Department for Business, Energy & Industrial Strategy

PEER REVIEW OF THE BUILDING ENERGY EFFICIENCY SURVEY ENERGY USE MODEL

November 2016

This document is available in large print, audio and braille on request. Please email correspondence@decc.gsi.gov.uk with the version you require.

PEER REVIEW OF THE BUILDING ENERGY EFFICIENCY SURVEY ENERGY USE MODEL

Any enquiries regarding this publication should be sent to us at nondomesticbuildings@decc.gsi.gov.uk

This publication is available for download at www.gov.uk/government/publications.

Contents

Acknowledgements

The Department for Business, Energy and Industrial strategy would like to thank the University College London (UCL) for carrying out this peer review exercise.

Glossary

Term	Definition/explanation
3DStock	A model representing buildings and premises in 3 dimensions, rather than just as footprints on a map, developed at UCL.
BEES	Building Energy Efficiency Survey.
CIBSE Guide A	A building services industry guide to energy use in buildings, used to provide certain pieces of input data to the energy use model, e.g. temperature set-points for different activity zones.
Candidate model	A specific set of parameter values considered when training the SimStock inference engine using a genetic algorithm.
DEC	Display Energy Certificate.
EnergyPlus	A whole building energy simulation engine used to model both energy use —for heating, cooling, ventilation, lighting, and plug and process loads—and water use in buildings. Its development is funded by the U.S. Department of Energy.
EUI	Energy Use Intensity in $kWh/m^2/year$.
Energy use model	The BEES energy use model developed by Verco for use in estimating energy end use intensities for a particular building based on a telephone survey response.
IDF	Input Data File used by EnergyPlus.
Inference engine	The inference engine is the core of the SimStock model. It directs exactly how source data is converted into EnergyPlus IDFs ready for simulation. The key feature of the inference engine is a set of parameter values which can be adjusted to best reproduce the known energy use of the building stock.
ND-NEED	Non-domestic National Energy Efficiency Data-Framework, a database of energy use created and managed by DBEIS to provide a better understanding of energy use and intensity in domestic and non-domestic buildings in Great Britain.
SCU	Self-Contained Unit. This is the basic unit defining premises in the 3D Stock data and in the SHU energy data used as source data for SimStock.
SHU	Sheffield Hallam University, the original source of a study of energy data

Term	Definition/explanation
	broken down by end use.
SimStock	A building stock modelling framework developed at UCL Energy Institute. SimStock uses EnergyPlus as a simulation engine for modelling the energy use of individual buildings.

1 Introduction and summary

The Building Energy Efficiency Survey (BEES) used a model to convert telephone survey answers for a particular building into estimates of energy use. The goal of this peer review project was to assess the ability of the BEES energy use model to represent the energy performance of non-domestic buildings, based on data obtained through the BEES telephone survey. The buildings in a sample of two building types (health centres and offices) were modelled using an independently-developed shadow model (the SimStock building energy stock model) to assess whether or not it would be possible to significantly improve the modelled results using the same data and help inform the department as to the level of confidence to place in the BEES energy use model.

Section 2 of this report presents the methodology used in the shadow modelling exercise. SimStock has three main differences from the BEES energy use model. First, a simple 3D representation of the building is produced rather than only using floor area. Second, a dynamic buildings simulation model is used rather than a static calculation. Finally, telephone survey responses are mapped to model parameters and values using a genetic optimisation approach rather than engineering judgement and manual calibration.

In Section 3, the outputs from the energy use model and SimStock shadow model are compared against matched energy meter data to assess the accuracy of their predictions. Both models performed fairly well at the aggregate level, with the energy use model predicting energy use almost as well or the same as the SimStock shadow model, depending on the sample used to test the models.

Section 4 presents the findings from a sensitivity analysis of uncertain inputs to individually constructed EnergyPlus models for a sample of site surveyed health centres and offices, to help assess the choice of variables within the energy use model. It was found that the energy use model included the key variables that impact the energy use in a building. It should be noted that uncertainty in the input telephone survey data could still lead to significant uncertainty in the energy use estimates, regardless of the model used to produce these estimates.

Finally, in Section 5, the key conclusions relating to the BEES energy use model are presented. Overall, the energy use model was found to be fit for purpose. That is, it produced reasonable estimates of energy use given the available input data. Although basing estimates of energy use on limited input data can lead to inaccuracy at a building level, outputs were robust at the sub-sector level and should provide a useful resource.

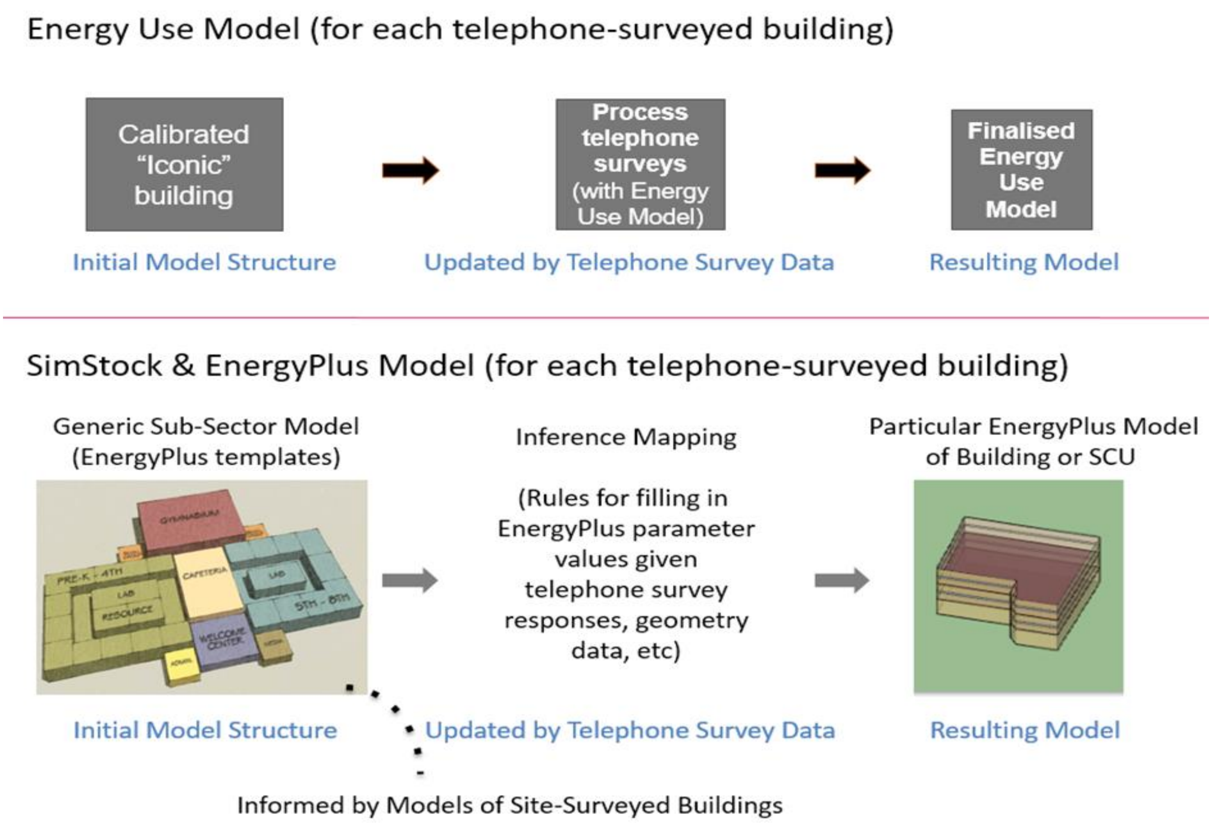
2 Methodology

2.1 BEES energy use model

Telephone surveys and site surveys formed the primary source data for the energy use model. The telephone surveys were conducted by GfK NOP, a market research company, with over 3500 non-domestic building occupiers. The majority of the questions were put to all respondents, as well as a small number of more specific questions for each sector. Respondents were also asked if it would be possible to conduct a site survey of the building in order to verify the telephone responses and gather more detailed information about the building.

Using information gathered during the site surveys, models were built for a typical building (known as an “iconic” building) in each sector. The results for this iconic building were then adjusted according to answers on the telephone survey in order to predict energy end uses for each building. The values used for the adjustments were calibrated in an attempt to match the known distribution of metered electrical and non-electrical energy use. A comparison of this methodology to that of the SimStock shadow model is shown in Figure 2.1.

Figure 2.1. Comparison of the BEES energy use model and SimStock approaches.



2.2 Peer review model

The methodological approach used for this review of the energy use model was a shadow modelling exercise. A shadow model is a second model which attempts to calculate the same outputs as the model under review, generally using a different modelling approach. Since the energy use model is a static spreadsheet model in a tree format as shown in Figure 2.2, the primary difference in modelling approach is that the SimStock model employs dynamic energy simulation rather than a static calculation (Coffey et al. 2015). This is better able to capture the interactions between factors which influence energy demand, as well as explicitly modelling the built form of each premise. The SimStock model has been developed further for this project leading to the second major difference; where the energy use model uses parameters based on engineering judgement and site surveys, in SimStock, engineering judgement is supplemented by an automated approach to fine-tuning the parameters.

A third difference is that the energy use model models a hypothetical square meter of premises, while SimStock explicitly models the whole premises before calculating the energy use per square meter.

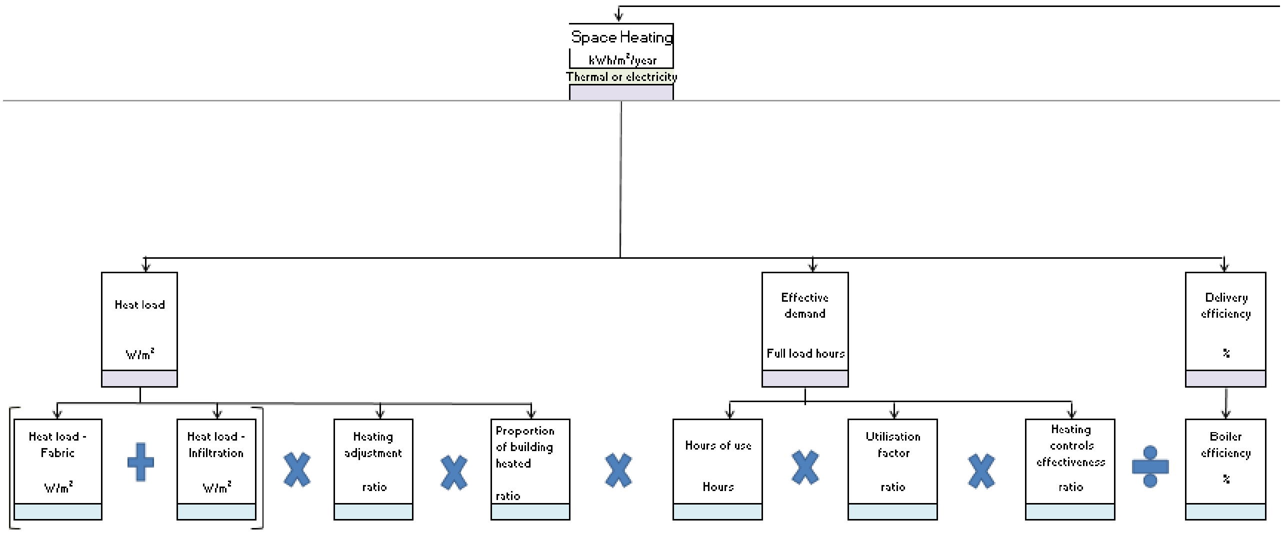


Figure 2.2. Section of the energy use model tree diagram used to calculate space heating demand.

The final outputs of both models are predictions of energy use intensity (EUI) in units of kWh/m^2/yr, for electrical and non-electrical energy use, and for end uses such as space heating and lighting. This common output format makes the models comparable.

The peer review shadow model was built in two stages. The first stage covered a small number of health centres and was used as a prototype to develop and test the modelling approach. The second stage covered offices and incorporated more detail and more automation into the modelling process.

2.3 Site survey models

The first step for each stage was to build models of a small number of site-surveyed buildings. The process was similar for health centres and offices, although the process of geometry building was automated for the offices.

- Hand built model geometry in OpenStudio energy modelling software (health centres only)
- Model geometry generated by SimStock (offices only)
- Demand and systems inputs informed by site surveys and telephone surveys (both)
- Model input calibration and sensitivity analysis in the SimStock framework (both)

These site survey models were not used directly in the SimStock predictions, but rather were used to test the sensitivity of the models to changing parameters, representing uncertainty in the telephone survey responses. This was an exploratory step used to understand the data in the telephone survey before moving on to the main shadow modelling exercise. Results of this sensitivity analysis, and what it tells us about the energy use model, are presented in Section 4.

2.4 Data cleaning

The amount of telephone survey data available for health centres was relatively small for a statistical approach to modelling. Once the data had been cleaned in a manner which made them usable in the SimStock model, only 22 records remained on which to train the model and derive weightings for the inference engine. The reasons for removal of records are given in Table 2.1 and

Table 2.2.

Table 2.1. Data cleaning for health centres.

Reason for removal	Number of records remaining
Starting no. of records	62
Unreliable floor area	61
No valid record found for premises energy data	43
No matched electricity meter data	30
No matched gas meter data (where expected)	23
Electricity outliers	22

Table 2.2. Data cleaning for offices.

Reason for removal	Number of records remaining
Starting no. of records	775
Unreliable floor area	754
Dropped from energy use model	637
No matched electricity meter data	450
No matched gas meter data (where expected)	409

Respondent did not know what building services were present	372
No unique identifier supplied	364
Respondent answered that building was not serviced	330

The number of records available for offices was much higher. This meant that even once the office data had been cleaned, a representative sample had to be taken of the remaining records since the number of records for which geometry details could be produced was limited to 200 to allow reasonable processing time. Criteria used in selecting the sample were:

- Minimise divergence in mean, median and standard deviation between the population and the chosen sample for:
  - Non-electrical energy use intensity (EUI)
  - Electricity EUI
  - Floor area
- Maintain the same ratio of whole building to part of building (premises) records
- Maximise combinations of boiler, HVAC and mechanical ventilation categories
- Include the site surveyed records in the sample

The records were sampled by randomly selecting 200 records from the remaining 330 records, evaluating them on the criteria above, and keeping track of the current best candidate population. The sampling was repeated 1,000 times and showed little further improvement after the first few hundred samples, giving confidence that the chosen sample was as representative as possible.

2.5 Input data sources

Three main sources of input data, other than the BEES telephone survey, were used in the SimStock model. These sources are described in the following sections.

2.5.1 3D Stock geometry data

The geometry is automatically generated for the 3DStock model by bringing together existing data from two national sources (Evans et al. 2014). The first source is the Ordnance Survey – the UK's national mapping agency – specifically, the Ordnance Survey Address Base (OSAB) and Mastermap digital map products (Ordnance Survey 2015a; Ordnance Survey 2015b). The second source is the Valuation Office Agency (VOA) of Her Majesty's Revenue and Customs (Valuation Office Agency 2015).

Heights for the majority of polygons that represent a self-contained unit (SCU) were derived from the Ordnance Survey product called Simple Building Heights. This product does not cover the whole of England and Wales and so for some regions, in particular more rural locations, heights were derived by examining photos of the individual buildings and adjoining properties on the internet, in particular using Google StreetView. In all cases, the heights were estimated to the eaves unless otherwise stated.

2.5.2 Line entries for premises end use electricity data

Between 1991 and 2000, the Resources Research Unit of Sheffield Hallam University (SHU) performed detailed internal energy surveys on approximately 700 non-domestic premises. The survey procedures, data processing and analyses carried out for the original SHU research have been published in various locations (Mortimer et al. 1999; Mortimer, Ashley, et al. 2000; Mortimer, Elsayed, et al. 2000; Penman 2000; Elsayed et al. 2002). Fundamental to this premises-level method was the reconciliation of bottom-up calculations, based on equipment, with the metered energy use (taken from billing data) of each premises. Subsequent work by Liddiard (2012) analysed electricity use – excluding heating and cooling – at the room scale. The line entries (room-level) data used in this study are based on that research.

To allow for changes in equipment profiles since the original SHU surveys, populations of computers have been increased and lighting has been updated, where deemed appropriate. Note that this was done after the data had been cleaned and otherwise processed, to avoid misrepresenting the original data.

2.5.3 National Calculation Methodology schedules

Default schedules defining the pattern of daily, weekly and annual occupancy and energy demand for each activity type (e.g. cellular office, circulation, etc.) were sourced from the National Calculation Methodology (NCM) activities database (Building Research Establishment 2015). These included heating, cooling, lighting and equipment demand schedules, and occupancy schedules. Each of these schedules was used to inform the shape of demand (or occupancy) across the year rather than the peak value of demand. These schedules were modified based on responses to telephone survey questions including number of operating days per week and number of holiday weeks per year.

2.6 SimStock model structure

The SimStock model structure consists of a mapping from sources of data about a building onto inputs in the EnergyPlus dynamic energy simulation model, as shown in Figure 2.3. Other than the geometry generation, which is a simple conversion, the main problem is to map the linguistic descriptions in the survey data, e.g. *age_of_boilers = 0To7Years*, to a quantitative boiler efficiency in the EnergyPlus input file, e.g. 89% efficient.

In order to overcome this, BEES telephone survey data responses were first mapped to particular variables in EnergyPlus. Then a genetic algorithm was used to establish appropriate values for these parameters, given particular survey responses. This algorithm used a set of buildings (the training set) to ‘learn’ what parameter values produced the closest energy use estimates to the known energy use for the building. Once these values had been learnt, they were applied to a new set of buildings (the test set) to see how well the model performed. This process is described in more detail in Section 2.7.

Because the model had access to the known energy use for buildings in the training set, it would be able to make better estimates of energy use for these buildings. In contrast, the model had to estimate the energy use for the test set based on the parameter values learnt from the training set, resulting in less accurate estimates but a better reflection of model performance.

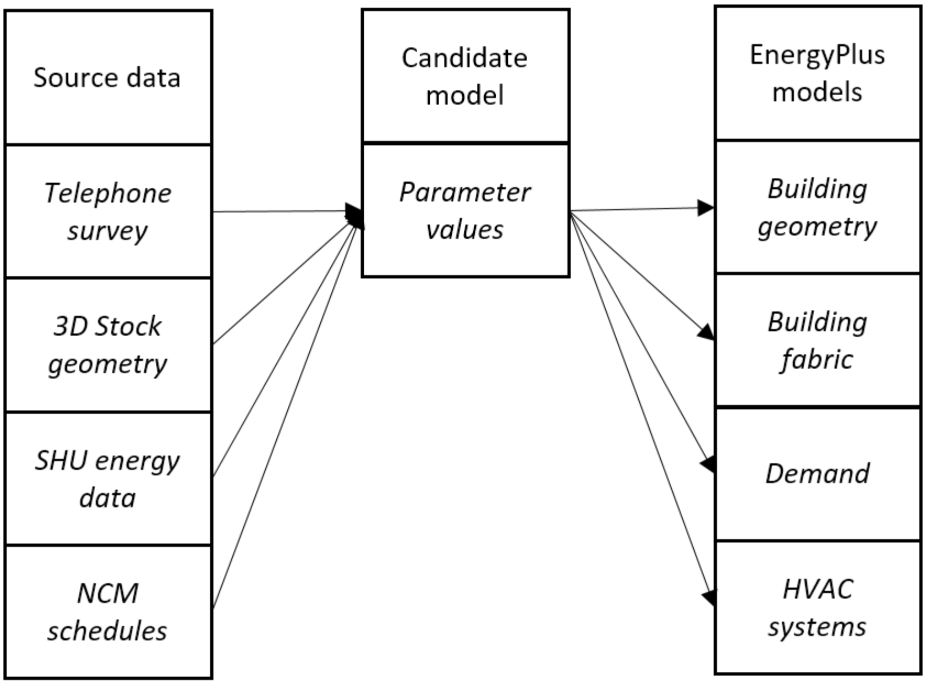


Figure 2.3. For each candidate model, source data were used to generate EnergyPlus models of each record, adjusted by that candidates set of parameter values.

2.7 Inference engine training

The process of inference engine development is shown in Figure 2.4 and attempted to minimise the differences between the end use intensity (EUI) predictions from the SimStock model and the known EUIs for electricity and gas. This involved setting the parameter values used in the SimStock model to better match the known meter reading values on a subset of the data (the training set). The values taken to represent known energy data were sourced from a combination of ND-NEED (DECC 2014) and Display Energy Certificate records.

Once the sets of records for health centres and offices had been selected, they were each split into two sets; one training set, and one test set. The training set was used to train the inference engine, using genetic optimisation to learn appropriate parameter values for the inference engine. In this process, the inference engine was seeded with randomised values for each parameter in the mapping, creating an initial candidate model. For example, the efficiency of boilers less than 8 years old would be mapped to a value randomly chosen between 60% and 95%. The resulting EnergyPlus models were then simulated, producing annual EUIs for electrical and non-electrical energy.

During this phase it was observed that a number of records were displaying very high variance between known EUI and predicted EUI, even while the other building models were improving. These outliers were investigated in more detail. In some cases miscoded values could be identified and corrected (for example where a building was coded as having electric heating but had high metered gas use the main heating fuel was changed to gas). Other buildings where the error was not seen to be rectifiable, including one recycling centre which was misclassified as an office, were removed from the training set.

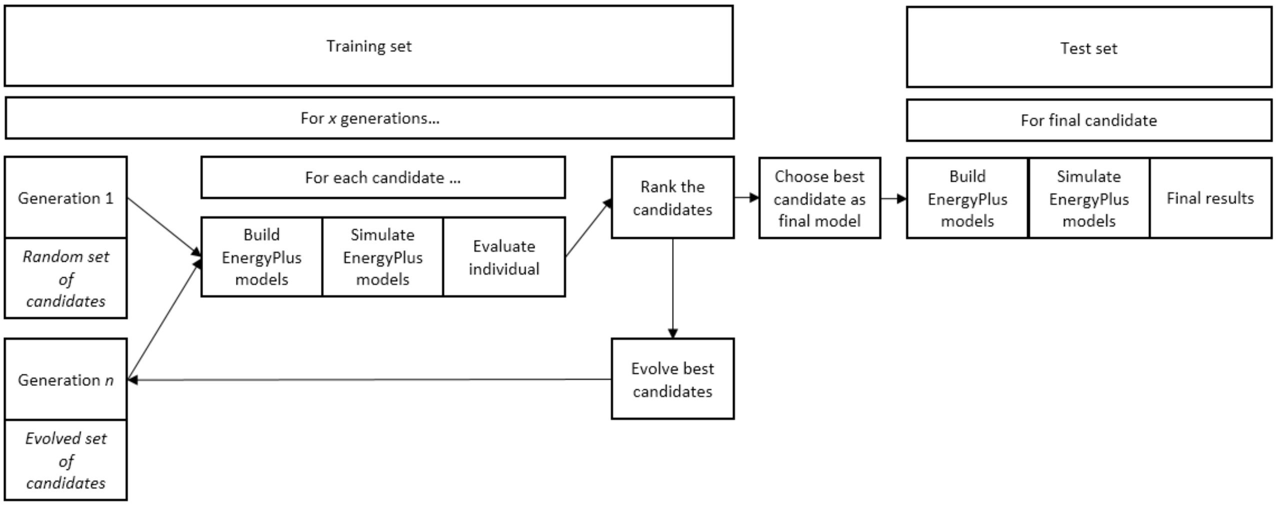


Figure 2.4. Flowchart showing the process of inference engine training and testing.

The genetic algorithm attempted to minimise two optimisation criteria. The first optimisation criterion was the widely-used root mean squared error (RMSE) of the amount by which the predictions for each record differed from the known energy use for each record. A larger number indicates a larger error. This drove the genetic algorithm to favour candidate models (sets of parameter mappings) which reduced the errors of the individual predictions, and in particular those which reduced the error on the greatest outliers. Other metrics such as mean absolute error or median-based criteria are less affected by outliers. The second criterion was median bias – whether, on average, the median EnergyPlus model prediction was over- or under-estimated relative to the known median energy use. Larger positive numbers indicate a larger overestimation of the median, while larger negative numbers indicate a larger underestimation of the median. This favoured candidate models which produced predictions centred on the median of the known distribution and so was less affected by outliers. These two criteria were summed for each fuel and averaged across all records to give each candidate model a “fitness score”.

The genetic algorithm kept candidate models with the lowest fitness score and discarded the others (survival of the fittest). The fittest candidate models were then bred together, by randomly combining their parameter values, to produce offspring candidate models which were then tested in the same way. As shown in Figure 2.5, this process continued for 25 to 30 generations, by which time the rate of improvement in the fitness of the best candidate model between generations had slowed down.

This process left a final mapping of telephone survey responses to parameter values, which could be used to make predictions about building data which were not used to train the model (the test set).

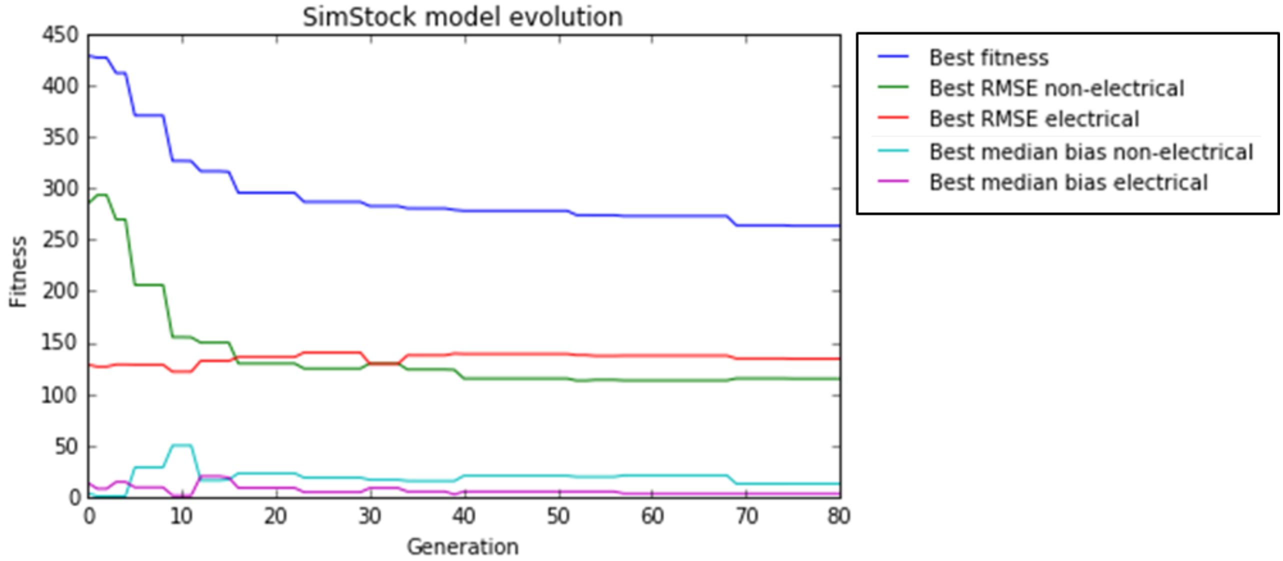


Figure 2.5. An example of model training showing the performance improvement over time, with the rate of improvement slowing significantly after approximately generation 25.

2.8 Final SimStock model run

Using the final candidate model (the best performing mapping of telephone survey responses to parameter values from the model evolution process), EnergyPlus models were generated for a test set consisting of the remainder of the records not used in model training. These EnergyPlus models were run and the resulting outputs converted from the total annual fuel use to end use intensity (EUI) by fuel (for both offices and health centres), and EUI broken down by end use (for offices only).

Due to the larger offices dataset, it was possible to carry out a second run using a slightly less representative set of records. In this second run individual outliers were not removed. This was done to understand the impact of a less labour- and simulation-intensive cleaning process. The results for both runs are presented in Section 3 and labelled ‘Run 1’ and ‘Run 2’ to distinguish between the two.

In cases where the energy use model record represented part of a building, there was a potential mismatch between the energy use model area and the 3DStock model area (and hence the EnergyPlus model area) for that building, since the 3DStock data for a record represented the entire building (excluding any residential activities such as upper floor and/or basement flats). To cope with this mismatch, the entire building was modelled in EnergyPlus and the results presented per m^2 of the floor area in the EnergyPlus model.

Note that producing a comparison of total energy consumption of the building, rather than intensity, would produce different results because the absolute prediction errors would be amplified by the floor area. For example, while both a small and large building may have an intensity error of 5 kWh/m^2/year, this error will have a limited impact on the total (kWh/year) consumption of the small building but a much larger impact on the larger building due to the area affected (though the proportional error would remain the same). Although the total energy consumption of the building was the key output from the energy use model, examining energy intensity facilitates comparison against the SimStock model.

3 Comparison with energy use model

3.1 Overview

For each sector, health centres and offices, scatter plots were generated for electrical and non-electrical energy use. These plots show the predicted energy use intensity (EUI) against the known EUI, which would be the same in a perfectly accurate model. This comparison showed whether it may be possible to produce more accurate energy use predictions using the BEES telephone survey data.

Box and whisker plots were also produced, showing the distribution of predictions alongside the known EUIs. The box shows the upper and lower quartile with a marker at the median. The whiskers extend to 1.5 times the interquartile range, and points beyond that are plotted as outliers. This gives a visual representation of the similarity between the simulated sample and the real sample.

Finally, a comparison of end use intensities from the two models was carried out for offices. This was used to provide more detail on the differences in energy use estimates, though it should be noted that there was no matched energy end use data to serve as a reference.

Overall, the BEES energy use model was shown to make good use of the BEES telephone survey data in predicting energy use, particularly at an aggregate level. However, the results of the comparison show significant variation in the energy use predictions of individual buildings from both models.

3.2 Health centres energy use

The charts in

Figure 3.1 and Figure 3.2 show the SimStock model predictions for the test set using the inference engine trained on the health centres training set alongside the predictions of the energy use model for the same records. The degree of scatter for both models reflects the fact that neither model was capable of making accurate predictions at the level of an individual building.

Figure 3.1. Health centres - comparison of the fit of predicted and known EUIs. The blue line indicates a perfect fit while the red line is a linear trendline.

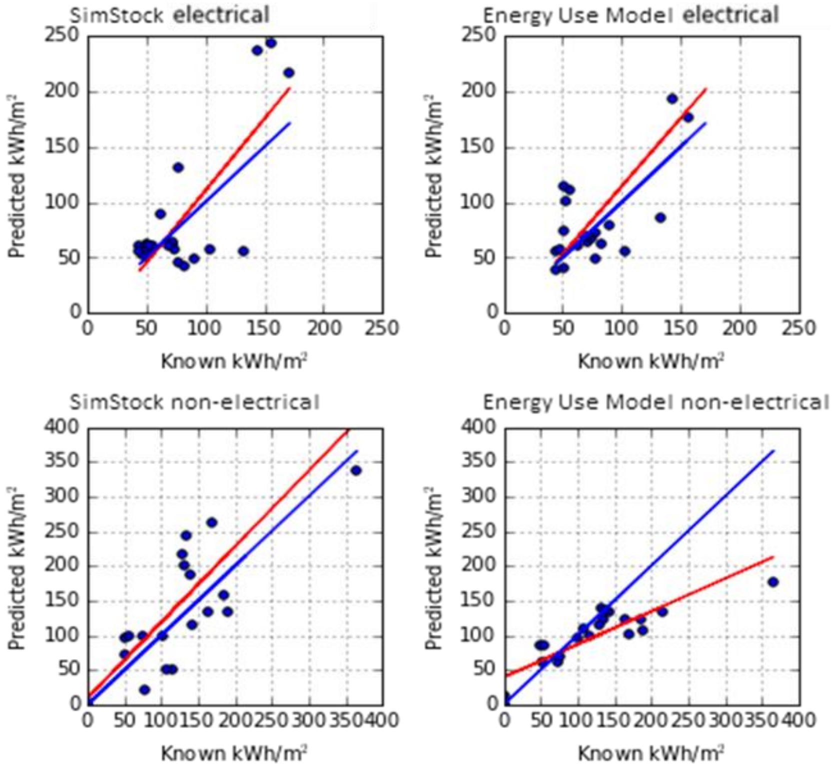


Figure 3.2. Health centres - comparison of distribution of predictions from SimStock and energy use model against known EUIs.

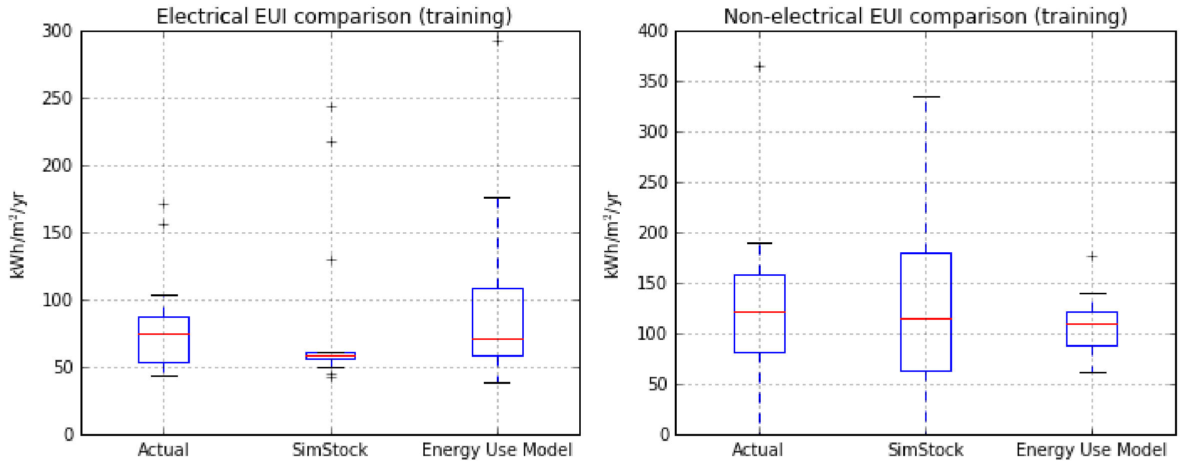


Table 3.1. Health centres - selection criteria. Root mean squared error (RMSE) normalised by the standard deviation of EUIs in SimStock and energy use model predictions for electrical and non-electrical energy use intensity (all records in the cleaned data set).

	SimStock		Energy use model	
	RMSE (train)	RMSE (test)	RMSE (train)	RMSE (test)
Electrical	0.97	1.16	2.60	0.57
Non-electrical	0.58	1.50	1.46	0.94

Note: A lower normalised fitness score reflects a lower RMSE and/or bias, indicating a better performing model.

The health centres results are based on too small a sample to draw any strong conclusions about the validity of either the SimStock model or the energy use model. The small sample size in the training set allowed the genetic optimisation to find a good fit to that part of the dataset, but the parameters do not appear to produce sensible results on the test set. A larger number of records with known non-electrical and electrical data would be required for this model to be improved upon in order to better test the energy use model.

The effect of sample size was particularly noticable in the predictions for non-electrical energy use. A small sample size meant that the genetic optimisation algorithm was prone to overfitting the data in the training set. The effect of this can be seen leading to poorer predictions on the test set in

Table 3.1 for both electrical and non-electrical energy use. However this should also be viewed in the context of the large variability between the training and test set for the Energy Use Model which provides further indication of the effect of a small sample.

In addition, since this sector was used as a prototype, the methodology for accounting for electricity use such as lighting and small power was not fully implemented. This, in combination with the fact that the great majority of space in each health centre was assigned to the same activity class, explains the low variation in electrical EUI in Figure 3.2.

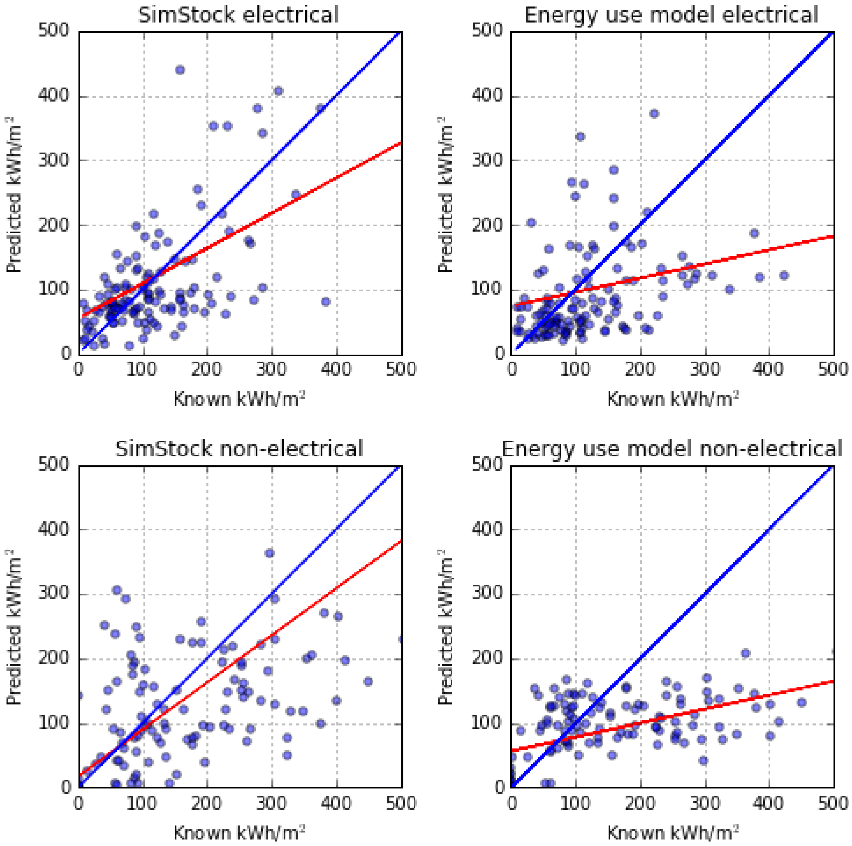
3.3 Offices energy use

As described in Section 2.8, the size of the offices dataset allowed two different runs to be performed. Run 1 contained a more representative set of records with outliners were removed. The charts in

Figure 3.3 and

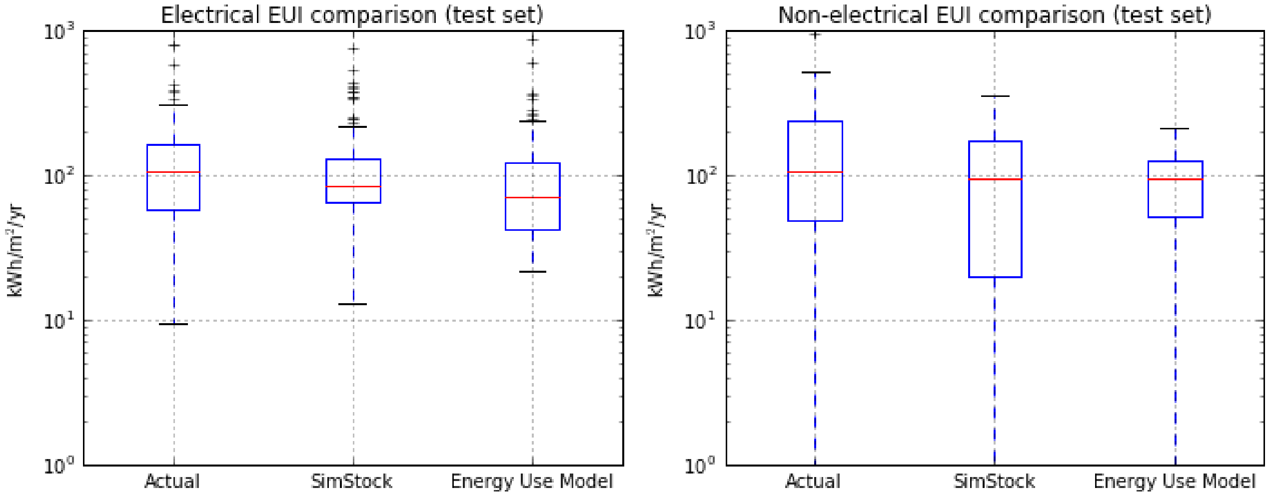
Figure 3.4 show the SimStock model predictions for the test set using the inference engine trained on the offices training set, alongside the predictions of the energy use model for the same records. These are results from Run 1.

Figure 3.3. Offices (Run 1) - comparison of the fit of predicted and known EUIs (test set only). The blue line indicates a perfect fit while the red line is a linear trendline.[1]



[1] Axes have been restricted to 500 kWh/m2/yr to improve readability, which excludes six outliers in the electrical plots and two in the non-electrical plots.

Figure 3.4. Offices (Run 1) - comparison of distribution of predictions from SimStock and energy use model against known EUIs. Axes have been restricted to 1,000 kWh/m^2/yr to improve readability, which excludes one outlier in the electrical plot.



In , though the difference in performance between the two runs was not as large as that in the SimStock model.

Table 3.2 we can see that the energy use model is almost as good or the same at predicting energy use as the SimStock model, depending on the model run. The SimStock model performed better when trained on the first set of records (Run 1) than it did on the second set. The energy use model also performed slightly better overall in Run 1 than in Run 2, though the difference in performance between the two runs was not as large as that in the SimStock model.

Table 3.2. Offices (Run 1) - selection criteria. Root mean squared error (RMSE) normalised by the standard deviation of EUIs in SimStock and energy use model predictions for electrical and non-electrical energy use intensity (all records in the cleaned data set).

Model (Run no.)	Electrical RMSE	Non-electrical RMSE	Electrical bias	Non-electrical bias	Normalised fitness
SimStock (Run 1)	0.90	0.77	-0.15	-0.08	1.91
Energy use model (Run 1)	1.14	0.99	-0.24	-0.07	2.45
SimStock (Run 2)	1.13	1.14	-0.12	-0.27	2.67
Energy use model (Run 2)	1.13	1.16	-0.21	-0.22	2.72

Note: A lower normalised fitness score reflects a lower RMSE and/or bias, indicating a better performing model.

As might be expected, the model which best predicted known energy use intensity (EUI) was the SimStock model run using the training sets, though this is an unfair comparison because the model had the known energy use available. While the energy use model had a worse RMSE than the SimStock model for both electrical and non-electrical energy use using the test set in Run 1, both models had a similar performance in Run 2. This makes it difficult to distinguish between the two models, though it could indicate that the process of carefully filtering the input data in the training set has an impact on SimStock model performance. The energy use model had a slightly larger bias in predicting electrical consumption and a slightly lower bias in predicting non-electrical consumption, though the difference between the two models was negligible.

The energy use model had a slightly worse non-electrical RMSE than SimStock in Run 1, indicating the energy use model could benefit from the addition of a dynamic non-electrical model (at the cost of requireing additional building geometry data). However, this finding is not substantiated by Run 2, making it difficult to draw conclusions. The non-electrical performance of both models was better in Run 1, which indicates that the handling of non-electrical energy use is sensitive to the model inputs.

The predicted electrical energy use from the energy use model was fairly robust between the two runs, providing some evidence that a robust methodology is used. Conversely, differences in electrical energy use were found between the two runs of the SimStock inference engine training (lighting and small power predictions were both almost 100% higher in Run 2, offset by lower space heating from electricity).

Table 3.3 shows the distribution of known and predicted EUIs within the test set for Run 1. Here it is possible to see the systematic under-prediction of non-electrical EUI in the energy use model. While the SimStock model performs better at matching the population statistics for non-electrical EUI at the upper end of the scale, it also has a problem with under-predicting non-electrical energy use at the lower end. This is shown by the non-electrical EUI at the 25th percentile which is around one third that of the known EUI.

Table 3.3. Variance from EUI of the known offices energy data (Run 1 test set only) for SimStock model and the energy use model.

	Electrical		Non-electrical	
Variance at EUI (measured)	**SimStock**	**Energy use model**	**SimStock**	**Energy use model**
mean	-4.1%	-28.1%	-17.4%	-58.9%
standard deviation	10.7%	-12.2%	8.2%	-143.4%
25th percentile	8.2%	-39.8%	-200.0%	-15.7%
50th percentile	-15.6%	-38.4%	-3.0%	-2.8%
75th percentile	-18.6%	-26.6%	9.6%	-23.7%

It is worth noting that, while the SimStock model matched the population statistics relatively well for electrical energy use as well as for non-electrical fuel use, as shown in Table 3.3, this is in spite of the model accuracy for individual records only being a little better than the energy use model (in Run 1). The energy use model also showed a much lower standard deviation than those of the known non-electrical EUI, confirming the observation that the model did not capture the true variation of heating energy use within the sample.

Underestimating energy use means that projected savings from measures in abatement modelling would be conservative. This has less of an impact than overestimating energy use as it significantly reduces the risk of recommending measures that are not warranted given their likely level of savings.

3.4 Offices end-use intensities

When considering the predictions of end uses, the largest components are space heating, lighting, and ICT/small power (these should be considered together since the two models make different assumptions about what constitutes ICT and small power). The SimStock model predicts more than 50% higher space heating loads from electricity and more than 50% lower cooling loads from electrical energy compared to the energy use model. The predicted hot water loads are similar between the two models, giving confidence in any measures targeted at saving energy on this end use.

While it has been possible to achieve some degree of agreement between the known energy use and that predicted by the EnergyPlus models, the degree of uncertainty in some of the parameters means that the same results could equally have been achieved with different combinations of parameter values, with little or no reason to select one combination over another. True model calibration would require much more detailed information on building fabric and operation, as well as known energy use at a higher level of granularity.

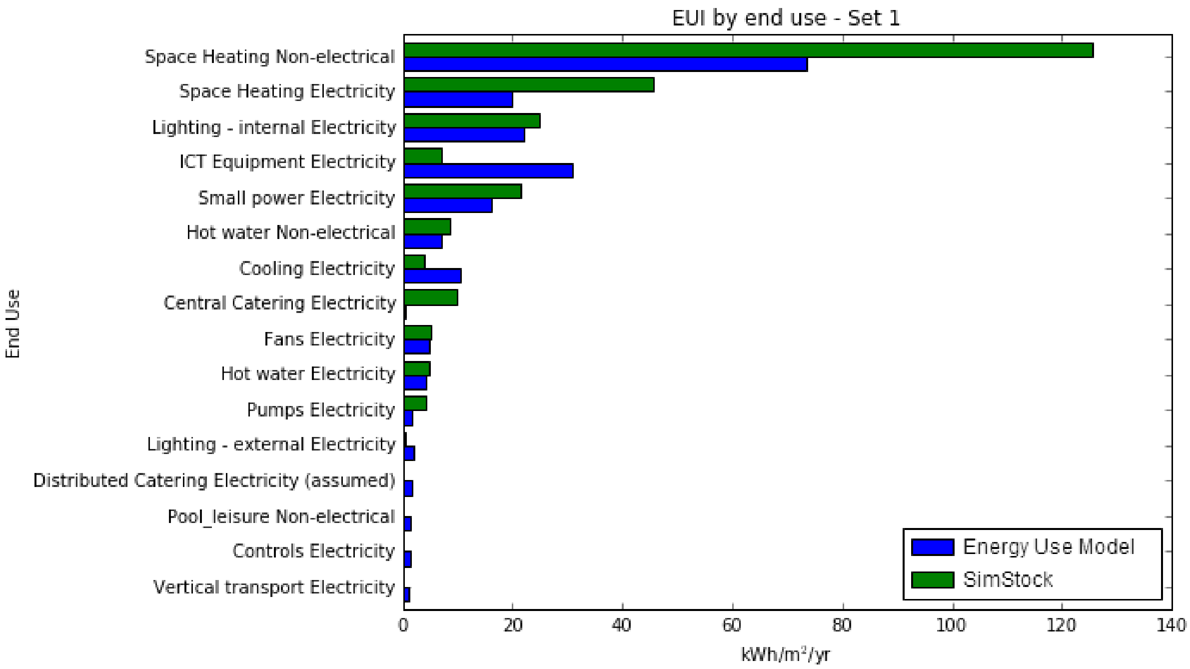

EUI by end use - Set 1
End Use
Space Heating Non-electrical
Space Heating Electricity
Lighting - internal Electricity
ICT Equipment Electricity
Small power Electricity
Hot water Non-electrical
Cooling Electricity
Central Catering Electricity
Fans Electricity
Hot water Electricity
Pumps Electricity
Lighting - external Electricity
Distributed Catering Electricity (assumed)
Pool_leisure Non-electrical
Controls Electricity
Vertical transport Electricity
Energy Use Model
SimStock
0
20
40
60
80
100
120
140
kWh/m2/yr

Figure 3.5. Mean predicted energy use intensity by end use for all records in the training and test sets using the learned parameters from Run 1

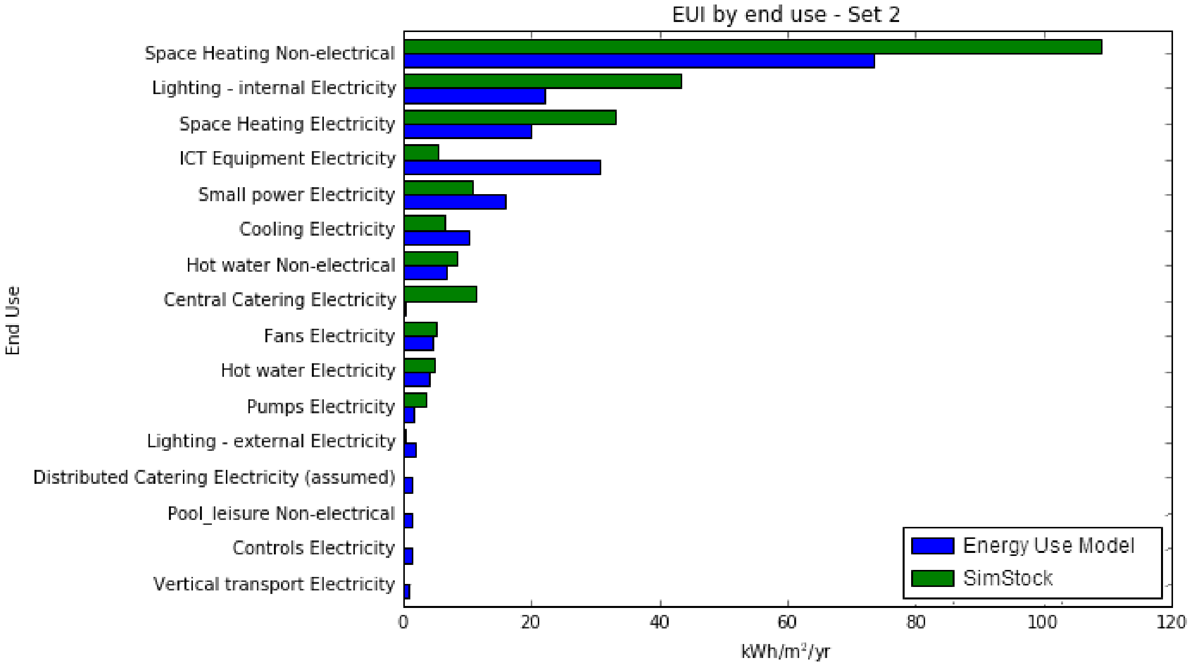


Figure 3.6 Mean predicted energy use intensity by end use for all records in the training and test sets using the learned parameters from Run 2

It should be noted that for both the energy use model and the SimStock model, the accuracy of predictions of end uses is uncertain since metered data for comparison with the values predicted by the models is not available. This is more of a concern for electrical energy use since fuel use is dominated by space heating, with hot water and catering uses accounting for only a small proportion of the total non-electrical energy use.

4 Sensitivity analysis

Sensitivity analysis was carried out to help show whether the BEES energy use model included variables shown to have a significant impact on energy use. The results indicated the model was appropriately specified, but that there may still be significant uncertainty in the outputs as a result of uncertainty in the BEES telephone survey. In general, the input parameters around which there was most uncertainty were found to have the largest influence on the modelled energy use.

4.1 Methodology

Detailed geometry models were created for the buildings which had been site surveyed. For the offices sector, this geometry was produced in an automated manner while for health centres the geometry was drawn manually based on site plans, or on StreetView images where plans were not available. Both methods produced IDF files which can be read by EnergyPlus, and also edited in a text editor. In order to prepare the IDF files for sensitivity analysis, placeholders were added which could then be substituted in each model run. The generated models did not describe systems such as ventilation rates and so these were also added, along with placeholders for the key variables such as ventilation and infiltration rates.

Once the models were set up, a baseline simulation was run with each variable taking on a typical value. Subsequently, each variable was varied to a high and a low value, representing the upper and lower range of uncertainty for that building based on engineering judgement. All other model aspects were kept constant[2] and the model was then run again, allowing a comparison of energy use intensities with the baseline simulation.

The results were produced as ranked tornado charts for the baseline gas and electrical energy use per m^2, showing the variance from the baseline simulation results when changing each variable individually (Figure 4.1).

[2] It would have been possible to carry out a more comprehensive sensitivity analysis of parameters, using a method which accounts for interactions between variables, such as the Morris sensitivity analysis. However, given the level of uncertainty already identified by the simple approach, this was not considered likely to improve understanding.

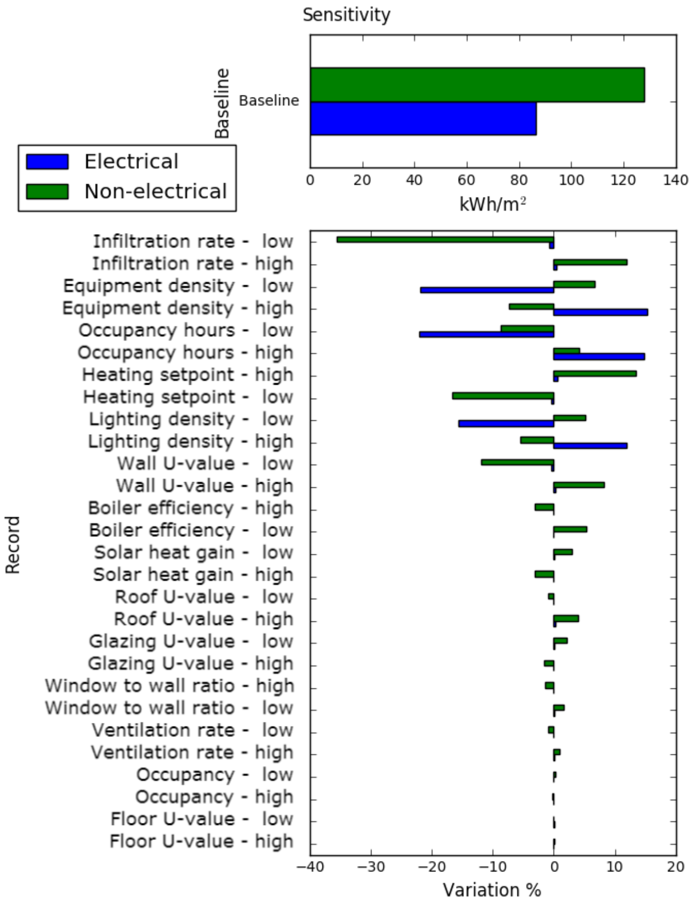


Figure 4.1. An example of the tornado charts produced to present EnergyPlus model sensitivity.

4.2 Sensitivity analysis results

The energy use model accounts for the variables that have a large impact on the energy use. However, due to the difficulty asking very detailed questions and limited time available during a telephone survey, some of the values for these variables have a high degree of uncertainty or are based on assumptions. So while the energy use model appears appropriately specified, significant uncertainty could be introduced at the record level due to the underlying data.

For example, ventilation rates and heating set-point temperatures had a large impact on heating energy use when varied over a reasonable range. These were particularly uncertain parameters since they are hard to measure quantitatively in a telephone survey or a short site survey. The infiltration values used in the energy use model were based on assumptions to convert qualitative survey responses (if present) into estimated infiltration rates, while heating set-points were based on CIBSE Guide A.

Boiler efficiency also made a significant difference to heating energy use, even within a relatively narrow range (±5% boiler efficiency). While boiler efficiency was included in the energy use model, it could only take one of three values. Again, given the uncertainty in the underlying data this was an appropriate modelling choice, however, it does introduce a lack of granularity and associated uncertainty at the record level.

Finally electrical energy use from lighting and equipment also had a relatively large effect on predicted electrical energy use and, to a lesser extent, non-electrical energy use. As discussed in Section 3, the energy use model accounted for aggregate level electrical energy use well. However, it does not explicitly take the relationship between thermal gains and space heating into account, which may help explain the poorer estimates of non-electrical energy use.

Small variations in operating hours were found to make little difference to energy use in the health centres. However, in offices, extended occupancy had a larger effect, partly due to the larger ranges used. This is appropriately accounted for in the energy use model and Verco's assessment that the doubling of occupied hours increases the heating energy use by approximately 30% is reasonable.

Glazing did not appear to have a large impact on non-electrical energy use uncertainty, however, the uncertainty may actually be higher than shown here. While values were varied within a reasonable range, both the surveys and energy use model only consider one glazing type for each building. Also, there was limited granularity in the way the energy use model accounted for glazing and assumptions had to be used when the glazing type was not known. Window to wall ratio also had a surprisingly small influence on energy use, especially given the influence of glazing U value (rate of heat loss). This may be attributed to solar gains offsetting some of the additional losses, though this has not been investigated in detail. The result helps justify the decision not to include a window to wall ratio in the energy use model.

The sensitivity analysis has shown that the uncertainty around some of the key parameters had a significant impact on the predicted energy use of a building. While the use of telephone surveys has allowed significantly more data to be collected than a direct metering approach, the nature of this input data means it is very difficult to predict the energy use of an individual building, regardless of the model used. However, we can derive more confidence in the aggregate sub-sector level analysis.

5 Conclusions

The Building Energy Efficiency Survey's (BEES) energy use model generally appeared to be appropriately specified and fit for purpose, given the available input data. Although SimStock performed slightly better overall, performance between the two models was close, increasing confidence that the energy use model was making the best possible use of the underlying survey data. In particular, the models had reasonable performance at the aggregate sub-sector level and reproduced the distribution of the known energy use fairly well. This showed that, although based on limited input data, the energy use model was well suited for making predictions and recommendations at a stock level. However, care should be taken interpreting the results for individual buildings.

The energy use model produced reasonable estimates of electricity use. Furthermore, performance was consistent across different building datasets indicating the methodology was robust. The relatively small number of parameters required to accurately calculate electricity use also helped increase confidence in this aspect of model performance, and the model appeared to be making good use of the telephone survey data to achieve this.

The energy use model also produced reasonable estimates of non-electrical energy. However, there was less consistency between the two office model runs indicating greater sensitivity to input data. The better performance of the SimStock model in the first office model run suggested explicit modelling of the built form of a building may be beneficial. However, the reason for this may also be attributed to other limitations in the energy use model's treatment of non-electrical energy use. For example, due to limitations in data collection through telephone surveys, the energy use model has no way to represent surface area to volume ratio (a measure of compactness of built form) which is known to affect the rate of heat loss from a building.

While space heating in the energy use model seemed to be appropriately specified given the underlying data, limitations in the data itself can result in significant uncertainty around outputs for individual buildings. Further investigation into the extent to which different types of built form modelling improve accuracy would be useful for future iterations of BEES. It may be that the inclusion of one or two key geometric parameters can increase performance without necessitating the additional complexity and data requirements of a full dynamic building simulation model such as EnergyPlus.

Evolutionary algorithms were useful tools for discovering good values for parameters in a model, once a mapping of survey responses to model parameters had been established. While the energy use model is currently calibrated using engineering judgement, it is likely to be very well suited to an evolutionary approach given the short model run time. However, care would need to be taken to ensure parameters resulted in robust results and sensible end use intensities across models trained with different samples.

The SimStock model needed a relatively large sample size in order for the genetic optimisation to be effective. The small sample size available for health centres severely limited the ability of the SimStock model to replicate and validate the energy use model results for that building type. While this approach could be easily and usefully extended to other sectors, the consideration of sample size must be taken into account. Careful selection of the building records in the training data also appeared important for model performance.

Finally, the accuracy of predictions of end uses is inherently uncertain since the surveyed buildings did not have metered data broken down by end use to compare with. It would be of value to test the model against a sample of buildings for which sub-metered electricity use data is available in order to inform the level of confidence to place in these disaggregated energy use predictions.

6 References

Building Research Establishment, 2015. NCM Databases for iSBEM_v5.2.g. Available at: http://www.uk-ncm.org.uk/download.jsp [Accessed June 4, 2015].

Coffey, B. et al., 2015. An Epidemiological Approach to Simulation-Based Analysis of Large Building Stocks. In *Proc. of IBPSA International*. Hyderabad.

DECC, 2014. The Non-Domestic National Energy Efficiency Data Framework (ND-NEED). Available at: https://www.gov.uk/government/statistics/the-non-domestic-national-energy-efficiency-data-framework-nd-need.

Elsayed, M., Grant, J.F. & Mortimer, N.D., 2002. *Energy use in the United Kingdom non-domestic building stock*, Sheffield.

Evans, S., Liddiard, R. & Steadman, P., 2014. A 3D geometrical model of the non-domestic building stock of England and Wales. *Building Simulation and Optimization 2014 Conference*. Available at: http://www.bso14.org/papers.html.

Liddiard, R., 2012. Characterising Space Use and Electricity Consumption in Non-domestic Buildings. , (February 2015), pp.37–41. Available at: http://hdl.handle.net/2086/6105.

Mortimer, N.D. et al., 1999. Developing a database of energy use in the UK non-domestic building stock. *Energy Policy*, 27(8), pp.451–468. Available at: http://www.sciencedirect.com/science/article/pii/S0301421599000440.

Mortimer, N.D., Ashley, A. & Rix, J.H.R., 2000. Detailed energy surveys of nondomestic buildings. *Environment and Planning B: Planning and Design*, 27(1), pp.25 – 32. Available at: http://www.envplan.com/abstract.cgi?id=b2572.

Mortimer, N.D., Elsayed, M.A. & Grant, J.F., 2000. Patterns of energy use in nondomestic buildings. *Environment and Planning B: Planning and Design*, 27(5), pp.709 – 720. Available at: http://www.envplan.com/abstract.cgi?id=b2573.

Ordnance Survey, 2015a. Address Base. Available at: www.ordnancesurvey.co.uk/business-and-government/products/addressbase-products.html [Accessed April 1, 2015].

Ordnance Survey, 2015b. MasterMap. Available at: www.ordnancesurvey.co.uk/business-and-government/products/mastermap-products.html [Accessed April 1, 2015].

Penman, J.M., 2000. A database and model of energy use in the nondomestic building stock of England and Wales. *Environment and Planning B: Planning and Design*, 27(1), pp.1–2.

Valuation Office Agency, 2015. Business rates. Available at: www.gov.uk/introduction-to-business-rates [Accessed April 1, 2015].

Department for Business, Energy & Industrial Strategy
3 Whitehall Place, London SW1A 2AW
www.gov.uk/beis