Climate change energy impact assessment: Information Commons

Jonathan Sykes
September 22, 2017
About the Author

I am a PhD student at the University of Sheffield, in both the School of Architecture and the School of Mathematics and Statistics. I am funded by the Grantham Centre for Sustainable Futures, here at the University of Sheffield, a joint project between the University of Sheffield and the Grantham Foundation. Its mission is to develop research for a sustainable future and push integration of this message into policy and industry.

My personal work focuses on bringing ways of using data into building design and management. This means developing statistical models capable of using the vast amount of data generated by modern simulations and smart metering.

Not only is climate change creating new challenges for our cities and buildings, but simulations and metering are making us more aware of existing problems. Many UK cities have been found to have high levels of pollutants breaching EU and UK air quality directives.

To tackle these challenges we need to harness the power of modern machine learning and data analytical methods to deal with datasets with 1,000,000s of entries that need such techniques to automatically extract patterns for inference and predictions.

The work presented here is a simple application of only a few hundred data points, but allows us to consider the likely impact of climate on a current asset.
Key Points

1. There are indications that the chiller energy use may increase in the future due to changing climate. See section 1.

2. Current energy use contradicts the fact we are within the baseline 1961-1990 scenario provided by the UK met-office. This supports the hypothesis we are experiencing climate change. See detail in section 1.

3. A data driven approach can be used to make conclusive forecasts of building energy use behaviour. For details of the approach see section 4.

Introduction

The purpose of this report is to demonstrate some of the potential and practicalities of data-based methods.

The Information Commons at the University of Sheffield is a joint venture between Corporate Information and Computing Services (CiCS) and the University Library. Delivering high quality IT-enabled study spaces and 24 hour access to student resources, the IC provides a platform for developing innovative learning and teaching techniques.

The Information Commons is designed to have a low environmental impact for a building of its size and function. Key design considerations were:

- Energy efficient conditional air module (CAM) climate control.
- High performance thermal insulation.
- Intelligent lighting that reduces lighting levels in areas where no students are working, and shelf lighting designed to switch on and off automatically.
- Participation in the Veolia district energy scheme for heating energy.

The Information Commons is a high thermal mass building with many technologies designed to reduce energy usage. This class of buildings attracts heavy investment, therefore it is of interest to determine the ability to predict future energy use, and its connection to climate and uncertainty associated with this.

With special thanks to: my supervisors Dr Chengzhi Peng, Dr Richard Wilkinson and Prof Helen Kennedy; Weston Park Museum and the Curator of Natural Science, Alister McLean; University of Sheffield Estates and Peter Webber, Energy Assistant; Eva Vichova in Grantham Centre for graphics.
1 Climate Change Impact

Figure 1: Graph of forecast chiller energy use by year. On the first plot, horizontal line, in red, is historical energy use, black plots: represent Met-office climate change scenario, grey plots: Met office pre-climate change (control) scenario. On the second plot, represent the uncertainty for each year.

This study is based on a machine learning model of the interactions between the Weston Park weather station data and the combined energy use of Chillers 1 and 2. This model was used to estimate the behaviour of the chillers on forecast weather days from the Met-office from 2021 to 2050.

- Central estimates of the energy use for years 2021 to 2050 are greater than current energy use taking into account climate change. **This leads to the conclusion of increased energy use from climate change.**

- Compared to the variation in estimates there is little difference in use from 05/2015 to 05/2016 compared to 05/2016 to 05/2017.

- The 95% intervals represent the uncertainty of:
  - lack of knowledge about the chiller’s behaviour or inherent randomness. (This is represented by the red portion of the decomposition of variances for each year.)
  - prediction of the future climate. (The blue portion of the decomposition of the variances.)

- A significant part of the difference between the different years is accounted for by inter-annual variations in the Met-office weather data. This could be derived from modelled effect such as solar cycles or El Nino effects.

- The grey bars in the first plot represent the Met-office ‘control’ or baseline scenario on the period 1961-1990. The usage that is predicted for climate for the ‘control’ period (grey error bars) is significantly less than the two years of of data we have for actual consumption. This is evidence that we are currently experiencing climate change.
2 Energy Use

This section aims to provide basic information as to the relative power usage in the different systems in the Information Commons.

![Pie charts showing usage of different metered areas in metered data. Other includes: CIS Alternate Supply, Fire Fighting Lift, Generator Incoming Meter, Smoke Extractor 100A](image)

Figure 2: Pie charts showing usage of different metered areas in metered data. Other includes: CIS Alternate Supply, Fire Fighting Lift, Generator Incoming Meter, Smoke Extractor 100A

2.1 Relative Proportions Energy Use

Figure 2 shows the relative proportions of energy usage in the Information Commons. Heating is not included due to being connected to the Veolia District Energy Network (a district heating system).

- In terms of total volume used the largest two systems of electrical energy usage, annually, are the Riser Stage 2 and Roof Top Heating and Ventilation Air Circulation (Roof Top HVAC) systems. In the two time periods considered the two comprise over 60% of power usage, each at least 30% of total usage.

- The Chillers 1 & 2 together comprise around 20% of total usage. After that the Cafeteria has significant usage at around 11-12% of usage. Lifts and other systems together account for less than 2% of total usage.
2.2 Variable Relationships

The different metered systems have distinct behaviours. Figures 3 and 4 show the usage of energy for different systems against date and daily maximum temperatures. Heating energy values are not directly comparable as they are district heating demand values.

Better results will be obtained if there is an apparent relationship between variables. Machine learning models are capable of learning behaviour that has no physical significance. A good guard against this is trying to learn patterns that are already known.

This would be exemplified by trying to learn the behaviour of Chiller No 1 in relation to date, but it is less clear what would be learnt in relation to the Cafeteria and date.

- Figure 3 shows there is significant structural variation in the behaviour of the Busbar Riser: Stage 2, Cafeteria, Chiller No 1, DB CIS Alternate Supply, both lifts and the Roof Top HVAC against time. Suggesting that there may be behaviour to learn.

- The Generator Incoming Meter (a standby generator) and the Fire Fighting Lift both show no usage at all so are of no interest to this analysis.

- Both Chiller No 2 and the Smoke Extractor show less clear behaviour. However as show in figure 4 there is a link for Chiller No 2 between energy use and maximum daily temperature.
• The clearest correlation between energy use and maximum temperature is with Chiller No 1, with a positive correlation (higher energy use relates to higher temperatures).

• Chiller no 2 is active significantly less of the time, which means, any trends are less obvious. There does appear to be some positive correlation between daily usage against higher maximum temperature.

• Heating energy used has a negative correlation with maximum daily temperature (high heating energy use relates to low max temperature).

• Of the other variables only Roof Top HVAC system shows a very weak positive correlation. (The smoke extractor runs at very low levels of consumption as such is removed due to being low priority.)

• As such the variables selected for modelling are Chillers No 1 & 2. Expanding this work to include heating energy would be a logical next step.
3 Evidence of Model Performance

The machine learning model trained has no knowledge of ‘real world’ physics. As such the model has possibilities of producing unrealistic estimates. This section aims to provide evidence of model performance.

3.1 Daily Predictions

Figure 5: Comparison of Training Data (05/2015 to 04/2017) to 2050 projections by Chiller 1 and 2 combined energy use against day of the year.

Figure 5 shows the metered data used to train the model and the predictions for 2050. Whilst direct comparison is not appropriate, some broad aspect are directly comparable.

- The median values (blue) do follow the annual trend and the summer period is clearly identifiable. The inactive winter period for chilling is also highly identifiable.

- The spread of values for the summer period is similar. The winter period has some parts of the 95% interval that do stray into negative regions. Whilst not ideal, there is evidence that there are so few values, they are not of significant concern.

The decision to combine Chillers No 1 and 2 was made on the basis that the modelling of each independently was a poorer model. Figures 6 and 7 show the results from modelling each individually.

- Chiller 1 (figure 6) is reasonably well modelled. There is more uncertainty and more negative values compared to the combined model.

- Chiller 2 (figure 7) has the poorest modelling of Chillers 1, 2 and the pair combined. This is shown by the fact the median values are nearly always near 0. This is a side effect of of how few non near zero values there are in the data.
Figure 6: Comparison of Training Data (05/2015 to 04/2017) to 2050 projections by Chiller 1 energy use against day of the year.

Figure 7: Comparison of Training Data (05/2015 to 04/2017) to 2050 projections by Chiller 2 energy use against day of the year.
3.2 Peak Values

Figure 8: Predicted peak days for forecast time and value period with training data peaks

The maximum chiller energy used gives us an idea as to the likely maximum daily usage which raise concern. As can be seen in figure 8 values exceeding 2500 kWh in one day would be an event that could warrant investigation.

The days of the year, when a peak is likely to occur, is clearly centred around the summer period. A high value would be unlikely in the winter period.

This sort of modelling can be used in automated building management systems. Clearly particular operational considerations are paramount. The number of peak events to be considered for investigation will be evaluated against available resources and risk appetite.

The peak values in the training data appear consistent with the modelled peak distributions. The peaks in the days of the year for the top 1 values shouldn’t be used to draw particular conclusions.
3.3 Energy Use Proportions

Figure 9: Plot of marginal distribution Chillers 1 and 2 combined energy use for training data (05/2015 to 04/2017) with predicted data for 2050.

The model aims to correctly apportion the when the chillers are actively cooling or inactive, such as in the winter months. Figure 9 shows the relative proportions of the different values in the training data and the predicted values for 2050.

Comparatively the two curves show:

- The modelled data is much smoother than the training data. This is demonstrative that the projected values are essentially averaged over many different weather samples (100 possible weather years), while the training data only covers the 2 years of real world weather data.

- The spikes near 0 correspond to the chillers being inactive on a cold day. The fit is reasonable considering the different weather years each data set corresponds to.

The plot shows a broadly good fit. It does show the predicted values can be negative, however it is a very small proportion. (The area under the curve represents the proportion of values in the corresponding range energy use values.)
4 Modelling Process

The modelling process is laid out in broad details below. In this case we are attempting to teaching
the learning machine the behaviour of the Information Commons in response to the weather data we
provide.

The data used is:

- Quarter-hourly meter readings from the Information Commons electrical energy used for each
  system analysed (not all systems where analysed).
- Half-hourly meter readings from Veolia district energy for heating.
- Daily Weather data recorded at the Weston Park weather station.
- UKCP09 Met-office: Weather Generator Simulations - Medium emission scenario, 2020-2049,
  Standard Weather Generator Variables. 100 samples.

4.1 Data Formatting

The data was formatted as follows:

- Quarter-hourly electrical energy used meter readings:
  - Reject any days that do not have 96 readings.
  - Reject any days for which there is no Weston Park weather data.
  - For remaining days calculate difference between first reading of that day and first reading
    of the next day.
  - The individual systems energy use are discussed in more detail in section 2.2.

- Half-hourly heating energy meter readings:
  - Reject any days that do not have 48 readings.
  - Reject any days for which there is no Weston Park weather data.
  - For remaining days calculate difference between first reading that day and first reading
    next day.

- Weston park weather station data:
  - The minimum and maximum daily temperatures, relative humidity and sunshine hours
    were extracted as weather variables.

- UKCP09 Data:
  - The minimum and maximum daily temperatures, relative humidity and sunshine hours
    were extracted as weather variables.
4.2 Modelling and Prediction

We are modelling the link between weather variables and energy use as follows:

\[ F(w) = e \]  

(1)

Where \( F \) is the function describing the link between the weather variable inputs, \( w \), and the energy use outputs, \( e \).

The modelling assumes independence between energy using systems, as such the model has separate internal models for each system modelled.

The model, \( F \), comprises several stages, evaluated separately for each system:

1. Clustering - Deciding how to place \( w \) values into different behavioural groups.
   
   (a) Decide how to cluster training data. Each training point is assigned to a class and the classes are designated either to be modelled as active or inactive.
   
   (b) Assign each training point its class.
   
   (c) Train a probabilistic classifier to classify new \( w \) into classes.

2. Modelling - Create models of each class
   
   (a) Train active class model (in this case only one active class is permitted) - Training Gaussian process with data based linear prior without nugget.
   
   (b) Train inactive class models - fit a univariate normal distribution, not dependant on \( w \), to the class’s values of \( e \).

3. Store the above trained models for each system (\( e \)).

The process used to make predictions using each system’s model is as follows:

- For each year of the period 2021-2050:

  1. For each year’s 100 weather samples:
     
        (a) Draw 2000 samples from each class predicted for each day for the whole year using forecast \( w \) values.
        
        (b) Calculate the predicted probabilities from classifier model for each day.
        
        (c) Draw the proportion of the 2000 samples from each class according to the predicted probabilities for each day.
        
        (d) The 2000 mixed class samples from previous step are used as predictions for this weather sample for each day.

  2. Group and save the year’s predictions. This gives 100 weather year samples of 365 days with 2000 samples (\( 100 \times 365 \times 2000 \)).

- Total predictions are 30 years of sample groups for each system predicted (\( 30 \times 100 \times 365 \times 2000 \)).