energies

MDPI

Article

Sensitivity Assessment of Building Energy Performance Simulations Using MARS Meta-Modeling in Combination with Sobol' Method

Amin Nouri *, Christoph van Treeck and Jérôme Frisch *

Institute of Energy Efficiency and Sustainable Building (E3D), RWTH Aachen University, 52074 Aachen, Germany; treeck@e3d.rwth-aachen.de

* Correspondence: nouri@e3d.rwth-aachen.de (A.N.); frisch@e3d.rwth-aachen.de (J.F.)

Abstract: Large discrepancies can occur between building energy performance simulation (BEPS) outputs and reference data. Uncertainty and sensitivity analyses are performed to discover the significant contributions of each input parameter to these discrepancies. Variance-based sensitivity analyses typically require many stochastic simulations, which is computationally demanding (especially in the case of the large number of input parameters involved in the analysis). To overcome these impediments, this study proposes a reliable meta-model-based sensitivity analysis, including validation, Morris' method, multivariate adaptive regression splines (MARS) meta-modeling, and Sobol' method, to identify the most influential input parameters on BEPS prediction (annual energy consumption) at the early building design process. A hypothetical building is used to analyze the proposed methodology. Six statistical metrics are applied to verify and quantify the accuracy of the model. It is concluded that the cooling set-point temperature and g-value of the window are the most influential input parameters for the analyzed case study.

Keywords: building energy performance simulation; validation; Monte Carlo simulation; meta-model; global sensitivity analysis

Citation: Nouri, A.; van Treeck, C.; Frisch, J. Sensitivity Assessment of Building Energy Performance Simulations Using MARS Meta-Modeling in Combination with Sobol' Method. *Energies* **2024**, *17*, 695. https://doi.org/10.3390/en17030695

Academic Editors: Xiaoshu Lu and Qunli Zhang

Received: 21 October 2023
Revised: 22 January 2024
Accepted: 23 January 2024
Published: 31 January 2024

1. Introduction

Over the past decades, building energy performance simulation (BEPS) has been particularly popular as a major approach to estimating the energy consumption of buildings and, subsequently, reducing the accompanying greenhouse gas emissions. However, the energy performance of buildings depends on a vast number of factors, including the building's characteristics, HVAC systems, control and maintenance, weather conditions, sociological factors, and occupants' behavior [1,2]. Due to the uncertainty in these parameters of a model, huge and sometimes serious discrepancies can arise between BEPS predictions provided by different dynamic BEPS tools and reference data. This mismatch is commonly referred to as the "Energy Performance Gap," which can be classified into three different groups: the gap between first-principle energy models and measurements, the gap between machine learning approaches and measurements, and the gap between predictions and display certificates in legislation [3,4].

Due to the proliferation of numerous BEPS tools, an acceptable degree of confidence in BEPS prediction has become a salient preoccupation for the building simulation community. Validation procedures help to ensure the accuracy and plausibility of the simulation results. The purpose of the validation is to evaluate the quality and reliability of the predicted and calculated performance simulations and to identify errors in the simulation outputs. Subsequently, by using calibration approaches, the uncertainties can be reduced or eliminated, and eventually, simulation results can be aligned with reference data to the greatest extent. Three different kinds of techniques are applied to the validation of dynamic thermal models of buildings: empirical validation, analytical verification, and software-to-software comparative testing [5–7].

To quantify the accuracy of building models, i.e., to characterize the discrepancies between simulation results and real data, statistical performance metrics are used. There are several types of statistical metrics that have been commonly used in the literature [8–22], including the Mean Absolute Error (MAE), Mean Bias Error (MBE), Normalized Mean Bias Error (NMBE), Root Mean Squared Error (RMSE), Coefficient of Variance of Root Mean Squared Error (CV-RMSE), and Coefficient of Determination (R^2). A large number of previous studies have used the mentioned validation techniques and statistical criteria to evaluate the level of prediction accuracy of energy performance models [23–33].

To discover the most influential building model parameters on the simulation outputs (i.e., parameters that contribute enormously to the discrepancies between the energy performance simulation output and the reference data), uncertainty and sensitivity analyses are conducted [34–38]. Table 1 summarizes the validation methods, sensitivity analysis and statistical performance indices used in the reviewed literature in the context of BEPS.

Table 1. Validation methods, sensitivity analysis, and statistical performance metrics found in the literature for the assessment of BEPS, where 'X' indicates methods/metrics involved in the reviewed literature.

Reference	Year	Validation Methods			Sensitivity Analysis	Metrics						
		Empirical	Comparative	Analytical		MBE	NMBE	RMSE	CVRMSE	R^2	MAE	Other
Judkoff et al. [7]	2008	X	X	X	X							
Loutzenhiser et al. [23]	2009	X	X		X						X	X
Tian et al. [32]	2009		X				X		X			
Raslan et al. [39]	2010		X									
Shrestha et al. [40]	2011	X										
Tabares-Velasco et al. [41]	2011		X	X								
Antretter et al. [42]	2011	X	X									
Tabares-Velasco et al. [24]	2012	X	X	X				X				
Spitz et al. [35]	2012	X			X							
Zhu et al. [43]	2013		X									
Kim et al. [44]	2013		X		X							X
Lauster et al. [8]	2014	X		X				X		X		
Mateus et al. [25]	2014	X			X	X				X		X
Strachan et al. [36]	2015	X			X							
Roberti et al. [9]	2015	X			X			X			X	
Alaidroos and Krarti [26]	2016	X				X		X	X			
Sandels et al. [27]	2016	X						X		X		
Ferroukhi et al. [10]	2016	X						X				X
Buonomano [45]	2016		X									
Moronis et al. [46]	2017			X								X
Cattarin et al. [28]	2018	X	X					X				X
Cattarin et al. [37]	2018	X			X							
Zhang et al. [11]	2018	X					X		X			
Nageler et al. [29]	2018	X	X			X		X				X
Kim et al. [12]	2019	X					X		X			
Lee and Hong [13]	2019	X					X		X			
Barone et al. [30]	2019	X	X								X	X
Im et al. [14]	2019	X					X		X			
Halimov et al. [15]	2019	X	X	X				X		X		
Kalogeras et al. [16]	2020		X		X			X		X		X
Gutiérrez et al. [17]	2020	X	X				X		X	X	X	X
Eriksson et al. [18]	2020	X			X		X		X	X		
Nouri et al. [47]	2020		X									
Shi and Chen [19]	2021	X	X				X		X		X	X
Yi et al. [20]	2021	X					X		X			

Table 1. *Cont.*

Reference	Year	Validation Methods		Sensitivity Analysis	Metrics						
		Empirical	Comparative Analytical		MBE	NMBE	RMSE	CVRMSE	R^2	MAE	Other
Alongi et al. [38]	2021	X	X	X							
Catto Lucchino et al. [33]	2021	X	X		X	X	X	X			
Nouri et al. [31]	2021		X		X	X	X	X	X		
Zakula et al. [48]	2021		X	X			X	X			X
Eguía-Oller et al. [21]	2021	X				X	X	X		X	
Magni et al. [22]	2022		X	X	X	X		X		X	X

Sensitivity analysis (SA) methods may be categorized into two major groups: local sensitivity analysis (LSA) and global sensitivity analysis (GSA) [49,50]. LSA is based on one-factor-at-a-time (OAT) approaches, where the inference about the model output variation is made by varying only one model input parameter at a time, and all other parameters are kept constant at their nominal values (or baseline points). Although LSA is computationally efficient [51] and easy to comprehend, it is predicated on the assumption that the model is linear (or at least additive) [52]. Furthermore, it cannot grasp mutual interactions between input parameters and is not feasible to explore multidimensional parameter spaces [49]. Therefore, the robustness and reliability of LSA cannot be guaranteed.

The GSA, on the other hand, investigates the model output where all the input parameters chosen for analysis are varied over the whole parameter space simultaneously. Therefore, GSA imparts further information regarding the correlation between input parameters and may be more reliable. Various methods have been proposed for GSA. The most commonly used approaches are a regression method [53], a screening-based method (such as Morris' method) [34,54,55], and a variance-based method (such as the Sobol' method and Fourier Amplitude Sensitivity Test (FAST)) [35,48,55,56]. Due to the non-linear, non-monotonic behavior of sophisticated building models, GSA has been widely implemented to rank the most important parameters [55,57]. However, the greatest shortcoming of GSA lies in its computationally demanding evaluations, especially in the early stages of the building design process, in which many design parameters are explored and a quick evaluation of the building performance is highly desirable. To implement GSA methods, a large number of stochastic simulations are typically required. For computationally expensive models, comprehensive GSA can be time-intensive and hence increase their cost [51,58].

To tackle the computational burdens of GSA for intricate models such as those used in dynamic BEPS (i.e., physics-based models), this study proposes a metamodeling-based SA method. Meta-models (also referred to as surrogate models or emulators) are approximate, fast models that can be substituted for highly sophisticated, time-consuming models (such as detailed engineering-based energy models) with acceptable reliability [59–61]. Thus, for a large number of simulation runs, meta-models are less computationally demanding in comparison with original, complex building models. A variety of statistical meta-modeling approaches have been employed in the context of BEPS, including regression analysis [62–64], support vector regression (SVR) [65–67], multivariate adaptive regression splines (MARS) [59,68,69], gaussian process regression (GPR) [60,61], Random Forest (RF) [69], neural networks [59,62], and ensemble models [63].

Rodríquez et al. [70] applied a regression method to calculate sensitivity indices using a case study of a dwelling with three ranges of weather data and three levels of occupancy. A sample size of 200 was investigated to create the set of parameters. Bucking et al. [71] proposed a methodology to identify the influential variables (design and operation parameters) on the building's performance (annual electricity consumed). A back-tracking search explored and ranked 8 of 26 variables that have a significant impact on the energy consumption in a net-zero energy house using 6000 EnergyPlus simulation runs. Kim [72] developed two meta-models (gaussian process emulator and polynomial chaos expansion) in MATLAB to alleviate the computational burden of BEPS tools, using a five-story office building located in South Korea. The uncertainty results were propagated by 200 simulation

runs using the EnergyPlus model, which has 47 input parameters. Hester et al. [64] presented a probabilistic framework using a regression-based energy meta-model to estimate the energy consumption of single-family buildings at the early stages of the design process. A total of 17 building attributes were selected, and Monte Carlo simulations were carried out to evaluate the variability in the meta-model output (energy consumption).

An extensive review of validation, uncertainty, and sensitivity analysis studies in the context of BEPS reveals several substantial research gaps. The choice of uncertain parameters for performing the SA is often based on either expert knowledge or existing previous studies, which can seriously impede an accurate and reliable SA. Furthermore, running stochastic simulations with a large number of samples in BEPS tools is very time-consuming or even impractical in the case of complex models. Finally, the application of the MARS meta-model in performing variance-based SA (in particular the Sobol' method) has rarely been carried out in the literature, and further studies are required. This study proposes a novel, four-stage SA framework, including validation, Morris' method, meta-modeling, and Sobol' method. The proposed framework improves the robustness and accuracy of the SA in the early building design process, considering a large set of uncertainties (105 uncertain parameters), to support robust decision-making under input uncertainties. The MARS machine learning algorithm is applied to develop a meta-model for the complex building model, and its prediction accuracy is scrutinized by comparison with three additional meta-models, namely Polynomial Regression (PR), Random Forest (RF), and Support Vector Regression (SVR) based on the radial basis function (RBF) kernel, using several statistical metrics.

The proposed methodology is applied to answer the main research question of this study: Which uncertain input parameters (in the analyzed case) drive the majority of the model output variation? The remainder of this paper is structured as follows: Section 2 introduces the proposed methodology. In Section 3, the investigated case study, the simulation tool used, and the weather data are presented. The results and discussion are described in Section 4. Finally, Section 5 gives concluding remarks.

2. Methodology

The methodology in this study is comprised of two constituent parts. The first part involves the validation approach, which is applied to evaluate and quantify the accuracy of dynamic BEPS results.

Two-time series datasets are created: reference and synthetic datasets. The reference dataset is built from the simulation results of the investigated case study (see Section 3) for one year, containing values of heating and cooling loads at an hourly resolution. To conduct the validation and evaluate the discrepancies, a virtual building is modeled. The virtual building uses the same simulation setup; however, the values of the input parameters, i.e., the thermophysical and optical properties of the building envelope, infiltration rate, internal heat gains, and HVAC system, are randomly varied using Latin Hypercube sampling (LHS). For each set of input parameters, the probability density function is given as a uniform distribution. The simulation results of the virtual building (heating and cooling loads) are collected in a synthetic dataset. The major advantage of the synthetic dataset over the measurement dataset is that there are no measurement errors, and all required information regarding the parameters of the building is attainable.

The main aim of creating the synthetic dataset is to assess the deviation of the simulation results and, subsequently, to inspect the most influential input parameters on the simulation outputs by SA, which will be discussed in the second part of the methodology.

Both qualitative and quantitative validations are performed. In quantitative validation, statistical metrics are applied to ascertain and quantify the mismatch between the simulation results of reference and synthetic datasets (see Section 2.1). The validation approach used in the first part of the methodology is schematically illustrated in Figure 1.

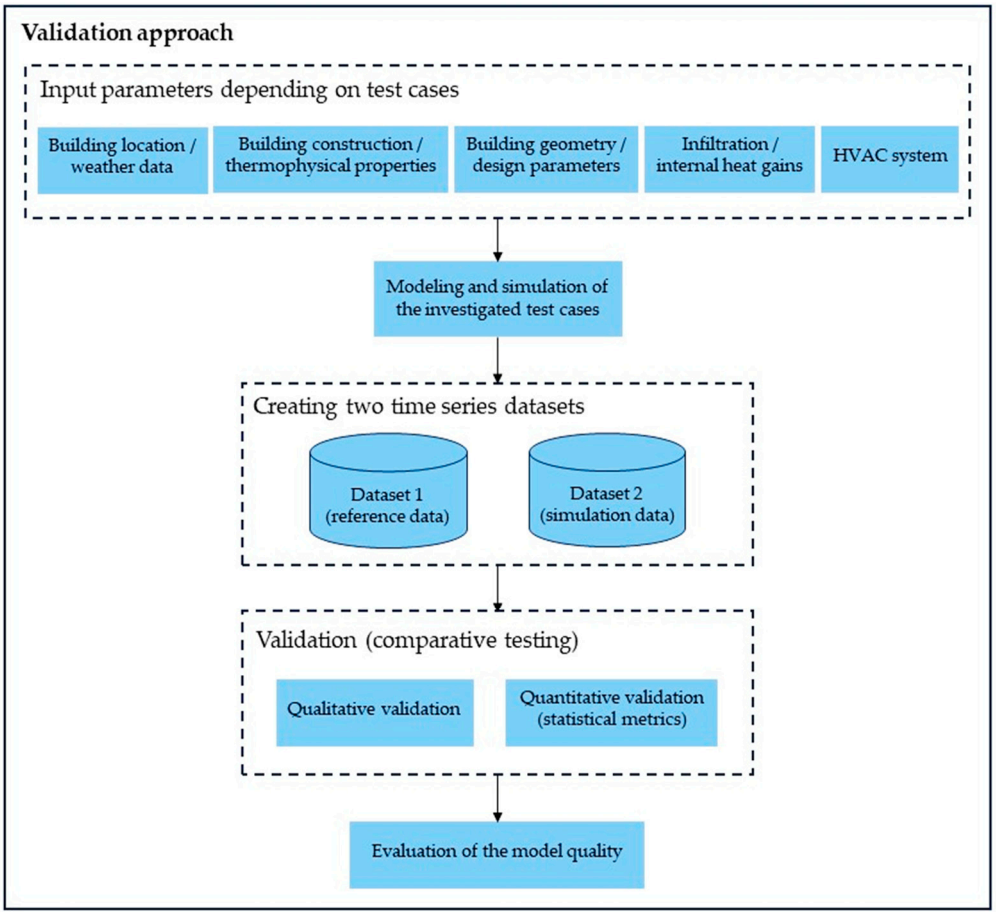


Figure 1. Schematic of validation approach.

The second part of the methodology investigated in this study presents the uncertainty and sensitivity analysis, consisting mainly of three sequential stages: screening-based SA (Morris' method), MARS meta-modeling, and variance-based SA (Sobol' method).

First, the screening-based SA by Morris [73] is performed to determine important input parameters and reduce the number of parameters that are later used in the metamodel-based SA. Since this study investigates a large number of parameters, implementing Morris' method prior to performing variance-based SA (Sobol' method) can substantially reduce the computational cost by selecting key parameters influencing BEPS results. Note that in this study, the term "computational cost" is interpreted as the number of times the investigated model has to be evaluated. 105 uncertain input parameters, including thermophysical properties, infiltration rate, internal heat gains, and HVAC system, are selected. Table A1 (Appendix A) shows the list of input parameters included in the analysis and the assigned uncertainty ranges for each parameter. Due to the large number of input parameters involved in this study, an uncertainty of $\pm 10\%$ is considered for all parameters. It is worth mentioning that the proposed SA framework can also consider parameters with different uncertainty ranges. To build samples of input parameter sets, the factorial sampling plan proposed by Morris [73] is used. All input parameters are uniformly distributed, and it is assumed that there is no correlation between input parameters. Two statistical measures from Morris' method are used to analyze the influence of each parameter on model outputs (annual heating, cooling, and total energy consumption): the mean of the absolute value μ^* and the standard deviation σ of the elementary effects (see Section 2.2). A total of 44 uncertain parameters are selected from an original set of 105.

The main objective of the SA in this study is applying the Sobol' method [74], which is one of the most robust and accurate variance-based SAs [55,75]. However, it requires a large number of samples and simulations, which incur considerable computational costs.

Before conducting the Sobol' method, meta-models are implemented in the second stage of the SA, which can be evaluated faster and reduce the computational cost of Sobol' indices in comparison with the direct simulation of the original BEPS model. To generate a dataset for developing meta-models, Latin Hypercube sampling (LHS) in combination with Monte Carlo simulation is applied in this study.

LHS is a particular form of stratified sampling method and is commonly used for uncertainty and sensitivity analysis in computationally demanding models. Due to its conceptual simplicity and efficient stratification over the range of each uncertain parameter, LHS is very popular [76]. The samples (N = 400) for each of the selected uncertain parameters (k = 44) are created using their probability distributions and stored in a matrix X as:

$$X = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{k,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{k,2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,N} & x_{2,N} & \cdots & x_{k,N} \end{bmatrix} \quad (1)$$

After generating the sample matrix, 400 Monte Carlo simulations (one simulation per row of the sample matrix) are carried out in Dymola (deterministic simulation) using a Functional Mock-up Interface (FMI) for co-simulation [77] to calculate the simulation output Y for each sample as:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} f(x_{1,1}, & x_{2,1}, & \cdots, & x_{k,1}) \\ f(x_{1,2}, & x_{2,2}, & \cdots, & x_{k,2}) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_{1,N}, & x_{2,N}, & \cdots, & x_{k,N}) \end{bmatrix} \quad (2)$$

The FMI provides a definition of an interface standard to facilitate co-simulation and integrate the models into different modeling environments. The models can be encapsulated into a Functional Mock-up Unit (FMU) and used in other simulation tools.

These simulation runs constitute the database, which is randomly partitioned into two sets for training and testing meta-models. The training set (70% of simulation runs) is used to develop meta-models. The remaining 30% of simulation runs are used as a testing set to assess the accuracy and reliability of the developed meta-models.

Four commonly used meta-models, i.e., MARS, PR, RF, and SVR based on the RBF kernel, are assessed. The performances of the developed meta-models are compared by calculating six metrics: MBE, NMBE, RMSE, CVRMSE, R^2, and GOF (see Section 2.1).

In the third stage of the SA, the variance-based SA by Sobol' is applied to impart more quantitative information concerning parameter significance. In the Sobol' method, the contribution of each uncertain parameter to the total variance of the output is evaluated by first-order (S_i) and total-order (S_{Ti}) Sobol' indices using the Sobol' sequence sampling (see Section 2.4). Approximately 23,040 simulation runs are carried out using the MARS model as a meta-model of the original complex Dymola model. The most sensitive parameters to the simulation model output (annual total energy consumption) are identified. The uncertainty and sensitivity analysis approach used in the second part of the methodology is shown in Figure 2.

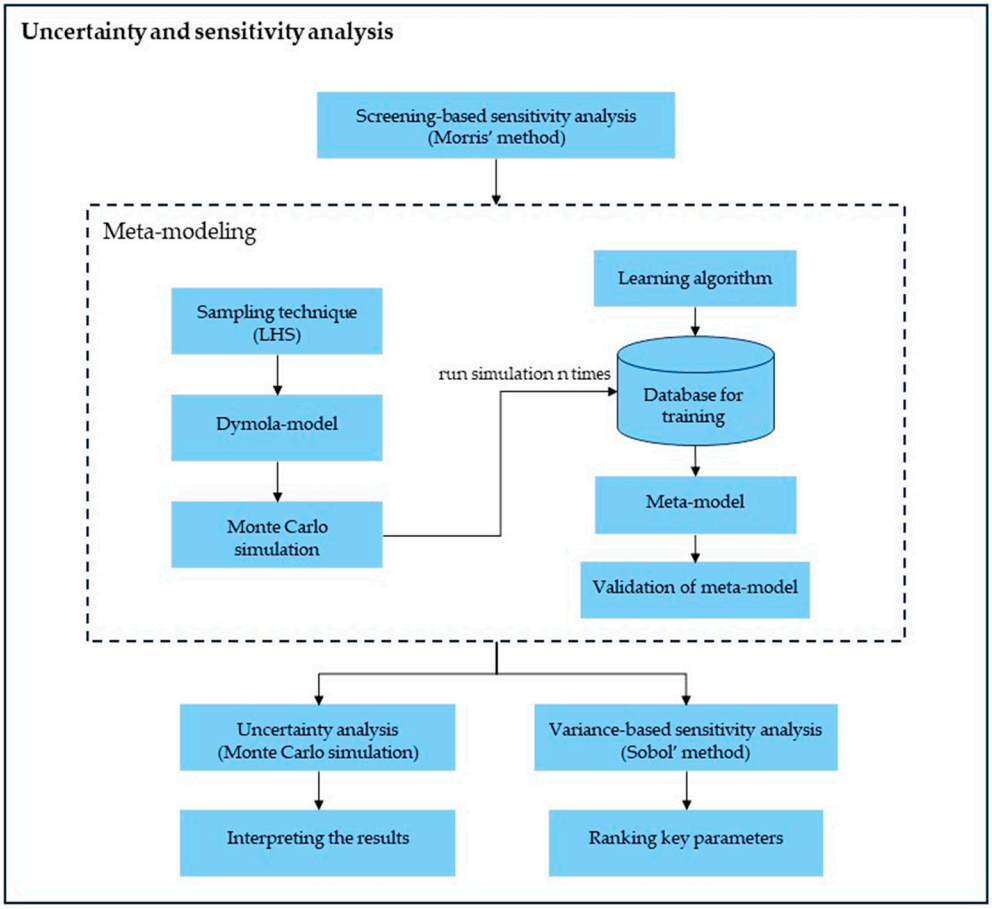


Figure 2. Schematic of uncertainty and sensitivity analysis approach.

2.1. Statistical Performance Metrics

Six statistical metrics are used to verify and quantify the accuracy of the model. Table 2 presents the list of metrics used in this study.

Table 2. Statistical performance metrics.

Metric	Formula	Equation
MBE	$\frac{\sum_{k=1}^{n}(r-p)}{n}$	(3)
NMBE	$\frac{\sum_{k=1}^{n}(r-p)}{n\times\bar{r}} \times 100$	(4)
RMSE	$\sqrt{\frac{\sum_{k=1}^{n}(r-p)^2}{n}}$	(5)
CVRMSE	$\frac{\sqrt{\frac{\sum_{k=1}^{n}(r-p)^2}{n}}}{\bar{r}} \times 100$	(6)
GOF	$\frac{\sqrt{2}}{2} \times \sqrt{\mathrm{CVRMSE}^2 + \mathrm{NMBE}^2}$	(7)
R^2	$1 - \frac{\sum_{k=1}^{n}(r-p)^2}{\sum_{k=1}^{n}(r-\bar{r})^2}$	(8)

MBE indicates the average of the model errors and is calculated using Equation (3), where r is the reference (or original) value, p is the predicted value, and n is the total number of data points. A value of MBE close to zero is desirable. NMBE is calculated by dividing the MBE by the mean of the reference values ($\bar{r}$) and is defined according to Equation (4). RMSE indicates the square root of the average of the squared model errors and is introduced in Equation (5). CVRMSE is a normalized form of RMSE and is calculated

by dividing the RMSE by the mean of the reference values ($\bar{r}$) and is defined according to Equation (6).

The CVRMSE and the NMBE are dimensionless metrics. ASHRAE Guideline 14 [78], Measurement and Verification of the Federal Energy Management Program (FEMP) [79], and International Performance Measurement and Verification Protocol (IPMVP®) [80] prescribe two metrics, NMBE and CVRMSE, for the evaluation of the model's accuracy. The Goodness of Fit (GOF) proposed by ASHRAE 1051-RP [81] combines the two normalized metrics NMBE and CV (RMSE) into a single metric, which is calculated according to Equation (7). The R^2 indicates how well prediction values fit reference values and is described according to Equation (8). The value of R^2 ranges between 0 and 1. The closer R^2 is to 1, the greater the accuracy of the prediction model.

2.2. Morris' Method

A generalized model is given by a mathematical function according to Equation (9):

$$Y = f(X) \tag{9}$$

where X is a $k \times N$ matrix with N samples of k input parameters under consideration (see Equation (1)), and Y is a vector of the corresponding model output (see Equation (2)).

To detect the most important input parameters from among a large number, the screening-based SA proposed by Morris in combination with a factorial sampling plan is implemented. The aim is to determine the subset of parameters that can be fixed at their baseline point values without considerably decreasing the output variance. Morris' method is generally used to discover which input parameters may be considered as (a) negligible or unimportant, (b) linear and additive, or (c) non-linear or involved in interactions with other parameters [82,83]. The considerable advantage of the screening-based SA by Morris is that the assessment of the influence of input parameters can be carried out with a small sample size. In cases of large numbers of parameters, Morris' method is a felicitous screening-based technique prior to performing variance-based SA to alleviate the computational cost [83].

The samples are representative of the parameter space Ω, which is discretized by dividing the range of each parameter into a number of p levels with a predetermined distance of equal size Δ. Each input parameter varies once with step Δ that builds t trajectories of $(k + 1)$ points in Ω. Evaluation of the influence of each parameter on the model output is carried out by calculating a so-called elementary effect (EE), according to Equation (10).

$$EE_i = \frac{f(x_1, \cdots, x_{i-1}, x_i \pm \Delta, \cdots, x_k) - f(x_1, \cdots, x_k)}{\Delta} \tag{10}$$

The computation of each elementary effect requires two sample points. Therefore, the number of required simulation runs (N) is given by $N = t \times (k + 1)$, where t is the number of elementary effects per input parameter and k is the number of input parameters. The values between 4 and 10 are recommended for t [34,73]. Due to the large number of parameters in this study (k = 105), the number of trajectories t = 4 is used. Therefore, $4 \times (105 + 1) = 424$ simulation runs are required to perform the screening-based SA by Morris.

The sensitivity of the model output to each input parameter is analyzed by two statistical measures: the mean value (μ) and the standard deviation (σ) of the elementary effects, as defined in Equations (11) and (12):

$$\mu_i = \frac{1}{t}\sum_{j=1}^{t} EE_i^j \tag{11}$$

$$\sigma_i = \sqrt{\frac{1}{t-1}\sum_{j=1}^{t}\left(EE_i^j - \mu_i\right)^2} \tag{12}$$

where EE_i^j represents the elementary effects of the input parameter i computed along trajectory j (EE_i^j, i = 1, 2, …, k and j = 1, 2, …, t). The mean value of elementary effects (μ) indicates a measure of the overall effect (sensitivity) of an input parameter, while the standard deviation of elementary effects (σ) indicates a measure of interactions of an input parameter with other parameters, i.e., non-linear effects of an input parameter. These two sensitivity measures are basically applied as indicators of which parameters can be considered important [73]. Campolongo et al. [82] proposed using the mean of the absolute value of elementary effects (μ^*), as presented in Equation (13), since positive and negative values of elementary effects can cancel each other out (especially where input parameters are involved in interaction effects).

$$\mu_i^* = \frac{1}{t}\sum_{j=1}^{t}\left|EE_i^j\right| \tag{13}$$

The μ^* is used as an indicator to rank parameters in order of importance.

2.3. Sobol' Method

Sobol' method is one of the variance-based GSAs proposed by Sobol' [74]. It is based on the decomposition of the total model output variance and can consider the main effects as well as the interaction effects of parameters (uncertainty in single parameters as well as uncertainty in combinations of parameters). The key aim of this method is that the contribution of each uncertain parameter to the total variance of the model output is quantified by specific order sensitivity indices, including first-order Sobol' index (S_i), second-order Sobol' index (S_{ij}), …, and total-order Sobol' index (S_{Ti}).

The variance of a model output $V(Y)$ can be defined as a sum of the variances of first-order and higher-order functions:

$$V(Y) = \sum_{j=1}^{k} V_i + \sum_{j=1}^{k}\sum_{i<j}^{k} V_{i,j} + \ldots + V_{1,2,\ldots,k} \tag{14}$$

where V_i represents the variance contribution of first-order interaction effects for the input parameter i and $V_{i,j}$ indicates the variance contribution of second-order interaction effects between two parameters.

The first-order Sobol' sensitivity index (S_i) is written as follows:

$$S_i = \frac{V_{X_i}\left(E_{X_{\sim i}}(Y|X_i)\right)}{V(Y)} \tag{15}$$

where $E_{X_{\sim i}}(Y|X_i)$ represents the mean of the model output Y taken over all values of the sample matrix of input parameters X when the parameter X_i remains fixed and all other input parameters are changed within the defined range. The term $X_{\sim i}$ indicates a matrix of all input parameters, but X_i. The first-order index S_i is applied to determine the most influential parameters on the variance of the model output. However, it does not describe the interactions of an input parameter with other parameters. Thus, a low S_i does not express that the parameter X_i is noninfluential, since there can be higher order interaction effects.

The total-order Sobol' sensitivity index (S_{Ti}) considers all interaction effects of the input parameter X_i with other parameters (first and all higher order interaction effects) and can be defined as:

$$S_{Ti} = \frac{E_{X_{\sim i}}\left(V_{X_i}(Y|X_{\sim i})\right)}{V(Y)} \tag{16}$$

The total-order index (S_{Ti}) is used to determine unimportant (negligible) parameters, i.e., parameters without significantly affecting the variance of the model output.

The sampling from parameter distributions is designed using the Sobol' sequence, which is a quasi-random low-discrepancy sampling technique, to generate uniform samples

of the parameter space Ω. Carrying out variance-based SA by Sobol' method for a model with k input parameters entails a computational cost of $N \times (2k + 2)$ simulation runs, where N is the sample size used for the Monte Carlo simulation [49]. In this study, to perform the robust SA, 23,040 simulation runs are executed using the MARS meta-model.

2.4. Meta-Modeling

The PR is an extension of the linear model that accommodates non-linear relationships between input parameters and simulation outputs. Since polynomial functions of parameters are included in the regression model, it is known as polynomial regression. In this study, the second-order polynomial model is applied, which can be described in Equation (17):

$$\hat{Y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \sum_{j=1,\, i<j}^{k} \beta_{ij} x_i x_j \tag{17}$$

where $\hat{Y}$ is the predicted value, x is the input parameter, and β is the regression coefficient. For k parameters, the second-order polynomial model of Equation (17) will have $p = (k + 1) \times (k + 2)/2$ terms [84].

The RF is a tree-based machine learning algorithm that applies trees as building blocks to construct prediction models [85]. The RF model breaks down the dataset into tree-structure forms and uses averaging to estimate the output of all trees. In this study, the number of trees in the forest was set at 600.

The Support Vector Machine (SVM) is an artificial intelligence (AI) algorithm developed by Vapnik [86]. The goal of the SVM is to minimize the generalization error bound and, furthermore, to obtain generalized performance. That is, the SVM attempts to minimize the deviation (ε) of the prediction function from the target. In this study, the RBF is used as a kernel function [87].

The MARS [88] is a nonparametric, automated regression-based machine learning algorithm that uses labeled samples to model the complex non-linear relationship between input parameters (independent variables) and simulation outputs (dependent variables) and, subsequently, to predict outputs (targets). It can be considered a generalization of stepwise linear regression and has an outstanding ability to tackle high-dimensional problems with a large number of input parameters [62,68,69]. The MARS algorithm can be used to identify the parameters that have a considerable influence on simulation outputs, which is essential in the early stages of the building design process.

MARS uses spline basis functions of the form $(x_v - t)_+$ and $(t - x_v)_+$, defined according to Equations (18) and (19).

$$(x_v - t)_+ = \begin{cases} x_v - t & if\ x_v > t \\ 0 & otherwise \end{cases} \tag{18}$$

$$(t - x_v)_+ = \begin{cases} t - x_v & if\ x_v < t \\ 0 & otherwise \end{cases} \tag{19}$$

where x_v is the v-th independent variable (input parameter), t is the knot of the spline, and the subscript "+" means $(.)_+$ takes the positive part of the inner argument. Each function is piecewise linear, with a knot at the value t. Figure 3 presents the spline basis functions ($t = 0.5$).

The general expression of the MARS model can be described according to Equation (20):

$$\hat{Y} = \hat{f}(X) = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(X) \tag{20}$$

where $\hat{Y}$ is the dependent variable (output) predicted by the MARS model by means of a function $\hat{f}(X)$, β_0 is an initial constant, β_m is the coefficient of m-th basis function, $h_m(X)$ is the m-th basis function, which can be a single spline basis function or a product of

two or more such functions, and M is the number of basis functions, h_m. The basis functions $h_m(X)$ take the form as follows:

$$h_m(X) = \prod_{k=1}^{K_M} \left[s_{k,m} \left(x_{v(k,m)} - t_{k,m} \right) \right] \quad (21)$$

where K_M is the number of knots (number of interaction order in m-th basis function), $s_{k,m}$ has the value of either +1 or -1, which determines the right/left sense of the associated step function, and $t_{k,m}$ indicates the knot location on each of the corresponding variables.

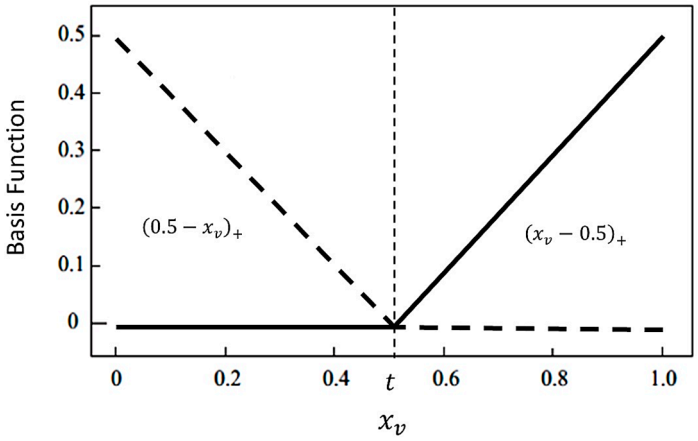


Figure 3. Basis functions (t = 0.5) used by MARS model.

3. Case Study

A hypothetical building with one single thermal zone (shoebox-shaped building) and one singular south-facing window without shading is considered an initial case study. The building has dimensions of 2 m in width, 5 m in length, and a floor-to-ceiling height of 3 m, totaling its interior volume at 30 m^3. The window is centered in the southern wall and has a total surface area of 3 m^2 (1.5 m $\times$ 2 m), as shown in Figure 4. All surfaces are modeled as exterior walls. The material properties of the building envelope used in this study can be found in Table 3. The same boundary conditions are applied to all surfaces, and an initial temperature of 20 °C is defined for the interior air volume and wall instances. The infiltration rate of 0.5 air changes per hour is assumed for the building air volume. No heat recovery is considered. The internal heat gains are set to 200 W. The radiative and convective portions of internal heat gains are taken as 60% and 40%, respectively.

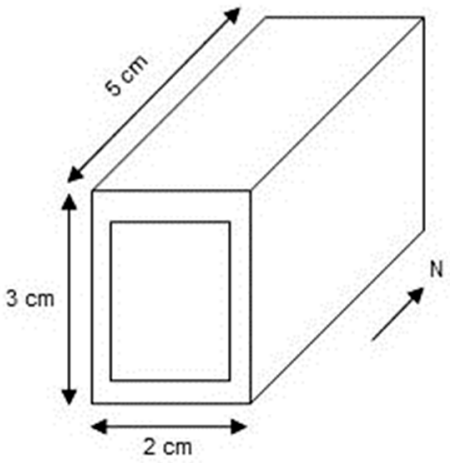


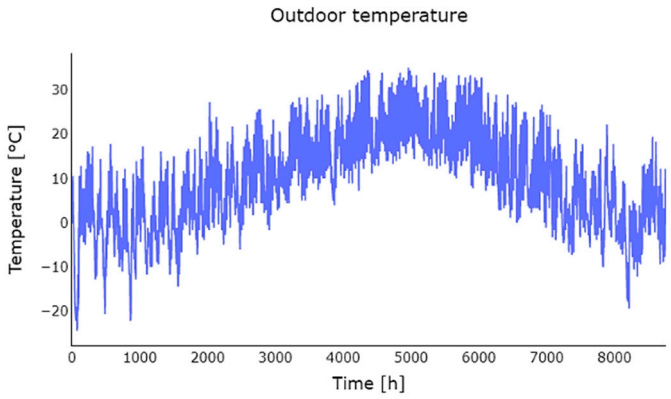
Figure 4. Geometry of the analyzed building.

Table 3. Material properties of the building envelope.

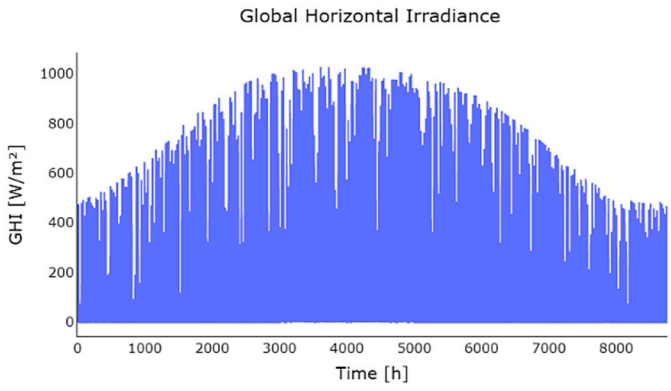
Layer Nr.	Components (Outside to Inside)	Thickness d [m]	Thermal Conductivity $\lambda \left[\frac{W}{m \cdot K}\right]$	Density $\rho \left[\frac{kg}{m^3}\right]$	Specific Heat Capacity $C \left[\frac{J}{kg \cdot K}\right]$
1	Plaster	0.005	0.14	800	1500
2	Thermal insulation	0.100	0.04	50	1000
3	Concrete	0.200	1.20	2000	1000

The building is assumed to be heated and cooled by an ideal HVAC system. The maximum heating power of the HVAC system is limited to 1000 W, while the maximum cooling power is capped at 1200 W. The heating set-point temperature is 20 °C, and the cooling set-point temperature is set to 28 °C. The proposed case study is modeled in Dymola using the Modelica® language (Version 3.2.2) [89] and simulated for an entire year using a time step resolution of one hour. The investigated building model is controlled based on the heating/cooling set-point temperature.

For this analysis, the simulations are carried out using the weather data of Denver-Stapleton (Denver, CO, USA), which is a cool-dry climate (zone 5B) according to the ANSI/ASHRAE/IES Standard 90.1 [90]. The original weather data were provided in the EnergyPlus Weather (*.epw) format, which is converted to the Modelica timeseries script (*.mos) before running the simulation in Dymola. The climate data consists of hourly values of solar radiation and meteorological conditions for a 1-year period. Figure 5 shows the outdoor dry bulb temperature and the global horizontal irradiance (GHI) for one year.



(**a**)



(**b**)

Figure 5. (**a**) Outdoor dry bulb temperature [°C]; (**b**) global horizontal irradiance [W/m^2] used in Dymola.

4. Results and Discussion

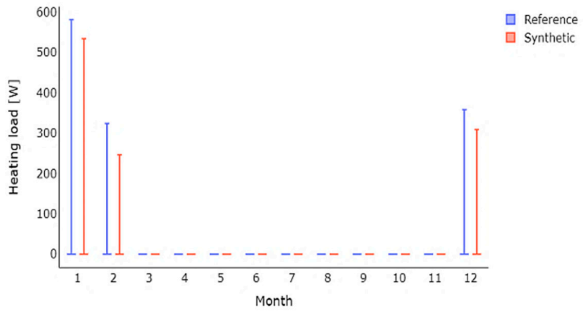
This section presents the results of the application of the proposed framework with respect to the investigated case study of a building with one single thermal zone.

4.1. Validation of Deterministic Simulations

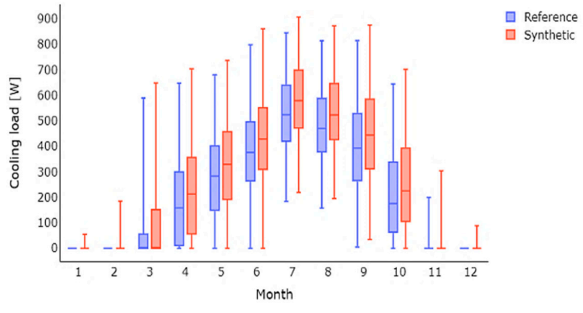
The scope of this section is to assess the results of two deterministic simulations (reference and synthetic), i.e., to reveal the inconsistencies between the two simulation results. In the reference simulation, the aforementioned case study was used to carry out the simulation. In the synthetic simulation, the virtual building model, which has different input parameters related to the thermophysical and optical properties of the building envelope, infiltration rate, internal heat gains, and HVAC system, was simulated.

Both deterministic simulations are performed for one year at an hourly resolution using the same simulation setup.

Figure 6 compares the results of heating and cooling loads. It can be qualitatively observed that there are discrepancies between reference and synthetic models for both heating and cooling load predictions. There are no heating loads from March to November. The maximum heating load predictions have values of 581.7 W and 534.5 W (in January) for the reference model and synthetic model, respectively. The cooling load predictions have maximum values of 906.8 W and 844.9 W (in July) for the synthetic model and reference model, respectively.



(**a**)



(**b**)

Figure 6. Deterministic simulation results of heating (**a**) and cooling (**b**) load [W] calculated with hourly dynamic simulation of reference and synthetic models.

To quantify and ascertain the discrepancies between reference and synthetic models, six statistical metrics, described in Section 2.1, were used. The results of the investigated performance metrics for the cooling load are shown in Table 4. It indicates deviations in reference and synthetic simulation results. To discover the most important model parameters that contribute enormously to these discrepancies, a meta-model-based SA was conducted.

Table 4. Performance evaluation for cooling load.

Metrics	Comparison of Reference and Synthetic Models
MBE [W]	−32.28
NMBE [%]	−15.18
RMSE [W]	43.10
CVRMSE [%]	20.27
GOF [%]	17.91
R^2	0.97

4.2. Sensitivity Analysis Using Morris' Method

The screening-based SA using Morris' method was performed to determine the most influential input parameters on annual heating, cooling, and total consumption. A total of 105 uncertain parameters, including thermophysical and optical properties, infiltration rate, internal heat gains, and HVAC system, were selected. The factorial sampling plan proposed by Morris was applied to build 424 samples of the input parameter.

Figure 7 illustrates the results of Morris' sensitivity analysis of building energy simulations for the investigated case study.

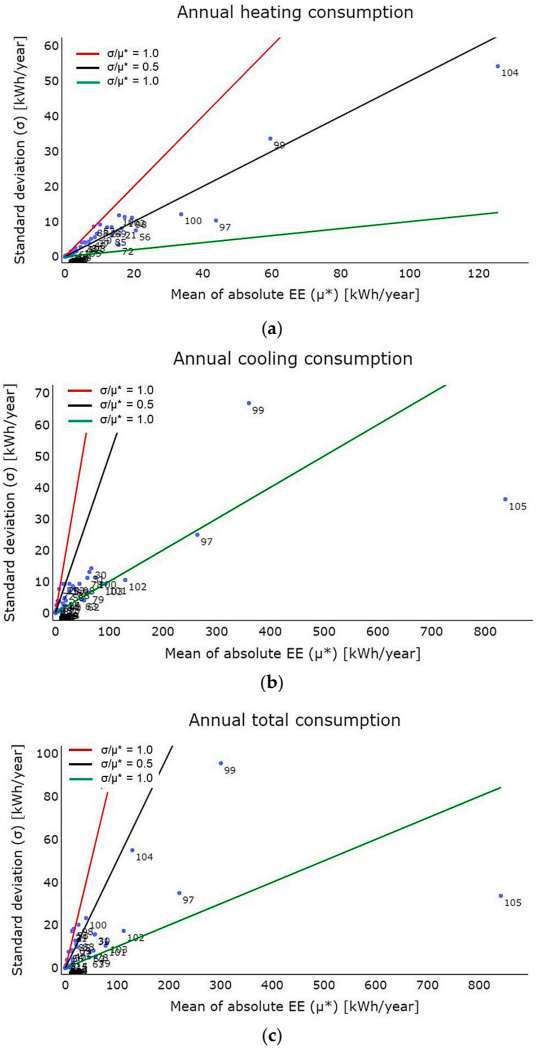


Figure 7. Morris' method—sensitivity analysis of 105 parameters for annual heating (**a**); cooling (**b**); and total (**c**) consumptions.

Each point in scatter plots indicates the impact of each input parameter included in the investigation on the selected simulation output. The horizontal axis represents the mean of the absolute value of the elementary effects (μ^*), which indicates the overall importance of an input parameter. The vertical axis represents the standard deviation of the elementary effects (σ), which describes the total interactions of an input parameter with other parameters involved in the analysis (non-linear effects of parameters). It is also worth pointing out that according to Equation (11) the elementary effect has the same unit as the simulation output, which is the annual energy consumption (kWh/year) in this study. The ratio σ/μ^* is used as an indicator for the assessment of the linearity (and monotonicity) of parameters. The ratio $\sigma/\mu^* = 0$ is considered for the linear behavior of parameters. Three slope lines ($\sigma/\mu^* = 0.1$, 0.5, and 1) can be plotted in the scatter plot (Figure 7) to classify parameters into four zones: almost linear (below $\sigma/\mu^* = 0.1$), monotonic ($0.1 < \sigma/\mu^* < 0.5$), almost monotonic ($0.5 < \sigma/\mu^* < 1.0$), and non-linear/non-monotonic ($\sigma/\mu^* > 1$) [54].

The mean of the absolute value of the elementary effects (μ^*) is used for ranking the input parameters in accordance with their importance. The higher the μ^* value, the more important (and sensitive) the parameter. Figure 8 shows the ranking of 44 input parameters from an initial set of 105 for the simulation output of annual total consumption. Table 5 shows the selected input parameters as well as their rank with the corresponding μ^* and parameter index. In this analysis, cooling set-point temperature was identified as the most influential parameter with the highest value of μ^*, having a monotonic effect on the annual total consumption, followed by the g-value of the window, window area, and heating set-point temperature.

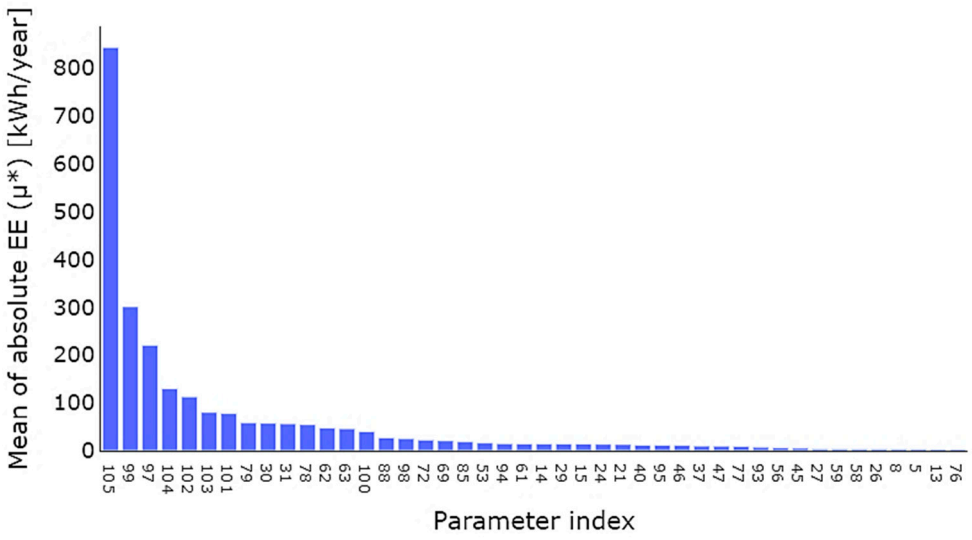


Figure 8. Ranking parameters by Morris' method, considering an uncertainty of ±10% for all parameters. Note that in this study, thermal bridges are not considered.

4.3. Meta-Modeling

In this section, four meta-models (MARS, PR, RF, and SVR) based on the RBF kernel were developed to evaluate their performance. LHS, in combination with Monte Carlo simulation, was used to generate a dataset for developing meta-models for annual total consumption.

Trading off the accuracy and the computational cost, 400 samples for each of the 44 uncertain input parameters (see Table 5) were created using uniform probability distributions. It is assumed that there is no correlation between input parameters. These simulation runs constitute the database, which is randomly partitioned into two sets for training (70% of simulation runs) and testing (30% of simulation runs) meta-models.

Six statistical metrics were applied to compare the performance of the developed meta-models. Table 6 shows that MARS considerably outperforms all other meta-models with respect to annual total consumption predictions.

Table 5. Selected input parameters and their μ^* by Morris' method.

Rank	Parameter	Unit	Index	μ^*
1	Cooling set-point	[°C]	105	841.97
2	g-value window	-	99	300.56
3	Window area	m^2	97	220.35
4	Heating set-point	[°C]	104	129.51
5	Internal heat gain (radiative)	W	102	112.71
6	Albedo	-	103	80.10
7	Internal heat gain (convective)	W	101	77.70
8	Solar absorptance ceiling	-	79	58.59
9	External convective coefficient wall west	$W/(m^2 \cdot K)$	30	57.46
10	Solar absorptance west	-	31	56.35
11	External convective coefficient ceiling	$W/(m^2 \cdot K)$	78	54.48
12	External convective coefficient wall east	$W/(m^2 \cdot K)$	62	47.37
13	Solar absorptance east	-	63	46.00
14	Infiltration rate	1/h (ACH)	100	40.04
15	Thickness floor layer 2	m	88	27.09
16	U-value window	$W/(m^2 \cdot K)$	98	25.39
17	Thickness ceiling layer 2	m	72	22.43
18	Conductivity ceiling layer 2	$W/(m \cdot K)$	69	20.78
19	Conductivity floor layer 2	$W/(m \cdot K)$	85	18.80
20	Conductivity wall east layer 2	$W/(m \cdot K)$	53	16.28
21	External convective coefficient floor	$W/(m^2 \cdot K)$	94	14.77
22	Internal convective coefficient wall east	$W/(m^2 \cdot K)$	61	14.35
22	External convective coefficient wall south	$W/(m^2 \cdot K)$	14	14.35
24	Internal convective coefficient wall west	$W/(m^2 \cdot K)$	29	14.25
25	Solar absorptance wall south	-	15	14.04
26	Thickness wall west layer 2	M	24	13.58
27	Conductivity wall west layer 2	$W/(m \cdot K)$	21	13.05
28	Thickness wall north layer 2	M	40	11.65
29	Solar absorptance floor	-	95	11.50
30	External convective coefficient wall north	$W/(m^2 \cdot K)$	46	11.21
31	Conductivity wall north layer 2	$W/(m \cdot K)$	37	9.94
32	Solar absorptance wall north	-	47	9.72
33	Internal convective coefficient ceiling	$W/(m^2 \cdot K)$	77	9.03
34	Internal convective coefficient floor	$W/(m^2 \cdot K)$	93	7.48
35	Thickness wall east layer 2	m	56	6.27
36	Internal convective coefficient wall north	$W/(m^2 \cdot K)$	45	5.54
37	Specific heat wall west layer 3	$J/(kg \cdot K)$	27	4.12
38	Specific heat wall east layer 3	$J/(kg \cdot K)$	59	4.03
39	Density wall east layer 3	kg/m^3	58	3.72
40	Density wall west layer 3	kg/m^3	26	3.71
41	Thickness wall south layer 2	m	8	3.52
42	Conductivity wall south layer 2	$W/(m \cdot K)$	5	3.30
43	Internal convective coefficient wall south	$W/(m^2 \cdot K)$	13	2.74
44	Thickness ceiling layer 3	m	76	2.55

Table 6. Performance evaluation of meta-models.

Meta-Model	MBE [kWh]	NMBE [%]	RMSE [kWh]	CVRMSE [%]	GOF [%]	R^2
MARS	−0.73	−0.04	33.60	1.74	1.23	0.99
PR	−13.32	−0.69	52.48	2.71	1.98	0.97
RF	−7.94	−0.41	104.32	5.39	3.82	0.88
SVR (kernel: RBF)	−9.28	−0.48	115.56	5.97	4.24	0.85

The R^2 values for MARS and PR are very close together at 0.99 and 0.97, respectively. However, the MBE value for MARS (−0.73 kWh) differs significantly from that for the PR model (−13.32 kWh). It can be interpreted that there may be a bias (i.e., systematic error) in the predicted values of the PR model. The negative value of MBE describes that the average of predicted values of the meta-model during the time period is higher than the average of Dymola values, which indicates the overprediction of the meta-model. Figure 9 graphically visualizes the values of CVRMSE, GOF, and R^2 for each meta-model examined, which confirms the superior performance of the MARS algorithm.

Figure 10 shows a scatter plot created between the MARS meta-model and the Dymola model for the whole generated sample. The *x*-axis represents Dymola simulation results, and the *y*-axis represents MARS predictions. The red line indicates the perfect fit line ($y = x$). It can be shown that the investigated MARS model provides an accurate prediction for the annual total consumption and is in very good agreement. The MARS is considered a meta-model of the original Dymola model for further SA, which is described in the next section.

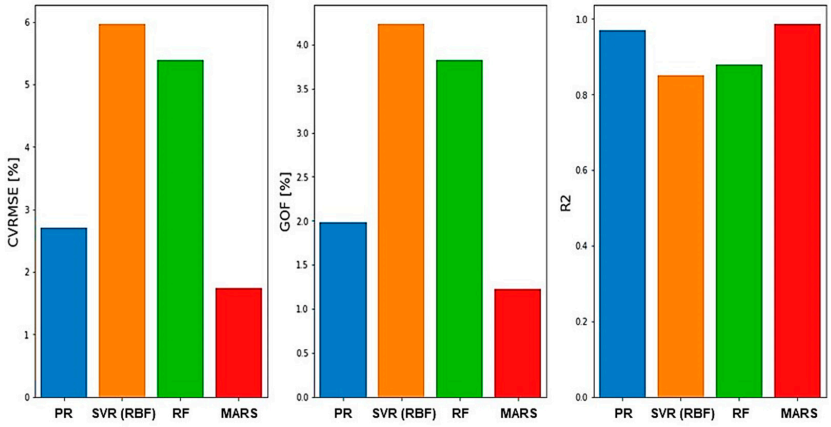


Figure 9. CVRMSE, GOF, and R^2 of investigated meta-models.

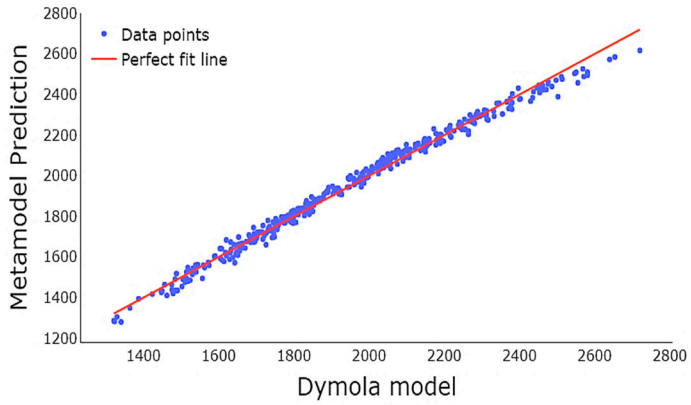


Figure 10. Scatter plot of MARS (R^2 = 0.99).

4.4. Sensitivity Analysis Using Sobol' Method

The Saltelli's extension of the Sobol' sequence was used to generate uniform samples of parameter space Ω. With a sample size of N = 256 and k = 44 uncertain input parameters (see Table 5), 23,040 evaluation runs ($N \times (2k + 2)$) were carried out using the MARS meta-model to perform the variance-based SA using the Sobol' method. The most sensitive parameters to the model output (annual total consumption) are shown in Figure 11 as total effects ST from the Sobol' method. The total sensitivity index ST indicates the total effects (first-order and all parameter interaction effects) of parameters on model output and can be considered a measure of negligible input parameters.

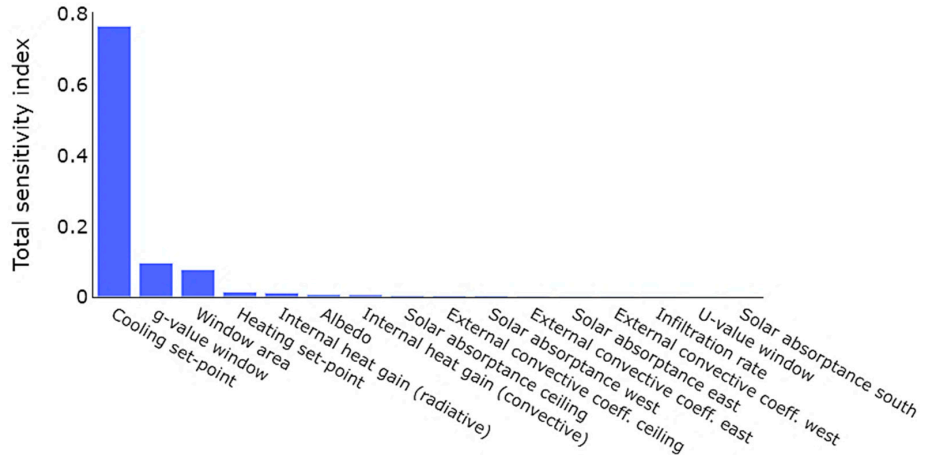


Figure 11. Ranking parameters by Sobol' method (total sensitivity index).

The ranking of the top 16 most influential input parameters (from an initial set of 105) with the corresponding total sensitivity index ST is given in Table 7. The remaining parameters exhibit very low ST values, i.e., they have no significant impact on the total variance of the model output.

Table 7. Input parameter rankings and their total sensitivity index by Sobol' method.

Rank	Parameter	Unit	ST
1	Cooling set-point	[°C]	0.7656
2	g-value window	-	0.0960
3	Window area	m^2	0.0768
4	Heating set-point	[°C]	0.0140
5	Internal heat gain (radiative)	W	0.0113
6	Albedo	-	0.0067
7	Internal heat gain (convective)	W	0.0063
8	Solar absorptance ceiling	-	0.0040
9	External convective coeff. ceiling	$W/(m^2\ K)$	0.0034
10	Solar absorptance west	-	0.0030
11	External convective coeff. east	$W/(m^2\ K)$	0.0025
12	Solar absorptance east	-	0.0021
13	External convective coeff. west	$W/(m^2\ K)$	0.0021
14	Infiltration rate	1/h (ACH)	0.0020
15	U-value window	$W/(m^2\ K)$	0.0016
16	Solar absorptance south	-	0.0012

It is observed that the first eight input parameters, i.e., cooling set-point temperature, g-value of the window, window area, heating set-point temperature, radiative portion of the internal heat gain, albedo, convective portion of the internal heat gain, and solar absorptance of the ceiling, are consistent with the SA from Morris' method (Table 5).

Nevertheless, there is considerable disagreement concerning the solar absorptance of the southern wall, which is ranked higher than by Morris' method. It can be inferred that the solar absorptance of the southern wall is involved in interaction effects with other input parameters or non-linear behavior.

5. Conclusions and Future Work

5.1. Conclusions

BEPS is a function of many input parameters. Due to large uncertainties and the complex non-linear behavior of parameters, a considerable mismatch can occur between BEPS output and reference data. Furthermore, carrying out stochastic simulations with a vast number of samples in BEPS tools is very time-consuming. In this paper, a novel four-stage SA framework based on the MARS meta-model is proposed to explore significant uncertain input parameters that the BEPS output (annual energy consumption) is most sensitive to and to provide an overview of critical parameters in a holistic context. The major contribution of this study can be formulated as follows:

- In the first stage, two deterministic simulations are carried out. The comparative testing validation method is used to assess the deviation of the simulation results.
- In the second stage, the screening-based SA using Morris' method is performed to identify the most important input parameters, which have a significant influence on the model output (annual energy consumption). In this stage, the number of uncertain parameters in the investigated case study is reduced from 105 to 44 prior to performing variance-based SA. The selected input parameters are used for constructing the meta-model.
- In the third stage, meta-modeling techniques are applied to tackle the computational costs of the variance-based SA, especially in cases of large numbers of time-consuming simulation runs. LHS, in combination with Monte Carlo simulation, is used to propagate uncertainties by running 400 simulation runs in the Dymola model and to generate a database for training and testing meta-models. A comparison study of the

four meta-models, MARS, PR, RF, and SVR (based on the RBF kernel), is conducted. The results show that MARS is the most accurate meta-model in this study and provides a very precise prediction for the annual energy consumption compared to the original Dymola model.

- Finally, in the fourth stage, the variance-based SA using Sobol' method is performed as a final, robust stage of SA. MARS is applied to implement uncertainty propagation based on Monte Carlo simulations (23,040 simulation runs). Running the MARS meta-model can be implemented faster and reduce the computational cost of Sobol' sensitivity indices in comparison with the direct simulation of complex building models.

Due to using meta-modeling techniques and implementing two GSAs (Morris and Sobol' methods), the proposed framework provides the fast, reliable SA of highly sophisticated building models, which is required in the early building design process for robust decision-making. In addition, this framework considers a large set of uncertain parameters, which increases the reliability of uncertainty and sensitivity analysis.

It can be concluded that the cooling set-point temperature has a strong influence on the annual total consumption with the total-order Sobol' index of $S_T = 0.76$ for the examined case study. Other highly influential parameters are those related to window parameters, i.e., the g-value and area of the window. In addition to this, this framework highlights that performing the screening-based SA (Morris' method) and MARS meta-model can reduce the computational cost of Sobol' sensitivity indices dramatically in comparison with the direct simulation of the original, complex building model. Therefore, the interpretation and evaluation of model outputs can be made faster and more robust.

5.2. Future Work

The proposed framework is based on weather data, which represents a cool-dry climate. The SA results could significantly differ if other weather data (such as hot-dry climate) is used. Further research work is required to improve the framework concerning several weather data points. Furthermore, MARS, in combination with Sobol' method, is used in this study to explore the impacts of uncertain input parameters on annual energy consumption. In future work, other meta-models in combination with different variance-based SAs (such as FAST) will be investigated to compare with the developed SA framework. However, machine learning algorithms are often uninterpretable black-box models and lack transparency. Indeed, for these algorithms, it is extremely hard to explain how predictions were made. Further research is required to apply explainable artificial intelligence (XAI) approaches, such as Local Interpretable Model-Agnostic Explanations (LIME) [91] and Shapley Additive exPlanation (SHAP) [92], to improve the interpretability of machine learning-based systems.

Author Contributions: Conceptualization, A.N., J.F. and C.v.T.; methodology, A.N.; software, A.N.; validation, A.N.; formal analysis, A.N.; investigation, A.N.; resources, A.N. and J.F.; data curation, A.N.; writing—original draft preparation, A.N.; writing—review and editing, J.F. and C.v.T.; visualization, A.N.; supervision, J.F. and C.v.T.; project administration, J.F.; funding acquisition, J.F. and C.v.T. All authors have read and agreed to the published version of the manuscript.

Funding: Parts of this research were funded by the German Federal Ministry for Economic Affairs and Climate Action under grant number 03ET1570B.

Data Availability Statement: Data are available upon request.

Acknowledgments: The authors would like to thank all project partners and contributors to the research project SimQuality (grant number 03ET1570B).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. List of input parameters and assigned uncertainty ranges used for the screening-based sensitivity analysis by Morris' method.

Index	Parameter	Baseline Point	Parameter Range (Lower and Upper Bounds)
	Wall south		
1	Conductivity layer 1 (plaster)	0.140 W/(m·K)	[0.126, 0.154]
2	Density layer 1 (plaster)	800 kg/m^3	[720, 880]
3	Specific heat capacity layer 1 (plaster)	1500 J/(kg·K)	[1350, 1650]
4	Thickness layer 1 (plaster)	0.005 m	[0.0045, 0.0055]
5	Conductivity layer 2 (insulation)	0.040 W/(m·K)	[0.036, 0.044]
6	Density layer 2 (insulation)	50 kg/m^3	[45, 55]
7	Specific heat capacity layer 2 (insulation)	1000 J/(kg·K)	[900, 1100]
8	Thickness layer 2 (insulation)	0.100 m	[0.09, 0.11]
9	Conductivity layer 3 (concrete)	1.200 W/(m·K)	[1.08, 1.32]
10	Density layer 3 (concrete)	2000 kg/m^3	[1800, 2200]
11	Specific heat capacity layer 3 (concrete)	1000 J/(kg·K)	[900, 1100]
12	Thickness layer 3 (concrete)	0.200 m	[0.18, 0.22]
13	Internal convective heat transfer coefficient	2.5 W/(m^2·K)	[2.25, 2.75]
14	External convective heat transfer coefficient	8 W/(m^2·K)	[7.2, 8.8]
15	Solar absorptance coefficient of external wall surface	0.6	[0.54, 0.66]
16	Longwave emission coefficient of internal wall surface	0.9	[0.81, 0.99]
	Wall west		
17	Conductivity Layer 1 (plaster)	0.140 W/(m·K)	[0.126, 0.154]
18	Density layer 1 (plaster)	800 kg/m^3	[720, 880]
19	Specific heat capacity layer 1 (plaster)	1500 J/(kg·K)	[1350, 1650]
20	Thickness layer 1 (plaster)	0.005 m	[0.0045, 0.0055]
21	Conductivity layer 2 (insulation)	0.040 W/(m·K)	[0.036, 0.044]
22	Density layer 2 (insulation)	50 kg/m^3	[45, 55]
23	Specific heat capacity layer 2 (insulation)	1000 J/(kg·K)	[900, 1100]
24	Thickness layer 2 (insulation)	0.100 m	[0.09, 0.11]
25	Conductivity layer 3 (concrete)	1.200 W/(m·K)	[1.08, 1.32]
26	Density layer 3 (concrete)	2000 kg/m^3	[1800, 2200]
27	Specific heat capacity layer 3 (concrete)	1000 J/(kg·K)	[900, 1100]
28	Thickness layer 3 (concrete)	0.200 m	[0.18, 0.22]
29	Internal convective heat transfer coefficient	2.5 W/(m^2·K)	[2.25, 2.75]
30	External convective heat transfer coefficient	8 W/(m^2·K)	[7.2, 8.8]
31	Solar absorptance coefficient of external wall surface	0.6	[0.54, 0.66]
32	Longwave emission coefficient of internal wall surface	0.9	[0.81, 0.99]
	Wall north		
33	Conductivity Layer 1 (plaster)	0.140 W/(m·K)	[0.126, 0.154]
34	Density layer 1 (plaster)	800 kg/m^3	[720, 880]
35	Specific heat capacity layer 1 (plaster)	1500 J/(kg·K)	[1350, 1650]
36	Thickness layer 1 (plaster)	0.005 m	[0.0045, 0.0055]
37	Conductivity layer 2 (insulation)	0.040 W/(m·K)	[0.036, 0.044]
38	Density layer 2 (insulation)	50 kg/m^3	[45, 55]
39	Specific heat capacity layer 2 (insulation)	1000 J/(kg·K)	[900, 1100]
40	Thickness layer 2 (insulation)	0.100 m	[0.09, 0.11]
41	Conductivity layer 3 (concrete)	1.200 W/(m·K)	[1.08, 1.32]
42	Density layer 3 (concrete)	2000 kg/m^3	[1800, 2200]
43	Specific heat capacity layer 3 (concrete)	1000 J/(kg·K)	[900, 1100]
44	Thickness layer 3 (concrete)	0.200 m	[0.18, 0.22]
45	Internal convective heat transfer coefficient	2.5 W/(m^2·K)	[2.25, 2.75]
46	External convective heat transfer coefficient	8 W/(m^2·K)	[7.2, 8.8]
47	Solar absorptance coefficient of external wall surface	0.6	[0.54, 0.66]
48	Longwave emission coefficient of internal wall surface	0.9	[0.81, 0.99]
	Wall east		
49	Conductivity Layer 1 (plaster)	0.140 W/(m·K)	[0.126, 0.154]
50	Density layer 1 (plaster)	800 kg/m^3	[720, 880]
51	Specific heat capacity layer 1 (plaster)	1500 J/(kg·K)	[1350, 1650]
52	Thickness layer 1 (plaster)	0.005 m	[0.0045, 0.0055]
53	Conductivity layer 2 (insulation)	0.040 W/(m·K)	[0.036, 0.044]
54	Density layer 2 (insulation)	50 kg/m^3	[45, 55]
55	Specific heat capacity layer 2 (insulation)	1000 J/(kg·K)	[900, 1100]
56	Thickness layer 2 (insulation)	0.100 m	[0.09, 0.11]
57	Conductivity layer 3 (concrete)	1.200 W/(m·K)	[1.08, 1.32]
58	Density layer 3 (concrete)	2000 kg/m^3	[1800, 2200]
59	Specific heat capacity layer 3 (concrete)	1000 J/(kg·K)	[900, 1100]
60	Thickness layer 3 (concrete)	0.200 m	[0.18, 0.22]
61	Internal convective heat transfer coefficient	2.5 W/(m^2·K)	[2.25, 2.75]
62	External convective heat transfer coefficient	8 W/(m^2·K)	[7.2, 8.8]
63	Solar absorptance coefficient of external wall surface	0.6	[0.54, 0.66]
64	Longwave emission coefficient of internal wall surface	0.9	[0.81, 0.99]
	Ceiling		
65	Conductivity Layer 1 (plaster)	0.140 W/(m·K)	[0.126, 0.154]
66	Density layer 1 (plaster)	800 kg/m^3	[720, 880]
67	Specific heat capacity layer 1 (plaster)	1500 J/(kg·K)	[1350, 1650]
68	Thickness layer 1 (plaster)	0.005 m	[0.0045, 0.0055]

Table A1. *Cont.*

Index	Parameter	Baseline Point	Parameter Range (Lower and Upper Bounds)
69	Conductivity layer 2 (insulation)	0.040 W/(m·K)	[0.036, 0.044]
70	Density layer 2 (insulation)	50 kg/m^3	[45, 55]
71	Specific heat capacity layer 2 (insulation)	1000 J/(kg·K)	[900, 1100]
72	Thickness layer 2 (insulation)	0.100 m	[0.09, 0.11]
73	Conductivity layer 3 (concrete)	1.200 W/(m·K)	[1.08, 1.32]
74	Density layer 3 (concrete)	2000 kg/m^3	[1800, 2200]
75	Specific heat capacity layer 3 (concrete)	1000 J/(kg·K)	[900, 1100]
76	Thickness layer 3 (concrete)	0.200 m	[0.18, 0.22]
77	Internal convective heat transfer coefficient	2.5 W/(m^2·K)	[2.25, 2.75]
78	External convective heat transfer coefficient	8 W/(m^2·K)	[7.2, 8.8]
79	Solar absorptance coefficient of external wall surface	0.6	[0.54, 0.66]
80	Longwave emission coefficient of internal wall surface	0.9	[0.81, 0.99]
	Floor		
81	Conductivity Layer 1 (plaster)	0.140 W/(m·K)	[0.126, 0.154]
82	Density layer 1 (plaster)	800 kg/m^3	[720, 880]
83	Specific heat capacity layer 1 (plaster)	1500 J/(kg·K)	[1350, 1650]
84	Thickness layer 1 (plaster)	0.005 m	[0.0045, 0.0055]
85	Conductivity layer 2 (insulation)	0.040 W/(m·K)	[0.036, 0.044]
86	Density layer 2 (insulation)	50 kg/m^3	[45, 55]
87	Specific heat capacity layer 2 (insulation)	1000 J/(kg·K)	[900, 1100]
88	Thickness layer 2 (insulation)	0.100 m	[0.09, 0.11]
89	Conductivity layer 3 (concrete)	1.200 W/(m·K)	[1.08, 1.32]
90	Density layer 3 (concrete)	2000 kg/m^3	[1800, 2200]
91	Specific heat capacity layer 3 (concrete)	1000 J/(kg·K)	[900, 1100]
92	Thickness layer 3 (concrete)	0.200 m	[0.18, 0.22]
93	Internal convective heat transfer coefficient	2.5 W/(m^2·K)	[2.25, 2.75]
94	External convective heat transfer coefficient	8 W/(m^2·K)	[7.2, 8.8]
95	Solar absorptance coefficient of external wall surface	0.6	[0.54, 0.66]
96	Longwave emission coefficient of internal wall surface	0.9	[0.81, 0.99]
	Window properties		
97	Total surface area (window glazing + frame)	3 m^2	[2.70, 3.30]
98	U-value (window glazing + frame)	1.10 W/(m^2·K)	[0.99, 1.21]
99	g-value (SHGC—Solar Heat Gain Coefficient)	0.6	[0.54, 0.66]
	Operation parameters		
100	Infiltration rate	0.5 1/h	[0.45, 0.55]
101	Convective internal heat gains	80 W	[72, 88]
102	Radiative internal heat gains	120 W	[108, 132]
103	Albedo (ground reflectance)	0.2	[0.18, 0.22]
	HVAC system		
104	Heating set-point temperature	20 °C	[18, 22]
105	Cooling set-point temperature	28 °C	[25.2, 30.8]

References

1. Fumo, N. A review on the basics of building energy estimation. *Renew. Sustain. Energy Rev.* **2014**, *31*, 53–60. [CrossRef]
2. Borgstein, E.; Lamberts, R.; Hensen, J. Evaluating energy performance in non-domestic buildings: A review. *Energy Build.* **2016**, *128*, 734–755. [CrossRef]
3. de Wilde, P. The gap between predicted and measured energy performance of buildings: A framework for investigation. *Autom. Constr.* **2014**, *41*, 40–49. [CrossRef]
4. Van Dronkelaar, C.; Dowson, M.; Burman, E.; Spataru, C.; Mumovic, D. A Review of the Regulatory Energy Performance Gap and Its Underlying Causes in Non-domestic Buildings. *Front. Mech. Eng.* **2016**, *1*, 17. [CrossRef]
5. Bowman, N.T.; Lomas, K.J. Empirical validation of dynamic thermal computer models of buildings. *Build. Serv. Eng. Res. Technol.* **1985**, *6*, 153–162. [CrossRef]
6. Judkoff, R. Validation of building energy analysis simulation programs at the solar energy research institute. *Energy Build.* **1988**, *10*, 221–239. [CrossRef]
7. Judkoff, R.; Wortman, D.; O'Doherty, B.; Burch, J. *Methodology for Validating Building Energy Analysis Simulations*; Technical Report; National Renewable Energy Lab (NREL): Golden, CO, USA, 2008.
8. Lauster, M.; Teichmann, J.; Fuchs, M.; Streblow, R.; Mueller, D. Low order thermal network models for dynamic simulations of buildings on city district scale. *Build. Environ.* **2014**, *73*, 223–231. [CrossRef]
9. Roberti, F.; Oberegger, U.F.; Gasparella, A. Calibrating historic building energy models to hourly indoor air and surface temperatures: Methodology and case study. *Energy Build.* **2015**, *108*, 236–243. [CrossRef]
10. Ferroukhi, M.Y.; Abahri, K.; Belarbi, R.; Limam, K.; Nouviaire, A. Experimental validation of coupled heat, air and moisture transfer modeling in multilayer building components. *Heat Mass Transf.* **2016**, *52*, 2257–2269. [CrossRef]
11. Zhang, R.; Sun, K.; Hong, T.; Yura, Y.; Hinokuma, R. A novel Variable Refrigerant Flow (VRF) heat recovery system model: Development and validation. *Energy Build.* **2018**, *168*, 399–412. [CrossRef]
12. Kim, J.; Frank, S.; Im, P.; Braun, J.E.; Goldwasser, D.; Leach, M. Representing Small Commercial Building Faults in EnergyPlus, Part II: Model Validation. *Buildings* **2019**, *9*, 239. [CrossRef]
13. Lee, S.H.; Hong, T. Validation of an inverse model of zone air heat balance. *Build. Environ.* **2019**, *161*, 106232. [CrossRef]

14. Im, P.; New, J.R.; Joe, J. Empirical validation of building energy modeling using flexible research platform. In Proceedings of the Building Simulation 2019: 16th Conference of IBPSA, Rome, Italy, 2–4 September 2019; pp. 4515–4521.
15. Halimov, A.; Lauster, M.; Müller, D. Development and validation of PCM models integrated into the high order building model of Modelica library—Aixlib. In Proceedings of the Building Simulation 2019: 16th Conference of IBPSA, Rome, Italy, 2–4 September 2019; pp. 4698–4705. [CrossRef]
16. Kalogeras, G.; Rastegarpour, S.; Koulamas, C.; Kalogeras, A.P.; Casillas, J.; Ferrarini, L. Predictive capability testing and sensitivity analysis of a model for building energy efficiency. *Build. Simul.* **2020**, *13*, 33–50. [CrossRef]
17. Gutiérrez, G.V.; Ramos Ruiz, G.; Fernández Bandera, C. Empirical and Comparative Validation for a Building Energy Model Calibration Methodology. *Sensors* **2020**, *20*, 5003. [CrossRef]
18. Eriksson, M.; Akander, J.; Moshfegh, B. Development and validation of energy signature method—Case study on a multi-family building in Sweden before and after deep renovation. *Energy Build.* **2020**, *210*, 109756. [CrossRef]
19. Shi, H.; Chen, Q. Building energy management decision-making in the real world: A comparative study of HVAC cooling strategies. *J. Build. Eng.* **2021**, *33*, 101869. [CrossRef]
20. Yi, Z.; Lv, Y.; Xu, D.; Xu, J.; Qian, H.; Zhao, D.; Yang, R. Energy saving analysis of a transparent radiative cooling film for buildings with roof glazing. *Energy Built Environ.* **2021**, *2*, 214–222. [CrossRef]
21. Eguía-Oller, P.; Martínez-Mariño, S.; Granada-Álvarez, E.; Febrero-Garrido, L. Empirical validation of a multizone building model coupled with an air flow network under complex realistic situations. *Energy Build.* **2021**, *249*, 111197. [CrossRef]
22. Magni, M.; Ochs, F.; Streicher, W. Comprehensive analysis of the influence of different building modelling approaches on the results and computational time using a cross-compared model as a reference. *Energy Build.* **2022**, *259*, 111859. [CrossRef]
23. Loutzenhiser, P.G.; Manz, H.; Moosberger, S.; Maxwell, G.M. An empirical validation of window solar gain models and the associated interactions. *Int. J. Therm. Sci.* **2009**, *48*, 85–95. [CrossRef]
24. Tabares-Velasco, P.C.; Griffith, B. Diagnostic test cases for verifying surface heat transfer algorithms and boundary conditions in building energy simulation programs. *J. Build. Perform. Simul.* **2012**, *5*, 329–346. [CrossRef]
25. Mateus, N.M.; Pinto, A.; da Graça, G.C. Validation of EnergyPlus thermal simulation of a double skin naturally and mechanically ventilated test cell. *Energy Build.* **2014**, *75*, 511–522. [CrossRef]
26. Alaidroos, A.; Krarti, M. Experimental validation of a numerical model for ventilated wall cavity with spray evaporative cooling systems for hot and dry climates. *Energy Build.* **2016**, *131*, 207–222. [CrossRef]
27. Sandels, C.; Brodén, D.; Widén, J.; Nordström, L.; Andersson, E. Modeling office building consumer load with a combined physical and behavioral approach: Simulation and validation. *Appl. Energy* **2016**, *162*, 472–485. [CrossRef]
28. Cattarin, G.; Pagliano, L.; Causone, F.; Kindinis, A. Empirical and comparative validation of an original model to simulate the thermal behaviour of outdoor test cells. *Energy Build.* **2018**, *158*, 1711–1723. [CrossRef]
29. Nageler, P.; Schweiger, G.; Pichler, M.; Brandl, D.; Mach, T.; Heimrath, R.; Schranzhofer, H.; Hochenauer, C. Validation of dynamic building energy simulation tools based on a real test-box with thermally activated building systems (TABS). *Energy Build.* **2018**, *168*, 42–55. [CrossRef]
30. Barone, G.; Buonomano, A.; Forzano, C.; Palombo, A. Building Energy Performance Analysis: An Experimental Validation of an In-House Dynamic Simulation Tool through a Real Test Room. *Energies* **2019**, *12*, 4107. [CrossRef]
31. Nouri, A.; Frisch, J.; van Treeck, C. Statistical methodologies for verification of building energy performance simulation. In Proceedings of the Building Simulation 2021: 17th Conference of IBPSA, Bruges, Belgium, 1–3 September 2021. [CrossRef]
32. Tian, Z.; Love, J.A.; Tian, W. Applying quality control in building energy modelling: Comparative simulation of a high performance building. *J. Build. Perform. Simul.* **2009**, *2*, 163–178. [CrossRef]
33. Lucchino, E.C.; Gelesz, A.; Skeie, K.; Gennaro, G.; Reith, A.; Serra, V.; Goia, F. Modelling double skin façades (DSFs) in whole-building energy simulation tools: Validation and inter-software comparison of a mechanically ventilated single-story DSF. *Build. Environ.* **2021**, *199*, 107906. [CrossRef]
34. Heiselberg, P.; Brohus, H.; Hesselholt, A.; Rasmussen, H.; Seinre, E.; Thomas, S. Application of sensitivity analysis in design of sustainable buildings. *Renew. Energy* **2009**, *34*, 2030–2036. [CrossRef]
35. Spitz, C.; Mora, L.; Wurtz, E.; Jay, A. Practical application of uncertainty analysis and sensitivity analysis on an experimental house. *Energy Build.* **2012**, *55*, 459–470. [CrossRef]
36. Strachan, P.; Monari, F.; Kersken, M.; Heusler, I. IEA Annex 58: Full-scale Empirical Validation of Detailed Thermal Simulation Programs. *Energy Procedia* **2015**, *78*, 3288–3293. [CrossRef]
37. Cattarin, G.; Pagliano, L.; Causone, F.; Kindinis, A.; Goia, F.; Carlucci, S.; Schlemminger, C. Empirical validation and local sensitivity analysis of a lumped-parameter thermal model of an outdoor test cell. *Build. Environ.* **2018**, *130*, 151–161. [CrossRef]
38. Alongi, A.; Angelotti, A.; Mazzarella, L. A numerical model to simulate the dynamic performance of Breathing Walls. *J. Build. Perform. Simul.* **2021**, *14*, 155–180. [CrossRef]
39. Raslan, R.; Davies, M. Results variability in accredited building energy performance compliance demonstration software in the UK: An inter-model comparative study. *J. Build. Perform. Simul.* **2010**, *3*, 63–85. [CrossRef]
40. Shrestha, S.S.; Maxwell, G. Empirical validation of building energy simulation software: EnergyPlus. In Proceedings of the Building Simulation 2011: 12th Conference of IBPSA, Sydney, Australia, 14–16 November 2011.
41. Tabares-Velasco, P.C.; Christensen, C.; Bianchi, M. Verification and validation of EnergyPlus phase change material model for opaque wall assemblies. *Build. Environ.* **2012**, *54*, 186–196. [CrossRef]

42. Antretter, F.; Sauer, F.; Schöpfer, T.; Holm, A. Validation of a hygrothermal whole building simulation software. In Proceedings of the Building Simulation 2011: 12th Conference of IBPSA, Sydney, Australia, 14–16 November 2011.
43. Zhu, D.; Hong, T.; Yan, D.; Wang, C. A detailed loads comparison of three building energy modeling programs: EnergyPlus, DeST and DOE-2.1E. *Build. Simul.* **2013**, *6*, 323–335. [CrossRef]
44. Kim, Y.-J.; Yoon, S.-H.; Park, C.-S. Stochastic comparison between simplified energy calculation and dynamic simulation. *Energy Build.* **2013**, *64*, 332–342. [CrossRef]
45. Buonomano, A. Code-to-Code Validation and Application of a Dynamic Simulation Tool for the Building Energy Performance Analysis. *Energies* **2016**, *9*, 301. [CrossRef]
46. Moronis, A.; Koulamas, C.; Kalogeras, A. Validation of a monthly quasi-steady-state simulation model for the energy use in buildings. In Proceedings of the 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus, 12–15 September 2017; pp. 1–6. [CrossRef]
47. Nouri, A.; Nicolai, A.; Krämer, B.; Hirth, S.; Agudelo, J.; Seifert, C.; Malhotra, A.; Madjidi, M.; Frisch, J.; van Treeck, C. Entwicklung von Qualitätsstandards für die energetische Gebäude- und Anlagensimulation als Planungswerkzeug. In Proceedings of the BauSim Conference 2020: 8th Conference of IBPSA, Graz, Austria, 23–25 September 2020; pp. 79–86. [CrossRef]
48. Zakula, T.; Badun, N.; Ferdelji, N.; Ugrina, I. Framework for the ISO 52016 standard accuracy prediction based on the in-depth sensitivity analysis. *Appl. Energy* **2021**, *298*, 117089. [CrossRef]
49. Saltelli, A.; Ratto, M.; Tarantola, S.; Campolongo, F. Update 1 of: Sensitivity Analysis for Chemical Models. *Chem. Rev.* **2012**, *112*, PR1–PR21. [CrossRef]
50. Tian, W. A review of sensitivity analysis methods in building energy analysis. *Renew. Sustain. Energy Rev.* **2013**, *20*, 411–419. [CrossRef]
51. Yang, J. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. *Environ. Model. Softw.* **2011**, *26*, 444–457. [CrossRef]
52. Saltelli, A.; Annoni, P. How to avoid a perfunctory sensitivity analysis. *Environ. Model. Softw.* **2010**, *25*, 1508–1517. [CrossRef]
53. Hygh, J.S.; DeCarolis, J.F.; Hill, D.B.; Ranjithan, S.R. Multivariate regression as an energy assessment tool in early building design. *Build. Environ.* **2012**, *57*, 165–175. [CrossRef]
54. McLeod, R.S.; Hopfe, C.J.; Kwan, A. An investigation into future performance and overheating risks in Passivhaus dwellings. *Build. Environ.* **2013**, *70*, 189–209. [CrossRef]
55. Kristensen, M.H.; Petersen, S. Choosing the appropriate sensitivity analysis method for building energy model-based investigations. *Energy Build.* **2016**, *130*, 166–176. [CrossRef]
56. Mechri, H.E.; Capozzoli, A.; Corrado, V. USE of the ANOVA approach for sensitive building energy design. *Appl. Energy* **2010**, *87*, 3073–3083. [CrossRef]
57. Pang, Z.; O'Neill, Z.; Li, Y.; Niu, F. The role of sensitivity analysis in the building performance analysis: A critical review. *Energy Build.* **2020**, *209*, 109659. [CrossRef]
58. Garcia Sanchez, D.; Lacarrière, B.; Musy, M.; Bourges, B. Application of sensitivity analysis in building energy simulations: Combining first- and second-order elementary effects methods. *Energy Build.* **2014**, *68*, 741–750. [CrossRef]
59. Van Gelder, L.; Das, P.; Janssen, H.; Roels, S. Comparative study of metamodelling techniques in building energy simulation: Guidelines for practitioners. *Simul. Model. Pract. Theory* **2014**, *49*, 245–257. [CrossRef]
60. Rivalin, L.; Stabat, P.; Marchio, D.; Caciolo, M.; Hopquin, F. A comparison of methods for uncertainty and sensitivity analysis applied to the energy performance of new commercial buildings. *Energy Build.* **2018**, *166*, 489–504. [CrossRef]
61. Wate, P.; Iglesias, M.; Coors, V.; Robinson, D. Framework for emulation and uncertainty quantification of a stochastic building performance simulator. *Appl. Energy* **2020**, *258*, 113759. [CrossRef]
62. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer Series in Statistics; Springer Science & Business Media: New York, NY, USA, 2009. [CrossRef]
63. Chou, J.-S.; Bui, D.-K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy Build.* **2014**, *82*, 437–446. [CrossRef]
64. Hester, J.; Gregory, J.; Kirchain, R. Sequential early-design guidance for residential single-family buildings using a probabilistic metamodel of energy consumption. *Energy Build.* **2017**, *134*, 202–211. [CrossRef]
65. Eisenhower, B.; O'neill, Z.; Narayanan, S.; Fonoberov, V.A.; Mezić, I. A methodology for meta-model based optimization in building energy models. *Energy Build.* **2012**, *47*, 292–301. [CrossRef]
66. Jain, R.K.; Smith, K.M.; Culligan, P.J.; Taylor, J.E. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Appl. Energy* **2014**, *123*, 168–178. [CrossRef]
67. Amasyali, K.; El-Gohary, N.M. A review of data-driven building energy consumption prediction studies. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1192–1205. [CrossRef]
68. Cheng, M.-Y.; Cao, M.-T. Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Appl. Soft Comput.* **2014**, *22*, 178–188. [CrossRef]
69. Østergård, T.; Jensen, R.L.; Maagaard, S.E. A comparison of six metamodeling techniques applied to building performance simulations. *Appl. Energy* **2018**, *211*, 89–103. [CrossRef]

70. Calleja Rodríguez, G.; Carrillo Andrés, A.; Domínguez Muñoz, F.; Cejudo López, J.M.; Zhang, Y. Uncertainties and sensitivity analysis in building energy simulation using macroparameters. *Energy Build.* **2013**, *67*, 79–87. [CrossRef]
71. Bucking, S.; Zmeureanu, R.; Athienitis, A. A methodology for identifying the influence of design variations on building energy performance. *J. Build. Perform. Simul.* **2014**, *7*, 411–426. [CrossRef]
72. Kim, Y.-J. Comparative study of surrogate models for uncertainty quantification of building energy model: Gaussian Process Emulator vs. Polynomial Chaos Expansion. *Energy Build.* **2016**, *133*, 46–58. [CrossRef]
73. Morris, M.D. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics* **1991**, *33*, 161. [CrossRef]
74. Sobol′, I.M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **2001**, *55*, 271–280. [CrossRef]
75. Burhenne, S.; Tsvetkova, O.; Jacob, D.; Henze, G.P.; Wagner, A. Uncertainty quantification for combined building performance and cost-benefit analyses. *Build. Environ.* **2013**, *62*, 143–154. [CrossRef]
76. Helton, J.C.; Davis, F.J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Eng. Syst. Saf.* **2003**, *81*, 23–69. [CrossRef]
77. Andersson, C.; Åkesson, J.; Führer, C. *PyFMI: A Python Package for Simulation of Coupled Dynamic Models with the Functional Mock-up Interface*; Technical Report; Lund University: Lund, Sweden, 2016. Available online: https://lup.lub.lu.se/record/961a50eb-e4a8-43bc-80ac-d467eef26193 (accessed on 22 January 2024).
78. ASHRAE Guideline 14. *Measurement of Energy and Demand Savings*; American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, GA, USA, 2002.
79. FEMP. *M&V Guidelines: Measurement and Verification for Performance-Based Contracts, Version 4.0*; U.S. Department of Energy Federal Energy Management Program: Washington, DC, USA, 2015.
80. IPMVP®. *International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings*; Efficiency Valuation Organization (EVO): Washington, DC, USA, 2002; Volume 1.
81. Reddy, T.; Maor, I. *ASHRAE Research Project 1051-RP: Procedures for Reconciling Computer-Calculated Results with Measured Energy Data*; American Society of Heating, Refrigerating and Air-Conditioning Engineers: Peachtree Corners, GA, USA, 2006.
82. Campolongo, F.; Cariboni, J.; Saltelli, A. An effective screening design for sensitivity analysis of large models. *Environ. Model. Softw.* **2007**, *22*, 1509–1518. [CrossRef]
83. Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. *Global Sensitivity Analysis. The Primer*; John Wiley & Sons, Ltd.: London, UK, 2007. [CrossRef]
84. Jin, R.; Chen, W.; Simpson, T. Comparative studies of metamodelling techniques under multiple modelling criteria. *Struct. Multidiscip. Optim.* **2001**, *23*, 1–13. [CrossRef]
85. Brieman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
86. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: New York, NY, USA, 1995.
87. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; Haussler, D., Ed.; pp. 144–152. [CrossRef]
88. Friedman, J.H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19*, 1–67. [CrossRef]
89. Modelica®. A Unified Object-Oriented Language for Systems Modeling, Language Specification, Version 3.2.2 [Computer Software]. Available online: https://www.modelica.org/ (accessed on 30 October 2022).
90. *ANSI/ASHRAE/IES Standard 90.1*; Energy Standard for Buildings Except Low-Rise Residential Buildings. American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, GA, USA, 2019.
91. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why Should I Trust You?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 1135–1144. [CrossRef]
92. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 4768–4777.