# A Bayesian Emulator Methodology to Support Evidence-Based Building Energy Model Parameterisation and Uncertainty Analysis

# 

Mohammad Royapoora, Dario Domingob, Hailiang Dub, Aaron Boranianc, Sara Walkera, Michael Goldsteinb

[a] School of Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU

[b] Department of Mathematical Sciences, Durham University, Durham, DH1 3LE

[c] Big Ladder Software LLC, Denver, CO 80202

**How many CV(RMSE) and MBE are in each energy subcat that qualify as calibrated model. Could it be that we chose a smaller parameter space that the emulator did not have enough information to be able to completely define the uncertainty space. Run and emulate over a bigger parameter space:**

**However, as the number of spaces increase the**

**Cut out the parameters that are not efficient in enabling expensive models to history match successfully,**

# **Abstract**

This work uses a systematic, evidence-based approach to parameterise a building energy model to be calibrated against monthly gas and electricity data and hourly temperatures. The emphasis is to avoid tuning the model is an ad-hoc manner and retain input parameters within the bounds of actual energy and temperature data, building thermophysical information and observed occupant activities. Expert judgement and product specifications are used to introduce uncertainty bands to train a Bayesian emulator in two annual waves of simulation each involving 1000 runs.

The results show:

Having put the deterministic occupancy and operation of the building as closely as possible, and having used local weather data, the probable variation in fabric properties and plant efficiencies could only bring 7 out of 12 months into a calibration band of +\_15% (CV(RMSE)), with the rest of uncertainties having to be left to micro-climate, weather, stochastic occupant behaviour, ???

# **Introduction**

Stating uncertainty bands and integrating them in building performance simulation results is essential for producing high-quality results that ~~also~~ adequately acknowledge~~s~~ modelling limitations. Both model uncertainty and performance gap between model predictions and actual operational data remain areas of active building research. A recent article reported that a large proportion of building modelling community lacks essential knowledge on what the most fundamental parameter inputs for buildings are and how these impact ~~the~~ model predictions [1]. This gap has major consequences as retrofit strategies are mostly derived in consultation with building energy models and as such techno-economic benefits of proposed solutions can be misleading. Similarly, future buildings are expected to be more responsive to other civic activities (i.e. power generation and storage, transport, etc.)[2]. This can only be assessed by modelling, simulation and near-real-time analytics of a cluster of buildings at district and potentially city level which in turn requires generating accurate building energy data.

Building model parameter space and predictions suffer of both inaccuracy and uncertainty due to:

1. The dynamic nature of building fabric thermo-physical properties (i.e., The U-value of a masonry wall changes as a function of its moisture content, while most models assign it a fixed value). Hygroscopic material can exhibit very wide thermophysical variance at different moisture and temperature conditions.
2. The stochastic nature of occupant behaviour and its interaction with the building (window opening, light and small power usage).
3. Uncertainties and variations in plant operational characteristics that are extensively simplified in energy models.
4. Exact zone air exchange figures and fabric infiltration values which are difficult to determine, with the later only available after performing a building pressure test that is logistically difficult and costly particularly for larger occupied buildings.
5. The weather files that impose large uncertainties in particular with solar irradiance data that quite often is partially or fully modelled (as opposed to measured). Micro-climatic variations are widely ignored and understudied. An example could be the difference in conventional airport weather station data used to model urban settings. Airports are exposed terrains often with proximity to water to facilitate emergency aircraft landing. Annual weather files compiled in these locations would therefore report higher wind velocities and miss the heat-island effect that is increasingly separating urban and open country micro-climates.

Most commonly in the energy literature [refs], ASHRAE guidelines [28] are followed to assess whether a model has been successfully calibrated against observed data. The process consists in computing the following discrepancy measures between model outputs and observations, and consider a model calibrated if their value is within pre-described bounds:

,

where is the sequence of measured energy consumption data and the sequence of simulated energy consumption data, *i=1, … , N*. ASHRAE guidelines suggest to consider a model calibrated if the associated MBE and CV(RMSE) satisfy:

1. -10% ≤ MBE ≤ 10% and CV(RMSE) ≤ 30% for hourly data
2. -5% ≤ MBE ≤ 5% and CV(RMSE) ≤ 15% for monthly data.

While the above criteria are easy to check, their use to assess model calibration presents inherent limitations.

1. Ideally, assessment of model calibration should depend on:

* the magnitude of measurement errors affecting the observed data; and
* the level of discrepancy between the model and the real-world process it simulates, caused by inherent modelling limitations and approximations.

The acceptance levels of the above criteria are instead independent of both the above sources of uncertainty.

1. While the computation of the MBE and CV(RMSE) is legitimate for quantities such as energy consumption, it is less (or not at all) appropriate for physical quantities such as temperature, for which model outputs and observations may be available.
2. The computation of MBE and CV(RMSE) can be carried out only after the model has been run for particular choices of inputs and its outputs have been observed. Computational resources and time hence limit the number of instances for which they can be computed. If model outputs are unavailable but predicted with associated uncertainty, there is no immediate way to include the latter into their computation.

This work introduces a Bayesian method of uncertainty quantification that is deployed in conjunction with an evidence-based building model development to characterise modelling uncertainties. Actual electricity, gas and space temperatures were compiled in a domestic property that together with logged occupant activities, detailed fabric properties, local weather and plant specifications and schedules of operation supported the parameterisation of an EnergyPlus model.

This work attempts to address three objectives; first to bring the principle of evidence-based model parameterisation as outlined in [3] to support the proposed Bayesian method while adjustments to the input parameters are only made according to available evidence and within the bounds of uncertainty as derived from product specification and literature review. Actual monthly energy data (natural gas and electricity) are used to enable initial history matching and subsequent forward uncertainty analysis, while hourly indoors temperatures from 2 locations are used to undertake inverse uncertainty analysis. The second objective is to highlight the relationship between zone energy consumption and the subsequent air temperature. This interactive relationship has rarely been examined in model calibration exercises and yet energy density parameters cannot be adjusted without assessing its reflections on zone air temperature (and vice versa). Finally, the extent to which such a detailed energy model can be successfully replicated by a Bayesian emulator so that the emulator could present the behaviour of a demand side energy system for wider system level of energy modelling.

# **Literature review**

**2.1 Building thermal property and HVAC**

A wide range of reasons exist that lead to building energy analysts (?) using assumptions in place of hard-to-measure building and occupant parameter inputs. These have a major impact on the prediction accuracy of the model. An analysis of two office buildings in Australia found cooling set-points, ICT and its schedule and lighting power density to impact the predictions of energy models most, leading to lighting retrofit paybacks that for the same building can range from 2.4 to 10.3 years[4]. Another area of large uncertainty is the fabric thermophysical properties that could arise from fabric composition (i.e., non-homogeneity, moisture content effecting thermal conductivity, etc.) or its surface properties (radiative or convective characteristics). Most notably the hypothesis of mono-dimensionality of heat flow which is fundamental to thermal resistance (U-value) calculations in ISO 6946:2007, ISO 9869-1:2014, CIBSE and ASHREA [5-8] is not entirely valid. Heat in solid bodies travels in a diffused and 3-dimensional manner which thus far building energy models are incapable of replicating. Existing techniques for in situ measurement of fabric U-Value show notable variations across seemingly uniform walls and floor surfaces [9-12]. Also the main standardised semi-stationary or dynamic U-Value measurement methods produce results that mostly reveal reasonable differences [10], and in some cases departures of up to 393% are reported between measured and calculated figures [9]. Traditional solid masonry elements [13, 14] as well as floors [15] have been found to perform better than calculation method suggest, whereas composite walls are reported to perform worse than model calculations. A. Marshall et al. found CIBSE model U-Values for brick walls, ceilings and doors to over-predict performance by 30.3%, 15.5% and 9.9% respectively while calculated window properties showed much better matches with measured figures [16]. A study of 57 properties found that while variations between similar wall types and even within the same dwelling’s walls existed, 44% of walls performed better than CIBSE predications, 42% were within acceptable bounds and only 14% of sampled walls performed worse than calculations suggest [17], however measured and calculated floor U-values were found to be in good agreement. Difficulties in identifying the composite material and their density and moisture content, internal and external air velocities, fabric non-homogeneity and cumbersome nature of in-situ fabric studies are notable reasons behind fabric U-value errors and uncertainty. Despite this, research suggests that with extensive field data and careful input parameter selection, building energy models can predict energy and environmental performance with high levels of accuracy [18].

**2.2 Weather, occupant and internal gains**

The exact micro-climatic condition never impacts a building more than once in its lifetime. Wind speed and direction, solar irradiance and its realisation on external fabric and within the building, external temperature and humidity, cloud cover and ground albedo, atmospheric pressure all interact in close to chaotic manner to dictate immediate micro-climatic condition that is represented in a simplified annual weather file for most energy simulations. While such weather files are normally extracted from decades of actual data (e.g. using Finkelstein-Schafer method), these remain a major source of energy simulation uncertainty for the following reasons: (i) building energy performance is affected by future and not past weather, (ii) a single weather file can hardly represent all meteorological conditions (iii) most weather files are not measured at the site but a nearby meteorological station (often exposed airports), leading to inadequate representation of urban heat island and sheltering effects [19].

Occupant behaviour is another simulation input that is drastically simplified to inform the current generation of simulation tools. This includes temporal schedules that describe the variations in occupant-related activities in a homogenous and deterministic form. More accurate representation of occupant and its dynamic interaction with the building remains an active area of research and advances in pervasive sensing and data collection has enabled better understanding of occupant presence and movements, its autonomy as an active agent, operation of windows and blinds and interventions with lighting and HVAC settings. Insights from improved monitoring has informed two categories of occupant representation, namely implicit and explicit models that seek to predict stochastic occupant activity in the form of control of windows, ICT and HVAC. Improvements are gained by seeking to understand the underlying logic behind occupant interaction with its surrounding at a behavioural level (explicit model) [20]. Challenges remain in adequately detailed monitoring of occupant behaviour and the integration of subsequent occupant behaviour into simulation models and finally validation of this approach against longitudinal field-studies. W. Tian et al. note the need for further research to provide clear guidance not only to characterise stochasticity but also on the number of sampling numbers required to arrive at acceptable statistical power in building energy analysis [21].

**2.3 Treatment and distribution of uncertainty**

Any rigorous treatment of an uncertainty problem starts with identifying all possible sources of uncertainty and their respective distribution type. Quantifying these uncertainties is only possible through the application of probability theory, for which an initial step is the probabilistic distribution of uncertainties. One of the most comprehensive attempts to quantify and characterise sources of modelling uncertainty was undertaken by I.A. Macdonald [22]. W. Tian et al. borrows from multiple previous works to report that Gaussian distribution is the most common technique for describing parameter uncertainty in energy modelling, while uniform distribution best describes the spread of building forms and design strategies [21]. Table 1 outlines various uncertainty distribution assumed for a range of building energy model parameters. Hypercube sampling method that supports the Bayesian method introduced in this work assumes a normal distribution for sources of uncertainty and underpins the consideration in this work.

**2.4 Co-dependency of energy and temperature predictions**

EnergyPlus (E+) is a collection of dynamic modules each simulating different environmental, climatic and operational conditions that define either the flow or the stored quantity of energy within building internal zones. The core of the programme is principally a heat balance equation that is solved for all zones using one of three methods (3rd order backward difference, Euler method or analytically) to converge zone loads and resultant temperatures to within a pre-defined tolerance range, using a predictor/corrector process. The uniqueness of E+ lies in being a physically based modelling solution that oversees a simultaneous calculation of radiative and convective heat and mass transfer processes, adsorption and desorption of moisture in building elements, iterative plant, building fabric and air interactions and accurate temperature and comfort predictions. This integrated and simultaneous simulation process is completed via several modules (and overseen by EnergyPlus simulation manager), with understandably multiple first-principle-based equations that are solved simultaneously and/or iteratively. This makes it very difficult to bring a sharp focus on any single or sets of expressions where model prediction uncertainties lie. However interestingly the zone air heat balance equation is one of the primary mechanisms that describes the connected nature of the loads within a zone, the corresponding plant duty to offset these loads as well as the zone mean air temperatures as follows:

[1]

Where is the rate of change of thermal energy stored in the zone air, is the sum of convective internal loads, is convective heat transfer from the zone surfaces, is the change of the room air enthalpy as a result of zone air mixing, is infiltration heat transfer and finally is the HVAC system input to achieve its target temperature. Given that E+ assumes a uniform zone air and surface temperature, uniform long and short-wave radiation and diffuse radiation and reflective surfaces (as opposed to direct or point-based), it is reasonable to regard zone air temperature as the interconnection where connective, radiative and convective heat balance and mass transfer are realised. Essentially each item on the right-hand side of Eq. 1 indicates a change of enthalpy due to environmental perturbations, while the left-hand side describes how these perturbations impact the zone air temperature and the enthalpy that it holds. In compiling calibration data and subsequent modelling runs this work attempts to explore the degree to which both temperature and energy data can assist reducing the space of modelling uncertainty.

# **Method**

**3.1 Case-study building**

A detached two-storey masonry construction built in 1994 was selected as the case-study building. Two occupants are the only residents of the dwelling and were asked to archive their gas cooker and shower usage each day across an annual cycle. Given a very predictable pattern of occupancy (both occupants had 8am-5pm working commitments), it was possible to limit the stochastic nature of occupant activity as far as practically manageable and use deterministic schedules to represent occupant activity. The building (with a gross area of 168.66m2 and 19.73m2 of unheated space) is in a built-up urban surrounding, is only partly shaded (on its west elevation) by another adjacent property (which was considered within the modelling work). Across the monitoring year (2016) the property had an observed annual gas (15,381 kWh) and electricity (2,991 kWh) consumptions that respectively reflect high and medium UK typical domestic consumption values [23]. Occupants only utilised shower facilities which at a measured flow rate of 4.37 l/min and recorded average eight 20-minute showers per week correspond to an average of 50 l/person/day. These recoded values are below UK average domestic hot water usage (reported as 142 l/person/day [24] and 122 l/person/day [25]), but primarily reflect their heavy use of gym washing facilities. Gas cookers (containing 3kW and 5kW hubs) were used 4 times a week for 1 hr per cooking session.

**3.2 Energy and temperature data collection**

A proprietary set of environmental and energy sensors were deployed to compile electricity and zone temperatures (Fig. 1). To reduce measurement uncertainty, each ~~one~~ of the two target zones was equipped with two separate air temperature sensors at 1.3m above floor level and set to log data at 30s intervals to achieve a moderated average. Therefore, space temperature was recorded by 4 sensors (two positioned in the south facing master bedroom and two in north facing kitchen). Whereas electricity and gas data required no imputation, overall kitchen and master bedroom temperature sensors had total annual losses of 5.7% and 2.7% that required imputation. Each missing hourly temperature cell was imputed by the average of the previous and successive available cells. Gas consumption data was manually recorded at monthly basis using mains gas meter. Electricity was logged at 10s intervals using two clip-on current sensors on the incoming live cable (to reduce measurement uncertainty) and the two set of similar readings were averaged and aggregated to form the measured power usage.



Figure 1 LSH: power monitor used to characterise household appliances, RHS: AC current sensor, monitoring transmitters and temperature sensors deployed in case-study building.

In order to parameterise the energy model more accurately, a plugin power monitor was also used to characterise instantaneous and time-average consumption of the main electrical devices (TV, washing machine, ICT) in the property.

**3.3 Input parameter selection**

By consulting manufacturers specification and the house builder’s literature, a detailed set of parameter inputs were compiled and where the greatest quantifiable uncertainty existed, lower and upper bands are imposed on the input value used. These bands were so far as possible derived from scientific literature and used to dictate the size of associated variations explored in batch-runs (Table 2). Ground floors are less prone to variations in internal and external air velocities that act on walls and roofs more robustly, and lead to dynamic heat transfer values that fail to be captured by standardised calculation methods. Therefore, a smaller floor uncertainty margin was derived from literature and imposed on floor thermal resistance (Tables 2 and 3).

The property’s glazing was updated in 2009 and manufacture’s literature set the G and U-value of the fenestration to 0.691 and 1.788 W/m2K respectively, with respective error bands of ±5% and ±2%. The compound upper and lower limit of these two values altered the gas consumption of the calibrated model by ±2.05 kWh (± 0.013%). Given its negligible nature, the error bands of the glazing were discounted in batch simulations. The compound effect of all other uncertainty bands created a lower boundary of 8,842 kWh and an upper boundary of 26,452 kWh with respect to an observed gas consumption of 15,381 kWh (i.e. -42.5% to +72%). Given that even a 5% increases in fabric U-value was reported to raise energy consumption of family homes by 0.3-2.5% [26], a uniformly distributed uncertainty band is imposed on elemental U-Values to reflect similar magnitude of variations reported in literature, as outlined in Table 2.

Actual building infiltration rates are difficult to arrive at and require convoluted air permeability tests. Table 4.16 of CIBSE guide A [6] outlines a range of 0.25 to 0.95 air change per hour (ACH) for various 2-storey buildings below 500m2 with a value of 0.5 ACH describing typical constructions similar to the case-study building. Therefore 0.5 ACH informed the calibrated base model with 0.25-0.95 ACH representing the range of possibilities that batch simulations in the Bayesian emulator explored (Tables 2-3). Local weather files compiled by a weather station approximately 3 miles away from the site was used to support the model development [27].

**3.4 Calibration**

A succession of 38 version-control models each with incremental adjustments paved the way to arrive at the final calibrated version used for Bayesian emulator training and batch simulations. Against actual hourly data, ASHRAE Guideline 14-2002 was followed to calibrate the building model [28]. This entails determining two dimensionless indicators of errors, MBE and CV(RMSE) values using formulae 1 and 2:

[1]

[2]

Where *Mi* and *Si* are respective measured and simulated data at instance *i*, and *Ni* is the count of the number of values used in the calculation. ASHRAE Guide 14 considers a building model calibrated if hourly MBE values fall within ±10% and hourly CV(RMSE) values fall below 30%. MBE and CV(RMSE) indices were constructed over monthly intervals to study monthly variations too.

It is worth noting that EnergyPlus calculates infiltration for each zone dynamically using an Air Flow Network component. This returns dynamic values at each time step as a function of two inputs: [i] a crack factor (a user-defined input) and [ii] air pressure differentials across all boundary surfaces with air to both sides [29]. HVAC-induced infiltration is also accounted for where mechanical air systems are modelled. In order to vary the input value for infiltration across batch simulation, dynamic infiltration airflow rates were used to create an infiltration intensity schedule (ranging from 0 to 1) for every zone. This schedule was then used with an overall value ranging from 0.25 to 0.95ACH to allow the infiltration to be dictated at each run.

The schedule (i.e., timetable) that governs the timing of heating in the building impacts the timing of the gas use greatly as it dictates when heat is requested in the occupied zone. Heating schedule presents a multitude of choices that are a product of any variation of 12 months by 7 days across 48 half-hourly timesteps which lead to a large number of possibilities. Since the injection of heating into the space should coincide with a rise in space temperature, the actual recorded data in two target areas were used to arrive at the closest match between the recorded and simulated space temperatures.

In UK, cooking as a proportion of total household energy demand has halved since 1970s, however the latest UK household energy survey reports that it has remained notably constant and equates to an average of 2.67% [-0.37% to +0.23%] of overall household primary energy demand [30]. This

**3.5 Bayesian emulator (Dario/Hailiang)**

**….**

# **Results and discussion**

* Calibrated model results

MBE figures provide an indication of errors averaged to the mean of measured values but suffer from the cancellation effect. CV(RMSE) index however is a measure of accumulated error normalised to the mean of the measured values. As such CV(RMSE) more closely reflects the accumulated magnitude of error and therefore is a better measure of the overall prediction accuracy of the model.

If a larger number of trials existed where the results of measured vs. calculated fabric U-Values were reported in relative terms, it would be possible for energy modelling communities to impose fabric error bands with a greater degree of uncertainty.

Despite a high-fidelity model that was populated by actual data that are of high integrity, the prediction of gas consumption remains less accurate than electricity. The roll out of smart gas meters in the UK are therefore a critical development if a more refined perspective of energy flow in the built environment is to be accomplished.

* Disparity reasons:

Authors of this paper reported on a previous model calibration where EnergyPlus temperature prediction was reported for a larger structure with openable windows to the north aspect of the building only. A smaller domestic property has a much smaller thermal mass within its insulated fabric and therefore large openable windows (as well as doors) provide the opportunity for very rapid purges of internal air that leads to space air temperature dropping dramatically. While within a large building with considerable mass, the core temperature of the mass acts as a moderating anchor and window opening events do not create major temperature swings, a smaller property remains more susceptible to notable swings following window opening events.

Overall, the actual room air temperature results show that in the absence of heating system input (i.e., in the freefloat mode), the building cools down more slowly than the model suggestion. This could be due to the absence of exact thermal mass contributed by furniture and furbishing that stores heat and slows down the internal convection currents that are a major cooling mechanism within the space. Secondly mature trees and shrubs around the house provide a greater degree of sheltering that cannot be represented within the model.

* Insufficient data is available to support assessment and variation of measured versus calculated fabric U-Values and in particular with external walls that are more instrumental in energy performance of a property, a wider range of field trials could inform better parameterisation of energy models.
* Floor and roof thermal values impose smaller uncertainty bands due to the thermally coupled nature of non-suspended floors with the ground, and in roofs due to the buffeting effect of the loft space. Better estimation of thermal values could only be produced by an in-situ measurement.

# **Conclusions**

Often random and continuous events in buildings are reduced to deterministic and discrete estimations in energy models. This work attempted to minimise any assumption in order to parameterise

What is the electricity and gas profile of 21 Prince’s meadow if all the schedules and activity descriptions were set to default or industry normal practices?

When using measured data to inform a Bayesian uncertainty quantification, the posterior distribution derived using Bayesian technique

Available observed data for building thermal and DHW load were aggregated at whole building level and at monthly intervals, leading to insufficient insight to infer subcategories of loads with greater confidence. Uncertainty remains an ongoing challenge in building simulations as both the magnitude as well as the shape of the distribution of uncertainty distribution for most import parameters are not yet uniformly defined.

Parameter input was assumed with a Gaussian distribution and posterior distribution of energy and environmental data was examined against actual measurements at monthly and hourly intervals.

One of the findings of this work could be that although the model produces initial results that are very close to the measured values, when looked at more closely, they differ widely from available benchmarks and area-weighted field measurements. This highlights the need to have a set of guidelines describing in more realistic terms the need for accurate parameter inputs of models.

Also quit often analysts do not have high resolution data of buildings, if we can demonstrate how much uncertainty is involved in the model if it is to be calibrated based on the monthly available data then it is of great use to the simulation community.

The purpose of this work is to use actual monthly energy and environmental data and reported uncertainty ranges of building fabric and plant efficiencies, as well as typical consumption data to train a Bayesian emulator to reduce the space of parameter input uncertainty.

We argue that the inclusion of zone environmental conditions (i.e., air temperature) is crucial to the quantification of building energy model uncertainty.

One of the conclusions can be that these models were not designed to consider the interaction and behaviour of the occupant with the building. Although the authors were completely aware of the details of how the building was used over the past year, the lack of this knowledge could have been predicted in this instance by an experienced analyst.

As buildings become more and more electrically powered, the stochastic nature of human comfort makes the predictions of models weaker.

References

1. Imam, S., D.A. Coley, and I. Walker, *The building performance gap: Are modellers literate?* Building Services Engineering Research and Technology, 2017. **38**(3): p. 351-375.

2. Royapoor, M., A. Antony, and T. Roskilly, *A review of building climate and plant controls, and a survey of industry perspectives.* Energy and Buildings, 2017.

3. Raftery, P., M. Keane, and J. O’Donnell, *Calibrating whole building energy models: An evidence-based methodology.* Energy and Buildings, 2011. **43**(9): p. 2356-2364.

4. Daly, D., P. Cooper, and Z. Ma, *Understanding the risks and uncertainties introduced by common assumptions in energy simulations for Australian commercial buildings.* Energy and Buildings, 2014. **75**(Supplement C): p. 382-393.

5. International Organization for Standardization, *ISO 6946:2007: Building components and building elements -- Thermal resistance and thermal transmittance -- Calculation method.* 2007.

6. CIBSE, *Guide A; Environmental Design*. 2015, The Chartered Institution of Building Services Engineers: London.

7. American Society of Heating, R.a.A.-C.E., *2017 ASHRAE Handbook—Fundamentals*. 2017, Atlanta, GA.

8. International Organization for Standardization, *ISO 9869-1:2014: Thermal insulation Building elements (In-situ measurement of thermal resistance and thermal transmittance).* 2014.

9. Rasooli, A., L. Itard, and C.I. Ferreira, *A response factor-based method for the rapid in-situ determination of wall’s thermal resistance in existing buildings.* Energy and Buildings, 2016. **119**(Supplement C): p. 51-61.

10. Deconinck, A.-H. and S. Roels, *Comparison of characterisation methods determining the thermal resistance of building components from onsite measurements.* Energy and Buildings, 2016. **130**(Supplement C): p. 309-320.

11. Meng, X., et al., *Factors affecting the in situ measurement accuracy of the wall heat transfer coefficient using the heat flow meter method.* Energy and Buildings, 2015. **86**(Supplement C): p. 754-765.

12. Ficco, G., et al., *U-value in situ measurement for energy diagnosis of existing buildings.* Energy and Buildings, 2015. **104**(Supplement C): p. 108-121.

13. Gaspar, K., M. Casals, and M. Gangolells, *A comparison of standardized calculation methods for in situ measurements of façades U-value.* Energy and Buildings, 2016. **130**(Supplement C): p. 592-599.

14. Desogus, G., S. Mura, and R. Ricciu, *Comparing different approaches to in situ measurement of building components thermal resistance.* Energy and Buildings, 2011. **43**(10): p. 2613-2620.

15. Hoffmann, C. and A. Geissler, *The prebound-effect in detail: real indoor temperatures in basements and measured versus calculated U-values.* Energy Procedia, 2017. **122**(Supplement C): p. 32-37.

16. Marshall, A., et al., *Domestic building fabric performance: Closing the gap between the in situ measured and modelled performance.* Energy and Buildings, 2017. **150**(Supplement C): p. 307-317.

17. P. Baker, *U-values and Traditional Buildings, Historic Scotland ConservationGroup,* . 2011: Glasgow.

18. Royapoor, M. and T. Roskilly, *Building model calibration using energy and environmental data.* Energy and Buildings, 2015. **94**: p. 109-120.

19. Yassaghi, H., N. Mostafavi, and S. Hoque, *Evaluation of current and future hourly weather data intended for building designs: A Philadelphia case study.* Energy and Buildings, 2019. **199**: p. 491-511.

20. Hong, T., et al., *Advances in research and applications of energy-related occupant behavior in buildings.* Energy and Buildings, 2016. **116**: p. 694-702.

21. Tian, W., et al., *A review of uncertainty analysis in building energy assessment.* Renewable and Sustainable Energy Reviews, 2018. **93**: p. 285-301.

22. Macdonald, I.A., *Quantifying the effects of uncertainty in building simulation*. 2002, University of Strathclyde Glasgow.

23. The Office of Gas and Electricity Markets. *Typical Domestic Consumption Values*. 2017 [cited 2017 17 November]; Available from: <https://www.ofgem.gov.uk/gas/retail-market/monitoring-data-and-statistics/typical-domestic-consumption-values>.

24. Energy Saving Trust, *At Home with Water*. July 2013, Energy Saving Trust: London.

25. Energy Saving Trust, *Measurement of Domestic Hot Water Consumption in Dwellings*. 2008, Department for Environment, Food and Rural Affairs,

: London.

26. Kragh, J., et al., *Possible explanations for the gap between calculated and measured energy consumption of new houses.* Energy Procedia, 2017. **132**(Supplement C): p. 69-74.

27. Newcastle University Urban Observatory. *King's Gate weather station*. Available from: <http://www.urbanobservatory.ac.uk/>.

28. ASHRAE, *ASHRAE Guideline 14-2002: Measurement of Energy and Demand Savings (2002)*. 2002.

29. US Department of Energy. *Air Flow Network Multizone Surface* 2018 [cited 2018 31 Jan]; Available from: <https://bigladdersoftware.com/epx/docs/8-5/input-output-reference/group-airflow-network.html#airflownetworkmultizonesurface>

30. Jason Palmer; Ian Cooper, *United Kingdom Housing Energy Fact File*. 2013, Department of Energy and Climate Change: United Kingdom (London).

31. Eguaras-Martínez, M., M. Vidaurre-Arbizu, and C. Martín-Gómez, *Simulation and evaluation of Building Information Modeling in a real pilot site.* Applied Energy, 2014. **114**: p. 475-484.

32. Wei, S., R. Jones, and P. de Wilde, *Driving factors for occupant-controlled space heating in residential buildings.* Energy and Buildings, 2014. **70**: p. 36-44.

33. Rodríguez, G.C., et al., *Uncertainties and sensitivity analysis in building energy simulation using macroparameters.* Energy and Buildings, 2013. **67**: p. 79-87.

34. Guerra Santin, O., L. Itard, and H. Visscher, *The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock.* Energy and Buildings, 2009. **41**(11): p. 1223-1232.

35. Huang, P., G. Huang, and Y. Wang, *HVAC system design under peak load prediction uncertainty using multiple-criterion decision making technique.* Energy and Buildings, 2015. **91**: p. 26-36.

36. Corrado, V. and H.E. Mechri, *Uncertainty and sensitivity analysis for building energy rating.* Journal of Building Physics, 2009. **33**(2): p. 125-156.

37. Wang, L., P. Mathew, and X. Pang, *Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building.* Energy and Buildings, 2012. **53**: p. 152-158.

38. de Wilde, P., W. Tian, and G. Augenbroe, *Longitudinal prediction of the operational energy use of buildings.* Building and Environment, 2011. **46**(8): p. 1670-1680.

39. Baldi, S., et al., *Real-time monitoring energy efficiency and performance degradation of condensing boilers.* Energy Conversion and Management, 2017. **136**: p. 329-339.

40. O’Brien, *Best Practice Guide, Energy Efficiency: Steam, Hot Water and Process Heating Systems* (2015), Sustainability Victoria.

41. Brian Anderson, *Conventions for U-value calculations*. 2006, Building Research Establishment Scotland: BRE Press, Garston, Watford.

42. Hou D., I.G. Hassan, and L. Wang, *Review on building energy model calibration by Bayesian inference*, Renewable and Sustainable Energy Reviews, 2021. **143,** 110930. https://doi.org/10.1016/j.rser.2021.110930.

Table 1: Sequence of uncertainty analysis on building energy models

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Field | Model inputs | Input range | Distribution of uncertainty | Reported variation in model output |
| Building energy model | Occupant behaviour | Presence, density, heat gain | Binary and non-binary | No reports found | 30% [31]  4-26% [32] |
| Building envelope thermal properties | U-values, thickness, surface and moisture properties | Range using Mean and SD [33] | Even [22]  Normal [33] | 42%[34] |
| Weather conditions | Wind speed, direction and pressure coefficients, solar irradiance, air humidity and temperatures | Cold, Med, Hot [33] | Normal [35]  Bivariate Normal [36]  Discrete Distribution [33] | -4% to 6.1% [37] |
| Site micro-environment | Wind-pressure coefficient, ground albedo | Range using Mean and SD [33] | Normal [33] | Not reported |
| HVAC | Values assumed for CoP, SEER and η | Best practice, typical | Normal Distribution [22]  Gamma Distribution [38] | -15.3% -70.3% [37] |
| Internal Gains |  | Low, Med, High [33] | Uniform discrete [33] |  |
| Operational regime | Controls and Scheduling of all HVAC, lighting and plug-in items | Good, average and poor practice [37] | Uniform Discrete [37] | -28.7%-79.2% [37] |
| Observational data | Gas [1] | - | - | Normal [22] | n/a |
| Electricity [1] | - | - | Normal [22] | n/a |
| Temp (Kitchen) | - | - | Normal [22] | n/a |
| Temp (master) | - | - | Normal [22] | n/a |
| Notes  [1] The gas and electricity meters’ accuracy were expected to comply with SI 684 (1983) and IEC 62053 respectively that allow +2.5% or −3.5% of compound instantaneous deviations. | | | | | |

Table 2 Parameter inputs for energy model development of the case-study building

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Description** | Uncertainty range |
| Heating | Natural gas boiler serving a radiator central heating system |  |
| Heating setpoint (setback) | 19°C (16°C) | 17.5°C-20.5°C |
| Heating schedule | 02:00-11:00 + 16:00-24:00 |  |
| Ventilation | Natural ventilation (mechanical extract to family bathroom and en suite) |  |
| Ventilation rate | Highly stochastic, controlled by occupants via openable windows |  |
| Gas boiler seasonal efficiency | 65% (15 years old non-condensing gas-fired system boiler – 77°C/55°C F+R) | 60% - 75% |
| DHW consumption | 0.59 litre/m2/day |  |
| Cooling setpoint (setback) | Uncontrolled |  |
| Nominal lighting power density | 1.4 W/m2 (manually controlled) to achieve 200 lux |  |
| Occupants | 2 people in total |  |
| Internal gains[a] | 6 W/m2 |  |
| Gross (conditioned) area | 168.66m2 (148.93m2) |  |
| Observed annual gas (electricity) consumption (2016) | 15,381 kWh (2,991 kWh) |  |
| **Fabric properties:** |  |  |
| Glazing (with low emissivity coating) | 1.788 W/m2K (3mm self-cleaning pane, 20mm Argon filled cavity, 3mm low emissivity pane) | |
| Glazing G Value (solar transmittance) | 0.691 |  |
| External walls [b] ( W/m2K) | 0.544 | ± 15% |
| Roof [c] (W/m2K) | 0.213 | ± 15% |
| Floor [d] ( W/m2K) | 0.335 | ± 5% |
| Infiltration (ac/h) [e] | 0.5 | 0.25 - 0.95 |
| [a] Electricity (ICT and appliances): 3 W/m2; Gas (catering): 3.3 W/m2 | | |
| [b] 100mm brickwork, 50mm Stone wool insulation, 100mm blockwork, 10mm plasterboards | | |
| [c] 25mm Clay tile roofing, loft space, 180mm glass fibre quilt insulation, 10mm plasterboards | | |
| [d] 100mm cast concrete, 7mm screed, 4mm high gauge polythene DPM, 5 mm foil-backed underlay, 15mm solid wood flooring | | |
| [e] Empirical values derived from table 4.16 (CIBSE Guide A) for a two-storey property on normally exposed site | | |

Table 3 Input parameter variations for Bayesian emulator development

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | V1 | V2 | V3 | V4 | V5 | V6 | V7 | V8 |
| Description | Heating setpoint [17.5°C-20.5°C] | Boiler seasonal efficiency | External wall U-value | Roof U-Value | Floor U-value | Infiltration rate (ach) | DHW consumption  (L/day/person) | Cooking |
| Base model input (1st wave) | 19°C | 65% | 0.544 | 0.213 | 0.337 | 0.5 | 300 (1st wave)  120 (2nd wave) | 3% of total domestic energy use |
| Base model input (2nd wave) | 17.5°C | 65% | 0.544 | 0.213 | 0.337 | 0.5 | 120 |  |
| Range of variation | ± 1.5°C | 60% – 75% | ± 15% | ± 15% | ± 5% | 0.25 to 0.95 | 70-250 L/day | 1.05% - 6.3% |
| Rational | [a] | [b] | [c] | [c] | [d] | [e] | [f] | [g] |
| Uncertainty quantification | Forward | Forward | Forward | Forward | Forward | Forward | Inverse | Inverse |
| Element varied in E+ batch simulations | Heating setpoint | Boiler seasonal efficiency | Wall cavity Insulation thickness [ 40mm-63mm] | Insulation thickness [150mm -210mm] | Insulation thickness [45mm- 55mm] | [e] |  |  |
| [a] Manufacturer’s room thermostat resolution reported at ± 0.5°C with an additional ± 1°C allowed for time-dependent drift degradation.  [b] Boiler insulation, heat exchanger and working fluid degradation, limescale and total dissolved solids leading to an accumulated min and Max performance degradation of 4% to 23% [39, 40]. These levels of degradation were imposed on boiler manufacturer’s quoted efficiency of 78%  [c] Although most literature report in-situ wall and roof measurements to be better than elemental method calculation suggestions [14, 16, 17], an equally distributed ± 15% imposed to first cater for all eventualities and enable the uncertainty emulator to assess the entire Latin hypercube space (including worst scenario range).  [d] as per [c] although the magnitude of variations reported for floors were smaller than those of walls/roofs [17] and non-suspended ground floors with no air cavities have much greater thermal unity [41] so a tighter band of ± 5% was imposed to reflect literature findings.  [e] As outlined in the last paragraph of sections 2 and 4.  [f] From field measurements of DHW consumption in the UK [25] where the mean DHW consumption per person in the UK is reported as 122 litres/day ± 18 litres/day (i.e. ±15% variation) leading to mean DHW energy consumption of 16.8 MJ/day ± 2.2 MJ/day (95% statistical confidence). In the 1st wave the model input was a much larger values of >300l/day and 53MJ/day, However the model predictions were calibrated to return close results to the observed energy consumption for the case-study building given its high fossil fuel consumptions.  [g] Cooking has been observed to currently account for an average of 3% of total household energy demand with historical data also indicating a maximum of 6% [30]. This observed data informs the average and maximum cooking demand with 1% also selected by the authors to represent a probable lower boundary. | | | | | | | | |