heated floor area from drawings. The variation was estimated to ± 2 per cent from the measured design value for the heated floor area. A triangular distribution was assigned and used at the global level. The used variation is smaller than the variation shown in the contest where this uncertainty was identified [31], or the study [33] evaluating the difference depending on the data source used, mentioned earlier, see section 1.3.4 Other factors. However, the effect of using various sources was exemplified by the two scenarios, where the design value differs between S1 and S2 while keeping the relative variation the same. S1 was estimated based on the building model's heated floor area, while S2 was based on the measured floor area from drawings. The difference in design value

between S1 and S2 is significant compared to the quantified uncertainty variation and showcases the difficulty of quantifying the heated floor area depending on its source and how it could impact the results. This parameter uncertainty was only used when normalising the simulations' output, as it usually would occur, and not for the sampling of input data or in the building model for the simulations.

External walls and roofs - mineral wool insulation, thermal conductivity – The cause of the included *design decided* uncertainty is the discrepancy between actual and declared performance for the insulation material's thermal conductivity - mineral wool in this case - quantified by Burke et al. [21] and described in section 1.3.1 *The building envelope*. The triangular distribution was kept and included at the global level. This model was unchanged between the scenarios.

Windows, thermal transmittance – The cause of the included *design decided* uncertainty is the discrepancy between actual and declared performance of the windows' thermal transmittance. The model quantified by Burke et al. [21] was used, as described in section 1.3.1 *The building envelope*, with the triangular distribution and included at the global level. This model was unchanged between the scenarios. The parameter's distribution is simplified so that the same thermal transmittance was used for all types of glazed components, including windows - both fixed and openable - and including glazed doors.

Windows, total solar heat gain coefficient (SHGC_{tot}) – Because this study focuses on a building stock, the simulated building's exact surroundings and the type of shading are unknown, which introduces uncertainties regarding the building's shading. Another uncertainty is how and when the building's occupants use the shading. For this parameter, three types of uncertainties were quantified and included from three sources of uncertainty. The *design decided* uncertainty for the SHGC properties of the building's windows' glazing – the *design undecided* uncertainty regarding the building's external and internal shading type and shape. The *scenario* uncertainty is dependent on how and when the building's occupants use the shading. These uncertainties are included in the BPS through a single numerical value for the parameter SHGC using a simplified model where the percentage of the solar heat gain that is transmitted is estimated. The model used to quantify this uncertainty was simplified to a combined numerical value for the total solar heat gain coefficient (SHGC_{tot}) based on two factors. These two factors were the SHGC of the window glazing (SHGC_{glazing}) and the windows' shading and how it is operated by the occupant ($k_{shading}$). Both were used at the global level and assigned a triangular distribution that was not changed between the scenarios; see Table 4. The total was calculated according to eq. (8).

$$X_{SHGC_{tot},i} = SHGC_{glazing,i} \cdot k_{shading,i} \quad (8)$$

Thermal bridges – The type of uncertainty quantified differs between the two scenarios; for S1, the thermal bridges are included using the simplified method; see description in section 1.3.4 *Other factors* – assumed to illustrate all types of uncertainties of the thermal bridges - as a standard value addition to the overall building envelope heat transmission on the local level. The variation was estimated based on experience and the variation shown in the study by Burke et al. [21]. The thermal bridges are connected to the overall internal area, assigned a uniform distribution and used globally. The type of uncertainty included in S2 originates from the *modelling* uncertainty, intending to quantify the impact caused by the difference in the abstraction of the problem and prior knowledge about how to model and simulate thermal bridges by the modeller. This problem was observed by Berggren [43], who showed significant variations depending on the modellers. The

modelling uncertainty was quantified using the outcome from a semi-controlled test. The test was performed using three modellers who quantified the thermal bridges independently of each other. The modellers were given the same documentation and used the same software for two-dimensional steady-state heat transfer calculations, HEAT 2 [44]. The models and output were reviewed for quality control to ensure no significant errors. The test attempted to quantify and include an additional parameter and another type of uncertainty – *modelling* – based on the resulting variation using the limited sample size from this experiment. A triangular distribution was assigned and used at a global level.

Airtightness of the building envelope (q_{50}) – The *design decided* uncertainty for the airtightness were kept - regarding the variation, triangular distribution and applied at a global level - as in a previous study by Burke et al. [45]. A similar performance variation was presented in [46], where the airtightness was measured in refurbished buildings, indicating the used variation's relevancy.

Specific Fan Power, supply and return – The type of uncertainty modelled was based on the contamination of ducts and filters, increasing the system's internal airflow resistance and lowering the supply and exhaust fans' efficiency. With time, the fans' efficiency will continue to decrease until the ducts are cleaned, and filters are changed, restoring the original efficiency level. The maximum value of the variation represents an AHU system that did not receive service in time, while the lower value represents a new or well-maintained system. The skewed triangular distribution was assigned based on the assumption that the performance cannot be better than clean filters and that most buildings follow the maintenance schedule during the first two years and used globally because of the central AHU.

Heat recovery efficiency – The type of uncertainty included was the *design decided* regarding the heat recovery efficiency parameter. The distribution and variation were kept from a previous study by Burke et al. [21] and not evaluated further. The uniform distribution was assigned based on the lack of data to estimate the shape and used globally because of the central AHU.

Supply air temperature setpoint and the supply-fan air-temperature rise – Modelled as a *design decided* uncertainty to account for the heating systems inability to achieve the decided design value. The temperature variation was estimated, assigned a triangular distribution, and used globally because of the central AHU.

The supply and return airflows – The supply and return airflows used in the building model are based on the building's design value. These values were defined depending on the type of zone, residential or non-residential, and the number of rooms for the residential zones. For S1, the return airflows' design value for residential zones was $30 \text{ l}\cdot\text{s}^{-1}$ and $0,35 \text{ l}\cdot\text{s}^{-1}\cdot\text{m}^{-2}$ for non-residential zones. For the design values of S2, see Table 3. There was an uncertainty regarding the unbalance between supply and return airflows in the buildings in S1. The uncertainty was included in the models; however, the size and precision were not known.

Thus, the parameter variations were based on expert estimation, assigned a triangular distribution and used at a local level. For S2, the airflows' unbalance was based on the designed unbalance; each zone's supply airflow is designed to be two litres per second less than the return airflow. Two sources of uncertainty - the measurement and adjustment uncertainty - were quantified based on the allowed variations in the standards [29] and were included in both scenarios, using eq. (9). Using these uncertainties and the design value for each zone resulted in a specific variation for each zone type used in the model on a local level; see types in Table 3. The measurement and adjustment uncertainty were assigned to all types of zones on a local level.

$$SAF_{Room,i} = SAF_{Room,design\,value} \cdot (1 + x_{Measurementunc,i})(1 + x_{Adjustmentunc,i}) \quad (9)$$

System efficiency, space heating: Plant and heat distribution losses - The case buildings use district heating as the heating source. Thus, a heat exchanger is required, which has conversion losses before distributing the buildings' heating. The heat distribution systems used to bring heat from the source to the point of use – pipes from the plant to the radiators and AHU in the case building – and the radiators' efficiency and ability to maintain the desired indoor climate leads to additional losses. The heat source, distribution system, and local heaters are simplified in the building model as ideal heaters – meaning the system always has enough power to ensure the indoor temperature setpoint without losses – and the heat source (plant) was modelled at an efficiency level which included the conversion losses. By modelling the system this way, conversion losses from the supply to the demand side and losses to distribute the heat throughout the building were not included in the modelled heating system. Instead, these losses were included as a system efficiency, which consists of the two sub-parameter uncertainties for the absolute **plant losses** ($x_{Plantlosses,i}$) and **distribution losses** ($x_{Distr.losses,i}$). The included *design decided* uncertainty variation was estimated and assigned a uniform distribution because of the lack of data and used globally. The system efficiency ($\eta_{p,i}$), is calculated according to eq. (10):

$$\eta_{p,i} = (1 - x_{Plantlosses,i})(1 - x_{Distr.losses,i}) \quad (10)$$

Occupants, number of - The internal heat gain from occupants were estimated at the apartment (local) level. The apartments were classified based on the number of rooms, either two-rooms and kitchen (2R&K) or three-rooms and kitchen (3R&K). Each class's design value was based on Sveby [42] with an estimated variation and assigned a triangular distribution.

Domestic hot water - The current Swedish regulation [39] dictates the expected value of the parameter to be $25 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$, with a possible reduction by up to 10 per cent if energy saving or flow-reducing measures are implemented. Based on this, the parameter's variation was estimated in S1. However, this standardised value has come into question lately by property owners because they seldom measure the DHW to be that high. A recent study [47] showed, based on measured data, that the DHW use can vary between 6 and $23 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$, compared to the standardised level of $22.5 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$, using a reduction of 10% based on the use of energy-efficient fixtures. For S2, a model was created based on the data from the study mentioned above. A uniform distribution was assigned and used on a global level.

Indoor temperature setpoint - The standardised input data in the Swedish building regulation defines the indoor temperature setpoint as $+21^\circ\text{C}$ for residential spaces [39]. When measured, there is a deviation from the standardised value that causes uncertainty regarding the occupants' actual indoor temperature setpoint. An example of this is found in a report by the Swedish National Board of Housing, Building and Planning that showed the average measured indoor temperature of multi-family houses to be 22.3°C , with a variation between 18°C and 24.5°C [48]. However, not all the variation is likely to be caused by the indoor temperature setpoint. This parameter uncertainty was modelled and quantified based on the two types of uncertainties, first was

the *design decided* uncertainty of the thermostats ability to maintain the defined setpoint for the indoor temperature and was used in all spaces and for both S1 and S2. Secondly, the occupant controlled setpoint's *scenario uncertainty* was estimated to $\pm 1.0^\circ\text{C}$, although the study mentioned above [48] could support a more significant variation. This parameter was only used for residential spaces in S2. Both types of uncertainties were assigned a triangular distribution, and the total deviation was determined in residential spaces by addition (the sampled values of both uncertainties were summarised to a total deviation). The design value for the indoor temperature in the residential spaces was changed between S1 and S2. From 21°C in S1, based on the standardised value, to 22°C in S2, based on the mode value in [48]. For all non-residential spaces, 18°C was used in both S1 and S2.

Household electricity - The amount of household electricity used in a building is highly dependent on the appliances available in the space and how the occupants use them. A share of the household electricity is converted into heat and accounted for as an internal heat gain. These variables are of the *scenario* uncertainty type, and an aggregated model quantifying these sources of uncertainty was used in this study. The current Swedish regulation [39] dictates the design value for household electricity to be $30 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$, with a 70% utilisable share for internal heat gains. For S1, the quantified *scenario* uncertainty was based on the standardised value with an estimated variation of $\pm 10 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$. A triangular distribution was assigned and used on a local level. For S2, a specific model – previously shown in section 3.4.1 *Building model resolution - Level of detail* – developed by Westin et al. [40] for quantifying the uncertainty for the household electricity at the local level was implemented. Another source of uncertainty impacting the household electricity was also added to S2, considering the uncertainty of the estimated share utilisable internal heat gain, estimated to $75 \pm 5\%$ based on the outcome from calculations of different scenarios in the study mentioned above.

4.2.2. External stochastic parameters

Built-in moisture losses - The amount of energy used to evaporate moisture from building materials during the heating season was calculated by simulating the amount of moisture lost in all the concrete during the first two years of operation, based on work by Tornberg and Stål [49]. The amount of moisture was then multiplied by the Heat of Vaporization for water. The triangular distribution was assigned based on the calculated estimation and used globally.

Domestic Hot Water Circulation losses (DHWCL) - The domestic hot water circulation system is used in buildings for two reasons. First, to ensure a maximum time until hot water arrives at the tap, regulated in Sweden to 10 s at a flow of $0.2 \text{ l}\cdot\text{s}^{-1}$ [19]. Second, to ascertain that the hot water temperature does not drop below 50°C to prevent the growth of the bacteria *Legionella pneumophila* in the system. Thus, the default solution to fulfil this in multi-family residential buildings is a circulation system. A consequence of using this system is that DHWCL occurs in the building. The system for distributing the DHW and the circulation system results in local thermal losses to the building from the pipes. These losses occur during the entire year; however, it is commonly assumed that the losses occurring during the heating season are usable as an internal heat gain in the building. Thus,

Table 3

S2 design values for the supply and return airflows of the different zone types used.

Zone types	2 R&K	3 R&K	Bicycle room	Laundry room	Storage room	Utility central	Stairwell	Unit
Supply air flow	21	23	10	19	40	15	58	l/s
Return air flow	23	25	10	19	40	15	58	l/s

Table 4

Summary of the stochastic parameters and the quantified types and sources of uncertainty, described by how included in building model (G = Global, L = Local), and the shape of the distribution (T = Triangular, T_{SR} = Triangular, skewed right, U = Uniform, or S = Specific) and the variation (min, design, max values) for both scenarios (S1 and S2).

		Scenarios		S1	S2
Parameter		G/L	Unit		
Building envelope	Heated floor area, A_{temp}	G	m ²	T [2399, 2448, 2497]	T [2519, 2570, 2621]
	Mineral wool, thermal conductivity	G	W·m ⁻¹ ·K ⁻¹	T [0.032, 0.038, 0.044]	T [0.032, 0.038, 0.044]
	Windows, thermal transmittance, U_w	G	W·m ⁻² ·K ⁻¹	T [1.00, 1.10, 1.30]	T [1.00, 1.10, 1.30]
	Shading	G	–	T [0.40, 0.50, 0.60]	T [0.40, 0.50, 0.60]
	Solar Heat Gain Coefficient (SHGC)	G	–	T [0.49, 0.52, 0.55]	T [0.49, 0.52, 0.55]
	Thermal bridges	–	–	–	–
	% of total thermal transmission of envelope	G	%/100	U [0.15, 0.20, 0.25]	–
	External slab / external wall	G	W·m ⁻¹ ·K ⁻¹	–	T [0.12, 0.23, 0.34]
	External wall / internal slab	G	W·m ⁻¹ ·K ⁻¹	–	T [0.08, 0.09, 0.10]
	External Windows & Doors perimeter	G	W·m ⁻¹ ·K ⁻¹	–	T [0.02, 0.03, 0.05]
	Roof / external walls	G	W·m ⁻¹ ·K ⁻¹	–	T [0.07, 0.12, 0.16]
	Balcony floor / external wall	G	W·m ⁻¹ ·K ⁻¹	–	T [0.28, 0.39, 0.51]
	External wall / external wall	G	W·m ⁻¹ ·K ⁻¹	–	T [0.13, 0.14, 0.15]
	Airtightness, q_{50} , external surface at ± 50 Pa	G	l·s ⁻¹ ·m ⁻²	T [0.10, 0.30, 0.50]	T [0.10, 0.30, 0.50]
	Built-in moisture	G	kWh·m ⁻² ·a ⁻¹	T [1.5, 3.3, 5.1]	T [1.0, 2.2, 3.2]
	Specific Fan Power - Supply air	G	kW·m ⁻³ ·s	T _{SR} [0.75, 0.75, 1.05]	T _{SR} [0.75, 0.75, 1.05]
	Specific Fan Power - Exhaust air	G	kW·m ⁻³ ·s	T _{SR} [0.75, 0.75, 1.05]	T _{SR} [0.75, 0.75, 1.05]
	Heat recovery efficiency	G	%/100	U [0.77, 0.80, 0.83]	U [0.77, 0.80, 0.83]
Air handling unit	Return airflow	–	–	–	–
	Measurement uncertainty	L	%	T [0.9, 1.0, 1.1]	T [0.9, 1.0, 1.1]
	Calibration uncertainty	L	%	T [0.9, 1.0, 1.1]	T [0.9, 1.0, 1.1]
	Supply airflow	–	–	–	–
	Unbalanced	L	% /100	T [0.03, 0.05, 0.07]	–
	Measurement uncertainty	L	%	T [0.9, 1.0, 1.1]	T [0.9, 1.0, 1.1]
	Calibration uncertainty	L	%	T [0.9, 1.0, 1.1]	T [0.9, 1.0, 1.1]
	Supply air temperature setpoint	G	°C	T [18.0, 19.0, 20.0]	T [18.0, 19.0, 20.0]
	System efficiency - Space heating	–	–	–	–
	Plant losses	G	%/100	U [0.05, 0.10, 0.15]	U [0.05, 0.10, 0.15]
Heating- and control- system	Distribution losses	G	%/100	U [0.03, 0.05, 0.07]	U [0.03, 0.05, 0.07]
	Domestic Hot Water circulation losses	G	kWh·m ⁻² ·a ⁻¹	T [2.0, 4.0, 6.0]	T [2.0, 4.0, 6.0]
	Indoor temp. setpoint - Residential	–	–	–	–
	Adjustment uncertainty	L	°C	T [-0.5, 0.0, 0.5]	T [-0.5, 0.0, 0.5]
	Occupant controlled setpoint	L	°C	21	T [21, 22, 23]
	Indoor temp. setpoint - Non-residential	L	°C	T [17.5, 18.0, 18.5]	T [17.5, 18.0, 18.5]
	Household electricity	L	kWh·m ⁻² ·a ⁻¹	T [20.0, 30.0, 40.0]	S [2, 23, 102]
	Share utilisable internal heat gain	L	%/100	–	T [0.70, 0.75, 0.80]
	Domestic hot water	G	kWh·m ⁻² ·a ⁻¹	T [15.0, 25.0, 35.0]	U [5.0, 14.5, 24.0]
	Kitchen ventilation losses	G	kWh·m ⁻² ·a ⁻¹	T [1.0, 2.0, 4.0]	T [1.0, 2.0, 4.0]
Occupant	Airing losses	G	kWh·m ⁻² ·a ⁻¹	T [3.0, 4.0, 5.0]	T [3.0, 4.0, 5.0]
	Heat gain - number of occupants	–	–	–	–
	Residential - 2 R&K	L	N/zone	T [1.19, 1.63, 2.07]	T [1.19, 1.63, 2.07]
	Residential - 3 R&K	L	N/zone	T [1.74, 2.18, 2.62]	T [1.74, 2.18, 2.62]

only the losses that occur outside of the heating season are accounted for as a loss. A standardised value commonly used in Sweden to account for the DHWCL is 4 kWh·m⁻²·a⁻¹ [50]. However, recent research has shown that this is often an underestimated loss. According to Kempe [51], calculations show using commonly used thermal insulation levels around the pipes; the variation could be between 6 and 25 kWh·m⁻²·a⁻¹. Another study [50] looking at 12 buildings found the DHWCL to be between 2.3 and 28 kWh·m⁻²·a⁻¹ when measured, and a recent study performing FM of the DHWCL in 134 multi-family residential buildings – from the 1920 s and more recent – found it to be between 0.5 and 76 kWh·m⁻²·a⁻¹ with an average of 6 kWh·m⁻²·a⁻¹ for more contemporary buildings [51]. Modelling the DHWCL based on measured data is difficult because the measured values are often only presented as total annual values, not elaborating on if and how the losses' utilisable share are included. Thus, the variation used in this study is an estimate based on the standardised value and the studies mentioned above. The parameter uncertainty is quantified as the share not utilised by the building. The parameter was assigned a triangular distribution, included globally, and through a fixed addition.

Kitchen ventilation losses - The air from the kitchen ventilation contains various cooking odours and fats that would reduce

the heat exchanger's efficiency if allowed to condense and pose a risk for leaking contaminants to the supply air. The cause of this loss is the activation of the function to bypass the heat recovery (rotating heat exchanger) in the central AHU when in use. The effect is that the heat from the kitchen fan airflow is not recovered and used to pre-heat the supply air. Thus, when in use, the ventilation heat losses and the exhaust fan's electricity use increases since the airflow increases. These losses are not added to the model dynamically. The kitchen ventilation losses are dependent on the following factors; airflow, indoor and outdoor temperature, the duration of use for the function, the number of residential units of the building and adjusted based on the energy that would otherwise have been recovered using the heat exchanger. When calculating the stochastic parameter's variation, the included and quantified uncertainty was the occupant's duration of use. In the base case simulation, the standardised value for the kitchen ventilation use of 30 min per day per residential unit from Sveby was used [42]. The design value was kept while adding a variation in the duration of use, assumed to be ± 15 min per day, to quantify the uncertainty with a triangular distribution.

Airing losses - The airing loss parameter's source is the opening of windows and doors in the building leading to an exchange between indoor and outdoor air. Modelling this parameter is diffi-

cult since the heat loss depends on what assumptions are used, such as the windows' duration and how much the window is open. A study by Hamid and Ibrahimovic [52] found that airing losses could be between 0 and 20 kWh·m⁻²·a⁻¹, although in many cases, they found it to be in the lower parts of this spread. A recent study [53] that measured the time and number of doors and windows that were open using magnetic strips found a correlation between outdoor temperature and the opening of windows (less open windows with lower outdoor temperature) and estimated the impact on the EP of the building to be between 3 and 8 kWh·m⁻²·a⁻¹. Instead of modelling the parameter dynamically with unknown input data, a simplified model - which tried to take all these parameters into account - was used in this project where the total annual airing losses were estimated based on a standardised value. The standardised value for this parameter in Sweden is 4 kWh·m⁻²·a⁻¹ [39]. Based on the standardised value and studies mentioned above, the airing losses' variation was estimated to be 4 ± 2 kWh·m⁻²·a⁻¹, with the triangular distribution, on the global level and added as a fixed addition.

4.3. Validation of results

The process for validating the results from the simulations follows the principles described for data collection in section 3.3.1 *Data collection, quality, and sources*. It consists of compiling field measurements, validating the field measurements, normalising both the simulated and measured data to enable comparison, and comparing the normalised predicted and actual EP. This process is described below and then applied, resulting in the density plots presented in section 5 *Results*. The objects in the field measurements are seen as independent observations from a population of buildings with the same design.

4.3.1. Field measurements

The predicted EP was compared to the actual EP using measured data from the buildings built according to the evaluated design to evaluate the prediction model's accuracy and validate the method. This process produced 90 identified objects, of which 32 remained for further analysis. Of these, 28 objects had high-quality and complete data sets after completing the verification and validation process. The four objects with complete data sets that were removed contained invalid data identified by visualising the data and identifying outlier data points. These objects were further investigated and found to contain invalid data - either measurement errors or missing data in the time series. The measured data from the 28 objects includes hourly data for the space heating, DHW and electricity for building services, compiled using either electricity, energy or flow meters. In addition to the geographical location difference, one other deviation exists between the modelled and measured buildings; for the measured buildings, the number of floors varies while the tower's overall design remains the same. The measured objects have between 5 and 8 floors compared to 8 for the simulated building, impacting the EP and requiring a normalisation process.

4.3.2. Normalisation process

To enable the comparison between predicted and actual EP and to account for the impact on energy used for space heating from differences in weather and climate for buildings in different locations and measured data from different periods, a normalisation process - for both predictions and observations - was needed and performed in steps. Each step's content varies depending on the influencing factors that need to be normalised.

The first influencing factor was the temporal variation of the outdoor temperature boundary condition, how the outside weather varies from year to year depending on the year when

the measurements were performed. This variability was compensated using the average year, where measured values are scaled based on how much the outdoor temperature deviates from a decided average temperature for the period. For simulations, the climate files used were already standardised to an average year. However, for the FM, weighting factors were applied for the measured data to correct how the actual weather during the measured period was compared to the normalised average year.

The second influencing factor to be normalised was the climate difference for various geographical locations, as shown by the average annual outdoor temperature difference in Fig. 7. Normalising for geographical locations was a significant problem since the evaluated buildings were built in various geographical locations, as shown in Fig. 8. In contrast, the simulations for the building was only performed in a specific location, also shown in Fig. 8.

The simulated object's geographical location was Malmö, while the measured objects' geographical locations were all over Sweden, with a variation in outdoor climate. The Swedish building regulation [19] has implemented a method to normalise the energy used for space heating for benchmarking buildings all over Sweden against one specific climate in Eskilstuna by using geographical location factors, F_{geo} . This method and the geographical location factor were used for the normalisation in this study. The factors were applied to both the measured energy use for space heating, $E_{Mea.SH}$, and the simulated energy use for space heating, $E_{Sim.SH}$, to calculate the normalised energy use for space heating used to quantify the EP, as shown by eqs. (6) and (7):

$$E_{Norm.sim.SH} = \frac{E_{Sim.SH}}{F_{geo}} \quad (6)$$

$$E_{Norm.mea.SH} = \frac{E_{Mea.SH}}{F_{geo}} \quad (7)$$

This solution's temporal resolution is currently low, performing the normalisation on the total annual energy use for space heating, which is a significant limitation.

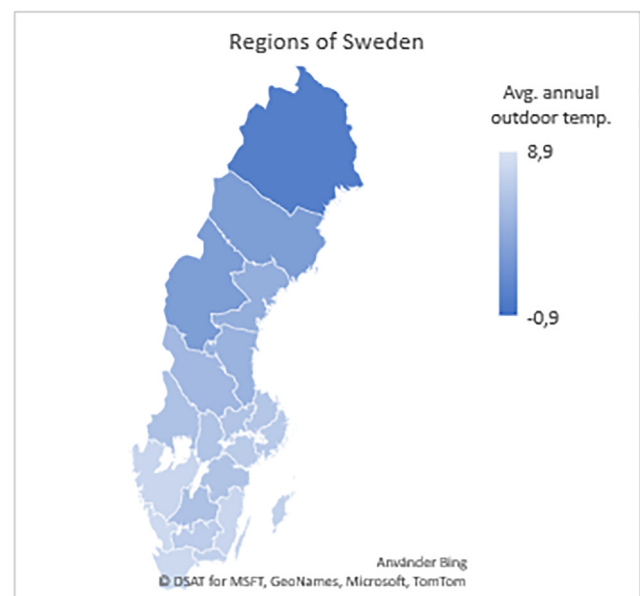


Fig. 7. The map of Sweden shows how the average annual outdoor temperature varies for the regions, illustrated by the colour gradient, based on Sveby and SMHI [54].

5. Results

This chapter begins with a summary of the quantified uncertain parameters used in the simulations. Next, the simulations' results are presented for the two scenarios compared to measured data from the as-built buildings to evaluate the predictive model's accuracy and data quality dependency. An alternate hypothesis was presented and analysed to evaluate how using external parameters impacted the accuracy based on the results.

5.1. Parameter uncertainty and distributions

A summary of all stochastic parameters is presented in Table 4. As mentioned in 3.3.2 *Input data – Stochastic parameters and probability density functions*, the values of the parameters are based on either *design decided*, *modelling* or *scenario* classification and the values (excluding user-behaviour parameters) are based on specific materials and components. The overall variation of some main parameters are based on more than one type of uncertainty; this way of modelling the uncertainty was chosen so that the variation could increase as new types and sources of uncertainties for a parameter were identified, quantified, and included in the study. Twenty-two main parameters were identified, with an additional six sub-parameters for S1 (total 28 variable parameters) and 13 sub-parameters for S2 (totalling 34 variable parameters). Some sub-parameters were assigned to multiple main parameters. The sampled input data includes additional distributions since several parameter uncertainties were included on the local level, sampling for each zone from the probability distribution. The parameters are described by the level of detail included in the building model (L = local, G = global), the units of the variation and the distribution shape and scaling used in each scenario. The distributions are described either by $U[a,b,c]$, a uniform distribution, or $T[a,b,c]$, a triangular distribution, with minimum value a , maximum value c , and the design value (mode) b , or - in specific cases - S , where the shape and distribution were based on models developed in other studies. The specific models, S , are explained in more detail below for the specific parameter.

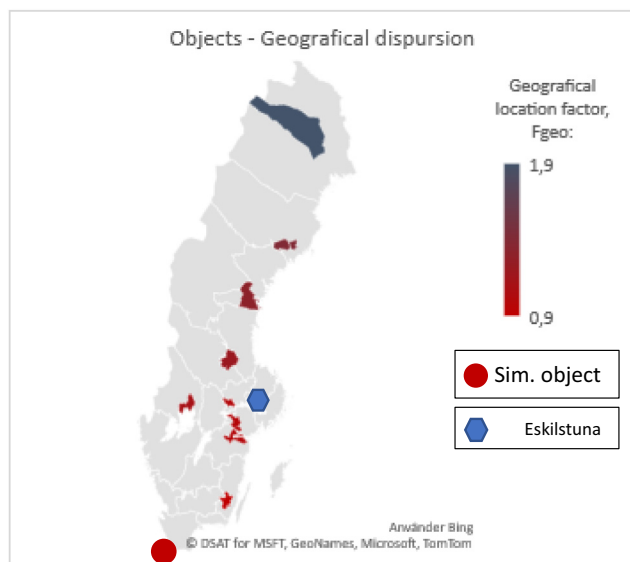


Fig. 8. Map of Sweden shows the 28 objects where measurement data was collected and how the geographical location factors differ, illustrated by the colour gradient. Also shown is the location of the modelled building in the BPS and the location used as the basis for the geographical location factors.

5.2. Accuracy of simulations and the impact on the probability of failure

The simulated results and FM from the objects are presented as density plots and CDFs to evaluate the impact on accuracy from the changed data quality of S1 and S2; see Figs. 9–14. Fig. 14 contains an example of how to determine the probability of failure, represented by vertical lines from the horizontal axis, starting from the design criteria for EP, up to the CDF curves and then horizontally out to the vertical axis. An example of how the decision-making process influences the probability of failure is presented using options for the design criteria, presented in Fig. 14 and Table 5.

The results for the EP show that compared to the actual EP, the predictive model's accuracy increased with S2 compared to S1, regarding both the trueness and precision (see Fig. 9). However, both S1 and S2 deviated significantly regarding the trueness compared to FM. The results for the EP were then divided into the sub-categories, presenting the energy used for space heating (Fig. 10), DHW (Fig. 11) and electricity for building services (Fig. 12), to increase the detail of the analysis and evaluate and identify the potential cause of the deviation. Analysing these categories identified that both the simulated results for electricity for building services and DHW showed similar results to the overall energy use; S2 improved the accuracy compared to S1. The trueness was good while the precision was off regarding the electricity for building services. The results were unclear for the energy use for space heating; initially, S1 showed higher trueness than S2; however, the trueness was still significantly off compared to measured data from the other two categories. The precision for both was still low.

A hypothesis about the cause was proposed: the cause could be the inclusion of external parameters – not included dynamically in the simulations – for the parameters impacting the energy use for space heating: airing losses, built-in moisture, and kitchen ventilation losses. These parameters had been included in the post-processing by adding the sampled value for each parameter to each simulated result. The new dataset was created by subtracting these parameters from the total results of S1 and S2; after the removal, another comparison was made (see Fig. 13). The new distribution for the predicted EP showed a significantly increased trueness compared to the distribution for the actual EP – since the mode of the distribution shifted closer to the mode of the distribution for the actual EP – while the precision was slightly reduced since the variation for the new distribution was reduced.

The probability of failure was calculated using the actual EP based on FMs, and the predicted EP based on the three models: S1, S2, and S2 – without external parameters and quantified using the options for design criteria (see Fig. 14 and Table 5). Comparing the probability of failure quantified using the predictive models compared to the FM indicates that the current predictive models

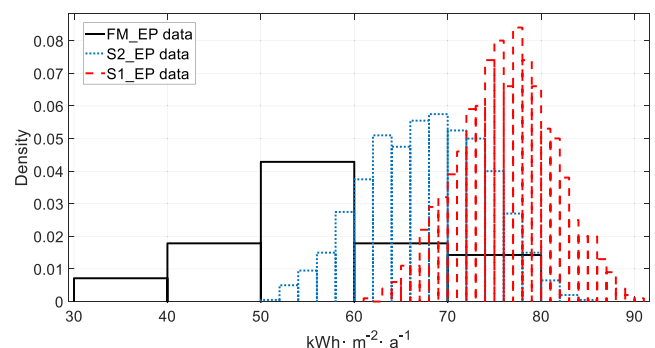


Fig. 9. Density plot of EP based on FM and simulated results from S1 and S2.

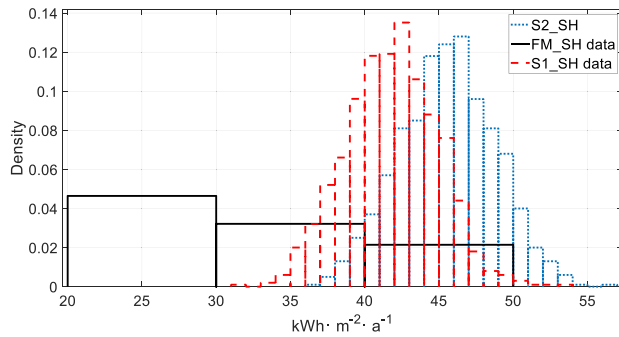


Fig. 10. Density plot of the energy use for space heating based on FM and simulated results from S1 and S2.

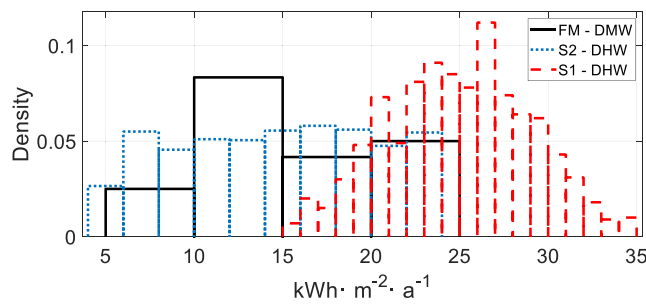


Fig. 11. Density plot of energy use for DHW based on FM and simulated results from S1 and S2.

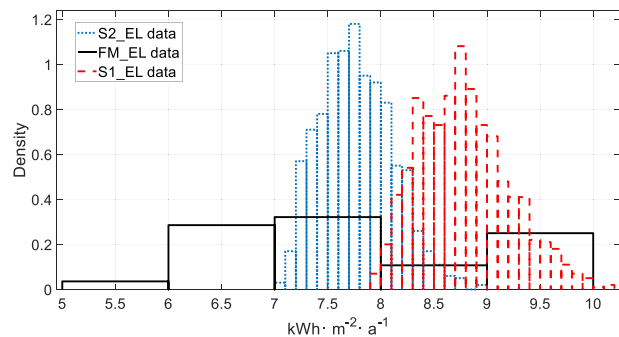


Fig. 12. Density plot of electricity for building services based on FM and simulated results from S1 and S2.

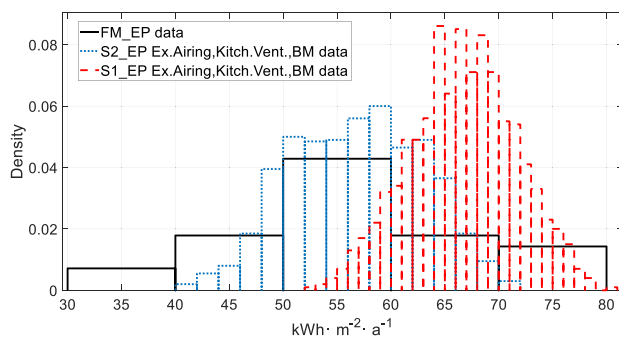


Fig. 13. EP's density plot based on FM and simulated results using S1 and S2 without external parameters compared to measured data, FM.

based on the two datasets are not accurate enough for predicting the probability of failure. S1 and S2 overestimate, while S2 without external parameters underestimates the probability of failure compared to the actual outcome.

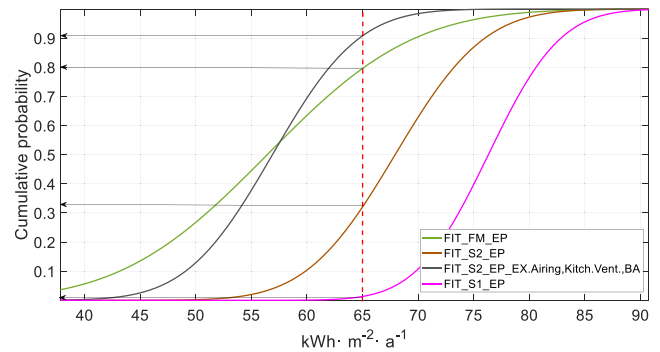


Fig. 14. Graph showing the CDF of EP based on data from FM and simulated results from S1 and S2 without external parameters. The probability of failure is calculated using the CDF, illustrated using the design criteria (dashed red line) and horizontal arrows from where the line crosses the different CDF curves. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5

Probability of failure depending on the chosen design criteria level and based on measurements or simulated results from S1 and S2.

Design Criteria [kWh·m ⁻² ·a ⁻¹]	FM	S1	S2	S2_Excl
65	20%	99%	68%	9%
70	10%	89%	37%	2%
75	4%	60%	13%	0%
80	1%	23%	3%	0%
85	0%	4%	0%	0%
90	0%	0%	0%	0%

6. Discussion

Although the application of the method in this case study was on multiple buildings with the same design with a different number of floors to enable the validation of the results, the primary purpose of the method is to improve future predictions using the probabilistic approach for BPS for individual buildings. For this, further calibration of input data models is required to increase the predictive model's accuracy.

The quantification of the probability of failure provides the foundation for using a probabilistic approach for risk analysis when combined with the consequences of failure, which would provide a data-driven decision support tool for decision-makers during the design phase. Thus, the results from the probabilistic BPS will be used as input for risk analyses regarding the probability and consequence that a building will fulfil the contracted design criteria for EP.

The results of this study provide data and knowledge that could support other practitioners using uncertainty analysis by showing problems with choosing and modelling uncertain parameters, the effect data quality has on the results of the BPS, and the impact on the aggregated EP of a building design and improve future simulations. The significance of these problems is amplified when investigating a single object with a known design. Using a known design eliminates the otherwise more influential design unknown category, which can utilize much larger spreads in the variation of the input models to determine which combinations of components result in a building that fulfils the design criteria. Interestingly that method and input data cannot show the probability of failure.

The variation in EP was significant, both for predicted and actual EP, where the predicted EP varied from 41 to 71 kWh·m⁻²·a⁻¹ the variation of the actual EP was even more significant, from 38 to 75 kWh·m⁻²·a⁻¹. The deviation between simulated results and FM indicates that the variations for uncertain input parameters in S1 and S2 were too conservative. The predictive model's

accuracy could likely be improved by including additional uncertain parameters while not underestimating the variation when quantifying uncertain parameters, as shown by the model for DHW in S2. However, both S1 and S2 deviated significantly regarding the trueness compared to measured data for the space heating.

A possible cause for the low accuracy in the precision between the predicted EP and actual EP could be that no dependencies between occupant dependent uncertain parameters were included. These parameters primarily affect the space heating of the building, which was the category with the lowest accuracy and largest variation. The actual EP and SH showed a larger variation, indicating more extreme variations in the input variables, as discussed before. One solution for this could be to include dependencies, i.e. if there are more occupants in a space, and the intensity per person is the same, the space with more occupants should also have a higher household electricity use, and vice versa. Currently, these parameters were treated as independent, so a zone with few occupants could also have a high household electricity use; although that type of outlier is not an impossible combination, the occurrence is less probable than that of more occupants using more household electricity. A model discussed – but not implemented – was that for each sampled occupant, a sample from an intensity profile for the household electricity per person could be used.

The option of using fixed additions – external parameters that are not included dynamically in the simulations – is commonly used when quantifying EP using a deterministic method, as exemplified by the recommendation of adding airing losses as a fixed addition of $4 \text{ kWh} \cdot \text{m}^{-2} \cdot \text{a}^{-1}$ to ensure that the predicted value does not overestimate the EP of a building. The simplification of including influential factors as a fixed addition is on the safe side if the parameters otherwise would have provided internal heat gains in the actual building resulting in reduced demand for space heating and there is no active cooling available in the building. When using a deterministic method with a margin of safety, the aim is not to predict the actual EP but rather a single value for the EP on the safe side; thus, applying this approach of fixed additions is more suitable. When using a probabilistic approach, as in this study, which aims to predict the possible outcomes as a distribution, fixed additions are less suitable, as the results show. For the probabilistic approach, removing the external parameters improved the predictive model's accuracy, indicating that the input dataset's uncertainty quantification may already include external parameters' effects. This indicates that interdependencies are lost when adding a parameter externally to the BPS result instead of incorporating them directly in the model. Thus, indicating the importance of including all relevant parameters impacting the energy balance of a building dynamically in the BPS. However, from a risk perspective, the predictive model based on the input data model in this study could not accurately predict the probability of failure and underestimated the probability and consequently the risk when these factors were removed. Further studies are needed to try and include these parameters dynamically as internal parameters.

There are two significant limitations with the developed method and its implementation in this study compared to measured observation for validation. Not all uncertainties associated with a parameter are known. Moreover, of those that are known, not all have been quantified. The identified parameters do not include the outlier behaviour often seen when measuring parameters impacted by occupants.

A method that can be applied to update the distributions when new data or knowledge is available is by using the Bayesian Theorem. However, this was not used in this study because most sources do not provide the data in raw forms. The sources commonly presented the data as graphs with a short description of how to interpret them. Finding multiple sources presenting data

for the same parameter and type of uncertainty was challenging to find.

For the input data, a significant obstacle during the uncertainty quantification phase was the last criteria for the data verification; that data “reflects the intended purpose”. During the data collection phase, data sources for the identified uncertain properties of specific parameters were found. However, when data and distributions were investigated further, problems appeared. These problems with data included: that the description of what the data represents was too limited to discern, data based on other types of uncertainties from the same category, data based on experiments performed under significantly different conditions, or data at a different level of detail. All of which led to the exclusion of that data source for that parameter. As exemplified by the floor area difference shown in section 1.3.4 *Other factors*, the source used could significantly impact the results. The difference showcases the importance of using the same source throughout the process since introducing another person's interpretations of a building could significantly alter how a parameter is quantified and modelled.

In specific cases, the compiled datasets contain the same apparent parameters, but the data was based on other types and sources of uncertainties or used a different parameter model. Interpreting the data and ensuring that the data and model fulfil the intended purpose was a significant obstacle. Showcasing the difficulty of quantifying and modelling this parameter is the variability in the data for indoor temperature in Annex 55 [46], which presented measured indoor temperature from several buildings, both single-family and multi-family buildings, from different periods, with various building systems, and in different countries.

Regarding the input data models and the level of detail, although an increased level of detail could increase the output's accuracy, this depends on if the building model's level of detail matches the input data used. If there is a mismatch between the level of detail, known or unknown, this could increase the uncertainty and reduce the model's accuracy. However, other aspects to consider when deciding on the building model's detail level are the needed simplifications when abstracting the real world into a building model. Secondly, the level of detail impacts the overall simulation time and the amount of time needed for the model's diagnostics – both regarding the building model's geometry and the accuracy of the results. Thus, determining the building model's level of detail is a crucial factor that requires deep knowledge of the modeller and balancing act between accuracy and time consumption. Using stochastic parameters adds another complexity level to model the building because of how the parameter distributions are defined. If a parameter model is created based on global level data, it is unsuitable for use at a local level. The aggregated distribution will have a reduced variation if using a global distribution on a local level, consequently reducing the accuracy of the outcome from the BPS. However, it is not always easy to discern at what level the compiled data from sources are. However, for most parameters, the available data are limited and often not at the level of detail needed to model the parameter as intended.

The input data's quality can be impacted by bias, confounding factors, and outliers. These deviations in the input data can be difficult and time-consuming to identify. Because of the need to process the data collected before using it to determine the variation and probability of parameters, there is a risk of bias when interpreting it. There could also be unknown confounding factors that influence the data. One confounding factor identified in the measured data by Burke et al. [21 14] from the 26 single-family houses was how the vestibule was used in at least one of the houses. This vestibule was uninsulated and not designed to be heated. However, in this case, the occupants used the area so that the interior door was open so that the area was heated to the indoor temperature,

increasing the building's energy use for space heating and seen as an outlier in the measured data.

The FM for some of the objects contained errors in the data. One example of this was the electricity for building services, the two lowest and the highest data points deviated from the other clustered data points. A limited evaluation was performed, where a comparison was made against the simulated values of the electricity needed for the central AHU ($>5 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$); this alone was higher than the lowest values ($2.9 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$), indicating a measurement error. For the highest value, some deviations were found; the most significant one was an error in the heated floor area reported for this building. Correcting the floor area improved the performance to match the other clustered data points (going from 13 to $9.3 \text{ kWh}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$). For one object, the DHW was significantly lower than the other objects. Further investigation identified that there were measurement errors in the data that could not be resolved. Two objects had a significantly higher DHW energy usage than expected, and the quality of the measurements could not be verified. Thus, these objects were removed from the analysis.

7. Conclusions

This paper has shown that data quality significantly impacts the predictive model's accuracy, primarily when used for a specific building design to predict the probability of failure. The results also showed that it is feasible to use a probabilistic approach to predict the EP of a building design; however, the quantified uncertainties used in this paper were too limited to give an accurate enough predictive model to use as a decision basis, for example, as input for risk analysis. The results and validation process show the difficulty of applying uncertainty analysis only based on quantified uncertainty in the input data. Without validation and calibration against the uncertainty in actual EP – quantified based on a building stock using the same design – the accuracy of the predicted EP is of low value as a predictor and thus not suitable for use in a risk analysis.

The FM showed that the real-world uncertainties based on known objects with similar designs resulted in a wide range of actual EP that was difficult to model and predict despite being a fixed design with a lot of measured data related to the buildings different energy use being available. While using a fixed design reduces the types and sources of uncertainties included in the data from the field measurements that need to be included and modelled in the BPS for the predictive model to be accurate. The causes for this wide range in EP for the same building design are still unknown and unquantified. Either significant uncertainties were missing in the models, the quantified uncertainties were too conservative, or the uncertainty model for occupant behaviour did not represent the users living in these buildings.

The role occupant behaviour has on the EP of a building is perhaps the most uncertain factors in this study. Since it is challenging to obtain such data in Sweden, the input parameters were based on literature and not possible to validate against the household electricity of the objects. More detailed occupant behaviour profiles may be needed to predict the EP of a building better. More detailed energy mapping is needed to determine the actual energy use dependent on specific factors within the different buildings, such as household electricity. This study shows that it may not be crucial to be conservative when quantifying uncertainties since the predicted range for all three categories – space heating, DHW, and electricity for building services – were more conservative than the FM' data and showed a higher risk of failure. The results are positive from the entrepreneur's point of view since the actual EP was much better than calculated; however, these same results are negative from a cost optimization point of view since the desired performance could have been achieved at a lower cost. This

discrepancy could affect an entrepreneur's chances of winning a bid if the bid is too high due to the building's design being better than required. Thus, further investigation and quantification of uncertainties for *design decided* parameters are needed.

An important conclusion from this study is that despite the level of knowledge about the physical building and building properties, and despite measurement data from 28 objects, the use of standardised data with general uncertainties were not good enough to result in a statistically accurate model in some cases. More object-specific data, S2, closed this gap. Perhaps one error throughout this study was the bias from the authors that data quality was partly related to how large the variation in the distribution of data was, with the assumption that smaller variation in the distributions being better than larger variations. In this study, we showed that the distributions used were too narrow and conservative, indicating faults with the input data for these specific buildings, thus indicating that some of the data quality for this building design would be better with *distributions with larger variation*, in particular with the occupant behaviour related data.

The results also indicate that using the approach with fixed additions for some parameters not included dynamically in the building model – commonly used for deterministic methods to increase the result's margin of safety – may not be necessary when using a probabilistic approach. This is because interdependencies within the model are lost when adding fixed additions to the BPS result. Instead, the focus should be on detailed general data, similar to the household electricity model used in S2, where measured data at a local level was used as a basis for the model. The level of detail and how data was modelled are critical aspects that significantly impact the uncertainty analyses' accuracy.

A necessary continuation to increase the predictive model's accuracy and thus close the performance gap is two-fold. Information on current parameters are complicated to find. Many studies state that they have collected input data and created probability distribution functions for their input parameters, but it is often difficult to discern the models' data, how they used the data, or where they obtained their data. However, without insight into how the data was collected and detailed description and categorisation of the data using the data as a basis for creating input models introduces possible errors in the predictions. This study shows examples of this, both in compiling data and creating input models, as discussed in section 3.4.1 *Building model resolution - Level of detail*, and the results when comparing predicted and actual EP. The design decided approach used in this study is different from studies that focus on a *design undecided* uncertainty in BPS – for which it is easy to generate input data variations – since the distributions include the variation in design values from the possible design options with a uniform shape. The second continuation would be implementing and quantifying additional uncertain parameters not currently implemented in this study. An example could include a measurement uncertainty in simulations to account for uncertainties in the data from FM to increase the accuracy between predicted and actual EP. Another step in the development would be to include parameters using probabilistic models, where for each time step of the simulation, the input value for the parameter was sampled from a PDF. This approach has been tried for window opening and thermostat setpoint by D'Oca et al. [55]. Another approach is to randomly select an occupant profile for each zone based on the monitored occupant behaviour and activity from a database, as performed by Fransson [56].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funding: The project was funded by SBUF, the Development Fund of the Swedish Construction Industry, and NCC AB.

References

- [1] EPBD 2002/91/EC, Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings. 2002, pp. 65–71.
- [2] A. Philippe, "Annex 53 - Total energy use in buildings - Energy performance analysis Separate Document Volume VI," 2013.
- [3] S. Imam, D.A. Coley, I. Walker, The building performance gap: Are modellers literate?, *Build. Serv. Eng. Res. Technol.* 38 (3) (2017) 351–375.
- [4] A. Nilsson, Energy use in newly built apartment buildings on the Bo01 area in Malmö (Energianvändning i nybyggda flerbostadshus på Bo01-området i Malmö) in Swedish, Lund University, Lund, 2003.
- [5] J. Wellholm, "Deviations between calculated and actual energy use in buildings - A case study of a property built, (in Swedish)", Uppsala, 2015, in 2012.
- [6] P. Filipsson, J.-O. Dalenbäck, Energy calculations Deviations between projected and measured energy needs (Energiberäkningar Avvikelser mellan projekterat och uppmätt energibehov), Göteborg (2014).
- [7] P. De Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, *Autom. Constr.* 41 (2014) 40–49.
- [8] H. Yoshino, T. Hong, and N. Nord, "IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods," *Energy Build.*, vol. 152, no. March 2013, pp. 124–136, 2017.
- [9] T. Hong, J. Langevin, K. Sun, Building simulation: Ten challenges, *Build. Simul.* 11 (5) (2018) 871–898.
- [10] W. Tian et al., A review of uncertainty analysis in building energy assessment, *Renew. Sustain. Energy Rev.* 93 (May) (2018) 285–301.
- [11] Y. Jiang, T. Hong, Stochastic analysis of building thermal processes, *Build. Environ.* 28 (4) (1993) 509–518.
- [12] I.A. Macdonald, Quantifying the Effects of Uncertainty in Building Simulation, University of Strathclyde (2002).
- [13] T.D. Pettersen, Uncertainty analysis of energy consumption in dwellings, Norges teknisk-naturvitenskapelige univ. Trondheim (Norway), Norway, 1997.
- [14] S. Sun, K. Kensek, D. Noble, M. Schiler, A method of probabilistic risk assessment for energy performance and cost using building energy simulation, *Energy Build.* 110 (2016) 1–12.
- [15] T. Ekström, S. Burke, L.-E. Harderup, J. Arfvidsson, Proposed method for probabilistic risk analysis using building performance simulations and stochastic parameters, *E3S Web Conf.* Jun. 172 (2020) 25005.
- [16] International Organization for Standardization, ISO 5725: Accuracy (trueness and precision) of measurement methods and results. 1994.
- [17] R. Rezaee, J. Brown, J. Haymaker, G. Augenbroe, A new approach to performance-based building design exploration using linear inverse modeling, *J. Build. Perform. Simul.* 12 (3) (2019) 246–271.
- [18] European Commission, "Directive (EU) 2018/844," *Off. J. Eur. Union*, vol. 2018, no. May 2010, pp. 75–91, 2018.
- [19] Swedish National Board of Housing, Building and Planning (Boverket), "The Swedish National Board of Housing, Building and Planning's regulations on amendments to the building regulations (Sv: Boverkets föreskrifter om ändring i verkets byggregler (2011:6)," pp. 1–16, 2017.
- [20] IEA, Final Report Annex 53. Total energy use in buildings Analysis and evaluation methods, Int. Energy Agency Program. *Energy Build. Communities* no. June (2016) 132.
- [21] S. Burke, J. Kronvall, M. Wiktorsson, P. Sahlin, A. Ljungberg, Method for Probabilistic Energy Calculations – Passive House Case Study, *Cold Climate HVAC 2019* (2018) 645–652.
- [22] Swedish National Board of Housing, Building and Planning (Boverket), "Mineral wool, thermal conductivity."
- [23] Swedish National Board of Housing, Building and Planning (Boverket), "Thermal transmittance of windows."
- [24] J. Wisth, "Airleakage test - Method evaluation and comparison of methods (Täthetsprovning - Metodutvärdering och jämförelse av metoder)," 2012.
- [25] E. Fahlén et al., "Vallda Heberg - Sweden's largest passive house area with renewable energy (Vallda Heberg - Sveriges största passivhusområde med förnybar energi)", in Swedish (2014).
- [26] D. Gawin, J. Kosny, A. Desjarlais, Effect of moisture on thermal performance and energy efficiency of buildings with lightweight concrete walls, *Proc. ACEEE Summer Study Energy Effic. Build.* 3 (2000) 3149–3160.
- [27] H.J. Moon, S.H. Ryu, J.T. Kim, The effect of moisture transportation on energy efficiency and IAQ in residential buildings, *Energy Build.* 75 (2014) 439–446.
- [28] A. Holm, H.M. Kuenzel, K. Sedlbauer, The hygrothermal behaviour of rooms: combining thermal building simulation and hygrothermal envelope calculation, *Eighth Int. IBPSA Conf.* (2003) 499–506.
- [29] SIS Swedish Standards Institute, Ventilation for buildings – Measurement of air flows on site – Methods. 2015.
- [30] D. Yan et al., International Energy Agency , EBC Annex 66 Definition and Simulation of Occupant Behavior in Buildings Annex 66 Final Report, no. May. 2018.
- [31] P. Levin and J. Snygg, "Results from energy calculation competition for an apartment building (Resultat från energiberäkningstävling för ett flerbostadshus)," 2011.
- [32] P. Levin, "Sveby's competition reveals uncertain energy calculations (Svebys tävling avslöjar osäkra energiberäkningar)," 2011.
- [33] O. Pasichnyi, J. Wallin, F. Levihn, H. Shahrokni, and O. Kordas, "Energy performance certificates – New opportunities for data-enabled urban energy policy instruments?," *Energy Policy*, vol. 127, no. October 2018, pp. 486–499, 2019.
- [34] B. Berggren, M. Wall, Calculation of thermal bridges in (Nordic) building envelopes - Risk of performance failure due to inconsistent use of methodology, *Energy Build.* 65 (2013) 331–339.
- [35] S. Burke, J. Kronvall, M. Wiktorsson, P. Sahlin, A. Ljungberg, "Method for probabilistic energy use in residential buildings (Beräkningsmetod för sannolik energianvändning i bostadshus), Swedish (2017).
- [36] T. Ekström, R. Bernardo, Å. Blomsterberg, Cost-effective passive house renovation packages? One-at-a-time optimization vs. Monte Carlo sampling, *Energy Build.* 161 (2018) 89–102.
- [37] T. Østergård, R.L. Jensen, F.S. Mikkelsen, The best way to perform building simulations? One-at-a-time optimization vs. Monte Carlo sampling, *Energy Build.* 208 (2020) 1–13.
- [38] EQUA Simulation AB, "IDA Indoor Climate and Energy (Version 4.8)." p. Detailed and dynamic multi-zone simulation applica, 2016.
- [39] Swedish National Board of Housing, Building and Planning (Boverket), BFS 2017:6 BEN 2 - The National Board of Housing, Building and Planning's regulations on amendments to the National Board of Housing, Building and Planning's regulations and general guidelines (2016: 12) on determining the building's energy use during nor. 2017, pp. 1–16.
- [40] R. Westin, "Household electricity in newly built multi-family buildings (Hushållsel i nybyggda flerbostadshus), in Swedish," 2019.
- [41] "Bernoulli's Theorem," in *The Concise Encyclopedia of Statistics*, New York, NY: Springer New York, 2008, p. 39.
- [42] P. Levin, "Sveby - Branschstandard för energi i byggnader (Sveby - trade standard for energy in buildings)," 2012.
- [43] B. Berggren, Evaluating energy efficient buildings Energy- and moisture performance considering future climate change Document Version : Publisher 's PDF, also known as Version of record Evaluating energy efficient buildings Energy- and moisture performance consider, Lund University, 2019.
- [44] "HEAT2." BLOCON AB, 2016.
- [45] S. Burke et al., Proposed method for probabilistic energy simulations for multi-family dwellings, *E3S Web Conf.* 172 (2020) 25011.
- [46] J. G. Carl-Eric Hagetoft, Nuno M. M. Ramos, N. M. M. Ramos, and J. Grunewald, Annex 55, Reliability of Energy Efficient Building Retrofitting - Probability Assessment of Performance and Cost, (RAP-RETRO): Stochastic Data, 2015.
- [47] E. Andersson and O. Larsson, "Normalization of domestic hot water use Preliminary study (Normalisering av tappvarmvattenanvändning Förstudie-Delrapport Version: 1.4)." 2019.
- [48] Swedish National Board of Housing, Building and Planning (Boverket), "Så mår våra hus - Redovisning av regeringsuppdrag beträffande byggnaders tekniska utformning m.m. (The status of the existing building stock - results from the BETSI project)," 2009.
- [49] J. Tornberg, "The impact of building moisture on energy needs (Bygghetens inverkan på energibehovet)." 2019.
- [50] Bebo, "Mapping of HVAC losses in apartment buildings - measurements in 12 properties (Kartläggning av VVC-förluster i flerbostadshus - mätningar i 12 fastigheter)," 2015.
- [51] S. Burke, J. Von Seth, T. Ekström, C. Maljanovski, and M. Wiktorsson, "Mapping of Domestic Hot Water Circulation Losses in Buildings - Preliminary Results from 134 Measurements," in *E3S Web of Conferences*, 2020, vol. 172, p. 12009.
- [52] A. Abdul Hamid and I. Ibrahimovic, "Energy losses due to ventilation in newly built apartment buildings (Energiförluster på grund av vädring i nybyggda flerbostadshus)," 2013.
- [53] P. Levin, D. Bergsten, and P. Kempe, "Airing behavior and impact on energy performance for a sabo combo house (Vädringsbeteende och påverkan på energiprestanda för ett sabo kombohus)," 2018.
- [54] Sveby and SMHI, "Climate data files for Sweden's municipalities 1981–2010 (Klimatdatafiler för sveriges kommuner 1981–2010, 20160217)," 2016.
- [55] S. D'Oca, V. Fabi, S.P. Corgnati, R.K. Andersen, Effect of thermostat and window opening occupant behavior models on energy use in homes, *Build. Simul.* 7 (6) (2014) 683–694.
- [56] V. Fransson, H. Bagge, D. Johansson, Impact of variations in residential use of household electricity on the energy and power demand for space heating - Variations from measurements in 1000 apartments, *Appl. Energy* 254 (2019) 113599.