Development of a diagnostic tool for evaluating the thermal performance of Nearly Zero Energy buildings

Mohamed Hany Abokersh [a], Marleen Spiekman [b], Olav Vijlbrief [b], T.A.J. van Goch [c], Manel Vallès [a], Dieter Boer [a, *]

[a] Departament d'Enginyeria Mecànica, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Spain

[b] Department of Building Physics and Systems, TNO, Leeghwaterstraat 44, Delft, the Netherlands

[c] BAM Bouw en Techniek, Bunnik, the Netherlands

** Corresponding author: Dieter.boer@urv.cat*

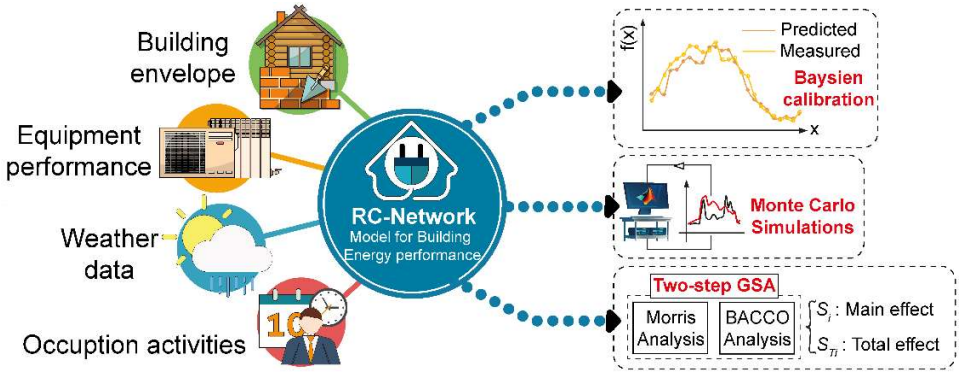
E-mail addresses: mohamed.abokersh@urv.cat (*M.H. Abokersh),* marleen.spiekman@tno.nl *(Marleen Spiekman*), olav.vijlbrief@tno.nl *(O. Vijlbrief),* dennis.van.goch@bam.com (*T.A.J. van Goch),* manel.valles@urv.cat *(Manel Valles)*

Abstract

The nearly zero-energy buildings (nZEB) presents a promising contribution to fulfill the EU sustainable future targets. However, the construction industry that leads the development of nZEB is facing challenges to guarantee its performance. In this context, this paper proposes a methodology framework based on Multizone Resistance and capacitance-Network to trace the nZEB performance challenges with quantifications for the time-dependent variables comprising occupant behaviors as well as the dynamic behavior of weather conditions and building operations. This approach incorporates Bayesian optimization for calibration purposes to minimize the required monitoring data. Finally, in order to quantify and assess the uncertainty associate with the performance of the developed digital dwelling to represent the actual building, the framework inherent the uncertainty analysis (UA) incorporating two-step global sensitivity analysis (GSA). The methodology application is demonstrated through a case study for a newly renovated two-story home located in a district of Emmen at the Netherlands. The results confirm a high accuracy for the digital dwelling performance where the model offers a prediction accuracy of 2.2% and 7.03% for the thermal energy consumption and indoor zone temperature, respectively. On the other hand, the UA confirms a high uncertainty associate with the nZEB performance where the total thermal energy consumption can increase up to 100 kWh/m^2/yr. This variation is driven by the infiltration rates followed by the building envelope characteristics. The proposed framework can serve a diagnostic tool to assist the construction and installation companies to maintain the performance of their products proactively.

Keywords: Nearly Zero Energy Building; Building performance simulation; Occupant behaviors; Bayesian calibration; Monte Carlo uncertainty analysis; Global sensitivity analysis

Graphical abstract



Highlights

- A diagnostic tool is developed to assess dwelling energy performance.
- A probabilistic approach is proposed for all the time-dependent uncertain parameters.
- Bayesian optimization can provide reliable calibration for dwelling energy model.
- Monte Carlo simulations quantify uncertainty associate with dwelling performance.
- The key uncertain parameters are defined based on global sensitivity analysis.

Nomenclature

$A(\theta(t))$	window opening area only for the ingoing airflow (m^2)
A_{glass}	effective solar collecting area (m^2)
$C.V$	coefficient of variation (%)
C_d	discharge airflow rate coefficient
C_p	wind pressure coefficient
C_v	air permeability coefficient
C	total thermal mass of the building (kW)
C_{int}	indoor thermal mass (kW)
C_{wall}	outdoor thermal mass (kW)
EE	elementary effect
f_c	correction factor of the thermal mass
$f_{shading}$	envelope shading reduction factor
G_{value}	solar energy transmittance (-)
$Infl._i$	Infiltration rate at zone i (m^3/s)
n	flow exponent
N_s	number of total simulations
$Occup_{sch_i}$	occupant schedule at zone i
$Occup_{int_i}$	internal heat gain at zone i (kWh)
ΔP	air pressure differential across a building envelope (Pa)
ΔP_{th}	difference of pressure across the window opening (Pa)
Q_{sol}	amount of global solar radiation (W/m^2)
Q_v	calculated volumetric airflow through infiltration (m^3/s)
$Q_{v_{10}}$	measured volumetric airflow through infiltration (m^3/s)
Q_{Heat}	total thermal energy consumption at each building zone (kWh)
r	number of elementary effects per parameter
R^2	R-squared (%)
RC_{roof}	thermal resistance of the roof ($m^2 \cdot K/W$)
RC_{floor}	thermal resistance of the floor($m^2 \cdot K/W$)
RC_{facade}	thermal resistance of the façade ($m^2 \cdot K/W$)
sf	solar fraction factor
$SMAPE$	symmetric mean absolute percentage error (%)
T_{int}	indoor zone temperature (°C)
T_{ground}	ground temperature (°C)
T_{wall}	envelope zone temperature (°C)
T_{out}	outdoor temperature (°C
T_{neigh}	neighbour temperature (°C)
T_{set_i}	setpoint zone temperatures (°C)
t	operation time (hour)
$t_{o/u}$	total over/underheating hours
U	heat loss coefficient ($W/m^2 \cdot K$)
U_{glass}	U-value of windows $W/(m^2 \cdot K)$
$vent._i$	mechanical ventilation at zone i (m^3/s)
'V_{wind}'	local air velocity (m/s)
$y_{data,i}$	actual value at time point i
$y_{predict,i}$	predicted value at time point i

Greek symbol

η_{fan}	heat recovery efficiency (%)
μ	mean value
θ	opening/closing the windows and doors (º)
σ	standard deviation

ρ	air density (m^3/kg)
$\emptyset_{int}$	internal heat gain by occupants and appliances (kW)
$\emptyset_{H}$	heat flow rate sourced from the heaters based on heat pump (kW)
$\emptyset_{exch}$	exchange heat flow rate with other zones (kW)
$\emptyset_{sol}$	heat flow rate sourced from solar (kW)
$\emptyset_{vent}$	ventilation and infiltration losses (kW)
$\emptyset_{neigh}$	heat flow rate exchange with the neighbour (kW)
$\emptyset_{trans}$	heat flow rate transmitted through the windows (kW)
$\emptyset_{gnd}$	heat loss to the ground (kW)

Abbreviations

BACCO	Bayesian analysis of computer code outputs
BES	Building energy simulation
COP	Heat pump performance factor
EPC	Energy Performance Coefficient
GEM-SA	Gaussian Emulator Machine Sensitivity Analysis Software
GSA	Global sensitivity analysis
MC	Monte Carlo
nZEB	Nearly zero energy building
PDF	Probability density function
RC	Resistance and capacitance
SA	Sensitivity analysis
UA	Uncertainty analysis
UC	Uncertainty Characterization

1. Introduction

From the Brundtland Report “Our Common Future” in 1987 [1] to UN Climate Conference of Parties, 21 of 2015 through the Kyoto Protocol stipulation in 1997, the increase of awareness regarding the sustainability can be identified on a global level. A vital footstep to spread on the sustainable communities concept relies on the energy efficiency in the building sector. A fraction of 40% of the total energy utilized in the Union [2] and 36% of the total emissions of CO_2 come from the buildings [3]. Thus, in the building sector, the EU environmental and energy policy has progressed, and it has covered the sustainability and resource efficiency for reducing the CO_2 emissions and energy consumption [4]. The objective of the European Commission is to reduce the CO_2 emissions from the building sector in 2050 by 88-91% compared to CO_2 emission 1990 levels [5].

The Netherlands is working on increasing the energy efficiency of buildings since 1975. They are limiting the transmission losses by controlling the insulation values [6]. In 1995,’EPC’ (Energy Performance Coefficient) was added to these limits. It is a non-dimensional quantity that uses the energy required for space heating, cooling, hot water, humidification, and lighting as an indicator of energy performance. The ratio of calculated energy demand to a standard energy performance yields the Dutch EPC calculation where the total heated area of the dwelling and the heat transfer surface defines the standardized energy performance [7]. In 1996, the value of EPC was set to 1.4 that could be achieved through the construction techniques. In 1998, the value of EPC was reduced to 1.2. It was further reduced to 1.0 in 2000, 0.8 in 2006, and 0.6 in 2011. With the introduction of the concept of the nearly zero energy buildings (nZEB), the EPC value was further reduced to 0.4 in 2015 [8].

The estimation of the heating systems impact and the accurate size of the renewable energy system with mutually and multiple different design criteria are the biggest challenges during the design of nZEB [9]. That is why; the literature has focused on learning the thermal performance of these buildings. In the research papers, the objective is achieved through the monitoring studies as given in [10] or through the model simulations as provided in [11,12]. The thermal performance of a building helps in foreseeing the consequence of the renovation measures of potential energy, and then, recognizing the optimized option for energy, comfort as well as addressing the economic issues [2].

In this context, building energy simulation (BES) can be utilized for studying the behavior of the nZEBs under many engineering and design conditions, such as thermal properties of the building [13], occupant’s behavior [14], different weather conditions [15], as well as energy supply systems [16]. There are several uncertainties and driving factors involved in building performance. Such aspects require detailed inputs to the BES tool for modelling the thermal performance. Software such as EnergyPlus, BLAST, TRNSYS, practices a transient method to simulate the cooling and heating loads of the building. Such tools require modelling expertise and professional knowledge because they provide reliable, detailed, and sophisticated modelling and simulations dynamically inherent into the calculations [17]. Another type of simulations is the Black box or statistical approach [18]. This approach is used in the case of incomplete information about the building. This approach uses the measured data for the identification of the parameters and selects a mathematical model (ARMAX, polynomial, transfer function) to represent the thermal behavior of the building [19]. The generation of the training database by using the simulation engines is the major challenge of this method because it consumes significant computational resources. Heuristic optimization has the same issues when it is used for the identification of the objective function specified in relation to the building energy use, the thermal comfort level, and economic benefits [20].

Apart from the above-described approaches, research has been conducted to define a lighter modeling method that can be utilized for the comparative study of the energy use and building heating and cooling loads. In one of these researches, RC (resistance and capacitance) model is used for modelling the building thermal behavior; this approach works on the phenomenon of electrical analogue [21]. In this context,

researchers have indicted the RC model for modelling the building's thermal reaction under the effect of outdoor and indoor thermal conditions. Berthoua et al.[22] simulated the heating and cooling needs of a multi-zone office building by using four RC models with different complexity. From the simulation results, it was observed that the two-order 6R2C semi-physical model is the best-compromised model out of all the tested models. The accuracy of the model for predicting the indoor air temperature and thermal needs during the cooling and heating periods was around 84%. Terés-Zubiaga et al.[23] used monitored data and developed a sophisticated RC model for the dwelling. In the first stage, an empty social housing dwelling was monitored for 3 months to analyze its thermal performance. Afterwards, the monitored data was used for the development of the proposed model. The development of the model and some of the results are evaluated and presented later. Asadi et al. [24] worked on the development of a multi-objective mathematical model that assists the stakeholders in making decisions regarding the energy-minimizing cost-effectively. They evaluate the effectiveness of all the combinations of the retrofit actions on a simultaneous basis by using an RC model. Liao and Dexter [25] simulated the dynamic behavior of the existing heating system by developing a second-order RC physical model. The results illustrated that the simple low-order RC model had been verified to have a good performance in modeling the heating and cooling need of a simple building that usually has one zone or single-use. The shortcoming of the RC model is that it presents the performance for a shorter period, such as a few days to several weeks. In this way, the long-term prediction is not possible from the RC model. Moreover, the simplified models show a lag in quantifications of occupant behaviors, building operations, and dynamic behavior of weather conditions. These factors strongly influence the RC network's performance [26]. Thus, the energy savings output from simulation models of existing buildings has indicated discrepancies often significant in some cases (up to 100% differences), between BES model-predicted and the actual metered building energy use [27].

In this context, the researchers have conducted several ways to reduce the variations between the real-time energy measurements and the model results. A detailed audit is conducted to get a real insight into the different characteristics of the building. It was the primary approach used for the calibration of the energy model [28]. All the parameters that affect the energy performance of the building must be included in these characteristics. Some of the essential parameters are material properties, building geometry, lighting and equipment schedules, and type of schedule and occupancy [29]. This approach is constructive for the correction of the model, but it involves many uncertainties, and it consumes a lot of time [27]. Because of these issues, statistical analysis methods such as Bayesian method is utilized to perform uncertainty analysis (UA) and calibrate the energy model [30]. For instance, Booth et al. [31] used the macro-level data to calibrate micro-level models by utilizing the Bayesian method. Bayesian inferences and a combination of regression models were used for the estimation of the average consumption intensity. Bayesian methods can illustrate the uncertainties involved in the consumption intensities, and that is why they are known as bottom-up methods. In simple words, the Bayesian approach predicts the range of consumption level with decent probability because it works on the basis of the detailed measured data.

By involving UA into the calculations, many other studies are convinced to use this method for the calibration of the energy model [32]. Over the last few decades, sensitivity analysis (SA) [33], optimization search algorithms [34], and objective penalty function [35] are utilized for calibrating the energy models. Monetti et al. [36] have presented an optimization-based calibration procedure. They used GenOpt [37] for developing a simulation-based optimization algorithm that minimized the variations in the simulation results and the measured data to calibrate the energy model. In another paper, Reddy and Sun [35] calibrated a building energy model by utilizing an analytical optimization approach. There were four steps in their calibration procedure: sensitivity analysis, identifiability analysis, optimization, and uncertainty analysis. The sensitivity analysis performed by them was very comprehensive, and their results contain the occupant's behavior during the calibration of the buildings.

Liu and Henze [38] calibrated the building energy model by utilizing a common optimization outline and GenOpt. They majorly focused on cooling loads. They exhibited that the model can perform energy model calibration with appropriate accuracy. However, for reducing the optimization problem dimensions, the number of parameters were kept lower. Therefore, occupants' behaviors' effect in the context of different factors of consumption is not undertaken in this study. As evidenced from the literature review, the majority of these researches focused on quantifying the uncertainty associate with building energy model based on the time-independent approach with ignoring the temporal variation for some parameters. In other words, the underlying occupancy behavior and weather variation are estimated in a deterministic manner instead of considering its stochastic nature [39,40].

In practice, the construction companies that lead the development of nZEB in the Netherlands indicate a substantial obstacle to guarantee the nZEB performance and thus to improve their market position. An additional challenge is that the guarantee of building data quality where monitoring a thousand of homes simultaneously isn't feasible since sensors may not function properly and it can be out of budget. Therefore, the need for a generic approach to monitor and guarantee the building performance is essential. Furthermore, this approach should be based on minimal monitoring data for individual homes to find out what is the cause for the current situation of the deviations between the actual and expected performance as analyses based on wrong data lead to wrong conclusions, wrong decisions, extra costs and loss of support.

Following the challenges facing the construction companies and the limitations associate with various strategies for BES regarding the verification and calibration processes, this study aims to propose a methodology framework to address these shortcomings via the following contributions:

i. The Development of a Multizone RC-Network on an hourly base. This model offers an extended period prediction for the heating demands and indoor air temperature with quantifications for the occupant behaviors as well the dynamic behavior of weather conditions and building operations.
ii. Trace the limitations associated with the current building energy models calibration through considering the temporal variation associate with the climate change projections and the occupant behaviors in the calibration process.
iii. The development of a modelling scheme for defining the difference between the predicted and actual energy and indoor performance at an individual nZEB based on uncertainty analysis (UA) incorporating Bayesian optimization to minimize the required monitoring data.
iv. The application of the proposed framework is demonstrated through a case study to insight the uncertainty importance using a global sensitivity analysis approach (GSA).

Hence, the main novelty is to develop a diagnostic tool to guarantee and proactively maintain the indoor climate performance of nZEBs with consideration for the temporal variation associate with the climate data and occupant behavior during the operation stage of the building. Furthermore, with the data analysis model based on UA incorporating GSA, additional knowledge in the future can be added to the builders, installers and occupants, which can bridge the nZEB concept with the client anticipations. The structure of the paper is: (i) The methodology for developing a Multizone RC-Network incorporating the Bayesian optimization for calibration purposes is given in Section 2. In addition, it shows the quantification for Uncertainty Characterization (UC) based on global sensitivity analysis approach. (ii) The case study of the paper for nZEB diagnostic tool is given in Section 3. (iii) The paper's results regarding the developed framework performance are presented and discussed in Section 4. Finally, (iv) the concluding marks are given in Section 5 can be useful for the construction and installation companies working with buildings performance diagnosis.

2. Methodology structure

Fig. 1 shows the outlines of the proposed methodology framework for diagnosing nZEB energy performance which consist of two main phases. Phase 1 includes the development of RC model and its calibration where step (A) consists of the development of a RC model for building energy performance. This model includes different building thermal properties, occupancy behavior, as well as the supply energy systems. In step (B), the Uncertainty Characterization (UC) that contains the uncertain RC model parameters are identified and described in a probabilistic manner. Once the RC model uncertain parameters are quantified, the dimension reduction is implemented to screen the important parameters that affect the energy model prediction. In step (C), the Bayesian calibration is carried out to minimize the gap between the field observations and the RC-Network simulations output. On the other hand, phase 2 comprises the estimation for the uncertainty associate with calibrated RC model and the parameters responsible for these uncertainties. Therefore, in step (D), Monte Carlo (MC) simulations are used to perform uncertainty analysis (UA) to estimate the variation in the RC Network output. In the end, step (E) provides the two-step Global Sensitivity Analysis (GSA) to highlight the important effects of uncertain on the energy performance of the nZEB.

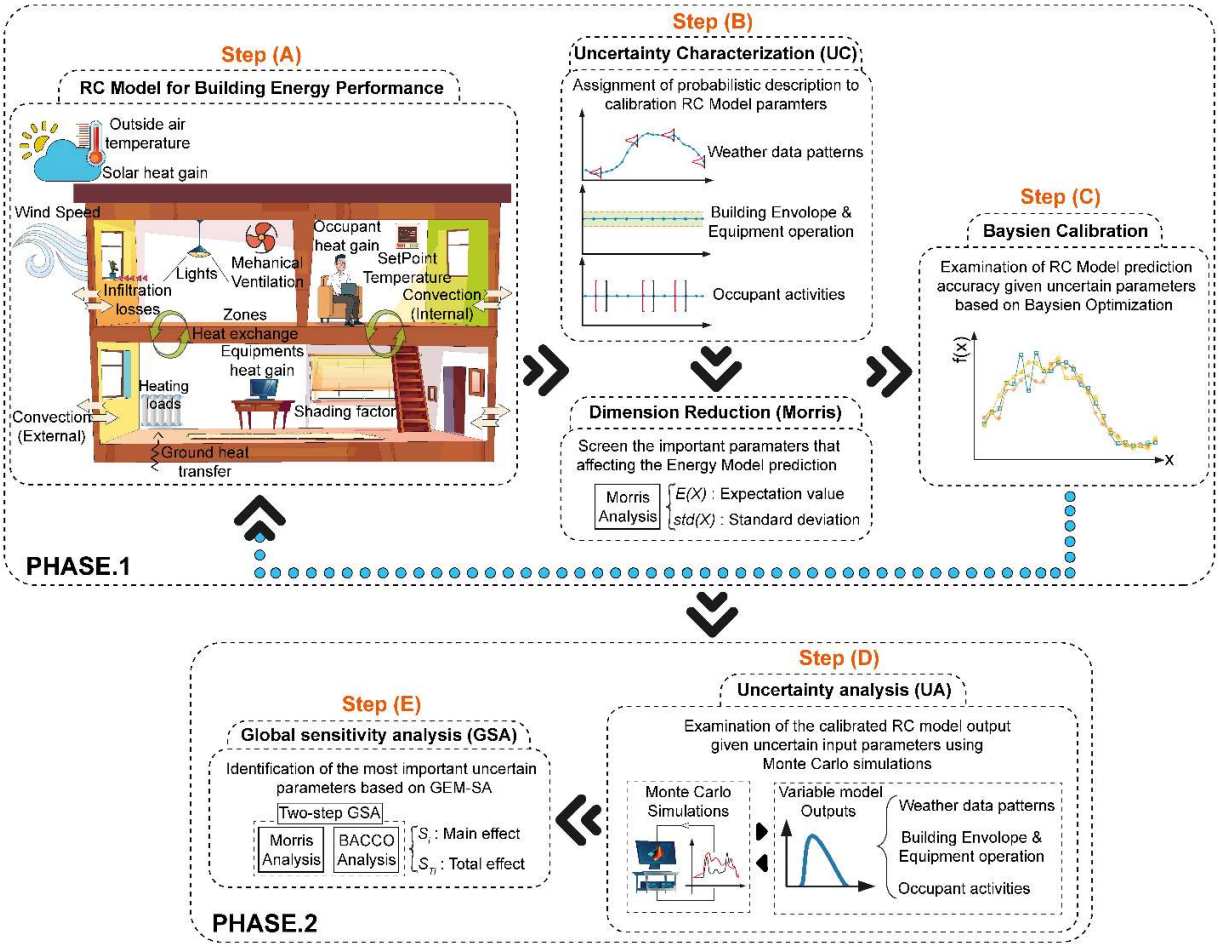


Fig. 1: General workflow for the diagnostic tool of nZEB energy performance

2.1. RC Model formulation

The core of the RC model relies on two-node model (6R2C) thermal-electric analogy that is programmed in MATLAB. This approach based on adapting the electric analog to simulate the thermal performance of multi zones dwelling in a dynamic manner. The RC model follows the Kirchhoff's law which is able to provide the temperature and the heat flow rate at each simulation time step [20].

The key parameters for the heat transfer in the modelling approach are $\emptyset_{int}$(internal heat gain by occupants and appliances, in kW) , $\emptyset_H$ (heat flow rate sourced from the heaters based on heat pump, in kW), $\emptyset_{exch}$ (exchange heat flow rate with other zones, in kW), $\emptyset_{sol}$ (heat flow rate sourced from solar, in kW), $\emptyset_{nei}$ (heat flow rate exchange with the neighbour, in kW), $\emptyset_{vent}$ (ventilation and infiltration losses of the zone, in kW), $\emptyset_{trans}$ (heat flow rate transmitted through the windows, in kW) and $\emptyset_{gnd}$, (heat loss to the ground, in kW). All energy flows through Zone ith are indicated in Fig. 2, where the blue dashed rectangle describes the inner and outer node.

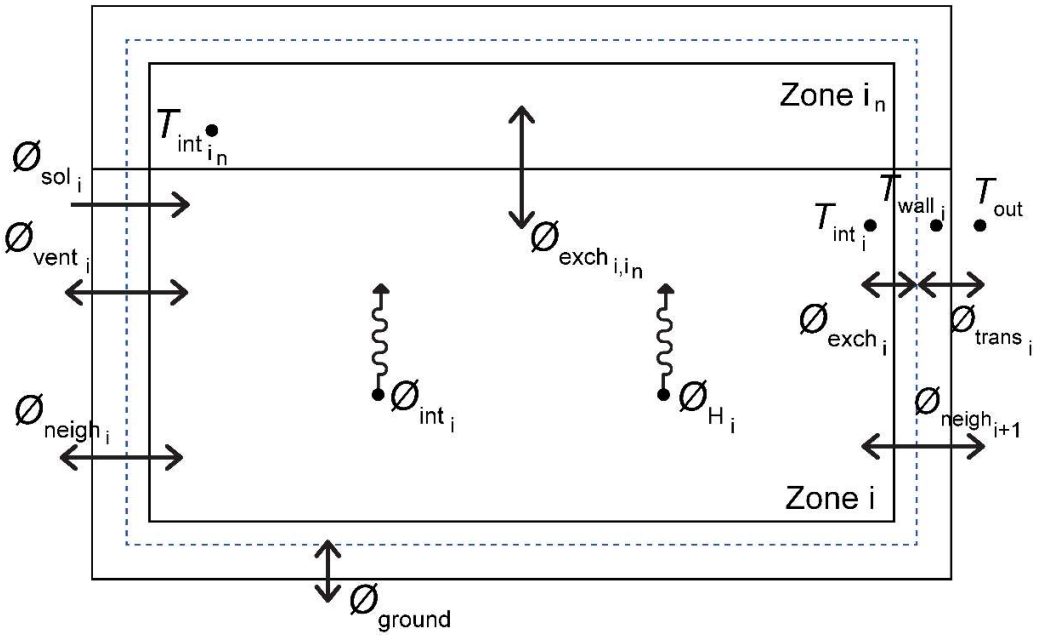


Fig. 2: Energy Flow Through Zone ith and its interaction with the surrounding

The basic idea of the proposed model relies on two specific thermal mass nodes. The first node is typically consisting of the indoor envelopes thicknesses including the walls, floor, and roof and it counts to a few centimeters in addition to the furniture. This part of the thermal mass is called the 'indoor mass', C_{int}. While the second node states for the heat exchange between the indoor thermal mass and the ambient through the R_{exc} , and it is called the 'outdoor mass', C_{wall}. The thermal capacity of the air is neglected since it is relatively small compared to the indoor walls' capacity. The RC Network representing the building by the two thermal masses is illustrated in Fig. 3. For estimating the total thermal mass of the building, a simplified formula derived using volume and density of the building materials is utilized, and it can be expressed as following [41]:

$$C = \alpha . V^{\beta} . f_c \quad (1)$$

where α & β coefficients are 0.18 and 0.92, respectively. While f_c is a correction factor which is defined as the ratio between the indoor mass and the total mass, which is initially set to 0.3.

To better take into account the heat transmitted through the envelope in each zone, the proposed model has three temperature nodes which are T_{int} (indoor zone temperature, in °C), T_{wall} (envelope zone temperature, in °C), and T_{out} (outdoor temperature, in °C). It can be described based on stochastic differential equations in the following way:

$$\frac{dT_{int_i}}{dt} = \frac{1}{C_{int_i}} (\emptyset_{int_i} + \emptyset_{sol_i} + \emptyset_{vent_i} - \emptyset_{exch_i} + \emptyset_{exch_{i,i_n}} + \emptyset_{neigh_i} + \emptyset_{neigh_{i+1}} + \emptyset_{H_i}) \quad (2)$$

Where: $\emptyset_{vent_i} = UA_{vent_i}(T_{out} - T_{int_i})$ (2.1)

$$\emptyset_{exch_i} = UA_{exch_i}(T_{int_i} - T_{wall_i}) \quad (2.2)$$

$$\emptyset_{exch_{i,\,i_n}} = UA_{exch_{i,\,i_n}} \left(T_{int_{i_n}} - T_{int_i}\right) \quad (2.3)$$

$$\emptyset_{neigh_i} = UA_{neigh_i}(T_{neigh_i} - T_{int_i}) \quad (2.4)$$

$$\emptyset_{neigh_{i+1}} = UA_{neigh_{i+1}}(T_{neigh_{i+1}} - T_{int_i}) \quad (2.5)$$

$$\frac{dT_{wall_i}}{dt} = \frac{1}{C_{wall_i}} (\emptyset_{trans_i} + \emptyset_{ground} + \emptyset_{exch_i}) \quad (3)$$

Where: $\emptyset_{ground_i} = UA_{gnd_i} \left((T_{ground} + T_{out})/2 - T_{wall_i}\right)$ (3.1)

$$\emptyset_{trans_i} = UA_{trans_i}(T_{out} - T_{wall_i}) \quad (3.2)$$

In general, the inverse of the heat loss coefficient (U) through different building skin is used to represent the thermal resistances in the proposed model. For the internal temperature node, the solar and internal heat gains tend to raise the indoor zone temperature. While the ventilation, infiltration and conduction heat losses tend to reduce it. The conduction heat losses to the ambient are represented through the sum of the R_{exch}. R_{exch_n} is presented to represents the transfer of heat between the floor and internal walls if the zone thermal coupling will be considered. Moreover, the ventilation and infiltration losses of the building are represented by the thermal resistance R_{vent}. On the other hand, the heat flow in the envelope zone temperature node is made up of the sum of the heat flow rate transmitted to the ground R_{ground}, windows R_{trans} and ambient R_{exch}.

In the proposed method, the internal heat gain is counting for heat gains by occupants, appliances, and lighting devices where occupants heat gain is estimated based on a schedule for the number of occupants in each zone. In order to incorporate the heat gain caused by the appliance, the net electricity consumption for the dwelling is determined by using the measured data of the electricity consumption from the grid, generation from PV panels, electricity consumption by a heat pump and the surplus electricity fed to the grid from the PV panels.

In calculating solar gains, solar radiation is not always perpendicularly oriented on the window. Thus solar fraction factor (sf) is introduced to covert the solar radiation on a horizontal surface to a vertical surface, and it can be expressed as [42]:

$$\emptyset_{sol} = f_{shading} . Q_{sol} . A_{glass} . sf . G_{value} \quad (4)$$

Where Q_{sol} is the amount of global solar radiation as a function of envelope shading reduction factor $f_{shading}$, effective solar collecting area A_{glass}, and G_{value} is the solar energy transmittance.

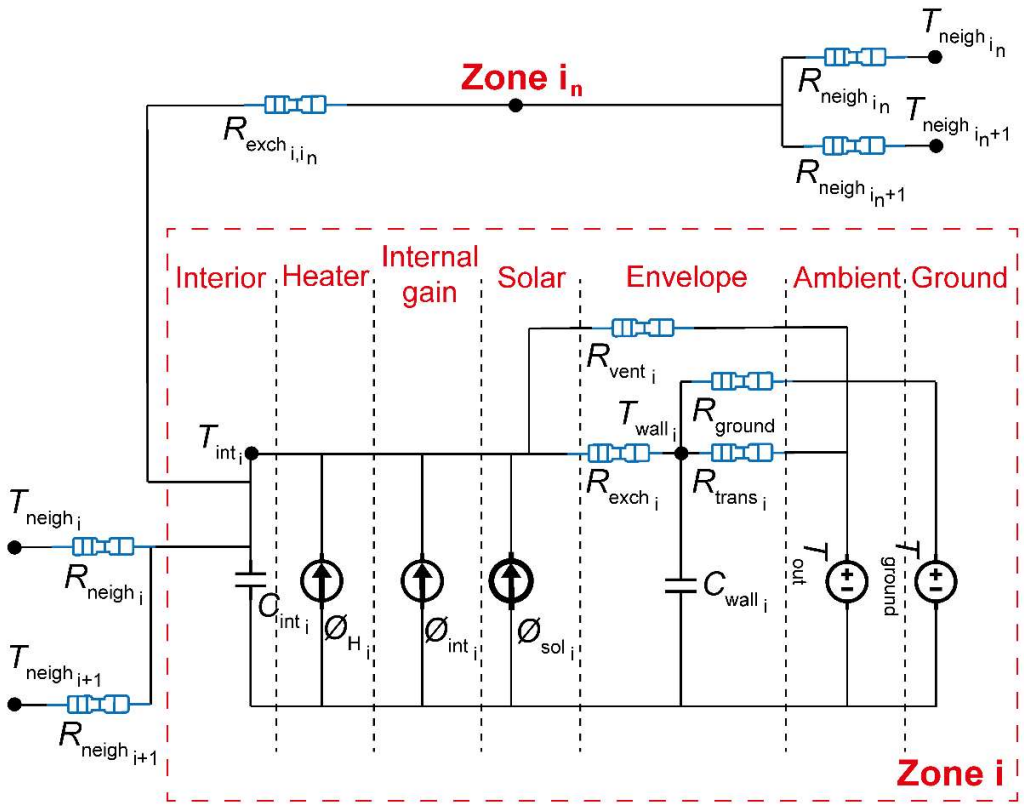


Fig. 3: Thermal network for the simplified model of Multizone dwelling

Based on the measured data, mechanical ventilation, and the natural ventilation comprising infiltration and occupant behavior regarding opening/closing windows and doors are quantified. For mechanical ventilation, it can be easily seen that it varies depending on the position of ventilation. However, there is no specific sensor data to obtain the ventilation position. Therefore, the energy consumption of the mechanical ventilation for all dwellings is used to estimate the position of ventilation in each zone, and subsequently, the mechanical ventilation volumetric flow rate.

For the occupant behavior regarding the windows and doors opening, the volumetric flow rate through windows that are opened on a single side is estimated through counting for the fresh air into buildings based on the pressure difference across the window, and it can be expressed as [43]:

$$q(t) = Opening\ sensor(t).C_d.A\big(\theta(t)\big).\sqrt{\frac{2|\Delta P_{th}(t)|}{\rho}} \quad (5)$$

where: $\Delta P_{th}(t) = \frac{1}{2}\rho.C_p.V_{wind}(t)^2$

Where C_d is the discharge airflow rate coefficient, which represents the fractional loss of airflow due to windows' geometry. $A\big(\theta(t)\big)$ is the window opening area only for the ingoing airflow, and it is a share of the total operable building's window area depending on the windows open percentage 'θ' at the given time 't', ρ is the air density, and ΔP_{th} is the difference of pressure across the window opening. The ΔP_{th} across the window opening is approximated estimated as a function of the local air velocity 'V_{wind}', for a given configuration of the building at a given time 't', and C_p is the wind pressure coefficient [44].

The volumetric airflow through infiltration can be modeled by Power Law below which is the function of an air pressure differential across a building envelope and the flow characteristic of the shell.

$$Q_v = C_v * \Delta P^n \quad (6)$$

where C_v is air permeability coefficient, ΔP is pressure difference and n flow exponent. Theoretically, the flow exponent lies between n = 1 and n=0.5. n=1 shows the fully developed laminar airflow and n = 0.5 given the fully developed turbulent airflow. Measured infiltration rate ($Q_{v_{10}}$) for the dwelling is 1 l/s.m^2 at 10 Pa pressure difference. Then $Q_{v_{10}}$ is extrapolated back to more typical pressures inside dwellings (1-5 Pa). The pressure difference inside the dwelling is assumed 1 Pa, and the n is assumed 0.5 in this model. Although the pressure difference inside the dwelling changes over the year based on wind orientation and wind speed, however, it is assumed constant in this model.

Finally, we estimate the energy provided by the heat pump. Firstly the temporary temperatures are calculated for the next time step by state equations (2) and (3). After that, the required heating powers for each zone are calculated based on setpoint zone temperatures (T_{set_i}) and temporary temperatures. Then the heating capacity is evenly distributed based on the heating needs in each zone. Finally, the temperatures in the next time step are calculated based on adjusted heating powers for each zone.

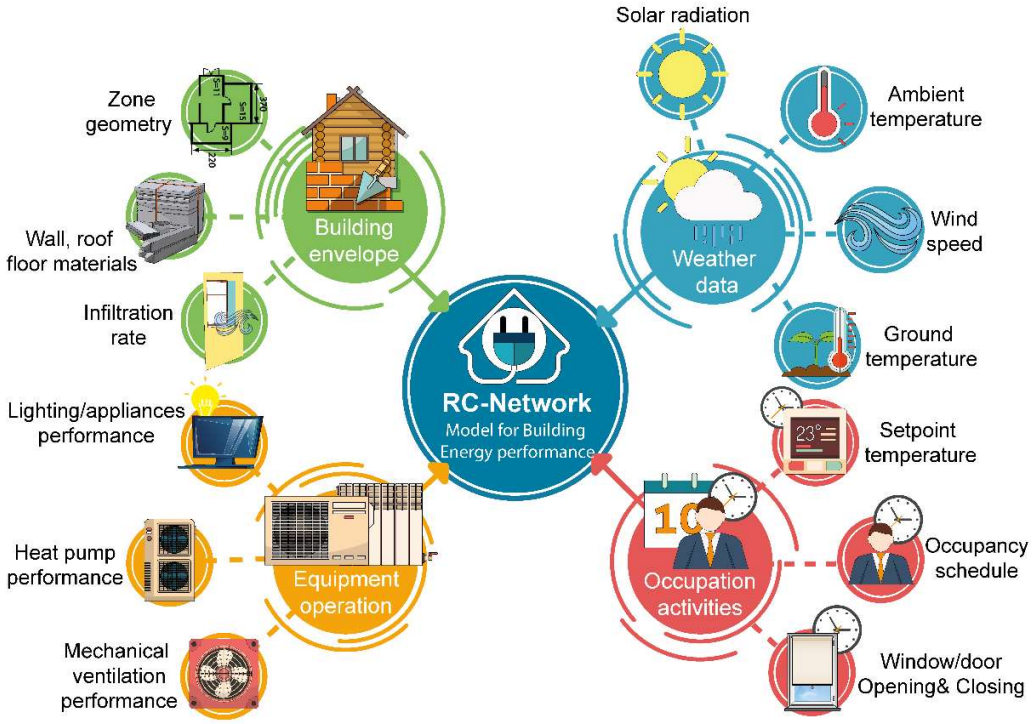


Fig. 4: A flow chart for the RC Model inputs

A flowchart summarizes the RC model inputs is shown in Fig. 4. This model is using the hourly bases for the metrological data, including the hourly-radiation on a horizontal surface, hourly-outdoor temperature, hourly wind speed, and average ground temperature. In terms of the building's envelopes, the zone geometry, the infiltration rates as well the building envelopes comprising the walls, floors and roof materials are defined. For the occupant behavior impact on the building performance, the RC model uses the hourly-setpoint temperature, the daily occupancy schedule, and the hourly-opening/closing schedules for windows and doors. Finally, in terms of the equipment operation, the mechanical ventilation performance as well the

heat pump performance and energy consumed by lighting and appliances are considered as inputs in the developed model.

2.2. Calibration of the building model

In the building industry, the simplified simulation models is a common approach used for the analysis of the building performance [45]. However, there are a lot of variations between the actual consumption of real buildings and the simulation results. Therefore, these models can be calibrated to identify savings and analyze retrofit options for supporting the investments grade Energy Conversion Measures [46]. When the forward energy simulation program that involves various numerical inputs are calibrated against the building energy data, a highly under-determined problem occurs with non-unique solutions [45]. Kaplan et al. [47], have noticed that sensitivity issues and the exact solutions of the calibration problems have primary significance in the field of the calibration. According to another perspective, a dynamic matching between the measured values and the computed data over one year is required for the calibration. Calibration cannot be performed at a static one at one condition [48].

To understand the concept of model calibration, model uncertainty is essential to understand, especially, for the indeterminate models of complex systems. The recent BES studies have neglected this vital issue, and it is not considered in any validation criteria of BES [27]. Østergård et al. [49] have classified the different uncertainty sources in BES as follows:

- **Specification Uncertainty:** occurs because of the inaccurate and incomplete specifications of the modelled building. Plant schedules, HVAC specifications, geometry, system, and material properties, etc. are the exposed model parameters that are included in the specification uncertainty.
- **Modelling Uncertainty:** It occurs because of the assumptions and simplifications of the complex models. These assumptions can be hidden from the tool (calculation algorithms) or maybe explicit to the modeler.
- **Scenario Uncertainty:** It occurs because of the external conditions that are imposed on the building. For instance, occupant's behavior and outdoor climate conditions.

In the current study, we are considering only the specification uncertainty and scenario uncertainty associated with the RC-model, and they are classified into four main categories comprises climate input data, building envelope, equipment system, and occupant behavior.

2.2.1. Uncertainty characterization (UC)

Uncertainty Characterization (UC) identifies the uncertain parameters associate with the RC model development and assign a mathematical description for their uncertainty. The uncertain parameters can be classified based on their time dependency, as shown in Fig. 5. In the time-independent uncertain parameters, the uncertainty associated with these parameters is constant over its whole simulation period. While the time-dependent uncertain parameters are temporal variables where they change at the time step of each simulation. The parameter uncertainty is described through the probabilistic methods. These probabilistic methods consider each parameter as a random variable by following the probability density function (PDF) [50]. Coakley et al. [27] gave a review of different UC approaches for describing the uncertain parameters of BES models. The next sections of the report discuss the UC approaches that are used for all the parameter categories of BES model. At each zone of the RC Model, 18 different uncertain parameters are found.

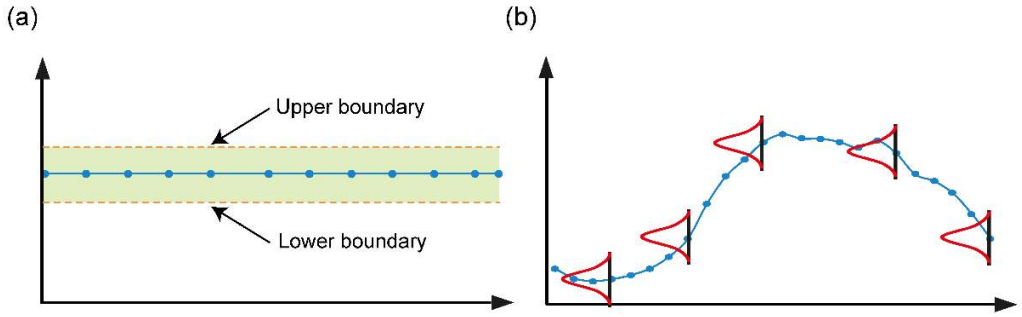


Fig. 5: Types of uncertain parameters: (a) time-independent parameters, (b) time-dependent parameter

a. Climate input data

Due to weather nature, there are high uncertainties in weather data. It further causes variations in building energy usage. In BES, Finkelstein-Schafer method is used for applying the chosen meteorological year from actual weather data over several decades [51]. Currently, concerns have arisen that a single weather file (that consists of 8760 hours values) does not contain sufficient and reliable information on credible weather conditions. In this way, it can't be reliable for assessing the building energy performance [52]. The simulated energy consumption based on typical weather data is not able to represent the weather data for actual meteorological years. Furthermore, historical weather conditions do not affect building performance. However, future climate affects building performance. Most of the weather data that is obtained from the historical data do not help forecast the building energy usage in the future [53]. The uncertainty associated with the weather data comprises; the solar radiation 'Q_{sol}', wind speed 'V_{wind}', the ambient temperature 'T_{out}', and ground temperature 'T_{ground}'.

Kaplanis and Kaplani [54] performed an outstanding analysis regarding solar radiation uncertainty. They used their observed data for analyzing the stochastic characteristics of global solar radiation based on an hourly basis. They deduct the following conclusions from their study:

- In hourly global solar radiation, the fluctuation is high in winter and lower in summer.
- In a day, the fluctuations are higher in the morning and late afternoon and lower around the solar noon.
- The normal distribution is experienced in hourly solar radiation.

A summary of the probability distribution of hourly global solar radiation is shown in Table 1. For depicting the frequency distribution of wind speed, probability density function (PDF) is a widely used technique. Several PDF approaches are proposed for wind speed estimation [55]. Among all the PDF approaches for wind energy, Weibull distribution is the most reliable and authentic approach [56]. It is illustrated in Table 1. Moreover, for the ground and ambient temperatures, a normal distribution is used by applying the mean equal to normal values as given in Table 1 [57,58].

b. Building envelope

The windows properties, building insulation materials thickness, and other properties such as building thermal correction factor and infiltration rate are the building envelope properties that affect the BES model's performance. The differences between the design specifications and the construction outcomes cause uncertainties in the insulation materials' thickness and subsequently changes its insulation capacity. Wei

Tiana et al. [53] followed the assumption that the parameter of materials thickness follows a normal distribution that contains a standard deviation of 5% of the mean value. Gaussian distribution is used for describing different parameters such as U-value of windows 'U_{glass}', G-value of windows 'G_{value}', and shading factor reflecting the blinds effect 'f_{shadi} ', and they were described based on a Gaussian distribution [59–61]. For the infiltration rate that depends on the construction quality, weather conditions, function of age, and building use [44]. It is counted as one of the most uncertain parameters because it is not easy to determine in buildings [62]. In this study, a Normal distribution that contains a standard deviation of 12.5% of its mean is considered. Finally, for the f_{c_i} at zone i which as always unknown parameters in the RC model, we assume it as a uniform distribution ranged between 10% and 90%. A summary for the uncertainty associated with the building envelopes parameters is shown in Table 1.

c. Equipment system

Based on the manufacturer data, engineering judgement, and experience and recommendations of Smith et al. [63], the systematic and random uncertainty values for the heat recovery efficiency 'η_{fan}' are taken as ±0.5% of its mean value. Following the recommendation by Zhang [61], the uncertainty associated with the mechanical ventilation '$vent._i$' at zone i is considered as a normal distribution with a standard deviation 2.5% of its mean value.

d. Occupant behavior

The relationship between the thermal performance of a building and its occupant behavior has emerged as an essential topic for research [64]. Occupant behavior is capable of introducing significant uncertainties in BES [65]. The detailed studies have shown that occupancy behavior in comparable buildings can lead to different energy consumption [26,66].

Most of the energy building simulation programs consider the factors linked to the occupant's behavior as deterministic by allowing the users to postulate fixed temporal schedules for occupancy-related variables, such as zone-heating set points, occupants, and closing/opening of the windows. These schedules can be easily implemented, but they are not responsible for the human behavior's stochastic nature or its interaction with the indoor environment of the buildings [53]. Following Sun et al. [67] recommendation, normal distributions are assigned to the occupant schedule '$Occup_{sch_i}$' and their relative internal heat gain '$Occup_{int_i}$' at zone i with a standard deviation 25% of its mean value. Furthermore, the same distribution is used with zone setpoint temperature 'T_{set_i}' at zone i with a 14% standard deviation to its mean value [53]. Finally, the angle of opening/closing the windows and doors 'θ_i' at zone i are of the most controversial issues regarding the BES. We assume that the fraction of opening windows is following a uniform distribution ranged between 1% and 90% with an initial value of 10%. In reality, many other occupant behaviors parameters such as the opening/closing windows schedule can contribute significantly to the BES performance especially in the nZEB case. However, due to the lack of information regarding the uncertainty associates with such these parameters, we define them only in a deterministic manner. A summary of the uncertainty associated with the occupant behavior parameters is shown in Table 1.

Table 1: Parameter-uncertainty distributions affecting the RC model performance

Parameter category	Uncertain parameter	Time frame	Probability distribution	Ref.
Input climate data	Q_{sol}	November–April, 9:00 am–3:00 pm	$N(\mu, 0.12\mu)$	[54]
		November–April, the rest of the day	$N(\mu, 0.25\mu)$	
		May–October, 9:00 am–3:00 pm	$N(\mu, 0.03\mu)$	
		May–October, the rest of the day	$N(\mu, 0.08\mu)$	

	V_{wind}	Time-dependent	$Wei(1.6,4.6)$	[56]
	T_{out}	Time-dependent	$N(\mu, 0.067\mu)$	[57]
	T_{ground}	Time-dependent	$N(\mu, 0.01\mu)$	[58]
Building envelope	RC_{roof}	Time-independent	$N(\mu, 0.05\mu)$	[53]
	RC_{floor}	Time-independent	$N(\mu, 0.05\mu)$	[53]
	RC_{facade}	Time-independent	$N(\mu, 0.05\mu)$	[53]
	U_{glass}	Time-independent	$N(\mu, 0.004\mu)$	[59]
	G_{value}	Time-independent	$N(\mu, 0.006\mu)$	[60]
	$f_{shading}$	Time-dependent	$N(\mu, 0.25\mu)$	[61]
	$Infl._i$	Time-independent	$N(\mu, 0.0125\mu)$	[61]
	f_{c_i}	Time-independent	$U(1\%, 90\%)$	-
Equipment system	η_{fan}	Time-independent	$U(75\%, 85\%)$	[63]
	$vent._i$	Time-independent	$N(\mu, 0.025\mu)$	[61]
Occupant behavior	$Occup_{sch}$	Time-dependent	$N(\mu, 0.25\mu)$	[67]
	$Occup_{int}$	Time-dependent	$N(\mu, 0.25\mu)$	[67]
	θ_i	Time-dependent	$U(0\%, 90\%)$	-
	T_{set_i}	Time-dependent	$N(\mu, 0.14\mu)$	[53]

2.2.2. Dimension reduction using Morris analysis

The dimension reduction strategy is used to identify the key uncertain parameters out of various uncertainty variables given in Table 1, which influence the performance of the BES based on RC model. It facilitates an accurate prediction within the BES performance. As shown in Fig. 6, the first step of the dimension reduction technique is to generate an input matrix. It is generated by grouping the design parameters on the basis of the distribution of the parameters with consideration of its time-dependent nature.

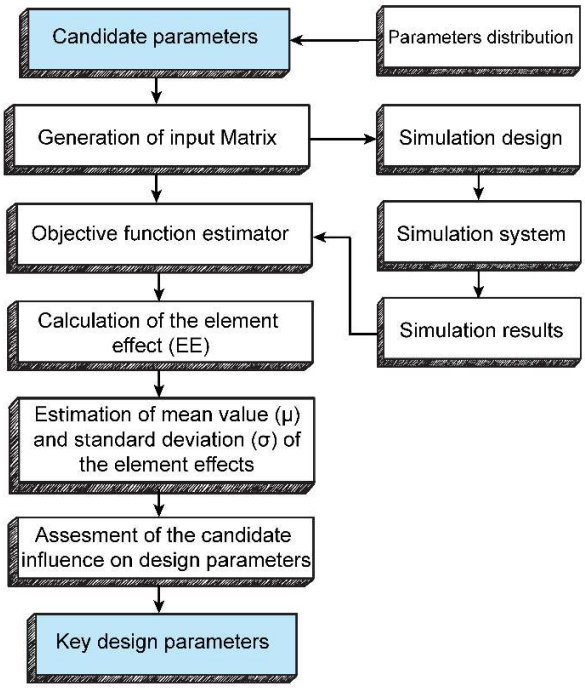


Fig. 6: Dimension reduction process

The simulation scenarios are designed by using the input matrix. The design is used for determining the BES performance data based on the simulation of RC model as given in section 2.1. The simulation results were then used to calculate the zone temperature and its thermal energy consumption to generate the element effects, and the mean values and standard deviations of the element effects. The final step was to compare the standard deviations and mean values for determining the effect of each uncertain parameter so that the key uncertain parameters can be found out.

Morris global sensitivity analysis is a dimension reduction process that is used in this study. It offers the handling of large samples at low computational costs. Furthermore, it helps achieve optimized cooperation between efficiency and accuracy [68]. The following equation is used for determining the minimum number of required simulations in Morris analysis [69].

$$N_s = r(k+1) \tag{7}$$

Where N_s is the number of total simulations, r is the number of elementary effects per parameter, and k represents the number of design parameters.

From Morris analysis, two sensitivity indicators, i.e. mean value (μ) and standard deviation (σ) of the absolute values of the elementary effects as defined in Eqs. (8) and (9) respectively [69]. The input parameter's influence on the output value can be determined with the help of the mean value. Moreover, the interactions among the non-linear effects or the parameters can be determined with the help of the standard deviation. A plane (μ, σ) is used for presenting these methods in Morris analysis. The presentation of the plane also helps in comparing their relative influence [70].

$$\mu = \sum_{i=1}^{r} |EE_i|/r \quad (8)$$

$$\sigma = \sqrt{\sum_{i=1}^{r} |EE_i - \mu|^2/r} \quad (9)$$

where EE is the elementary effect, and r shows the total number of elementary effects for each parameter.

The elementary effect EE is derived from a model $y = y(x1, \dots, xj)$ with j input parameters, i.e. $x_1, \dots, x_j$. The EE for the i^{th} input parameter at the k^{th} sampling point is calculated by Eq. (10).

$$EE_i^{(k)} = \frac{y\left(x_1^{(k)}, x_2^{(k)}, \dots., x_{i-1}^{(k)}, x_i^{(k)} + \Delta, x_{i+1}^{(k)}, \dots, x_j^{(k)}\right) - y(x_1^{(k)}, \dots, x_j^{(k)})}{\Delta} \quad (10)$$

The proper sampling of each input variable within its defined range is required for the success of the Morris sensitivity analysis. This study used the Latin hypercube sampling method that can generate a specific amount of discretized values within the range of every parameter. This method helps improve the efficiency of the Morris method [71].

2.2.3. Bayesian optimization

Following the proposed framework, the Bayesian calibration [72] tends to incorporate the uncertainty of the parameters in the calibration process, and correct any inadequacy associate with them to keep the difference between the model prediction and observed values minimum.

In this study, a Bayesian approach is used for calibrating the uncertain parameters θ. The calibration of energy simulation models is performed in accordance with the building specifications and the observed data y. The calibration process aims to find the closest match between the observed data and the energy simulation data by deriving the values for θ. The method of Bayesian calibration (mathematical formulation) devised by Kennedy and O'Hagan [73] is used in this study. Three types of uncertainties are captured by the statistical formula: (i) discrepancy between the true behavior of building and the simulation model, (ii) parameter uncertainty in the energy simulation model, and (iii) errors because of the observations. Known conditions *x* are used for quantifying these uncertainties. Eq.(11) gives the relationship between the model outputs and the observations:

$$y(x) = \eta(x, \theta) + \delta(x) + \varepsilon(x) \quad (11)$$

$y(x)$ denotes the observations, the simulation model outputs that are computed at the given conditions x are given by $\eta(x, \theta)$. Known occupancy and external temperatures are examples of the known input conditions x. Even with the help of the calibration parameters, simulated models cannot perfectly predict the energy consumption of a building. Indeed, building energy models are based on approximations of the heat transfer processes occurring in a building. $\delta(x)$ represents the discrepancies between the physical behavior and simulated behavior of the building. This term helps in the over-estimation of the calibration values. It also indicates the shortages of the energy model. All the errors that occur during the recording of observations are given by $\varepsilon(x)$.

In the Bayesian paradigm, expert judgement is used for assigning the prior distributions $p(\theta)$ to the uncertain parameters. Several sources can be used for deriving it. Observations obtained from the formal set up are used for updating the prior distributions. In this system, the chances of obtaining observations

from the simulated model drive the updating process of the prior distributions. As a result of this, plausible distributions of calibration parameters that are also known as posterior distributions are collected.

For characterizing the performance of the developed model, the accuracy is evaluated by utilizing a group of performance metric of the metamodel. Following standard evaluation measures are included in the performance metrics: (i) R-squared (R^2), (ii) Coefficient of Variation ($C.V$), and (iii) Symmetric mean absolute percentage error ($SMAPE$), their formulas are expressed in Eqs (12) to (14).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{predict,i} - y_{data,i}\right)^2}{\sum_{i=1}^{n}\left(y_{data,i} - \overline{y}_{data}\right)^2} \tag{12}$$

$$C.V(\%) = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{predict,i} - y_{data,i}\right)^2}{\overline{y}_{data}}} \times 100 \tag{13}$$

$$SMAPE(\%) = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|y_{data} - y_{predict}\right|}{\left(\left|y_{predict}\right| + \left|y_{data}\right|\right)/2} \times 100 \tag{14}$$

In the above equations, $y_{predict,i}$ gives the predicted value at time point i. $y_{data,i}$ is the actual value at time point i, n represents the total number of the data points, and k gives the number of regressors. There are two main reasons for which the $C.V$ is used as an objective function in the model set. The first reason is that its index allows to determine how well a model fits the data by capturing offsetting errors between measured and simulated data. Therefore, it does not suffer from the cancellation effect [27]. The second reason is the recommendation of ASHRAE for using $C.V$ in evaluating the BES models [74].

2.3. Uncertainty analysis (UA)

The uncertainty analysis requires two basic functions; the deterministic model $y = g(x)$ of the inspected system (section 2.1) and a probabilistic framework for the uncertain variables of the model (Section 2.2.1), The initial step of UA is to generate N samples for every k uncertain variables by utilizing their probability distributions. Then, the samples are kept in a matrix $\boldsymbol{X}$ Eq. (15). The j^{th} sample of a parameter X_i is given by $x_{i,j}$ based on Latin hyper cube design, In the space-filling design, we consider the range [0,1] divided into N intervals of the equal length $1/N$. One point is selected at random from each interval forming a sequence of N points in $H^1\{1\}, 1,..,i\,x\,i = N$. Similarly, but independently we construct another sequence $\{x_i^2\}, 1,..,i\,x\,i = N$. The two sequences $\{x_i^1\}, 1,..,i\,x\,i = N\ and\ \{x_i^2\}, 1,..,i\,x\,i = N$ can be paired to populate a bidimensional space. These N pairs can in turn randomly be combined with the N values of $\{x_i^3\}, 1,\ldots,i\,x\,i = N$ to form N triplets, and so on until an n-dimensional sequence of N is formed. The Latin hypercube has better space-filling properties than another random sampling since it allows us to obtain a full distribution of the output in less time than Monte Carlo; that is why; it is used for sampling from the parameter distribution [71].

$$\boldsymbol{X} = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{k,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{k,1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,N} & x_{2,N} & \cdots & x_{k,N} \end{bmatrix} \tag{15}$$

After developing the sample matrix, the next step is to evaluate N Monte Carlo of the deterministic model. It is evaluated one per row of a matrix to get the model response, $\boldsymbol{y}$, see Eq. (16). The uncertainty impact can be quantified through the evaluation of the statistical and variability characteristics.

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} g(x_{1,1}, x_{2,1}, \ldots\ldots, x_{k,1}) \\ g(x_{1,2}, x_{2,2}, \ldots\ldots, x_{k,2}) \\ \vdots \\ g(x_{1,N}, x_{2,N}, \ldots\ldots, x_{k,N}) \end{bmatrix} \quad (16)$$

While using the Monte Carlo simulations, a random question arises regarding the number of iterations for the sufficient accuracy of the results. Based on the availability of the competitional resources, 10,000 simulations are selected as a thumb rule for the current study.

2.4. Sensitivity analysis (SA) for building performance

After identifying the uncertainty associate with the parameters in the calibrated BES model based on Monte Carlo simulations, SA is performed to identify the most influential uncertain parameters in the building performance. The main output indicators are; the total energy thermal consumption at each building zone (Q_{Heat}) and the total over/underheating hours ($t_{o/u}$) in which the operating temperature at each zone gets out of the thermal comfort conditions (20:24 °C) identified by the ISO7730 [75].

Given the drawbacks in the local sensitivity analysis approach [76], a two-step global sensitivity analysis (GSA) approach is proposed in this study, as shown in Fig. 7. In the first step, the Morris analysis is applied to screen the insignificant uncertain parameters and rank them in a low computational cost, as mentioned in Eq.(7) in terms of two objective outputs (Q_{Heat} and $t_{o/u}$). Noting that this ranking and screening phase is only qualitative approach, thus it cannot quantify the important of each parameter. Following this limitation, another GSA technique is proposed in the second step to provide more quantitative information regarding the uncertain investigated parameters.

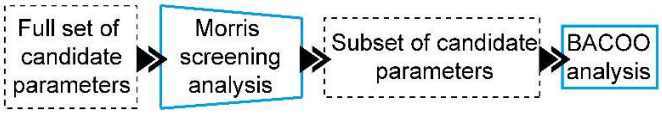


Fig. 7: Two-steps GSA approach including screening phase (first step) based on Morris analysis and a second step using BACOO method using the subset of candidate parameters from the first step

In the second step, the Bayesian analysis of computer code outputs (BACCO) [73] is employed. The advantage of this method includes investigating a full range of the input parameters and their distribution. In addition, it offers a substantial declination in the computational cost compared to other GSA methods [77]. The BACCO approach has two main key stages. In the first stage, a set of training data are utilized to develop a statistical model derived from the RC model simulations. The training data tends to cover the whole problem domain based in a multidimensional filling algorithm (Latin hypercube design). In order to verify the model accuracy, cross-validation is applied to the developed model. In the second stage, the statistical model is proposed to in quantifying the importance of parameters in interest and their relative interaction due to its low computational expenses [78].

3. Case study

3.1. Building general description

The application of the proposed framework is illustrated through a two-story dwelling located in a district of Emmen at the Netherlands. A layout for the dwelling is shown in Fig. 8, where the total floor area is 60 m^2, and the ceiling height is 3 m. The considered dwelling has two external façades with a total area of 33.7 m^2, oriented to the East and West, whereas the opaque part represents 19.6 m^2. The first floor consists of

living room, kitchen and the entrance. While the second floor has 3 bedrooms, one storage room, one bathroom and the hallway.

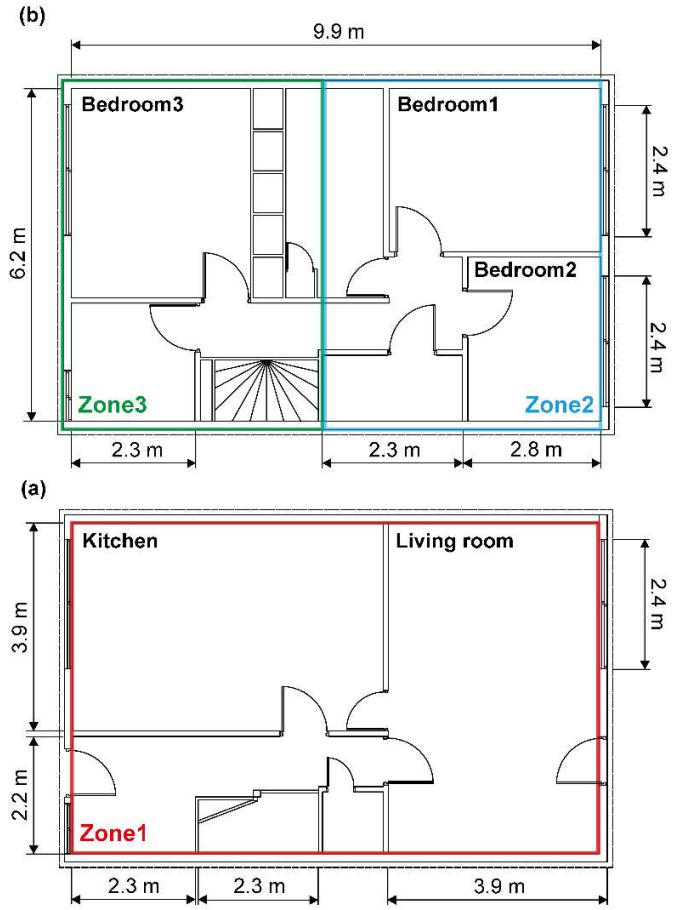


Fig. 8: A layout for the two-story Emmen's dwelling: (a) first floor (Zone 1) highlighted in red, and (b) second floor (Zone 2&3) highlighted in blue and green, respectively

A 3-zone model is applied in Emmen's dwelling. The first floor is an open kitchen and living room with two temperature sensors. Therefore, as it is seen in Fig. 8, all first floor is defined as Zone 1. All bedrooms have temperature sensors, and the orientation of the bedrooms is different. Whereas bedroom 1 and 2 face towards the west and bedroom 3 faces towards east. Moreover, it can be possible that 2 bedrooms are heated, and one is unheated and vice versa. Therefore, to distinguish the conditioned and unconditioned area and for the more accurate consideration of the solar gain, floor 2 is divided into 2 zones as it is seen in Fig. 8. 60% of the total floor area is considered as Zone 2, and the rest is Zone 3.

3.2. Building construction feature

This building is a newly renovated home where all insulation of roof, floor, walls and glazing are reconstructed in passive level to enhance the dwelling energy-efficiency. The external walls are composed of two main walls; the outer walls comprise hollow bricks covered with layer insulation and plaster. While the inner walls are gypsum boards over light concrete. The ground surface composed of two main surfaces; the ground surface which comprises stones, concrete and insulation materials, whereas the inner floor surface composed of light concrete, and floor layer over gypsum boards. Finally, for the roof surface, it composes of three layers, including concrete and plaster over insulation materials.

As far as the windows are in concern, double-glazing windows with a total area of 12.1 m^2 are constructed in the western façade and 11 m^2 for the eastern façade. A summary of the construction material properties is shown in Table 2.

Table 2: Thermal properties of the building envelopes

Parameter name	Unit	Value
RC_{roof}	[m^2K/W]	5
RC_{floor}	[m^2K/W]	5
RC_{facade}	[m^2K/W]	4.7
U_{glass}	[W/m^2K]	1.1
G_{value}	-	0.7

3.3. RC Model boundary conditions

For the development of the RC model, several monitoring data are required to trace the occupant behavior and equipment operation effects in the development of an accurate predictive model. The acquisition was performed for nine months (September 2017 to May 2018). This data collection guarantees to provide a quality data set during the whole heating period (autumn, winter and spring). On the other hand, the summer period was excluded in the data acquisition process since these dwelling doesn't have a cooling system. Thus, the summer period does not influence the annual energy thermal consumption of the building. The dwelling measurements were carried out while the dwelling was occupied in order to reflect the occupant behavior interaction with the dwelling. The measurement was taken within 10 min frequency and integrated over 1 hour, which is a more suitable frequency for building measurements over long periods. The measurement comprises the climate input data, the equipment operation and occupant activities.

3.3.1. Climate input data

To obtain real-time weather data for the dwelling, the climate data are obtained from a KNMI weather station. The nearest KNMI weather station to the dwellings in Emmen is in the Hoogeveen (STN:279, LON(east):6.574, LAT(north):52.750, ALT(m):15.80) which is approximately 25-30 km away from the dwellings. The monthly average ambient temperature and its relative incident solar radiation on a horizontal surface are shown in Fig. 9.

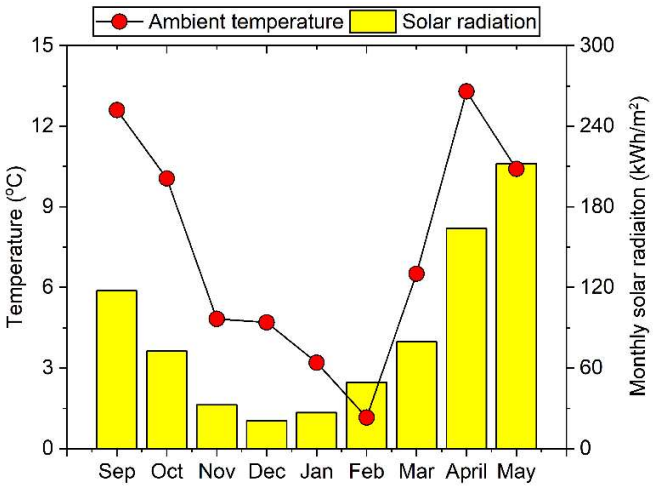


Fig. 9: The climate conditions in the dwelling in Emmen from September 2017 to May 2018

In addition, the wind speed, as well, was extracted from the KNMI weather data. While the wind pressure coefficient on the building envelope is assumed to be 0.1. On the contrary, the ground temperature data is not available in the KNMI weather data file. Therefore T_{ground} is assumed as 10 °C and constant during the all year (although it varies from 7 °C to 13 °C during the year). However, the dwellings have a crawlspace which is naturally ventilated. Therefore, the average of the ground and outside temperature $\left(\frac{T_{ground}+T_{out}}{2}\right)$ is used in the model to calculate the transmission loss through the ground floor. Finally, to consider the shading by surrounding objects and the dwelling itself, the shading factor (f_{shadi}) is introduced. The monthly shading factor values are in Table 3.

Table 3: Monthly shading factor

Month	Sep	Oct	Nov	Dec	Jan	Feb	Mar	April	May
Shading factor [%]	0.3	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.3

3.3.2. Equipment operation

The building is heated up through radiators on all levels that are primarily provided by an air source heat pump (PUHZ-SW50VHA). The heat pump compressor controller managed the required flow rate based on the thermostats and the zone setpoint (user adjusted). Therefore, a variation in the heat pump performance factor (COP) is recognized with the variation in the ambient air temperature where the observed COP is up to 5, More details regarding the heat pump model is found in S1. Base on the time-dependent of the COP and its relative heat pump electricity consumption ($HP\ Power_{electric}$), the thermal heat pump power ($HP\ Power_{therm}$) is estimated as following:

$$HP\ Power_{thermal} = COP(time) * HP\ Power_{electric} \quad (17)$$

As shown in Fig. 10, the maximum monthly thermal energy consumption occurs during February (1028.9 kWh_{th}).

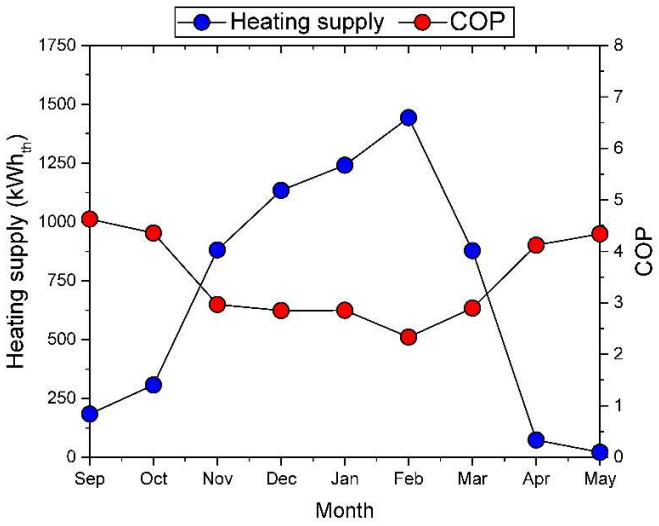


Fig. 10: The pump heat performance during the heating period

For the mechanical ventilation, the hourly flow rate at each zone is estimated through the measurement of ventilation position throughout the operation period, and flow rate for the different ventilation positions at all the three zones. Since there is no specific sensor data to obtain the ventilation position, the energy consumption of the heat-recovery/ventilation for all dwelling is utilized for ventilation position. The mechanical ventilation system has three main positions during its operation, as shown in Fig. 11-(a). The relative heat recover, and the mechanical ventilation flow rate at each zone is shown in Fig. 11-(b),(c),(d).

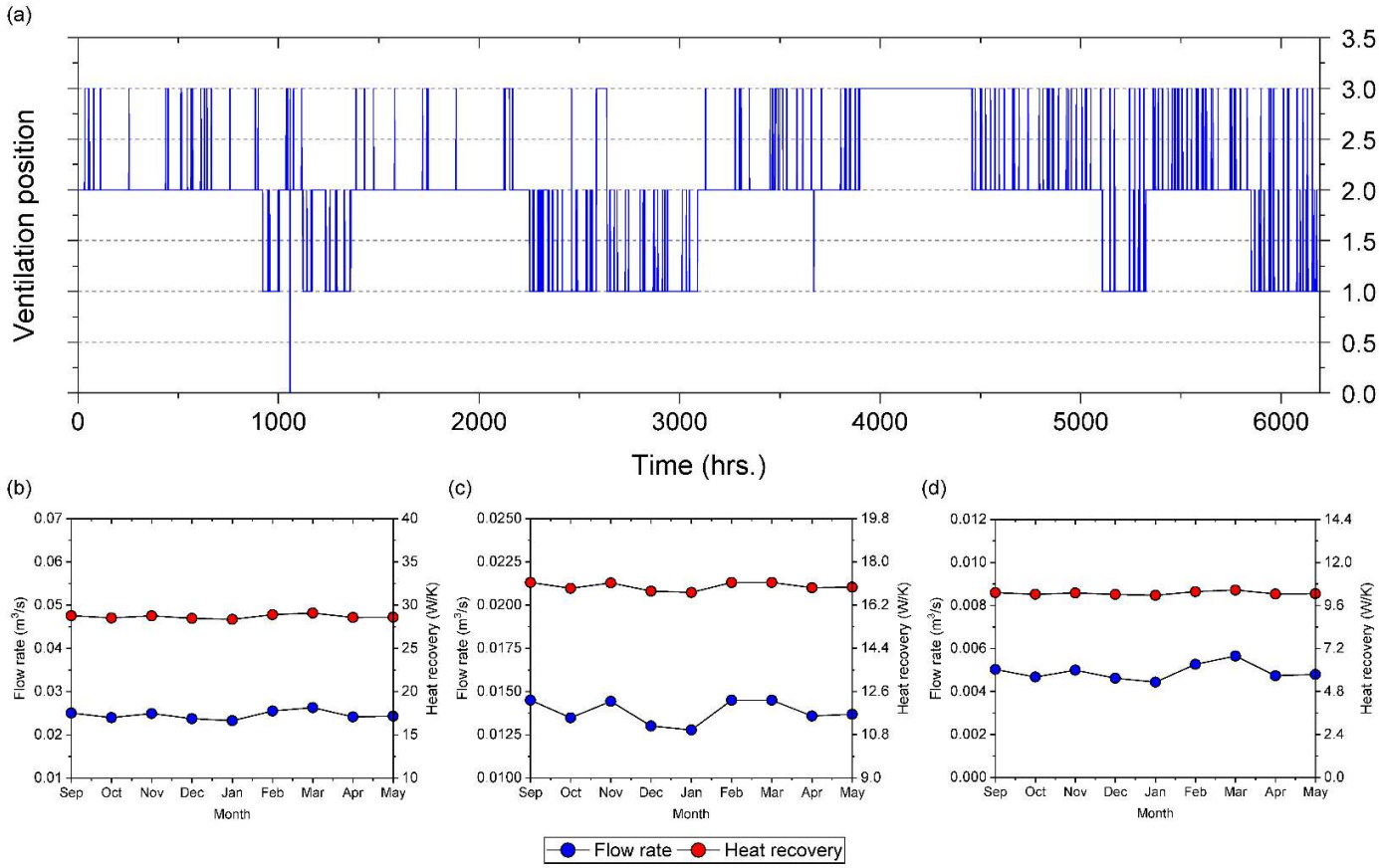


Fig. 11: The mechanical ventilation position throughout the heating season aligning with its relative flow rates and heating recovers at each zone where; (a) The mechanical ventilation position, whereas (b), (c) and (d) mechanical ventilation flow rate and its relative heat recovery in Zone 1, 2 and 3, respectively

3.3.3. Occupant activities

Since the occupant activities have a significant effect in the building energy performance, the measurements are taken under normal operating conditions in order to calculate the thermal response of the building under the different possible realistic operating conditions accurately. The occupant activities, including the setpoint temperature, the occupancy schedule, neighbour temperature and open/closing of window and doors are utilized as inputs in the developed RC model.

Occupancy data and their relative internal heat gain for each zone-type were determined based on human resources interviews, personnel counts, and multiple day/night occupancy surveys, as shown in Table 4. Furthermore, the internal heat gain is estimated to be 100 W per person during daylight and 70 W per person during night period where the distribution of occupants is 75% in zone 1 and 25 % in zone 2&3 during the daylight. While this percentage changes to 0% for zone 1 and 100% in zone 2&3 during the night period.

Table 4: Occupancy schedules for dwelling in Emmen

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Morning (08:00 - 13:00)	5	3	3	3	3	3	5

Afternoon (13:00 - 18:00)	5	5	5	5	5	5	5
Evening (18:00 - 23:00)	5	5	5	5	5	5	5
Night (23:00 - 08:00)	5	5	5	5	5	5	5

In the dwelling at Emmen, the opening sensors are available in all windows and doors. Thereby, it is possible to detect when the window and door are closed or open. The results of opening sensors are available for Emmen's dwelling in Fig. 12. The opening and closing windows/doors operation include the detection of its opening and its relative period where 1 means that the window/door is opened for 1 hour while zero represent a complete closing for an hour.

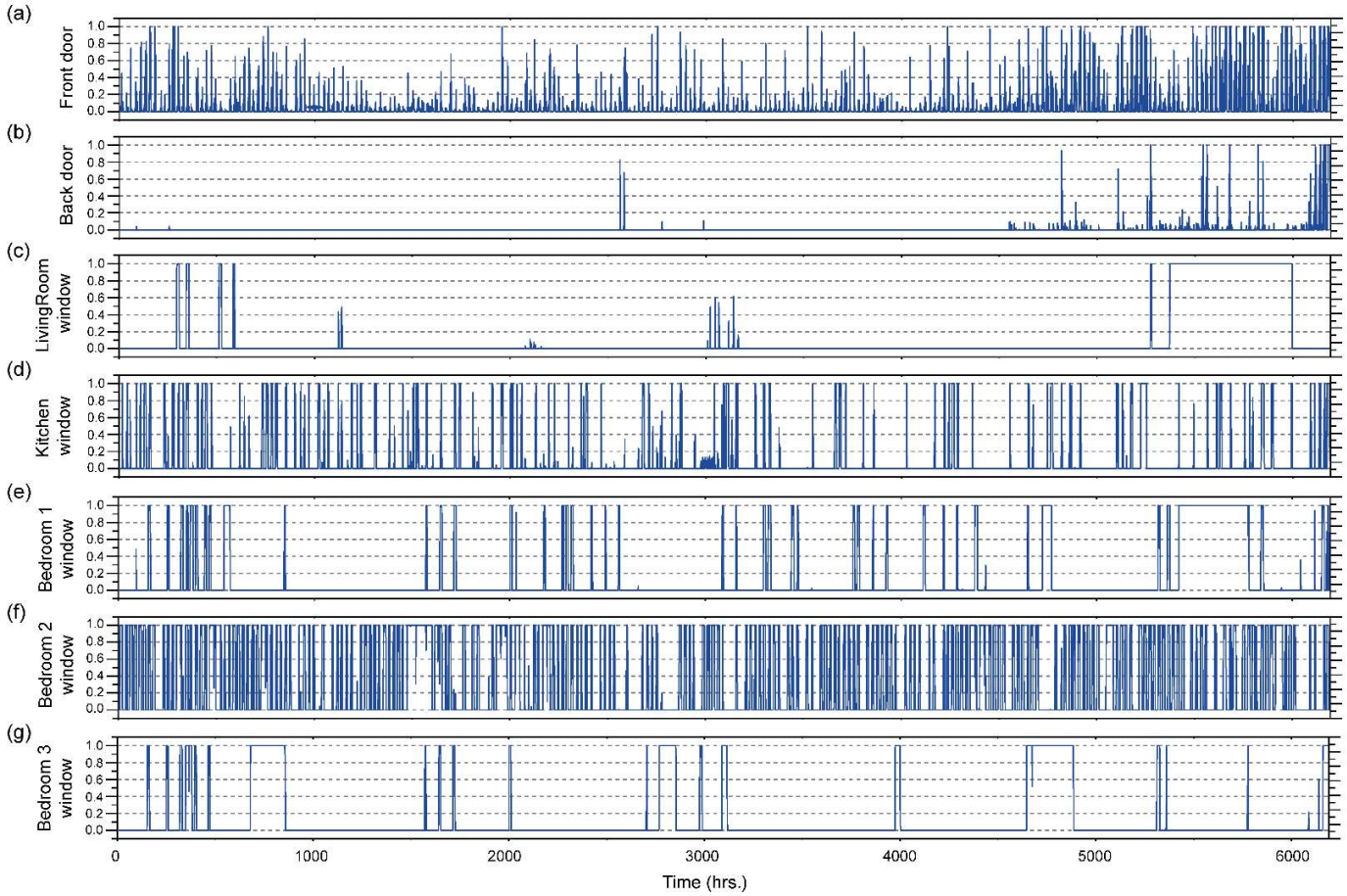


Fig. 12: Windows and doors opening schedule for dwelling in Emmen

In addition to the occupancy and window operation schedules, the neighbour temperature is projected based on it is the neighbour's indoor temperature. It is used to calculate the heat loss or gain through the neighbour/s. Finally, the heating set-point temperature is also measured for zone 1, and it is utilized to calculate the required heating power. In contrast, the setpoint temperature for zone 2&3 is considered in a fixed proportional band to zone 1. A summary for the utilized equipment in the measurements and their error is shown in Table 5.

Table 5: Measurement equipment and their relative error

	Sensor Name	Measurement error
Temperature	Wireless M-Bus - CMa12w	0.2 degrees
Power	Energy meter - SDM630	< 1% of range / Power factor 0.01
Open/close windows	Window Sensor - FGK-101	NONE / Binary sensor
Weather data	KNMI (Dutch national weather institute)	

4. Result and discussion

The results in the current study are divided into three main parts. The first part depicts the calibration of the RC model based on the Bayesian optimization approach incorporating Morris analysis for dimension reduction. The main purpose of this phase is first to define the most relevant uncertain parameters in the RC model and then calibrate them with consideration for the temporal variation (Time dependency) inherent in the relative uncertain parameters. Once the RC model performance is verified and calibrated in terms of the energy performance and zones temperature, the second part purposes a quantification for the uncertainty associate with RC model based on Monte Carlo simulations. Finally, a two-steps global sensitivity analysis is purposed to express the contribution of the uncertain parameters to the variation in the dwelling performance adequately.

4.1. Model calibration

4.1.1. Dimension reduction results

In order to carry out the dimension reduction based on Morris analysis for the developed RC model, 31 uncertain candidates are select for the three zones in Emmen's dwelling, as shown in Table 6. The Morris analysis in this stage is utilized to estimate all parameters that affect the accuracy of the RC model performance in terms of predicting the total thermal energy consumption and the zones temperatures. Given a total number of parameters k =31, the number of elementary effects r is selected as 100, leading to $r(k+1) = 3200$ model performance evaluation. For all candidate design parameters, the performance of Morris analysis was quantified based on the mean value (μ) for the element effect and the standard deviation (σ) for its potential for interaction with other parameters.

Table 6: Candidate uncertain parameters at Emmen zones classification based on the parameter category

	Uncertain parameter			
Parameter category	Whole building	Zone 1	Zone 2	Zone 3
Input climate data	1. Q_{sol}	-	-	-
	2. V_{wind}	-	-	-
	3. T_{out}	-	-	-
	4. T_{ground}	-	-	-
Building envelope	5. RC_{roof}	-	-	-
	6. RC_{floor}	-	-	-
	7. RC_{facade}	-	-	-
	8. U_{glass}	-	-	-
	9. G_{value}	-	-	-
	10. $f_{shading}$	-	-	-
	-	11. $Infl._1$	12. $Infl._2$	13. $Infl._3$

	-	14. f_{c_1}	15. f_{c_2}	16. f_{c_3}
Equipment system	17. η_{fan}	-	-	-
	-	18. $vent._1$	19. $vent._2$	20. $vent._3$
Occupant behavior	21. $Occup_{sch}$	-	-	-
	22. $Occup_{int}$	-	-	-
	-	23. $\theta_{Living\ room}$	27. $\theta_{Bedroom\ 1}$	29. $\theta_{Bedroom\ 3}$
	-	24. $\theta_{Kitchen}$	28. $\theta_{Bedroom\ 2}$	-
	-	25. $\theta_{Entrance}$	-	-
	-	26. θ_{Exit}	-	-
	-	-	30. T_{set_2}	31. T_{set_3}

All the Morris analysis results for the total thermal energy consumption and temperature at each zone are shown in Fig. 13. In the plot analysis, we consider the most 14 influential parameters from the Morris method and label them according to their μ and σ values. The important parameters are marked with circles, whereas the remain parameters are considered less important and marked with X. For the prediction of the total thermal energy consumption, it can be seen that the infiltration rate in zone 2 (i.e. factor 12) followed by the infiltration rate in zone 1 (i.e. factor 11) are the most influential parameters due to their high μ. While the correction factor of thermal mass fraction in zone 1 (i.e. factor 14) followed by the window opening angle of kitchen room (i.e. factor 24) have the highest possibility for interactions with other parameters due to their high σ. In term of the temperature at zone 1, the window G-value (i.e. factor 9) and the correction factor of thermal mass fraction in zone 1 (i.e. factor.14) are the most influential parameters since they have the highest μ and σ. This is followed by the infiltration rate in zone 1 (i.e. factor 11). Furthermore, the window opening angle of kitchen room (i.e. factor 24) followed by wind speed (i.e. factor 2) have a high potential for interaction with other parameters due to their high σ. For the temperature prediction in zone 2, the correction factor of thermal mass fraction in zone 2 (i.e. factor.15), followed by the window G-value (i.e. factor 9) and the infiltration rate in zone 2 (i.e. factor 12) have the highest influence in the temperature in Zone 2. Furthermore, the window opening angle of bedroom 1 (i.e. factor 27) have high interaction potential. Finally, for the temperature in zone 3, the most influential parameter is the correction factor of thermal mass fraction in zone 3 (i.e. factor.16). The second most influential parameter is the window G-value (i.e. factor 9), followed by the window opening angle of bedroom 3 (i.e. factor 29). Noting that occupant behavior has a significant impact on the performance of the model accuracy, as mentioned by [64]. Therefore, including more occupant behavior parameters such as opening/closing window schedule can significantly change the contribution of each uncertain parameter to the model accuracy. However, due to the lack of extensive knowledge regarding the uncertainty associate with these parameters, we keep them as deterministic parameters.

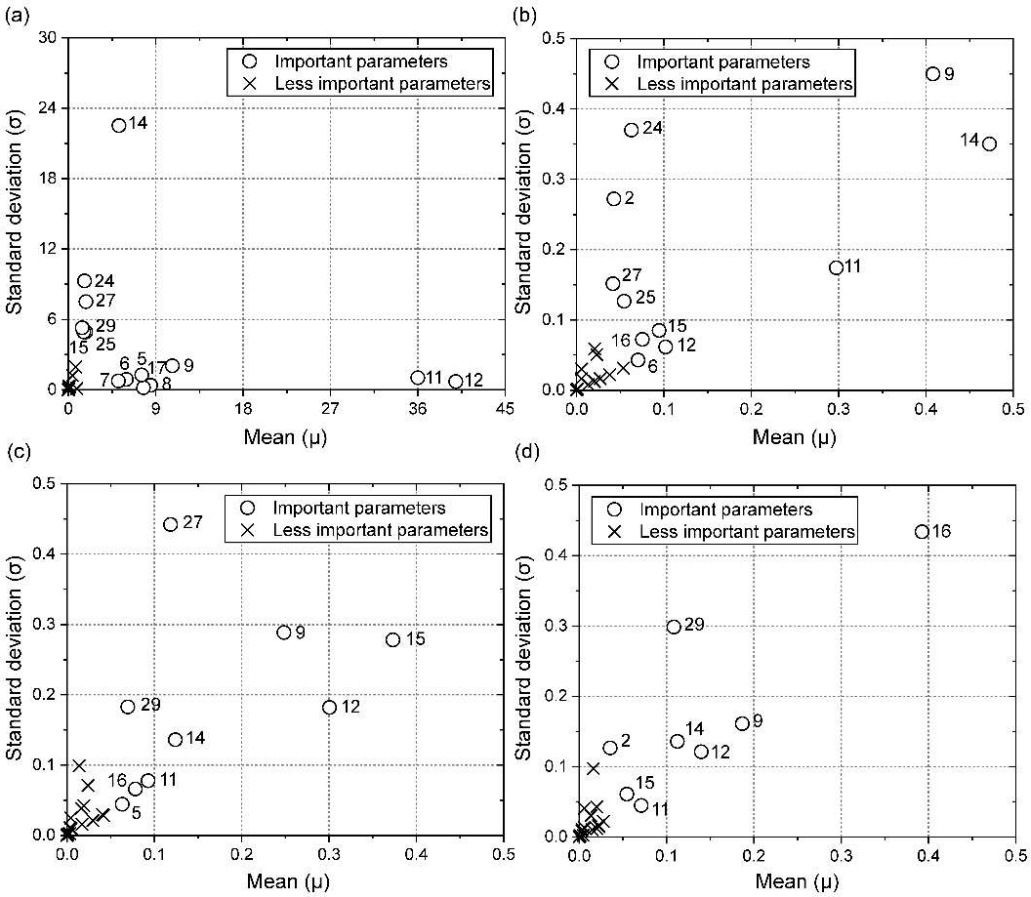


Fig. 13: Results of the Morris sensitivity analysis to evaluate the parameters affect the RC model accuracy in predicting; (a) the total thermal energy consumption, (b) temperature at zone 1, (c) temperature at zone 2, and (d) temperature at zone 3

The remaining uncertain parameters can be considered as the parameters with a less impact on the prediction of the RC model performance, and they set as constant values (see Table 7) based on the case study data entry mentioned in section 3.

Table 7: Low sensitivity parameters and values used

Parameter	**Value**
1. Q_{sol}	KNMI weather station
3. T_{out}	KNMI weather station
4. T_{ground}	10 °C
10. $f_{shading}$	Monthly shading factor (Table 3)
13. $Infl._3$	0.31 l/s
18. $vent._1$ / 19. $vent._2$ / 20. $vent._3$	Ventilation position profile (Fig. 11)
21. $Occup_{sch}$	Occupant questionnaire (Table 4)
22. $Occup_{int}$	100 W (daylight) & 70 W (night)
23. $\theta_{Living\ room}$ / 26. θ_{Exit} / 28. $\theta_{Bedroom\ 2}$	10%
30. T_{set_2} / 31. T_{set_3}	T_{set_1} profile

4.1.2. RC Model setting (Bayesian optimization)

Following the screening phase using Morris analysis to eliminate the less important uncertain parameters, the Bayesian optimization is proposed in this stage to calibrate and produce an accurate RC model that can predict the dwelling thermal energy consumption as well as the zones temperature more accurate. Based on 500 iterations in the Bayesian model, the optimal values for uncertain parameters are estimated for the time-independent and dependent uncertainty parameters.

A summary of the time-independent parameters is shown in Fig. 14. For the building envelope parameters, a large diversion from the initial construction materials design is observed in the optimal parameters where the thermal resistances are higher by 38% and 34% for the roof and floor surfaces, respectively. While the thermal resistances of the façade surface is lower by 17% from its initial value. The difference between the initial and the Bayesian optimal values are due to the aging and the difference between the design specifications and the construction outcomes. For the windows construction properties, a limited deviation is detected for the U_{glass} where it is lower by 19% from the initial values, whereas the G_{value} is higher by 32%. In terms of the infiltration rates, the deviation between the initial value and optimal estimate values is about 80% for zone 1 and 43.3% for zone 2. This deviation in the infiltration rate is a function of several parameters comprising the construction quality, the building use as well as the weather conditions. For the equipment operation presented by the η_{fan}, a deviation of 5% is estimate from the equipment specification. Finally, for the correction of the thermal mass parameter, this value was set to 0.3 for all zones. However, a big deviation from the initial values is observed for all f_c where the optimal value is 0.81, 0.14 and 0.45 for zone 1, 2 and 3, respectively.

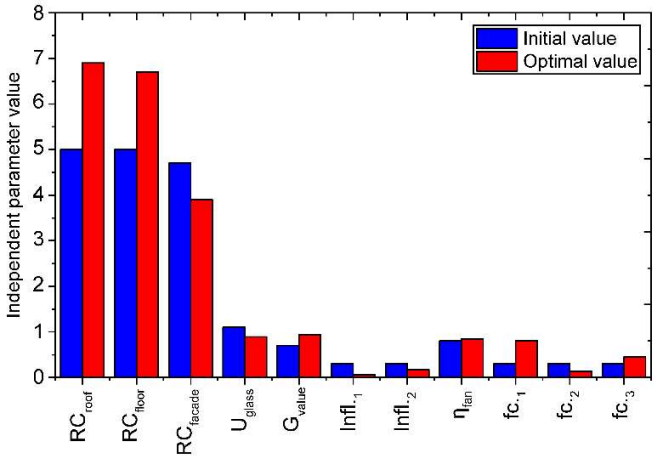


Fig. 14: The initial vs optimal values for time-independent parameters

On the other hand, a summary of the time-dependent uncertain parameters is illustrated in Fig. 15. For the wind speed, the optimal scenario over the 36 weeks (heating seasons) has a limited difference from the data extracted from the KNMI weather station. The maximum deviation is about 9% in week 24 (February). While for the parameters related to the opening angle of windows and doors 'θ_i', a big deviation is indicated in all windows and doors from the initial proposed values. This deviation is expected since the θ_i is one of the most controversial uncertain parameters in the BES models. For the $\theta_{Entrance}$, the angle of opening is stayed around 45.5%±3.5 for the whole 36 weeks. Moreover, the windows $\theta_{Kitchen}$, $\theta_{Bedroom\,1}$, $\theta_{Bedroom\,3}$ stay around for the 45%±3 over all the operation period.

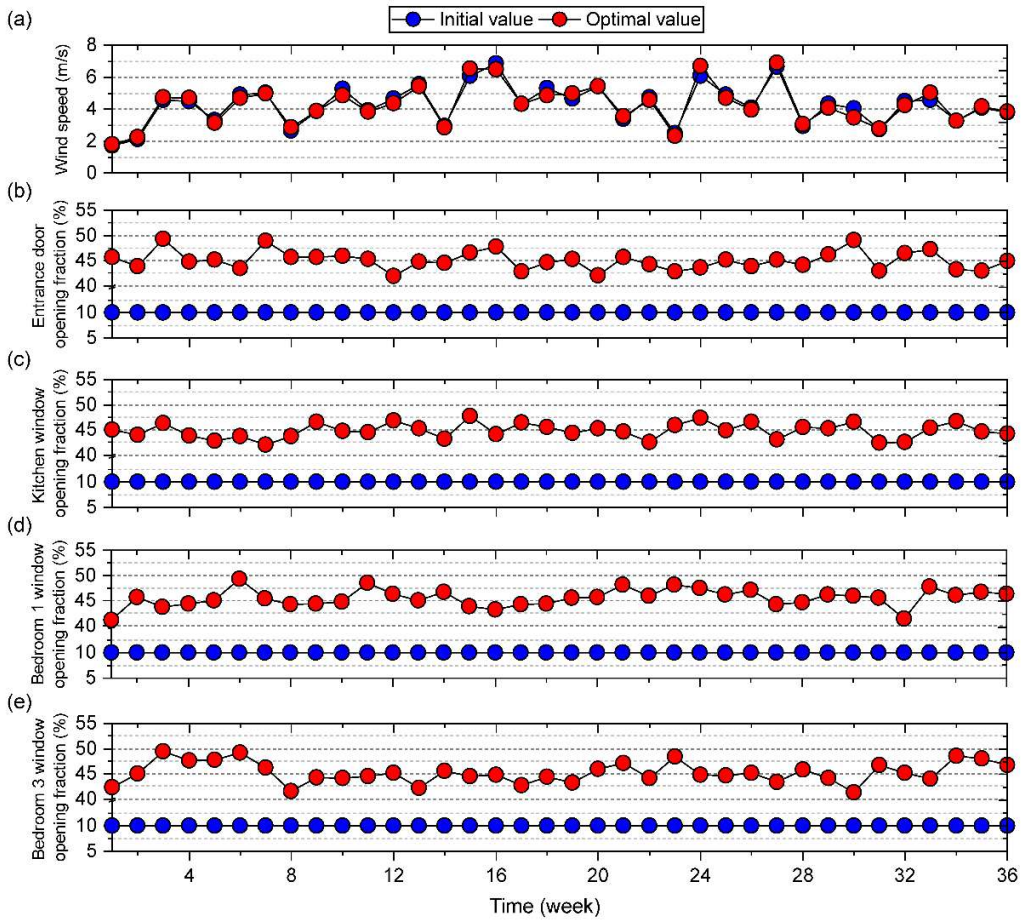


Fig. 15: The initial vs optimal values for time-dependent parameters

4.1.3. Calibration performance

Following the calibrated parameters based on the Bayesian optimization approach, the developed RC network for the digital dwelling using zone thermal coupling method is compared with the measured hourly data of 9 months of operation (heating seasons). The developed model performance is validated in terms of the indoor air temperature at the three investigated zones and the cumulative thermal energy consumption by the dwelling. According to the indoor air temperature profiles shown in Fig. 16, a good agreement is indicated between the prediction of the RC model and the indoor measured temperatures at all the dwelling zones. In zone 1, the $C.V$ is 3.63% and $SMAPE$ is 5.04%. While the maximum error of 4.5 ^{o}C is indicated during the end of the spring season (April). For upper floor (Zone 2 and 3), a higher error is indicated within the RC model prediction for indoor temperature where the C.V is 5.53%, and 7.03% for zones 2, and 3, respectively. While the SMAPE stayed around 3.54% and 4.35% for both zones 2 and 3, respectively. Furthermore, the maximum error of around 9 ^{o}C for the prediction of the indoor temperature is zone 2, and 3 is indicated during the winter season (end January).

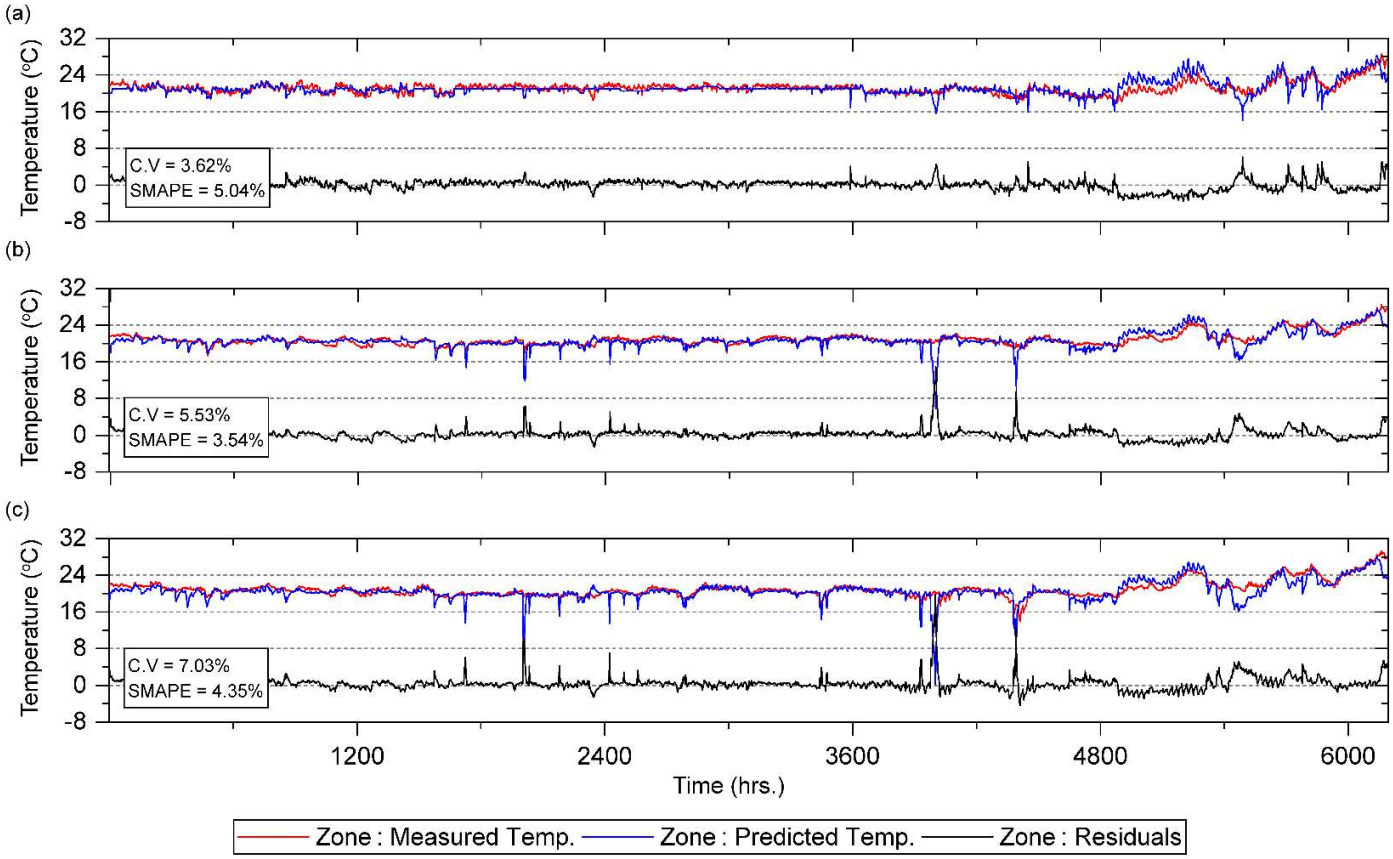


Fig. 16: Hourly indoor air temperature in °C for the RC model prediction in comparison to the Emmen's dwelling monitoring data

Fig. 17 represents the cumulative heating demand results of the RC model in the Emmen's dwelling during 9 months of operation. In this plot, the results of the RC model (Predicted $Q_{heat\ total}$) are compared to the measured heating demand for dwellings (Measured $Q_{heat\ total}$). A good performance is indicated in the RC model for the prediction of the $Q_{heat\ total}$ where the predicted value is 4.05 MWh with an error of 2.2% from the measured $Q_{hea\ \ total}$. In addition, the C.V and SMAPE show that the prediction error doesn't excess 2%. Finally, the residual analysis shows that the maximum cumulative thermal energy consumption difference between the measured and the predicted $Q_{heat\ total}$ is around 0.25 MWh.

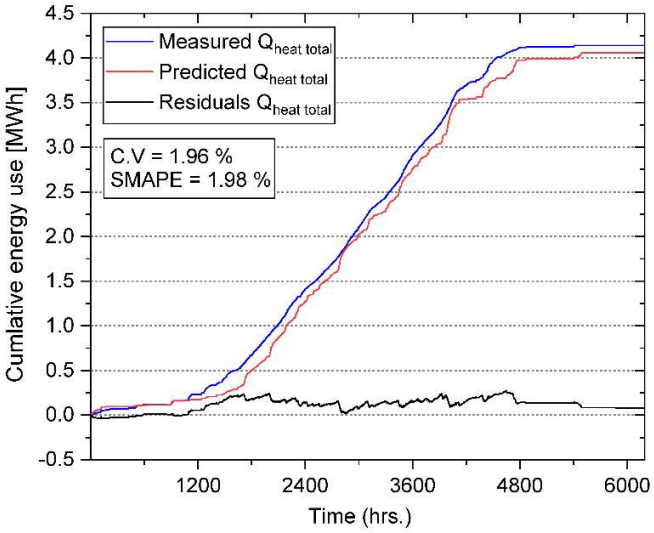


Fig. 17: The predicted cumulative thermal energy consumption at Emmen's dwelling in comparison to the monitoring data

In order to test the performance of the RC model at a different scale of prediction, the energy signature results for the Emmen's dwelling is presented in Fig. 18. In this plot, the results of the RC model (predicted thermal power) are compared to the measured thermal power of dwellings. The energy demand and the outdoor temperature is averaged on a weekly basis. The dots in the figure represent the data points, and the lines are obtained from these data points by fitting where the energy demand decreases with the increase in outside temperature. The results confirm an agreement between the measured and predicted weekly thermal power where the R^2 equal to 93%. The validation of the hourly indoor air temperature and the building thermal energy performance reflect the ability of our methodoly approach cover the limitation associate with the RC models in predicting the thermal energy performance for a long-term period successfully.

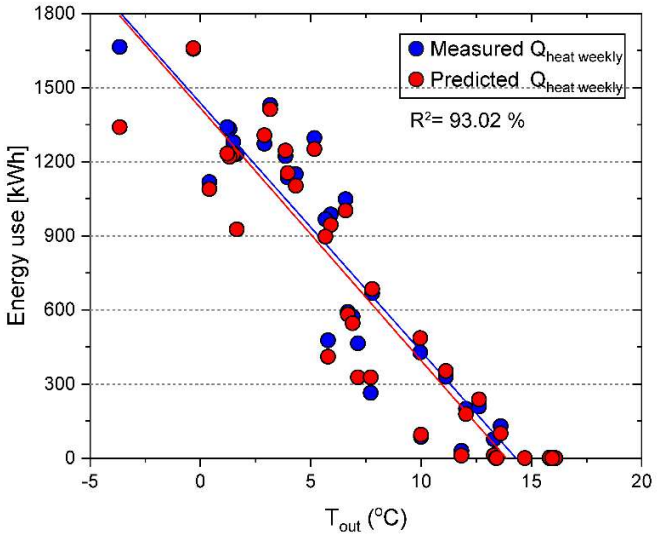


Fig. 18: The weekly predicted vs measured energy signature for the Emmen's dwelling

4.2. Uncertainty analysis (UA) results

In recent years, several contractors in Netherlands (e.g. Van Wijnen and BAM) offer energy-neutral housing in both the construction market and existing construction market. However, there is a burden of proof with the builder for zero on the gauge of the housing due to the uncertainty associate with the performance of the building. This burden increases due to the residents who have not yet confidence regarding the building performance guarantees. Therefore, it is important to guarantee of performance of these building independently on uncertain dwelling parameters. The long-term guarantee of the quality of the home, both in terms of thermal energy consumption and indoor climate, will lead to forecast the demand for these homes, and subsequently will boost energy savings of the housing, both new construction and renovation.

In this context, once the RC model is calibrated, and its performance is validated in comparison to the monitoring data. The UA based on Monte Carlo simulations is performed to quantify the uncertainty associate with the performance of the developed digital dwelling in terms of the indoor temperature at each zone and the total thermal energy consumption which is reflected in the building energy efficiency rating. For the UA step of the methodology, 10,000 MC evaluations of the developed model are performed using the uncertain parameters proposed in Table 6.

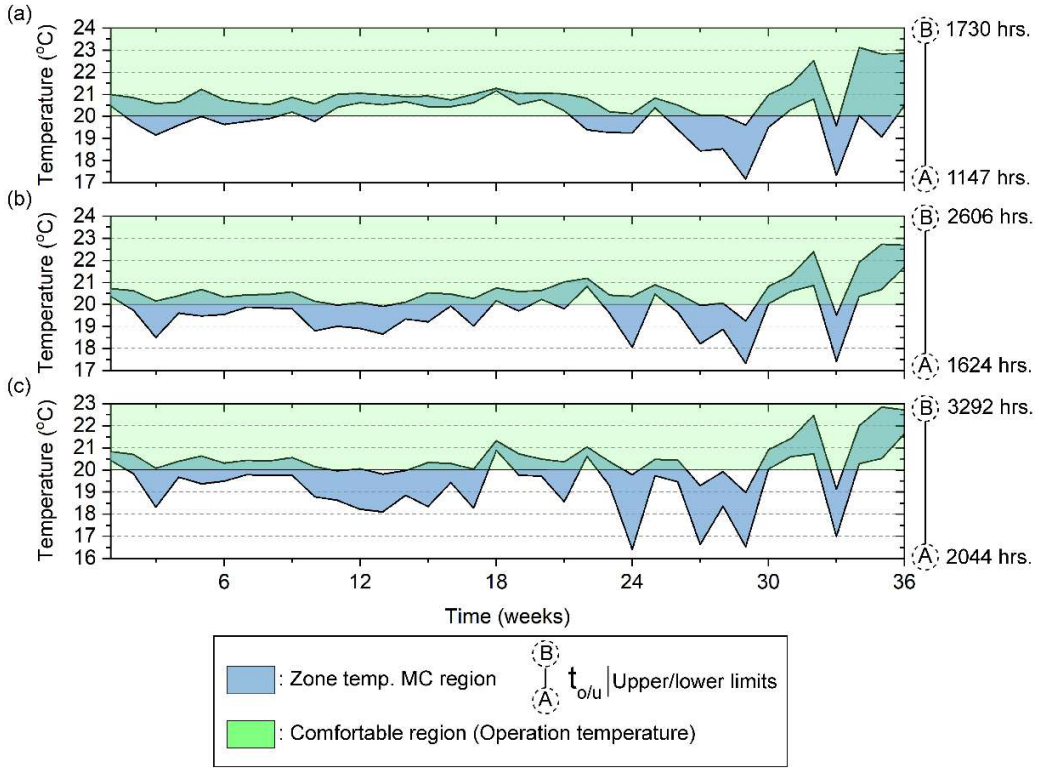


Fig. 19: The indoor temperature range at different zones based on the UA where (a) indoor temperature in zone 1, (b) indoor temperature in zone 2, (c) indoor temperature in zone 3. The blue filled areas represent a range for the indoor temperature at different zone under uncertainty, the light green region represents the temperature thermal comfort conditions (20:24 °C), and the dotted circles between A and B anchor points represent the upper and lower limited number of hours out of the thermal comfort conditions

Fig. 19 illustrates the variation in the zone indoor temperature using the blue plots and its relative $t_{o/u}$ using the dotted circles. In zone 1, a limited variation is indicated in the indoor temperature throughout the heating

operation seasons where the maximum variation reveals by the end of the operation period (May). This limited variation is reflected in $t_{o/u}$ in which the zone indoor temperature gets out of the thermal comfort conditions mentioned by ISO7730 [75].The over/underheating hours varies between 1147 hrs to 1730 hrs throughout 9 months of operation. In zone 2, almost the same behavior is indicated as zone 1 except for two regions. First during weeks between 10 and 18 (November and December). While the second region appears by the end of the operation (May) where a limited variation in the indoor zone temperature is indicated in comparison to zone 1. This variation is reflected in the $t_{o/u}$ and it ranged between 1624 hrs. and 2606 hrs. Besides, zone 3 follows the zone 1 as well as 2 behaviors with a larger range of variation in the indoor temperature, especially during the winter season. This variation expansion is reflected in $t_{o/u}$ and it varies between 2044 hrs. and 3292 hrs. In general, the increment in the $t_{o/u}$ associated with zone 2 and 3 demonstrates a high impact for the uncertain parameters on the thermal comfort conditions at these zones.

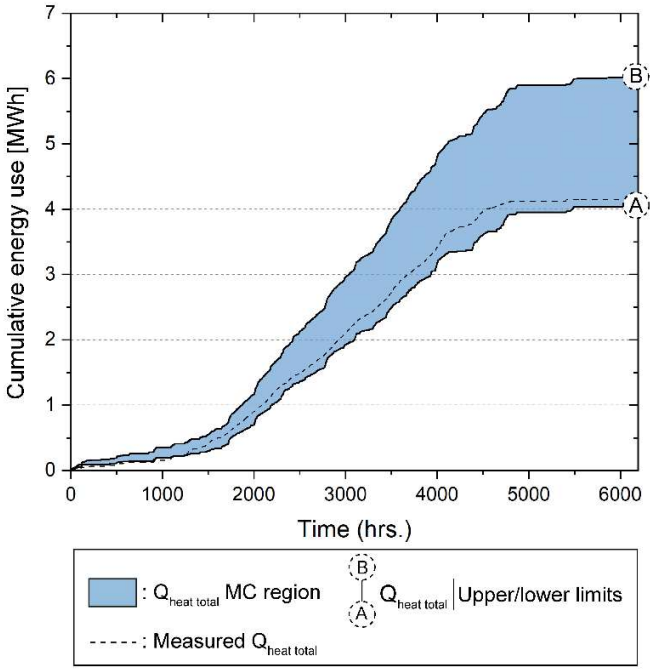


Fig. 20: The total thermal energy consumption variation range where the dotted circles between A and B anchor points represent the upper and lower limits. The blue filled areas represent a variation range for total thermal energy consumption vs the monitoring data in dotted line

Fig. 20 illustrates the variation in the total dwelling thermal energy consumption ($Q_{heat\ total}$) using the blue plots and the dotted circles. The UA shows a significant impact for the uncertain parameters in the thermal energy building performance where the $Q_{heat\ total}$ ranges between the 4.03 MWh (67 kWh/m^2/yr) to 6.01 MWh (100 kWh/m^2/yr) due to the uncertainty associate with climate data, the building envelope and occupant behavior. Accordingly, the high thermal energy consumption due to the possibility of uncertainty can decline the motivate the prospective owners or new tenants of the building to purchase or rent the building and subsequently reduce the market for energy-efficient renovation [79].

4.3. Global sensitivity analysis results

The UA results demonstrated the significant effects of uncertainty on the RC model output scopes comprising the indoor zone temperature and its relative $t_{o/u}$ and $Q_{heat\ total}$. The GSA is able to produce

additional understanding by recognizing the uncertain parameters that are majorly accountable for the output variations. The GSA technique can be used for investigating the main sources behind the variability in the RC model output. In this paper, the GSA technique is used for identifying the significant factors that are related to the variations in the total thermal energy consumption and indoor zone temperature, as shown in Fig. 21.

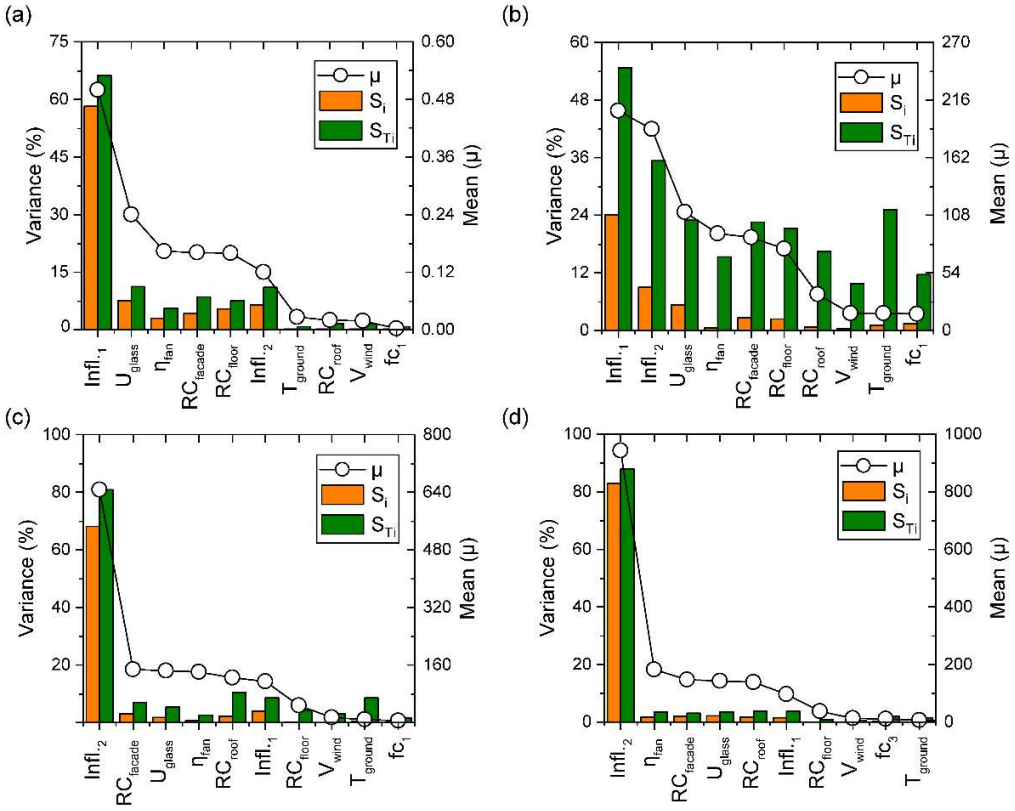


Fig. 21: The Morris analysis mean and the BACCO analysis method indices to specify the most influencing uncertain parameters with regards to the Emmen's dwelling where (a) the total thermal energy consumption in the Emmen's dwelling, (b) $t_{o/u}$ at zone 1, (c) $t_{o/u}$ at zone 2, (d) $t_{o/u}$ at zone 3

According to Table 6, there are total 31 uncertain parameters in the two-step GSA, the Morris analysis setting is selected, as mentioned in section 4.1.1. Then, for the ten most influential parameters, 400 sample points are used for the BACCO analysis. Latin hypercube sampling design is used for obtaining these sample points. The analysis is performed by formulating the sensitivity analysis problem and then using the Gaussian Emulator Machine Sensitivity Analysis Software (GEM-SA).

Fig. 21 shows the GSA results of all the model outputs. In each graph, μ value is used for presenting and ranking the ten most uncertainty parameters obtained from the Morris method. The BACCO analysis is quantified based on the main effect (S_i) and the total effect (S_{Ti}). The initial observation of the graphs revealed the fact that the ranking of parameters obtained from the Morris method is in accordance with the BACCO analysis results where the μ value and the BACCO main effect (S_i) follow the same decreasing trend. Therefore, the Morris method is applicable for performing the GSA analysis in case of the quantitative ranking. Although the essential parameters are identified with the help of the Morris method, the BACCO analysis tells about the contribution of these parameters towards the output's variance that is also given by adding the first-order indices and total-order indices.

For the thermal energy consumption in the dwelling ($Q_{heat\ total}$). The total effect indices are almost following the first-order effect indices in the BACCO analysis where the infiltration rate in zone 1 is the most significant parameter in the building performance, followed by the U_{glass}. In addition, the total effect indices showed a noticeable effect for the infiltration rate in zone 2 followed by the building envelope materials of the floor and façade (RC_{floor}, RC_{facade}). In terms of the $t_{o/u}$ at zone 1, the total effect indices show that the infiltration rate at zone 1&2 are the most influential parameters followed by the T_{ground} and building envelope characterization comprising the U_{glass}, RC_{floor} and RC_{facade}. On the other hand, the $t_{o/u}$ in the zone, 2&3 is only affected by the infiltration rate in zone 2. In general, the two-step GSA encompass a significant influence for the air tightness represented by the infiltration rate at each zone. This influence is a function of various parameters, including the building age, construction materials, weather conditions and building use. Since in the current study, the parameters related to the climate conditions and building construction are already included in the sensitivity analysis, which approve the building performance independency on these parameters. Therefore, the uncertain parameters regarding the building use presented by opening/closing doors& windows for natural ventilation may be responsible for the significant impact for the infiltration rate.

In general, the minimum ventilation rate should be around 41 L/s based on the Dutch standard (NEN 1087) [80]. However, it completely changes with the occupant behavior (opening/closing window schedule). The lack of extensive knowledge regarding such these parameters present a constraint in the performance of our framework when we preform uncertainty analysis. It can significantly change the contribution of each uncertain parameter to the model energy performance. Therefore, we keep them as deterministic parameters. Besides the most important parameters, the GSA results also showed that the uncertainty associate with the occupant behavior including the occupant schedule and zone temperature setting have less impact on the performance of the dwelling in terms of the $Q_{heat\ total}$ and $t_{o/u}$ objectives. In general, the two-step GSA is a valuable decision support technique that can offer the builders and construction companies an evaluation for the main drivers of the uncertainty associate the dwelling performance.

5. Conclusion

Improving the building energy efficiency is widely considered as a promising technique to meet the European Commission 2050 targets in reducing the CO_2 emission. In Netherland, Association of Dutch Municipalities is working on expanding this concept through the introduction of the nearly zero energy building (nZEB) approach [81]. However, the construction companies that lead the development of nZEB in the Netherlands indicate a substantial obstacle to guarantee the nZEB performance and thus to improve their market position. This work tends to propose a methodology framework to benefit the guarantee and assist proactively in maintaining the indoor climate performance of nZEBs. This methodology is implemented through developing a calibrated RC (resistance and capacitance) model for estimating the building energy performance. The first step in the calibration process is the dimension reduction strategy using Morris sensitivity analysis to estimate the influential parameters that contribute to the proposed model accuracy. Then the calibration process is implemented based on Bayesian optimization with consideration for the temporal variation associate with the climate data as well as the occupant behavior. Once the calibrated model is established, the uncertainty analysis (UA) incorporating global sensitivity analysis (GSA) is performed to quantify and assess the uncertainty associate with the performance of the developed digital dwelling. In this context, the developed methodology is applied for a newly renovated two-story home located in a district of Emmen at the Netherlands to close the gap between the initial retrofitting design and the monitoring operation throughout the heating seasons (9 months), a summary of the methodology framework key findings is the following:

- The Morris sensitivity analysis for the RC model accuracy drivers shows that the infiltration rates are main drivers for the prediction of the energy performance. While the indoor temperature zones mainly depend on the zone correction factor of the thermal mass fraction.
- The developed digital dwelling based on RC model illustrates a high accuracy in predicting the dwelling performance where the model offers a prediction accuracy of 2.2% for the total thermal energy consumption. Moreover, the developed model provides acceptable estimations for the indoor zone temperature with $C.V$ of 7.03%.
- The UA results show a large variation associate with digital dwelling output due to uncertain parameters including the climate data, the building envelope and occupant behavior. The uncertainty associates with indoor temperature can increase the over/underheating hours up to 3292 hrs. While the uncertainty associates with total thermal energy consumption can change the building performance from 4.03 MWh (67 kWh/m^2/yr) to 6.01 MWh (100 kWh/m^2/yr) .
- Finally, the GSA based on the Morris analysis and BACCO analysis reveals that the variation in building energy performance is due to the infiltration rates followed by the construction properties comprising the U_{glass}, RC_{floor} and RC_{facade}. Without losing the generality of the results, including more in parameters regarding the occupant behavior can contribute significantly to the GSA results.

In the real application, the developed methodology can serve as a diagnostic tool to improve and guarantee the performance of nZEB buildings during its operation stage which can be useful for the construction and installation companies to maintain the performance of their products proactively. Furthermore, it can promote a clear statement toward the EU approach to sustainable development.

Acknowledgements

The work is funded by the Spanish government RTI2018-093849-B-C33 (MCIU/AEI/FEDER, UE). This work is supported by the Ministerio de Ciencia, Innovación y Universidades – Agencia Estatal de Investigación (AEI) (RED2018-102431-T). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 713679 and from the Universitat Rovira i Virgili (URV). In addition, this work was part of the Dutch project TKI Optimaal and co-funded by TKI Urban Energy from the Surcharge for Top Consortia for Knowledge and Innovation (TKIs) of the Ministry of Economic Affairs of the Netherlands.

Appendix A. Supplementary material

Supplementary information SI.

References

[1] G. Brundtland, M. Khalid, S. Agnelli, S. Al-Athel, B. Chidzero, L. Fadika, V.Hauff, I. Lang, M. Shijun MM de B. The Brundtland Report: “Our Common Future.” vol. 4. 1988. doi:10.1080/07488008808408783.

[2] European Energy Agency. Final energy consumption by sector and fuel. Denmark: 2017. doi:CSI 027/ENER 016.

[3] European Environment Agency. Annual European Union greenhouse gasinventory 1990–2016 and inventory report 2018. Copenhagen: 2018.

[4] Dodd N, Donatello S, Garbarino E, Gama Caldas M. Identifying macro-objectives for the life cycle environmental performance and resource efficiency of EU buildings. 2015. doi:10.2791/975886.

[5] European Commission 6317. Towards an Integrated Strategic Energy Technology (SET) Plan: Accelerating the European Energy System Transformation. vol. 151. 2015. doi:10.1145/3132847.3132886.

[6] Boot PA. Energy efficiency obligations in the Netherlands: A role for white certificates? 2009.

[7] NEN (Nederlands Normalisatie-instituut). NEN7120+C2 Energieprestatie van gebouwen – Bepalingsmethode. 2017.

[8] Beuken R. mplementing the Energy Performance of Buildings Directive (EPBD) - Netherlands Country Report 2012. 2012.

[9] Lu Y, Wang S, Yan C, Huang Z. Robust optimal design of renewable energy system in nearly/net zero energy buildings under uncertainties. Appl Energy 2017;187:62–71. doi:10.1016/j.apenergy.2016.11.042.

[10] Guerra-Santin O, Tweed C, Jenkins H, Jiang S. Monitoring the performance of low energy dwellings: Two UK case studies. Energy Build 2013;64:32–40. doi:10.1016/j.enbuild.2013.04.002.

[11] Stazi F, Vegliò A, Di Perna C, Munafò P. Experimental comparison between 3 different traditional wall constructions and dynamic simulations to identify optimal thermal insulation strategies. Energy Build 2013;60:429–41. doi:10.1016/j.enbuild.2013.01.032.

[12] Terés-Zubiaga J, Campos-Celador A, González-Pino I, Escudero-Revilla C. Energy and economic assessment of the envelope retrofitting in residential buildings in Northern Spain. Energy Build 2015;86:194–202. doi:10.1016/j.enbuild.2014.10.018.

[13] Hillary J, Walsh E, Shah A, Zhou R, Walsh P. Guidelines for developing efficient thermal conduction and storage models within building energy simulations. Energy 2017;125:211–22. doi:10.1016/j.energy.2017.02.127.

[14] Monteiro H, Fernández JE, Freire F. Comparative life-cycle energy analysis of a new and an existing house: The significance of occupant’s habits, building systems and embodied energy. Sustain Cities Soc 2016;26:507–18. doi:10.1016/j.scs.2016.06.002.

[15] Spandagos C, Ng TL. Equivalent full-load hours for assessing climate change impact on building cooling and heating energy consumption in large Asian cities. Appl Energy 2017;189:352–68. doi:10.1016/j.apenergy.2016.12.039.

[16] Shen P, Lior N. Vulnerability to climate change impacts of present renewable energy systems

designed for achieving net-zero energy buildings. Energy 2016;114:1288–305. doi:10.1016/j.energy.2016.07.078.

[17] Bertagnolio S. Evidence-based Model Calibration for Efficient Building EnergyServices. University of Liège, 2012.

[18] Eisenhower B, O'Neill Z, Narayanan S, Fonoberov VA, Mezić I. A methodology for meta-model based optimization in building energy models. Energy Build 2012;47:292–301. doi:10.1016/j.enbuild.2011.12.001.

[19] Jiménez MJ, Madsen H, Andersen KK. Identification of the main thermal characteristics of building components using MATLAB. Build Environ 2008;43:170–80. doi:10.1016/j.buildenv.2006.10.030.

[20] Shen P, Braham W, Yi Y. Development of a lightweight building simulation tool using simplified zone thermal coupling for fast parametric study. Appl Energy 2018;223:188–214. doi:10.1016/j.apenergy.2018.04.039.

[21] Ji Y, Xu P, Duan P, Lu X. Estimating hourly cooling load in commercial buildings using a thermal network model and electricity submetering data 2016;169:309–23.

[22] Berthou T, Stabat P, Salvazet R, Marchio D. Development and validation of a gray box model to predict thermal behavior of occupied office buildings. Energy Build 2014;74:91–100. doi:10.1016/j.enbuild.2014.01.038.

[23] Terés-Zubiaga J, Escudero C, García-Gafaro C, Sala JM. Methodology for evaluating the energy renovation effects on the thermal performance of social housing buildings: Monitoring study and grey box model development. Energy Build 2015;102:390–405. doi:10.1016/j.enbuild.2015.06.010.

[24] Asadi E, Da Silva MG, Antunes CH, Dias L. Multi-objective optimization for building retrofit strategies: A model and an application. Energy Build 2012;44:81–7. doi:10.1016/j.enbuild.2011.10.016.

[25] Liao Z, Dexter AL. A simplified physical model for estimating the average air temperature in multi-zone heating systems. Build Environ 2004;39:1013–22. doi:10.1016/j.buildenv.2004.01.034.

[26] De Wilde P. The gap between predicted and measured energy performance of buildings: A framework for investigation. Autom Constr 2014;41:40–9. doi:10.1016/j.autcon.2014.02.009.

[27] Coakley D, Raftery P, Keane M. A review of methods to match building energy simulation models to measured data. Renew Sustain Energy Rev 2014;37:123–41. doi:10.1016/j.rser.2014.05.007.

[28] Asadi S, Mostavi E, Boussaa D, Indaganti M. Building energy model calibration using automated optimization-based algorithm. Energy Build 2019;198:106–14. doi:10.1016/j.enbuild.2019.06.001.

[29] Lam KP, Zhao J, Ydstie EB, Wirick J, Qi M, Park J. An energyplus whole building energy model calibration method for office buildings using occupant behavior data mining and empirical data. 2014 ASHRAE/IBPSA-USA Build. Simul. Conf., 2014, p. 160–7.

[30] Riddle M, Muehleisen RT. A guide to Bayesian calibration of building energy models. 2014 ASHRAE/IBPSA-USA Build Simul Conf 2014:276–83. doi:10.13140/2.1.1674.9127.

[31] Booth AT, Choudhary R, Spiegelhalter DJ. A hierarchical bayesian framework for calibrating micro-level models with macro-level data. J Build Perform Simul 2013;6:293–318. doi:10.1080/19401493.2012.723750.

[32] Wang S, Sun X, Lall U. A hierarchical Bayesian regression model for predicting summer residential electricity demand across the U.S.A. Energy 2017;140:601–11. doi:10.1016/j.energy.2017.08.076.

[33] Eisenhower B, O'Neill Z, A. Fonoberov V, Mezic I. Uncertainty and sensitivity decomposition of building energy models. J Build Perform Simul 2012;5:1–18. doi:10.1080/1940149YYxxxxxxxx.

[34] Yang T, Pan Y, Mao J, Wang Y, Huang Z. An automated optimization method for calibrating building energy simulation models with measured data: Orientation and a case study. Appl Energy 2016;179:1220–31. doi:10.1016/j.apenergy.2016.07.084.

[35] Sun J, Reddy TA. Calibration of building energy simulation programs using the analytic optimization approach (RP-1051). HVAC R Res 2006;12:177–96. doi:10.1080/10789669.2006.10391173.

[36] Monetti V, Davin E, Fabrizio E, André P, Filippi M. Calibration of building energy simulation models based on optimization: A case study. Energy Procedia 2015;78:2971–6. doi:10.1016/j.egypro.2015.11.693.

[37] Wetter M. GenOpt. Generic Optimization Program. User Manual. California: 2011. doi:10.2172/962948.

[38] Liu S, Henze GP. Calibration of building models for supervisory control of commercial buildings. IBPSA 2005 - Int Build Perform Simul Assoc 2005 2005:641–8.

[39] Stoppel CM, Leite F. Integrating probabilistic methods for describing occupant presence with building energy simulation models. Energy Build 2014;68:99–107. doi:10.1016/j.enbuild.2013.08.042.

[40] Neill ZO, Niu F. Uncertainty and sensitivity analysis of spatio-temporal occupant behaviors on residential building energy usage utilizing Karhunen- ve expansion Lo e. Build Environ 2017;115:157–72. doi:10.1016/j.buildenv.2017.01.025.

[41] Koene FGH, Bakker LG, Lanceta D, Narmsara S. SIMPLIFIED BUILDING MODEL OF DISTRICTS. Fifth Ger. IBPSA Conf., Aachen: RWTH Aachen University; 2014, p. 152–9.

[42] Heo Y. BAYESIAN CALIBRATION OF BUILDING ENERGY MODELS FOR ENERGY RETROFIT DECISION-MAKING UNDER UNCERTAINTY. Georgia Institute of Technology, 2011.

[43] Larsen TS, Heiselberg P. Single-sided natural ventilation driven by wind pressure and temperature difference 2008;40:1031–40. doi:10.1016/j.enbuild.2006.07.012.

[44] Heo Y, Choudhary R, Augenbroe GA. Calibration of building energy models for retrofit analysis under uncertainty 2012;47:550–60. doi:10.1016/j.enbuild.2011.12.029.

[45] Mustafaraj G, Marini D, Costa A, Keane M. Model calibration for building energy efficiency simulation 2014;130:72–85. doi:10.1016/j.apenergy.2014.05.019.

[46] Raftery P, Keane M, Costa A. Calibrating whole building energy models : Detailed case study using hourly measured data. Energy Build 2011;43:3666–79. doi:10.1016/j.enbuild.2011.09.039.

[47] Kaplan MB, McFerran J, Jansen J PR. Reconciliation of a DOE2.1c model with monitored end-use data for a small office building. ASHRAE Trans, 1990, p. 981–93.

[48] Reddy TA MI. Procedures for reconciling computer-calculated results with measured energy data. Research Project 1051-RP. 2006.

[49] Østergård T, Jensen RL, Maagaard SE. Building simulations supporting decision making in early design – A review. Renew Sustain Energy Rev 2016;61:187–201. doi:10.1016/j.rser.2016.03.045.

[50] Chowdhury S, Champagne P, Mclellan PJ. Uncertainty characterization approaches for risk assessment of DBPs in drinking water : A review. J Environ Manage 2009;90:1680–91. doi:10.1016/j.jenvman.2008.12.014.

[51] Eames M, Ramallo-gonzalez AP, Wood M. An update of the UK ' s test reference year : The implications of a revised climate on building design. Build Serv Eng 2015;37:316–33. doi:10.1177/0143624415605626.

[52] Hong T, Chang W, Lin H. A fresh look at weather impact on peak electricity demand and energy use of buildings using 30-year actual weather data. Appl Energy 2013;111:333–50. doi:10.1016/j.apenergy.2013.05.019.

[53] Heo Y, Wilde P De, Li Z, Yan D, Park CS. A review of uncertainty analysis in building energy assessment. Renew Sustain Energy Rev 2018;93:285–301. doi:10.1016/j.rser.2018.05.029.

[54] Kaplanis SÃ, Kaplani E. A model to predict expected mean and stochastic hourly global solar radiation I (h ; n j) values 2007;32:1414–25. doi:10.1016/j.renene.2006.06.014.

[55] Vela S. A review of wind speed probability distributions used in wind energy analysis Case studies in the Canary Islands 2009;13:933–55. doi:10.1016/j.rser.2008.05.005.

[56] Cai YP, Huang GH, Yang ZF, Lin QG, Tan Q. Community-scale renewable energy systems planning under uncertainty — An interval chance-constrained programming approach 2009;13:721–35. doi:10.1016/j.rser.2008.01.008.

[57] Huang P, Huang G, Wang Y. HVAC system design under peak load prediction uncertainty using multiple-criterion decision making technique. Energy Build 2015;91:26–36. doi:10.1016/j.enbuild.2015.01.026.

[58] Silva AS, Ghisi E. Uncertainty analysis of user behaviour and physical parameters in residential building performance simulation. Energy Build 2014;76:381–91. doi:10.1016/j.enbuild.2014.03.001.

[59] CORRADO V, MECHRI HE. Uncertainty and Sensitivity Analysis for Building Energy Rating. J Build Phys 2009;00. doi:10.1177/1744259109104884.

[60] Hoes P, Hensen JLM, Loomans MGLC, Vries B De, Bourgeois D. User behavior in whole building simulation 2009;41:295–302. doi:10.1016/j.enbuild.2008.09.008.

[61] Zhang S, Sun Y, Cheng Y, Huang P, Olaide M, Lin Z. Response-surface-model-based system sizing for Nearly / Net zero energy buildings under uncertainty. Appl Energy 2018;228:1020–31. doi:10.1016/j.apenergy.2018.06.156.

[62] Tian W. Identification of key factors for uncertainty in the prediction of the thermal performance of an office building under climate change 2009:157–74. doi:10.1007/s12273-009-9116-1.

[63] Smith A, Luck R, Mago PJ. Analysis of a combined cooling , heating , and power system model under different operating strategies with input and model data uncertainty. Energy Build 2010;42:2231–40. doi:10.1016/j.enbuild.2010.07.019.

[64] Hong T, Taylor-lange SC, Oca SD, Yan D, Corgnati SP. Advances in research and applications of energy-related occupant behavior in buildings &. Energy Build 2016;116:694–702. doi:10.1016/j.enbuild.2015.11.052.

[65] Re F, Manfren M, Chiara L, Luigi A, Ciribini C, Angelis E De. Probabilistic behavioral modeling in building performance simulation : A Monte Carlo approach. Energy Build 2017;148:128–41. doi:10.1016/j.enbuild.2017.05.013.

[66] Bahaj ASÃ, James PAB. Urban energy generation : The added value of photovoltaics in social housing 2007;11:2121–36. doi:10.1016/j.rser.2006.03.007.

[67] Sun Y, Huang P, Huang G. A multi-criteria system design optimization for net zero energy buildings under uncertainties. Energy Build 2015;97:196–204. doi:10.1016/j.enbuild.2015.04.008.

[68] Tian W. A review of sensitivity analysis methods in building energy analysis. Renew Sustain Energy Rev 2013;20:411–9. doi:10.1016/j.rser.2012.12.014.

[69] Saltelli A, Tarantola S, Campolongo F RM. Sensitivity analysis in practice: a guide to assessing scientific models. John Wiley & Sons, Inc; 2004.

[70] Sohier H, Piet-lahanier H, Farges J. Acta Astronautica Analysis and optimization of an air-launch-to-orbit separation. Acta Astronaut 2015;108:18–29. doi:10.1016/j.actaastro.2014.11.043.

[71] Kucherenko S, Albrecht D, Saltelli A. Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques. 8th IMACS Semin. Monte Carlo methods, 2015, p. 1–32. doi:10.1016/j.ress.2017. 04.003.

[72] Chong A, Menberg K. Guidelines for the Bayesian calibration of building energy models. Energy Build 2018;174:527–47.

[73] Kennedy MC, O'Hagan A, Higgins N. Bayesian Analysis of Computer Code Outputs. Quant Methods Curr Environ Issues 2011:227–43. doi:10.1007/978-1-4471-0657-9_11.

[74] Amasyali K, El-gohary NM. A review of data-driven building energy consumption prediction studies 2018;81:1192–205. doi:10.1016/j.rser.2017.04.095.

[75] ISO. Moderate thermal environments- Determination of the PMV and PPD indices and specification of the conditions for thermal comfort: ISO7730. 1994.

[76] Abokersh MH, Vallès M, Cabeza LF, Boer D. A framework for the optimal integration of solar assisted district heating in different urban sized communities : A robust machine learning approach incorporating global sensitivity analysis. Appl Energy 2020;267:114903. doi:10.1016/j.apenergy.2020.114903.

[77] Uusitalo L, Lehikoinen A, Helle I, Myrberg K. An overview of methods to evaluate uncertainty of deterministic models in decision support. Environ Model Softw 2015;63:24–31. doi:10.1016/j.envsoft.2014.09.017.

[78] Petropoulos G, Wooster MJ, Carlson TN, Kennedy MC, Scholze M. A global Bayesian sensitivity analysis of the 1d SimSphere soil-vegetation-atmospheric transfer (SVAT) model using Gaussian model emulation. Ecol Modell 2009;220:2427–40. doi:10.1016/j.ecolmodel.2009.06.006.

[79] Koo C, Hong T, Lee M, Seon Park H. Development of a new energy efficiency rating system for existing residential buildings. Energy Policy 2014;68:218–31. doi:10.1016/j.enpol.2013.12.068.

[80] NEN. NEN 1087:2001- Ventilatie van gebouwen: Bepalingsmethoden voor nieuwbouw. 2001.

[81] Zeiler W, Gvozdenović K, de Bont K, Maassen W. Toward cost-effective nearly zero energy buildings: The Dutch Situation. Sci Technol Built Environ 2016;22:911–27.

doi:10.1080/23744731.2016.1187552.