University College London (UCL)

QUANTIFYING AND MITIGATING DIFFERENCES BETWEEN PREDICTED AND MEASURED ENERGY USE IN BUILDINGS

Chris van Dronkelaar

2018

Dissertation submitted in fulfilment for the degree of
Engineering Doctorate in Virtual Environments, Interaction and Visualisation

In collaboration with:
UCL the Bartlett School of Environment, Energy and Resources
UCL Computer Science, Virtual Environments, Interaction and Visualisation
BuroHappold Engineering

Declaration of Authorship

I, Chris van Dronkelaar confirm that this thesis entitled 'Quantifying and mitigating differences between predicted and measured energy use in buildings' and the work presented in it is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Simulation is commonly utilised as a best practice approach to assess building performance in the building industry, and can help facility managers and engineers identify energy saving potentials, forecast future scenarios and evaluate the energy and cost performance of energy saving measures. However, the built environment is complex and influenced by a large number of independent and interdependent variables, making it difficult to achieve an accurate representation of real-world building energy in-use. This gives rise to significant discrepancies between simulation results and actual measured energy consumption of real buildings, termed 'the performance gap'. This is partly fuelled by a lack of understanding of the procedural differences between national calculation methodologies and energy certificates commonly employed in presenting energy use. As such, a classification was adhered to, which distinguishes between three different performance gaps; the regulatory gap (predictions from compliance modelling), static gap (predictions based on performance modelling), and dynamic gap (calibrated predictions taking a longitudinal perspective). This research added to knowledge by making three separate contributions. The first contribution was the exploration of industry practices and stakeholders, which identified common barriers to delivering high building performance, and made suggestions on how to overcome such barriers. Through semi-structured interviews and round-table discussions with industry experts, five key factors were suggested for delivering better building performance. The second and third contributions emerged from case research, for which an overarching methodology was developed, aiming to quantify and mitigate differences between predicted and measured energy use. Fundamental tasks within the methodology were based upon previous research efforts, while new techniques were introduced to include the uncertainty of typically static input parameters to improve the calibration process. In particular, the second contribution was the quantification of the underlying causes of the performance gap and mitigation of differences between predicted and measured energy use in four case study buildings, through the application of sensitivity, and uncertainty analysis and manual calibration. Subsequently, the third contribution investigated the effect of data granularity on model calibration accuracy through meta-model based optimisation.

Impact statement

Previous research on the energy performance gap has been reviewed, and was published in an academic journal paper. It presented a classification of different performance gaps, analysed the magnitude according to the regulatory gap, highlighted the underlying causes for a performance gap, and proposed measures to overcome these issues in order to mitigate the performance gap. Furthermore, it identified several research gaps that have led to three contributions to knowledge in this thesis. Contributions to knowledge fall under the category of *"empirical work which has not been done before, developing and explaining a new synthesis of empirical observations and/or theoretical argument*s" (Francis, 1976; Phillips & Pugh, 1994). These contributions are:

1. **Identification of common barriers to delivering reliable building performance, with key recommendations on how such barriers can be overcome.**

This research has contributed to a report that explored industry perspectives on how reliable building performance can be delivered by adhering to several identified key principles. It discussed different perceptions of building performance amongst building industry stakeholders and how these relate to each other, identified common barriers in the supply chain that prevent performance from being delivered, and established several key principles that need to be followed to overcome these barriers. This work developed a practical guideline that outlines several important principles that need to be taken into account during building design, construction and operation. The report was published by the UK Green Building Council in May 2016, named "Delivering Building Performance", and was followed by several promotional debates and talks on the topic to communicate this to the wider construction industry.

2. **Mitigation of differences between predicted and measured energy use and quantification of the impact of underlying causes on the regulatory performance gap.**

Case research highlighted common issues in data collection and performance predictions in four existing buildings. Building performance modelling was utilised to create representable models of the existing building operations, using uncertainty and sensitivity analysis to mitigate differences between predictions and measurements. Calibrated models were used to explore the impact of typical assumptions on the regulatory performance gap. This contribution was supported by developing a calibration methodology that compared predicted and measured energy use in existing buildings, for the application of both manual and automated calibration to mitigate their discrepancies. It introduced several parametric techniques that improved the calibration process over previous research.

3. **Quantification of the effects of data granularity on model calibration accuracy using meta-model based optimisation.**

Building on previous research in the area of model calibration, the research sought out to understand how the granularity of data affects model accuracy. Such knowledge is useful as it determines the relationship between model accuracy and need for quantitative data and information, establishing a trade-off between time consumption, cost and accuracy. Subsequent to the second contribution, which used manual calibration processes, meta-models were constructed based on the relationships between inputs and outputs created through parametric simulation. Higher levels of accuracy were achieved using these meta-models for optimisation (i.e. automated calibration), but were found to introduce additional complexity.

Contents

Acknowledgements

Thanks to all that made it possible for me to take on the pursuit of obtaining a doctorate (especially my parents) and those that helped me along the way. In particular, I like to commend several people on their efforts. Professor Dejan Mumovic, for continuous support throughout the engineering doctorate (EngD), expeditious and uncomplicated, albeit busy. Dr. Catalina Spataru for her valuable feedback throughout the years. Dr. Mark Dowson, industrial supervisor at BuroHappold Engineering who guided me significantly throughout the process, keen, capable and straightforward. Duncan Price, who supported the EngD in its second half and who recognised an opportunity for me to get involved with the UK Green Building Council (UKGBC) during my doctorate. The UKGBC and its employees for allowing me to support their research, in particular I like to commend Lynne Ceeney for her valuable feedback and easy collaboration during the process, and of course, those involved in the interviews and round-table discussions, their knowledge was fundamental to the delivering building performance report. Dr. Esfand Burman for assisting me in the data collection process, and conversations on the performance gap. Dr. Ivan Korolija for helping me with the difficult bits in building energy simulation and diagnosing my models. Special thanks to Efstathios Eleftheriadis, Yair Schwartz and Christodoulos Aspromallis, a.k.a. EngD legends, whom started the EngD in Virtual Environments, Interaction and Visualisation concurrently, in pursuit of knowledge, but ended up in conversation in the pub. The doctoral students of BSEER, with whom I had the pleasure of having many lunches with and other get-togethers. The Sustainability & Physics team at BuroHappold Engineering, a clever bunch of people trying to make the world a better place, bit by bit. Friends, those whom I have met along the way and provided me with the necessary distractions, many thanks! Furthermore, I would like to thank UCL Computer Science, Centre for Virtual Environments, Interaction and Visualisation and UCL the Bartlett, Institute for Environmental Design and Engineering (IEDE), both of which I felt part of. UCL Estates, in particular Joanna Marshall-Cook, Chris Marshall and John Ilott, who provided me with invaluable data and information on two of the case study buildings. UCL IT and security, who provided me with (anonymised) Wi-Fi and swipe card access data. I would also like to acknowledge the use of UCL Legion High Performance Computing Facility (Legion@UCL) for running thousands of building simulations. The Engineering and Physical Sciences Research Council (EPSRC) and BuroHappold Engineering funded this research programme.

Acronyms

AHU	Air handling unit
AMR	Automatic Meter Reading
ANOVA	Analysis of Variance
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BMS	Building management system
CIBSE	Chartered Institution of Building Services Engineers
COP	Coefficient of Performance
CTK	Contribution to knowledge
CV(RMSE)	Coefficient of Variation of Root Mean Square Error
DB	Distribution board
DEC	Display Energy Certificate
ECM	Energy Conservation Measure
ESOS	Energy Savings Opportunity Scheme
EPBD	Energy Performance of Buildings Directive
EPC	Energy Performance Certificate
EUI	Energy use intensity
FCU	Fan coil unit
FM	Facilities management
GA	Genetic algorithm
GRESB	Global Real Estate Sustainability Benchmark
HVAC	Heating, Ventilation and Air Conditioning
IDF	Input data file
IoT	Internet of Things
IPMVP	International Performance Measurement and Verification Protocol
LV	Low voltage
LED	Light-Emitting Diode
LHS	Latin Hypercube Sampling
LPHW	Low Pressure Hot Water
L&P	Lighting and power
LR	Linear regression
MAE	Mean absolute error
MEES	Minimum Energy Efficiency Standards
MSE	Mean squared error
NCM	National Calculation Methodology
NMBE	Normalised mean bias error
NSGA	Non-Dominated Sorting Genetic Algorithm
OFAT	One-factor-at-a-time
O&M	Operation & Maintenance Manual
PCC	Pearson correlation coefficient
PV	Photovoltaic
POE	Post Occupancy Evaluation
REIT	Real Estate Investment Trust
RFP	Request for Proposal
SP	Setpoint
SRCC	Spearman rank correlation coefficient
SRC	Standard regression coefficient
SVM	Support Vector Machine
VRV/VRF	Variable refrigerant volume/flow

1 INTRODUCTION

This chapter sets the context for the research, introducing the energy performance gap between predicted and measured energy use buildings, then presenting the aims and objectives of this thesis and the research approach taken to achieve them. The energy performance of the vast majority of buildings in the United Kingdom (UK) do not align with the challenge of meeting carbon targets (UKGBC, 2016). Binding targets set out in the Climate Change Act 2008, require the UK to ensure that the net carbon account for the year 2050 is at least 80% lower than the 1990 baseline (HM Goverment, 2008). To achieve such targets, the operational energy consumption of new buildings and existing buildings should be minimised. Building Regulations, Energy Performance Certificates and other sustainability certification schemes offer a form of quality assurance, but generally only focus on design elements. It is then not surprising, that measured energy use in buildings during occupation frequently shows major differences with respect to predicted energy performance. This phenomenon has previously been termed 'the performance gap' (Cohen, et al., 2001; Menezes, et al., 2012; Carbon Trust, 2012; Burman, et al., 2014; de Wilde, 2014). The discrepancy in predicated and actual performance undermines the credibility of building designers, planners and consultants. Large corporations, universities, building owners and occupiers also suffer from higher than anticipated utility bills and rising energy costs. To understand the performance gap in more detail, building energy modelling was utilised to compare predictions of existing buildings with measured building performance. The use of calibration techniques can quantify and minimise differences and help to better understand underlying causes of a discrepancy. However, due to the inherent complexity of building energy models and the myriad parameters that predictions are based on, many of these parameters cannot be accurately defined. As such, many potential solutions exist and more accurate solutions can be masked by inaccuracies in the model and level of detail of data available. To alleviate this issue, performance modelling predictions were compared and calibrated with measurements at a higher level of temporal, spatial, and hierarchical data granularity, to understand how this affects model calibration accuracy.

1.1 Background

Designed performance is generally determined through *compliance modelling* [1]. In the UK this is currently implemented by the use of simplified (steady-state) or dynamic thermal modelling, to calculate the energy performance of a building under standardised operating conditions (e.g. occupant density, setpoints, operating schedules, etc.), set out in the National Calculation Methodology (NCM, 2013), for the Energy Performance of Buildings Directive (EPBD). Compliance modelling is useful to assess the energy efficiency of buildings under standardised conditions to determine if minimum performance requirements are met, by comparing with a typical notional building. However, because designed performance is typically determined for regulatory purposes, based on standard operating conditions, it does not offer a like-for-like comparison with the performance of the completed building. This results in a deviation between regulatory predictions and measured energy use, which creates a significant risk to designing and operating low energy buildings. Conflating compliance modelling with measured energy use is one of the reasons for the popularisation of the 'perceived' energy performance gap (i.e. regulatory performance gap). For example, in England and Wales, the Energy Performance Certificate (EPC) rating is used to indicate the energy efficiency potential of a building (not intended to represent operational energy use), whilst the Display Energy Certificate (DEC) provides a performance rating based on measured energy use of the building. Hogg & Botten (2012) analysed 128 office buildings and highlighted EPC ratings as a poor indicator of actual energy use, due to unaccounted complexities and variations that affect energy use. EPCs are asset ratings, aimed to compare buildings to one another in order to assess their efficiency. They are calculated based on standard weather conditions and building use.

Theoretically, a gap is significantly reduced if predictions are based on actual operating conditions, also known as *performance modelling*. Performance modelling includes all energy quantification methods which aim to accurately predict the performance of a building. Thus, there is a need for design stage calculation methodologies to address all aspects of building energy consumption for whole building simulation (Norford, et al., 1994; Torcellini, et al., 2006; Diamond, et al., 2006; Turner & Frankel, 2008). However, the built environment is complex and influenced by a large number of independent and interdependent variables (Coakley, et al., 2011), making it difficult to achieve an accurate representation of real-world building energy in-use. Although a margin of error between any type of prediction and measured energy use is inevitable due to uncertainties in design, quality of construction and limitations of measurement systems, investigation of predicted and measured energy use is necessary in order to understand the significance of the underlying causes of the energy performance gap.

Presently diagnosis techniques can identify performance issues in operation; trend analysis, energy audits and traditional commissioning of systems can highlight poor performing processes in a building. An integrated approach is the calibration of virtual models to measured energy use. Calibration can pinpoint differences between how a building was designed to perform and how it is actually functioning (Norford, et al., 1994). A concern with calibration models is that they can mask modelling inaccuracies when focusing on the building or system level (Clarke, 2001). Raftery et al. (2011) showed that even the most stringent monthly acceptance criteria do not adequately capture the accuracy of the model with measured

[1] Compliance modelling refers here to as-built performance (e.g. EPC) including any changes after design (i.e. value engineering), different to as-designed compliance.

data on an hourly level. Furthermore, the reliability and accuracy of calibrated models depends on the quality of the measured data used to create the model, as well as the accuracy and limitations of the tools used to simulate the building and its systems (Coakley, et al., 2012). Therefore, it is essential to calibrate energy use at a high level of data granularity to increase model accuracy. By employing calibration using hourly data and component level information, dynamic relationships between building processes are maintained, this essentially reduces the solution space for the calibrated model to comply with. An accurately calibrated model is more reliable in assessing the impact of Energy Conservation Measures (ECMs) and their feasibility by forecasting its potential savings on energy use from implementation.

1.2 Aim and objectives

This research aimed to **quantify and mitigate the energy performance gap and its underlying causes in four buildings.** It demonstrates how a discrepancy can be mitigated by: utilising operational data at a high level of data granularity to inform building performance simulation, improve models through calibration, and quantify underlying causal factors of the performance gap. The objectives were:

- To explore industry perspectives on how to deliver better building performance in buildings,
- To utilise operational data to inform building performance simulation assumptions,
- To mitigate discrepancies between predicted and measured energy,
- To quantify the impact of the underlying causes of the performance gap,
- To determine the effects of data granularity on building model calibration accuracy.

The aim of this research is similar to some of the other research on the topic of the energy performance gap, but it distinguishes itself in its specific objectives and research approach. In particular, it makes use of calibrated models and meta-model based optimisation to investigate this problem. Previous research which has focused on different aspects of the energy performance gap include: Maile (2010), who developed a comparison methodology to identify performance problems from a comparison of measured and simulated energy performance data; Menezes (2013), who investigated in particular the impact of equipment loads and their effects on the performance gap; Burman (2016), who developed a M&V framework for verifying actual performance in relation to regulatory calculations. These are only a few examples, whereas de Wilde (2014) gives a detailed overview of ongoing efforts to bridge the gap.

1.3 Research approach

This thesis presents both qualitative and quantitative studies into building performance and the performance gap. Within the qualitative study, the author partnered with the UK Green Building Council (UKGBC) and brought together a group of industry experts to discuss and highlight process improvements that design, construction and property developers, as well as occupiers, might adopt to deliver buildings which perform as expected in operation. Interviews and round-table discussions were also conducted to explore perspectives on delivering better building performance in non-domestic buildings. This study viewed building performance simply as how a building functions against its needs, which helped to understand other elements in the building procurement process that also affect the energy performance of a building, and indirectly the energy performance gap.

Together with the literature review, this formed a key input in the design of the methodology for the quantitative study, which investigated the discrepancy between predicted and measured energy use, using model calibration in four existing case study buildings. This approach enabled the development of conceptual models to study the behaviour of interrelated variables in existing buildings. The investigation of multiple case studies showed differences in influential variables. However, it was not possible to generalise any building specific findings over a population (large amount of similar buildings), which was outside of the research scope. Rather, it aimed to establish a methodology for investigating the discrepancy between predicted and measured energy use; introducing new approaches to parametric simulation and calibration techniques at a high level of data granularity, in addition to quantifying the impact of assumptions in building energy modelling. Furthermore, the case study buildings provide a platform for determining the significance of operational data granularity on model calibration accuracy and investigation of potential benefits of automated calibration.

A major task has been the collection of data and acquisition of appropriate case studies and their respective data sources to support the research at a high level of data granularity. Predicted energy performance was established through detailed modelling of the existing buildings and comparison against measured energy performance. The four case study buildings were selected from a sample of more than ten buildings, based on the availability of data and accessibility to these buildings. However, it is recognised that the unsuitable buildings, those not selected, are typical in the existing building stock (i.e. those without comprehensive data). The cost of sensors and meters and rise in popularity of the internet of things (IoT), is starting to provide opportunities to measure and improve datasets on the performance of buildings, systems and the indoor environment. Calibrated models were employed to identify underlying causes for the discrepancy. This was supported by sensitivity and uncertainty analysis to determine the impact of assumptions on the energy performance of the buildings. A workflow of the research approach and contents of this thesis is visualised **Figure 1.1**.

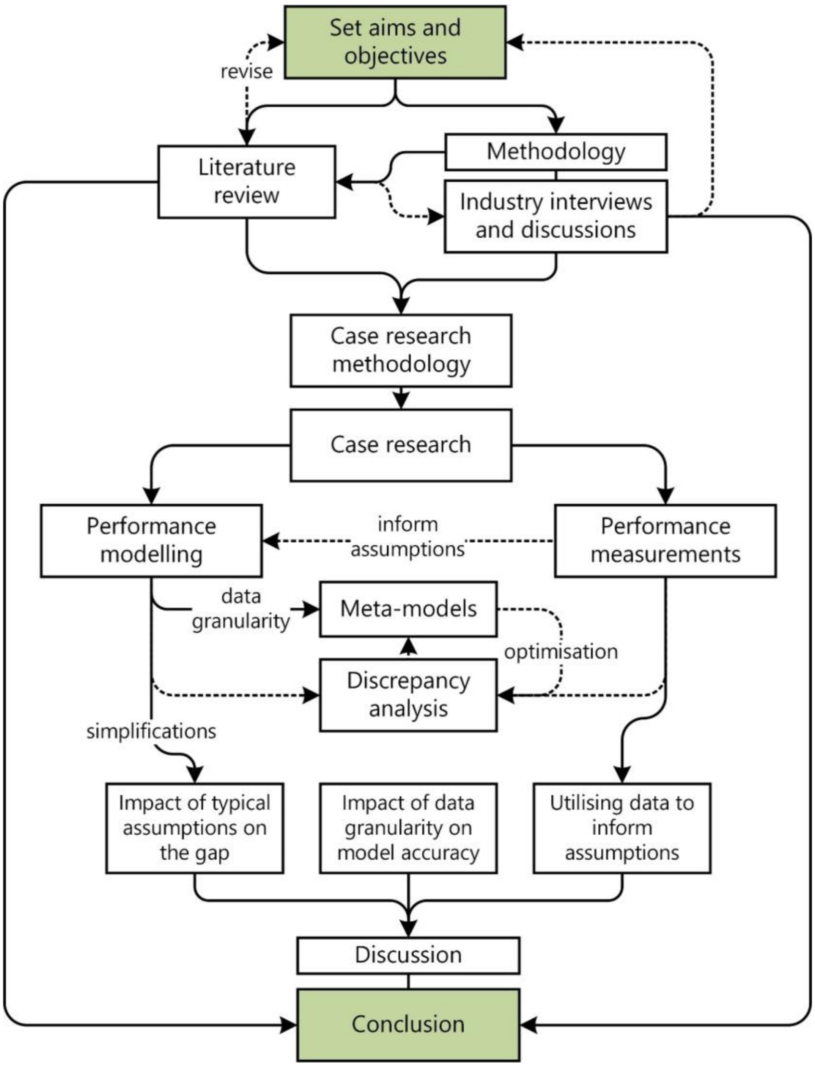


Figure 1.1: Workflow of the research approach and contents of the thesis.

The research used both deductive (theory to confirmation) and inductive (observation to theory) processes. A deductive approach was applied in validating the existing hypothesis that there is an energy performance gap in buildings. Literature findings were compared to observed reality to verify the established theory and the hypothesis was confirmed in the case studies. Limited evidence and understanding of the underlying causes for the energy performance gap did not allow for a well-founded hypothesis, instead an inductive approach was used to develop a hypothesis based on observations.

1.4 Publications

[1] van Dronkelaar, C., Dowson, M., Burman, E., Spataru, C., & Mumovic, D., 2016, A review of the regulatory energy performance gap and its underlying causes in non-domestic buildings. Frontiers in Mechanical Engineering, 1:17.

[2] UK Green Building Council, 2016, Delivering Building Performance (Project researcher within task group at the UKGBC in partnership with BuroHappold Engineering and UCL)

[3] van Dronkelaar, C., Dowson, M., Spataru, C., Mumovic, D., 2018, Utilising meta-model based optimisation to determine the effects of data granularity on building model calibration accuracy. Conference on Building Simulation and Optimization (BSO 2018), Cambridge, UK.

1.5 Thesis overview

Chapter 2 - Literature review

Literature review of the performance gap, its magnitude, underlying causes and approaches for mitigation. Furthermore, it gives an overview of existing methods for comparing predicted and measured performance. Part of this literature review has been published as a journal paper in Frontiers in Mechanical Engineering (van Dronkelaar, et al., 2016).

Chapter 3 - Industry perspective on delivering building performance

Interviews and round-table discussions have identified barriers and correlated key factors and processes for delivering building performance. This work is published as an industry focussed report (UKGBC, 2016).

Chapter 4 - Case research methodology

Describes the developed methodology employed for the case study research. It describes the process of data collection, synthesis, modelling and subsequent calibration and analysis techniques.

Chapter 5 - Utilising operational data to inform building modelling assumptions

Operational data collected on the case study buildings are analysed and subsequently utilised to inform assumptions for building performance modelling.

Chapter 6 - Quantifying the impact of underlying causes of a discrepancy

Results from the measured data collection and modelling are compared according to the established hierarchy. The building models were calibrated towards measured performance, used to quantify the impact of typical assumptions.

Chapter 7 - Quantifying the effects of data granularity on model accuracy

Meta-models were created based on the relationship between inputs and outputs, then utilised to investigate how a higher level of data granularity affects model calibration accuracy.

Chapter 8 - Case research discussion

The results from the case research are discussed, with an emphasis on data collection in supporting model assumptions, quantification and mitigation of the discrepancy between predictions and measurements and the effect of data granularity on model accuracy supported by uncertainty and sensitivity analysis.

Chapter 9 - Conclusion

Concluding remarks on the main findings of the literature review, exploratory study on industry perspectives and the case research. Contributions to knowledge are summarised and additional areas for further work are proposed.

2 LITERATURE REVIEW

This chapter reviews the energy performance gap and its implications on the building construction industry in a UK context with global outlook. It reviews three different types of gaps, the regulatory- (i.e. the perceived gap), static- and dynamic gap. Highlighting the importance of reducing the discrepancy between predictions and measurement. Existing studies were then reviewed to quantify the magnitude between regulatory predictions and measured energy use, it proceeds in describing the underlying causes for the performance gap, existent in all stages of the building life cycle, and identifies that dominant factors are related to specification uncertainty in modelling, occupant behaviour and poor operational practices. Other contributing factors are related to the early design decisions, heuristic uncertainty in modelling and setting an initial energy performance target. Measures and feedback processes are categorised in order to understand how the performance gap can be reduced, indicating the need for energy in-use legislation, insight into design stage models, accessible energy data and expansion of research efforts towards building performance in-use in relation to predicted performance according to regulations and performance modelling purposes. The classifications highlight that different processes can be utilised for predicting and measuring performance. Therefore, energy performance quantification and performance assessment methods were reviewed. Calibration of building energy models and performance modelling are seen as two methods that could support in understanding and mitigating differences between predictions and measurements by better representing the performance of buildings. Several research gaps were identified concerning the calibration of energy models. Furthermore, the literature review emphasised the segmentation of disciplines involved throughout the building life cycle stages and that there are split incentives between stakeholders in regards to the energy efficiency of buildings and the energy performance gap.

2.1 Introduction

According to the International Energy Agency (IEA) the buildings sector is the most significant worldwide energy consumer at around one third of total final energy consumption throughout the period to 2035 and is growing at an average of 1% per year (IEA, 2010). In the UK, buildings accounts for up to 40% of total energy consumption (DECC, 2013). Since the oil embargo in 1973, achieving better energy efficiency has become one of the world's major challenges (Hong, et al., 2000). One of the largest sources of energy and carbon emission savings can be achieved through more energy efficient buildings, a fundamental part to address climate change (Commission, 2011). Regulation of building energy use will have a critical role in meeting energy and emission targets in developed countries (Heo, et al., 2012). In the UK for example there is the CRC (Carbon Reduction Commitment) developed by the DECC (Department of Energy and Climate Change), recently reorganised under the Department for Business, Energy and Industrial Strategy (BEIS), which is set to be closed in 2018-19 and be replaced with an increase in the Climate Change Levy (CCL). Cutting emissions in the public and private sector is necessary to meet requirements of the 2008 Climate Change Act. In London in particular, the London Plan requires a 35% reduction in carbon dioxide emissions for non-domestic buildings compared with UK Building Regulations 2013. However, even though most of these regulations focus on energy and carbon, there has been a trend towards thinking in terms of building performance, which extends to include economic, social and environmental aspects, such as occupant well-being, thermal comfort, operational costs, public image, etc., historically regulation has been mostly about safety. In order to understand how energy performance is influenced and can be reduced, it is important to distinguish building energy performance and building performance, the latter affects the former.

2.2 Building performance

Fragmentation of the UK construction industry is a key influential factor of building performance (Construction Task Force, 1998; Cox & Townsend, 1997; House of Commons, 2008). Interrelations between stakeholders are complex and typically not well integrated in the supply chain, more apparent in some procurement methods (Korkmaz, et al., 2010). Fundamental to this integration are the underlying incentives for procuring a building in the first place, incentives which themselves can be defined as buildings performance. This could mean that delivering building performance for a capital provider is their return on investment and yield, for a designer it is the provision of a safe, resilient and sustainable building as predicted and for facilities management this is the assurance of an operable building in which occupants can be comfortable. Building performance thus describes how well a building functions against stakeholder's needs.

2.2.1 Stakeholders influence on performance

Understanding the interrelationship between stakeholders gives insight into delivering reliable building performance. Although there is a close relationship between some stakeholders, decisions made at the start of a project do not necessarily involve those further down the line, resulting in performance requirements that are not robust enough and potentially missing out on opportunities (House of Commons, 2008). Whereas decisions made during design are disconnected from building performance in operation (Fellows & Liu, 2012). **Figure 2.1** shows the typical communication lines between different stakeholders involved in different stages (this is not exhaustive and other relationships might exist in certain projects, e.g. a property management company may be involved in-use, nor are planners depicted in this diagram).

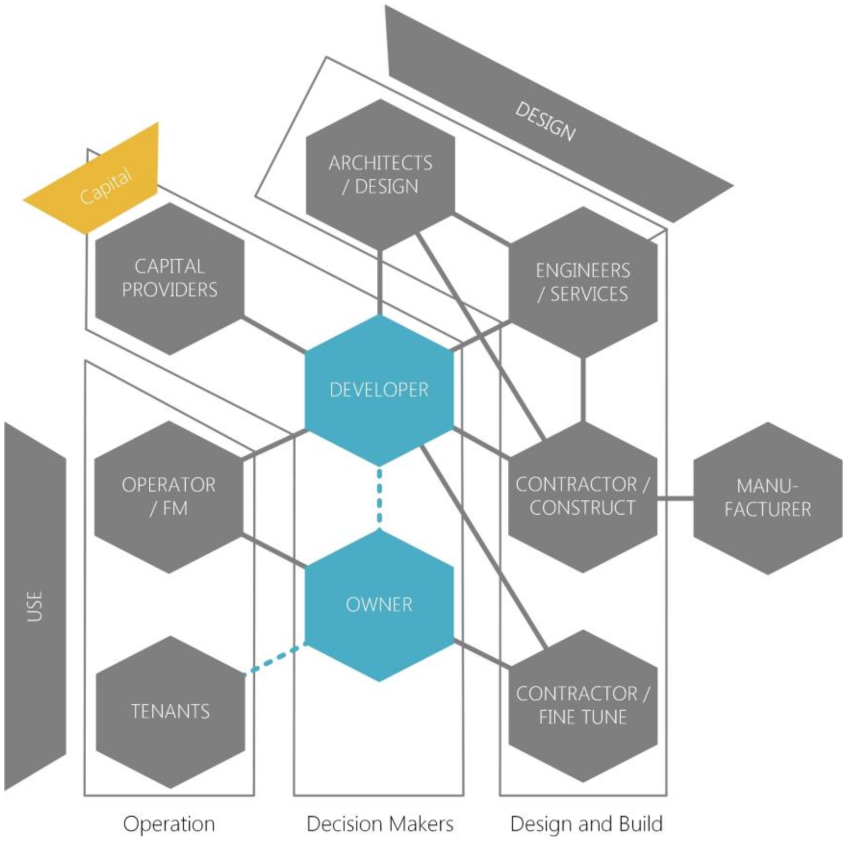


Figure 2.1: Stakeholders involved in different stages and with different connections.

The developer and owner are here depicted as those communicating with most of the parties. There are different models possible, where sometimes the owner and developer are the same party or where for example managing agents are employed by the owner for the operation and maintenance of buildings, sometimes with in-house facilities management.

Decision makers

Project financing is attracted by the developer through a variety of funding structures based on the business case made. Capital providers provide this financing, but typically lack understanding of the economic benefits of sustainable building, for this stakeholder it is a challenge to obtain financial support if there is no visible market value attached to it. They are primarily motivated by maximizing yield or return on investment and the building's market value and any associated risks (Atink, et al., 2014). The increased market value of sustainable buildings needs to be recognized if investors are to provide financing for them. Demonstrating and communicating to capital providers that there is a business case for building performance is necessary in order to induce a market transition towards better buildings. The owner is the most diverse stakeholder, fulfilling different roles the owner can have a large influence on building performance outcomes. Most favourably, the owner is also the operator and/or occupier of a building and sometimes also the capital provider, as they would be incentivised to have a high performing building (Ellison & Sayce, 2007). Public authorities, universities and hospitals often take up such a role. They have a longitudinal focus and have a significant control of the building procurement process through setting specific performance criteria which are often more stringent than the required prescriptive regulatory framework. An owner who takes up a role only as developer with the with the intention of selling the property upon completion is typically less concerned about performance aspects in relation to wellbeing and sustainability of a building, this type of owner is usually interested in enhancing their profit (Lim & Mohamed, 1999). Furthermore, they often procure buildings without information about prospected tenants. Similarly, the owner-landlord, not necessarily being the occupier rents out the property and is interested in with performance aspects that affect renting rates, seeking profitability (Fellows & Liu, 2012).

Design and build team

Building design has a large influence on building performance, from building shape to system controls to renewable technologies, especially decisions during concept development can have a large effect on performance. Integrated design at this stage is therefore encouraged to include important performance aspects in combination with functionality and aesthetics. Designers need to understand how buildings are actually being used, what this means for performance and need mechanisms to translate this to new building designs. Finally, architects and engineers have only limited influence on the building projects in terms of performance outcomes or criteria set by the client i.e. the developer / owner, the design is mostly guided by client's requirements. Tendering of construction works to the developer is mostly awarded based on capital cost, upon receiving work a contractor is primarily concerned with risk and schedule of delivery. They are typically unaware of performance requirements unless these are written as deliverables in contracts. Any additional features or new features related to performance will increase their fee as limited knowledge and experience will mean increased risk for the contractor.

Commissioning and operation

When a building is constructed, it is handed over, at this point contractors come in to deliver the installation and commissioning of building services, done poorly this results in reduced system efficiency and compromising the air tightness and ventilation strategies. Commissioning is an activity that needs to be carried out throughout the use phase of the building, ensuring that systems continue operating as expected. Furthermore, effective personnel and investment in operational management is crucial in achieving a reliable performance in operation (Finch & Zhang, 2013), currently this segment is struggling to keep up with the development of the industry and wide adoption of innovative technologies. Tenants benefit from performance by being provided with a healthy, safe and comfortable environment where they can be productive and feel good. This group is rarely involved during the design of a building, while having a large influence on its performance. Only when the owner is also the tenant will they be involved and have good incentives to make sure building performance is being delivered.

2.2.2 Business case for better performance

The better the performance of a building, the stronger the benefits and business case to deliver performance. The benefits of better building performance were comprehensively reported by the World Green Building Council (WGBC, 2013) and consist of environmental, financial and social aspects.

Financial reasons

High performing buildings positively affect capital and operating expenses, return on investment is higher in both new construction and existing building projects compared to conventional construction (Chegut, et al., 2013). This is due to reduced construction costs, operating costs, and increased in value. Payback improves as energy prices rise, energy efficiency improvements become a better investment and save owners and tenants money, reducing expenses improves profitability (Pless, et al., 2012). It also increases the competitive edge and return on investment for capital providers (O'Mara & Bates, 2012). Recently, more interventions have come into place that link energy or other sustainability related criteria to the right to develop, sell or let a property, directly affecting asset value and valuations (GCB, 2014).

According to the World Green Building Trends 2016 survey, the percentage of companies expecting to have more than 60% of their building projects certified green is anticipated to more than double by 2018 (Jones, et al., 2016). A certified building (BREEAM, LEED etc.) is becoming the norm for most global corporate occupiers. Furthermore, the expectation is that a growing number of occupiers will expect building developers and owners to be able to demonstrate the impact of the building on occupant health and productivity.

In 2015, a report commissioned by the Carbon War Room and GRESB presented findings on the relationship between sustainability investment and financial returns of real estate investment trusts (REITs) (Fuerst, 2015). The study found that a higher sustainability ranking in the annual Global Real Estate Sustainability Benchmark (GRESB) REIT survey correlated to a superior financial performance; in respect of both returns on assets and returns on equity. It also found a significant link between portfolio sustainability indicators and REIT stock market performance and was able to establish that investing in sustainability enhances business performance and lowers risk exposure and volatility.

Mitigate climate change and energy security

The largest part of our energy sources remain CO_2 related, creating grave consequences with rising global temperatures. Decision makers who invest in delivering building performance will reduce the impact on climate change, preserve the quality of human life, improve business performance, and meet governmental regulations. Increasingly stringent government energy efficiency regulations open up a major transformation in processes, technology and human innovation.

Global energy demand is growing faster than current production capacity, resulting in diminished supplies, increased risk of fallouts and increased energy prices (OECD, 2012). Furthermore, lack of access and reliability of energy can disrupt daily life, negatively impact the economy. Buildings need to play an important role in both generating and storing electricity, alleviating grid pressures and supporting the needs of the community by ensuring access and reliability. High performance buildings are a risk mitigation strategy as they reduce energy consumption, as such they can modify the patterns to avoid expensive peak rates and avoid risks associated with volatile energy prices (HEEPI, SUST and Thirdwave, 2008). Through local generation, they can furthermore feedback electricity to the grid.

Social aspects

High-performance buildings increase well-being and productivity through better daylighting, outdoor views, indoor air quality and thermal comfort, which in turn help both attracting and retaining employees (WGBC, 2014). The increased well-being of building occupants result in reduced illness, lower rates absenteeism and lower employee turnover (WGBC, 2016). Not only is a sustainable place to work more likely to attract and retain talent, it is also more likely to get the most from that talent in terms of productivity. In addition, a well performing building that provides a good living space while being resource efficient can strengthen the public image of those involved in the supply chain (HEEPI, SUST and Thirdwave, 2008).

2.3 The energy performance gap

To design and operate more efficient buildings many classification schemes have been established, providing a means to communicate a building's relative energy efficiency and carbon emissions (Wang, et al., 2012). These assessment schemes are related to the energy consumption of a building, and can be quantified using different methods, both during design (e.g. Asset Ratings, Energy Performance Certificates (EPC) and Part L calculations in England and Wales) and operation (e.g. Operational Ratings, Display Energy Certificates (DEC) in England and Wales) of a building. Accredited performance assessment tools, ranging from steady-state calculations to dynamic simulation methods are utilised to predict the energy consumption of a building, to comply with regulated targets using standardised procedures. Both classification schemes (such as the EPC and DEC) and standard calculation procedures have shown significant discrepancies to measured energy use during occupation, which risk not achieving regulated targets. Although a margin of error between any type of prediction and measured energy use is inevitable due to uncertainties in design and operation, as well as limitations of measurements systems, explaining its magnitude and underlying causes are necessary to more confidently forecast and understand energy use in buildings.

2.3.1 Classification of the gap

Building energy modelling is an integral part of today's design process, however research has shown that buildings can use twice the amount of their regulatory energy performance (Norford, et al., 1994; Pegg, et al., 2007). This makes it unlikely that the building industry achieves model-based targets (UKGBC, 2007). The theoretical performance is generally determined through *compliance modelling* (i.e. normative models), which is presently the implementation of thermal modelling to calculate the energy performance of a building under standardised operating conditions (occupant density, setpoints, operating schedules, etc.), set out in national calculation methodologies. Compliance modelling is useful to assess the energy efficiency of buildings under standardised conditions to determine if minimum performance requirements are met. They are used as standards for evaluation, and are non-descriptive in many of their aspects as to be comparable to other similar buildings. As such, its predictions should not be compared to actual operating conditions (Burman, et al., 2014), which apparently occurs in reality. This results in a deviation between regulatory predictions and measured energy use, which creates a significant risk for energy-related issues to go unnoticed, as they are understood as expected differences in operating conditions. Comparing compliance modelling with measured energy use is one of the reason for the popularisation of the perceived energy performance gap. Theoretically, a gap could significantly be reduced if a building is simulated with actual operating conditions, in other words, when care is taken to predict the actual performance of a building, referred to as *performance modelling*. Performance modelling includes all energy quantification methods which aim to accurately predict the performance of a building. The difference between compliance modelling and performance modelling is further illustrated in **Figure 2.2**. In this figure, although the DEC is supposed to represent actual energy performance of the building, the procedure for creating DECs can be either based on bills or measured energy use. However, due to the rating methodology, many DECs of the same building in different years tend to give different results; differences in floor areas, but also significant differences in energy end-uses such as heating, electricity or renewables, even though these have not necessarily changed throughout the years (Hong & Steadman, 2013; Henderson, 2011). This was also evident in the case research presented in this thesis, where the DECs showed significant differences throughout the years.

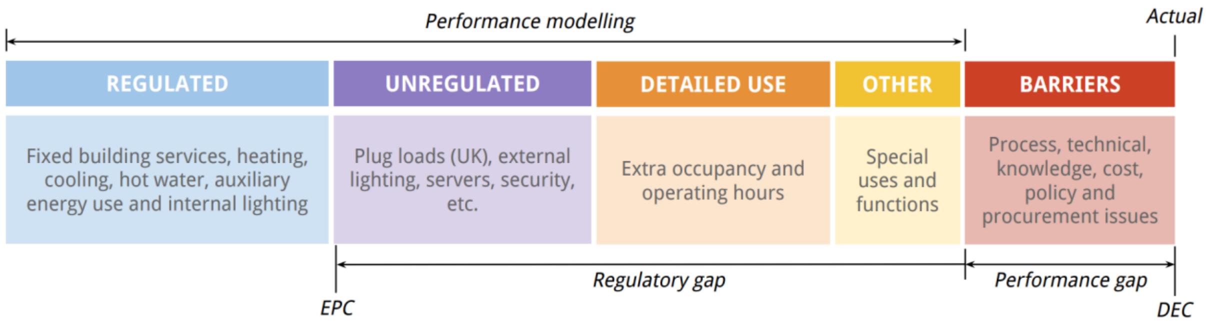


Figure 2.2: Representation of how compliance and performance modelling include different end-uses and level of detail in an energy calculation of a building.

On-going efforts to understand the energy performance gap have utilised calibration techniques to fine-tune a building energy model to actual operating conditions and energy use, ideally over a longer period of time. This method gives insights into the operational inefficiencies of a building and can pinpoint underlying reasons for differences between design estimations and actual use. Subsequently, a calibrated model could reintroduce design assumptions to quantify impacts of any underlying causes and their effect on energy performance. As such, a distinction can be made between three types of modelling efforts, these can be classified in three different ways to interpret the energy performance gap. These are; the gap between compliance modelling and measured energy use, performance modelling and measured energy use and calibration and energy use with a longitudinal perspective (Burman, 2016):

- *Regulatory performance gap*, comparing predictions from compliance modelling to measured energy use.
- *Static performance gap*, comparing predictions from performance modelling to measured energy use.
- *Dynamic performance gap*, utilising calibrated predictions from performance modelling with measured energy use taking a longitudinal perspective to diagnose underlying issues and their impact on the performance gap.

A similar classification is used by de Wilde (2014), but includes machine learning approaches as a separate classification of the energy performance gap. Here, machine learning is seen as an approach that is generally used for performance modelling, but might be used for compliance purposes in the future through analysing energy use data and associate design parameters for a large number of buildings, such as proposed by (Hawkins, et al., 2012). These techniques are under development, but have high potential in analysing the impact of building specifications on energy use, including factors such as operating conditions. Their accuracy is largely dependent on high quality measured data.

2.3.2 Importance of reducing the gap

There is a need for design stage calculation methodologies to address all aspects of building energy consumption for whole building simulation, including regulated and unregulated uses and predictions of actual operation (Norford, et al., 1994; Torcellini, et al., 2006; Diamond, et al., 2006; Turner & Frankel, 2008). Building energy simulation models need to closely represents the actual behaviour of the building under study for them to be used with any degree of confidence (Coakley, et al., 2011). These models contain the design goals, and should therefore be the basis for an assessment to determine whether the completed product complies with the design goals (Maile, et al., 2012). Underperformance in design may soon be met by legal, financial implications (Daly, et al., 2014) and demands for compensation and rectification work (ZCH & NHBC Foundation, 2010), as future regulatory targets may be based around in-use performance. The Minimum Energy Efficiency Standard (MEES) is a good example of this in the UK, which will regulate landlords of any properties rented out in the private sector not to renew existing tenancies or grant new tenancies if a building has less than the minimum Energy Performance Certificate rating (EPC) of E from 2018 (HM Goverment, 2017).

Investigation of predicted and measured energy use is necessary in order to understand the underlying causes of the performance gap. Furthermore, feedback helps improving the quality of future design stage models by identifying common mistaken assumption and by developing best-practice modelling approaches (Raftery, et al., 2011). This also guides the development of simulation tools and identifies areas requiring research (Raftery, et al., 2011), such as uncertainty and sensitivity analysis, parametric modelling, geometry creation and system modelling. Furthermore, it can help policy-makers define performance targets more accurately (HM Government, 2010; ZCH, 2010), which then assist in mitigating climate change. In operation, methodologies that analyse a discrepancy and its related issues can help in understanding how a specific building is operating, highlighting poor-performing and well-performing buildings and identifies areas where action is required. Investigating a discrepancy between design and operation can also support in identifying retrofit options to reduce energy use. In order to more accurately predict energy conservation from a set of proposed retrofit technologies, the simulation model must represent a building as operated (Heo, et al., 2012).

2.4 Magnitude and underlying causes

Different procedures can be used for calculating the energy performance in both the design and operational stage of a building, affecting the magnitude in discrepancy. Wang et al. (2012) provide an extensive overview of energy performance quantification and assessment methods. These methods use mathematical equations to relate physical properties of the building, system and equipment specifications to its external environment. They can help prospective occupiers, building owners, designers and engineers in giving an indication of building energy use, carbon dioxide emissions and operational costs. Furthermore, it allows a better understanding of where and how energy is used in a building and which measures have the greatest impact on energy use (CIBSE, 2013). In operation it can identify energy savings potential and evaluate the energy performance and cost-effectiveness of energy conservation measures to be implemented (Pan, et al., 2007). However, the built environment is complex and influenced by a large number of interdependent variables, making it difficult to represent real-world building energy in-use (Coakley, et al., 2011). Thus, models represent a simplification of reality, therefore, it is necessary to quantify to what degree they are inaccurate before employing them in design, prediction and decision making processes (Manfren, et al., 2013). Comparing measured values to modelled or estimated values for regulatory purposes does therefore not offer a valid comparison and should be avoided whenever possible (Fowler, et al., 2010). This perception is by ASHRAE (2004) as it states in its Energy Standard 90.1 Appendix G —"neither the proposed building performance nor the baseline building performance are predictions of actual energy consumption, due to variations such as occupancy, building operation and maintenance, weather, and the precision of the calculation tool". Indeed, the modelled or baseline performance here refers to compliance modelling which is not a representation of reality. Nevertheless, it is useful to signify the perceived performance gap and to understand how compliance modelling is different to measured energy use. Especially since performance modelling is rarely used to predict the actual energy use of a building.

2.4.1 Magnitude

Table 2.1 gives an overview of reported discrepancies between regulatory energy use predictions and measurements found in literature. An indication of the discrepancy is given by a percentage deviation from the predicted baseline value for a range of different non-domestic buildings. The magnitude of the performance gap is typically reported using percentages, given as an increase or decrease from predicted. Case study buildings are located in different climates and have been predicted through different assessment methods using various simulation software. In this analysis averages from the CarbonBuzz database were used as reported by Ruyssevelt et al. (2014), who analysed 408 buildings. A more in depth analysis is given in (Robertson & Mumovic, 2013). CarbonBuzz (2015), is a platform established in order to benchmark and track energy use in project from design to operation. The type of prediction method used are given in **Table 2.2**.

Table 2.1: Magnitude in discrepancy reported in literature (perceived gap incl. equipment energy use).

	% from predicted (no. buildings)**	Gap average	Source	Method ***	Modelling software
Office	47%		Austin (2013)	B	VisualDOE 3.0
	11%		Bertagnolio (2012)	C	ISO 13790
	-27% to 13% [4]	-14%	Calderone, (2011)	B	IES VE
	-2%		Daly et al. (2014)	B	ECOTECT
	-82% to 74% [9]	-10%	Diamond et al. (2006)	B	
	-30%		Kimpian et al. (2014)	F	DSM / TM22
	40%		Korjenic & Bednar (2012)	H	BuildOpt_VIE
	72% and 113% [2]	93%	Menezes et al. (2012)	E	TM22
	-17%		Murphy & Castleton (2014)	B	SBEM
	160%		Norford et al. (1994)	B	DOE-2.1C
	-32% to 148% [15]	30%	Piette et al. (1994)	A	DOE 2.1
		63% [30]	Ruyssevelt et al. (2014)	n/a	
	31% and 73% [2]	52%	Torcellini et al. (2006)	B	DOE 2.2
	25 offices	16% {σ 53}*			
Laboratory	1% to 95% [2]	32%	Diamond et al. (2006)	B	
Restaurant	-13% to 71% [4]	31% {σ 31}	Piette et al. (1994)	B	DOE 2.1
Schools	29% to 124% [4]	71%	Kimpian et al. (2014)	F	DSM / TM22
	111% to 127% [3]	117%	Pegg et al. (2007)	D	
	-3% and -6% [2]	-5%	Piette et al. (1994)	A	DOE 2.1
		37% [58]	Ruyssevelt et al. (2014)	n/a	
	11 schools	67% {σ 48}*			
Multipurpose	13% to 48% [4]	5%	Ahmed & Culp (2006)	I	DOE-2.1E
	13% to 142% [5]	99%	Piette et al. (1994)	A	DOE 2.1
	69%		Salehi et al. (2013)	B	IES VE
	95% to 132% [3]	113%	Torcellini et al. (2006)	B	DOE 2.2
	8 multipurpose	45% {σ 53}*			
University	22%		Diamond et al. (2006)	B	
	8%		Knight (2008)	B	SBEM
		156% [13]	Ruyssevelt et al. (2014)	n/a	
	3 universities	62% {σ 66}			
Retail		12% [5]	Ruyssevelt et al. (2014)	n/a	
	62%		Torcellini et al. (2006)	B	DOE 2.1E
	2 retail	37% {σ 25}			
Supermarket	-25% and 5% [2]	-10%	Piette et al. (1994)	B	DOE 2.1
Library	-32% and 48% [2]	8%	Diamond et al. (2006)	B	

*The averages per sector exclude the case studies where the prediction method used is A
Actual = predicted ± % over/under-predicted. i.e. +160% means that actual energy use is 160% higher than predicted. ***see **Table 2.2

Table 2.2: Type of prediction used in the reviewed literature case studies, denoted by a letter.

	Type of prediction
A	Design stage calculation, excluding unregulated loads
B	Design stage calculation, including equipment energy use, standard operation
C	Quasi steady-state hourly simulation relying on simple normative models (EN-ISO 13790)

	Type of prediction
D	CIBSE building energy code 1 (1998) using monthly average temperatures and included unregulated loads, no thermal modelling, similar to CIBSE TM22 bottom-up approach
E	CIBSE TM22 Bottom-up approach
F	National Calculation Methodology (NCM) in thermal modelling + equipment, external lighting and lift using TM22
G	National Calculation Methodology (NCM) in thermal modelling + benchmarking for DHW and Auxiliary loads +equipment, external lighting and lift using TM22
H	Monthly balance method (EN-ISO 13790) + equipment energy use
I	Design stage calculation partially based on built information, including equipment energy use

In the UK, regulatory compliance calculations include the equipment loads solely to estimate the heating and cooling loads, but are excluded from the final evaluation of energy use. From a compliance perspective, plug loads/equipment loads are excluded as they are strongly dependent on occupant behaviour and fit-out specification, which in many shell and core designs are not well known, making them difficult to predict. Excluding them however, makes it easier and arguably fairer to compare the efficiency between buildings, i.e. comparing EPCs. On the other hand, it makes sense to include them as they make up a large portion of total energy use, which otherwise leads to the 'perceived' gap where compliance results are compared to actual energy use, even though some energy end-uses are not taken into account and the methodology was never intended for this purpose. In the end, it comes back to the classification of the performance gap(s).

From 62 reviewed case studies the average discrepancy between regulatory energy use predictions and measurements is +34%, with a standard deviation of 55%. These studies include a prediction of equipment energy use and thus exclude studies with prediction method A for a fair comparison. In the reviewed studies a baseline is often calculated according to national calculation methodologies to comply with Building Regulations. The analysis highlights schools and university buildings to have a large average gap of 67% and 62% respectively. On the contrary, offices show a much lower average difference of only 16%, whereas most averages show a standard deviation of about 50%, indicating a large variation in between individual buildings.

Reviewed case studies are primarily from the United Kingdom (18) and the United States of America (34). When comparing by country, the UK shows an average discrepancy of +65% compared to +17% in the USA, both have a similar standard deviation of about 50%. A more detailed look at the specific studies identifies potential reasons for such a difference. First, Diamond et al. (2006) analyse LEED certified buildings in their study, and hypothesise themselves that such buildings tend to have better agreement between design and actual energy use. Second, Ahmed & Culp (2006) present 4 buildings where their design model is partially based on as-built information, resulting in a more careful consideration of energy use in the building compared to compliance models. Excluding case study buildings from these two authors would increase the average discrepancy to 46%. Finally, there are some fundamental differences between the UK and US compliance modelling methods, NCM and ASHRAE respectively, these influence design stage predictions through different standard assumptions. However, without a large sample size no well-supported claims can be made.

In absolute values for the case studies are visualised in **Figure 2.2** and **Figure 2.4** illustrates the percentage differences from measured to predicted energy use, segmented by building function. The +100% means that measured energy use is twice the amount predicted, while -100% means that measured energy use is 0 kWh/m^{2}a.

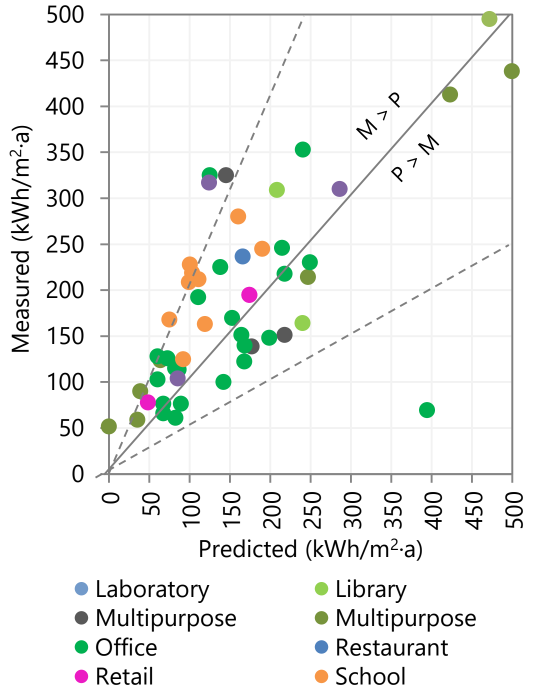


Figure 2.3: Predicted and measured energy use intensities of reviewed case studies for different building functions

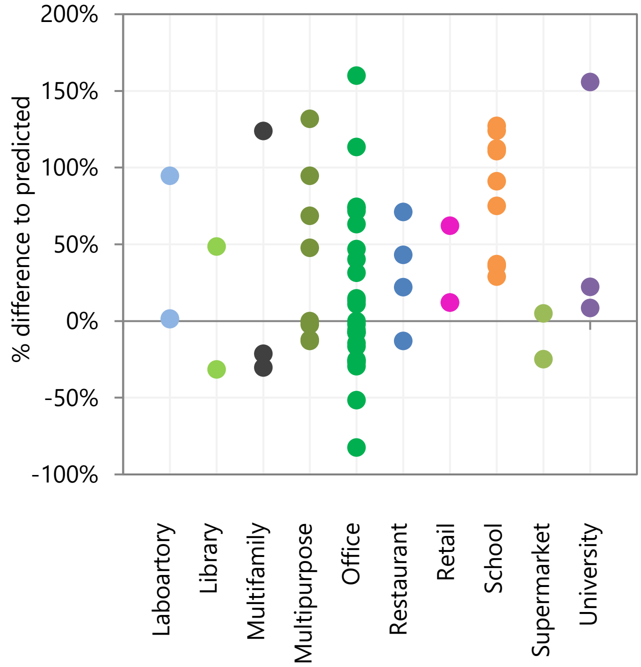


Figure 2.4: Difference in percentages of measured to predicted energy use of reviewed case studies for different building functions

In **Figure 2.3** several of the buildings have high energy use intensities (>1,000 kWh/m²a) and are therefore not shown. Two dash lines indicate if measured energy use in specific case studies is double the predicted value (top left) or is less than half the predicted value (bottom right). This visualisation emphasises that measured energy use in most case studies is higher than predicted energy use. While 15% (9 of 62) of the case studies use double the amount of energy initially predicted in contrast to one significant outlier that uses less than half the energy initially predicted. In particular, Ruyssevelt et al. (2014) reported university buildings to use 156% more energy than initially predicted, an average based on 13 individual studies. Knight (2008) and Diamond et al. (2006), who report a difference of only 8% and 22% for university buildings respectively, do however not support this percentage. Similarly, Ruyssevelt et al. (2014) reported schools to use 37% more energy than initially predicted, based on an average of 58 individual studies. Whereas Kimpian et al. (2014) and Pegg et al. (2007) report much higher average percentages of 117% (3 schools) and 71% (5 Schools) respectively.

Figure 2.4 shows that the number of case studies for the different building functions is not well distributed, the dataset consists mainly of offices, schools and multipurpose buildings. Although the sample size is small, analysis indicates that schools generally consume more energy than predicted. Pegg et al. (2007) reasoned it to be, amongst others, due to 24h security and the need for lighting to be on for security cameras, facilities management that had little relative experience with building services and systems not optimised to respond to changing requirements. The discrepancies for offices are much more variable with an average of +22% and standard deviation of 50%.

Most studies use compliance modelling with the inclusion of equipment energy use, and use model calibration to further understand the underlying causes of a difference. It therefore remains unclear how significant the energy performance gap would be when performance modelling is compared with measured energy use.

2.4.2 Underlying causes

In **Figure 2.5** an overview is given of some of the underlying causes of the performance gap existent in the different stages of the building life cycle according to the Royal Institute of British Architects' (RIBA) plan of work (RIBA, 2013) drawn in relation to building performance influenced by such underlying causes as proposed by Bunn & Burman (2015). Although they propose a smooth 's-curve', perhaps more realistically, building performance should be represented by a jagged line, where performance is affected continually instead of continuously. This model allows for visualising performance issues and clarifies the classification of the different types of performance gaps. Addressing every source available will help in assessing evidence on the impact of these issues, it is therefore also useful to also look at domestic experiences in this area.

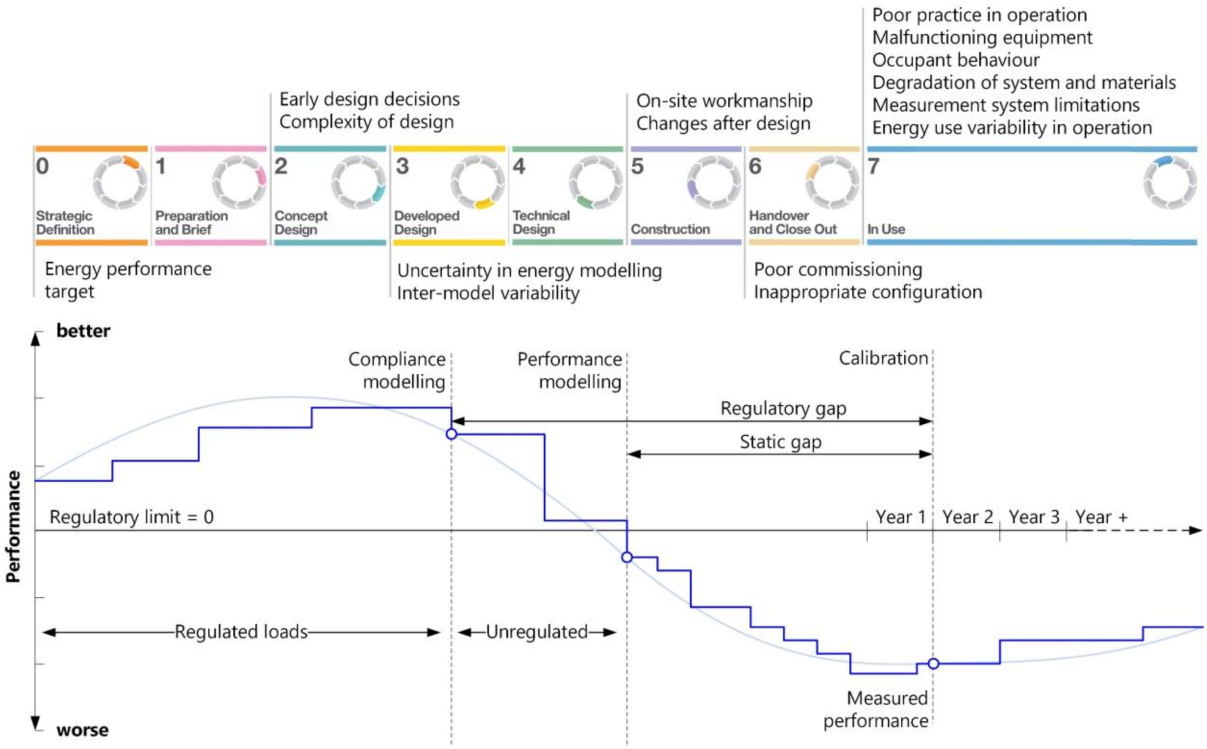


Figure 2.5: Underlying causes aligned to RIBA stages (adapted from RIBA (2013)) and S-curve visualisation of performance throughout the life cycle (adapted from Bunn & Burman (2015))

Limited understanding of impact of early design decisions

During the early design stage there is a lack of focus and understanding on the energy implications of design decisions (ZCH, 2014), requiring a need to educate clients and end-users in how buildings perform against initial design specifications (Morant, 2012). Choices such as, form, orientation, materials, use of renewables, passive strategies, innovative solutions and others should be critically addressed during the concept design. Uncertainty and sensitivity analysis that determine the impact of design parameters can guide the design process through identifying and preventing costly design mistakes before they occur (Bucking, et al., 2014).

Complexity of design

Another reason for introducing a discrepancy between design and operational energy performance, is the buildings' complexity. For example, mistakes in construction become more frequent and complex systems are less well understood (Bunn & Burman, 2015). Simplicity should be the aim of the design as many of the underlying issues relate to the complexity of the building (Williamson, 2012).

Uncertainty in building energy modelling

A building energy model represents the speculative design of a building and is a simplification of reality. It is therefore important to quantify to what degree it is imperfect (Manfren, et al., 2013). In the detailed design stage, building energy modelling requires a high level of detail in order to accurately predict the energy use of a building. Myriad uncertain parameters can have a large effect on the final performance due to the aggregated effect of uncertainties. Among uncertainties in design, those related to natural variability, such as material properties are relatively well covered (de Wit & Augenbroe, 2002). Other uncertainties are less well understood and need a strong basis for research to be established in modelling procedures. Investigation towards well-defined assumptions can assist in more accurately and confidently predict performance of a building (Heidarinejad, et al., 2013). Different sources of uncertainty exist in the use of building simulation, de Wit (2001) classified specification, modelling, numerical and scenario uncertainty, where heuristic uncertainty has been added to describe human-introduced errors as reported by (Kim & Augenbroe, 2013):

- **Specification uncertainty:** Arise from incomplete or inaccurate specification of the building or systems modelled. This refers to the lack of information on the exact properties and may include model parameters such as; geometry, material properties, HVAC specifications, plant and system schedules, casual gains, etc. Parameters related to specification uncertainty are often 'highly unknown' during the early design stage, and can have a large effect on the predicted energy use, assumptions for such parameters are often not representative of actual values in operation. Olivier (2001) reports that UK figures for construction U-values using in calculations are optimistic and theoretical savings are not achieved in practice, due to the exclusion of many types of thermal bridges and neglecting adverse effects of wind on heat loss and little correction for the effects of moisture on thermal conductivity. Burman et al. (2012) identified that a constant value assumed for the specific fan power was actually changing radically in operation. Similarly, Salehi et al. (2013) identified that underlying assumptions for plug loads and lighting were significantly underestimated.
- **Modelling uncertainty:** Arise from simplifications introduced in the development of the model. These include system simplification, zoning, stochastic process scheduling, but also calculation algorithms. Wetter (2011) asserted that mechanical systems and their control systems are often so simplified that they do not capture dynamic behaviour and part-load operation of the mechanical system or the response of feedback control systems. Salehi et al. (2013) were unable to model the unconventional heating system of a building in utilised modelling software, which may lead to wrong performance prediction, further support this. Also Burman et al. (2012) found that pump auxiliary power could not be modelled for compliance purposes and had to apply default values based on HVAC system type. Such tool limitations are extensively reported and contrasted and highlight that certain systems and its configurations are not supported by building simulation software (Crawley, et al., 2008).
- **Numerical uncertainty:** Errors introduced in the discretisation and simulation of the model. Modelling implies a simplification of the real physical processes in real buildings, differences between measured and calculated results can arise due to the application of different numerical solution techniques (Judkoff, et al., 2008).
- **Scenario uncertainty:** Uncertainty related to the external environment of a system and its effects on the system. The specification of weather, building operation and occupant behaviour in the design model. Accuracy of design weather data can have a large effect on the predicted energy performance of a building. According to Bhandari et al. (2012) the predicted annual

building energy consumption can vary up to 7% as a function of the provided location's weather data. While (Wang, et al., 2012) showed that the impact of year-to-year weather fluctuation on the energy use of a building ranges from -4 to 6%. Knowing the uncertainty of related microclimate variables is necessary to understand its impact on energy prediction (Sun, et al., 2014). Similarly, occupants are an uncertainty external to the system and play a major role in the operation of a building. In design calculations occupancy is normally accounted for through a fraction profile, which determines their presence in the model and separately determine when they can operate building equipment. This profile is simplified by taking the average behaviour of the occupants, and therefore neglects temporal variations and atypical behaviour (Kim & Augenbroe, 2013). Furthermore, occupant effects are related to specification uncertainty through assumed base loads (e.g. lighting and equipment), which make it difficult to determine how occupant profiles or wrong base load assumptions impact the energy performance.

- **Heuristic uncertainty:** Human-introduced error in the form of modeller's bias or mistakes. User errors are inevitably quite common due to the complexity of building energy simulation and its tools, these errors range from modellers setting up a building system in different ways, forgetting to apply operation or occupancy profiles to the correct zones or can be related to geometry creation. Guyon (1997) investigated the influence of 12 energy modellers on prediction of energy consumption of a residential house, and found a 40% variability in their final predictions. A similar observation was made by the Building Research Establishment (BRE) where 25 users predicted the energy consumption of a large complex building and found that their results varied from -46% to +106% (Bloomfied, 1988).

Inter-model variability

Energy use prediction is performed using different tools, developed in different countries, for different reasons and as such introduce variability in the results when modelling the same building, i.e. inter-model variability. This is directly related to uncertainties in building energy simulation, especially model simplification, user error and numerical uncertainties will drive the variability between different tools. These tools are utilised for the purpose of prediction and thus have to give credible and relatively accurate results. Neymark et al. (2002) compared seven different tools and indicated a 4-40% disagreement in energy consumption. Raslan & Davies (2010) compared 13 different accredited software tools. They highlighted a large of variability in the results produced by each of the tools in their consistency in granting approval with Building Regulations. In a more recent study, Schwartz & Raslan (2013) performed an inter-model comparative analysis of three different dynamic simulation tools using a single case study and found a 35% variability in the total energy consumption.

On-site workmanship

As Building Regulations become more stringent, the quality of construction has to be improved and new technologies are introduced. On site workmanship needs to adapt and be trained to these increasing levels of complexity in building construction. New skills such as air tightness detailing for limiting air infiltration give rise to performance issues as airtightness is compromised during construction by discontinuous insulation or punctured airtight barriers (Williamson, 2012). Installation of services, such as drainage, air ducts and electrical pipe work can often leave gaps which also reduce airtightness and induce thermal loss (Morant, 2012). Other common issues related to on-site workmanship are eaves to wall junction insulation, incorrect positioning of windows and doors which reduce the actual performance of the thermal envelope (ZCH, 2014). These issues are more prone to affect the energy performance in domestic buildings, where usually the performance of the thermal envelope is more significant.

Changes after design

After the building is designed, often products or changes are value engineered and directly influence energy performance gap as such changes are often not fed back to the design team for evaluation against the required performance standard (ZCH, 2014). These changes can occur during design due to site constraints, not well thought of integration of design modules problems with detailing and budget issues. Murphy & Castleton (2014) reported in their case study that the roll-out of unspecified low-energy equipment affected final unregulated loads, and indirectly cooling energy use due to lower internal gains. Similarly, Morant (2012) reported inconsistencies between design specified and installed lighting loads in an office, which had a considerable impact on the discrepancy between predicted and measured electricity use. Good communication and coordination by the contractor is essential to prevent changes in design changes to influence the energy performance.

Poor commissioning

Piette et al. (1994) reported poor commissioning of control measures, which were not set-up for proper control, and operation. Kimpian et al. (2014) identified that inverters for supply and extract fans were provided to AHUs, but were not enabled during commissioning, resulting in the fans to operate at maximum speed at all times. In operation, such issues persist and require continuous commissioning.

Poor practice and malfunctioning equipment

The actual operation of a building is idealised during design by making assumptions for temperature setpoints, control schedules and general performance of HVAC systems. In practice however, it is often the case that many of these assumptions deviate and directly influence a building's energy use. Kleber & Wagner (2007) monitored an office building and found that failures in operating the building's facilities caused higher energy consumption, they underline the importance of continuous monitoring of a building. Wang et al. (2012) showed that poor practice in building operations across multiple parameters results in an increase in energy use of 49-79%, while good practice reduces energy consumption by 15-29%. Piette et al. (1994) suggest that building operators do not necessarily possess the appropriate data, information, training and tools needed to provide optimal results. As such, operational assumptions made in the design stage may not be met by building operators (Moezzi, et al., 2013).

Occupant behaviour

Another dynamic factor for a building in use are occupants. They have a substantial influence on the energy performance of a building by handling controls, such as those for lighting, sun-shading, windows, setpoints, and office equipment, but also through their presence, which may deviate from assumed schedules. People are very different in their behaviour through culture, upbringing and education, making their influence on energy consumption highly variable. One of the major factors that has been reported to have a large influence on the discrepancy is night-time energy use from leaving office equipment on (Masoso & Grobler, 2010; Zhang, et al., 2011; Mulville, et al., 2014; Kawamoto, et al., 2004). Both related to occupant behaviour (not turning off equipment) and assumptions for operational schedules, extended working hours not taken into account in the design model. In an uncontrolled environment (not extensively monitored) it is difficult to determine how one or the other is influencing the discrepancy. Azar & Menassa (2012) investigated 30 typical office buildings and found that certain occupancy behavioural actions influenced energy use by 23.6%. Parys et al. (2010) reported a standard deviation of up to 10% on energy use to be related to occupant behaviour. Martani et al. (2012) studied two buildings and found a 63% and

69% variation in electricity consumption due to occupant behaviour. Using modelling, Hong and Lin (2013) investigated different work styles in an office space and found that an austere work style consumes up to 50% less energy while a wasteful work style consumes 90% more energy. Similarly, Clevenger and Haymaker (2006) studied an elementary school with varying types of occupant behaviour, whereas high-end values affected energy use by up to 150%.

Measurement system limitations

Similar to predicting energy use using building energy models, measured energy use obtained from measurement systems needs to be validated to ensure accuracy of the data. Limitations of measurement systems make adequate assessment of energy use inaccurate (Maile, et al., 2010). Typical sensor accuracies to lie within 1-5% for normal operating conditions, whereas incorrectly placed sensors will have increased levels of error (Maile, et al., 2010). Most common sources are calibration errors, or the absence of calibration (Palmer, et al., 2016). Fedoruk et al. (2015) identified that measurements were not accurately representing the performance of buildings systems due to mislabelling, incorrectly installed and calibrated. They report that simply having access to large amounts may actually result in more confusion and operational problems.

Longitudinal variability in operation

Finally, commonly the energy performance gap is assessed for a year of available data. However, longitudinal performance is affected factors such as building occupancy, deterioration of physical elements, climatic conditions, and building maintenance processes and policies (de Wilde, et al., 2011). Brown et al. (2010) present a longitudinal analysis of 25 buildings in the UK and found an increase of 9% in energy use on average per year over 7 years, with a standard deviation of 18%. Similarly, Piette et al. (1994) analysed 28 buildings in the US and found an average increase of 6% between the third and fourth year, with no average increase during the fifth year. Thus, a longitudinal variability in operational energy use has to be taken into account when investigating the energy performance gap.

2.4.3 Assessing the underlying causes

All of these causes combined can have a large influence on the final energy performance of a building. **Table 2.3** shows a risk matrix that defines the potential associated risks of the discussed underlying causes based on general consensus in literature. An overview of several of the reviewed case studies and their reported underlying reasons for a discrepancy are given in Appendix A.

Table 2.3: Potential risk on energy use from underlying causes assessed based on general consensus in literature.

	Underlying cause	Evidence from literature*	Rated impact on energy use	Estimated quantitative effect	Compliance modelling related
Context	Energy performance target	Low	High		Yes
	Impact of early design decisions	Medium	High		
	Complexity of design	Low	Medium		
Model	Specification (geometry, material, equipment)	High	High	20-60%	Yes
	Modelling (simplification)	Medium	Medium	<10%	Yes
	Numerical (discretisation)	Low	Low	<5%	
	Scenario (weather, schedule, operation)	High	Medium	10-30%	Yes
	Heuristic (user)	Low	High	<70%	

	Underlying cause	Evidence from literature*	Rated impact on energy use	Estimated quantitative effect	Compliance modelling related
	Inter-model variability	Medium	Medium	5-40%	
Construction	On-site workmanship	Medium	Low		
	Changes after design	Low	Low		
Commissioning	Poor commissioning	Medium	Medium	<20%	
Operation	Poor practice in operation	High	High	15-80%	
	Occupant behaviour	High	High	10-80%	
	Degradation of system and materials	Low	Low	<10%	
	Measurement system limitation	Low	Low	<10%	
	Energy use variability in operation	Low	Medium	5-15%	

*Based on the number of mentions in literature and their consensus of the impact on performance.

Important underlying causes identified in literature are those that have a high impact and high evidence rating, these are specifically related to specification uncertainty in building modelling (20-60%), occupant behaviour (10-80%) and poor practice in operation (15-80%), with percentage effect on energy use. Other important factors that are likely to have a high rated impact are the energy performance target, impact of early design decisions and heuristic uncertainty in modelling.

An assessment of the underlying causes of the energy performance gap has shown that there is a need for both action and further research to be undertaken. Detailed building prediction methods and post-occupancy evaluation have proven to be essential in the understanding of building assumptions, occupant behaviour, systems and the discrepancy between predicted and measured energy use.

2.5 Reducing the energy performance gap

A major concern in the built environment is the fragmentation of disciplines involved in the building life cycle stages. Traditionally designers, engineers and contractors are involved in the building development process, but leave once the building is physically complete, leaving the end-users with a building they are unlikely to fully understand. The design community rarely goes back to see how buildings perform after they have been constructed (Torcellini, et al., 2006). Feedback mechanisms on energy performance are not well developed and it is generally assumed that buildings perform as designed, consequently there is little understanding of what works and what does not, which makes it difficult to continuously improve performance (ZCH & NHBC Foundation, 2010). Gathering more evidence on both the performance gap and its underlying issues can support feedback mechanisms and prioritise principle issues. For this, the primary requirement is the collection of operational performance data, which can be fed back to design teams to ensure lessons are learnt and issues are avoided in future designs. It can help policy makers understand the trend of energy use and support the development of regulations. Finally, operational data is valuable to facilities management in order to efficiently operate the building. This feedback process is illustrated in **Figure 2.6**.

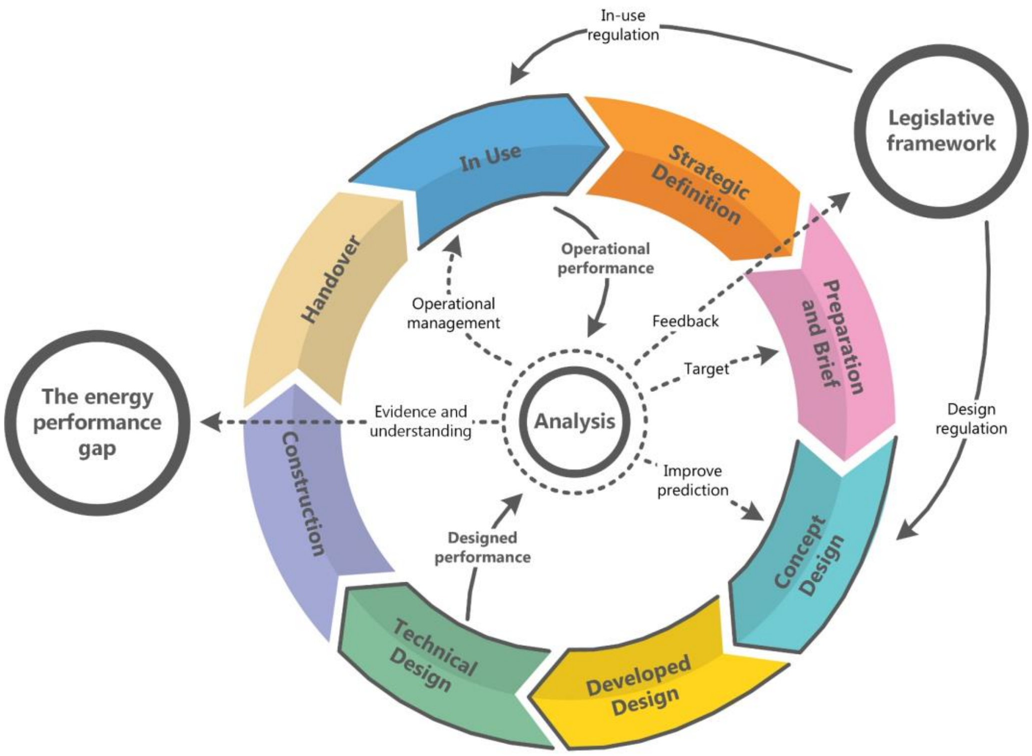


Figure 2.6: Feedback process in relation to the RIBA plan of work stages, adapted from RIBA (2013).

2.5.1 Legislative frameworks

In 2014, the UK department of Energy & Climate Change introduced the Energy Savings Opportunity Scheme (ESOS) in order to promote operational management in buildings, an enacted policy under the Energy Performance of Buildings Directive (EPBD). A mandatory energy assessment to identify energy savings in corporate undertakings that either employ more than 250 or have an annual turnover in excess of ~38 million pounds (50 million euros). An assessor should calculate how much energy could be saved from improved efficiency. How these energy savings are predicted is however left open and could entail simple hand calculations instead of the more detailed dynamic thermal simulations, furthermore implementing proposed energy savings are voluntary. Following a similar process are Energy Performance Contracts, these contracts legally bind a third party for predicted savings to be realised, otherwise equivalent compensation needs to be provided. It then becomes important to make accurate predictions of energy conservation measures as their reliability directly influences the profit of the businesses providing these contracts.

For new buildings, regulatory limits become ever more stringent in order to mitigate climate change and achieving them requires new energy efficient technologies and higher quality construction materials to be proposed. Although such limits are theoretically engineered, evidently such targets are not achieved in practice, without any repercussions for modellers who produce inaccurate predictions, this can foster a lack of confidence in simulation in the building industry and may soon be met by legal and financial implications (Daly, et al., 2014). Burman et al. (2014) proposed a framework that measures any excess in energy use over the regulatory limit set out for the building. This excess in energy use could cause disproportionate environmental damage and it could be argued that it should be charged at a different rate or be subject to an environmental tax. Kimpian et al. (2014) suggest mandating the disclosure of design stage calculations and assumptions as well as operational energy use outcomes in Building Regulations, such data would significantly support the understanding of the energy performance gap. Governments continue to face the difficult task of balancing the principal of not interfering in the affairs of businesses with the recognition of serious consequences of energy waste and climate change (Jonlin, 2014).

2.5.2 Data collection

Accessible meter data is invaluable to confirm that building really do achieve their designed and approved goals (UKGBC, 2007). A continued lack of such data is likely to lead to a progressive widening of the gap between predicted and measured energy use (Oreszczyn & Lowe, 2009). Energy performance data can be used by design teams to enable them to deliver better designs, clients to enable benchmarking and develop a lower carbon building brief, building users to drive change and management in operation, policy makers to target plans and incentives and monitor the trend of energy use (HM Government, 2010). Without data collection there would be no feedback loop to inform future policy and regulation (ZCH, 2010).

2.5.3 Design improvements

Negating the performance gap starts at the beginning of a project. At this point it is important to set a high energy performance target, which can assist in a more rigorous review of system specifications and operational risks (Kimpian, et al., 2014). With high performance expectations it becomes necessary to carry out performance modelling, validate assumptions made in the building model, make sure that building fabric is constructed to a high standard, properly commissioned systems and to operate the building as efficiently and effectively as possible. Making an accurate prediction of building energy performance becomes an integral part of the design process. A building design however can be based on thousands of input parameters, often obtained from guidelines or Building Regulations, some of which have extensive background research while others are only best-guess values. Pegg et al. (2007) argue for the use of feedback to inform design and need for realistic and relevant benchmarks. However, such feedback is often very case specific. Research on occupant behaviour is extensive because many factors, such as control of lighting, equipment, windows, etc. are dependent on how occupants interact with them. Martani et al. (2012) propose a method to measure occupancy using Wi-Fi connections to determine its relation to HVAC levels and electricity supply. Whereas, Mahdavi & Pröglhöf (2009) suggest the collection of occupancy behaviour information to derive generalised (aggregate models) and utilise such models in building energy simulation. Capturing user-based control actions and generalising these as simulation inputs can provide more accuracy in prediction. Using operational data Rysanek & Choudhary (2014) used an open-source software to generate hourly profiles of occupancy services demand for use in common building energy models. Simulation can be further supported by introducing well-defined uncertainties in design, improving the robustness of the building design, reliability of energy simulation and enable design decision support, in particular when supported by sensitivity analysis (Hopfe & Hensen, 2011).

2.5.4 Training and education

During construction, robust checking and testing is necessary to ensure that the quality of construction is maintained (Morant, 2012). Furthermore, changes during design and construction have to be communicated, the supply chain has to be informed in time to make sure everything is integrated appropriately. The real performance of building elements is underestimated as they are taken from lab-tests and omit, for example, the occurrence of thermal bridge mistakes during construction. Clear guidance on thermal bridging should be therefore be provided to the construction industry (ZCH, 2014). Training and education need to increase the skills in the construction industry to ensure better communication and quality of construction. Similarly, training and education should be enhanced for facility managers, to more strictly perform maintenance and operation of buildings, and in the design stage it is important to create awareness to energy modellers of the energy performance discrepancy, while promoting skills, innovation and technological development in order to deal more appropriately with creating a robust design.

2.5.5 Operational management

Post occupancy evaluation studies have shown that buildings are often poorly commissioned and that there is a lack of continuous commissioning during operation (Kimpian, et al., 2014). Frequent re-commissioning exercises can help maximise the efficiency of building services, avoiding unnecessary energy use (Morant, 2012). For guidance in this process, the Soft Landings framework was developed in order to provide extended aftercare, through monitoring, performance reviews and feedback. However, its purpose has been extended to provide a relevant guide to building procurement throughout the building life cycle. While Soft Landings provides useful guidance, Kimpian et al. (2014) found it ineffective in practice and only when data started to be collected and analysed, procurement related problems began to surface. Aftercare, but also professional assistance is required as technologies and solutions made during the design often prove too complicated to be manageable (Way, et al., 2014). Continual monitoring of the performance during operation is thus important in order to ensure that design goals are met under normal operating conditions (Torcellini, et al., 2006). It is essential that facilities managers take ownership of energy consumption in buildings as they have detailed information of operational issues (CIBSE, 2015).

2.6 Predicting and measuring energy use

Three types of performance gaps (regulatory, static, dynamic) have been identified. To understand why there are differences and how they can be mitigated, an overview is given on how predicted and measured energy use can be compared through different performance quantification and assessment methods. The magnitude and causal factors that influence the discrepancy between predicted and measured energy performance can only be identified if detailed information from both the design and operational stages is analysed. The magnitude of a discrepancy can be obtained through comparison, but its coherency depends on the type of quantification method applied, as they determine the energy performance of a building. Thus it is important to understand the fundamental differences of different performance assessment methods. Furthermore, the introduction of classification schemes and benchmarking have further allowed the comparison of building energy efficiency, likewise these assessment methods have inherent differences between which need to be understood in order to make like for like comparisons.

2.6.1 Energy performance quantification

Energy use for a building can be obtained through three different approaches: calculation-based, measurement-based and a hybrid approach. These differ in cost, time, effort, level of detail, accuracy and availability. Where utility bills can provide quick and accurate results, they are only relevant in existing buildings where enough data has been collected over a longer period of time, but do not provide the level of detail as a monitoring-based method could. Dynamic simulation on the other hand can include an even higher level of detail about the predicted energy performance of a building, but its relative accuracy to a utility bill will be lower. A hybrid approach is a combination of calculation and measurements, an example of this is calibrated simulation. **Figure 2.7** illustrates three different approaches and their underlying methods and applications.

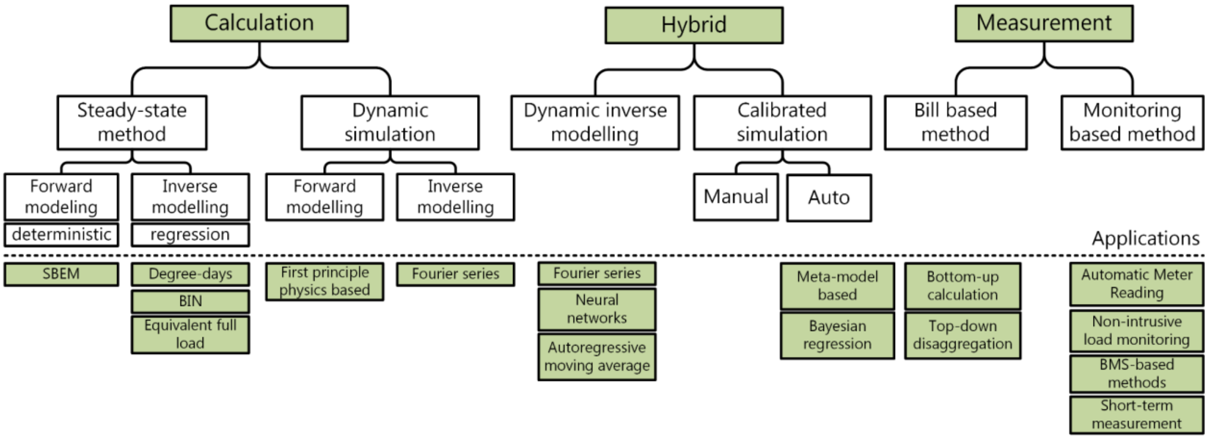


Figure 2.7: Energy performance quantification methods, adapted from (Wang, et al., 2012).

Calculation-based method

Building energy performance calculations consist of defining inputs to determine the outcome of energy performance. Inputs are required in the simulation engine that describe detailed mathematical models and represent the interaction between actual building physics and services. Outputs from the calculations typically include the annual energy use and carbon emissions with a more detailed energy breakdown in end-uses and graphical building services results. Calculation-based methods can be divided into:

- *Steady-state methods*, ignore or simplify dynamic effects by correlation factors.
- *Dynamic simulation,* is more suited to functional and volumetric complexities as they allow for more detailed input options (Raslan & Davies, 2010) and are capable of capturing building dynamics of the envelope and systems (Wang, et al., 2012).

Both these calculation methods can be established using a *forward (deterministic)* or *inverse (regression) modelling* approach. A forward modelling approach describes the building services and defines the building according to its physical configuration. The inverse modelling approach relates the energy performance indicators (outputs) to the influential factors (inputs), a physical configuration of a building or system is assumed and a model coefficient is identified by statistical analysis.

Under the EPBD both of these methods are used for energy performance calculation procedures for regulatory compliance (e.g. Part L compliance and EPCs in the UK) described in more detail in ISO 13790 (ISO, 2008) and ISO 52016-1 (ISO, 2017). The dynamic approach differs from the steady-state method by taking into account the non-linear and interactive heat transfer within a building, thus capturing periodic changes such as temperature within the building. The application of the methods used may lead to different compliance results, another factor that will influence the static gap between predictions and measurements.

Hybrid-based method

Hybrid methods combine the use of computational analysis while being supplemented by measurements to identify model parameters or calculation discrepancies. Two typical hybrid methods to quantify energy performance are calibrated simulation and dynamic inverse modelling.

- *Dynamic inverse modelling*, broadly known as machine learning is a method that uses correlation between input parameters and output parameters to predict energy use without explicitly modelling the systems and physical processes of a building (de Wilde, 2014). These types of models are complex and need detailed measurements to fine-tune the model (Haberl

& Culp, 2005). Typical examples of dynamic inverse models include autoregressive moving average (ARMA) models, artificial neural networks (ANN), and Fourier series.

- *Calibrated simulation,* calibration is the process of adjusting numerical or physical modelling parameters in the computational model for the purpose of improving agreement with real-world data (Oberkampf & Roy, 2010). Calibrated simulation can be used to quantify the impact of energy conservation measures using simulation models. Due to its detailed analysis procedure, it can provide feedback to improve the quality of future design stage models by identifying common mistakes in assumptions (Raftery, et al., 2011). Iterative efforts can be reduced by treating calibration as an optimisation problem by using sensitivity analysis to identify which values can be mathematically tuned to their reference values (Sun & Reddy, 2006; Manfren, et al., 2013).

Measurement-based method

Whenever energy quantification is established using calculation, a certain discrepancy will exist between the predicted and actual energy performance, therefore the credibility of calculated results should always be questioned. For new buildings, energy quantification using calculation is the only method. For existing buildings however, the use of measured energy performance is more accurate. Measurement methods can be divided into energy-bill based methods and monitoring-based methods and its accuracy depends on the temporal resolution of the system and measurement limitations.

- *Energy bills* can provide some easily attainable measurement data in most existing buildings, but are often based on estimates. Furthermore, monthly bills provide insufficient information for detailed energy performance assessment. To acquire more information for end-uses, energy bills can be disaggregated to provide a better understanding of the energy use in a building.
- *Monitoring based-methods* provide more accurate and detailed energy use consumption. There are several technologies available. Automatic meter reading (AMR) obtains energy use of individual loads by placing separate metering hardware on relevant circuit branches. Non-intrusive load monitoring (NILM) uses pattern recognition that is capable of gathering detailed energy-use without sub-metering. Building management systems (BMS) collect data to get a clear picture of the energy use of typical HVAC systems and the whole building and have control over these systems.

2.6.2 Performance assessment methods

The most commonly used indicator for building energy use is the annual energy use per unit area, determined by the relation between a building and its occupants, systems and external environment. To help buildings reduce their energy consumption, energy performance assessment assists in communicating a building's energy efficiency. Energy performance assessment can be classified into performance-based and feature-specific approaches. Performance-based approaches compare quantifiable performance indicators against established benchmarks and feature-specific approaches are awarded credits when criteria of specific features are met. Figure 2.8 gives an overview four different performance assessment methods with a breakdown of different techniques and applications.

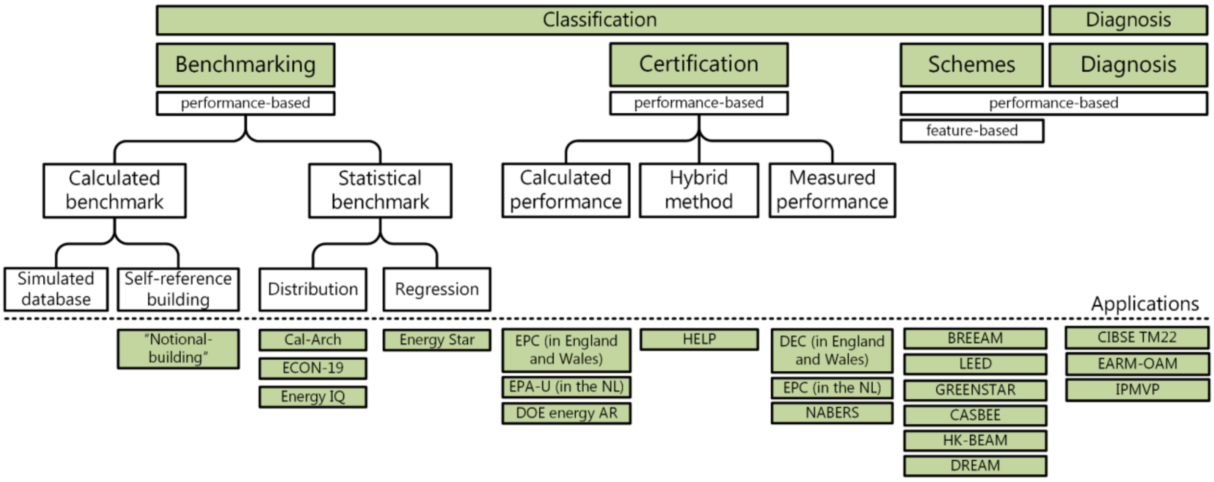


Figure 2.8: Energy performance assessment applications for existing and new buildings.

Whole-building benchmarking

Benchmarks can be established by applying statistical analysis of building stocks that have a coherent building function with the to be compared building, simple normalised and regression-based models employed for this purpose. Alternatively, benchmarks can be calculated by using building energy simulation to a variety of building functions and parameters, or by using a self-reference (i.e. notional building) building as the comparison criterion.

Energy certification

The performance of a building is certified by an authorised institute or person, and generally includes and "energy rating" process to quantify energy use, and an "energy labelling" scale to classify the corresponding performance (Wang, et al., 2012). There are many different building energy certificates available, resulting in different calculation methodologies. These calculation methodologies can be divided into three different methods; calculation-, measurement- and hybrid-based methods. For new buildings, the calculation-based method uses simulation to determine the energy performance in the design stage and provides the building with an "asset rating". Existing buildings can be provided with a similar certificate based measured data to provide an "operational rating". A combination (hybrid method) using calculation and measurements is also applied in some cases to provide an energy certificate.

Building environmental assessment schemes

Environmental assessment schemes encourage the development of sustainable "green" buildings. Rating systems generally assess environmental, social and economic impacts of various aspects of building development. Environmental aspects concern itself with the efficiency of energy use, water, waste, material, set use, etc. They were used to address the asset of new buildings, but are also available for existing building and even communities. Many assessment schemes exist around the world with most countries having developed their own system (e.g. BREEAM and LEED).

Hierarchical assessment and diagnosis tools and guidelines

Diagnosis tools assess the performance of systems or facilities; they provide specific information through detailed energy audits to identify root causes of energy problems. Furthermore, they were used to detect whether and where energy inefficiencies occur to subsequently fix problems and enhance energy efficiency. Such assessment are specifically carried out in existing building, for example the International Performance Measurement and Verification Protocol (IPMVP) (EVO, 2009) defines standard terms and suggests best practice of among others energy efficiency.

2.7 Modelling and analysis

The regulatory performance gap seems to arise mainly from the misconception that a compliance model is supposed to predict the actual energy use of the building. Whereas the static performance gap is not well understood, because performance modelling has not been common practice among practitioners, and where it is undertaken, predictions are often not validated. Finally, the dynamic performance gap is a more common approach to investigating the performance of existing buildings, because it can identify operational issues, improvements, and determine typical behaviour in buildings, which in turn can support design assumptions. As such, model calibration is seen as one of the key methodologies to develop more accurate energy models, it involves bringing predictions closer to measurements. Therefore, model calibration can only be employed on existing buildings to investigate the dynamic performance gap. Model calibration is the process of changing input parameters of the model within the margins of uncertainty in order to obtain a model that lies within agreed boundary criteria and is therefore likely to predict future options more closely to the actual situation. It is becoming more important as it is increasingly used in activities such as commissioning and energy retrofitting scenarios of existing buildings (Fabrizio & Monetti, 2015).

2.7.1 Existing calibration methodologies

Reddy et al. (2007) provide an extensive framework for building model calibration and give initial insights into automated calibration in building energy models. Raftery et al. (2011) proposed an 'evidence-based' methodology, which clarifies the importance of different data sources and the need for evidence when iterative changes are made to reliably calibrate a model. Maile et al. (2012) developed a methodology, which employs calibration processes on a detailed level and uses a building object hierarchy that represents both a spatial and systems perspective for relating components in the buildings and supports identifying performance issues. Although their methodology proposes to estimate the impact of identified issues and includes a feedback process to both the design and operation, they present this as an area for future research. Both evidence based decision-making and uncertainty analysis techniques can improve the reliability of calibrated models. It does this by making changes according to evidence in a clearly defined hierarchy of priorities and identifying the most influential input variables to guide the model improvement process, while using uncertainty to capture a range of possible results.

Eisenhower et al. (2012) suggest the use of meta-models for optimisation (i.e. automated calibration), which proved to be essential with large models of existing buildings. More recently, several studies aimed to achieve a higher accuracy in model calibration. Yang & Becerik-Gerber (2015) compared HVAC related energy consumption and considered several energy conservation measures. Whereas Yin et al. (2016) describe the calibration of a building model using system performance metrics and energy use. Sun et al. (2016) present an automated calibration approach to tune energy models at a monthly basis through pattern recognition implemented in a web-based building energy retrofit analysis toolkit. Finally, there is also an interest in calibrating model towards predicting environmental parameters. Roberti et al. (2015) calibrated a building energy model with measured indoor air and surface temperatures. Such an approach, in combination with energy, increases complexity, but can also increase the accuracy of calibrated models.

2.7.2 Manual and automated calibration

Existing methodologies for model calibration in building simulation are presented by Coakley et al. (2014), where they broadly classify two approaches; manual and automated calibration. Manual calibration involves changing the input parameters manually, most often decisions are made based on juxtaposing model predictions and real data through graphical means, sometimes supported by assessing the statistical relation between the input parameters and outputs. Variables with a larger significance are adjusted to advance the predictions towards the real solution, i.e. measured data. For large complex models with many parameters, this process is time-consuming, as many iterations need to be run and analysed to get to a potential solution (Sun, et al., 2016). A caveat to model calibration is that it is impossible for the model to exactly represent the existing building, many input parameters are unknown and even for lab-controlled environment getting the exact values for input parameters is impractical (Garrett & New, 2016). Automated calibration takes away some of the iterative decisions from the modeller and instead they are made by algorithms, which are however created and controlled by the modeller, so user error is not necessarily eliminated. Therefore, a good degree of understanding is necessary to guide how the model is adjusted both in manual and automated calibration, to accurately represent the real situation. Automated calibration is a form of optimisation, many simulations are run based on a pre-determined input parameter space and solutions are then analysed and optimised towards the measured data. Due to the complexity of building energy models, simulation time can be considerable, limiting the possibility of running optimisation algorithms using building energy modelling software directly, as the optimisation procedure requires thousands of model runs. Instead, a meta-model of the detailed simulation model can be constructed based on the calculated solution space. The meta-model is a simplified mathematical representation of the detailed building energy model, which reduces computation time from tens of minutes to several milliseconds.

2.7.3 Calibration criteria

Previously, calibration of building energy models has focussed mainly on total monthly energy use, potentially masking the different types of energy end-uses. As such, a calibrated model of total energy use is unlikely to produce accurate results for each end-use (Yin, et al., 2016). Therefore, a higher level of data granularity hypothetically increases the accuracy of calibrated models, as suggested by Garrett & New (2016) and Yin et al. (2016), but has so far been difficult to determine, due to the sparsity of detailed measured data. The use of automated calibration could potentially increase the time efficiency of calibrating energy models. However, the benefits of automated calibration have not been quantitatively investigated. Various guidelines propose acceptable calibration tolerances that consider a model to be calibrated when they fall within pre-scribed values according to the above statistical measures, see **Table 2.4**. However, as pointed out by Chaudhary et al. (2016), these are likely to be outdated as new research efforts in particular using automated calibration can achieve discrepancies of less than 1% at these levels.

Table 2.4: Statistical limits for when a model is considered calibrated.

Statistical measure	Monthly		Hourly	
	ASHRAE	IPMVP	ASHRAE	IPMVP
NMBE (%)	±5	±20	±10	±5
CV(RMSE) (%)	<15	<5	<30	<20
ASHRAE Guideline 14 (ASHRAE, 2002)				
IPMVP Vol I (EVO, 2009)				

2.7.4 Uncertainty and sensitivity analysis

Uncertainty and sensitivity are particularly useful analysis that typically accompany modelling and in particular used for the purpose of calibration. Uncertainty analysis quantifies uncertainty in the output of the model due to the uncertainty in the input parameters. Whereas sensitivity analysis apportions the uncertainty of the model output to the input. Uncertainty is ascribed to a parameter in the form of a distribution of likelihood of the parameter, typically a normal or triangular distribution is used as a certain value is more likely to occur. Whereas design parameters are typically uniformly distributed, because they are considered design options.

Sensitivity analysis techniques can be categorised in local and global methods. Local methods calculate the effect of uncertainty in parameters independently, whereas a global method analyses sensitivity concerning the entire parameter distribution.

Local methods

Local methods use numerical approximations of local derivatives between output and input to estimate parameter sensitivity. Generally good for studying a small number of uncertain parameters. Differential sensitivity analysis is one example of an often-used method. It does however assume that parameters are independent of each other, which is generally often not the case in building physics, as illustrated by Macdonald (2002). To take into account interaction of parameters, factorial analysis can be used, which works by simulating all possible combination of parameter values. Typically, only efficient for a small number of uncertain parameters as the number of simulation (N) grows factorially with the number of parameters (k). Another method is one-factor-at-a-time (OFAT), which is generally less efficient than the factorial method, as it requires more runs and cannot estimate interactions between parameters.

Global methods

Global methods are often implemented using the Monte Carlo method, they use a set of generally randomly determined samples to explore the design space, corresponding model outputs are then statistically analysed to determine its variance. Differences between global methods are related to sampling and analysis of the results. The sensitivity to a parameter is measured as the proportion of the model variance that can be explained by changes in that parameter (ten Broeke, et al., 2016), this allows comparing the sensitivities of different parameters.

Correlation and regression analysis methods

Correlation methods determine a correlation coefficient between the input parameters and the outputs as a sensitivity measure. Several techniques exist which can be applied, mainly depending on the type of data under investigation. Pearson correlation coefficient (PCC) is used when a linear relationship exists between inputs and outputs. Spearman rank correlation (SRCC) is effective for nonlinear but monotonic relationships.

Regression methods derive the sensitivity as a by-product of regression analysis, for example Standard regression coefficient (SRC) can be used when input factor have different units of measurement (Pianosi, et al., 2016). Equations for the computed coefficients are as follows:

(PCC)
$$\rho\left(X_j, Y\right) = \frac{\sum_{i=1}^{N}(X_j^{(i)} - E\left(X_j\right))(Y_i - E(Y))}{\sqrt{\sum_{i=1}^{N}\left(X_j^{(i)} - E\left(X_j\right)\right)^2}\sqrt{\sum_{i=1}^{N}(Y_i - E(Y))^2}} \quad (1)$$

where E is the expectation

(SRCC)
$$\rho_{rg_X, rg_Y} = \frac{cov(rg_X, rg_Y)}{\sigma_{rg_X}\sigma_{rg_Y}} \quad (2)$$

where ρ is the correlation

(SRC)
$$SRC_j = \beta_j\sqrt{\frac{Var\left(X_j\right)}{Var(Y)}} \quad (3)$$

Variance-based global sensitivity analysis

Variance-based methods allow exploring the full input space, whilst account for interactions, and nonlinear responses. Decomposition of the uncertainty of a model quantifies how the uncertainty of the input affects the uncertainty of the output. It can explain how certain subsystems contribute to uncertainty at the building level, this information is both useful for model calibration, quantifying the impact of the underlying causes of the performance gap, understanding the effects of design assumptions, and highlighting influential parameters in existing buildings for operational management.

Variance-based methods decompose the output variance into parts that can be attributed to input parameters and combination of parameters. A square-integrable function, which represents the building model with uncertain parameters can be decomposed into a sum of functions (Sobol, 2001):

$$f(x) = f_0 + \sum_{i=1}^{k} f_i(x_i) + \sum_{j>i}^{k} f_{ij}(x_i, x_j) + \cdots + f_{12\cdots k}(x_1, \cdots, x_k) \tag{4}$$

Sensitivity is then measured as the variance in the output caused by specific inputs. First-order sensitivity indices (sometimes referred to as Sobol' indices) were then used to quantify how sensitive a particular output is to variation of a parameter by measuring the effect of varying a single parameter on its own, but averaged over variations in other input parameters. They are defined by calculating the partial variance relative to the total variance. Whereas higher-order sensitivity indices were defined by calculating the partial variance over two or more parameters. The sensitivity index for parameters is always between 0 and 1, a high value signifying an important variable. Adding first and higher-order indices together will result to 1. The difference between a parameter's first and total order indices represents the effects of its interactions with other parameters (Herman, et al., 2013). Finally, there is the total sensitivity index, which is the sum of all indices for particular parameters, calculated as follows:

$$S_{T_m} = S_m + \sum_{\substack{j>i \\ i\ or\ j=m}}^{k} S_{ij} + \sum_{l>j>i}^{k} S_{ijl} + \cdots + S_{1\cdots m\cdots k} \tag{5}$$

In non-additive models the total sensitivity index will be larger than 1, due to the fact that interaction between different parameters is counted in both total effects. It is typically used to eliminate insignificant parameters.

2.8 Summary

Predicted and measured energy use has been shown to deviate significantly, also termed 'the performance gap'. This gap can be classified as; (1) a difference between compliance and measured energy use (the regulatory gap); (2) as a gap between performance modelling and measured energy use (the static performance gap); (3) a gap between calibrated predictions and measured energy use with longitudinal perspective (the dynamic performance gap). Literary sources have been reviewed in order to understand the significance of the regulatory energy performance gap and its underlying causes have been assessed on their impact on energy use:

- From 62 case studies buildings the average discrepancy between predicted and measured energy use is +34%, with a standard deviation of 55%. These studies include a prediction of equipment energy use.
- The most important underlying causes identified in literature are specification uncertainty in building modelling, occupant behaviour and poor practice in operation, with an estimated

effect of 20-60%, 10-80% and 15-80% on energy use respectively. Other important factors are the energy performance target, impact of early design decisions, and heuristic uncertainty in modelling.

Understanding and mitigating differences between predicted and measured energy use requires an expansion of research efforts and focus on its underlying causes. Detailed energy audits and model calibration are invaluable techniques in order to quantify these causes. Furthermore, tools are necessary to support intuitive visualisation and data disaggregation to display energy uses at detailed levels and for different time granularity, comparing predicted and measured energy use taking a longitudinal approach. It identified several research gaps on which this research will focus:

4. There is a lack of studies investigating the different aspects and meanings of building performance, how stakeholders in the procurement and delivery process view building performance, and what common barriers are to achieving better building performance.
5. There is a need to better understand why differences between predicted and measured energy use exist, how design and operational improvements can be made and how existing buildings can be improved using building energy models.
6. There is a lack of understanding on how the level of operational data granularity affects the accuracy of model calibration and the benefits of automated model calibration.

The first research gap was addressed by exploring industry perspectives on delivering building performance, through interviews and group discussion on what building performance means to different stakeholders, understanding their incentives and the barriers they come across. The second and third research gaps were addressed through case research, investigating the discrepancy between predicted and measured energy use in four case study buildings. Which focussed on understanding how data granularity affects the accuracy of calibrated models and explored the potential benefits and drawbacks of automated calibration, with a direct comparison to manual calibration of building energy models.

The literature review identified that there is a discrepancy between predicted and measured energy performance and concludes that widespread views of its magnitude are generally exaggerated and unjustified, it proposes the use of a classification of three different ways of comparing predicted and measured energy use. It acknowledged that a static or regulatory performance gap exists, which is of great concern to the industry. Therefore, it reviewed and analysed different underlying causes for these performance gaps. It identified several necessary steps that need to be taken to reduce the energy performance gap, but simultaneously recognised that there is a link between building energy performance and other environmental, economic and social aspects of performance, which are of importance to a variety of stakeholders in the building life cycle. This led to a partnership between the author and the UK Green Building Council to investigate how better building performance can be delivered from an industry perspective. This exploratory study incentivised a detailed study of differences between predicted and measured energy use.

3 INDUSTRY PERSPECTIVE ON DELIVERING BUILDING PERFORMANCE

This chapter presents an exploration of industry practices and stakeholders that identified common barriers to delivering high building performance throughout the building life cycle, how such barriers can be overcome and how stakeholders need to be engaged. This work expands upon the main research aim, which focusses on building energy performance. It does this to highlight the fact that the energy performance gap is a far broader issue than just energy and is influenced itself by many other less technical factors as were established during the literature review. In particular, it looks into economic, social and environmental performance aspects in delivering building performance taking a wider view of building performance and what it means to different stakeholders in the building life cycle. Delivering building performance leads to increased asset value, productivity, well-being and energy efficiency. The author in partnership with the UK Green Building Council and a group of its members explored how the construction industry currently designs, constructs and operates non-domestic buildings. It did this through semi-structured interviews with industry experts and round-table discussions to examine industry approaches, tools and behaviours. It focussed on how to maximise building performance, not just in energy terms, but also other aspects of performance that impact both the building user and the wider environment. It found that there are several key factors that determine the success or failure of project to deliver reliable building performance; (1) the aspiration for delivering performance in a project, (2) control of the delivery process, (3) need to design for performance, (4) feedback loops and (5) an increase in knowledge on all levels.

3.1 Methodology

The main research focussed on building energy performance, but the term building performance itself is far broader, including economic, social and environmental performance aspects. As such, building performance simply describes how well a building functions against its needs, often not effectively being met. So, *"how can the construction industry deliver better building performance and more reliable outcomes?".* The author partnered up with the UK Green Building Council (UKGBC) and brought together a group of industry experts to seek out and highlight process improvements that design and construction professionals, property developers, as well as occupiers might adopt to deliver buildings which perform as expected in operation. The UKGBC membership represents all stages of the project life cycle, so it can provide an important role in connecting different stakeholders, ensuring they understand the challenges, and encourage them to adopt good practice solutions. Through semi-structured interviews and round-table discussions with industry experts supported by desk-based research, behaviour and processes were examined across the built environment that affect building performance. The work aimed to identify key barriers to performance and how to overcome these barriers. Focussing mainly on the commercial sector, as opposed to public sector, but much of the analysis is relevant to both.

This chapter tackles the first objective of this thesis by providing an exploration of industry practices in relation to building performance. This was necessitated by the underlying findings in the literature review. Although, the identified causes of the energy performance gap are predominantly of a technical nature, it can be inferred that many causes related to the interaction between stakeholders in the building life cycle and the meaning of, and aspiration for building performance. These are qualitative aspects related to building performance and indirectly the performance gap. This chapter discusses building performance in a much broader sense, but does highlight several key factors that need to be adhered to, to deliver building performance. Of which two (design for performance and feedback) in particular support the need for the other objectives of this thesis.

3.1.1 Literature review

An initial literature review was carried out to start answering the research question in advance of the interviews. The literature review focussed on understanding typical barriers and success factors, and tools and processes that support delivering building performance. In addition, it aimed to provide an overview of all the stakeholders involved in the building life cycle and how they are related and all have an important role to play in delivering building performance. The literature review resulting in several handouts focussing on these aspects which were used during the interview process to map additional insights from the interviewees. Part of the literature review carried out within the UKGBC project has been incorporated in the wider thesis literature review in Chapter 2.

3.1.2 Interviewee selection

Criteria for selection

In accordance with the purpose of this study, the UKGBC were looking for a range of interviewees to cover participating stakeholders in the whole building life cycle. In addition, the interviewees were to be in a position of seniority which helped to ensure they had a good understanding of the concept of performance within buildings and were familiar with interdisciplinary processes and practices.

Methods of contact

Due to the UKGBC being a non-profit member organisation, its first point of contact were its members, whom are familiar and interested in concepts such as ‘Green’ building and sustainability within buildings. This resulted in some form of inevitable bias that was slightly alleviated by further reaching out to non-members, but which were less likely to participate. As a result, more than 70% of interviewees were UKGBC members. In total, over 50 people were approached by email, 15 of those responded positively, they were interviewed either at the UKGBC office or the interviewers travelled to the interviewee’s choice of location.

Interviewee sample

Interviewees were invited from a range of backgrounds, in total 15 industry experts were interviewed of which three were architects/designers, three were investors/developers, three project managers/contractors, three owners/occupiers, one facilities manager, an independent consultant and a proprietary software provider. The sample size was deemed to be sufficient for the purpose of this research, this was supported by ensuring that the interviewees were the holders of knowledge in the area of investigation. In particular our qualitative sample size was based on the idea of “saturation”, the point at which no new information or themes are observed in the data (Guest, et al., 2006). During the interviews it was realised that common patterns were surfacing, as the established research questions gave rise to similar answers. The recurring themes allowed the author to establish several key factors to answer the main research question based on the recurring barriers and solutions mentioned. Furthermore, due to limited resources within the research team and the amount of time available to carry out interviews, the total sample size had to remain feasible, whilst ensuring that a broad range of backgrounds was covered.

3.1.3 Interview process

Semi-structured interviews of 15 industry experts from different stakeholder groups in the building life cycle were carried out. Asking some pointed questions about specific subjects, whilst maintaining a common theme throughout, encouraging two-way communication. Industry perspectives were explored by asking about different aspects in the building life cycle, more specifically topics and questions explored were;

1. Skills, people, culture, tools, processes and practices
    a. Who are the key stakeholders during the formulation of a project brief?
    b. What does best practice look like when buildings perform as desired?
2. Business case and barriers
    a. What would be the business case to pursue a high performance?
    b. What do you think are major barriers to delivering performance?
3. Definition and aspiration for building performance
    a. How do you describe building performance?
    b. What needs to happen in order for people to start considering performance based design over performance driven by regulation.
4. Building life cycle processes with regards to performance
    a. Where do you see the key risk areas in a building life cycle?
    b. Do you think the procurement route influences the pursuit for delivering performance?

In addition, some questions were asked that focussed on understanding what tools, processes, performance indicators, case studies are out there and were being used that can effectively benefit delivering better building performance.

During the semi-structured interviews, the interviewers used handouts, which showed different barriers and success factors, tools and processes aligned to the different life cycle stages based on previous research (literature review). This proved useful in talking about different topics, and helped mapping their experiences and knowledge to the pre-made diagrams on the handouts. Rudimentary diagrams were used of both the s-curve visualisation shown in the previous chapter in **Figure 2.5**, and an overview of the RIBA stages as shown in **Figure 3.3**, which highlighted some of the barriers to delivering performance based on initial findings. Talking around these diagram helped in further developing an overview of barriers and success factors, which resulted in identifying several common themes throughout, in particular the 5 "key factors". In addition, other questions focussed more on the broader understanding of what building performance means to different stakeholders and how that related to these key factors. Quotes have been used to highlight certain points throughout, to provide context, references to the quoted interviewees are shown in Appendix B, describing their role and backgrounds.

3.1.4 Round-table discussions

The round-table discussions involved a group of 10 senior industry members from different stakeholder groups to debate on previous research and findings from the interviewees throughout the project. The industry members were from a range of backgrounds; two architects/designers, a developer, two sustainability consultants, a product manufacturer, a contractor, an asset manager, a software provider and a UKGBC consultant. The project researcher prepared and presented findings and proposed different concepts on how to frame the findings in several diagrams and suggested the initial structure of the report, the former were implemented and the latter was adjusted in accordance to UKGBC style of reporting.

3.1.5 Involvement

The author was involved as a project researcher and was supported by an external consultant to the UKGBC. The project researcher and consultant have led this research by inviting UKGBC members for interviews and setting up regular meetings with a 'task group', which consisted of 10 senior industry members. The project researcher, with the consultant, conducted 15 interviews with external industry experts. The project researcher conducted a literature review prior to the interviews and round-table discussions on performance based procurement and benefits of delivering better building performance. Based on the literature review, interviews and round-table discussions, the project researcher wrote the initial draft of the report. Both the external consultant and UKGBC's employees thereafter revised the draft and produced the final publication. The author's background and work on the energy performance gap may have resulted in some inherent bias in carrying out this research with the UKGBC.

3.2 Key factors to deliver building performance

By exploring what companies are already doing to address the topic of building performance and seeking out best practice, different gaps and barriers were identified that need to be overcome across the whole industry. This chapter is structured around five key factors that determine the success or failure of projects to deliver performance that is more reliable. Analysis of the research and interviews has established the following five inner-related key factors:

1. **Aspiration**

   Expect a building that performs as required in use. Setting a simple target – at the very least for energy use (kWh/m^2) – should help create a common language and shared aspirations across the delivery process. This is relevant to all sectors, but it is particularly incumbent upon investors and developers to drive this.

2 **Control**

Contractual control throughout the delivery process is crucial. Collaborative contracting, with performance guaranteed and control maintained throughout the delivery process helps to ensure predictable outcomes. Again, investors and developers can set expectations, but those in the supply chain should take greater ownership during procurement. There is a role for lawyers to support these aspirations, rather than revert to a default position of least-risk.

3 **Design for performance**

Do not design simply for compliance. Performance improves when aspirations are not limited to compliance or, in other words, “going for the ceiling, not the floor”. There is a responsibility on architects and engineers, not just their clients, to educate and advocate, making the business case for higher performance – including the benefits of sustainable design on staff productivity.

4 **Feedback**

Reciprocal links and a commitment to monitor and feedback, particularly during the handover process, is vital. So too is giving time for well-documented building commissioning. Links have to be made between operational facilities management and the design team, and between FM and building occupiers. By definition, there are shared responsibilities across the value chain, particularly during the handover process.

5 **Knowledge**

Improved knowledge is needed across all professions in order for each part of the supply chain to play its part in delivering building performance. Every organisation has a responsibility to assess the knowledge levels of staff, involve HR teams and identify training needs. Organisations also need to participate more openly in lesson-sharing activities.

The applicability of these factors across the RIBA life cycle stages is mapped in **Figure 3.1**.

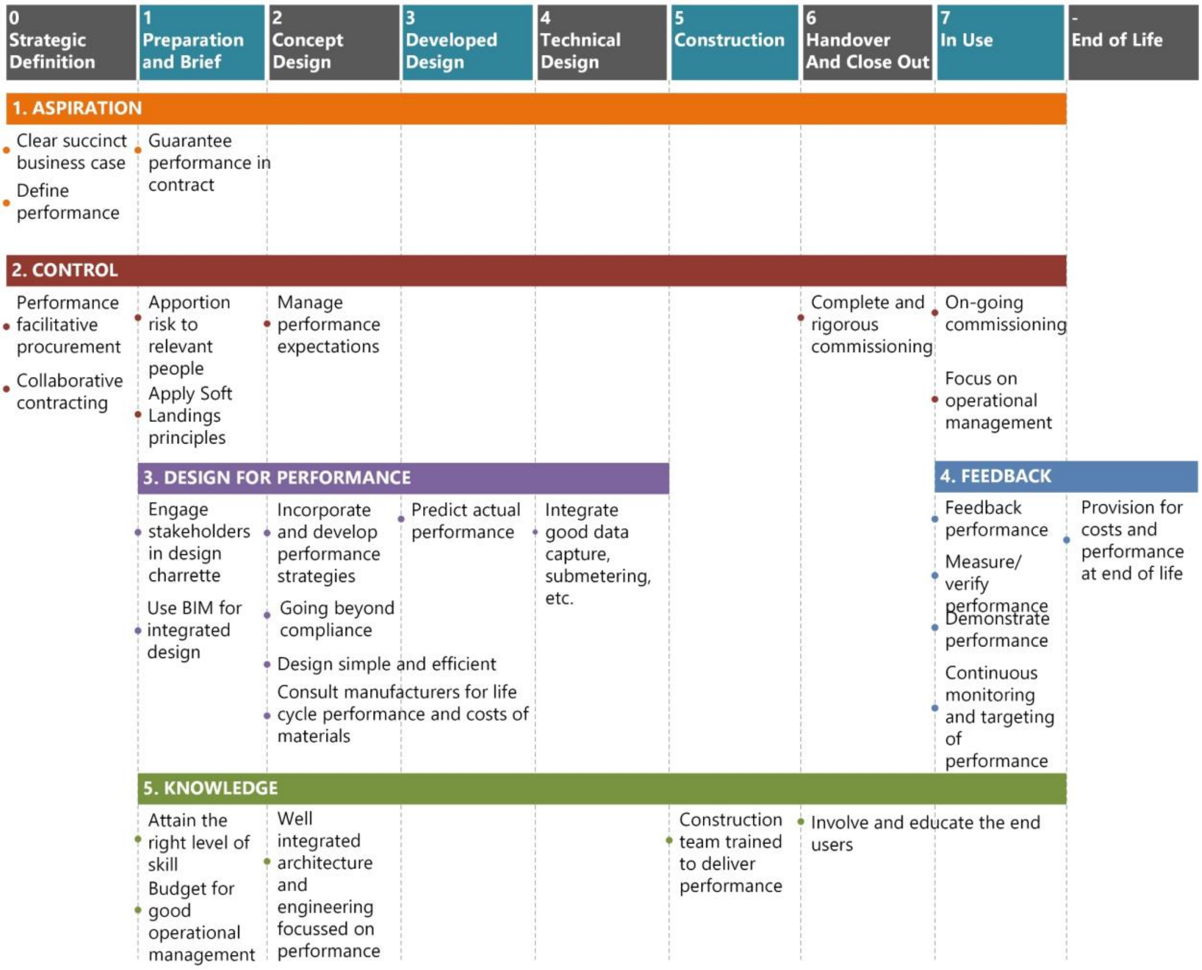


Figure 3.1: Factors ensuring delivery of reliable building performance, aligned with RIBA life cycle stages.

3.2.1 Aspiration

Perception of risk and value

Several of the interviewees noted that there is misconception in some design teams that a good building performance specification is more expensive than a standard specification. Buildings that perform well do not need to be complex. This has repeatedly been found in post occupancy evaluation studies, from the early PROBE programme (Cohen, et al., 2001) to the recent Innovate UK Building Performance Evaluation programme (Palmer, et al., 2016). Burman (2016) suggests that building performance can decrease with building complexity, and that more expense does not necessarily lead to better performance.

There is also a need to deal with perceived risk in the supply chain. Sustainable buildings are generally perceived to have increased risks due to lack of experience in using new technologies and processes, which risk projects going over budget, taking longer to complete and creating room for mistakes (HEEPI, SUST and Thirdwave, 2008). Without certainty for enhanced returns or profit, stakeholders throughout the supply chain are conservative and stick with what they know. The business case for delivering building performance is therefore critical to raising aspirations. Enhanced returns are, after all, one of the most important motivators for many business decisions.

The Green Construction Board (2014) has suggested there is little belief in differential market pricing, especially in buoyant markets where 'anything will let'. Interviewees for this report believed that many decision makers are not actually aware of the benefits or are not convinced of them. In addition, one interviewee (INT1) in particularly pointed out that *"Decision makers are hanging back to see what their competitors are going to do, nobody is really embracing change as it is not the nature of our industry."*

Whilst the industry seems reluctant in moving towards an environment with better building performance, new regulations keep taking shape. The Minimum Energy Efficiency Standards (MEES) regulations will start having an impact on the performance of the building stock. Properties with the lowest EPC energy ratings of F and G cannot be let anymore, subject to certain exemptions (HM Goverment, 2017). Although this relates to design rather than operational performance, it is raising the overall profile of energy performance in the minds of owners and investors. In addition, agents and tenants understand that a good building shell is more likely to lead to better operational performance, so price chipping is beginning to occur on poorer performing assets which are in need of substantial investment to bring them up to minimum standards.

Defining performance

Building performance simply describes how a building functions against needs. Those needs vary – for each of the parts of the building life cycle, building performance means different things. **Figure 3.2** shows the aspects of performance that the task group representatives felt to be relevant at different stages of the life cycle. Stakeholders have different perceptions of building performance, but they can complement each other, for example, better environmental and social aspects in a building will support economic values, creating the business case for decision makers.

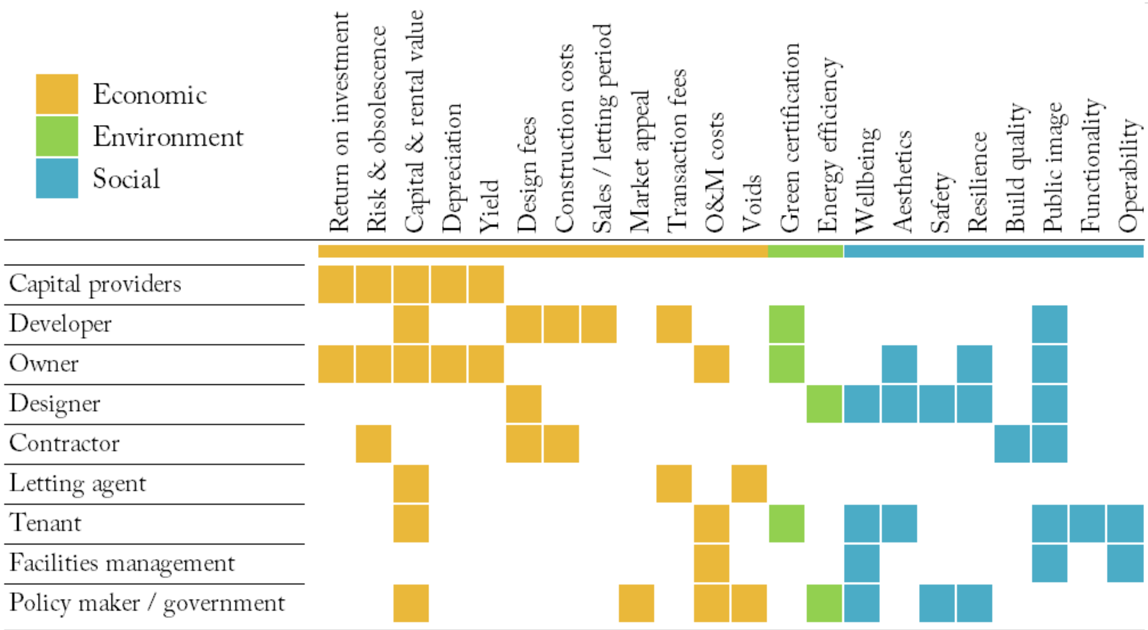


Figure 3.2: Building performance for different stakeholders, developed during round-table discussions according to task group representatives and based on consensus from interviews and literature.

Demonstrating performance

Demonstrating the value of good building performance will be an important factor in order for tenants/occupiers to ask the right questions and request the right standards for the building spaces they procure (Hsu, 2014). One interviewee noted that occupiers ask about visible performance elements – e.g. cycle space, showers. There is a need to explain the occupier benefits of the invisible building performance elements as well, e.g. air quality, energy efficiency, and to point to tangible benefits. Whereas for some sustainability is a clear path and incorporating it benefits them directly as an interviewee (INT2) at a large development corporation put it, *"Sustainability is being used to secure business, to separate us from other* *developers. Sustainability is fundamental in an increasing challenging environment with increasing resource scarcity and highly complex social and demographic issues."*

A business case exists for delivering better building performance that can benefit all stakeholder in the building life cycle. The main obstacle is communicating the benefits sufficiently compellingly and to the right audiences within mainstream owner, financier and occupier sectors. Case studies targeted at these sectors and cooperation from within these sectors are highly desirable.

Setting a target

The expectation and requirement for building performance needs to be driven through the building delivery supply chain by an organisation that can influence the whole supply chain in terms of both objectives and purpose, and contractually (Pless, et al., 2012). There was agreement within the round-table discussions that the drive for building performance needs to be led by the client – usually the owner, developer or occupier, this is supported by Swarup et al. (2011). The consensus was supported by the interviewees, they regularly cited that architects and engineers have limited influence on the building projects in terms of performance outcomes or criteria set by the client. The design is mostly guided by client's requirements. One of the interviewees stressed the importance of support at the C-suite level as it gives others the confidence to press for the required performance levels and leads to productive engagement with the entire supply chain.

The Boards/Executives/C-Suite of most organisations procuring space (as owners, funders or occupiers) are not sufficiently engaged in the building performance debate, and in many cases nor are their agents and other advisors whose technical knowledge is varied. Yet investors and occupiers are key to driving building performance. They sanction procurement processes, empower their advisors to act in particular ways, and have the ultimate sign-off on decisions. In essence, the building owner and funder have a key role to "push" the requirement for building performance down through the supply chain, and the building occupier provides the "pull" as the end-user.

3.2.2 Control

Delivering good building performance requires control of the procurement and delivery process, collaborative procurement and client commitment to validating operational performance. During the interviews and literature review, the project researcher took note of potential barriers to delivering better building performance in the delivery process. Some of the questions focussed specifically on what these barriers are and where they might occur. During the round-table discussions, the identified barriers were discussed and typical barriers in concurrence with finding from literature are mapped in **Figure 3.3** along the RIBA building life cycle stages. Fragmentation of the supply chain can amplify these barriers and creates many difficulties for the responsibilities stakeholders face.

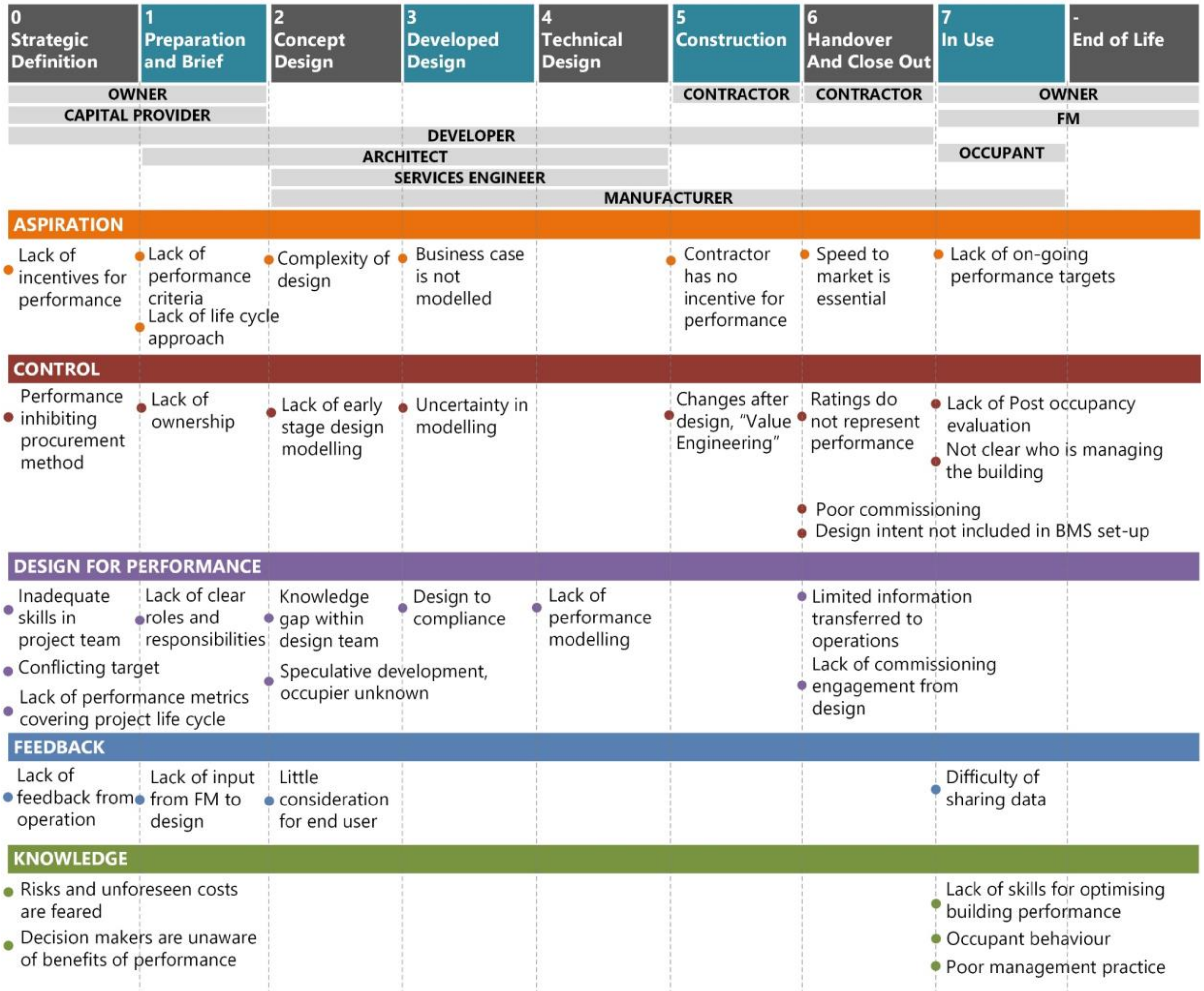


Figure 3.3: Barriers to delivering performance throughout the life cycle stages, mentioned during interviews and identified in literature. These were categorised in relation to the key factors to overcome such barriers.

Fragmented industry

Several interviewees observed that energy consultants are increasingly only required for Part L compliance and energy assessments, and are not asked to provide technical advice, the responsibility for which is passed to contractors. Several interviewees believed that some tend to see technologies as risky, so price them out or pass the specification onto manufacturers who see their specification only in isolation. This is supported by findings from HEEPI, SUST and Thirdwave (2008), who see last minute substitutions of materials or equipment on site as one of the main reasons that significantly affect environmental performance. Involving suppliers as early as possible in the design phase enables them to understand where their contribution fits in to the wider project and to provide the optimal response. It also helps the project to gain the most value from the whole supply chain.

Another interviewee pointed out that during lease negotiations, where legal teams and prospective occupiers do not understand the requirements or benefits of (sustainable) fit-outs, or where green leases are specified, that they may be struck out if considered a risk to the occupier. Additionally, a misfit-out can undermine all of the design done for the shell and core systems.

After a building is constructed, commissioning often falls short and as a result the design of a building is not provided as intended (Palmer, et al., 2016). Systems that are not working correctly deteriorate faster, use more energy and negatively impact occupant comfort when heating and cooling loads cannot be met. Arguably commissioning is one of the most cost-effective strategies to reduce energy use and costs in buildings today (Mills, 2011). Commissioning needs to be done rigorously when the building is handed over, instead it is often done quickly and poorly due to time constraints. In addition, sub-metering and the BMS should be set-up correctly, calibrated and tested in order to monitor and measure performance and be able to identify excess consumption and operational issues. An interviewee (INT3) from a real estate developer observed: *"The biggest hole that we see is from the time that (a property) is handed over from the contractor to the recipient, what you find is that the system is not tested and set up properly. That gets exacerbated when people move in and do the fitting out. This situation effectively erodes any thinking you did during the design." (real estate developer).*

In both design and operation, clients and their design teams should take into account the occupiers' capabilities to manage a building and its engineering systems (Bordass, et al., 2001). Whereas occupiers should not take a building for granted, but need to be aware that management is an important aspect of a good performing building as they ensure that the building its systems are operating and maintained. Most interviewees emphasised the importance of Post-Occupancy Evaluation (POE), and co-operation between the occupier and building owner as being extremely important to get the most out of the design work.

Speculative developments

According to several interviewees, the UK has a higher proportion of speculative developments than Europe, Australia and the US. Speculative development often has a short-term focus, and developers are mostly concerned with minimising cap-ex costs to increase yield and profit from an early sale. The occupier is unknown and building performance above legal compliance is generally not demanded since actors in this type of procurement are not involved in the operation of the building. Location is considered key, yet the building is likely to have a considerable lifespan. This reinforces the function of building regulation to act as a back stop. Systematic gathering of feedback and data from buildings in-use will reveal the financial benefits of good building performance and increase demand from the buyers of speculative buildings and their occupiers.

Value engineering

There was particular concern about problems with value engineering from several of the interviewees and in the literature. A number of interviewees pointed out that it is no different to 'penny pinching'. The purpose of value engineering in construction is to re-design or change construction build-up to remove challenges or save cost whilst retaining the same performance and functionality, therefore following these principles, value engineering should take place before on site construction.

The Green Construction Board found that many schemes initially aspire high performance, but that design characteristics are compromised during the development process to reduce cost (GCB, 2014). In the case of buildings with complex systems, such as low carbon technologies, value engineering does not tend to take out the main technology, but it often takes out the controls or associated design details (e.g. thermal stores) to make it work well. This is partly because the associated equipment was not detailed in the design and tender documents. Additionally, an interviewee (INT1) indicated that at the point where the contractor comes on board, they often get very scant information. *"They look at the information they have got and make an assessment as to how the building is put together, they price it on that basis. They take a risk because they do not have enough information to ensure themselves that compliance can be achieved. Especially in Design and Build (D&B) contracts, where they are required to take on the responsibility from scratch again."*

Procurement models

Academic reasoning indicates the effects a procurement method can have on performance, especially design-build and design-build-operate (e.g. private finance initiative) should provide the right relations and processes for a performance incentivised delivery (Korkmaz, et al., 2010). Regardless of the procurement method taken, a collaborative approach, i.e. Collaborative Contracting or at least its features should be employed, as it promotes sharing of information, risks and responsibilities between parties. In addition, there is common agreement that owner engagement in a project is pivotal to drive and deliver performance, from setting targets and involving stakeholders early on to guiding them throughout the supply chain (Swarup, et al., 2011). As an interviewee observed, no individual party wants to take the risk on their own, it needs to be shared across the project. Another interviewee (INT4) took this further; suggesting, *"It needs to be about empowering different people in the supply chain to take ownership, rather than just saying it needs to be M&E design or architectural design – it needs to be more than that. It is more about embedding the approach within the design process. It needs everyone."*

3.2.3 Design for performance

Limitations of regulations and rating schemes

Legislation on energy performance may not be a sufficient driver for occupiers. Under schemes such as the Carbon Reduction Commitment (CRC) and the Climate Change Levy (CCL) it is often seen as easier to just pay the tax as energy, as energy costs are such a small proportion of total running costs in non-industrial building sectors. The interviews and literature review suggest that there is a much more powerful case to be made through wellbeing and productivity benefits of buildings that perform well, in particular this should be communicated to owners, occupiers and letting agents. Given their lack of detailed knowledge there is an understandable situation in which they assume that good building performance automatically follows from compliance with building regulation or with accreditation schemes such as BREEAM whose sliding scale of accreditation is straightforward to understand. An interviewee noted that there is a tendency to treat regulations as a ceiling, rather than a floor. Even if projects delivered results with a percentage improvement above Part L, most people outside the technical community would not understand what this meant or the relevance to them. However, even when Part L conformance or improvement is delivered, it does not directly relate to operational building performance. Another interviewee (INT4) at a large real estate business, thinks that rating schemes such as *"BREEAM and EPCs might result in unintended consequences. Some of the recommendations that come out, might up diverting the right intentions." (real estate developer).*

3.2.4 Feedback

The final process is the reciprocal link that needs to be established between the end and start of a project, a recurring theme that can directly and indirectly support the other processes. This reciprocal link is the feedback of performance, it is essential that a building performance is measured, verified and demonstrated (Preiser & Vischer, 2005). This raises awareness, understanding and is fundamental for improving building performance in future projects. Feedback mechanisms on performance are not well developed. The design team is therefore limited in its ability to predict the actual use of the building. Presently, there are few methods that allow verification of performance in operation, these are a necessity for the designer to understand how buildings actually operate and how occupants behave in a building.

Measurement and verification

Most interviewees stressed the importance of POE. It can assist in verifying other metrics not related to energy consumption, such as indoor environmental quality metrics and the well-being of occupants (Clements-Croome, 2014, p. 92). Measuring of performance is essential to verify the design goals, comparing initial predictions with actual operation can identify areas that are underachieving, and such information should be fed back to the design and be reconfigured where possible by operational management. Public and portfolio data collection will support benchmark development, fundamental for setting out performance requirements. Although design has a major influence on building performance, as design features become more efficient, occupant behaviour and operational management becomes more significant (Li & Lim, 2013). Tenants are rarely involved in the design of a building. Clients and their design teams must take into account the occupiers' capabilities to manage a building and its engineering system (Bordass, et al., 2001). Conversely, occupiers should not take a building for granted, but need to be aware that commissioning and management is a crucial aspect of a well performing building. One of the interviewees (INT5) noted that FM managers need to understand and be willing to use available technology. She said *"For one of our new developments I am trying to work out if the building is performing to the initial design. The AMR system, consisting of 300 meters including optimum front-end for monitoring and targeting was turned off by FM employed by the managing agents because 'They don't normally do things like this', instead they manually red the meters and key that into their system." (investor / developer).*

Demonstrating performance

Demonstrating performance is another factor that could prove very valuable in the pursuit for delivering better building performance. One of the primary benefits of public disclosure of performance is education of decision-makers, owners and tenants (George & Garrod, 2017). It provides valuable feedback in terms of benchmarks that they can use to inform new projects, support valuation of a property as it becomes more integral to the real estate market and highlight benefits of living and working in a particular premise. This results in property developer to be more aware of its effect on market value and include performance aspects in new designs (Frankel, et al., 2015). Practices should consider the public disclosure of performance on a voluntary basis and need to understand that an underperforming building does not necessarily mean it is failing, rather the monitoring of performance raises awareness for improvements. In addition, data can be used to evaluate the performance of the building stock and enable benchmarking to become more credible and reliable.

The lack of benchmarks in the UK makes it difficult to set credible performance targets, a similar approach of collective data collection and benchmarking used in the U.S. is needed. Here the Commercial Building Energy Consumption Survey (CBECS) provides a comprehensive dataset characterizing the performance of the U.S. building stock. In addition, there is the Building Performance Database (BPD), which is the largest publically available source of building performance data in the U.S.

However, one interviewee (INT6) noted that *"More often than not, in particular in new buildings, people are not willing to hand over data to a collective source."* Most interviewees also stressed the importance of case studies, yet compiling transparent case studies, with genuinely useful lessons learned and data is notoriously difficult.

3.2.5 Knowledge

Knowledge and skills are intrinsic to the key factors discussed above. All stages of the building process have a role to play. The provision of training is crucial for all parts of the supply chain in order that they have the knowledge and capability required to drive higher performance standards. Those already seeking a leadership position will provide some 'pull' to the industry, but to 'push' the majority it needs to ensured that key players have a solid understanding around sustainability in the built environment, and the relevance to their role and organisation in delivering solutions.

Financer/owner/occupier

The financer/owner/occupier can have a large effect on the performance outcome of a building. The business case for good building performance needs to be better communicated, and a better understanding of how building performance can be defined in the brief and what can be achieved is necessary. This research and the interviews align with Carbon Trust recommendations that the financer/owner/occupier should consider appointing a specialist with responsibility for driving performance throughout the project (Carbon Trust, 2012). To support the uptake of well performing buildings, developers/owners need to be engaged to adopt new methods for capturing a property's value.

Letting agents

Letting and other transactional agents such as surveyors play a critical role in building performance as it is they who value and market buildings, an interviewee (INT7) observed that they *"shape the appeal of buildings to potential investors/owners and occupiers."* Other interviewees noted that there is a need to build knowledge about the business case within the letting agent community and to get them on board. One interviewee specifically observed that letting agents advise owners how to refurbish space and build new space, and then advise tenants about what to ask for. They therefore drive the market from both sides, and could really change thinking if they started to give advice which improved building performance.

Operations staff

Operations staff need to be motivated, skilled and incentivised to realise the desired levels of building performance. They are an important stakeholder in the supply chain, but tend to be under represented. It is essential to communicate the business case for good building performance to those in the organisation responsible for budgeting building operation. However, as Frankel et al. (2015) suggest, operational teams in organisations are given limited resources to carry out the needed activities to attain a running well-performing building, and budgets available in organisations do not reflect the actual investments required for effective operations.

An interviewee remarked on challenges associated with the consistency of training across Facilities Managers, particularly with personnel changes following initial building handover. Another interviewee identified a similar issue and now record and video commissioning and training so a permanent record is available and accessible to staff.

Tenants

Considering tenants are the layman in terms of building performance, it is important that they are educated in using the building and its systems to satisfy their own needs of health, comfort and safety, effectively, efficiently and safely. There is a lack of understanding by occupants/operators of how systems are designed to be used, especially so when designers and contractors are not involved after the building is completed. In many cases, the performance of a system is largely dependent on the engagement by building operators and tenants. However, one interviewee noted that tenants are increasingly setting their own energy performance targets, although reporting on metrics of indoor environmental quality and occupant satisfaction remain relatively rare.

Another interviewee noted that it would be ideal if every business had an environmental manager or dedicated resource for environmental issues. Many small businesses do not and this makes them harder to engage on building performance. Occupants have a substantial influence on the performance of a building by handling controls, such as those for lighting, sun shading, windows, setpoints, and office equipment. People are different and their behaviour in relation to energy consumption varies. Managing this requires tenant organisations to undertake awareness raising of their own – and to make their own business case for staff to accept that building commissioning will create a better medium term result than a quick "DIY fix" of altering settings or bringing in under desk heaters.

Construction

On site workmanship needs to adapt and be trained to increasing levels of complexity in building construction. An interviewee, who was one of several who noted that BMSs were rarely set up properly. This is repeatedly found in post-occupancy evaluation studies, including the PROBE (Cohen, et al., 2001) and Innovate UK's BPE programme (Palmer, et al., 2016). An interviewee (INT8) jokingly mentioned that *"buildings that do not report problems with the BMS are typically those that do not have one." (software provider).*

The whole supply chain requires collaborative design skills in order to determine the performance targets for the building and to work together to deliver them. They also need to establish feedback mechanisms to better understand how buildings performance and how technologies and processes affect this performance. 'Soft' influencing skills will be key to encouraging clients to undertake data sharing. Related to this a number of the interviewees noted the "translation" issue – being able to talk a language that clients and occupiers understand and making the building performance debate comprehensible and relevant across the industry.

Overcoming complexity

Information needs to be in the right language for owners, financiers and occupiers. Interestingly several of the people that were interviewed made the point that those engaged in the building performance debate tended to speak to each other rather than the whole industry, and to use a complex approach and language that was hard for those not engaged – but in the same industry - to understand. Soft landings was described as too complicated, some research critiqued it as making a theoretical point rather than relating to understandable situations, and as one interviewee (INT1) told us: *"Industry struggles with non-quantitative issues that keep changing. There are layers of guidance, documentation and policy. We need something that is going to make all of that easier, not add to it." (contractor).*

3.3 Summary

An exploratory study of industry perspectives of delivering building performance identified common barriers through 15 semi-structured interviews and separate round-table discussion with industry experts. As such, this study contributed to knowledge by collating relevant first-hand experience from industry experts on common barriers to delivering building performance. In addition, it developed five key factors that need to be adhered to deliver reliable building performance. Published as an industry-focussed report, it should help prospective stakeholders in the building procurement process understand the common barriers to performance, help set goals for delivery, and improve relevant stakeholder processes in line with these goals.

This exploratory study viewed building performance simply as how a building function against its needs, which helped to understand the broader aspects in the building procurement process, which also affect the energy performance of a building and indirectly the energy performance gap. This study has supported some of the initial findings from literature, especially concerning the collection of data, need for case studies, demonstration of performance, measurement and verification processes and operational management. However, it also identified many other facets of the building procurement process that will influence energy performance and the difference between regulatory or performance modelling predictions and measurements.

The case research, the main research in this thesis aligns with several of identified key-factors, in particular around design for performance and feedback. First, it does this by providing comprehensive examples of measured energy performance of existing buildings compared to detailed performance modelling, in addition to utilising calibration (i.e. demonstrating performance and measurement and verification). Second, it investigates the effect of simplifications of modelling assumptions on energy performance predictions and explains the limitations of regulations in the context of the energy performance gap. Third, it explores the use of detailed measurements to identify patterns in energy use and to inform design assumptions (i.e. feedback). It demonstrates how energy performance can be better predicted during the design stage by proposing techniques for incorporating uncertainty in assumptions and more representative assumptions of actual use. Four case study buildings are presented which reiterate the value of complete and rigorous commissioning and importance of operational management in the operational stage of a building and highlights some key aspects of the disconnect between design and operation of buildings. The case research investigates a few of the key factors identified, further work will be necessary in particular around some of the more qualitative aspects, such as knowledge, aspiration and control to ensure more reliable building performance can be delivered.

4 CASE RESEARCH METHODOLOGY

This chapter describes the development of a calibration methodology that can be employed to investigate and mitigate the discrepancy between predicted and measured energy use. It advances the state of the art in building model calibration by incorporating meta-model multi-objective optimisation at a high level of data granularity. The methodology explains the processes used for data collection, synthesis and establishment of different levels of data granularity for comparing energy use. It goes on to describe the creation of building energy models for calibration to measured energy use. Furthermore, it outlines how input parameters are sampled for randomised parametric simulation, fundamental to global sensitivity analysis approaches, meta-model development and automated calibration. It then explains how different sensitivity and uncertainty analysis techniques were utilised in order to quantify the uncertainty in the model, investigate the impact of the underlying causes of the energy performance gap and determine the influence of typical design assumptions. Finally, it presents different data visualisation techniques to improve the calibration process and understanding of differences between predicted and measured energy use.

4.1 Case research

Several types of methods exist for testing the performance of building energy modelling. Analytical verification compares modelling results with an exact known solution, and an inter-model comparison method compares predicted energy use using different software, while empirical validation compares predicted with measured experimental (idealised) or field monitored data (realistic). The analytical verification approach is only able to provide a comparison to an exact known solution for a specific problem, whereas an inter-model comparison can only identify differences between simulation software, which can indirectly justify differences with measured energy use. An idealised validation method would be to model a test cell in order to study both the performance of a building and used modelling tools, idealised validation however does not take into account occupant behaviour in a realistic manner. The only approach that is able to study and capture all facets of a real building is the empirical or realistic validation method. This method was employed to investigate differences between predicted and measured energy use.

The disadvantage of using empirical validation however, is that it does not allow for rapid investigation of a multitude of buildings. This is due to the large amount of detailed information that needs to be gathered, compared and analysed. This process can be made more time efficient by employing software to support in automating the comparison and analysis of data. Energy auditing is one of the key methods in order to establish data collection. It is used to identify how buildings are used, and how parameters, such as occupancy presence, differ from typical design assumptions. It aims to provide detailed information of energy flows for the whole building, including major sub loads such as lighting, heating, ventilation, air-conditioning and equipment. A discrepancy can be identified by representing the operation of a building as accurately as possible by using advanced and well-documented simulation tools.

Previous research efforts have primarily focussed on monthly calibration for electricity and gas consumption for the whole building (Pan, et al., 2007), although others have proposed the use of hourly values instead (Raftery, et al., 2011). When measurement data is available on a floor or zone level, calibration can be performed at that level. Furthermore, instead of using aggregated monthly values, it is hypothesised that hourly or even half-hourly calibration improves the accuracy of a calibrated model. The availability of additional information from the Building Management System (BMS) provides the opportunity to calibrate a model on specific parameters such as setpoints, operational schedules and room temperatures instead of solely on energy use, further increasing the accuracy of the virtual model. This increases the accuracy of predicting energy conservation measures, which are generally based on calibrated models. A methodology has been established for comparing predicted and measured building energy use, it involves the following tasks:

1. Data collection, both design information and actual operational procedures need to be identified in addition to performance data from sub-metering, building management systems, environmental monitoring and potentially other monitoring systems;
2. Data synthesis, large amounts of data can help establish a good understanding of how the building works and is operated, but requires data synthesis and quality assurance. A hierarchy is established to compare data at different levels;
3. Modelling, after collected data is synthesised, the initial model is created, predicted performance from the model is compared to measured performance to identify discrepancies. In addition, parametric simulation and sampling of inputs is performed;
4. Sensitivity and uncertainty analysis is performed to understand how input parameters affect the outputs and to quantify the uncertainty in the model output given the uncertainty in the input, respectively;

5. Manual calibration of the initial model, to assure that the model is defined accurately according to building design specifications identified through data collection and synthesis.
6. Quantify the impact of underlying causes of the regulatory performance gap through introducing NCM assumptions to the manually calibrated building energy models.
7. Automated calibration utilising meta-model based optimisation to quantify the impact of data granularity on model calibration accuracy.

The whole methodology or part of the methodology was applied to four case study buildings, where virtual models of the buildings were created based on collected information and energy audits, predicted energy use was compared with measured data.

4.2 Case study buildings

Four case study buildings were used within the case research; two university buildings (referred to as CH and MPEB) and two office buildings (referred to as Office 17 and 71). The buildings are located in London, the United Kingdom. Access was provided by University College London (UCL) estates and BuroHappold Engineering, both parties were directly involved with the research. The buildings consist primarily of open-plan office space, where both university buildings include some teaching spaces and MPEB includes workshops and laboratories, in addition to two large server rooms. Office 17 is a naturally ventilated building with a provision for air conditioning in the basement and reception. CH and Office 71 are also naturally ventilated, but have air conditioning throughout the building. MPEB has a variety of mechanical systems in place to provide ventilation and air conditioning, in addition to openable windows for natural ventilation in perimeter spaces. The buildings have very dissimilar building physical properties, as they are a variety of ages. For each building, a short description is given of its building materials, layout, function and HVAC system, with further detail provided in the Appendices.

Building selection

Buildings were selected based on their accessibility and availability of both design and measured data at a high level of granularity, which is in line with the aims and objectives of this research. In particular, this proved essential to quantify the impact of different levels of data granularity on model calibration accuracy, in addition to understanding the underlying causes of the regulatory performance gap. The sample used in this study is not meant to be representative for a part of the UK building stock. Instead, the selection aligns to the main objectives of the study, which are concerned with the feasibility of the calibration methodology to understand the effect of high-granularity data collection and analysis, in addition to understanding how such data can be used to juxtapose predicted and measured energy use. The selected buildings are located in London, the United Kingdom, listed in **Table** 4.1.

Table 4.1: Summary of selected case study buildings

Building	Built	Space types
CH	~1900	Open-plan offices, library, lecture rooms
MPEB	2005	Teaching, workshops, CS, offices
Office 17	<1930	Open-plan offices
Office 71	1950-60s	Open-plan offices

The selected case study buildings have varying functions. The two office buildings consist primarily of office space. CH is similar to an office building except that it also provides student working areas and a library. MPEB houses workshops, teaching spaces, computer clusters, laboratories and office spaces.

4.3 Data collection

The collection of data consists of two main categories; (1) modelling information, necessary for the creation of the as-built model, and (2) operational data, the measurements used to calibrate the as-built model and any other measurements that used to identify patterns of use and performance of systems.

Modelling information

Ideally a design model would be available as a starting point, this will allow comparing initial assumptions with the calibrated model, however these are often not available as these are generally not disclosed by the design engineer. In the buildings under study these have not been looked into and models were built from the ground up based on collected information. Missing information from as-built documents were supported by measurements, surveys or otherwise using building standards and codes or typical benchmarks. Much of the modelling information was obtained from Operation & Maintenance (O&M) manuals, which contain architectural and building services drawings, sub-metering schematics, equipment inventory lists, and commissioning data.

Operational data

Information on the buildings was reviewed prior to the auditing period. Reviewing the collection of such information is necessary to be acquainted with the building and its systems, making the auditing process more efficient. The auditing process helped in further understanding the general state of the building and its systems in order to verify if it works as intended and described in the operation and maintenance (O&M) manuals. It also ensures the collection of missing information. In the case of insufficient metering data, a bottom-up approach such as the CIBSE TM22 (CIBSE, 2006) method can be used to determine end-use consumption. This approach was taken to establish some of the assumptions, for example for lighting and power, the number of fittings, lamps and appliances were counted to determine the installed load. This approach is time consuming, but proved to be necessary for estimating the small power load and verifying the installed lighting. Several tasks that were essential during the energy audit of the building:

- Check if previously obtained information deviates from the actual building, such as building drawings, HVAC system and components,
- Interview the building operator/facilities management, whom can provide detailed information on building operation aspects and system performance,
- Taking photographs and notes.

Most modelling information can be determined from the O&M manuals and energy audits, however to capture more sophisticated stochastic processes (such as occupancy presence and system operation) operational data is necessary to accurately represent these in the models. For the two university buildings, Wi-Fi and swipe card access data was obtained and used to establish when people are present in the buildings. Building management system data was analysed to understand when systems are in use and how they perform. Additional short-term monitoring of electricity was necessary where the existing metering systems did not account for certain energy end-uses. Environmental monitoring of temperature and humidity in some of the buildings was analysed to understand how spaces are conditioned and when systems are operated. The high level of data granularity established in this study was essential in understanding how it affects model calibration accuracy.

4.4 Data granularity

To make sense of the large amount of data that is available a spreadsheet was created for each building separately to be used an index for information. Data for every space in a building can then be easily filtered and adjusted. Building input parameters inferred from the as-built documentation were then ordered for each building similarly to create an efficient workflow for analysis. This proved helpful to automate certain processes, such as input parameter adjustment and parametric simulation. Ordered data was assessed according to different levels of data granularity; hierarchical, temporal and spatial.

Determining the necessary level of data granularity depends on the purpose of model calibration. Typically, calibrated models are seen as obtaining the highest level of accuracy for determining energy savings, option D as proposed by the IPMVP (EVO, 2009). However, this depends on what type of savings can be made and potential measures implemented to make these savings. If these are fairly straightforward or easy to estimate, model calibration might not be the best solution, because it is a time-consuming and therefore costly process, so any improvements in its use are helpful in reducing costs. Finally, calibrated simulation is typically applied by multifaceted energy management programs affecting many systems and where energy use data for calibration is available (EVO, 2009, p. 22).

Data was organised systematically to compare measurements with predictions and assumptions from energy modelling. Representing all this data requires a hierarchy. Predicted and measured energy use can be compared at the top-level by aggregating all consumption in a building. A building however consists of floors, zones and spaces and has distribution boards that provide electricity to these parts, disaggregating energy use by different end-uses. For example, it is common to see electricity being measured separately for lifts, HVAC systems and perhaps even a kitchen, preferably electricity for lighting and equipment is distinguished at a per floor basis. When automatic metering readings were available at this level of detail, comparisons will provide more insights in how the virtual model is performing and helps identifying any occurring performance issues. However, energy use at the zone-level is generally not available. It is however more common to see indoor environmental parameters being measured in zones or spaces through a BMS.

All this information is processed at a spatial level, environmental parameters such as temperature and light levels are important to a zone, but are less useful at building level. Whereas, equipment and lighting energy use can be both compared by zone, floor or building level. Finally, building systems and their components aim to provide the right environmental conditions in the building and its zones, for which they use energy. Systems are located in a separate plant area of a building, whereas system components can be located in a specific zone. Measured energy use of system components can then be compared like-for-like in the virtual model at the zone-level and be aggregated at the system level. The building hierarchy (at the top) and two examples of comparing predicted and measured data are illustrated in **Figure 4.1**.

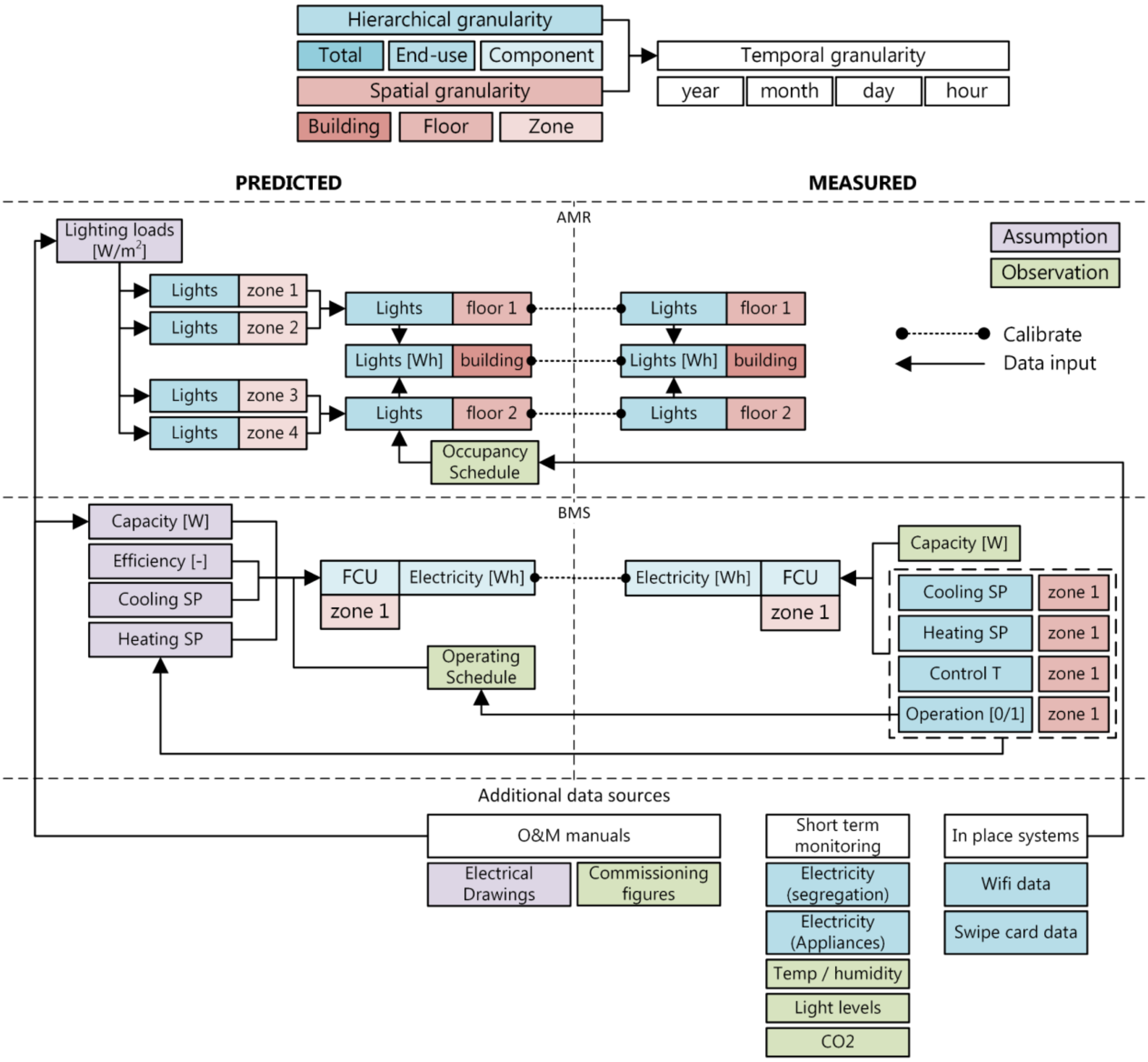


Figure 4.1: Data granularity where two examples illustrate how predicted and measured data is compared and where different data sources were used to inform model inputs.

The figure shows how automatic meter readings for lighting were compared to predicted data on a floor basis and if available at a zone basis. The predicted lighting energy use is dependent on the lighting load and occupancy schedule design assumptions. Also shown is a comparison of electricity use of a fan coil unit in a particular zone. In addition to comparing electricity use, the BMS provides information on the environmental of the zone and performance of the system component. Based on a comparison, any performance issues were identified at a detailed level, whereas any constraints for the design parameters can be set in order to guide the calibration process. It is thus important to understand how data granularity affects the accuracy of the calibration model at different levels, a distinction is made between temporal, hierarchical and spatial data granularity, as shown in **Table 4.2**.

Table 4.2: Types and levels of data granularity in building modelling and measurements

Granularity type	Low to high level of granularity						
Temporal	year	>	month	>	day	>	hour
Hierarchical	total	>	end-use	>	component	>	sensor
Spatial	building	>	floor	>	zone		

4.5 Modelling and calibration

4.5.1 Procedure

A methodology is developed based on previous work as described in the literature review and is further extended to fit the research objectives. There was a need to include a way to identify and estimate the impact of the underlying causes of a discrepancy, whilst considering the building hierarchy to allow juxtaposing specific predicted and measured data points. The manual calibration of the model uses an evidence-based decision making process whilst employing sensitivity analysis to estimate the impact of design assumptions and to quantify their effect on a discrepancy. Furthermore, automated calibration was used in conjunction with meta-models to replace the first-principle software. A flowchart of the modelling, calibration and analysis process is shown in **Figure 4.2**.

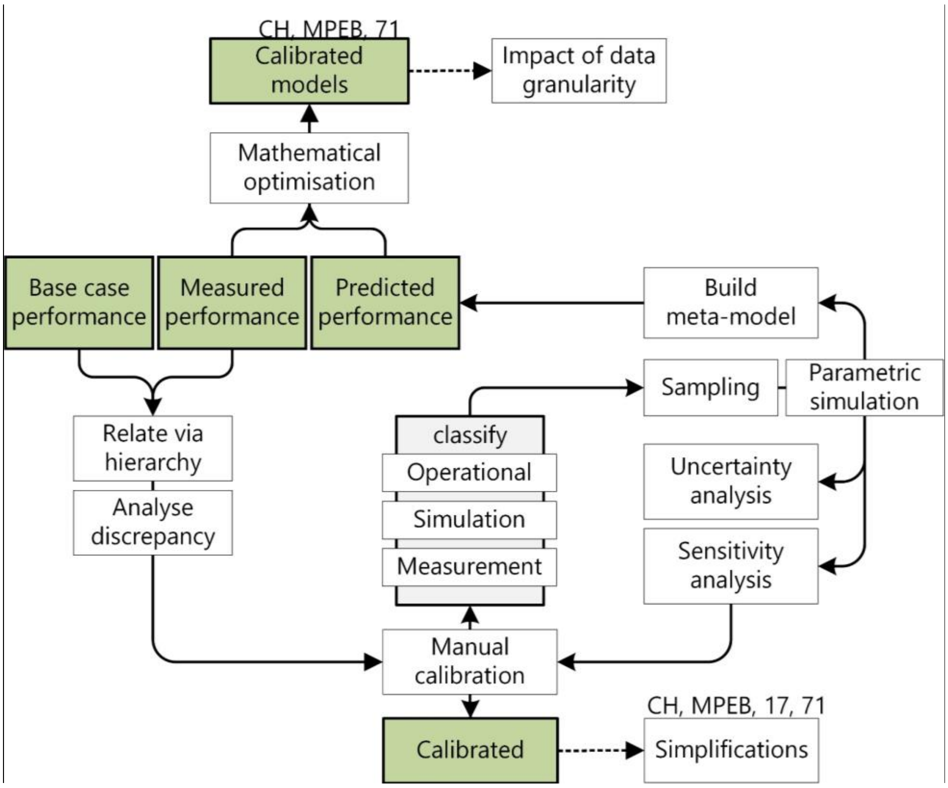


Figure 4.2: Modelling, calibration and analysis process.

4.5.2 Modelling

After collecting and synthesising building information, input parameters were established to create a virtual representation of a building. 3-dimensional geometrical models were created in SketchUp (2016), all building spaces were included in the model and zoning of the spaces is identified during audits. SketchUp[2] incorporates a plugin for OpenStudio[3] (1.14), which is a graphical interface for the EnergyPlus[4]

[2] https://www.sketchup.com/

[3] https://www.openstudio.net/

[4] https://www.energyplus.net/

(8.6) simulation engine. Both EnergyPlus and OpenStudio are open-source software with an increasingly supportive community, they allow good integration with other software tools, have additional integrated features, such as scripting to pre-process and post-process simulations, where OpenStudio is packaged with tools for parametric simulation. The graphical models were created in SketchUp, and then further defined in OpenStudio by changing the input parameters for construction properties, internal loads, system profiling, etc.

4.5.3 Parameter classification

An initial model will have myriad input parameters that are determinant of the final results of the simulation, at this point in the process it is however unclear what influence each of these parameters would have on the result. This influence can be determined with sensitivity analysis by quantifying the uncertainty in the input and running multiple simulations. However, varying all possible input parameters will result in many different scenarios that need to be run, which for large models is too time consuming due to its computational intensiveness. In addition, varying all input parameters is not necessarily required as most parameters can be observed, are well known or a change does not significantly affect the output. Therefore, the models need to abide to a classification scheme, which segregates important parameters from less important parameters. Yang & Becerik-Gerber (2015) described such a classification by distinguishing observable and non-observable parameters, this classification is adapted as shown in **Figure 4.3.**

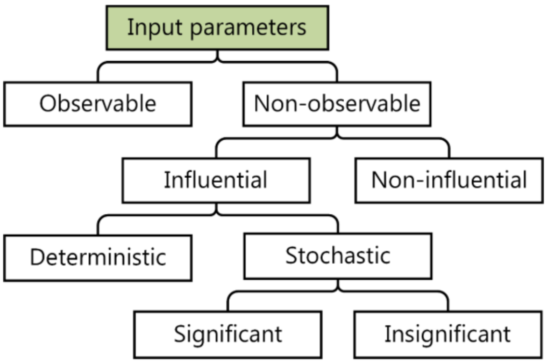


Figure 4.3: Classification scheme of input parameters, adjusted from Yang & Becerik-Gerber (2015).

Observable parameters are those whose values can be determined through available information, e.g. window sizes and lighting fixtures from drawings. Non-observable parameters cannot be determined by available information from data collection. To determine if these are influential, sensitivity analysis was employed. Due to the model size and simulation run time, the amount of parameters was reduced before running simulations to avoid the curse of dimensionality. It is therefore necessary to neglect insignificant parameters identified in previous studies as non-influential. As an example, certain material properties cannot be observed, but many are non-influential in, in particular in non-domestic buildings. Potentially influential parameters were split into deterministic and stochastic parameters. Deterministic parameters inhibit no randomness, where stochastic parameters are those that vary in their respective domains and cannot be measured exactly. Potentially influential parameters were varied during parametric simulation and those actually significant were determined through sensitivity analysis to be used for mathematical optimisation of a model.

4.5.4 Analysis of discrepancy

Differences between predicted and measured energy use need to be compared on a detailed level to identify performance issues. Maile et al. (2012) use data graphs to explain performance issues by comparing predicted and measured data. They and others (Chaudhary, et al., 2016; Kim & Park, 2016; Sun, et al., 2016; Yang & Becerik-Gerber, 2015; Kim, et al., 2017) use statistical variables proposed by Bou-Saada and Haberl (1995), such as the coefficient of variation of the root mean square error (CV(RMSE)) and normalised mean bias error (NMBE). The NMBE and CV(RMSE) in equations (6) and (7) were used to compute the residuals for the different data sets.

$$NMBE\ (\%) = \frac{\frac{\sum_{i-1}^{n}(m_i - s_i)}{n}}{\bar{m}} \cdot 100 \tag{6}$$

$$CV(RMSE)(\%) = \frac{\left[\frac{\sum_{i-1}^{n}(m_i - s_i)^2}{n}\right]^{\frac{1}{2}}}{\bar{m}} \cdot 100 \tag{7}$$

where m_i and s_i are the measured and simulated data points, ,$\bar{m}$ is the average of the measured time series data and n is the number of data points in the time series (i.e. $N_{monthly} = 12$, $N_{hourly} = 8760$).

Differences relate to either simulation-, operational- or measurement issues. Operational issues, for example, were encountered by looking at HVAC components and seeing if they are active when they are not supposed to, while simulation assumptions were distinguished by higher temperature setpoints than expected throughout the year or longer occupancy hours. Measurement issues were related to erroneous data points or missing data. This process is important as it identifies if performance issues relate to the simulation or operation of the building. Identifying where discrepancies exist highlights areas of importance and screens potential parameters that have a significant effect on model outputs.

4.5.5 Manual calibration

The calibration of building energy models is an underdetermined problem and its accuracy relates to the granularity of operational data used for model creation and comparison. A model is 'calibrated' when predictions differ by a certain range from the measured data, according to ASHRAE (2013) this is a 5% monthly mean bias error on total yearly energy use and 10% for hourly mean bias error on yearly energy use. In addition, CV(RMSE) is used, ASHRAE sets their criteria at <15% for the months, and <30% for the hours. It is however questionable if a calibrated model predicting within this range is an accurate model for representing reality, as the ranges mask higher levels of data granularity by only representing total energy use. Therefore, not taking into account energy end-uses, individual end-uses may not necessarily abide to the same statistical criteria, for example, lighting energy use may be significantly over predicted, while chiller energy use is under predicted, in total however they fall within the range and can be considered calibrated. Inherently, the input parameters determine the amount of energy used, as such, their assumptions need to be supported where possible by, building design- or commissioning data, as a last resort, benchmarks figures should be used. A larger uncertainty arises when such information is not available, this can be represented by the uncertainty of input parameters, which increase the variability in the output and potential solutions. There will be many models, among those calibrated, that are less accurate at a higher level of data granularity.

4.5.6 Sampling

The accuracy of sensitivity analysis and the meta-models depends on the amount of simulations and number of inputs. The necessary amount of simulations can be determined by calculating the sensitivity indices for increasing sample sizes, when results start converging the number of simulations can be determined to be sufficient, depending on the convergence criteria. Similarly, a larger sample will increase the accuracy of a meta-model as they can be trained on more data, however at a certain point, the increase in accuracy will be marginal.

A sample size (N) is in effect the number of different combinations of input parameters that are run through the deterministic model, whereas the number of parameters (k) included in a sample determine the volume of the parameter input space (Ω). It is therefore important to reduce the sample size as much as possible to avoid having to run many simulations and avoid the curse of dimensionality. Burhenne et al. (2011) analyse the performance of four different sampling techniques, pseudo-random, stratified sampling, Latin hypercube sampling (LHS) and sampling using Sobol sequences in the building simulation context. They found that Latin hypercube sampling and Sobol sequences had the fastest convergence on the mean estimates. Furthermore, Sobol sequences showed the least variations in the cumulative distribution functions, indicating that it produced the most robust results. Sobol sequences is a quasi-random sampling method, which obtains a faster convergence rate, meaning that fewer simulations are needed in order to reach the same accuracy that other methods offer. The numerical error of quasi-random sampling method is theoretically evaluated by $(\log N)^k/N$ in contrast to Monte Carlo methods which scale as $1/\sqrt{N}$. Eisenhower et al. (2011) and Peles et al. (2012) have used this quasi-random sampling to sample parameters in the input parameter space in the context of building modelling. In this research, Latin hypercube sampling is used for creating randomised designs, because of simplicity of its implementation and the availability of a computer cluster running simulations, which significantly reduced simulation time, compared with a personal laptop.

Parameter ranges were set at a 20% variation for all considered variables used in the samples, to allow for enough variation and understand parameter influences as done by (Eisenhower, et al., 2011). A range of variation for all parameters is determined by setting the standard deviation (σ) of a normal distribution at 20% of the typical value, with upper and lower bounds of the mean defined by adding and subtracting three times the standard deviation, mathematically in equation (8):

$$[lower, upper] = \left[\mu - (3 \cdot (\mu \cdot 0.2)), \mu + \left(3 \cdot (\mu \cdot 0.2)\right)\right] \tag{8}$$

Discrete distributions were used for input parameters such as the heating- and cooling setpoints, and an offset in schedules. Ideally, in detailed calibration, uncertainty ranges should be validated, for example for lighting loads in the building, the average load and standard deviation can be established based on design data, and for materials, these ranges are well-defined from previous research. The range of a particular variable influences the significance it has on the output, therefore the possible input parameter values have to be realistic. For example, if a deliberately narrow range of say 15.5 to 16 W/m^2 for equipment power density in a zone is chosen, this will still allow the model to be calibrated and conclude that this particular variable is not significant compared to other variables, which have a proportionately larger variance. However, in this research, the uncertainty of most variables could not be determined, as such, the 20% distribution as described above is used.

4.5.7 Parametric simulation

Input parameters were sampled to generate random values within a pre-defined range of variability. Each sample of inputs was exported to a separate simulation file, which was then simulated using EnergyPlus on Legion, UCL's computer cluster. For each building, numerous simulations were run in order to analyse the variance in predictions. Sampling, the parametric simulation process and file generation is set-up in Python. For sampling, pyDOE (pyDOE, 2017), an experimental design package for Python, was utilised for creating Latin hypercube designs. An initial base case simulation file is overwritten with new input parameter values from the samples. Although software is available for parametric simulation using EnergyPlus, the use of Python made the process more flexible and allowed unsupported features to be included, such as introducing schedule variations automatically, control of output parameters and direct integration with analysis and data visualisation.

4.5.8 Uncertainty and sensitivity analysis

Uncertainty analysis quantifies uncertainty in the output of the model due to the uncertainty in the input parameters to the building models. The Monte Carlo method is applied to repeatedly random sample distributions of inputs to obtain the distribution of the energy consumption. Input parameters are similar for the case study buildings, with minor exceptions. For scheduling, Office 71 does not include any variability, instead lighting and equipment schedules are the same for each simulation, based on the typical weekday and weekend day of respective end-uses. This enables the comparison of different methodologies for representing occupancy presence. In both MPEB and CH the occupancy, lighting and equipment schedules are variable.

For sensitivity analysis, typical correlation and regressions coefficients were calculated using the SciPy library, scientific tools for Python. In addition, a variance-based method was investigated by calculating Sobol' indices using SALib, an open source Python library for sensitivity analysis (Herman & Usher, 2017). Variance-based sensitivity analysis is computationally demanding, as it requires numerous simulation runs for the sensitivity results to converge. The meta-model is therefore used to compute a large amount of samples and perform variance-based sensitivity analysis.

4.5.9 Meta-model development

The base case building energy models were simulated for a range of uncertain input parameters, resulting in a range of outputs, these are referred to as the search space (possible input parameters, created from the design of experiments) and solution space (possible outputs). When enough samples (sets of input parameters) are simulated, the relationship between inputs and outputs can be represented mathematically, and new samples can then be predicted without using the building energy simulation software. This mathematic representation is classified as supervised learning, where the inputs and outputs are training data, used to fit an estimator to predict future data for new samples. An estimator is a rule for calculating an estimate based on the training data (inputs and outputs), i.e. it is the method or machine learning algorithm used to calculate new predictions, also known as a meta-model or surrogate model.

A meta-model is a simplified model of the energy model based on a mathematical relation between the inputs and outputs from Monte Carlo simulation, approximating component functions of the building model. This allows analysing the variance of the output and identifying the model response for different parameters that differ from the sampling points. A meta-model can be created using different techniques, Eisenhower et al. (2011) and O'Neill & Eisenhower (2013) use Support Vector Machines (SVM) with Gaussian kernels, whereas Peles et al. (2012) and Mara & Tarantola (2008) use orthonormal polynomials for static parameters and stochastic processes (Ahuja & Peles, 2013). For the case research, use is made of generalised linear regression techniques (or general linear models, i.e. multivariate multiple linear regression). In addition, Artificial Neural Networks (ANN) were compared with the performance of the generalised linear regression techniques. Other algorithms, such as support vector regression, Gaussian process regression, and principal component regression for supervised learning are becoming increasingly popular as they can represent complex relations between inputs and outputs. During development of the meta-models for the case study buildings it was found that with simpler models and a limited number of input parameters, the linear regression techniques were sufficient to accurately predict new values.

However, with increasing complexity there was a need to implement a more sophisticated machine learning algorithm.

Meta-models were created using the Python programming language with open source libraries available for implementing different estimators. The scikit-learn library (Pedregosa, et al., 2011) is used to test different multi-variate regression techniques, such as ordinary least squares regression, partial least squares regression, ridge regression, Bayesian regression, Lasso regression, etc. A general linear model can be written as:

$$Y_{nxp} = X_{nx(k+1)}\beta_{(k+1)xp} + \varepsilon \quad (9)$$

where Y is the measured data (multiple dependent variables), X are the samples or variable input parameters (multiple independent variables) and β are the parameters to be estimated from the relation between the data, with ε being the independent distributed normal error. The regression techniques then aim to minimise the differences between the general linear model and measured data by adjusting the β-coefficients and ε-error (also known as intercept). Izenman (2008) gives an introduction on linear regression and machine learning, where Rencher (2002) provides an extensive overview of methods for multivariate analysis.

In pseudocode, the following steps were taken for developing the meta-models:

```
Data preparation # manipulate samples and simulation results
        Retrieve inputs and outputs from parametric simulations # search and solution-
        space
        Model order reduction if too many variables
Randomly split solution space in training and test data #[75% train/25% test]
Loop through set of estimators: # different types of machine learning algorithms
        Scale input and output data if necessary
        Train or fit inputs and output relationship
        Predict using the test data
        Calculate regression performance scores and loss functions
        Save scores
Evaluate estimator scores and decide on best meta-model
Save meta-model for future predictions
```

Since the relationship is only an approximation of the first-principle energy modelling software, it is important to determine its accuracy, which is dependent on the number of simulations and input parameters that make up the training data. Training data were split into a training set for learning the relation, and a testing set in order to validate the accuracy of the meta-model. Statistical error measures test the regression performance and help in determining a suitable estimator for constructing the meta-model, using the following statistical measures:

- Mean Absolute Error (MAE), refers to the amount by which predicted values differ from those being estimated (i.e. the test data excluded from the training sample).
- Mean squared error (MSE), similar to the MAE, the MSE emphasizes outliers, i.e. larger errors have greater influence than smaller ones due to the error being squared.
- Coefficient of (multiple) determination (r^2), generally measures how well future samples are predicted. It is the proportion of variability in the data explained by the statistical model.

4.5.10 Mathematical optimisation

After constructing and validating the meta-models, they were used for optimisation (i.e. automated calibration), minimising the difference between predicted and measured energy use at different levels of data granularity. Optimisation using a meta-model is significantly faster than using building energy modelling software. Computation times of the larger first-principle models were over 20 minutes to run for a yearly simulation on a personal laptop, while the meta-model computes a new sample in a fraction of a second. The added benefit of fast computation times is that different optimisation techniques and their variables were quickly investigated.

Meta-model based optimisation aims to find calibrated models by adjusting the variable input parameters within their range of uncertainty. When the outputs or predictions of the meta-model more closely resemble (fit certain criteria) the measured data, the combinations of input parameters are stored and improved iteratively to get even an even closer match between predictions and measurements (i.e. objective minimisation). This iterative process uses the Non-dominated Sorting Genetic Algorithm-II (NSGA-II), an evolutionary optimisation algorithm, used to solve multi-objective optimisation problems by finding sets of solutions. A multitude of optimisation algorithms exist, such as genetic algorithms, evolutionary algorithms, particle swarm, simulated annealing, stochastic tunnelling and so forth. The choice was made to focus on only one optimisation algorithm here, due to the ease of implementation within Python through available open source code and because it is a popular and proven method for building performance optimisation. Evins (2013) reviewed the use of optimisation algorithms in the context of building simulation and found that more than half of the reviewed works used a genetic algorithm. Similarly, Nguyen et al. (2014) found that stochastic population-based algorithms, primarily genetic algorithms were most frequently used in building performance optimisation. They noted that there are a number of reasons for the popularity of genetic algorithms in building performance simulation, in particular its capability at handling continuous and discrete variables, concurrent evaluation of n individuals (Deb, 1999), and robustness in handling discontinuity, multi-modal and highly constrained problems without being trapped at a local minimum (Colorni, et al., 1999). The distributed evolutionary algorithms in python (DEAP) library developed by Fortin et al. (2012) was used for implementing the NSGA-II multi-objective optimisation.

Pseudocode for the genetic optimisation algorithm:

```
Set optimisation parameters #pop size, n generations, probabilities of crossing and mutation
Set fitness/weights #the more energy use the more important to minimise
Create initial population of individuals #samples
Load meta-model
Run through generations
    Generate offspring
    Crossover individuals within population
    Mutate offspring #change parameters within individuals
    Evaluate objective function #predictions vs. measurements, minimise RMSE
        Save individuals that fit optimisation criteria #CV(RMSE) / NMBE
    Select next generation population # incl. good offspring
```

The objective function implemented is to minimise the RMSE between predictions and measurements, shown in equation (10). Weights were introduced based on the amount of energy use by normalising measured energy use, meaning that the more energy use for an objective (energy end-use) the higher the weight and importance to minimise.

$$\min(f_{obj} = \frac{1}{n}\sum_{i=0}^{n} w_i(p_i - m_i)^2) \quad (RMSE) \tag{10}$$

where p_i are predictions, m_i measurements and w_i weights.

4.6 Data visualisation techniques

Throughout the research and in this thesis, operational data was analysed and measured energy use was compared with predictions. Several data visualisations were used which have certain underlying assumptions to how the visualised data was calculated, and is therefore explained here, to be used as a reference.

4.6.1 'Typical' weekday and weekend day profiles

Typical weekday and weekend day profiles of energy use were used to compare predicted and measured energy use and understand the hourly variation in energy use, profiles were created for either a per month or per year basis. The weekday and weekend day denotations can be used interchangeably with a working day and non-working day respectively, however in this research, the 'typical' weekend day excludes holidays, whereas a non-working day would not. A typical weekday and weekend day are determined by first separating the holidays, working days and weekend days in a specific time series (these are year dependent). The separated day types are then grouped together and the average, mean, standard deviation or specific quantile during a specific period is determined. The mean represents a typical day and is in some instances accompanied by the standard deviation throughout the same period. In certain cases, the profile is normalised to unity (i.e. scaled to bring the values into the range [0, 1]), which is helpful for comparing solely the trends in data. Energy use data was available at both 15 and 30-minute intervals, reporting the kWh used from one data point to the next. To allow like-for-like comparison at a sub-hourly basis, the time series data were first converted to kW. An example of a typical weekday and weekend day of electricity use for lighting and power from different meters is given in **Figure 4.4**.

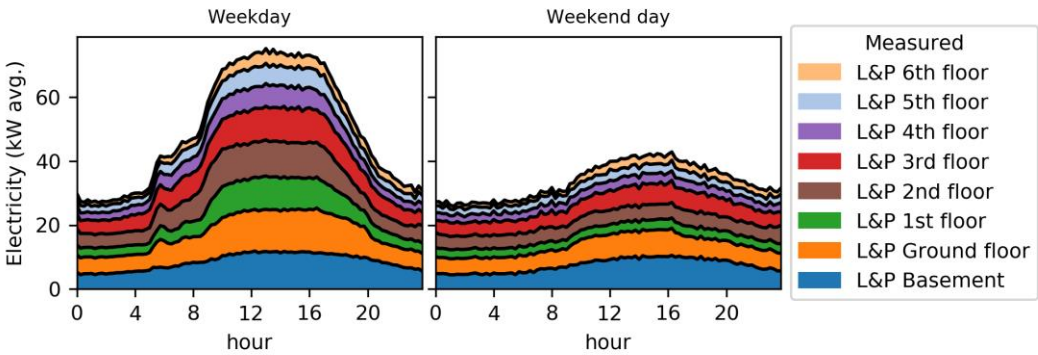


Figure 4.4: Example of a typical week and weekend day showing the mean of electricity use for lighting and power on different building floors, based on a year's worth of data.

4.6.2 Representative load patterns for benchmarking

Similar to the typical weekday and weekend day profiles are the representative load patterns. These were used for benchmarking purposes of (in this case) electricity use. A representative load pattern is calculated by normalising the load for each time step by the average daily near-peak load, where the near-peak load is the average of the daily load at 95% quantile for all working days. Then, the normalised vector for each working day is averaged over a month (or potentially a season). Mathematically, expressed in equation (11), as explained by (Luo, et al., 2017), an example of a representative load profile is shown in **Figure 4.5.**

$$RLP = \frac{1}{n}(WD_{day})\left(\frac{L_i}{\frac{1}{n}(W_1 pct_{95}, W_2 pct_{95}, \dots, W_n pct_{95})} for\ i\ in\ \{0,1,\dots,n\}\right) for\ day\ in\ \{0,1,\dots,n\} \quad (11)$$

where,
WD_{day} = working day (24 hours as hourly/sub-hourly values electricity use in kW)
L_i = measured sub-hourly electricity use (kW),
pct_{95} = near-peak load (load at quantile 95%) for all working days (W).

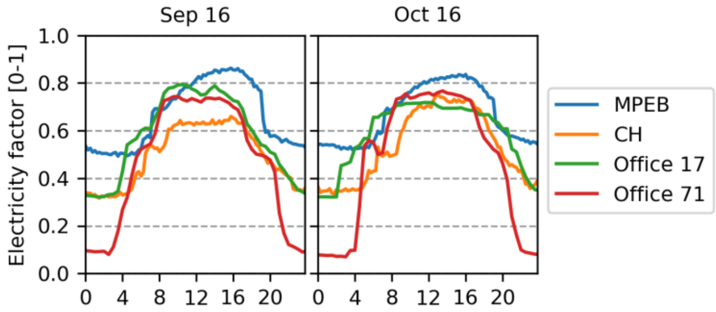


Figure 4.5: Example of a representative load profile based on electricity use.

4.6.3 Heat map

The heat map visualises energy use for a short period (few weeks to a few months), but with the underlying sub-hourly fluctuation. The heat map shown in Figure 4.6 presents 2 months' worth of total energy use data, in the top plot, each square represent a data point measured at a one hour interval. The y-axis represents the hours and each line of blocks on the x-axis represents a full day. The heat maps are accompanied by a bottom plot, which represents the total daily energy use, visualising the variation between the week and weekend.

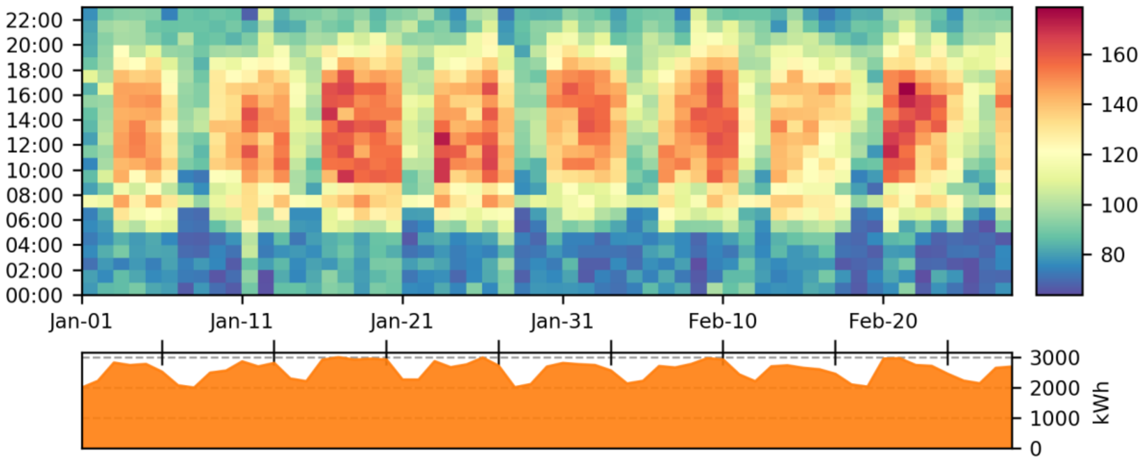


Figure 4.6: Example of a heat map of sub-hourly energy use (kW) at the top accompanied by a visualisation of daily energy use (kWh) on the bottom.

4.6.4 Discrepancy metrics analysis

As explained in Section 4.5.3, a discrepancy between predicted and measured energy use is analysed by using the normalised mean bias error (NMBE) and coefficient of variation of root mean square error (CV(RMSE)), these metrics indicate the error or difference between two datasets. Used to indicate the error on a monthly (differences between energy use on a monthly interval) and hourly level (differences between energy use at an hourly interval). Counterintuitively, the bar graph in **Figure 4.7** indicates the differences on an hourly interval per month, whereas both metrics can also be calculated on a yearly basis. The difference between two datasets at an hourly interval over the whole year, which is given by $NMBE_{hourly}$ and $CV(RMSE)_{hourly.}$ Differences between the months on a yearly basis are given by $NMBE_{monthly}$ and $CV(RMSE)_{monthly}$. Finally, the orange lines are the criteria set by ASHRAE that deem a model calibrated, however these criteria should actually be compared at a yearly basis (the metrics shown at the top of the graph). Nevertheless, they were used in the bar graph here to understand how the model is performing on a monthly basis, which would be more difficult to achieve (a smaller difference between the two datasets on an hourly level).

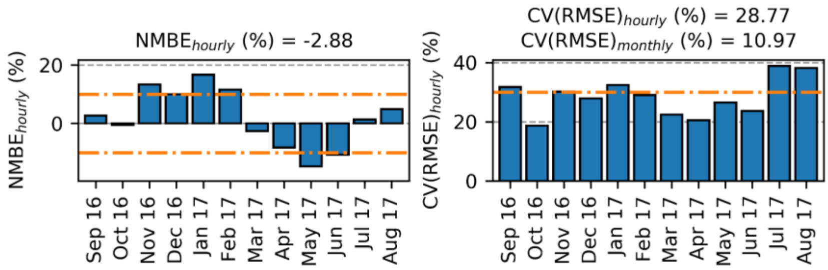


Figure 4.7: Example of a discrepancy metrics analysis.

4.6.5 Kernel distribution

Kernel density estimation has been used to estimate the probability density function of certain variables. Although the probability density function was estimated to resemble a normal distribution in most cases, in some the kernel distribution provided additional insight in the actual distribution of variables. Bandwidth selection is done automatically as the kernel density estimation was applied through the SciPy library in Python, which applies Scott's rule (1992), bandwidth influences the estimate obtained. An example of kernel distribution using kernel density estimation is shown in **Figure 4.8.**

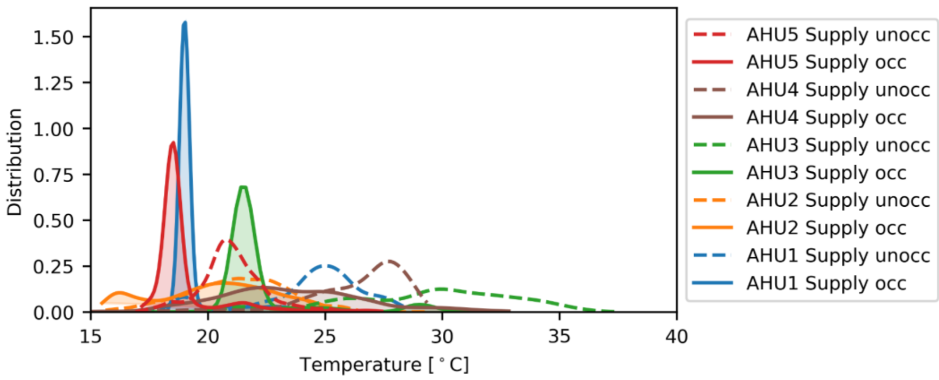


Figure 4.8: Example of kernel distributions for air handling unit supply temperatures.

In this graph in particular a distinction is made between occupied (7am-7pm) and unoccupied hours, created for temperatures in MPEB, the kernel estimation is then determined based on these specific hours.

4.7 Summary

A methodology was developed to compare predicted and measured energy use at a high level of data granularity. It describes how data was collected and synthesised, how virtual models were created to represent existing buildings, how sensitivity and uncertainty analysis was performed and describes how manual and automated calibration techniques tune the models towards measured energy use. The methodology builds on previous research, referred to throughout the methodology. The vast amount of previous modelling and calibration research proves the importance of the subject, but also highlights that primarily minor improvements are made in this area of research. Improving the efficiency and effectiveness of the calibration process is helpful as it is used to make design decisions based on predicted savings for retrofitting existing buildings. In addition, it can provide a link between simulation and measurement systems in existing buildings to enable real-time forecasting of building performance in the future, in this case model calibration accuracy and application of state-of-the-art techniques such as machine learning become essential. Finally, calibrated models can be used as a tool to diagnose building performance and understand differences between predictions from modelling and measurements.

In essence, the methodology improves upon existing knowledge by introducing several new processes and techniques. First, the parametric modelling process uses variable scheduling in occupancy presence, equipment and lighting loads, to allow for automated adjustment of schedules to improve the accuracy of the calibration model. Second, the methodology introduced seasonal variability in profiling and uncertainty in heating and cooling setpoints in addition to uncertainty in typical static parameters. This allows the model to replicate seasonal trends of an existing building, important for certain building types (e.g. university buildings). Finally, the methodology was developed in order to determine if a higher level of data granularity increase model accuracy through model calibration.

5 UTILISING OPERATIONAL DATA TO INFORM BUILDING MODELLING ASSUMPTIONS

This chapter discusses the operational data collection and analysis for the case study buildings, which was used to inform input assumptions and support the development and calibration of building energy models. Additionally, an overview is given of how the case study buildings perform against each other and typical UK benchmarks. Several techniques were used to establish input parameters for the building energy models. Electricity use for lighting and power per floor was analysed to create typical schedules of use and determine out-of-hours baseloads. Wi-Fi and swipe card data were used to create typical occupancy schedules, but were also used to inform lighting and power use schedules. In addition, both datasets were used to understand seasonal variation, which proved to be significant in the university buildings. Strong correlations were identified between occupancy presence and lighting and power energy use, and a less strong correlation between occupants and systems energy. Space temperatures were used to make evidenced assumptions about heating and cooling setpoint temperatures in spaces, which were compared to design specifications from O&M manuals. Finally, systems operation, in particular air handlers and fan coil unit operation were investigated to match their operation in the building energy model.

5.1 Introduction

Operational data, i.e. data collected during the operation of the buildings, was used to inform building modelling assumptions, an integral part of calibrating energy models. Consecutively, this data was used to compare predictions from building energy models with measurements to support the calibration process. The collected operational data consisted of energy use data on a sub-hourly basis and gas data on a monthly and in some cases sub-hourly basis. Building management systems in two of the buildings log data regarding the operation of building systems. Additionally, anonymised occupancy data through Wi-Fi connections and swipe card access points was obtained for the university buildings. **Table 5.1** provides an overview of the data collected for the four case study buildings.

Table 5.1: Data availability for the case study buildings

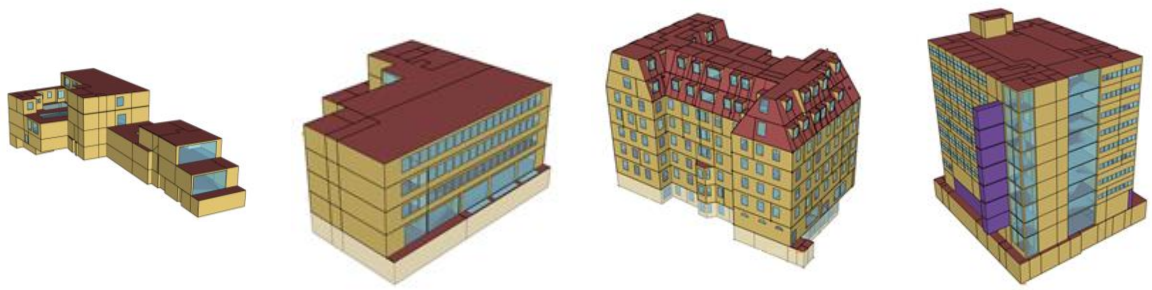
Building	BMS	AMR	STM	Gas / heat	Wi-Fi	Environmental data
CH	Y (local)	Y	Y	Monthly	Y	Temperature
MPEB	Y (online)	Y	Y	n/a	Y	Temperature
Office 17	N	Y	N	Sub-hourly	N	n/a
Office 71	N	Y	N	Sub-hourly	N	n/a

BMS: Building management system, AMR: Automatic meter reading, STM: Short term monitoring

This chapter describes findings from analysing energy use data of the case study buildings and then describes how operational data was used to inform building modelling assumptions. In particular, this data was used to; (1) develop occupancy, equipment and lighting schedules based on occupancy and energy use data; (2) establish building services operation and set-point temperatures analysing the environmental performance and system performance given by the BMS.

5.2 Building modelling

In this research, four existing buildings have been investigated, their geometrical representation is shown in **Figure 5.1**. For each case study building a virtual model is created to predict performance aspects (e.g. energy use, internal conditions, etc.) of a real building. The initial model is based on best-guess parameter assumptions, and is referred to as the base case. The base case is then either manually and/or automatically adjusted to ensure a better representation of reality, which are referred to as calibrated models.



Office 17 | Office 71 | CH | MPEB

Figure 5.1: Geometrical representation of the four case study buildings.

Common practice among building modellers is to classify thermal zones as different space types. This allows for creating templates for specific space type, which in turn assigns the same parameters to included zones. However, it is important, especially for model calibration, to represent each space as accurately as possible. Two zones that are very different from each other should not be assigned to the same space type. For example, if two kitchens or office spaces differ considerably in their installed plug loads, then it might be more accurate to separate these into two space types with differing parameters or separately assign these plug loads to accurately represent the plug loads. Where possible, assumptions were based on observed conditions through energy audits, measured data, and available design and commissioning data from O&M manuals. When parameter values could not be established through collected information, they were based on typical assumptions from guidance documents.

Lighting and equipment loads are based on the number of lights and appliances in each space summed over an agglomeration of spaces under one space type. People presence was determined using Wi-Fi data, occupancy profiles and number of people were derived from this data. HVAC systems parameters, such as heating and cooling capacities and design airflow rates were hard-sized where data was available to decrease uncertainty in the model. Natural ventilation, available through the opening of windows is available in all buildings and was controlled dependent on the indoor and outdoor temperature and time of day. Weather files were obtained covering the measurement periods to represent external weather conditions to coincide with the data collection period. Building modelling inputs are further described in **Appendix D** and materials in **Appendix E**.

5.3 Energy use and occupancy presence

The collection and collation of data proved to be time-consuming, which is one of the limiting factors in applied detailed building model calibration. In particular, energy use data often contained missing and erroneous values. In addition, electricity meters were often not labelled correctly, excluded important and significant energy end-uses and run on outdated software. In contrast, the building management system, which was readily accessible for MPEB proved to provide easy data analysis through an online platform. In CH however, this system was only accessible offsite and obtaining useful data was difficult, as most points were not being logged for longer than a month. This research focussed on energy use and establishing a high level of data granularity, which was achieved by gathering disaggregated sub-hourly energy end-use data for four case study buildings. For some buildings, energy for lighting and power was also disaggregated on a floor by floor basis. However, due to faulty meters, certain energy end-uses were neglected, a detailed breakdown was not always established. In particular, the heat meter in MPEB was malfunctioning and the heat pump and air handling unit in Office 71 were not sub-metered. An overview of energy use for each building is given in **Appendix F**.

5.3.1 Energy use benchmarking

Yearly energy use intensity

Energy use for each building is compared on a yearly basis per floor area in **Figure 5.2**. For both CH and MPEB, the available 8 months of data was extrapolated to a year by taking the average per month for an extra 4 months. For Office 17 and Office 71 the year of 2013 is used as comparison. A considerable difference exists between the case studies, MPEB is the most energy intensive building, closely followed by CH and Office 17, which are both less than half the energy use intensity (EUI). Note here that district heating in MPEB is not included, as it was not measured, similarly so for the VRF system energy use in Office 17 and 71. However, district heating in MPEB was determined to account for 3.4% of yearly energy use from predictions, whereas predicted system energy use in Office 17 and Office 71 accounted for 6.5% and 13% respectively.

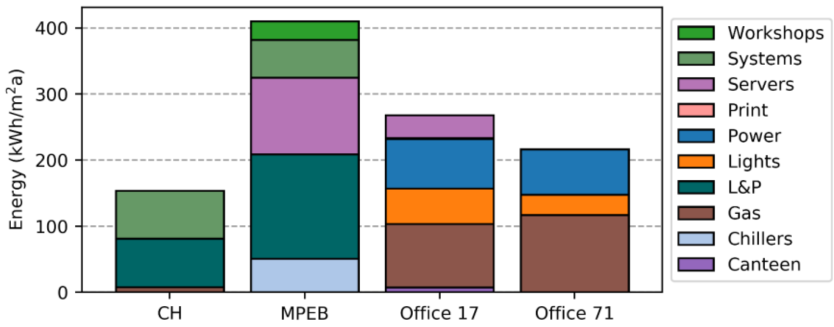


Figure 5.2: Energy use for during September 2016 to September 2017.

MPEB is an energy intensive building due to the high internal loads from the server and large amount of power appliances on the first three floors in the building. In addition, MPEB is a fully air conditioned building, nearly every space in the building is provided with heating, cooling and mechanically supplied fresh air besides providing operable windows in most working spaces around the perimeter of the building. CH is mainly naturally ventilated, but does provide heating and cooling through the VRF system in all offices and meeting rooms, system energy use is similar to that of MPEB. Office 71 is also mechanically ventilated and has a VRF system to provide heating and cooling, the VRF heat pumps and AHU in the building are however not captured by the sub-metering system and therefore system energy use seems negligible in this building. Finally, Office 17 is naturally ventilated except for the basement and reception area, which are provided with some air conditioning, but the main contributor to the high intensity in energy use in Office 17, is the unnecessary gas use throughout the year and outside of occupied hours. Furthermore, lighting and power is more intense than that of CH and Office 71 as the occupancy density and therefore equipment gains are much higher.

Energy use intensity for each building is compared to benchmarks from the Energy Consumption Guide 19 (ECG019), shown in **Table 5.2.** This guide provides benchmarks from common four common office types, derived from surveys of a large number of buildings and divides them into typical (T) and good practice (G) benchmarks. MPEB is identified as type 4 as 'prestige' air conditioned building, while CH and Office 71 are type 3, 'standard' air conditioned buildings, CH is a type 2, open plan naturally ventilated.

Table 5.2: Energy use intensity (kWh/m²a) compared to benchmarks from ECG019 (Department of the Environment, 2003)

	MPEB	CH	71	17	G	T	G	T	G	T
Type	4	3	3	2	2	2	3	3	4	4
Cooling	41				1	2	14	31	21	41
Fans, pumps and controls					4	8	30	60	36	67
Humidification					0	0	8	18	12	23
Systems	55	76	4	5	5	10	52	109	69	131
Lighting			36	30	22	38	27	54	29	60
Power			52	110	20	27	23	31	23	32
L&P	213	67	88	140	42	65	50	85	52	92
Computer room	79	4	0	32	0	0	14	18	87	105
Catering [electricity]	0	0	9	10	3	5	5	6	13	15
Catering [gas]	0	0	0	0	0	0	0	0	7	9
Other electricity					4	5	7	8	13	15
Total gas/oil	N/A	12	53	185	79	151	97	178	114	210
Total electricity	388	147	101	187	54	85	128	226	234	358

Note: Type 2: open plan naturally ventilated; Type 3: 'standard' air conditioned; Type 4: 'prestige' air conditioned
G: Good practice, T: Typical

Looking at total electricity use, MPEB has a very similar EUI compared to the typical type 4 benchmark. Heat data was not available from the district heating system for comparison. Disaggregated into different energy end-uses, MPEB has smaller computer room (servers) and systems energy use, but a much higher lighting and power EUI. CH and Office 71 consume nearly as much energy on total, however their distribution to gas and electricity is slightly different. Compared to type 3 typical benchmarks, both electricity and gas use is significantly lower, even good practice benchmarks are more energy intensive. Systems energy use in CH falls within the typical and good benchmarks of type 3, however the data for Office 71 doesn't include all systems and is lacking a significant portion of energy use. Lighting and power are similar to the typical type 3 benchmarks levels. Finally, Office 17, which is naturally ventilated, and heated through radiators connected to two gas boilers is significantly more electricity intensive than the typical type 2 benchmark. The main difference is due to the higher power loads, in addition to the server which is not present in the benchmark. Gas use was however identified to be used inefficiently in this building as shown in Office 17.

The benchmark comparison does not provide a very meaningful comparison, first, the benchmarks are outdated as they are derived from an older building stock, an updated version is being created, but unlike in the U.S. where EUIs in certain states have to be made publicly available, the UK does not have such a system in place for private buildings. Second defining the buildings according to a building type is not always straightforward. Where perhaps Office 17 is a typical naturally ventilated open plan office, Office 71, is likely in between that and a 'standard' air conditioned office. The university buildings however are even more difficult to assign a certain building type or function as they provide many different utilities. CH has mainly office spaces in the building, in addition to some lecture spaces, but the pattern of use was found to be very different from a typical office building, having a significant influence on energy use. MPEB serves many purposes, but from the 2nd floor up, is again mainly office space. Finally, each individual building can be significantly different, not just in the internal functions they provide, but also in the way systems condition the building, additional servers or small power equipment, patterns of use and efficiency in operation. The latter ideally identified through looking at benchmarks, however, to understand if a building is actually energy efficient it will be more fruitful to analyse energy use on a higher temporal granularity, perhaps looking at typical weekday and weekend day patterns, broken down into energy end-uses.

For the case study buildings several Display Energy Certificates (DEC) were available, shown in **Table 5.3**. DECs should be available for all public buildings in the UK, they are based on analysing measured energy use of a building, renewed yearly. As is evident from the logged DECs, annual assessments can be considerably different. For example, for CH the floor area is determined to be different between 2014 and 2015, significantly affecting the final rating. For MPEB, there is a significant difference between reported annual energy consumption for both heat and electricity. These values can however not be verified as there was no sub-metered data available for these years (DECs can be based on energy bills).

Table 5.3: Display Energy Certificates of the four case study buildings.

Building	Floor area (m^2)	Issue date	DEC	Heating	Electricity	Renewables
CH	3973	2014	F (138)	83 (135)	211 (126)	0
CH	5664	2015	D (76)	8 (237)	137 (93)	0
CH	5665	26-10-2016	D (85)	8 (227)	145 (93)	0
MPEB	9020	2014	F (128)	111 (237)	170 (80)	18.4% (e)
MPEB	9580	29-03-15	E (113)	122 (242)	210 (131)	40.2% (e)
MPEB	9580	25-05-16	G (181)	93 (258)	376 (131)	30.2% (e)
MPEB	9590	27-01-17	G (152)	93 (247)	257 (131)	28.1% (e)
Office 71	2621	2015	D (90)	33 (126)	113 (94)	0
Office 17	1924	2015	F (137)	81 (103)	151 (95)	0

Electricity use load shape benchmarking

Although the certificates are helpful in understanding the cumulative energy use compared to other buildings, it is evidently not an entirely fair comparison if the EUIs, which are based on floor areas, are not representative for the whole building. Furthermore, EUIs mask the underlying fluctuations of energy use over time, however, with the uptake of smart metering and availability of sub-hourly energy consumption data, energy use patterns can be identified and compared as benchmarks among buildings. These energy use patterns (or load shapes) contain information on how electricity changes over the day, as a composite of end-uses such as lights, appliances and heating, ventilation and air conditioning (HVAC) (Luo, et al., 2017). Potentially, in the future, when data collection systems improve and accurately capture disaggregated energy use, energy end-uses can also be compared among buildings.

There are different metrics that can be employed to represent the behaviour of energy use patterns, the metrics proposed by Luo et al. (2017) were compared among the four case study buildings. More specifically, the peak-base load ratio, weekday-weekend day load ratio and on-hour duration have been calculated for electricity use, where gas use was not always available on sub-hourly basis. In addition, a representative load pattern (RLP) is determined for each month of the year, used to identify irregular shapes and were compared among the case study buildings, the representative load profiles are shown in **Figure 5.3**.

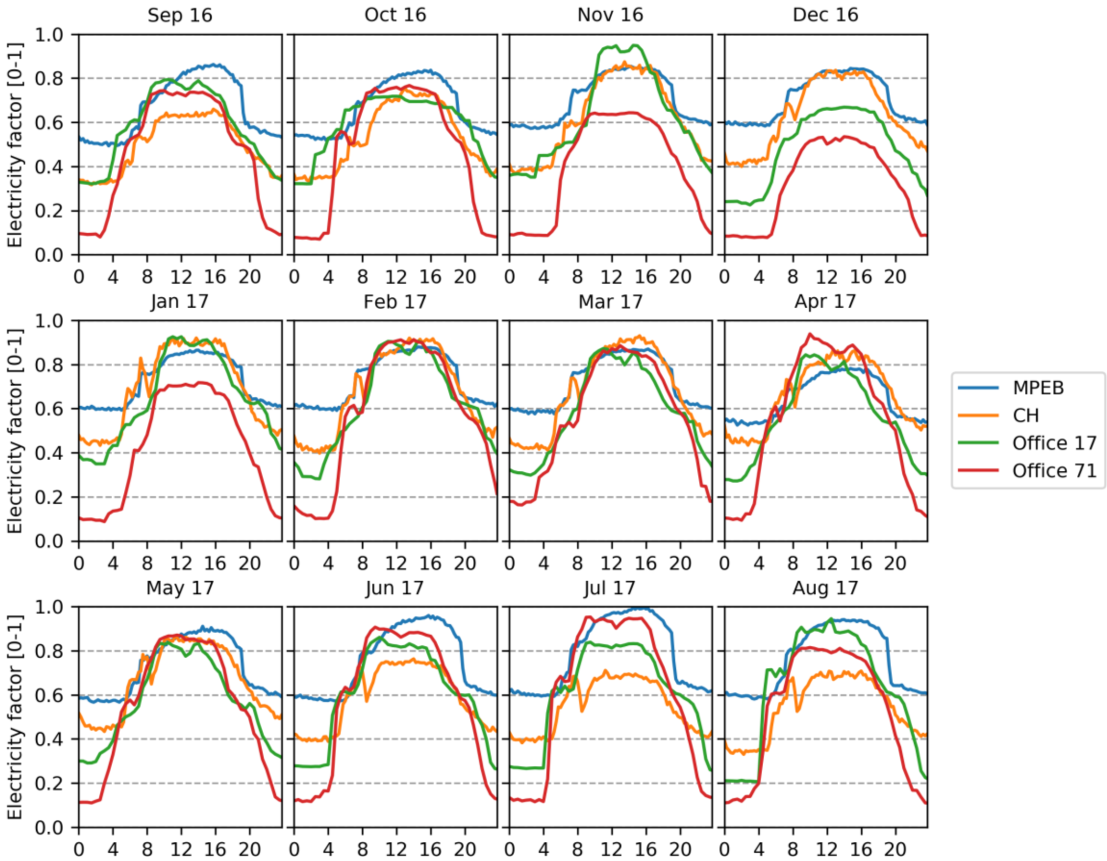


Figure 5.3: Representative electricity load profiles for the four case study buildings (typical weekday).

The representative electricity load profiles enable comparing typical patterns of use between the buildings on a monthly basis. The main observation that can be made when comparing the profiles is the large difference between day and night time electricity use. Furthermore, as the profiles were normalised based on the average annual near-peak load, a difference in magnitude for energy use throughout the months can also be identified. For CH, there is a significant difference between summer and early autumn (June to October), which have a much lower load than for the other months of the year. MPEB in contrast has a higher load during the summer months (May to August). Office 17 and 71 show a somewhat smaller load during the month of December in comparison to other years, indicating that these two offices are closed or partially closed, most likely during Christmas and New Year. Another observation is the difference between the operating hours, which seem to be slightly longer for Office 17 and 71 in comparison to the two university buildings. Finally, MPEB shows a significant increase and decrease (rise and fall time) of electricity use during the occupied hours, which is more fluent for CH. This was identified to be due to the strict time-operated systems in MPEB, whereas CH has individual control on rooms. Office 17 and 71 on the other hand also show a significant rise and fall between occupied and unoccupied hours, however, this is not the same throughout the year.

The representative load patterns are supported and further explain electricity use by the load shape metrics. The peak to base load ratio is calculated by taking the average working or non-working day during a certain period, then dividing the near-peak load (95% quantile) by the base load (15% quantile). The on-hour duration is calculated by taking the average working or non-working day and determining the number of hours above a threshold (5% quantile + ¼ x 95% quantile). The weekday-weekend day load ratio is calculated by taking the average weekday and weekend day and dividing their 15% quantile loads.

Results for the peak to load base ratio, on-hour duration and weekday to weekend day ratio are shown in **Figure 5.4**, **Figure 5.5** and **Figure 5.6** respectively.

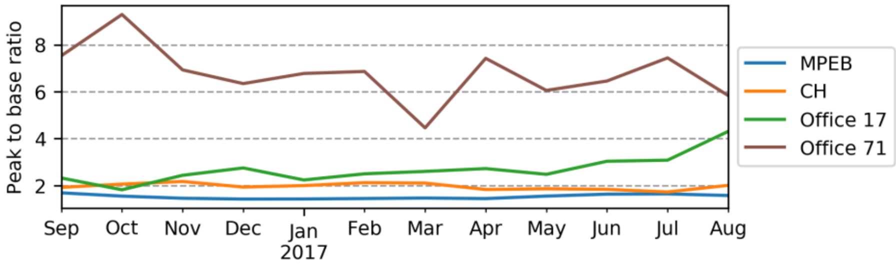


Figure 5.4: Peak to base ratio.

The peak base load ratios for the university buildings are very similar, whereas Office 17 is slightly higher during the months and Office 71 is significantly higher, with values ranging from 4 to 9. A higher ratio indicates a large difference between day and night electricity use, where a low peak to base ratio results from potentially appliances and lighting being left on. In MPEB and Office 17, the server that is continuously operating is one of the causes for their low ratio. The ratios can also be inferred from the representative load profiles, where a distinct difference was seen during the day and night.

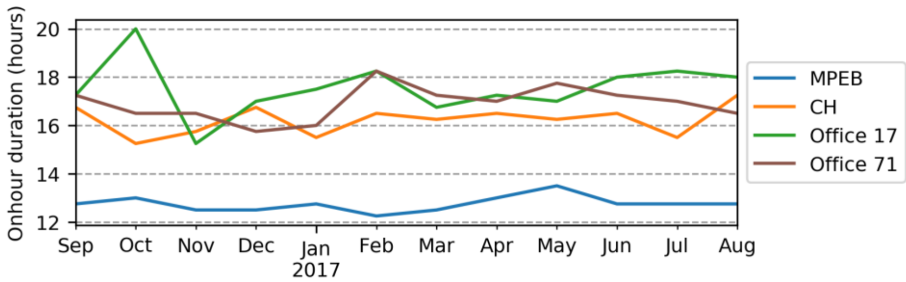


Figure 5.5: On-hour duration.

CH, Office 17 and 71 are operated between 16-18 hours according to the on-hour duration metric, whereas MPEB is on for about 13 hours. For MPEB this does however not necessarily mean that the building is occupied for only 13 hours. The large difference in electricity use during the day and night determine the threshold, which determines the on-hour duration, the large difference is however mainly due to system energy use, where lighting and power is nearly constant throughout the day. Therefore, occupation hours might be longer than the on-hour duration.

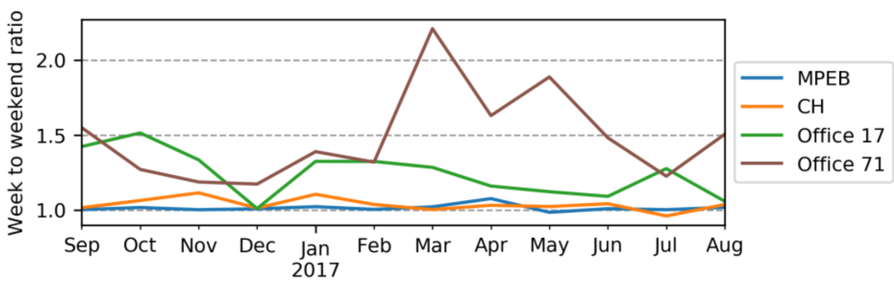


Figure 5.6: Weekday to weekend day ratio.

The weekday and weekend day ratio indicate that the university buildings are occupied and utilised 7 days a week, in contrast to the two office buildings.

5.3.2 Occupancy presence

Occupants can have a large influence on the energy consumption of a building, they use equipment, use hot water and are able to adjust their internal environmental conditions by changing lighting conditions, opening windows and changing thermostat settings. It is therefore necessary especially in the calibration of building energy models to capture their presence and understand their behaviour. Several techniques can be employed to determine occupancy presence, Davis III & Nutter (2010) use building security cameras, doorway electronic counting sensors, semester classroom scheduling and personal observation as ways to collect such data, Yun et al. (2012) use occupancy sensors in spaces and Martani et al. (2012) use Wi-Fi connection data. For the university buildings, CH and MPEB, both Wi-Fi connection data and swipe card access data is collected. The Wi-Fi data shows when a connection to a particular router is made, this can then be used to understand when people are present in a building and on which floor. In addition, and in similar fashion, students and staff have an ID card to gain access to a building, building floor or particular zone. When and where they gain access is the retrieved anonymised swipe card data, this was analysed to determine occupancy presence. There are several limitations to both methods of data collection, for the Wi-Fi data, the number of connections made does not directly represent the number of people as each occupant can have more connections to the Wi-Fi network. Nevertheless, connections and occupancy presence are correlated, which gives an understanding of when and relatively how many people are present. Swipe card data is likely to be a better representation for the number of people that enter the building, as everyone is supposed to swipe in for access. On the different floors where the access points are mainly at door entrances instead of a main gate, it is more likely that doors are held open for other people and actual people flow is more difficult to determine. Furthermore, there is no need to swipe out when leaving the building, which makes it difficult to determine if people are still present.

A year’s worth of data is collected and analysed to understand daily, weekly and seasonal occupancy presence and is used to create typical occupancy profiles for the building energy models. Swipe card data is available on a minute basis, each swipe is logged individually, whereas Wi-Fi data is available at a 5-min interval on a cumulative basis, showing the total number of connections at a particular time step. In **Figure 5.7** and **Figure 5.8**, show swipe-in and Wi-Fi data for a weekday in CH. The left graph shows the number of connections made or lost (negative) and the number of swipe-ins on 15-min interval, showing strong fluctuations during the day. The right graph shows the cumulative Wi-Fi connections, which are not available for the swipe-ins as its unclear (not logged) when people leave the building.

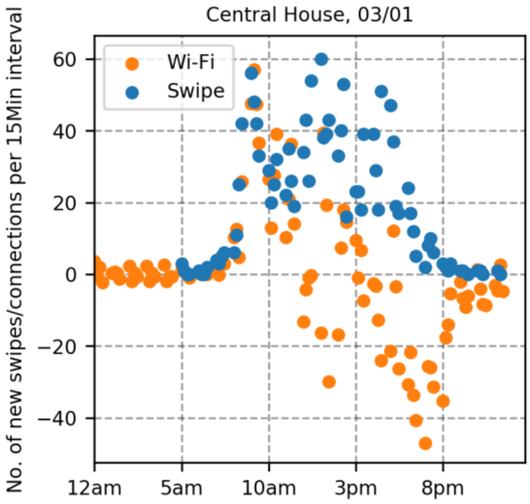


Figure 5.7: Number of new Wi-Fi connections and swipe-ins for a single day (03/01/2015) (negative value indicates people leaving the network compared to previous value).

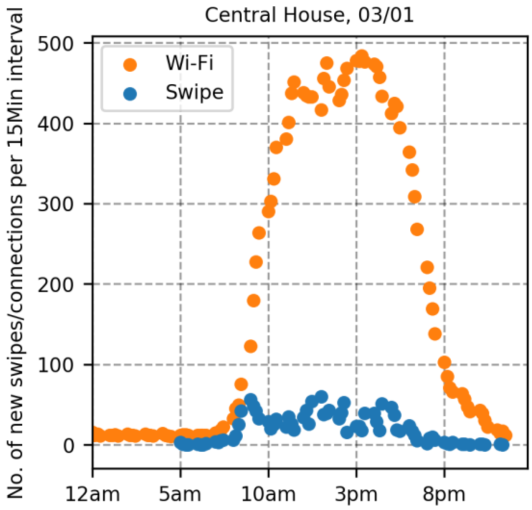


Figure 5.8: Number of Wi-Fi connections and swipe-ins for a single day (03/01/2015), where Wi-Fi connections are cumulative.

Occupancy influence on energy

Occupants have a large influence on energy consumption, after all, it is for them that spaces are conditioned and they are the ones that use most of the small power equipment and lighting in a building. Their significance was determined by correlating the total energy use obtained from the sub-metering and short term monitoring of energy use against the number of Wi-Fi connections, for a 3 month period, as shown in **Figure 5.9** for MPEB and CH.

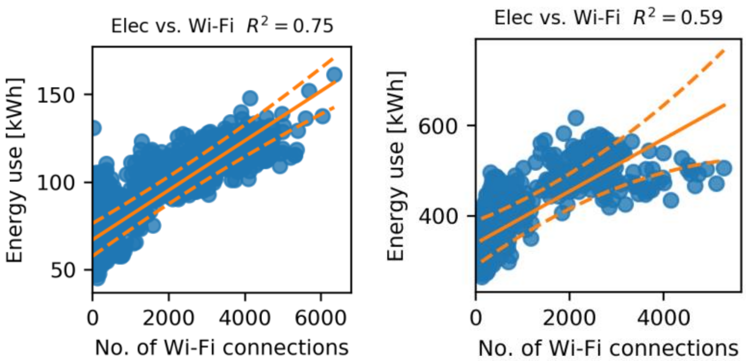


Figure 5.9: Occupancy correlated to total electricity use over a period of 2 months for CH (left) and MPEB (right), each point represent an hour of data.

The data indicates a significant correlation between occupancy and electricity use, in fact when the variation is calculated, it can be shown that 75% and 59% of the variation in electricity use is accounted for by the occupancy for CH and MPEB respectively. Similar results were shown by Martani et al. (2012), where they analysed two buildings and report variations of 69% and 63%.

5.4 Developing typical schedules of use

Lighting and equipment use and occupancy presence in building simulation are based on schedule of use, represented using values between 0 and 1. These schedules are then applied to certain spaces or space types, which are then multiplied by a power and lighting load (W/m^2) or occupant density (m^2/p) assigned to these spaces, to calculate the final load (W) or no. of people in a space at certain time of the day. For compliance modelling in the UK standard schedules are used, which may not necessarily be good indicators of reality. Schedules of use directly affect energy use and can have a large impact on the discrepancy between predicted and measured energy use. As such, for performance modelling during design, it is essential to determine future use of spaces, potentially based on previous experience or data from existing buildings. In existing buildings, operational data can be utilised to develop these schedules of use, either on; (1) measured occupancy presence data or directly based on (2) measured electricity use.

Logically, the use of lighting- and equipment electricity consumption will create more accurate lighting- and equipment use schedules, whereas occupancy data will create more accurate occupancy presence schedules. The use of occupancy data to represent lighting- and equipment schedules assumes that occupancy presence has a large influence on these types of energy use. Although this holds true in most buildings, it can differ significantly per building and is not always as strongly correlated, introducing a certain margin of error. Lighting- and equipment schedules are more important than the occupancy schedules as they will have a more significant influence on energy use, using this reasoning, the use of electricity data to develop use schedules would be the preferred option. Nevertheless, in both cases, it needs to be ensured that the developed schedules do not apply for all space types (depending on the granularity of data collected; whole building/floor/space). Using whole building collected occupancy data would be a better proxy for the schedules of use in office spaces, than in storage-, toilet- and kitchen spaces. In the case study buildings, the first approach was used for MPEB and CH, where Wi-Fi data was available, but an accurate breakdown of lighting- and equipment electricity was not. Whereas for Office 71, solely lighting and equipment electricity data was available and the second approach was used to develop schedules of use.

5.4.1 Developing typical schedules of use based on occupancy data

In building energy modelling, occupancy presence is represented by schedules with values between zero and unity, which are subsequently multiplied by the occupancy density assigned to different space types. To develop these, typical weekday and weekend day schedule were calculated using the swipe card and Wi-Fi datasets, shown in **Figure 5.10**. There is a large difference between the typical weekday and weekend day, occupancy presence during the weekend in MPEB and CH is 1/3rd and 1/6th of the weekday respectively. Wi-Fi data exhibits a much more continuous profile as it is cumulative (people stay connected to the Wi-Fi routers), and is therefore a better representation of occupancy presence. An interesting observation are the nearly identical profiles for occupancy presence in the two different buildings. Both are university buildings at the same university, but they have distinctly different space types.

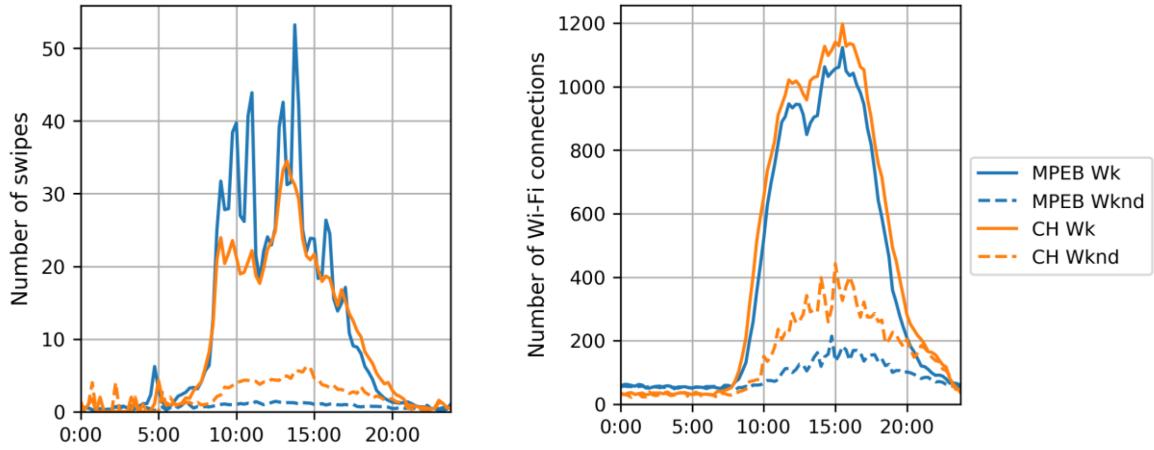


Figure 5.10: No. of swipes (left) and no. of Wi-Fi connections (right) for an average weekday and weekend day at a 15-min interval for MPEB and CH (over a whole year).

The typical Wi-Fi profiles were scaled to between 0 and 1, to be used as occupancy schedules in the building energy models for CH and MPEB. However, whilst analysing occupancy data, it became clear that a large variation exists in occupancy throughout the seasons as MPEB and CH are university buildings where occupancy is affected by university terms or semesters, evident from **Figure 5.11**.

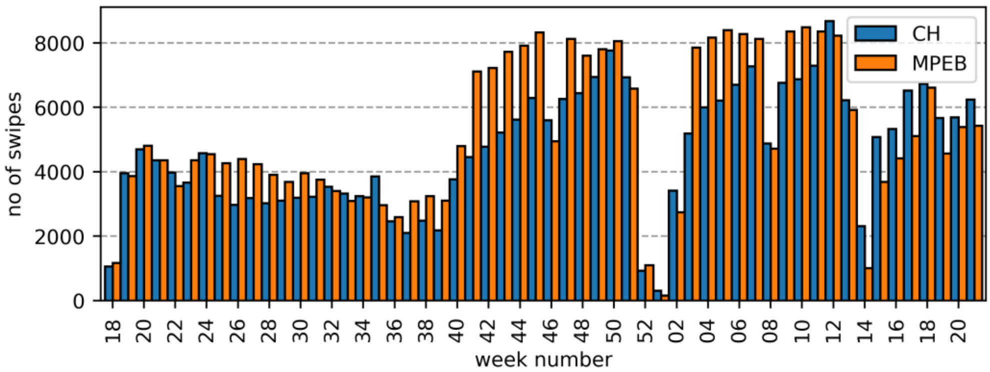


Figure 5.11: Daily number of swipe-ins at main entrance in CH and MPEB.

There is a low occupancy during the summer and early autumn, in particular during September, with increasing occupancy from October to December and January to the end of March (Easter). Furthermore, almost no Wi-Fi connections or swipe-ins occurred during the holidays (Christmas and Easter). This seasonal variation in occupancy needs to be accounted for in building energy modelling as occupancy has a significant effect on energy use. A monthly seasonal factor was calculated by taking the average daily maximum number of swipes per month, as shown in **Figure 5.12**, these values were then scaled to between 0 and 1. The seasonal factors were then multiplied by the typical weekday and weekend day schedules, creating a total of 12 different weekday and weekend day occupancy schedules, one for each month.

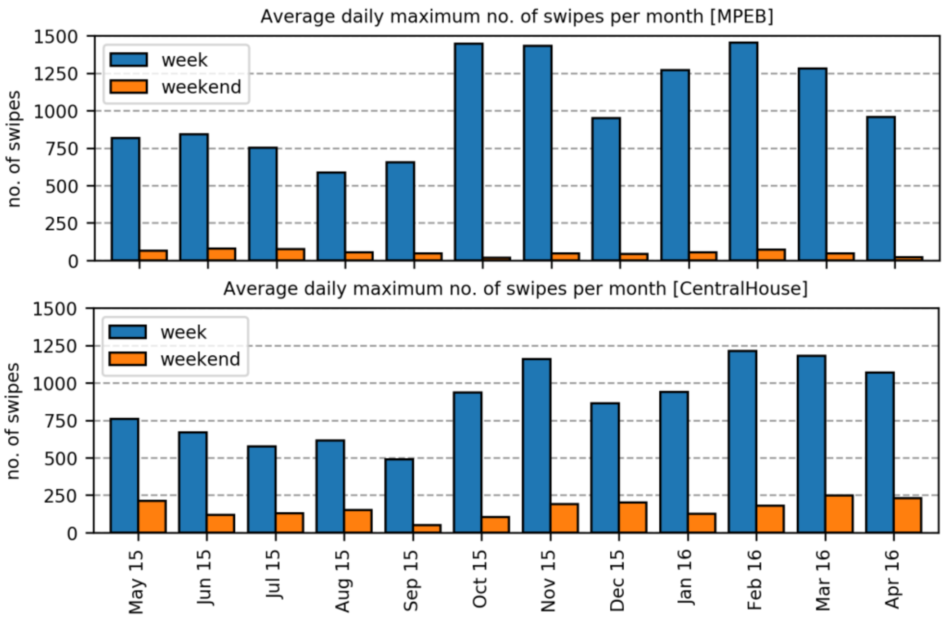


Figure 5.12: Average daily maximum number of swipe-ins per month for both MPEB and CH.

It would be possible to incorporate the collected occupancy data directly into the energy modelling software, i.e. instead of creating 12 monthly schedules, each half-hour can be represented by the collected data. This would however make the occupancy schedule considerably longer, would over fit the model and therefore only apply to a specific year. Instead, the schedules can take into account uncertainty within the data to understand the effect of variability in the schedules on performance.

At this point, the created schedules represent only the occupancy presence within a building. These schedules were used to create lighting and equipment schedules by introducing an out-of-hours baseload, which is the lighting or equipment electricity use during unoccupied hours relative to the average peak during the day. This requires making an assumption about the typical baseload for lighting and equipment use. By analysing typical lighting and equipment electricity profiles for the case study buildings, it was found that lighting and equipment energy use baseloads for Office 17, Office 71, CH and MPEB were; 20/25%, 15/20%, 20/65% and 65/85% respectively. These figures were used to create the lighting and equipment schedules, by taking the occupancy schedules and applying the baseload where the occupancy schedule factor is lower than the baseload. The resulting occupancy-, equipment- and lighting schedule for a typical weekday and weekend day in CH are shown in **Figure 5.13**. It was assumed here that the lighting and equipment schedules are slightly wider than the occupancy schedule. This is the base schedule to which the monthly seasonal factor was applied.

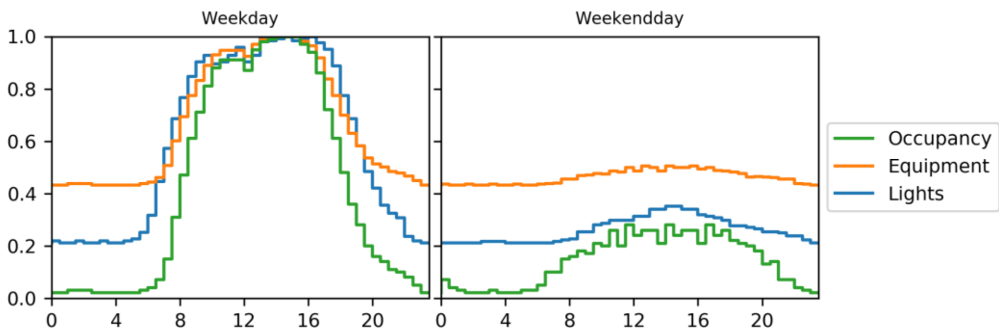


Figure 5.13: Occupancy, equipment and lighting schedules for a single simulation run for CH.

The schedules of use for CH and MPEB were created as previously described, whereas for Office 71, these were based solely on electricity data, and for Office 17 these were based on adjusted typical schedules taken from the National Calculation Methodology (NCM).

Variability in scheduling

The developed schedules of use based on the described method above were varied during the Monte Carlo simulations (each simulation had a slightly different schedule). Each schedule consisting of 48 (half hours) x 2 (weekday and weekend day) values were adjusted for each simulation, based on three parameters; horizontal offset, baseload, and the seasonal factor. In reality, occupancy presence, equipment and lighting energy use it not exactly the same each day and exhibit some form of variability, it is either higher or lower, or varies at different times of the day. As such, this margin of variability or uncertainty within the created schedules can be quantified by introducing variability in the three parameters for each simulation run. A **horizontal variability** is included by introducing an offset for both occupancy, lighting or equipment schedules. This offset is created by taken the previous or next value in the schedule and offset by 0, 1, 2 time steps, where a time step is 30 minutes. In addition, the equipment and lighting profiles are assumed to largely follow the occupancy profiles, but are furthermore based on a **baseload** factor, which is assumed to vary within 20% of the base value. Finally, the **seasonal variation** factor is also varied within 20% of the base value. The schedules of use directly affect the occupant density and equipment and lighting power density, therefore a vertical variability in schedules was neglected, as this is represented by the variation in these variables.

5.4.2 Developing typical schedules of use based on electricity data

An alternative to using occupancy data, is to use electricity data. This would require at least a breakdown into lighting and power for the building. For Office 17 and 71, equipment and lighting electricity use was disaggregated and available per floor. For Office 71, **Figure 5.14** and **Figure 5.15** show lighting- and equipment (power) electricity use disaggregated per floor. The 1st to 3rd floors consists mainly office spaces, where the ground floor includes a reception, meeting rooms and toilets, the basement consist of a large canteen, plant room and shower facilities. The 3rd floor is distinctly different from the nearly identical 1st and 2nd floors in terms of lighting and power use. The baseloads for lighting are around 30% of the daily peak load for the ground floor and 10% for the other floors, while the base loads for power are 30% of the daily peak load for the 3rd floor and 20% for the other floors.

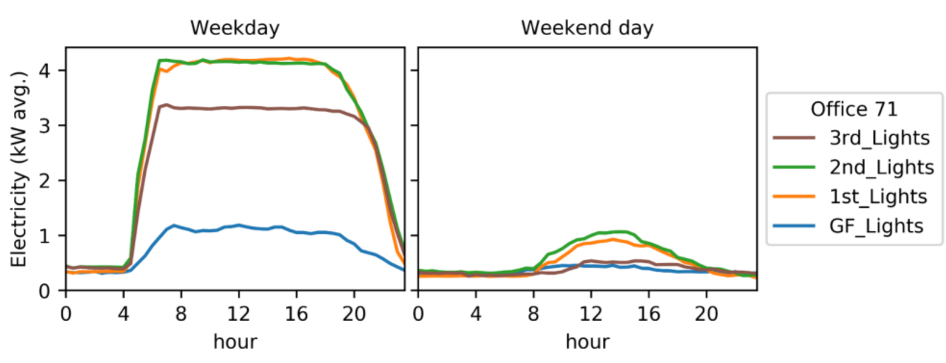


Figure 5.14: Lighting energy use on several meters for a typical weekday and weekend day (2013).

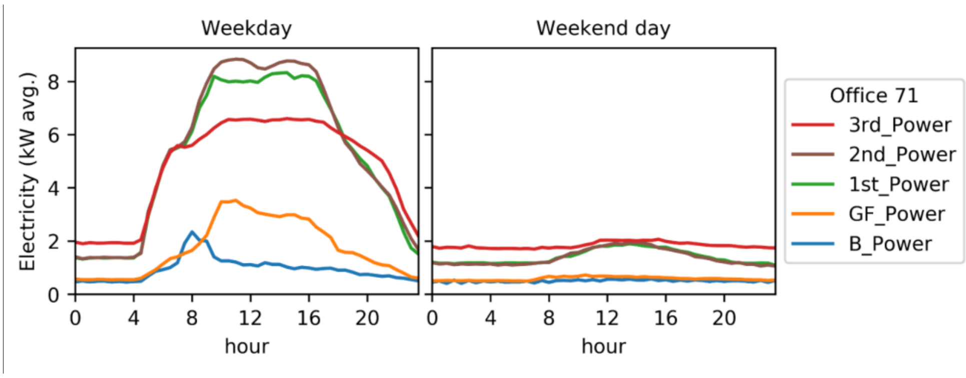


Figure 5.15: Power energy use on several meters for a typical weekday and weekend day (2013).

Besides different baseloads for lighting and equipment, the shape of the electricity profile is different. Lighting electricity use follows a much more on/off pattern, which can be due to the fact that lighting in large open office spaces are either on/off, whereas personal equipment may be turned off when people leave work.

The absolute electricity use schedules are scaled to between zero and unity as to be used in the building simulation software. Either separate schedules can be used for each floor, as shown in **Figure 5.16** or an average for the building can be created.

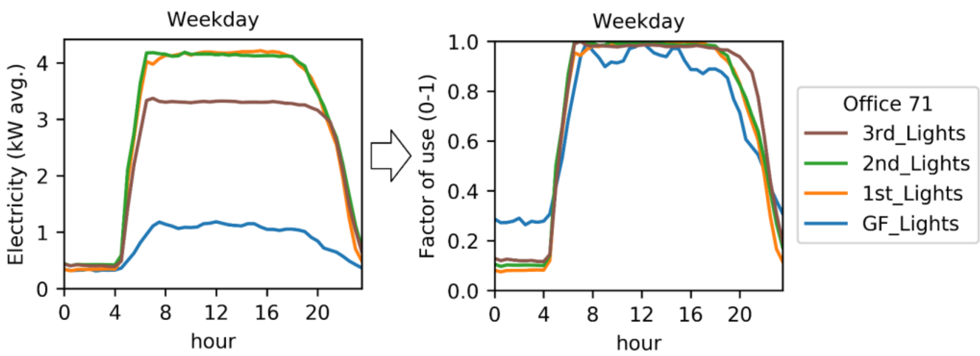


Figure 5.16: Lighting electricity use per floor scaled to between zero and unity, as a use profile.

In contrast to CH and MPEB, Office 17 and 71 did not show seasonal variation in energy use. As such, no seasonal variation for these buildings was taken into account. Similar to the seasonal factors calculated for the occupancy data, this can be done based on electricity data by calculating the average daily peak load per month.

Electricity as a proxy for occupancy presence

In many existing buildings, the availability of occupancy data is minimal and assumptions about their presence have to be made. To support these assumptions, it might be useful to base occupancy profiles on other available data, such as electricity use. Ideally, electricity for lighting or power in contrast to total electricity is available, as L&P correlates more significantly to occupancy presence. In this research, the difference between lighting and power could not always be accurately determined, due to the limited amount of data available for lighting and power separately and due to the quality of data. Therefore, combined lighting and power were compared with occupancy data from Wi-Fi to understand how schedules can be established as input for building simulation. **Figure 5.17** shows the correlation between measured Wi-Fi connections and lighting and power electricity use for different floors in CH.

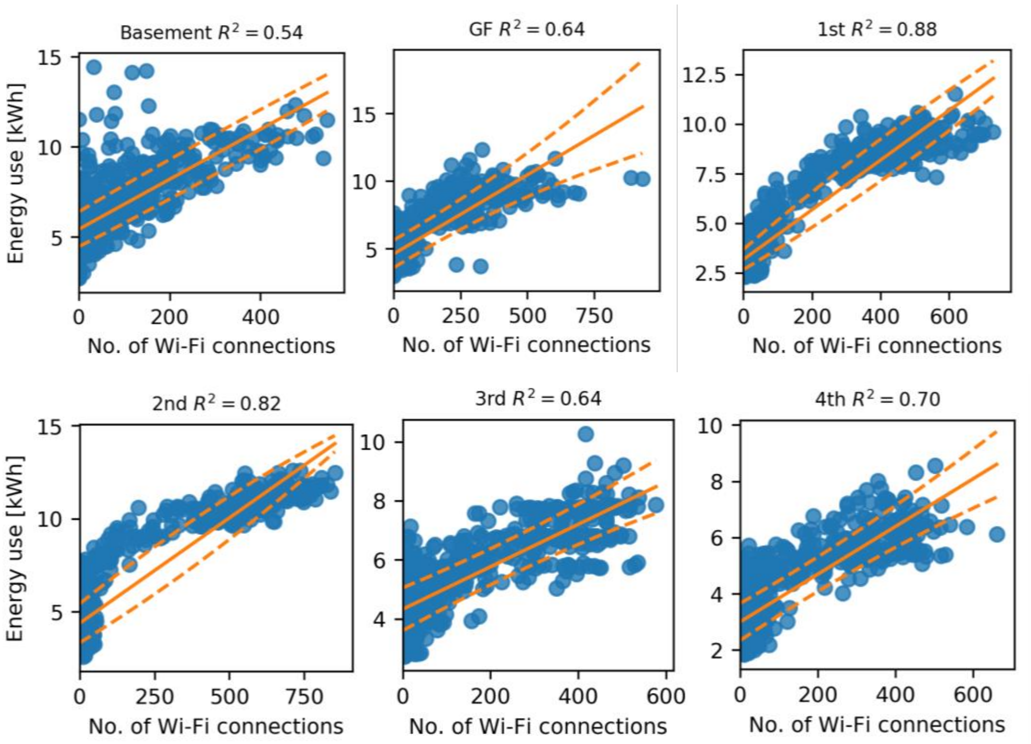


Figure 5.17: Lighting and power energy use for different floors in CH in relation to number of Wi-Fi connections on these floors for the month September 2016 (at an interval of 15-mins).

There is a significant correlation between occupants and lighting and power on all floors, with some variability, which arises mainly due to the type of loads on the lighting and power meters (some loads are less determinant on occupancy). High correlations indicate that occupancy presence and lighting and power energy use follow a similar trend, this is shown in **Figure 5.18**, where both are compared by plotting them as a typical weekday and weekend day and scaling their values to 0 and 1.

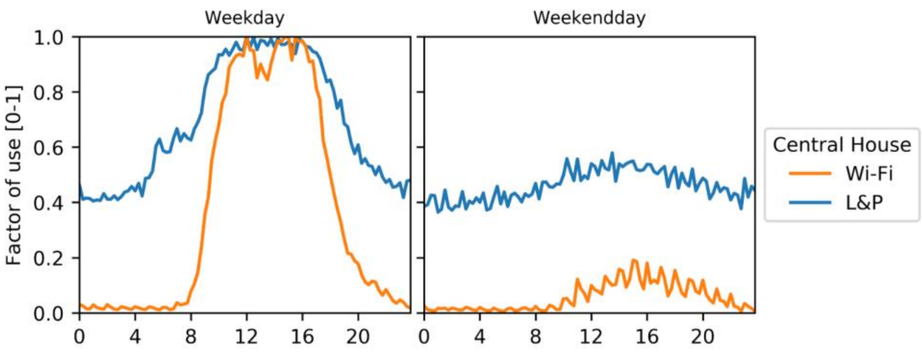


Figure 5.18: Scaled lighting and power electricity use and Wi-Fi connections for a typical weekday and weekend day in CH.

The graph shows that the datasets follow a similar pattern, but that occupants, as expected, are not present during the night in a university building, in contrast, L&P electricity use is still significant, with a baseload of around 40%. The baseload in MPEB is even more significant for L&P, around 80%, as shown in **Figure 5.19**.

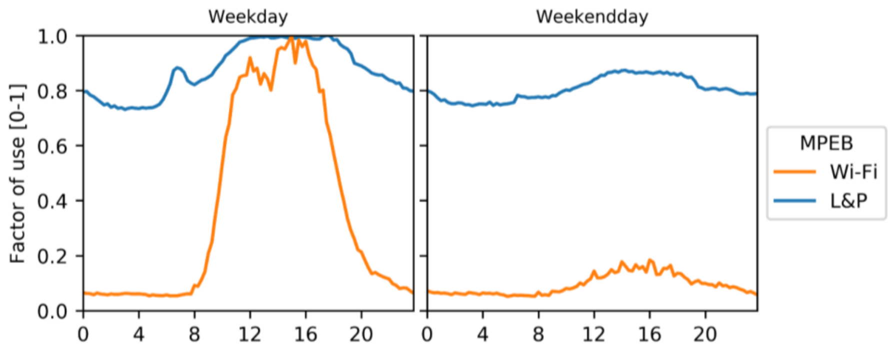


Figure 5.19: Scaled lighting and power electricity use and Wi-Fi connections for a typical weekday and weekend day in MPEB.

To recreate the Wi-Fi patterns solely based on the lighting and power electricity use patterns, the L&P profiles were first scaled to between 0 and 1 (see previous graphs, but then for the whole time series), the baseload is then subtracted from the values in the time series and negative values are set to 0. Then, the time series is again scaled to between 0 and 1, and a typical weekday and weekend day were calculated, as shown in **Figure 5.20**.

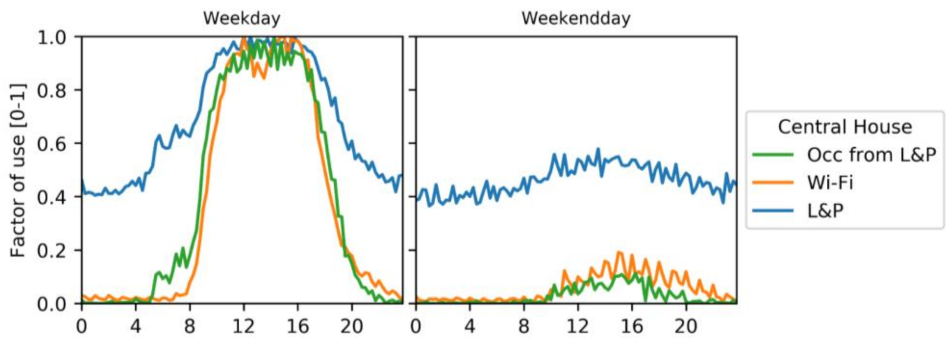


Figure 5.20: Scaled recreated occupancy profile based on L&P, L&P electricity use and Wi-Fi connections for a typical weekday and weekend day in CH.

As can be seen, the newly created occupancy profile solely based on lighting and power electricity represent the actual Wi-Fi data surprisingly well. However, the accuracy of this method is very dependent on the baseload, how large it is in proportion and what its assumption in the calculation. For CH, the baseload is half that in terms of proportion to that of MPEB. Therefore, some of the underlying aspects of the profile are retained through this transformation. This becomes evident when applying different baseloads for the same scaling calculation in MPEB, as shown in **Figure 5.21**.

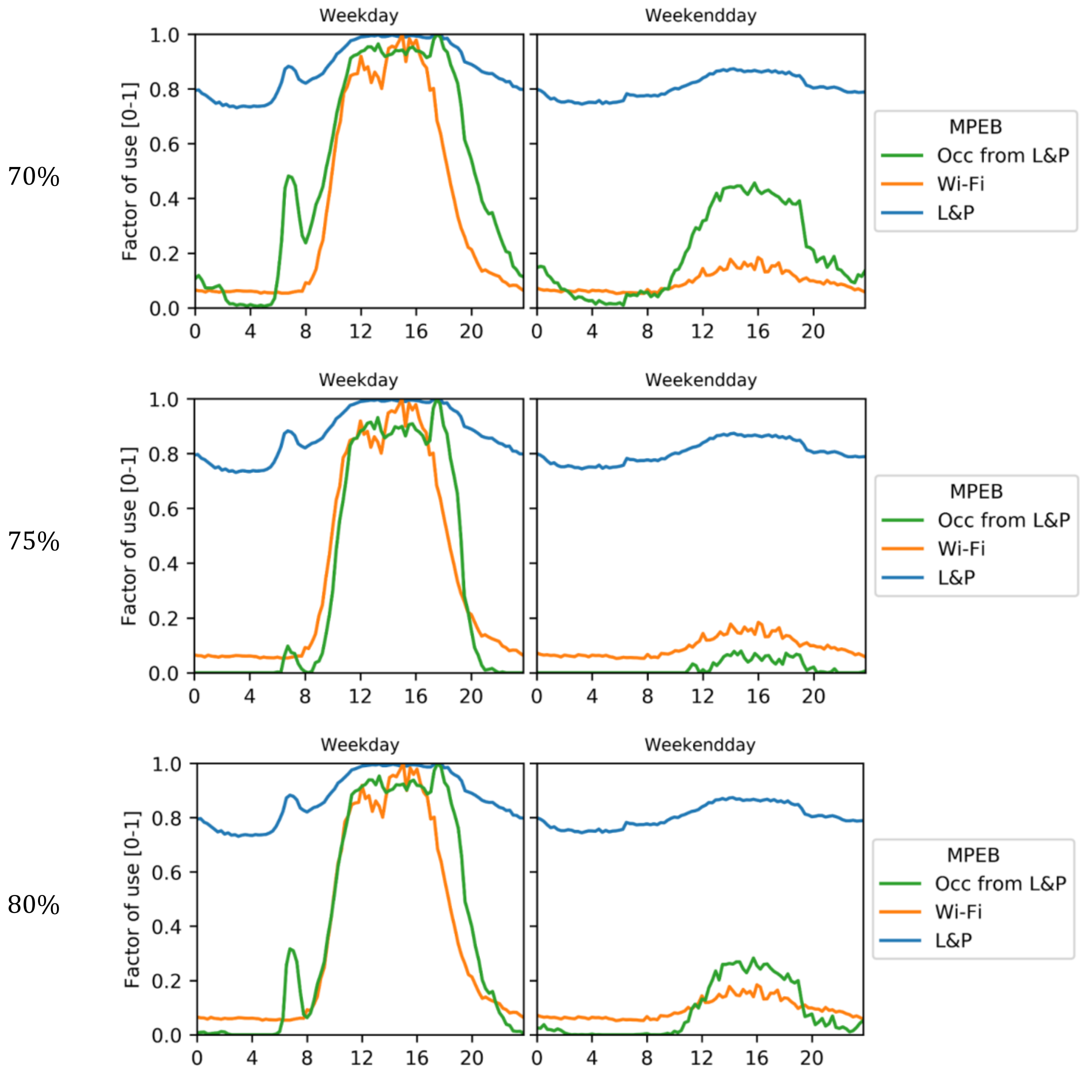


Figure 5.21: Scaled recreated occupancy profiles using different baseloads (from top to bottom 70%, 75%, 80%) based on L&P, L&P electricity use and Wi-Fi connections for a typical weekday and weekend day for MPEB.

Using 75% as the baseload in MPEB gives the best representation of the Wi-Fi connection profile, but the baseload is strongly determinant of the variability between the profiles. In addition, the calculation assumes negative values to be 0, which is likely to be more accurate than the calculated baseload based on the Wi-Fi connections, which is around 8% in MPEB, due to computers/servers that are continuously connected to the Wi-Fi. Furthermore, in MPEB there is a spike in the lighting and power electricity use that should be removed and interpolated, this spike is likely due to the power surge in turning on parts of the system or electric water heaters that turn on during the early morning. In conclusion, occupancy presence can be represented by lighting and power electricity use profiles considerably well, but care should be taken where high baseloads exist.

5.5 System and environmental performance

Only for MPEB, a building management system is in place logging data for thousands of sensors placed throughout the building. The building management system is connected to the DemandLogic[5] platform, which is a software service that provides data analytics and an online collaboration platform for insights into building performance and is used for improving staff comfort, condition-based maintenance and reducing energy consumption. The platform provided easy access to the data, which proved to be a valuable resource in understanding system behaviour, but has many more capabilities in place that could prove useful for more in-depth studies focussed on system calibration and real-time building energy forecasting.

5.5.1 Space temperature set-points

The building management system in MPEB logs data regarding the temperatures measured by the fan coil units (setpoint temperatures and space temperatures). Such data provides valuable information about achieved comfort levels and setpoint temperatures for different space types. It was found that typically, setpoint temperatures in the spaces are not being met, to illustrate, **Figure 5.22** shows the space and setpoint temperatures and heating and cooling demand of a fan coil unit in a computer lab for two weeks in June 2017.

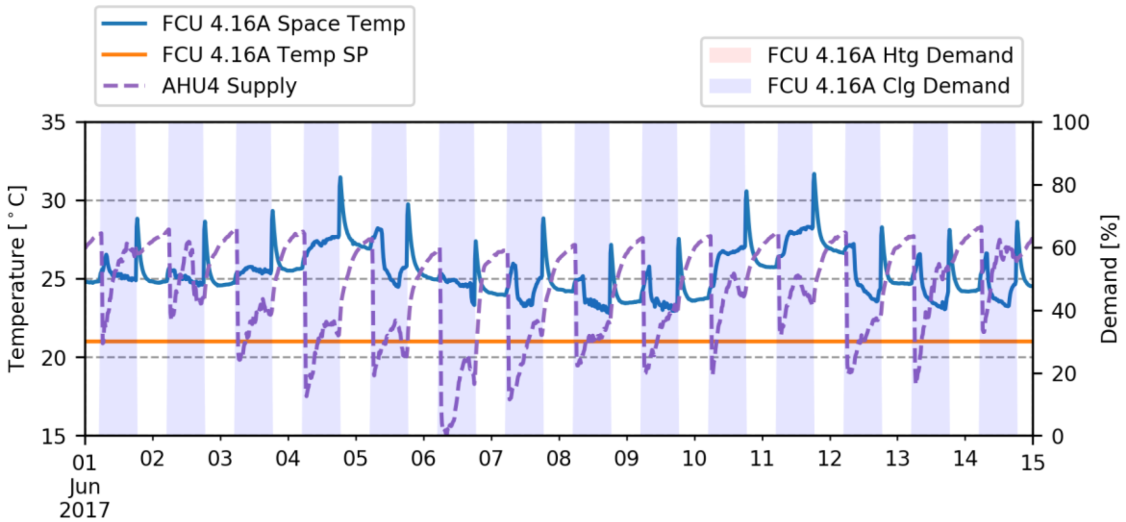


Figure 5.22: FCU measured variables for space, setpoint temperature and required heating and cooling demand in a lift lobby.

The space temperatures exceed the setpoint temperature by more than 2 degrees at its minimum and are on average around 25 °C, 4°C higher than the setpoint temperature. During occupied hours, the fan coil system is operated in cooling mode, trying to bring down the temperature to the setpoint, but is ever failing to do so. Interestingly shortly after the fan coil unit turns off (which is operated on a time schedule), the temperature quickly rises within the space, before dropping down during the night to an

[5] http://www.demandlogic.co.uk/

even lower temperature than during the day. The quick drop in temperature is likely due to the lower supply temperature from AHU 4, which is also supplying conditioning to the space. This behaviour is similar for the other lift lobbies, which are located at the front of the building on each floor. During the building audits, it was found that the lift lobbies originally did not have any provision for FCUs, they were added later due to the unexpected high internal gains from people working in these space (which they were initially not designed for).

In contrast to the lift lobbies, the temperature in meeting rooms is controlled much more tightly, as shown in **Figure 5.23**. The setpoint temperature is set at 19 °C and increased to 20.5 °C in the second week of June 2017, coincidentally the space temperatures increased. This space has primarily a heating demand, even during the summer months due to low internal gains.

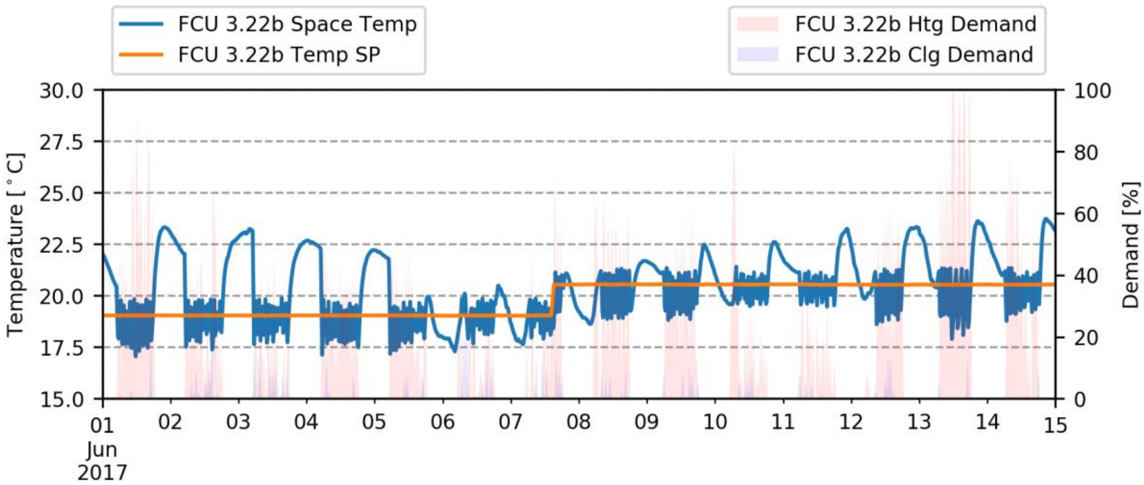


Figure 5.23: Space-, and setpoint temperature, heating- and cooling demand in a meeting room.

Similarly to the previous graph, **Figure 5.24** shows an office space where three fan coil units provide heating and cooling to the space. The three FCUs, their respective temperatures, and heating and cooling demand are plotted separately to illustrate that small variations exist between both the space and setpoint temperatures measured at the different fan coil units. At FCU 4.07a there is a large difference between the setpoint and space temperature achieved, it is continuously in cooling mode, whereas FCU 4.07c is much closer to the setpoint temperature and its cooling demand is a lot smaller the plotted days. The setpoint temperatures between the three FCUs differ as well, however they are never working against each other, heating and cooling does not occur at the same time.

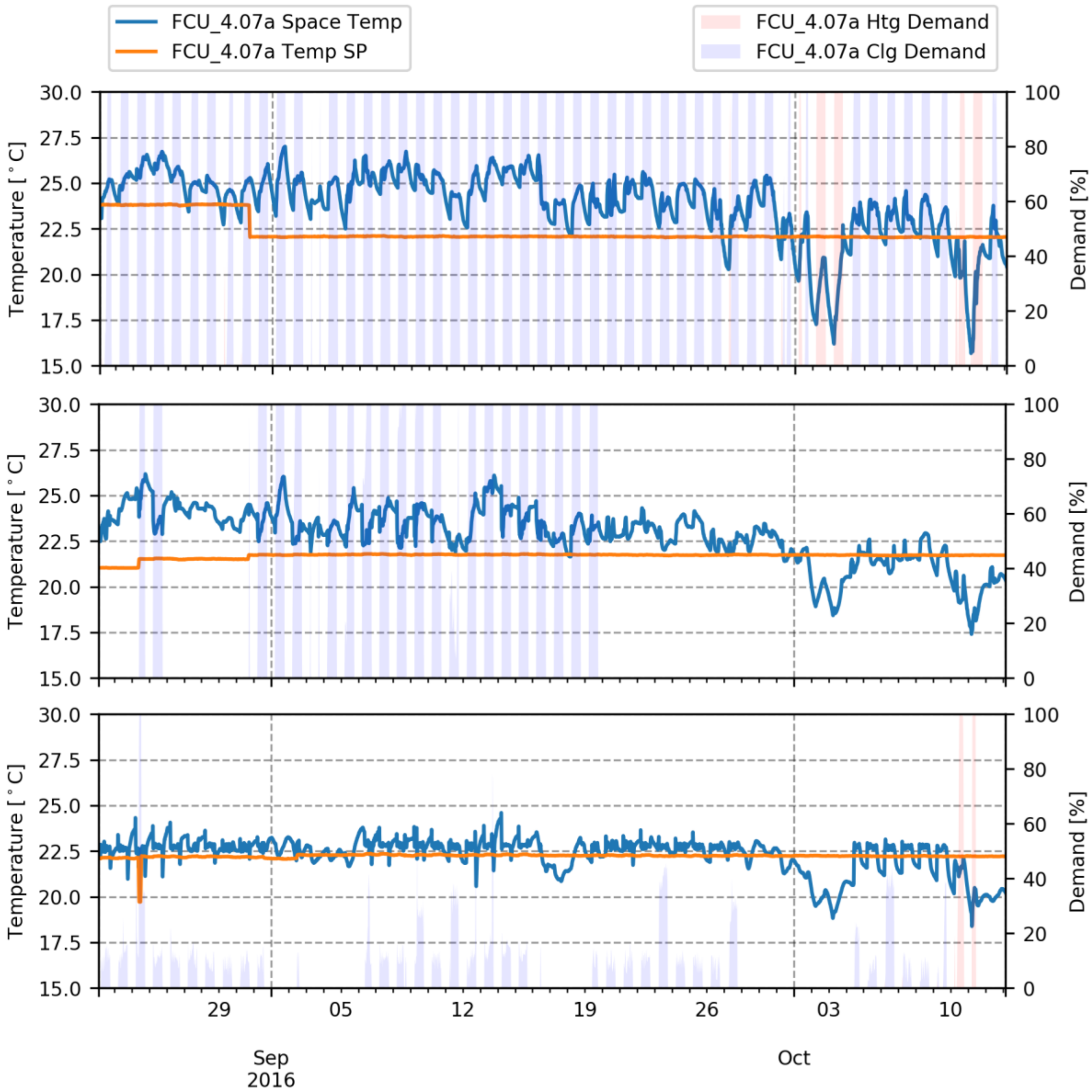


Figure 5.24: FCU 4.07a-c measured variables for space, setpoint temperature and heating and cooling demand in space 420 Systems staff office (from top to bottom; FCU 4.07a, b and c)

An example of where systems seem to operate against each other is shown in **Figure 5.25**, where the space temperature in G01B Machine room is compared to the supply air temperature of AHU 3. G01B is a server room, where computer clusters are located that dissipate large amounts of heat. AHU 3 supposedly provides tempered fresh air at around 22 °C (higher temperatures indicate that the AHU is off), during occupied hours to several spaces on the ground floor, including G01B, the space temperature in the server room is however controlled at 18 °C. An additional three fan coil units 209A, C and D are located in this space, distributing the cooling load among them, cooling down the incoming air supplied by the AHU.

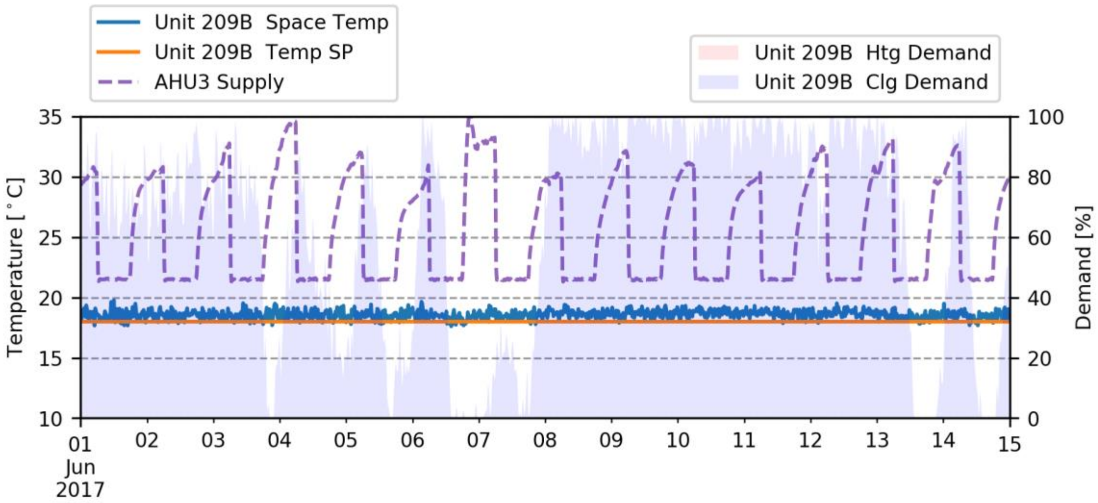


Figure 5.25: FCU measured variables for space, setpoint temperature, heating and cooling demand in G01B Machine room and AHU 3 supply air temperature.

To represent the actual situation with the building energy model, it is important to resemble the achieved space temperatures. Although supply air design temperatures for the air handling units are available, and design setpoint temperatures are given for different space types, the actual situation seems to differ slightly from those. Setpoint temperatures are not met in many of the spaces, to replicate the behaviour, space temperature distributions are analysed during occupied and unoccupied hours over a longer period to understand what the actual achieved temperatures are in the spaces, shown for computer labs and offices as shown in **Figure 5.26** and **Figure 5.27** respectively. Both computer labs and office spaces show relatively stable space temperatures throughout the year, fluctuating within 20-26°C and with several degrees of difference between individual spaces.

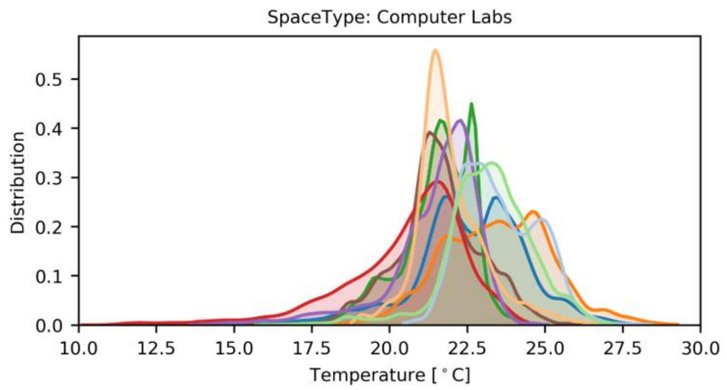


Figure 5.26: Kernel density estimation of the space temperatures in 9 computer labs for a year.

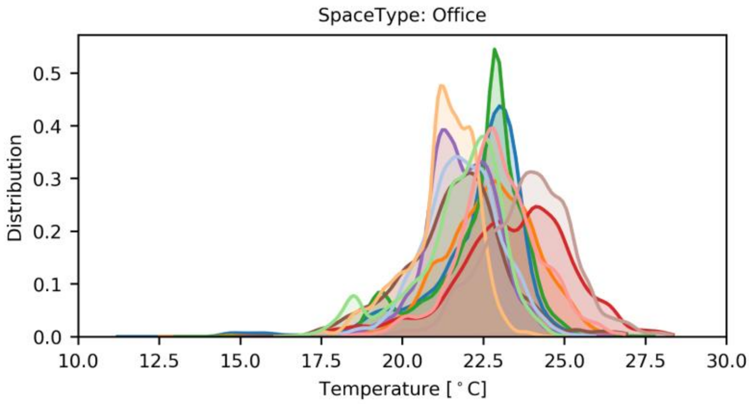


Figure 5.27: Kernel density estimation of the space temperatures in 11 office spaces for a year.

Although the distribution plots are helpful in understanding how different spaces compare to each other, it is difficult to discern the monthly variation of temperature in these spaces. The multiple peaks in the profiles are an indication that this variation exists, due to changed setpoint temperatures, which follows from analysing monthly boxplots of space temperatures against average setpoint temperatures during those months, as shown in **Figure 5.28.**

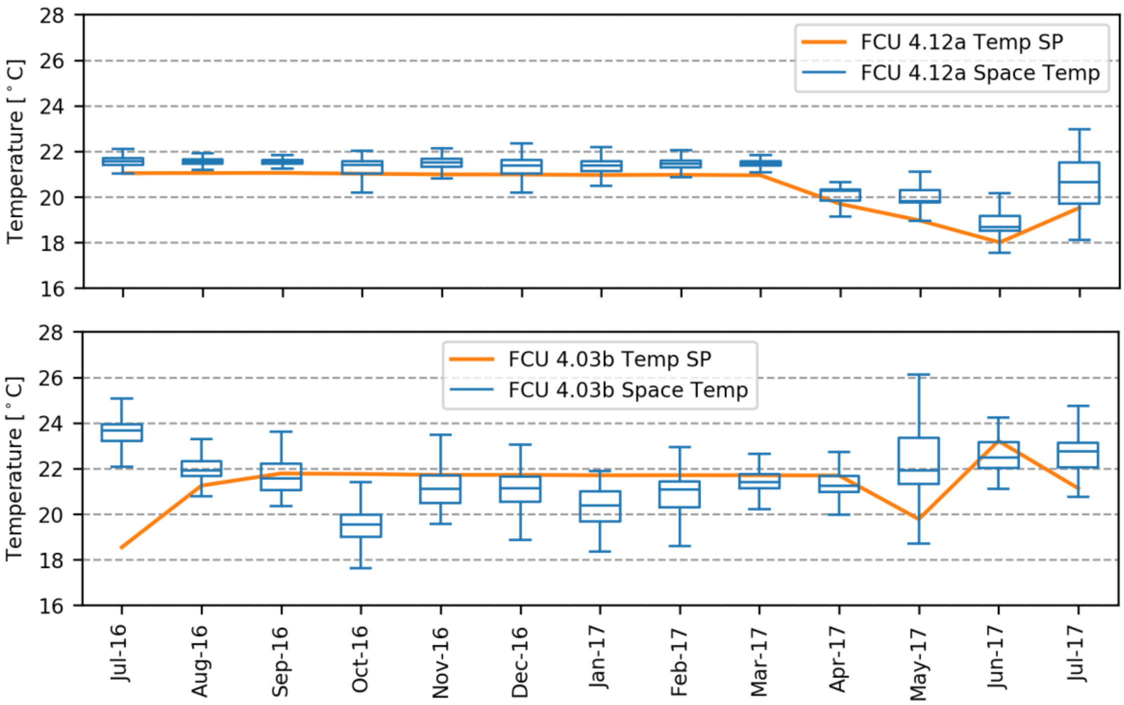


Figure 5.28: Space temperatures as boxplots distributions per month for two computer lab spaces (406 and 417) for a whole year, with the setpoint temperature taken as average per month.

The measured space temperatures show a large variation between individual spaces and space types. Typical setpoint temperatures for the different space types can therefore not be defined with certainty, therefore it was opted to use the design setpoint temperatures where applicable and introduce uncertainty within the setpoint temperatures for parametric simulation, in order to understand the effect of such variations through sensitivity analysis.

Similarly, for Office 17, 71 some air conditioning exists in the offices, meeting rooms and canteen, the setpoint temperatures in the simulations are fixed. The space heating and cooling setpoint temperatures are shown in **Table 5.4**.

Table 5.4: Space heating and cooling temperature setpoints (°C) in Office 17, 71 and MPEB.

Space type	Office 17	Office 71	MPEB
Office, Meeting, Reception	22 – 24	22 - 24	22 - 26
Computer cluster, Lift lobby, Lecture theatre			22 - 26
Server	22 - 22		22 - 22
Canteen	22 - 24	22 - 24	
Lavatory, Shower, Circulation			18
Laboratory, Workshop			24 - 26

In CH, occupants have manual control over the operation of fan coil units (heating/cooling) in the spaces. Temperatures varies widely in spaces with fan coil units. To replicate this behaviour and the uncertainty of setpoint control the temperature setpoints for heating and cooling is varied during the parametric simulation. The setpoint is randomly selected by temperature control where the interval between the heating and cooling setpoints is at most 2 degrees, varied between 21 and 25 degrees, heating and cooling setpoints respectively, feasible options are shown in **Table 5.5**.

Table 5.5: Setpoint (SP) temperatures options for heating and cooling for CH.

SP (°C)	0	0.5	1	1.5	2	2.5	3
H, C	23.5, 23.5	23.5, 24	23.5, 24.5	23.5, 25	23.5, 25.5	23.5, 26	23.5, 26.5
H, C	24, 24	24, 24.5	24, 25	24, 25.5	24, 26	24, 26.5	24, 27
H, C	24.5, 24.5	24.5, 25	24.5, 25.5	24.5, 26	24.5, 26.5	24.5, 27	

The schedule for the heating and cooling setpoints are typically controlled from 7am to 7pm.

5.5.2 Supply and return air temperatures

Finally, data is available on the performance of the five air handling units in MPEB. Providing information on the demand profiles of the heating, cooling and frost coils, supply and return temperatures and fan speeds, **Figure 5.29** shows this data for AHU 1 for the first week of June 2017. The air handler operates under time control, from 7am to 7pm, both during the week and weekend, as are most of the systems in MPEB. Noticeable here are the supply and return temperatures, which fluctuate significantly during the day and night. AHU 1 is controlled to maintain a supply temperature of 18 °C according to the design data found in the O&M manuals, the graph however clearly shows that it is instead supplying a temperature of 19 °C.

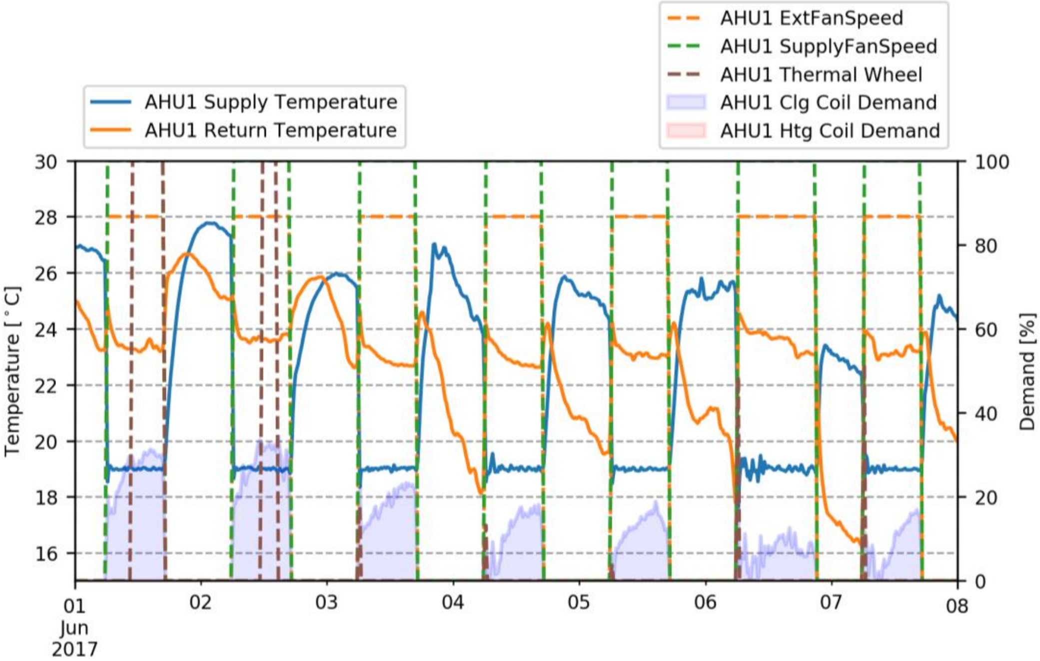


Figure 5.29: Performance data for AHU 1, including both data on demands and fan speeds for the first week of June 2017.

According to the design data, AHU 1 supplies fresh air to the lecture rooms, study rooms and computer labs on the first floor. AHU 2 supplies a modulated supply air temperature to maintain a return air temperature of 22 °C, supplying spaces in the basement, such as workshops and restrooms. However temperatures can be limited through a setpoint manager, set to 20-24 °C as observed during the audit. AHU 3 supply air temperature is modulated to maintain a return air temperature of 22 °C, supplying laboratories on the ground floor and AHU4 supply air is modulated to maintain a return air temperature of 20 °C, supplying lavatory spaces. Finally, AHU 5, the largest air handler, is controlled to constantly supply 20 °C to spaces on the 2nd to 8th floors. The actual supply and return air temperatures for the AHUs are shown in **Figure 5.30** and **Figure 5.31** respectively. The green shade in the graphs illustrate the weekend and the blue shade illustrates the occupied hours during the weekdays, set at 7am to 7pm. MPEB operates under the same conditions during the weekend, although with significantly less occupancy.

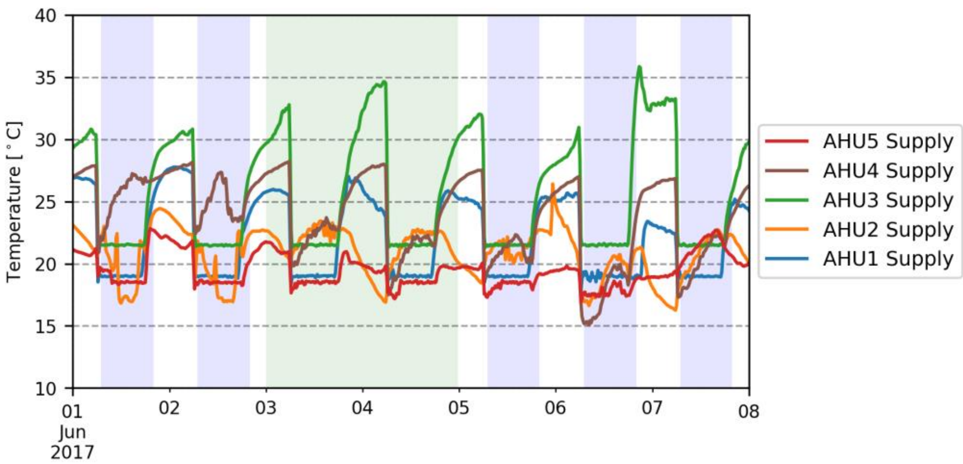


Figure 5.30: AHU supply air temperatures in MPEB for the first week of June 2017.

The supply air temperatures for the AHUs fluctuate significantly between occupied and unoccupied hours. More so for those that are controlled based on the return air temperature (AHUs 2 to 4). AHU 1 and 5 seem to be maintaining a steady supply temperature close to 19 °C during the day, whereas AHU 3 supplies a steady temperature at around 22 °C. Supply temperatures for AHU 2 and 3 however vary significantly more during the day. AHU 3 does have a higher return air temperature than that defined in the O&M manual, 23 °C instead of 22 °C. While, AHU 2 maintains its design return air temperature of 22 °C, with slight fluctuations.

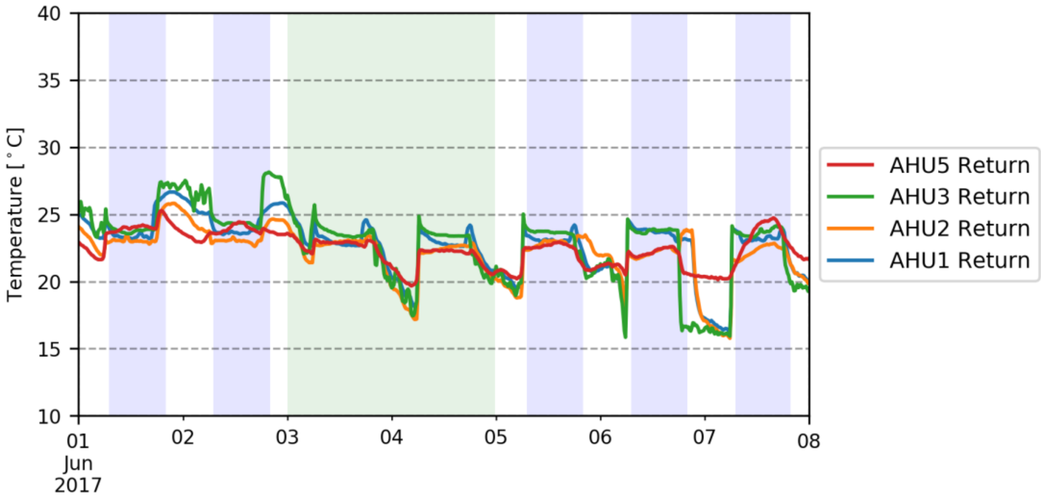


Figure 5.31: AHU return air temperatures in MPEB for the first week of June 2017.

A longer period is analysed to make sure the plotted temperatures are not a weekly anomaly. **Figure 5.32** and **Figure 5.33** show kernel density estimations of the AHU supply and return air temperatures respectively. The graph explains where temperatures most often occur, data is separated as occupied (7am to 7pm) and unoccupied to see how the system controls the supply air temperature. Supply and return air temperatures were analysed for several months in different seasons to see if there is any change. Similar to the previous graphs, the temperatures are stable during occupied hours for AHU 1 and 5, which are explicitly controlled on their supply air temperature, and AHU 3, which is modulated to maintain a return air temperature of around 23 °C.

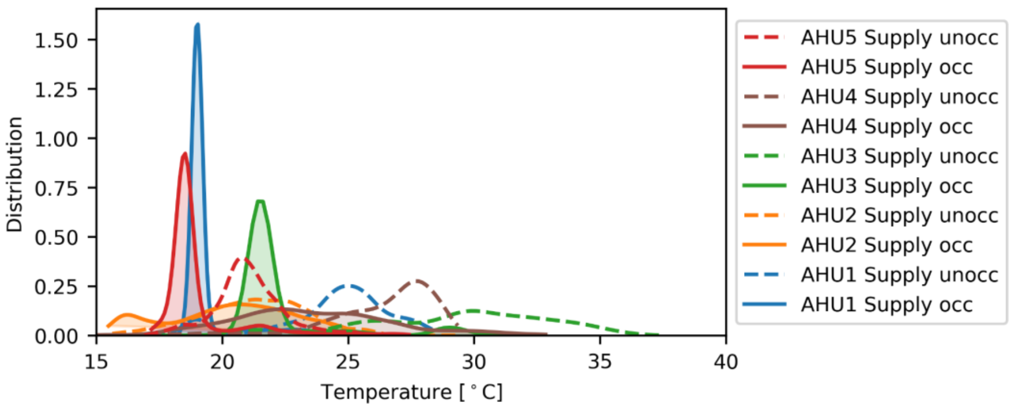


Figure 5.32: Kernel density estimation of AHU supply air temperatures in MPEB for June and July 2017.

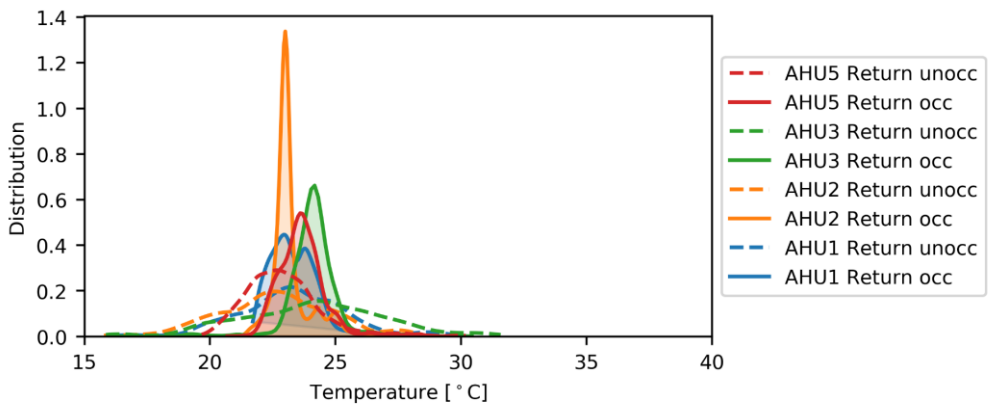


Figure 5.33: Kernel density estimation of AHU return air temperatures in MPEB for June and July 2017.

Analysis of system performance data illustrated the potential benefit of including such data for model calibration purposes, a model calibrated not solely on energy use, but also on system performance. This is done by hard-sizing the system components using available design data or commissioning data from O&M manuals and comparing supply and return temperatures, and the cooling and heating demand of systems with those predicted by the model. The model will then more accurately represent the actual situation in terms of systems performance, and indirectly so for the energy consumption of systems.

5.6 Summary

Collected data proved valuable in supporting assumptions in the building modelling process. In particular, the development of occupancy, equipment and lighting schedules is essential for model calibration. In addition, environmental and system performance data was analysed to establish system settings and set-point temperatures. These variables can be included as targets during model calibration, aligning for example the indoor space temperatures predicted by the model with those measured will further increase the accuracy of a model. However, this does add another layer of complexity, which may not be worthwhile depending on the objective of the model calibration process.

Energy use

There are large differences in energy use between the four case study buildings, specifically Office 71 and the two university buildings indicate that building type can have a significant effect on use patterns. CH and MPEB are operated every day of the week with a large variation during the seasons, whereas Office 17 and 71 are not conditioned during the weekend, even though it is typically occupied by a small amount of people. Equipment energy use in both Office 17 and 71 were found to have smaller equipment power baseloads than the university buildings, around 20-30%, compared to 50% and 60-80% for CH and MPEB respectively. For lighting energy use, the baseloads are very similar for Office 17, Office 71 and CH, around 10-15%, whereas MPEB showed baseloads of around 50%, this was found to be due to equipment being connected to the lighting meters (different from the distribution schedules) as it was observed that lighting was not being left on during the night.

The case study buildings were compared to typical, albeit outdated, benchmarks, which did not provide a very meaningful comparison and highlighted the need for more comprehensive benchmarking techniques. First, the university buildings, although very similar to office buildings, provide additional space function, and cannot be classified as office buildings. Second, each building was significantly different.

Understanding if the buildings fall within the typical or good benchmarks as presented is not straightforward as they have additional services such as the large server rooms in MPEB. The office buildings were more similar to the benchmarks (as the benchmarks are also intended for offices), but determining their efficiency in operation or performance against the stock was difficult. As such, their electricity patterns were also compared at a higher level of data granularity by computing their representative load pattern and other load shape metrics. These proved to more effectively identify operational differences within the buildings. Total energy use baseloads were significantly different, from high to low, MPEB, CH and Office 17 and 71, with baseloads of about 60%, 45%, 30% and 10% respectively. In addition, on-hour duration for MPEB was determined to be around 13 hours, while the other buildings fluctuate around 16-18 hours. This was mainly due to the steep difference in energy use between the day and night in MPEB, whereas the profiles for the other buildings are much smoother and therefore calculate longer on-hour durations.

Occupancy presence

Occupancy presence was analysed from collected swipe card access and Wi-Fi data for MPEB and CH. The datasets were utilised to build occupancy profiles. A strong seasonal variation was identified and accounted for in the models by including a seasonal variability parameter, which adjusts the schedules on a monthly basis. Furthermore, a strong correlation exists between lighting and power with occupancy presence from Wi-Fi data. Based on this observation, lighting and equipment profiles were based on the occupancy profiles. It was also determined that occupancy profiles could be derived from lighting and power electricity use alone, by replicating them and validating them against the Wi-Fi data. However, determining the right baseload of electricity was essential in the computation of these profiles.

System performance and environmental data

System performance was available for mainly MPEB, which contains a multitude of systems. An online platform was utilised to analyse system performance such as the air handlers, chillers and fan coil unit operation. Environmental data was available for both MPEB, in particular space temperatures and heating setpoints on FCUs were analysed to understand distributions and operation. These variables proved useful in understanding if the systems operated according to design specifications. Space temperatures from FCUs were analysed, although it was difficult to discern in which spaces the FCUs were located, it showed that spaces were conditioned very differently, with some being controlled at a narrower range than others. Some of the data was invaluable in determining typical set points (e.g. the setpoint temperature in the server rooms significantly affect chiller energy use).

General applicability

The data collected and the assumptions they have informed is case study specific. Nevertheless, the approach used here is applicable to other buildings. Furthermore, there is potential for such approaches to be carried out on a larger scale for the purpose of collecting typical use profiles in different building types, which can then be directly used to inform or compare to during both the design of new buildings or operation of existing ones. For example, the calculation and collection of typical energy use profiles (e.g. using the representative load pattern approach) can inform which building types have a high baseload, this can then be taken into account during performance modelling approaches during the design of new buildings.

6 QUANTIFYING THE IMPACT OF UNDERLYING CAUSES OF A DISCREPANCY

Initial base case models were built for the case study buildings, subsequently examined to see if initial assumptions were close to measured data. Several iterations were necessary to adjust the models to achieve a closer representation of the existing buildings. Manual calibration focussed on removing modelling errors and mitigating discrepancies between predictions and measurements, due to differences between design specifications and observations during energy audits and analysis of measured data. More specifically, this involved including specific holidays in the prediction models, establishing base loads of lighting and equipment electricity use, disaggregating energy end-uses for juxtaposition, determining efficiency of system components and introducing seasonal use factors. Aligning predicted and measured energy use, such as replicating system behaviour and resolving modelling errors are tasks that are difficult to implement automatically, although, automating the analysis can improve time efficiency when carrying out these tasks. Subsequently, parametric simulation was employed using the base case models, and variability in the input parameters was taken into account and thereafter quantified using uncertainty analysis. The numerous simulation runs form a solution space based on computed sets of input parameters sampled using Latin hypercube sampling. Measurements fall within the computed uncertainty ranges of the solution space at a low level of data granularity, but in certain cases fall outside of these ranges for specific monthly energy end-uses. Sensitivity analysis was utilised to determine the impact of inputs on outputs. Lighting and equipment power densities were typically the most significant, where a large uncertainty exists for the equipment power density in a space. In addition, heating set-point temperatures (in Office 71 and CH) and cooling set-point temperatures in CH were significant in influencing energy use. In MPEB, the large server rooms contributed to most of the energy use in the building. Finally, the calibrated models were used to assess the impact of using typical assumptions as specified by the National Calculation Methodology (NCM) on model predictions. These assumptions give an understanding of the influence on energy use during the design stage of a building when modelling to comply with Building Regulations Part L2 (HM Government, 2013). Simplifications to the calibrated models quantified the effect of design stage assumptions on predicting energy use and indirectly the regulatory performance gap.

6.1 Introduction

Hypothetically, utilising extensive knowledge of the building, its systems and patterns of use will improve model accuracy considerably, in contrast to when such information is scarcely available. To prove this, data was collected for four case study buildings to inform model development and calibration. In this chapter predicted and measured energy performance are compared and the manual calibration of the initial models towards the base case model is discussed, in the next chapter these base case models were used in the automated calibration process. This chapter also discusses results from sensitivity and uncertainty analysis. The parametric simulations were utilised for performing correlation and regression analysis, whereas the meta-models (of which their development is described in the next chapter) were used for variance-based global sensitivity analysis, as it required the computation of a large amount of samples, many more than which were simulated. In the calibration methodology, these activities are represented by the white and green shaded boxes shown in **Figure 6.1**.

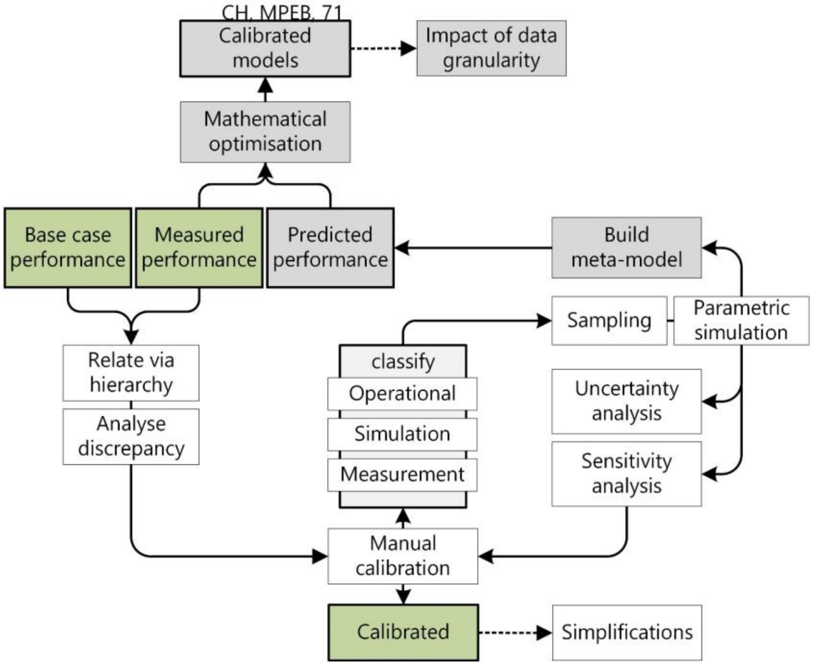


Figure 6.1: Calibration methodology, activities in grey are not discussed in this chapter.

Manual calibration rationale

In Office 17, the objective was to achieve a calibrated model solely through manual adjustment of the input parameters. For this case study, changes made to the model can sometimes seem arbitrary, especially when not enough data is available to justify making a change to the model. For example, if a large under prediction of power energy use is identified, there are then several options for changing the model to align to actual situation, as power energy use is dependent on a range of input parameters. There is then no premise for changing one parameter over the other when detailed measurements are not available, even though they will affect the model in different ways. Changing equipment power density in one space with space conditioning opposed to one without will have different effects on heating and cooling loads, whilst achieving the purpose of aligning power energy use. Under this rationale, it becomes clear that a higher level of data granularity can support in developing a more accurate model, but that a lack of information can cause these parameter changes mask the real situation. Choices made in changing input parameters are not extensively described, but will be explained when significant changes were necessary or when certain limitations were identified that could drastically affect the accuracy of the model.

6.2 Office 17

6.2.1 Predicted vs. measured performance

An initial comparison of the predicted and measured monthly energy use of the case study building is shown in **Figure 6.2.** A direct distinction is visible between energy use types in both datasets. Predicted energy use was calculated by EnergyPlus and is broken down in different types of energy use. In this case study there are 8 different types of measured energy use. Predictions and measurements have the same dominant energy use types (gas, equipment and lights). In Office 17, the server and lifts were not taken into account in the model and the print room and canteen electricity energy uses were seen as an extension to the equipment and lighting loads.

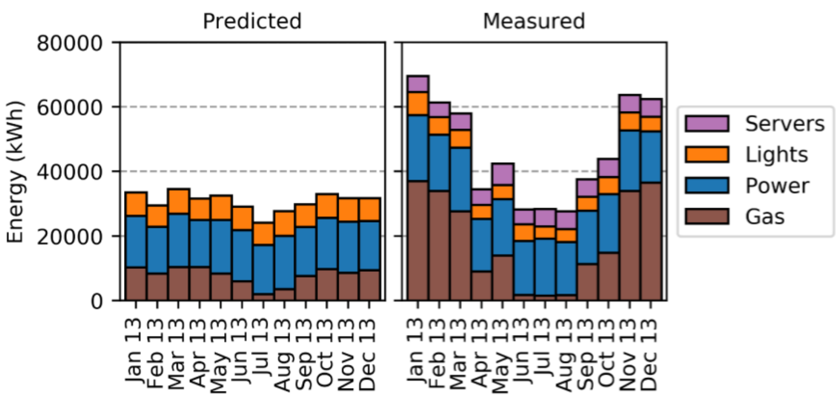


Figure 6.2: Base case predicted and measured monthly energy use (kWh) for the case study building.

In **Figure 6.3** the predicted and measured energy use were aggregated to total energy use for a year. It shows that the base case model underestimates measured energy use by 186.000 kWh (32). For the dominant energy uses, actual gas use is 57% higher, equipment electricity use 11% higher and lighting electricity use is 45% lower than predicted. Other energy use types account for 7% and 16% respectively of the total predicted and actual energy use. Measured server energy use accounts for 65% of the remaining 16% in energy use. Server energy use is not accounted for in predicting regulatory energy performance of buildings and can be a major cause for a discrepancy between regulatory predictions and measurements. The total energy consumption for electricity and gas differ by 2.9% and 2.3% CV(RMSE) respectively.

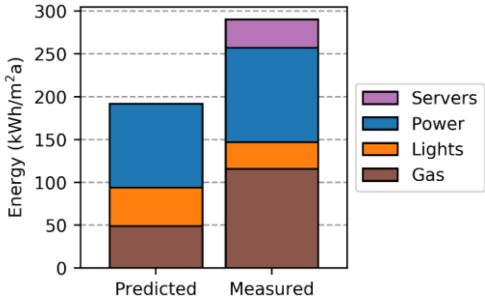


Figure 6.3: Base case predicted and actual energy use.

The former two figures show that there are large differences between predicted and measured energy use. However, they only highlight this on an aggregated scale. As the results are based on half-hourly data, a more detailed picture can be given by showing a typical weekday for the predicted energy use of the base case model against measured energy use, shown in **Figure 6.4.**

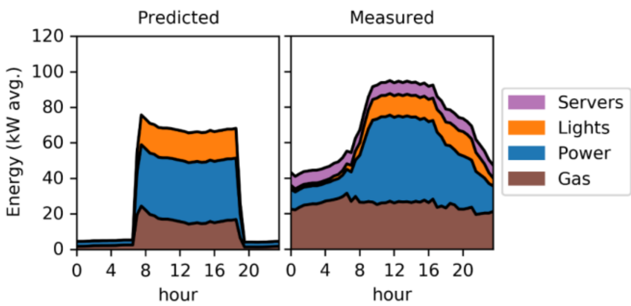


Figure 6.4: Predicted (base case) and measured energy use for a typical weekday.

The building simulation model predicts energy use according to a daily pattern with a small amount of variation, where it is assumed that the building is occupied between 7am and 7pm. The arrival and departure of occupants in the building is different, it people come in between 8am and 10am, leaving between 6pm and 9pm. Lights and equipment show a much smoother profile due to the changing occupancy density in the actual building. However, this does not explain night-time equipment and gas energy use, especially a constant gas use during the whole day is a major factor that is not accounted for in model predictions. In the model gas energy use differs during the week and weekend, whereas measured gas energy use does not demonstrate this pattern. This indicated that heating was left on during the weekends. Whereas, gas use during the summer months was minimal, because the boiler was turned off. Several major differences between predicted and measured energy use were identified that influence and contribute to the total discrepancy.

6.2.2 Sensitivity analysis

In total, 1100 parametric simulations with randomised inputs were carried out using EnergyPlus, which was determined to be enough (as explained in the methodology) to accurately calculate sensitivity indices. Sensitivity analysis was used to determine correlation coefficients and rank the influence of 30 input parameters on the output. Regression analysis is applied using Pearson and Spearman correlation coefficients to show the relationship between input variables on the output. Spearman's rank correlation coefficient is found to be similar to Pearson's correlation coefficient. Similarity indicates linearity in the data, therefore only Spearman Rank correlation coefficients in relation to electricity and gas use is analysed, shown in **Figure 6.5**.

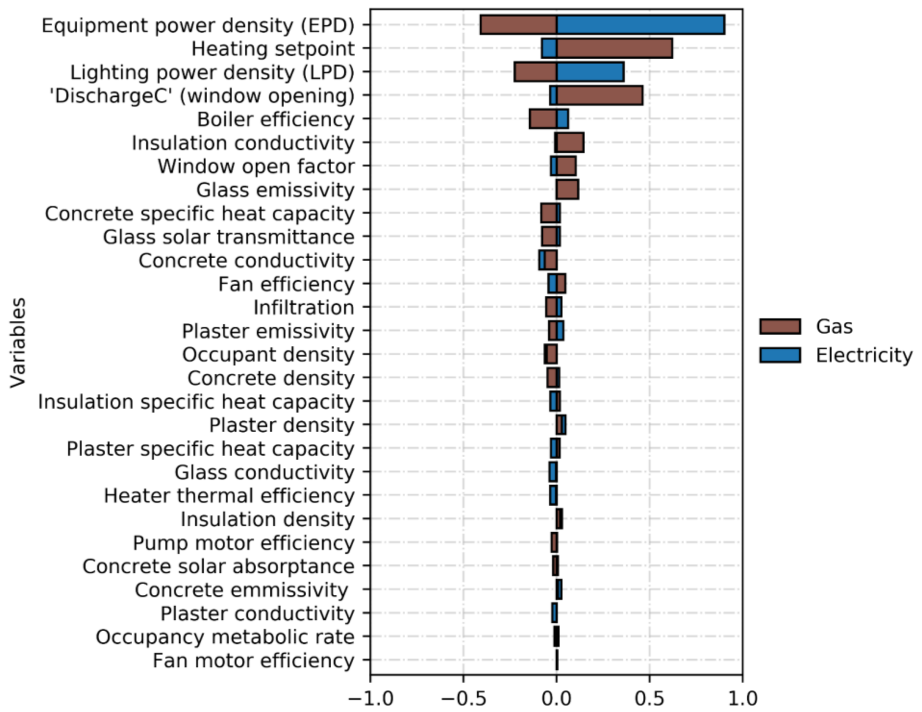


Figure 6.5: Spearman's rank correlation for varied building parameters.

Several parameters are shown to have significant correlations. For electricity in particular, equipment power density (EPD) has a very strong positive correlation (ρ 0.90), whereas lighting power density (LPD) has strong positive correlation (ρ 0.36). Gas use is primarily driven by the heating setpoint, with a very strong positive correlation (ρ 0.62), and opening of windows (discharge coefficient) with a strong positive correlation (ρ 0.46). The discharge coefficient specifies the airflow effectiveness through windows and doors and is related to the operation of windows by occupants. Correlation coefficients have also been calculated on a monthly basis for all energy use types, see further in Appendix B. Sensitivity analysis also indicates negative correlations between equipment power density (ρ -0.41), lighting power density (ρ -0.23) and boiler efficiency (ρ -0.14) on gas use. A negative correlation means a decrease of gas use with an increase of the input variable. The boiler efficiency and discharge coefficient correlate less with gas during the summer months, due to smaller heating loads. However, the significant correlations for electricity are stable during the seasons.

Regression analysis identified significant statistical correlations among inputs and outputs, these are necessary to assist in an effective calibration of the building energy model to actual building energy use. Sensitivity analysis shows that calibrating the model should mainly focus on adjustment of the equipment and lighting power density, heating temperature setpoint and discharge coefficient for windows. These parameters have a large influence on the dominant energy end-uses in the building and should only require relatively small changes.

6.2.3 Manual calibration

Sensitivity analysis indicates that the dominant energy use types (gas, equipment- and lighting electricity) primarily influenced by the heating setpoint, opening of windows, equipment- and lighting power density. The monthly and hourly mean bias error and CV(RMSE) signify the discrepancy between predicted and measured gas and electricity energy use, as shown **Figure 6.6** and **Figure 6.7** respectively.

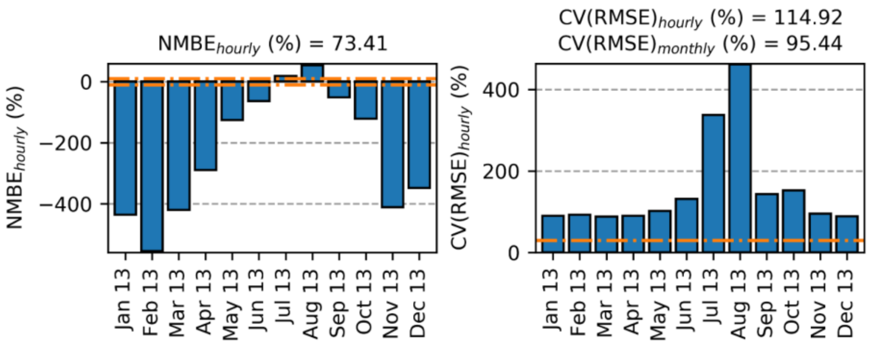


Figure 6.6: Differences between predicted and measured **gas energy use** represented by the NMBE and CV(RMSE) for both monthly hourly values and total month and hour over the year.

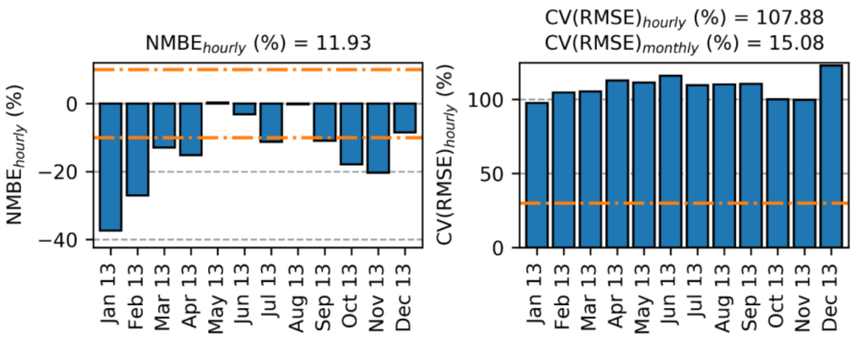


Figure 6.7: Differences between predicted and measured **electricity use** represented by the NMBE and CV(RMSE) for both monthly hourly values and total month and hour over the year.

The statistical measures indicate a significant difference between predicted and measured energy use. As can be seen from previous figures, absolute differences between the months are considerable, but at an hourly interval they are even more significant. The orange lines indicate the ASHRAE guidelines for deeming a model calibrated, the base case model does not come close to satisfying these requirements and calibration is therefore necessary.

Due to a large difference in gas use, the heating temperature setpoint has been increased to 23°C, and discharge coefficient for windows is increased simultaneously. The calibration procedure involved changing input parameters over several iterations. Adjusted inputs are shown in **Table 6.1.**

Table 6.1: Adjusted input parameters for the calibrated model

Input variable	Base case	Calibrated
Equipment power density (W/m^2)	30	25.5
Lighting power density (W/m^2)	15	7
Discharge coefficient (opening of windows)	0.68	0.77
Heating temperature (°C)	22	23 (Winter) and 18 (Summer)
Heating temperature unoccupied (°C)	15	23 (Winter) and 18 (Summer)

The typical weekday and weekend day energy use show a large discrepancy during unoccupied hours. This was reflected by changing schedules for equipment, lighting, occupancy and heating. Schedules for system operation and occupancy have a large influence on the energy consumption of a building. Schedules for equipment, lighting and occupancy were widened as to represent the actual building use profiles. Furthermore, to account for heating during unoccupied hours, some heating was allowed for in the model. Finally, it was chosen to calibrate the building simulation model to the measurements excluding the server energy use as it was already identified as a major source of discrepancy.

The calibrated model is adjusted to the aforementioned changes, predicted and measured energy use is then compared a second time. In **Figure 6.8,** predicted annual energy use for the calibrated model is compared to measurements. Server and canteen energy use have not been taken into account in the simulation model, resulting in the remaining gap between predicted and actual energy use. The total energy consumption for electricity and gas between the calibrated model and measurement now differ by 2% and 0.02%, down from 11.2% and 4% CV(RMSE) respectively.

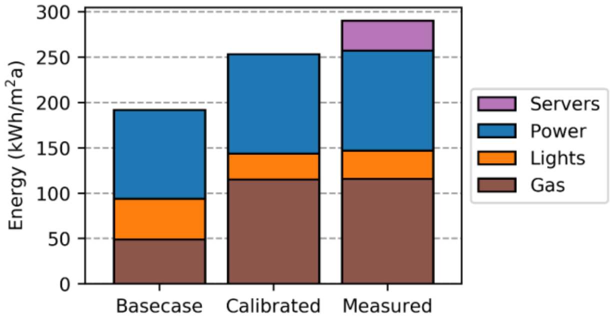


Figure 6.8: Calibrated building model predicted energy use and measured energy use.

If the average day profile in **Figure 6.9** is juxtaposed with the base case profile shown in the initial comparison in **Figure 6.4**, it becomes clear that a large amount of energy use was not accounted for in the initial model due to night-time energy use. This is clear when focusing on the dominant energy uses as shown in **Figure 6.9**. Such an assumption has a large influence on predicted energy use and in this case leads to an underestimation of actual energy use.

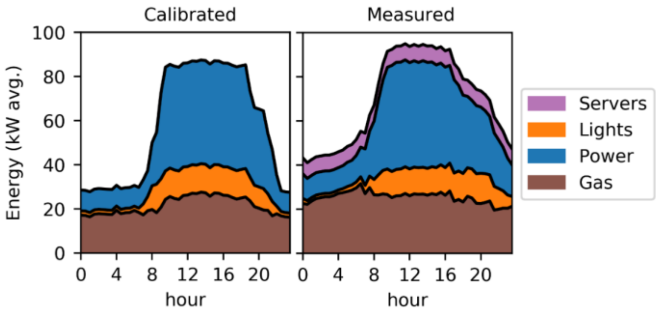


Figure 6.9: Typical weekday energy use predicted by the calibrated model compared with measured energy use.

Several parameters were changed to calibrate the model and account for the variation in predicted and measured energy use. However, the initial building simulation model is a bespoke model already set up to represent the actual building in terms of its occupancy schedule and initial best guess input parameters to closely predict its actual energy use. Thus, expecting a minimal discrepancy in results. From analysis of both the initial comparison and calibration of the model, many underlying causes for the discrepancy have been identified.

6.2.4 Impact of assumptions

Several adjustments were introduced to the calibrated model, they are related to equipment power density, the heating setpoint temperature, equipment, lighting, heating and occupancy schedules. Assumptions were simulated per simplification and in combination as shown in **Table 6.2**, simplifications are numbered, which relate to later explanation of the results.

Table 6.2: Effect of simplifications on the calibrated model as a percentage difference per yearly energy end-use.

	Simplification	Gas	Equipment	Lights	Fans	Cooling	Pumps	Total
S1	Equipment power density from 25.5 to 11.75 W/m²	29	54	0	50	50	69	37
S2	Changed equipment schedule	2	18	0	6	6	22	9
S3	Changed occupant schedule	4	4	0	7	7	8	4
S4	Changed lighting schedule	6	4	31	11	11	18	8
S5	Heating setpoint from 23 to 22 °C	19	4	0	8	8	23	10
S6	Heating setpoint unoccupied from 22 to 12 °C	67	4	0	14	14	70	32
S7	Simplification 6 and summer heating SP from 18 to 22 °C	60	4	0	14	14	62	29
S8	Combination of all	60	62	26	68	67	65	57

Results in **Table 6.2** are visualised in **Figure 6.10,** the simplifications are shown on the x-axis with on the left-hand side the calibrated model.

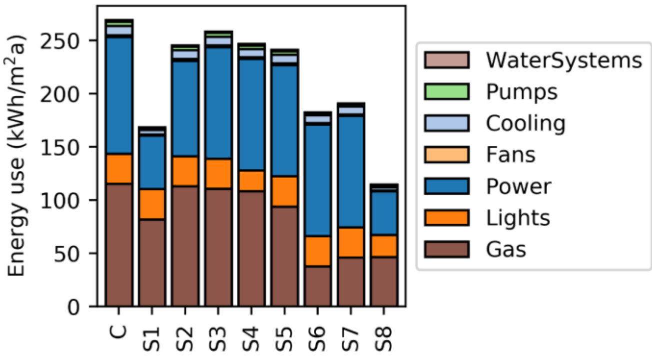


Figure 6.10: Energy use for the calibrated model with simplifications as numbered in **Table 6.2.**

Simplification 1 reduces equipment power density from 25.5 W/m² to 11.75 W/m², resulting in a significant reduction of equipment, gas and cooling energy use. In total, this reduces predicted energy use by 37% from the calibrated model. **Simplifications 2, 3 and 4** introduce changes to the equipment, occupancy and lighting schedules respectively, these changes are shown in **Figure 6.11**. Schedule related assumptions have a smaller influence on the energy use, changes result in a reduction on total predicted energy use of -9%, -4% and -8.3% for equipment, occupancy and lighting schedules respectively.

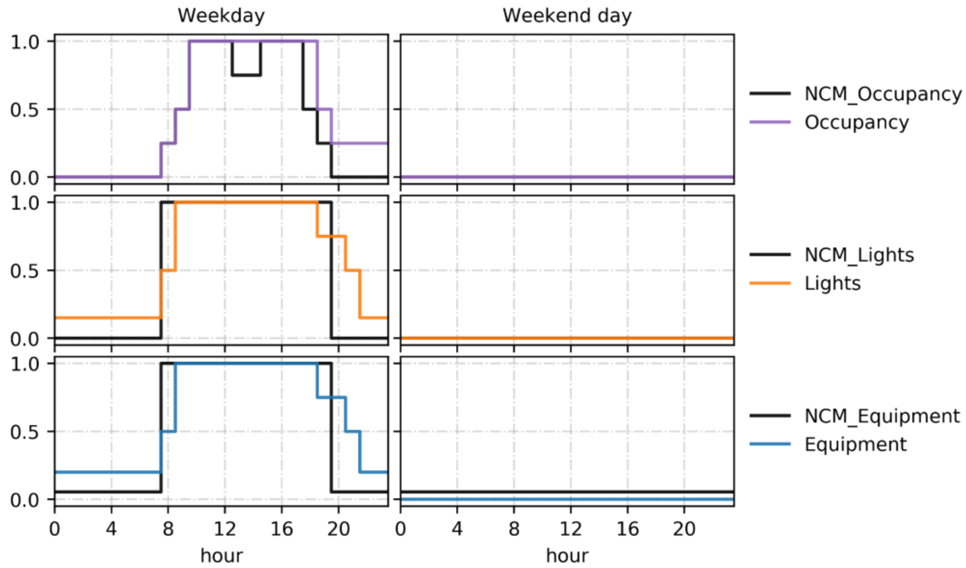


Figure 6.11: Calibrated schedules for occupancy, lighting and equipment compared to NCM schedules used for the simplifications.

Simplification 5 and **6** involve a change in the heating temperature setpoint. For simplification 5, the heating setpoint is changed from 23°C to 22°C, this then leads to a 18% reduction in gas use. For Simplification 6 the heating temperature setpoint during unoccupied hours is changed from 22°C to 12°C to prevent night-time heating, this leads to a 67% reduction in gas energy use and 32% total energy use. The temperature setpoint is set to 22°C, which in combination with **simplification 6**, results in a 60% reduction in gas energy use. Finally, **simplification 8** is a combination of aforementioned individual assumptions and leads to a 57% reduction of predicted energy use by the calibrated model, reducing gas use by 60%, equipment by 62% and lighting by 26%.

The model simplifications introduced to the calibrated model are typical assumptions that have a significant influence on the predicted energy use of a building. The predicted model based on initial assumptions underestimates the measured energy use by 23%, the building simulation model based on a combination of design stage assumptions underestimates calibrated energy use by 57%. Notably, this underestimation still includes equipment energy use, which for compliance modelling would be excluded.

6.3 Office 71

6.3.1 Predicted vs. measured

Office 71 has a high level of disaggregation in the sub-metering system, which made it possible to compare predicted and measured energy use at a per floor basis. The level of disaggregation at which predicted and measured energy use was compared is shown in **Table 6.3.** Both lighting and power electricity use were separated per floor, where power electricity use includes small power equipment, FCU fans and electric water heaters.

Table 6.3: Energy end-uses as defined for comparison.

End-uses	Model disaggregation
Systems	Pumps, Toilet exhaust, Lifts
Lighting	Lighting per floor
Power	Power per floor, DHW (WaterSystems) per floor, FCU fans, canteen
Gas	Gas
Excluded	Electrical heating, Cooling, AHU fans

The main systems (AHU and VRF) were not measured by the sub-metering system. The AHU was determined to be off as identified during the building audit. Systems energy use excludes the electrical heating and cooling provided by the VRF system and fans in the air-handling unit. These components are however included in the model based on system design specifications, which decrease their uncertainty. However, this is not validated and is therefore a limitation as the accuracy in predicting system energy use cannot be determined. System energy use was predicted to present 13% of the total energy use.

An initial model was set up based on available data from building design specifications and the building audits. Several adjustments were made to the initial model to resemble measured energy use more closely:

Model adjustments

- The boiler turns off during the summer months of (May to September), as such, the boiler is turned off in the model to replicate this behaviour.
- Created typical profiles for lighting and power on each floor based on measured electricity use.
- Determined baseload electricity use for lighting and power at 15% and 20% respectively.
- Set holidays in the model based on measured daily energy use.

Limitations identified

- Excluded system energy from comparison to measured energy use as the metering system is not measuring electricity use from the VRF systems and air-handling unit, however the system is included in the model based on design specifications.
- Electrical heating is available through zip taps and in showers and cannot be distinguished from measured power energy use, but is likely to be a large contributor.
- Material properties were determined during the walkthrough and based on previous assessment, but no design specifications were available.

Total monthly predicted and measured electricity use for Office 71 is shown in **Figure 6.12** for 2014, the reference year used for comparison and calibration. Compared to the two previous case study buildings, Office 71 is a relatively small building which reduced simulation time significantly and allowed quick iterations of the model. Slight differences exist between the predictions made by the base case and measured monthly energy use. In contrast to the other buildings, electricity use is constant throughout the months with peak energy use occurring during the month of November 2014, interestingly predicted monthly electricity use is significantly lower during this month and a large discrepancy occurs. The large differences were found to be due to November having 10 weekend days, where some of the weekends saw much higher occupancy than during other weekends of the year.

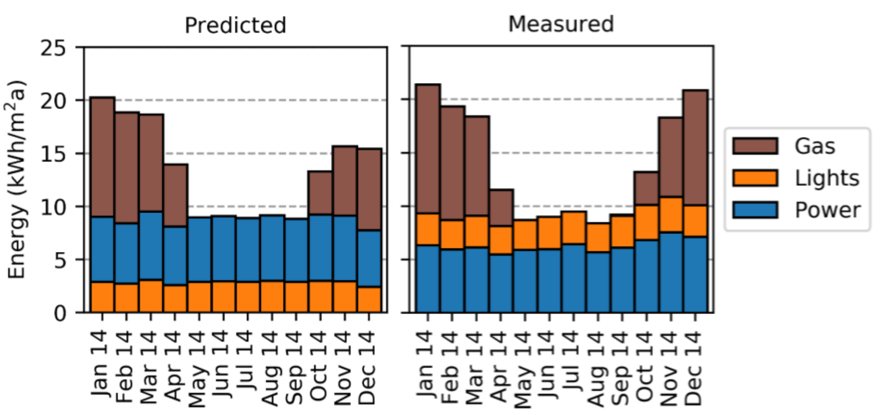


Figure 6.12: Monthly predicted (base case) and measured energy use for Office 71 in 2014.

System energy use was predicted, but not measured. Monthly energy use for heating and cooling supplied by these systems and their auxiliary energy use is shown in **Figure 6.13.** Cooling and fan energy use from the air handlers represent about 30-50% of total system energy use during the summer, whereas it is a significantly smaller proportion during the winter months (~10%). Detailed information of these systems was available from O&M manuals and site inspection, used as input in the building simulation model.

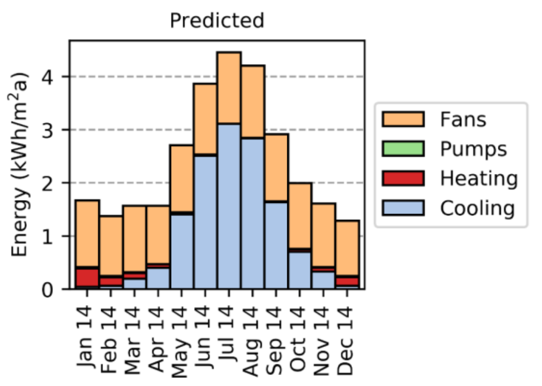


Figure 6.13: Predicted monthly energy use from the heating and cooling systems for Office 71.

The monthly comparison of energy use shows that the base case model predicts monthly energy use accurately. The exact error on a monthly and hourly basis is further explicated by the statistical measures; NMBE and CV(RMSE), shown in **Figure 6.14.** The hourly NMBE for April and November show significant differences due to the over and under prediction in gas energy use for these months respectively. The total hourly NMBE is however considerably low at -1.71%, in contrast the CV(RMSE) hourly values are somewhat higher than the NMBE hourly values, in many cases the 30% threshold represented by the orange line is exceeded. The actual criteria set by ASHRAE however, is 30% for the whole year, which in this case is not met. This clarifies that large differences on an hourly level need to be mitigated and are masked by the monthly deviations of energy use, because the monthly values are a good indication of a calibrated model.

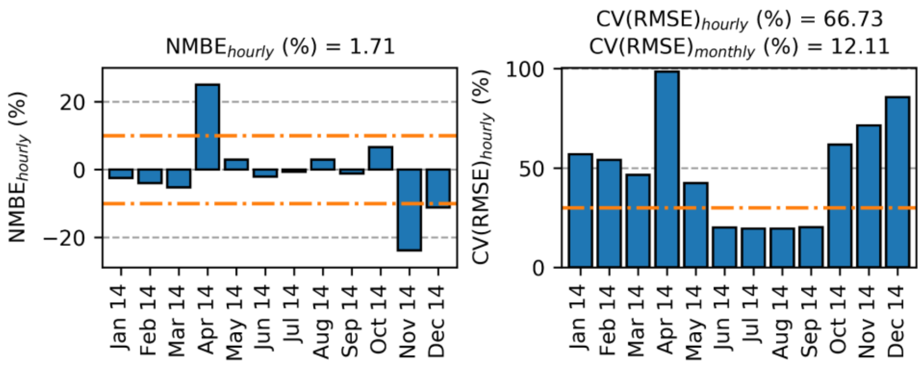


Figure 6.14: Statistical measures denoting difference between predicted and measured energy use, given by the average monthly NMBE and CV(RMSE).

To understand model predictions, some additional comparisons between predictions and measurements are necessary. As such, a typical week and weekend day can be helpful, this is shown for total electricity use in **Figure 6.15**. High values for CV(RMSE) can be explained by the large hourly variation in measured electricity use, as indicated by the shaded region, which signifies one standard deviation from the mean. In contrast, the predicted standard deviation is negligible, i.e. predictions are similar every day as there are no factors influencing hourly use.

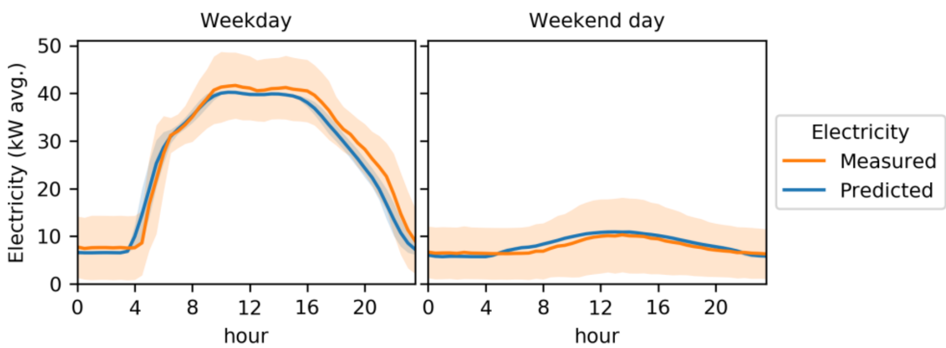


Figure 6.15: Predicted and measured electricity use for a typical week and weekend day, shaded region represents the standard deviation around the average.

Breaking down the typical profile in actual hourly values clarifies that there is a large deviation during the weekends and peaks during the weekdays. This is explicit when half-hourly predicted and measured electricity use are compared, as shown for the first three months of 2014 in **Figure 6.16.** The building is not always occupied during the weekend and when it is, it is higher than predicted as the predicted occupancy, lighting and equipment profiles are based on the mean electricity profile throughout the year.

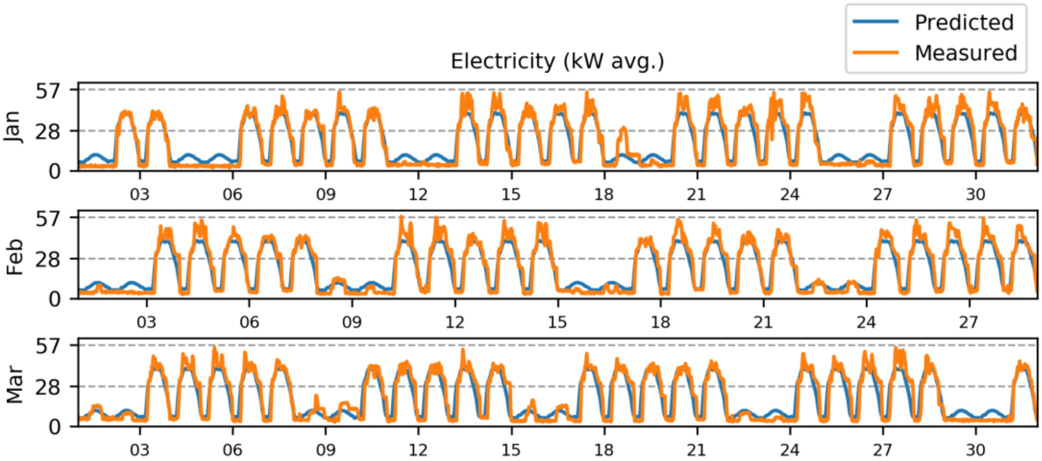


Figure 6.16: Predicted and measured electricity use for the first three months of 2014.

Occupancy during the weekends is very intermittent and will be difficult to replicate, it does however highlight the limitations of using typical weekday and weekend profiles to represent use throughout the year. Although schedules could potentially be created for the whole year and as such exactly replicate the measured behaviour, the model predictions would start overfitting the measured data. It would be more realistic to use average profiles where uncertainty is introduced to reflect the actual situation.

Lighting and power were disaggregated in the model to compare with measured data for each floor. A higher level of granularity can potentially increase the accuracy of a calibrated model as behaviour on different floors are captured separately. This was the case for the 1st, 2nd and 3rd floors, which are used as office space, their profiles of use differed significantly and disaggregation allowed for replicating the actual behaviour separately. Ground floor power and third floor lighting energy use for a typical weekday and weekend day are shown in **Figure 6.17** and **Figure 6.18** respectively.

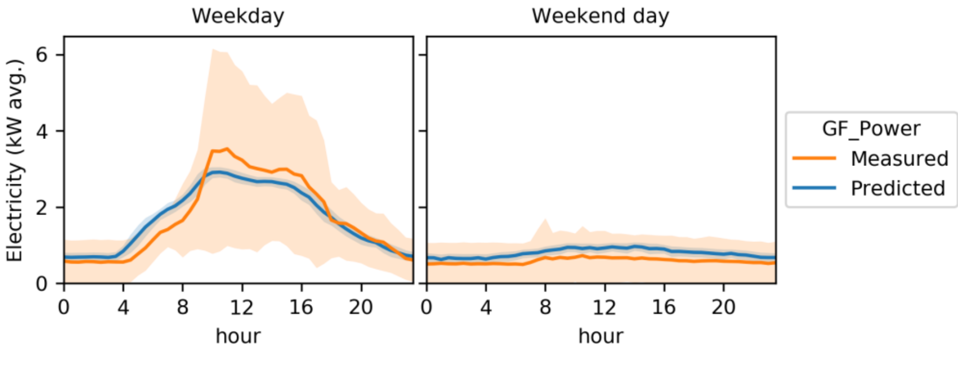


Figure 6.17: Predicted and measured ground floor power energy use for a typical week and weekend day, shaded region represents the standard deviation around the average.

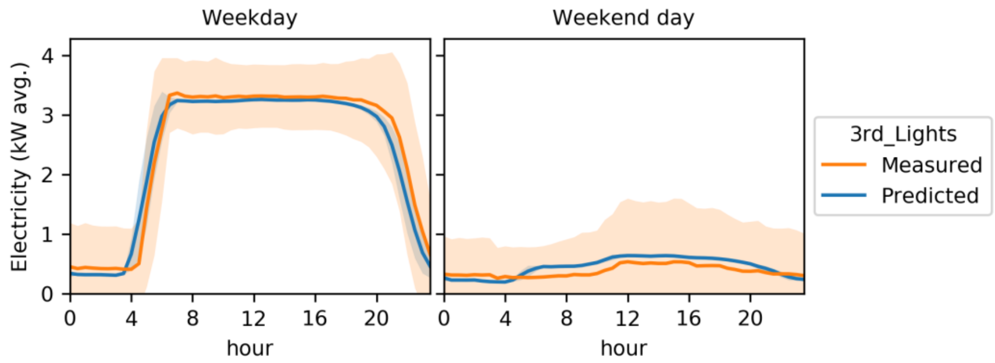


Figure 6.18: Predicted and measured third floor lighting energy use for a typical week and weekend day, shaded region represents the standard deviation around the average.

Total monthly predicted and measured gas energy use is shown in **Figure 6.19**. Differences exist between the months, during the winter months, actual energy use is higher, whereas during the summer months, predicted energy use is higher. This balances out over the year, making the total yearly consumption nearly identical. The summer months show no measured gas use, and it was identified that radiator heating is turned off during these months, this was therefore replicated by the model to represent the actual situation.

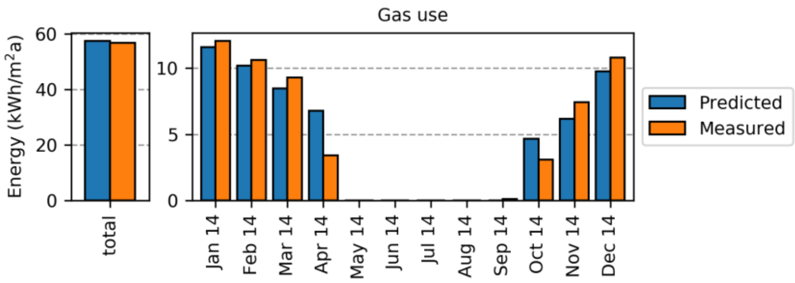


Figure 6.19: Total monthly predicted and measured gas use for the base case model.

6.3.2 Uncertainty analysis

The base case model is run 1500 times while varying input parameters, generating a solution space that predicts total energy use within 140 to 170 kWh/m^{2}a, as shown in **Figure 6.20**, red dots represent individual simulations and the orange dot indicates measured energy use. Total measured energy use and disaggregated energy end-uses (lighting, power and gas) fall within the solution space. A higher level of spatial granularity was achieved by disaggregating lighting and power per floor, as shown in **Figure 6.21.**

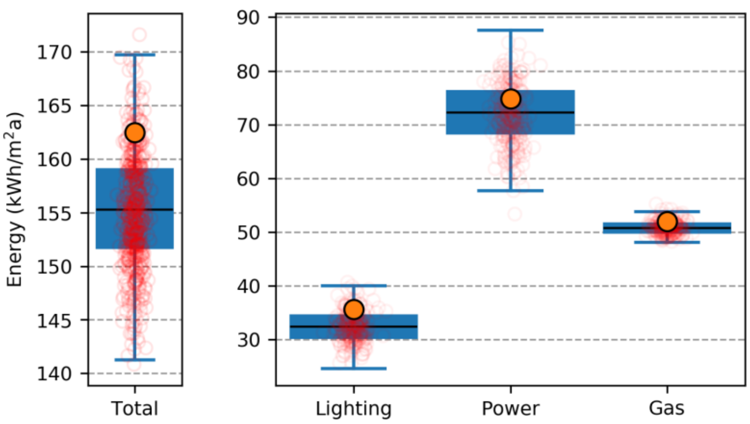


Figure 6.20: Total predicted (boxplots) and measured (orange dots) energy use for 1500 simulations.

The base case model replicates typical use by incorporating schedules based on the energy use profiles per floor, this resulted in the predicted lighting and power to be very similar on an hourly level and measurements fall within the predicted solution space.

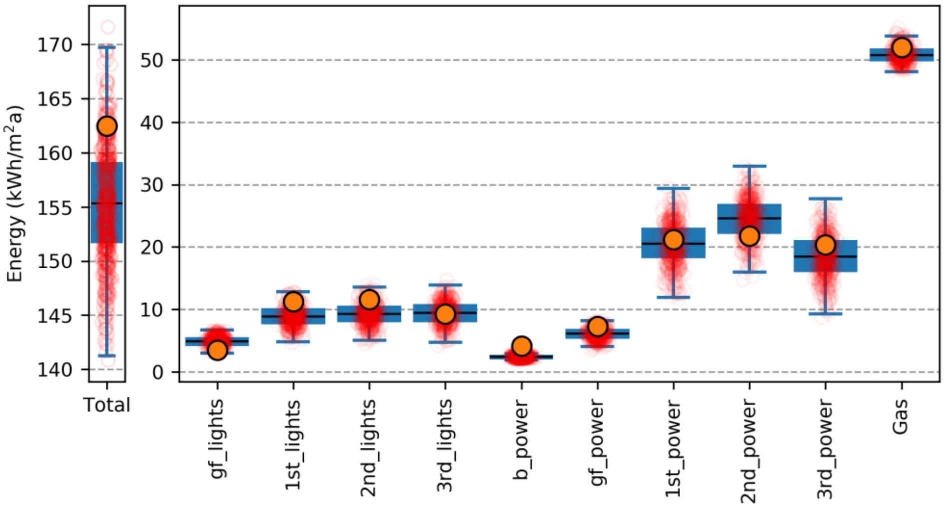


Figure 6.21: Total predicted (boxplots) and measured (orange dots) energy use for 1500 simulations.

Finally, a monthly temporal granularity is compared for the end-uses, as shown in **Figure 6.22**. There are significant variations within the months for both lighting and power, predicted energy use follows a similar trend for some of the months, but measured energy use falls outside of the prediction range for many of the months. Initially, gas energy use seems to show a surprisingly small range in monthly and total predictions, indicating that the changing variables have a small impact on gas energy use. However, as the monthly variation in gas energy use is larger than lights and power, the variation is actually similar, but looks more condensed. Nevertheless, other input parameters could be included to increase this variation on a monthly basis to allow automated calibration to converge towards possible solutions.

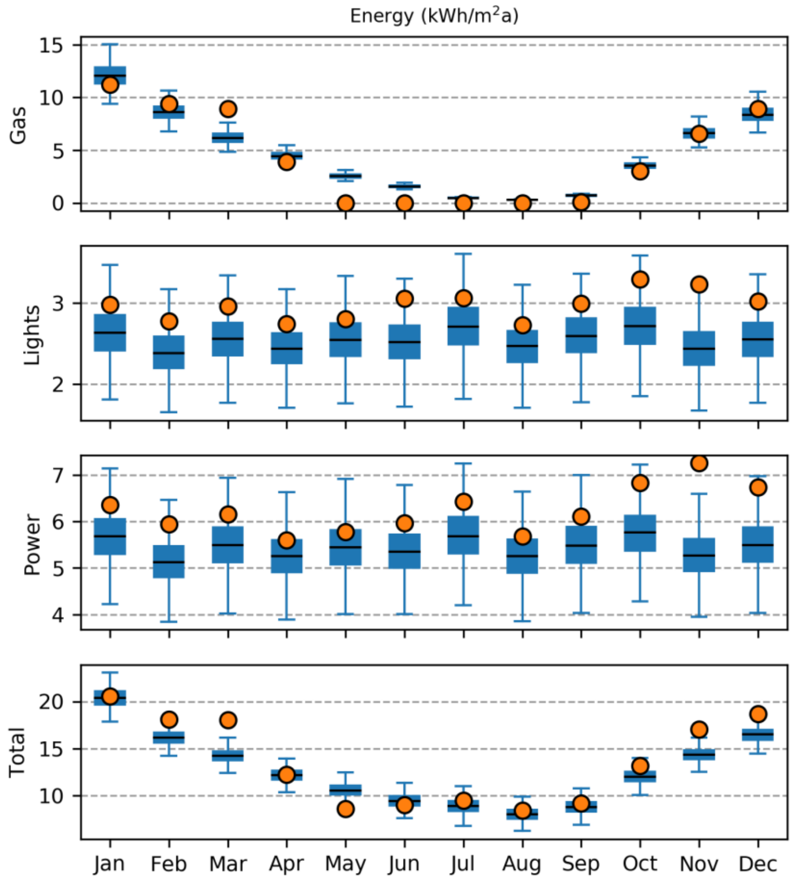


Figure 6.22: Monthly predicted (boxplots) and measured (dots) energy use for 15000 simulations.

Analysing the distribution of predicted energy use from numerous simulations is used to assess if the measured values fall within the solution space, it does however not give a good understanding of the relation between inputs and outputs. Instead, relationships between input parameters and outputs is presented using sensitivity analysis.

The previous figures are configured for specifically the energy end-uses that were comparable to those measured. Although a higher level of granularity could not be obtained, an energy model is able to disaggregate all components of energy use and as such, more analysis on the different energy end-uses can be performed using the model. End-uses were disaggregated as shown in **Figure 6.23**, which presents the uncertainty in annual energy use for Office 71, represented by the standard deviation in total energy use (orange) and percentage of the mean (blue) or coefficient of variation x 100%.

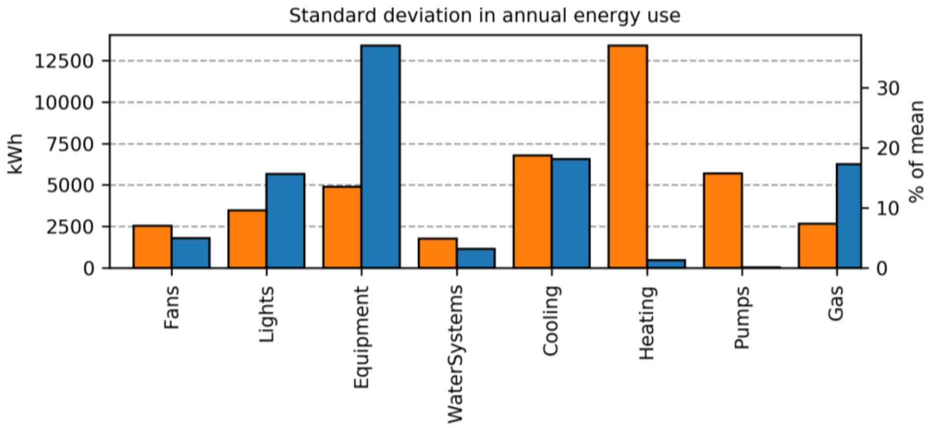


Figure 6.23: Uncertainty in annual energy use represented by the standard deviation of the total (blue - kWh) and coefficient of variation (orange - % of mean) for Office 71.

In Office 71, equipment energy use is one of the main contributors to total energy use, typical for an office building. Whereas, cooling is closely followed by lighting and gas energy use. The coefficient of variation allows comparing the uncertainty of the different end-uses, indicating that in particular electrical heating and pump energy use are relatively uncertain. In contrast, gas, fan and L&P energy use from electrical water heating show a small coefficient of variation, the simulation are predicting similar values.

6.3.3 Sensitivity analysis

Sensitivity of parameters on energy use was determined by calculating the Spearman rank correlation coefficients for each input parameter and energy end-use predicted. **Figure 6.24** shows the Spearman rank correlation coefficients of input parameters on the eight energy end-uses on the left, input parameters with a coefficient is greater than $\rho = 0.25$. There is a distinct difference between input variables and their effect on different types of energy use. The most significant positive correlation ($\rho = 0.97$) is between the DesignOutdoorAir flowrate, and fan energy use, which is the amount of mechanically ventilated air provided to the office spaces, distributed within ($\mu = 8$, $\sigma = 1.2$) litres per second. Implying that a variation in fan energy use is mainly driven by the amount of ventilation provided to the office spaces, while an increase in equipment power density (in the offices) and the heat build-up in spaces has a slight influence on fan energy use. In contrast the strongest negative correlation is between hot water temperature of the boiler ($\mu = 82$, $\sigma = 4$) and pump energy use. Although both highlight strong effects on their respective energy end-uses, their proportionate effect on total energy use is considerably smaller. However, when taking into account the percentage of energy end-use on total energy, it becomes clear that some of the significant coefficients become less important. Most important variables are those related to end-uses with a high amount of energy use, such as lights, equipment and gas, and their dependent parameters; power density of lighting and power in the offices, boiler efficiencies and seasonal variation factors. In Office 71, variations in occupancy, lighting and equipment profiles were not considered, which would have a large impact on the variance in energy use.

Coefficients near zero imply that there is no linear correlation between the inputs and outputs. Although filtered out in the figures, parameters such as the conductivity of materials, number of people in zones, infiltration rate and heat pump COPs are not significant and therefore have little influence on the energy use in Office 71.

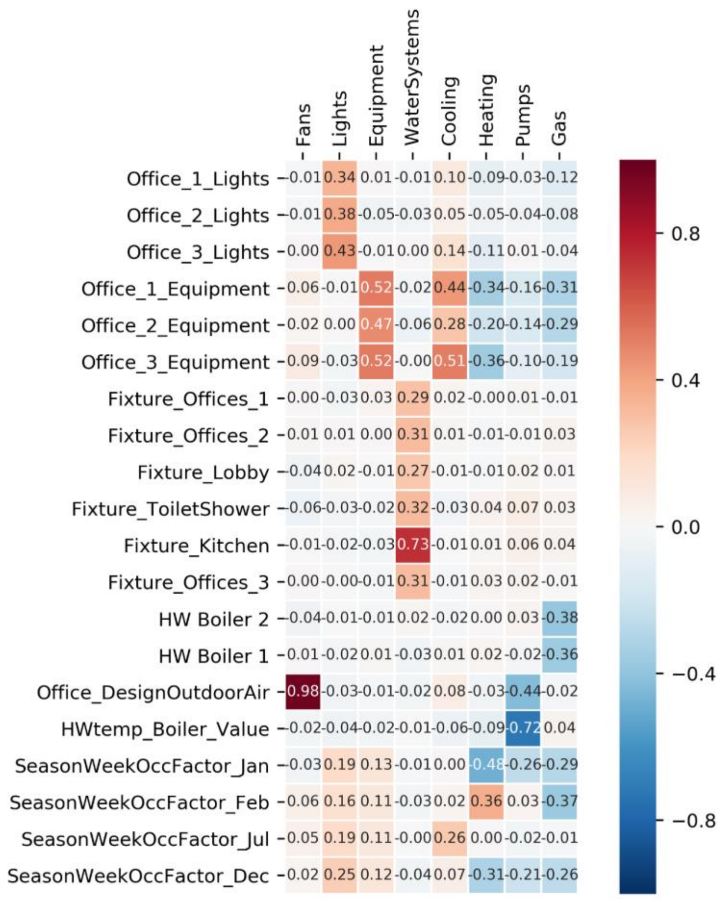


Figure 6.24: Spearman correlation coefficients per energy end-use for Office 71.

The coefficients here are calculated for the annual energy end uses; potentially they could be calculated for each month or even each day. Per month, the correlation coefficients for input parameters that have seasonal dependency will change. For example, during the winter, variation in boiler efficiencies (and other variables that affect gas energy use) will have more significant correlations than in summer, the calculated correlation for the year therefore lies within those that would be calculated for the winter and summer months. Similarly, this would be true for cooling, the larger the seasonal variation in energy use, the larger the difference of the coefficients during the seasons when input parameter values are the same during the year.

The Spearman correlation coefficient describes the relationship between two variables, but is unable to explain if any underlying output variance is caused by the interaction between multiple inputs. Sobol' first-order, second-order and total-order indices were therefore calculated, where total-order indices measure the contribution of both first and higher-order interactions. They are compared for several variables to both Spearman rank and Pearson correlation coefficients as shown in **Figure 6.25**.

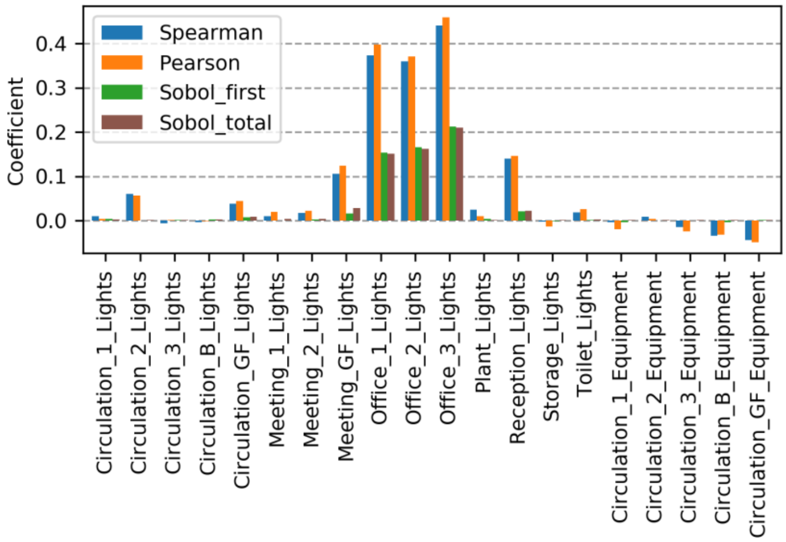


Figure 6.25: Comparison of Spearman correlation coefficients and Sobol' first and total indices, for Office 71.

Spearman rank and Pearson correlation coefficients have similar values, showing that the linear correlation is larger than the rank correlation, indicating that the influential observations in the ends of the distribution have a larger influence relative to their ranked values. Sobol' first-order and total-order indices are very similar, indicating strong linearity between inputs and outputs. Sobol' indices can therefore be disregarded as no further interactions need to be explained between the inputs (Iooss & Lemaitre, 2015).

6.3.4 Impact of assumptions

The impact of typical (NCM) assumptions on the base case model have been assessed, the applied simplifications and their impact in percentage difference per energy end-use are given in **Table 6.4.** In contrast to Office 17, the manually calibrated model is referred to as the base case for Office 71, CH and MPEB, as the automated calibration procedure was applied for these building, resulting in a final calibrated model. Input parameters that were previously identified as significant through sensitivity analysis are implemented as NCM assumptions, where other less significant factors have not been further investigated.

Table 6.4: Effect of simplifications on the calibrated model as a percentage difference per yearly energy end-use.

	Simplification	Cooling	Equipment	Fans	Gas	Heating	Lights	Pumps	WaterSystems	Total
S0	Equipment power for space types based on NCM	-20	-7	-1	1	30	0	2	-1	-6
S1	NCM schedule for equipment, lighting and occupancy	-26	-17	-2	14	24	-34	11	-11	-20
S2	NCM schedule for heating and cooling	-1	0	0	-4	1	0	-4	0	0
S3	Infiltration (12 to 8 m³/m²h @ 50Pa)	0	0	0	-1	-1	0	0	0	0
S4	Combination of S0 and S1	-41	-19	-3	16	54	-34	14	-12	-23
S5	Combination of S0 to S3	-41	-19	-3	11	54	-34	9	-12	-23

The model simplifications and their effect on total energy use per floor area are shown in **Figure 6.26. Simplification 0** implements the NCM assumptions for equipment power density, found to be very similar in Office 71 to NCM assumptions, although slightly underestimated for the Offices. Determined to be in the range of 14-18 W/m² for the different floors, compared to the NCM assumption of 11.77 W/m², reducing total energy use by 6%.

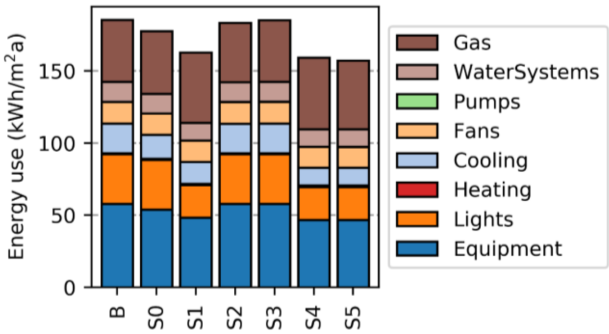


Figure 6.26: Energy use for the calibrated model with simplifications as numbered in **Table 6.4.**

Simplification 1 applies the NCM schedules for equipment, lighting and occupancy to the base case model, the base case model schedules for Office 71 were determined through the use of sub-metered electricity use per floor, a comparison between the schedules for the first floor and NCM simplifications are shown in **Figure 6.27**. Changes in the schedules reduced total energy use by 20%, directly affecting equipment and lighting energy use and indirectly system energy use due to lower internal gains.

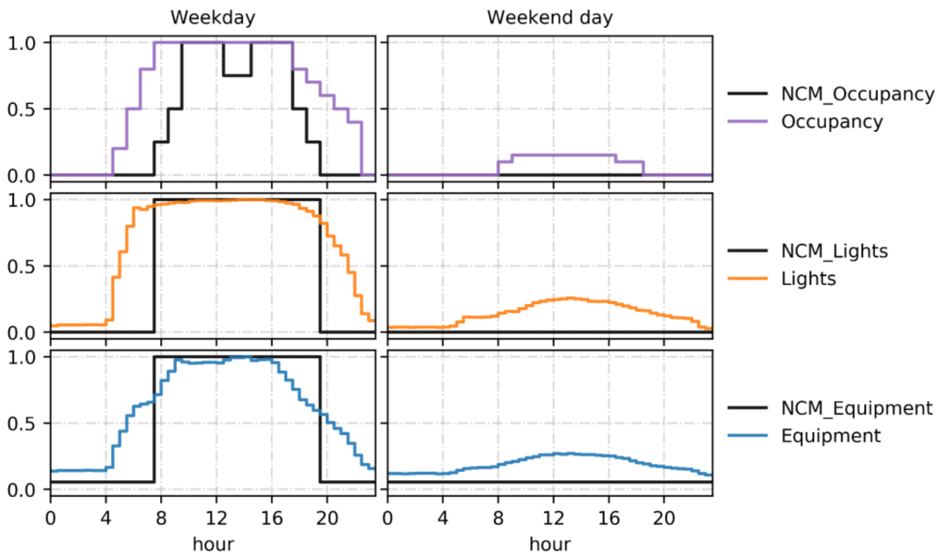


Figure 6.27: Calibrated schedules for occupancy, lighting (1st floor) and equipment (1st floor) compared to NCM schedules used for the simplifications.

Initial heating and cooling temperatures in the base case model were assumed to be very similar to the NCM assumptions, and as such have not significant impact on energy use. Infiltration, as part of **simplification 3** was reduced from 12 to 8 m³/m²h @ 50Pa, this had a negligible effect on energy use, expected, as sensitivity analysis indicated it to be an insignificant parameter. Energy use in Office 71 is mainly dependent on the schedules of use and power density assumptions for equipment and lighting, typical for office buildings.

6.4 CH

6.4.1 Predicted vs. measured

Energy use measured in CH includes the VRF systems, where two separate meters measure electricity use for all roof condensers, gas use for radiator heating in circulation spaces and lighting and power. Energy use for the lifts and a server is measured separately in the building, both were included in lighting and power for comparison. Although lighting and power was available for all floors, they could not be accurately separated and are therefore considered as a total. **Table 6.5** shows the energy end-uses used for comparison. Heating, cooling and pumps & fans, as typically predicted by the simulation software, were not separated in the measurements and are therefore grouped under 'Systems', L&P includes lighting, power and electric water heating (WaterSystems). Even though the end-uses were grouped, they can still be analysed separately to understand how the model behaves and what changes may be necessary for a more representative base case model.

Table 6.5: Predicted energy end-uses as defined for comparison.

End-uses	Model disaggregation
Systems	Heating, Cooling, Pumps, Fans
L&P	Lighting, Power, DHW (WaterSystems), Server
Gas	Gas

An initial model was set up based on available data from design specifications and the building audits. Consecutively, some additional changes were necessary to align the actual situation as the design specifications were not always in line with observations. Several adjustments were made to the initial model to resemble measured energy use more closely:

Model adjustments

- Air hander unit on ground floor was observed to be out of order and was disabled within the model, this air handling unit should have been providing fresh tempered air to three spaces in the basement.
- Out-of-hours baseloads were calculated based on analysing available lighting and power consumption on several floors and were determined to be 30% and 65% respectively.
- The systems in CH are using a constant amount of energy use throughout the week, with a similar pattern during the week and weekend. The heating and cooling schedules assume low occupancy during the weekend, but it was clear that systems were still operating, a seven-day operational cycle was therefore implemented.
- The systems were being operated almost continuously throughout the day and night at nearly the same baseload. To replicate this, the heating and cooling setpoint allowed night-time conditioning, increasing night-time setpoint temperatures.
- Introduced monthly adjustment factor to account for large deviations in measured electricity use between the months, due to large variations in measured occupancy levels.

These model adjustments brought the model predictions closer to the actual energy consumption in the building. A further refinement of the model was however infeasible due to the following limitations:

Limitations identified

- Heating and cooling could not be separated as the VRF systems provides both electrical heating and cooling, which made it difficult to understand underlying behaviour of system energy use. In addition, heating and cooling set-points are set individually for each room, making it difficult to represent this behaviour in the model.

- Electrical heating is available through zip taps and in the showers and could not be distinguished from measured power energy use, but is likely to be a large contributor to total power energy use.
- Material properties were determined during the walkthroughs and based on previous assessment of the physical structure, but no detailed design specifications were available. However, a recent refurbishment introduced internal insulation and secondary glazing to the building, for which some detailing is available.
- Labelling of electrical meters for lighting and power were unclear and could therefore not be disaggregated on each floor, some indications of their trends were however identified and used as a basis for the establishing separate lighting and power baseloads.

Predictions from the baseload model after model adjustments are compared to measured energy use, total energy use for the **measured** months is shown in **Figure 6.28**. A final breakdown between systems, L&P and gas energy use is shown, energy use is mainly from the systems (46%) and lighting and power (49%), with a smaller amount of gas energy (5%) use during the measured year. Differences between some of the months are larger than others, in particular January, March and April show differences of several kWh/m^{2}a.

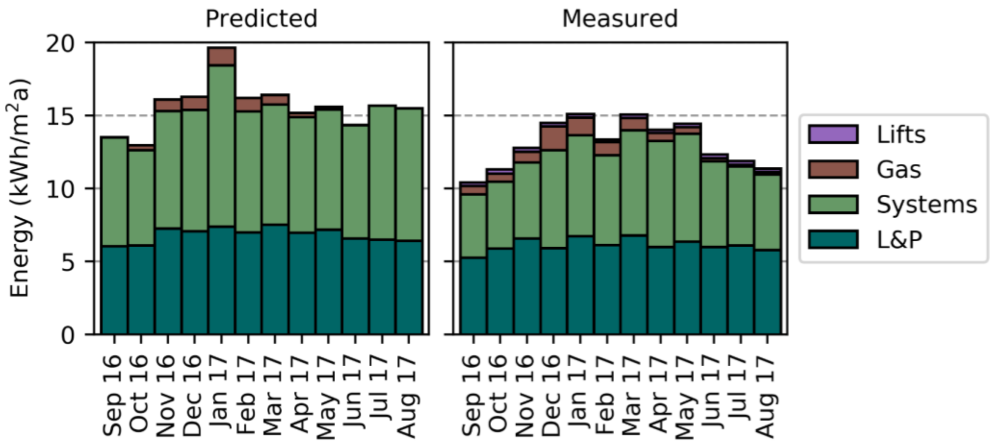


Figure 6.28: Total predicted and measured energy use for the measured months.

Differences use are somewhat clearer when looking at the percentage errors, shown by the NMBE and CV(RMSE) in **Figure 6.29**. A negative mean bias error indicates that the model under predicts measured energy use, the graph indicates the hourly variation per month, whereas the total numbers for the whole year are given by the values on top of the graph. In this case, the hourly NMBE for March, April and May differ by more than 20% where the total hourly NMBE criteria according to ASHRAE is 10%. All months taken together, the percentage difference comes down to -17.37%, which falls outside of the criteria. The CV(RMSE) based on hourly data per month is given on the right, with the total yearly CV(RMSE) per month on top, which is 26.86% which is lower than the 30% set by ASHRAE. On a monthly basis the hourly error is given, where most are higher than 30%. Indicating that some further calibration would be necessary to achieve a 'calibrated' model.

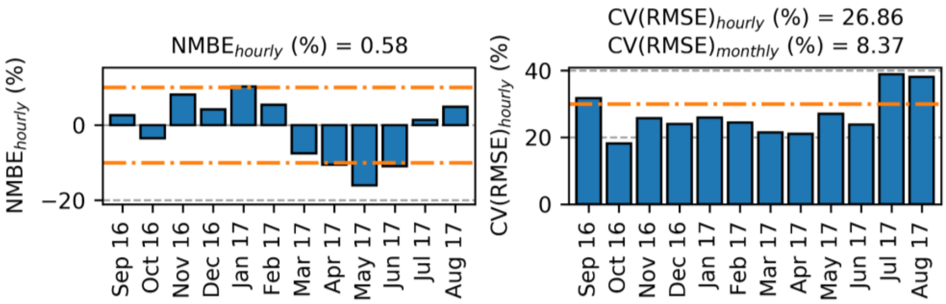


Figure 6.29: NMBE and CV(RMSE) based on hourly electricity use (excl. gas) calculated per month, and totals above.

It was determined that there is significant energy demand during the night, specifically from the VRF air conditioning system (meters R1 and R2) and L&P. To represent this, behaviour profiles and setpoint temperatures were adjusted to allow for some night-time conditioning, as can be seen in **Figure 6.30**. The flexibility of the system (user-control of the space temperatures) makes it difficult to define a specific control strategy in the energy model. For calibration purposes, a number of control strategies is therefore simulated through parametric simulation.

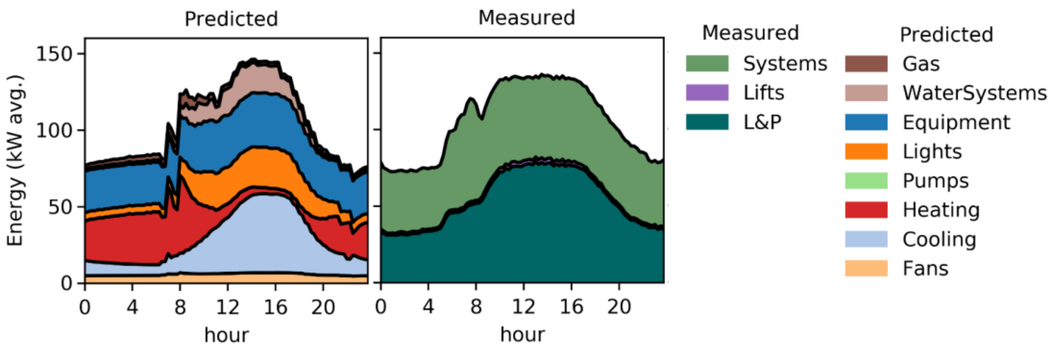


Figure 6.30: Predicted and measured typical weekday energy use for CH.

This initial insight gives an understanding of how temporal granularity of data can affect the accuracy of a model. If the model is calibrated on a yearly basis it would mask some of the underlying monthly data, and even more so for the hourly data. Hierarchical granularity presents another level of detail, the disaggregation of total electricity use in lighting and power (L&P) and system energy end-uses is however not detailed enough to understand the underlying behaviour of the building. It does not give a good understanding of when certain spaces are being heated or cooled, or when and how much domestic hot water energy contributes to the total power energy use. All of which are large contributions to energy use, the accuracy of the model in predicting these end-uses can have a large effect on the type of savings that might be calculated based on a calibrated model.

Predicted and measured systems and L&P electricity use are compared in **Figure 6.31** for a typical weekday and weekend day by overlaying their profiles for the whole measurement period. The base case model is able to represent the measured energy use profiles, peak loads are occurring at the same time during the day, while some differences exist during other periods, especially so for systems electricity. The shaded regions represent the standard deviation around the average. The differences represented by the error bar are mainly due to the seasonal factor that changes occupancy presence, lighting and power throughout the months, which directly influence systems energy use. However, for systems electricity use it can be seen that midday, especially for the weekday, there is a large standard deviation in predictions in contrast to the measurements. Going back to **Figure 6.30**, this seems to be due to the interaction of heating and cooling loads, where a significant amount of cooling is needed during the day, likely to deviate with reality. Due to the aggregation of measurements for systems, mitigating this difference is difficult, as the actual operation cannot be analysed.

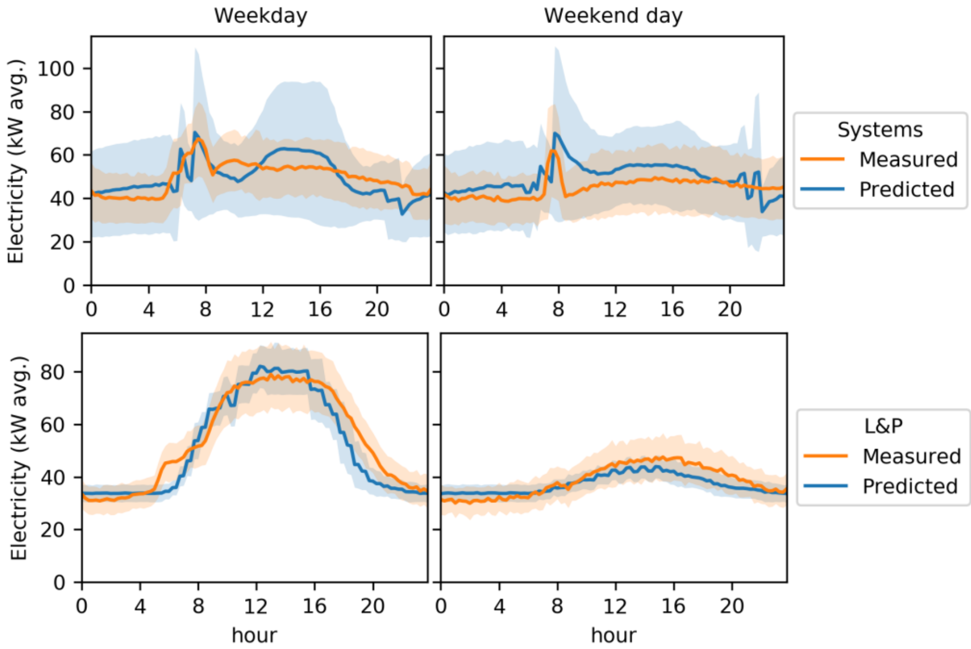


Figure 6.31: Predicted and measured system and lighting and power (L&P) energy use for a typical week and weekend day, shaded region represents the standard deviation around the average.

Although the typical weekday and weekend day graphs proved very useful in understanding trends in energy use, they mask some of the underlying behaviour, which can be further analysed by plotting actual data points, instead of averages. **Figure 6.32** shows predicted and measured system electricity use for two months, which reveals some discrepancies, masked by the typical day graphs. Predictions during September and October are very dissimilar, in contrast to measurements.

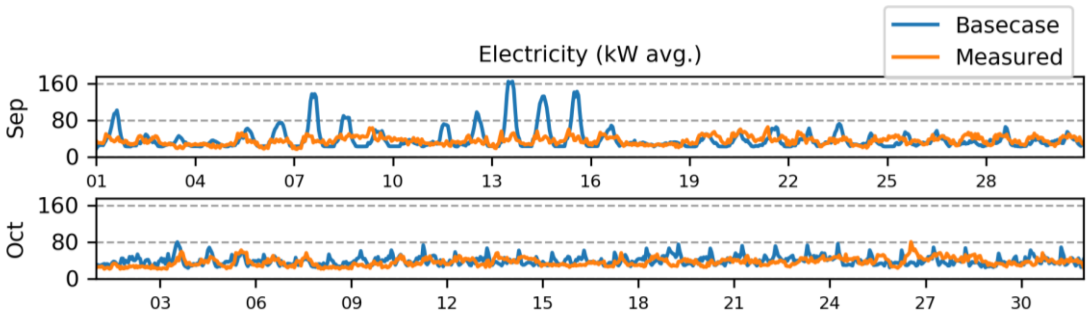


Figure 6.32: Hourly predicted and measured systems electricity use for September and October 2016 in CH.

Looking at only predicted heating and cooling for these months, it becomes clear that a large amount of cooling is required in September, with almost no heating load, whereas October more heating is needed and much less cooling. In contrast, measured system energy use is mostly unaffected by the change in weather (5°C drop, see **Figure 6.40**). An explanation for this could be that the temperature control as previously mentioned is highly fluctuating in the spaces and electricity use for heating and cooling is balanced. Whereas, predictions are based on determinant input parameters, e.g. space set-point temperature are the same in an agglomeration of spaces (based on space-type), having a much stronger effect on either heating or cooling loads.

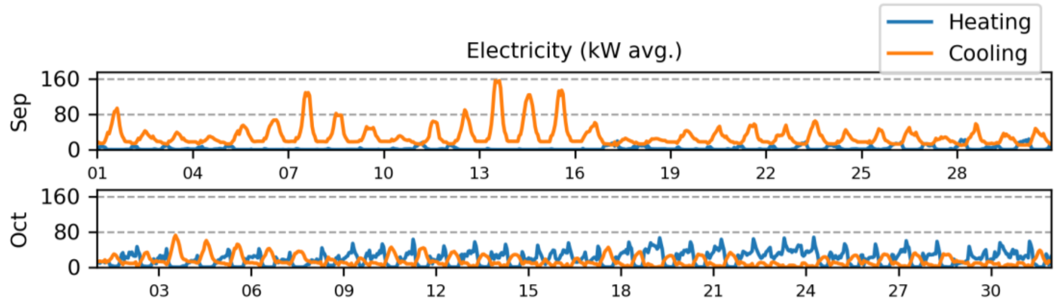


Figure 6.33: Hourly predicted heating and cooling energy use for September and October 2016 in CH.

As previously mentioned during the analysis of operational data, there is a large difference between lighting and power loads due to large variation in occupancy, occurring solely for the two university buildings. For this purpose, an additional seasonal factor was introduced that takes into account the seasonal variation, which will be used by the automated calibration process to minimise the monthly differences both at the monthly and hourly level.

6.4.2 Uncertainty analysis

Predicted energy use from parametric simulations is shown for the total and for each energy end-use separately in **Figure 6.34**. A large distribution exists for the predicted energy use as a total, predominantly inherited from the large variation for lighting and power energy use, whereas systems and gas energy use deviate from the median. From this, and the fact that the input variables are all changed at an equal percentage it follows that the variable parameters have less effect on systems and gas energy use. This was expected, as more input parameters related to lighting and power were variable, such as lighting and equipment power densities, which directly influence their energy use.

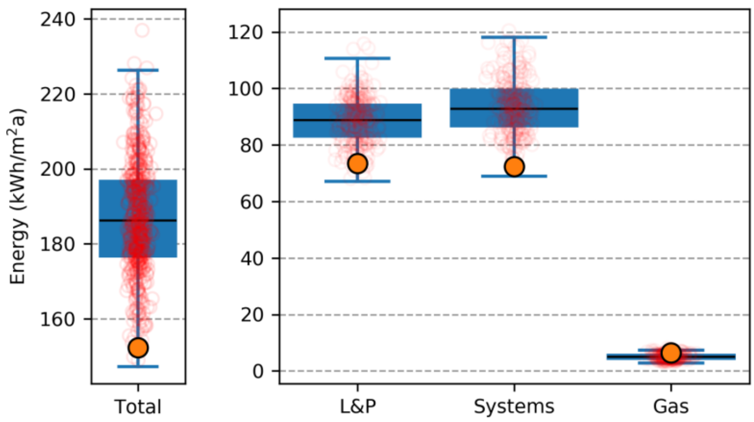


Figure 6.34: Total predicted (boxplots) and measured (orange dots) energy use for 3000 simulations.

Total predicted energy use for all runs is broken down per month to understand how well seasonal variation is presented in the predictions, see **Figure 6.35**. For most months, predictions follow a similar pattern to the measured energy use, while for some, measured energy use lies at the far end of the distribution of predictions. Even though these trends were already established in previous graphs when comparing the base case with measured energy use, they give and understanding of when the variation in inputs are more significant. For example, in the month of January, systems energy use shows a significantly larger variation than in other months, similarly so for the month of March for lighting and power energy use.

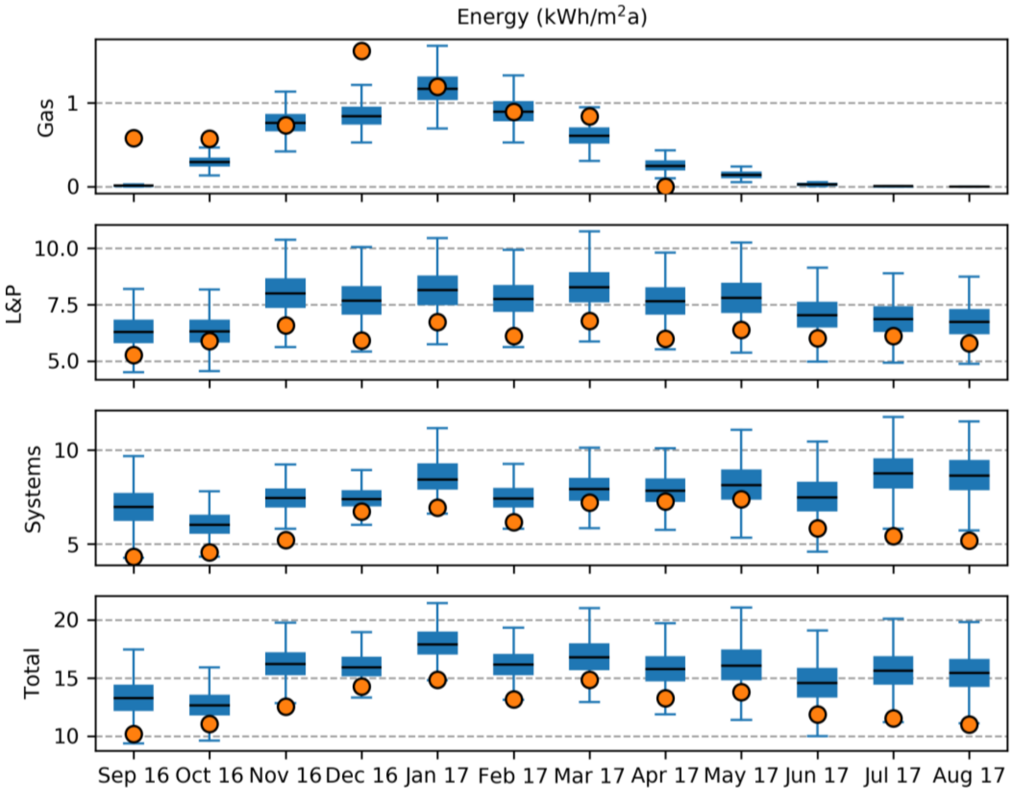


Figure 6.35: Monthly predicted (boxplots) and measured (orange dots) energy use for 3000 simulations.

6.4.3 Sensitivity analysis

Spearman's rank correlation coefficients were calculated for the inputs and outputs based on the 3000 simulations. In total 85 variables were varied during parametric simulation, the most significant parameters, those with a coefficient of $\rho > 0.25$ and $\rho < -0.25$ are shown in **Figure 6.36**. Insignificant parameters were:

- The CCOP and HCOP of the VRF heat pumps (max $\rho < \pm 0.1$), which were varied at a 5% standard deviation as a normal distribution,
- Natural ventilation rate ($\rho < \pm 0.05$),
- Lighting and equipment power density for space types of which there are only a few spaces, such as the library ($\rho < \pm 0.1$), reception ($\rho < \pm 0.05$), computer labs ($\rho < \pm 0.08$)
- Exhaust fan efficiency ($\rho < \pm 0.05$), they have a relatively low energy consumption.
- Boiler no. 2 ($\rho < \pm 0.01$), indication that it is not used in the model.
- Seasonal weekend factor for all months ($\rho < \pm 0.03$), whereas the seasonal week factor ($\rho < \pm 0.2$) is much more significant as energy use during the week is much higher. These were varied for automated calibration purposes and adjust the monthly lighting, power and occupancy schedules.

Sensitivity indices are relative, dependent on the other varied input parameters included and their ranges of variation. They will be different when a different amount of parameters is included in the parametric simulation. Furthermore, they will vary for different buildings. However due to the linearity between typical uncertain parameters and predicted energy use, sensitivity indices can to some extend be predicted prior to simulation, based on a general understanding of the interrelations between inputs and outputs. For example, when two spaces are modelled, one is twice as large as the other, with the same equipment power density inputs and range of uncertainty. Its sensitivity can then be inferred to be twice as strong, i.e. their equipment power use in the larger zone is twice as large and will have a stronger influence on total equipment energy use. Nevertheless, this is more complex when many parameters are included and even more so when input parameters are included that affect for example the shape of a schedule, or one that determines the temperature difference between set-points.

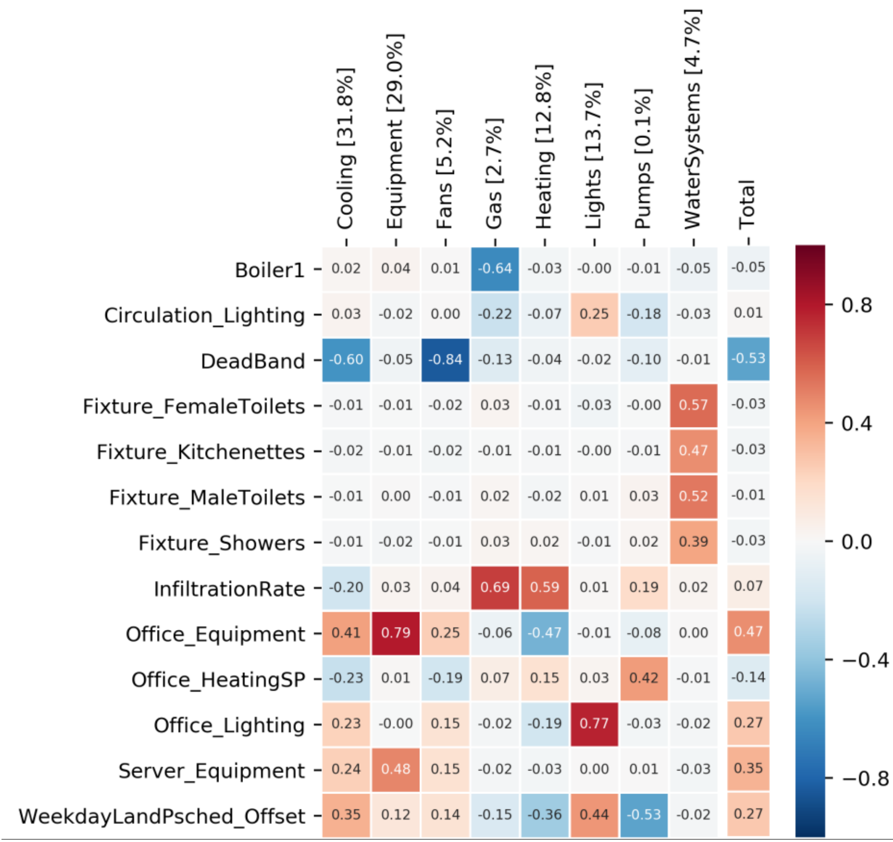


Figure 6.36: Spearman rank correlation coefficients per energy end-use for CH.

The significant parameters in certain cases have a large influence on multiple parameters, such as the office equipment and lighting power density, infiltration rate and set-point temperatures. Whereas others, such as the boiler efficiency and flow rates of hot water fixtures influence only gas energy use ($p = -0.64$) or water systems energy use (DHW) ($0.39 < \rho < 0.57$) respectively. The heating setpoint in the offices and 'DeadBand', which determines the cooling temperature set point, have a large influence on their respective heating and cooling energy use, but also significantly influence auxiliary system equipment such as fans and pumps. Interestingly, the heating set point has a larger influence on cooling energy use ($\rho = -0.23$) than heating energy use ($\rho = 0.15$). Sensitivity analysis is helpful in understanding how different components in a building affect energy use, however if it is to be used as an indication for where to make savings, it is important that relevant input parameters are taken into account as uncertain variables. However, when taking into account the proportion of the different energy end-uses to total energy use, some of the parameters that were considered significant previously, such as the hot water fixtures, boiler efficiency, becomes less important. As cooling and equipment energy use are the largest contributors to total energy use, the variables that have a significant effect on those become more relevant. These are the heating and cooling set points, office and server equipment power densities and also the lighting and power offset. The lighting and power offset determine the width of the lighting and power schedules. Their effect on cooling energy use is shown visualise in **Figure 6.37**, where it can be seen that a large temperature difference between the heating set point and cooling set point decreases cooling energy use, while an increase in the width (increase in power density of lighting and equipment) of the L&P profiles will increase cooling energy use. The strength of the relationship for the L&P profile offset is dependent not only on its own value, but also on the equipment and lighting power density variables, similarly so for the cooling temperature set-point. If this were not the case, their r-squared value would be likely be higher as there would be stronger relationship between the two variables.

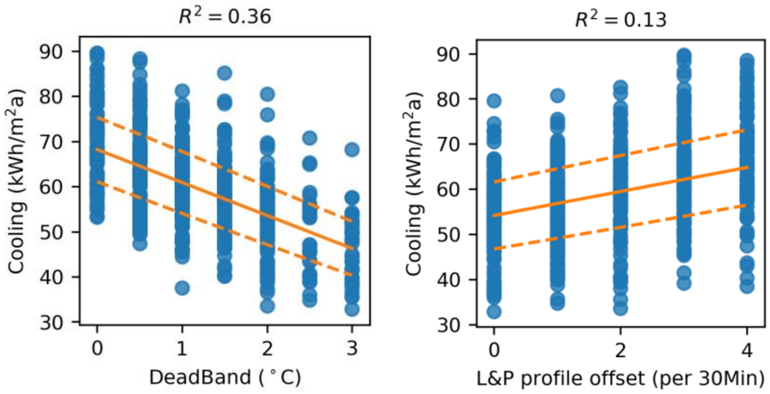


Figure 6.37: Relationship between input parameters 'DeadBand', which determines the cooling set point temperature and L&P schedule offset against cooling energy use.

Several of the input parameters, mainly boiler efficiency, the heating- and cooling set points and infiltration rate will vary in their sensitivity throughout the seasons (or even the week and weekend). However, as previously remarked, most of this can be understood by having a general understanding of the fundamentals of building energy modelling and how energy changes throughout the year based on different assumptions. **Table 6.6** shows the spearman correlation coefficients of four input parameters and their effect on monthly energy use. Giving an indication of how significant they are throughout the seasons.

Table 6.6: Spearman correlation coefficients of several input parameters on monthly energy use.

Parameter	Sep-16	Oct-16	Nov-16	Dec-16	Jan-17	Feb-17	Mar-17	Apr-17
Boiler 1 (0-1)	0.01	-0.02	-0.06	-0.1	-0.13	-0.1	-0.05	-0.01
DeadBand (°C)	-0.19	-0.16	-0.21	-0.22	-0.37	-0.27	-0.18	-0.17
Office heating SP (°C)	-0.16	-0.07	0.01	0.04	0.21	0.05	-0.03	-0.06
Infiltration rate	-0.05	0.02	0.17	0.14	0.38	0.26	0.14	0.04

As can be expected in the British climate, the infiltration rate has a more significant effect on energy use during the winter months, ranging from ρ -0.05 to 0.38, whereas the significance of the office heating setpoint and boiler efficiency ('Boiler 1') also strongly fluctuate during the seasons. The meta-models developed in the next chapter rely heavily on the relations between inputs and outputs, understanding their significance is then important to decide if these should be taken into account or can be discarded.

6.4.4 Impact of assumptions

The impact of typical (NCM) assumptions on the base case model have been assessed, the applied simplifications and their impact in percentage difference per energy end-use are given in **Table 6.7**.

Table 6.7: Effect of simplifications on the calibrated model
as a percentage difference per yearly energy end-use.

	Simplification	Cooling	Equipment	Fans	Gas	Heating	Lights	Pumps	DHW	Total
S0	No seasonality	72	37	14	-21	-40	76	-10	-1	34
S1	Occupancy density	-5	0	-1	-2	4	0	-2	0	-1
S2	Equipment power density	35	50	11	1	2	0	1	0	27
S3	NCM schedule for equipment, lighting and occupancy	-21	-22	6	19	59	-27	6	0	-4
S4	NCM schedule for office heating and cooling	2	0	-15	-28	-71	0	-11	0	-14
S5	Infiltration (16 to 10 m³/m²h @ 50Pa)	8	0	-2	-21	-19	0	-3	0	-1
S6	Combination of S0 to S2	100	87	20	-22	-36	76	-10	0	58
S7	Combination of S1 to S3	5	28	13	17	56	-27	5	0	19
S8	Combination of S0 to S4	76	48	2	-40	-68	29	-14	0	28
S9	Design weather	-17	0	4	25	30	0	11	1	1

In addition to the percentage differences, the total annual energy use per floor area for the base case model and each simplification applied to it is given in **Figure 6.38.**

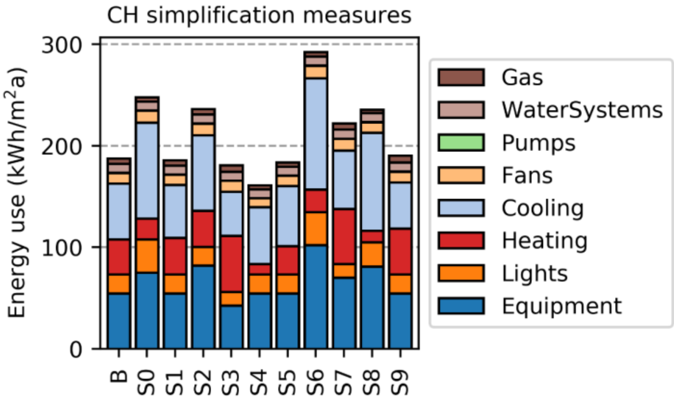


Figure 6.38: Implications of applying typical assumptions on energy use.

In contrast to Office 17 and 71, a seasonal factor based on the identified occupancy patterns was implemented in both CH and MPEB. It was found that the number of occupants in a building between term periods can differ by a 100%, found to have a significant correlation to energy use in university buildings. **Simplification 0** (S0) differs from the base case model by discarding the monthly seasonal variation factor, assuming a constant occupancy presence throughout the year, which directly affects the equipment and lighting profiles, resulting in a 34% increase in total energy use. An increase in both occupants, equipment and lighting leads to a higher cooling load and lower heating load conditioned from the VRF system and gas boiler. In non-domestic buildings, equipment power density is a dominant energy end-use, which has a significant influence on system energy use.

Making accurate prediction of equipment power density is important for making accurate predictions of energy use, both during design and for calibrated models in operation. **Simplification 2** introduces NCM assumptions for equipment power density (W/m²) to the base case model, in particular notable differences were identified for the large library, computer clusters (IT workspace) and server in CH. For NCM, a library has a typical equipment power density is 2 W/m², however the library is actively used by students who bring their laptops to this space, it was therefore assumed to be much higher, counting equipment in this space led to an assumption of 12 W/m² in the base case model. Similarly, the computer cluster, which is a high density IT space is 30 W/m^2 according to NCM, but was found to be about 14 W/m^2. The server room was found to be much less intense than the NCM assumption, with 250 W/m^2 instead of the typical 500 W/m^2. Nevertheless, for the offices, the typical 12 W/m^2 was very similar to that determined through the energy audit, which makes up most of the floor area in CH. Lighting power density was found to be very similar to NCM assumptions. Lighting can be predicted with more certainty as it is typically less influenced by people presence and in most space types a certain lux level is to be achieved, which correlates to its power density.

Equipment, occupancy and lighting are determined in building performance simulation through defining their density per space or typically space type. Actual equipment or lighting power and no. of people is then calculated through the multiplication with unity scaled profiles (i.e. profile between 0 and 1). For NCM, these profiles were not representative of the actual building as shown in **Figure 6.39**. Although occupancy follows a very similar pattern as the NCM schedule, lighting, and in particular equipment NCM profiles are very conservative in their baseload. A load of around 5% is assumed as the equipment baseload for the NCM profile, where it was found that in reality this is about 60%, a significant difference. Furthermore, NCM assumes that the building is unoccupied during the weekend, which was not the case. **Simplification 3** compares these profiles, showing a 4% reduction in energy use when applying the NCM schedules.

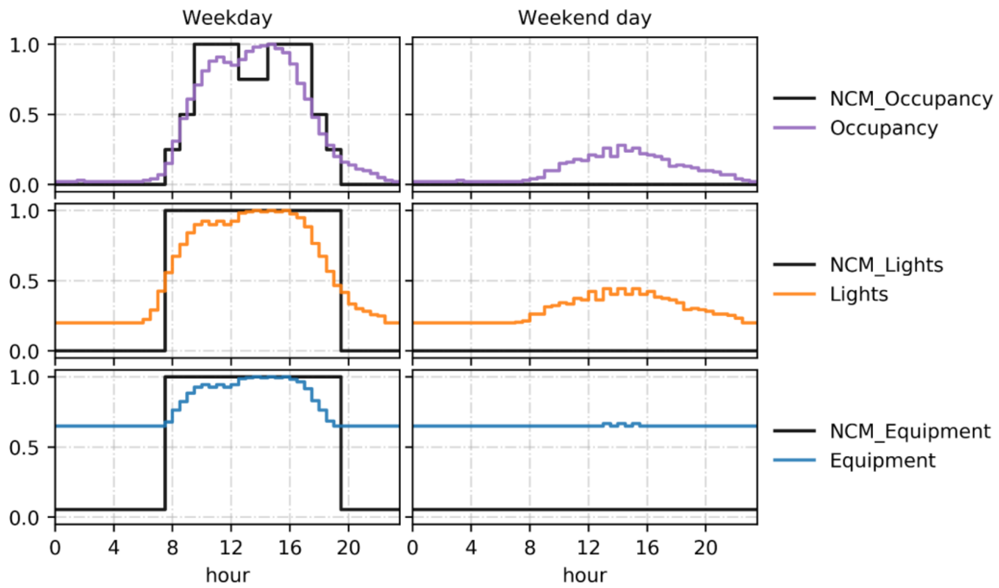


Figure 6.39: NCM schedules for occupancy, equipment and lighting and assumed base case profiles.

Simplification 4 compares heating and cooling set-point temperatures for the offices and similar spaces, which NCM assumed to be H: 22, C: 24 and 12°C during unoccupied hours. In reality, it was found that the VRF heat pumps were operating 24/7, with a small peak during the morning. Furthermore, temperatures in the spaces were found to be very stable instead of fluctuating between these temperatures, to replicate energy consumption of the VRF systems, the set-point temperatures were therefore assumed to be; H:24, C:24, and 20°C during unoccupied hours. This simplification led to a decrease of total energy use by 14%, predominantly due to a 71% reduction in heating energy use. **Simplification 5,** a change in the infiltration rate, has a negligible effect on energy use.

Simplifications 6 to 8 combine the previous simplifications. **Simplification 6** combines the assumption of no seasonality and an increase in equipment power density, increasing total energy use by 58%. **Simplification 8**, which combines the first four simplifications increases total energy use by 28%, these simplifications do however not represent all NCM assumptions in typical compliance modelling, but do cover the ones that have the largest impact on energy use.

Finally, **simplification 9** compares the local and design weather file. The design weather file for the location of London Gatwick is used. This is an IWEC[6] weather file, originally developed by ASHRAE. Gatwick is the closest available weather station from which publicly available weather files are created. **Figure 6.40** shows the differences between these two weather files, the dry-bulb temperature is shown as the monthly mean and standard deviation. There is a large difference throughout the year, the temperature in the centre of London is significantly higher than Gatwick which is located on the outskirts of London. Such a difference has an effect on the predicted energy use.

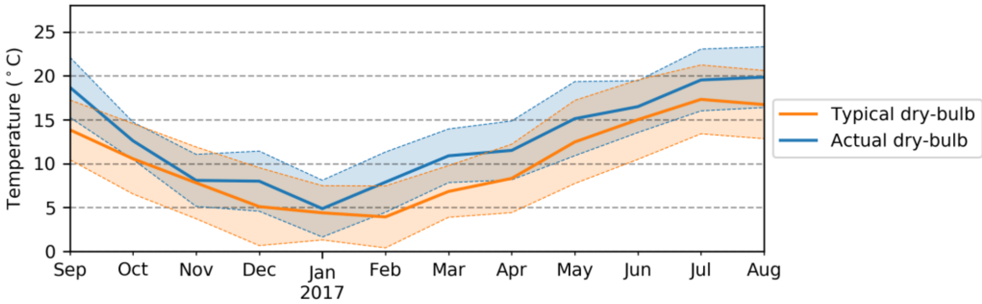


Figure 6.40: Monthly mean and standard deviation of external dry-bulb temperature from the local weather station (actual) and Gatwick weather station (typical).

The base case model is simulated using both the local weather file and typical weather file and evidently the is a large difference between predictions, mainly for systems and gas energy use, which are highly dependent on external weather conditions. Lighting and power energy use is ignored as these don't change depending on outside weather (although more people might stay home when it's too cold to go outside!). In **Figure 6.41** monthly weather use for both simulations are compared, there are significant differences between the months for both gas and system energy use. For the months of March and April gas energy use is twice as much for the typical weather file compared to the local weather file. For systems energy use differences are most significant during the early months of the year, February to April 2017.

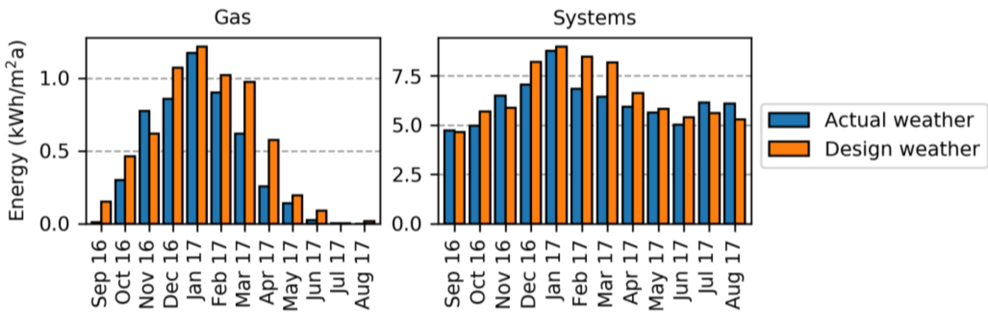


Figure 6.41: Monthly energy use for actual and design (Gatwick) weather files

A counter argument to the reasoning that a local weather file is essential for model calibration is that weather conditions mostly affect systems electricity and gas energy use, both of which

[6] https://energyplus.net/weather/sources

are typically hard sized in a building energy model (besides some other parameters, such as the opening of windows). Thus, parameters for these systems (e.g. a VRF system or boilers) are not actually supposed to change when calibrating the model to measured data, because these elements carry the least uncertainty. However, this would only hold true when measured energy use from lighting and power is available. Because, if measured system energy use is higher than predicted, the modeller might determine that L&P power density loads were incorrectly defined and need to be higher in order to increase system energy use (e.g. higher cooling loads for an increase in equipment power density). It is therefore difficult to determine the exact effect a different weather file might have on the accuracy of model calibration, comparing differences between predictions of two weather files is evidently not enough to understand the intrinsic effect on calibration. Nevertheless, it was determined that the total predicted energy use between the two weather files differed by 4% for a single year.

6.5 MPEB

6.5.1 Predicted vs. measured performance

The sub-metering system in MPEB measures many different parts of the building, some of which include a mix of components (e.g. electric water heaters, FCU fans on power meters). As such, the system does not always allow for separating different end-uses. Different meters were combined to form typical end-uses, these were then replicated in as much detail as useful by the energy model, to enable a like-for-like comparison. For MPEB, this meant combining pumps, heating and certain fans as ‘Systems’ energy use, as these are not separated by the sub-metering system. Nevertheless, even the disaggregation of systems energy use needs to be analysed when calibrating the model as different parameters may influence different types of energy end-uses, thus also for sensitivity analysis a higher level of energy use disaggregation is analysed. **Table 6.8** shows the energy end-uses used for comparison, lighting and power were measured only measured on several floors.

Table 6.8: Energy end-uses as defined for comparison

End-uses	Model disaggregation
Systems	Plant Fans, Pumps, Heating
L&P	Lighting, Power, FCU Fans, DHW (WaterSystems)
Servers	409 Machine room, G01b Machine room
Cooling	Chiller 1, Chiller 2

An initial model was set up based on available data from building design specifications and the building audits. Consecutively, some additional changes were necessary to align the actual situation as the specifications were not always in line with was observations. Several adjustments were made to the initial model to resemble measured energy use more closely:

Model adjustments

- Two servers DB409 (large server on fourth floor) and GP03 (server on ground floor) were separated as server energy use.
- Out-of-hours equipment and lighting schedule input parameter that sets the baseload was adjusted to 85% and 65% respectively, based on analysis of the available separate lighting and power meters.
- Introduced monthly adjustment factor to account for the large deviations in measured electricity use between the months.

Limitations identified

- Separately measured L&P on floors did not coincide with total consumption on bus bars, it was identified during talks with UCL estates that some of the server load is on the L&P bus bars instead of on a separate meter. This made it difficult to distinguish the exact server load. Total L&P is therefore based on separate metering of L&P on several floors.
- Electric water heaters were added to L&P energy use as they are connected to sockets in the spaces, similarly so for the FCU fans.
- District heating was not measured due to a faulty heat meter throughout the measurement period, nevertheless some information was available in regards to district heating supply temperatures and heating system were hard-sized where information was available.

Again, major limitations in the metering system were identified that made it difficult to get an accurate representation of disaggregated energy use within the building. It proves that metering systems and commissioning of these systems need to performed rigorously to ensure that exhaustive data is available, which is becoming more important with the increasing demand for operational energy efficiency.

Predictions from the baseload model after model adjustments are compared to measured energy use, total energy use for the measured months is shown in **Figure 6.42**. Energy use is mainly from lighting and power on the floors (37%), servers (27%), systems (13%), chillers (12%) and workshops (7%), district heating was determined to be another 3%. Predicted and measured monthly energy use align considerably well, with some months exhibiting a higher fluctuation than others.

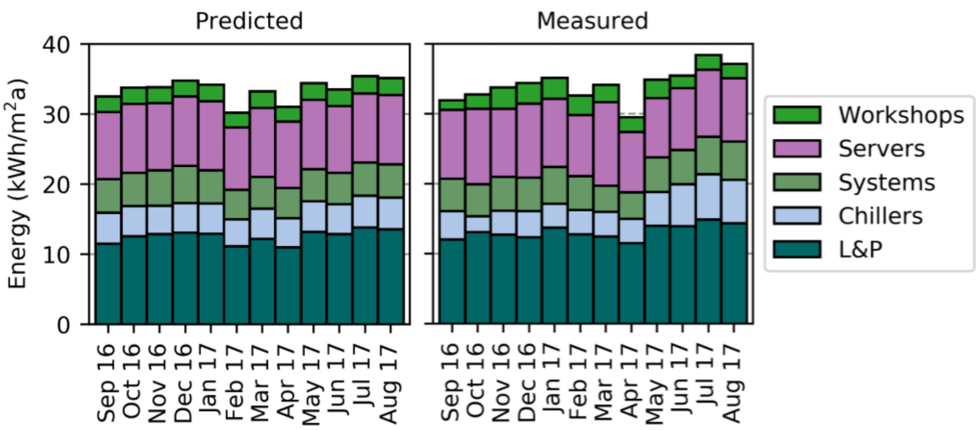


Figure 6.42: Predicted and measured monthly energy use for the energy end-uses in MPEB.

Comparing energy use for a typical weekday and weekend day in **Figure 6.43** shows that the distinct measured energy use profile with high baseload was replicated with the model. The server load is a large part of the high baseload in this building.

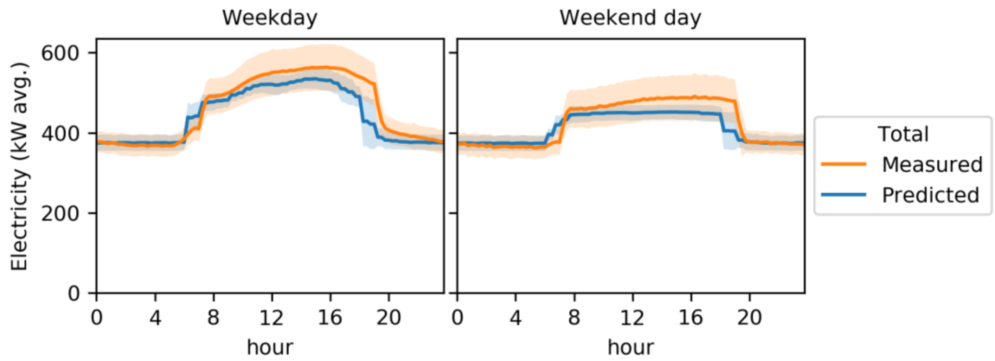


Figure 6.43: Predicted and measured total energy use for a typical week and weekend day, shaded region represents the standard deviation around the average.

Calibration of the model initially focussed on determining the right lighting and power loads in the building by adjusting the equipment and lighting power densities in addition to determining the typical profiles of use. Other energy end-uses, such as those related to conditioning of the building are strongly influenced by these loads. System components that consume heating, cooling, fan and pump energy use are ideally hard sized, however these types of energy uses are dependent on many variables and need detailed investigation to calibrate accurately, some of this information might not be available from building design specification and would need to be determined through observation.

The server equipment load has been adjusted iteratively to achieve measured energy use levels, this in turn significantly increased cooling loads and systems energy use, both of which were then found to be over predicted. Server energy use, lighting, and power on the different floors is shown in **Figure 6.44** and **Figure 6.45** respectively. Server energy use is significant and it was determined that some of the load was actually connected to L&P meters on the upper floors, some of these meters were not measured separately and it was therefore difficult to determine the exact magnitude of this load. Therefore, total energy use was taken from the bus-bar, by subtracting and extrapolating typical lighting and power loads for the measured floors then determined the additional server load. In addition, the large server room (DB409), shows a large increase of energy use during occupied hours, indicating that perhaps other types of equipment are connected on this meter. Whereas the smaller server room (GP03) exhibits a continuous pattern of use.

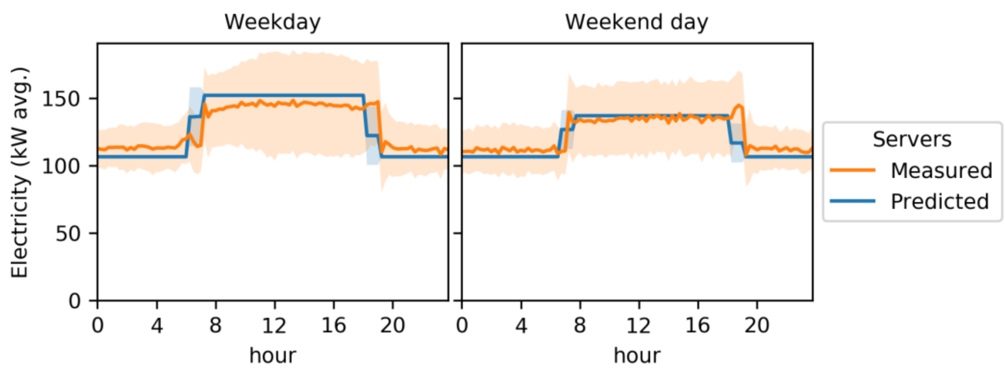


Figure 6.44: Predicted and measured servers energy use for a typical week and weekend day, shaded region represents the standard deviation around the average.

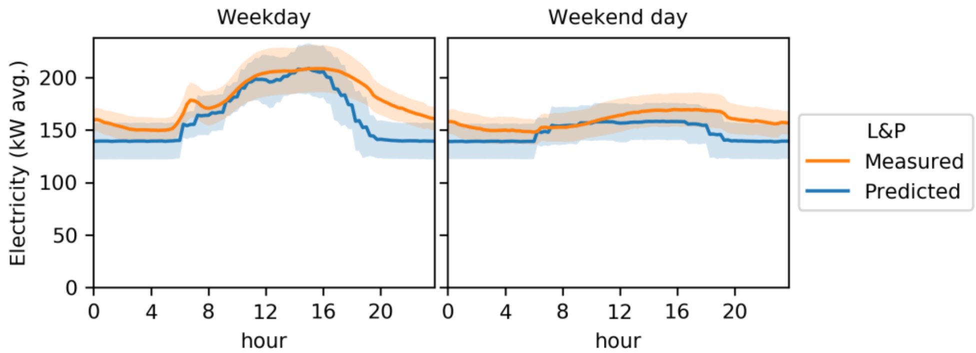


Figure 6.45: Predicted and measured lighting and power energy use on all floors for a typical week and weekend day, shaded region represents the standard deviation around the average.

The model predicts lighting and power energy use on the floor reasonably accurate, the profiles are smoother for the actual situation and lies somewhat higher. The difference per month is around 1-2 kWh/m^2, and in total a difference of 7.6 kWh/m^2 for the measured months.

Sizing the systems

With lighting and power defined, the hard sized systems for conditioning should be predicting the demand for system energy use more accurately. Systems were incorporated into the model, based on design specifications and commissioning data, reducing their uncertainty within the model. A representation of AHU1 in the OpenStudio software is given in **Figure 6.46**. OpenStudio provides an interface for creating different system types, the whole building set-up in this software and then exported to an EnergyPlus simulation file (.idf file). Predicted system performance is analysed by looking at specific nodes for the supply and return air temperatures, other measurement data is not available for the air handling units.

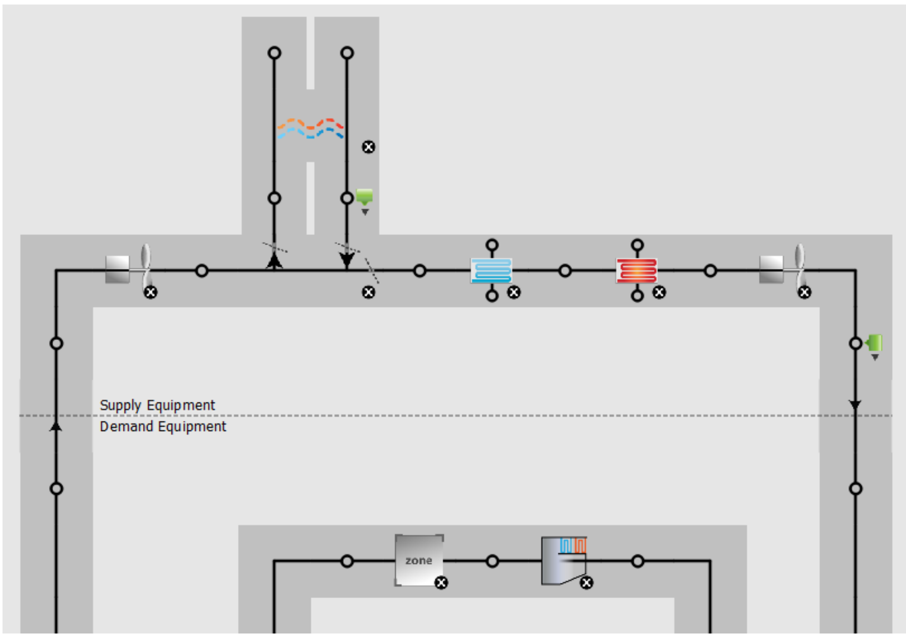


Figure 6.46: Virtual representation of AHU1 in building simulation software.

The air handling unit supply and return air temperatures are compared to measurements as shown for AHU1 in **Figure 6.47** and **Figure 6.48** respectively. Inside the software, different types of strategies or schedules for conditioning can be set to either control the supply air temperature, 18°C for AHU1. Which, as can be seen in **Figure 6.47,** is replicating measured behaviour. The higher measured temperatures here are when the air handler is out of operation during unoccupied hours. The return air temperature however reveals some additional patterns, simulated and measured temperatures do not match, simulated return air temperatures are around 25°C, while measured temperatures are around 22°C for the two analysed weeks. The return air temperature is based on 10 different zones (including lecture theatres and computer 'labs' and is uncontrolled. Indicating that differences exist between predicted and measured space temperatures.

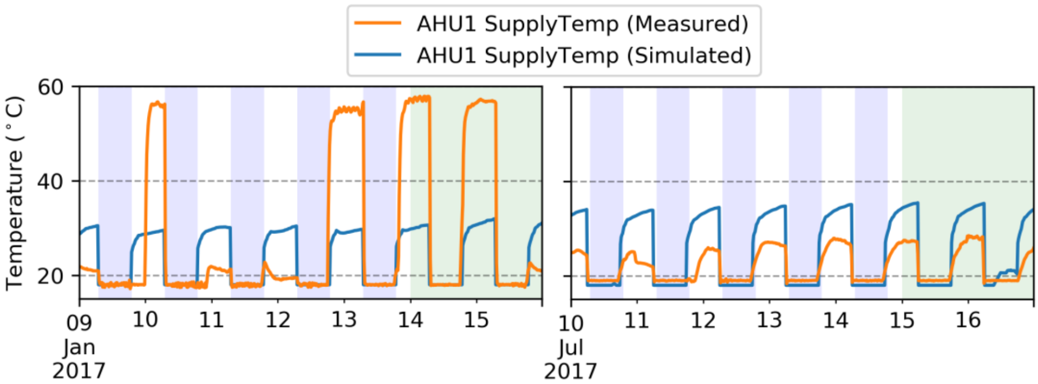


Figure 6.47: Simulated and measured air handling unit supply air temperatures for the 2nd week of January and 2nd week of July 2017.

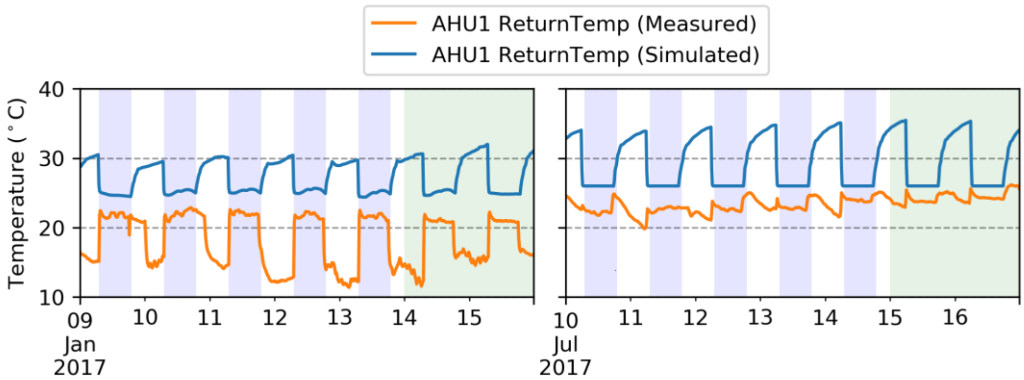


Figure 6.48: Simulated and measured air handling unit 1 return air temperatures for the 2nd week of January and 2nd week of July 2017.

Figure 6.49 show the predicted and measured systems energy use for a typical week and weekend day respectively. Even though profiling between the model predictions and measured data follows a similar pattern, some difference exists. The hours of operation of the systems seem to be slightly off by about an hour and the baseload is measured to be slightly lower, while energy use during occupied hours is higher than predicted. However, monthly predictions and measurements align considerably well, as shown in **Figure 6.50**.

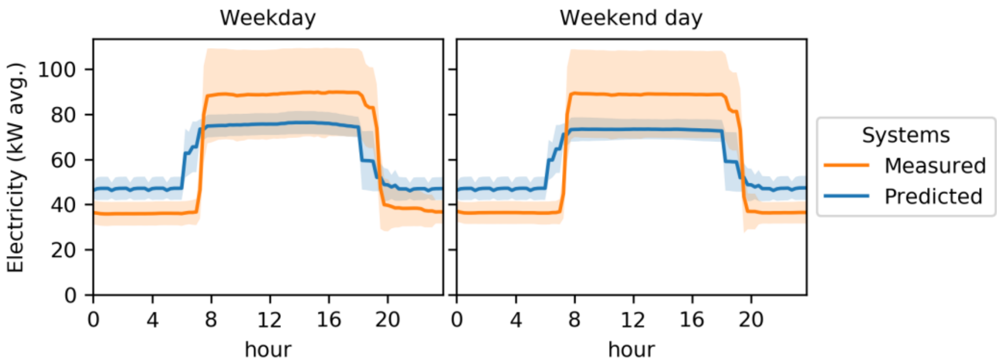


Figure 6.49: Predicted and measured systems energy use for a typical week and weekend day, shaded region represents the standard deviation around the average.

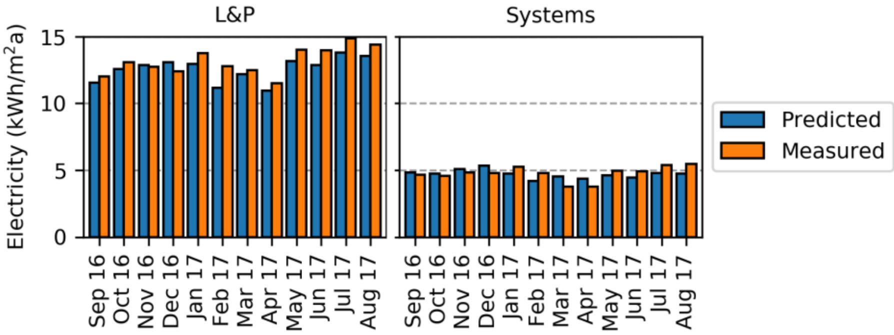


Figure 6.50: Monthly predicted and measured lighting and power and systems energy use for MPEB.

Finally, chiller energy use was not replicated accurately, measured chiller energy use showed large fluctuations throughout the year, whereas the model predicts a very constant use of energy throughout the day and night. This indicates that some of the underlying physics in the building were not accurately captured, even though predicted and measured energy use are 'calibrated' as shown in **Figure 6.51**.

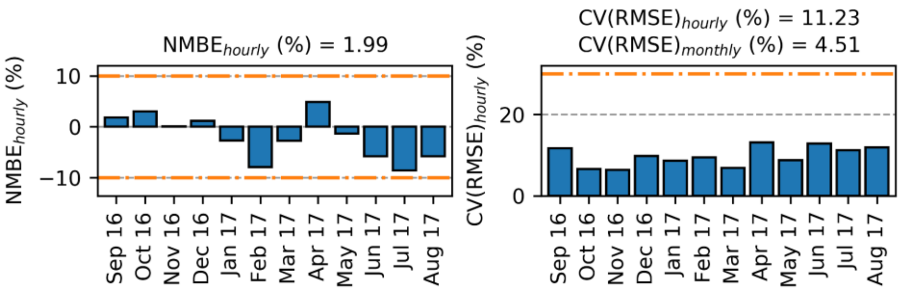


Figure 6.51: Statistical measures denoting differences between predicted and measured energy use, given by the average monthly NMBE and CV(RMSE).

Chiller energy use energy use follows a similar pattern during the winter where some cooling is still necessary (mainly the servers), while during the summer the measured chiller energy use is highly fluctuating and more dependent on the weather and increase in cooling loads, the model however does not display this behaviour. It is clear that the model needs additional adjustment to represent the actual building operation. More importantly, MPEB and the previous case study buildings have shown similar results, monthly energy use can relatively easily be replicated by a model, but the underlying hourly levels of energy use are not necessarily in line with measured data. The underlying physical processes at this level of temporal granularity are masked through wrong assumptions.

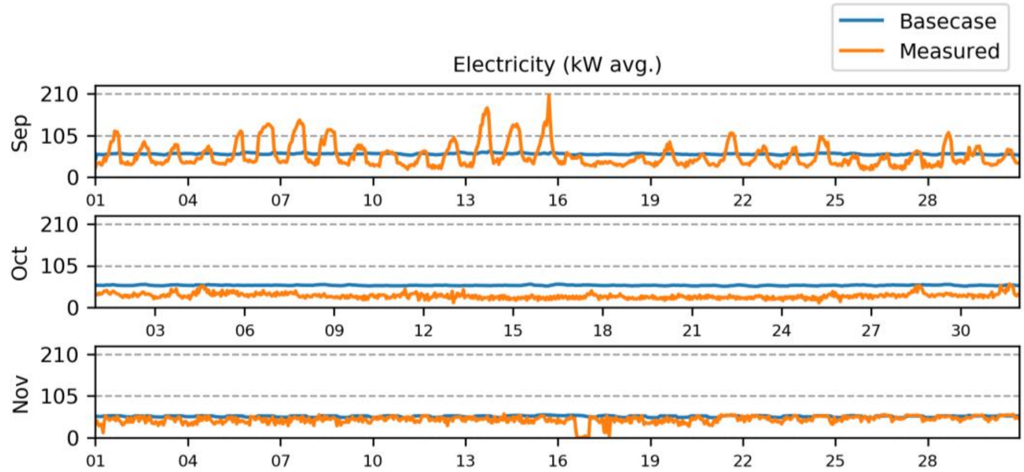


Figure 6.52: Chiller energy use during September to November 2016 for MPEB.

Monthly predictions and measurements of energy use for the different end-uses are compared by calculating the NMBE and CV(RMSE) statistical measures at a monthly and hourly interval as shown in **Figure 6.53**.

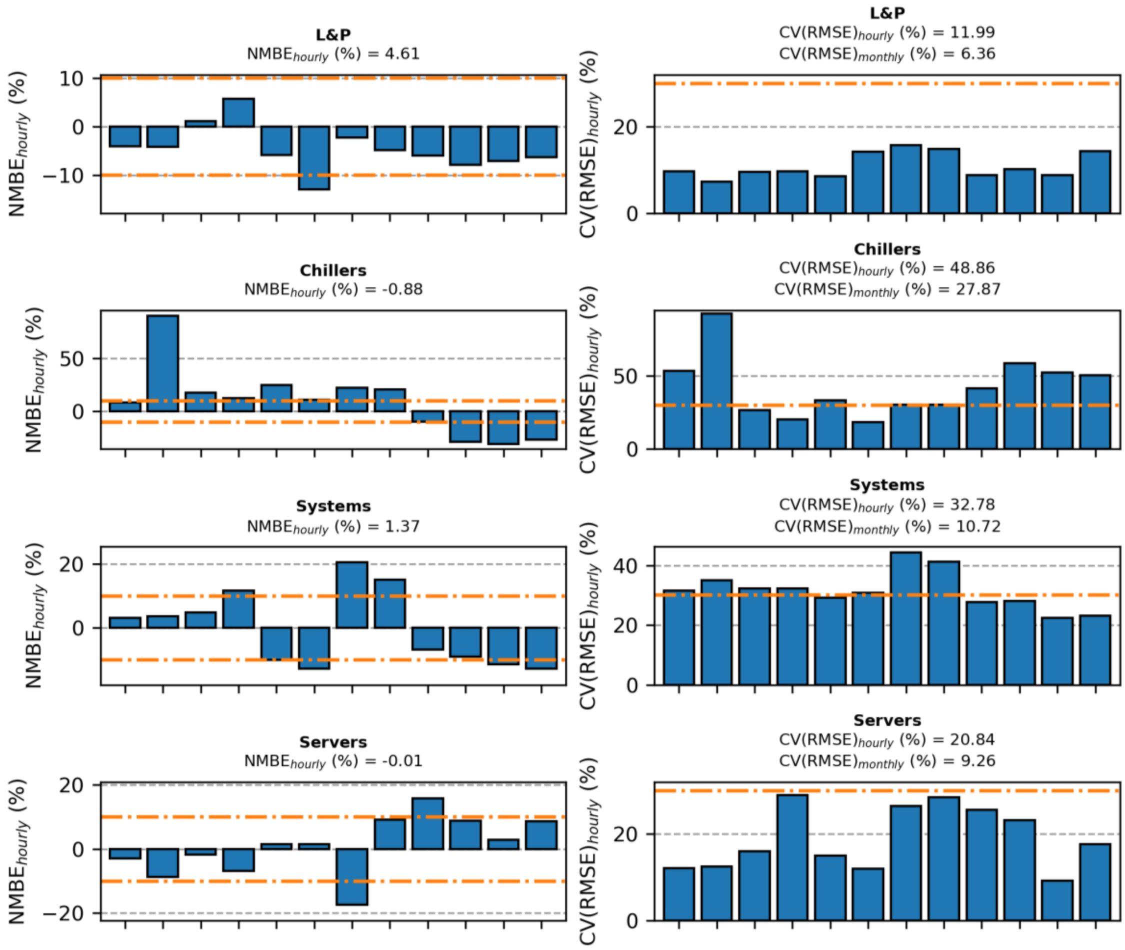


Figure 6.53: Statistical measures denoting differences between predicted and measured energy end-uses, given by the average monthly NMBE and CV(RMSE).

For the different energy end-uses, the total monthly and hourly NMBE values are all within the ±5% and ±10% criteria. However, the CV(RMSE) values tell a somewhat different story, total monthly and hourly criteria, <15% and <30% are only satisfied by L&P and servers. In conclusion, the model is 'calibrated' at the total monthly energy use, but differences become clearer for the energy end-uses. While some can still be considered calibrated, others could still be improved, mainly at an hourly level.

Space temperatures

The discrepancy between simulated and measured return air temperatures from the air-handling units indicated that a difference exists between the simulated and measured space temperatures. Previously the space temperatures have been analysed and showed that there is a large fluctuation both during the year and between spaces for both the same and dissimilar space types. The analysis showed that the set-point temperatures in some of the spaces are changed throughout the year, which is difficult to replicate for each space individually. As such, a comparison is made between simulated and measured space temperatures in an office space 221A and lift lobby 470 (people work in this space as well), as shown in **Figure 6.54** and **Figure 6.55**.

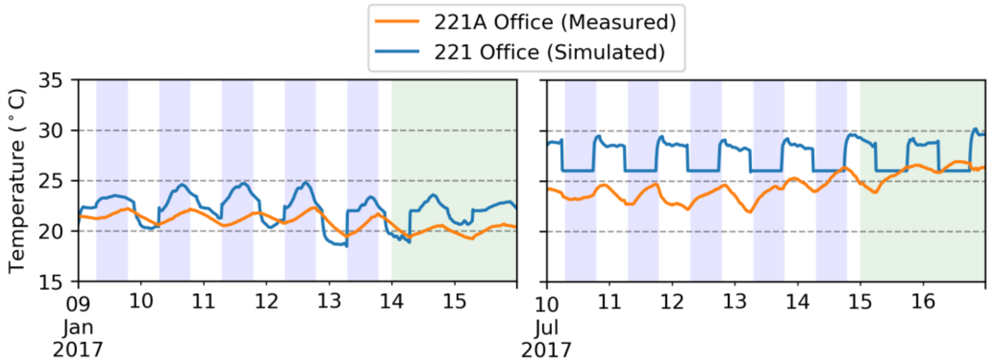


Figure 6.54: Simulated and measured space temperatures in 221A office, for the 2nd week of January and 2nd week of July 2017.

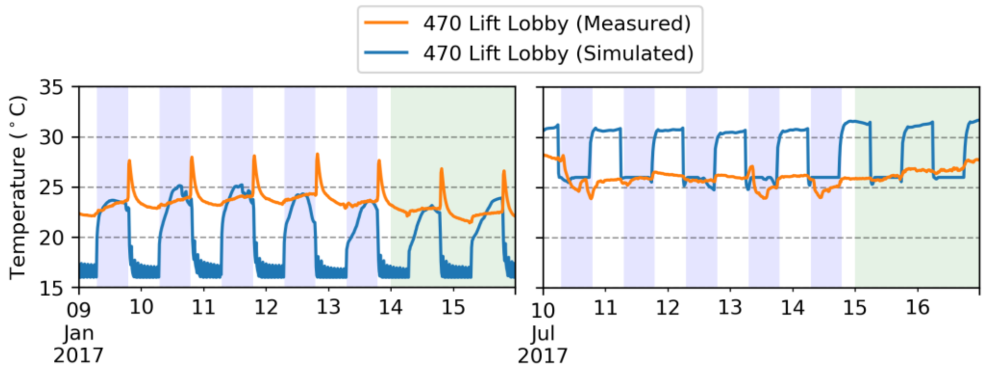


Figure 6.55: Simulated and measured space temperatures in 470 lift lobby, for the 2nd week of January and 2nd week of July 2017.

Office space 221A and many of the other office spaces show a similar trend in space temperatures between what is simulated and what is measured. Temperatures fluctuate during the day and night, however in office 221A there is a large difference between occupied and unoccupied. During the winter similar temperatures are achieved, for the summer however the simulated temperatures are significantly higher than measured. During occupied hours, the office is conditioned at 26 °C and increases to around 28 °C during unoccupied hours, in contrast the actual building shows that the spaces are conditioned, but not as strictly, with fluctuating temperatures between 23-26°C, slowly increasing during out-of-hours, but decreasing significantly during the night. The significant increase in simulated temperatures is likely due to the high internal gains in these spaces out-of-hours, which may have been assumed incorrectly. For lift lobby 470, differences are more drastic. During winter simulated space temperatures reach night-time heating set-point of 16 ^{0}C and are heated up to 24-25 ^{0}C during the day. In contrast, measured temperatures are heated slowly during the day to about the same temperature, then there is a significant surge in temperature after the systems shut down and a steady temperature for the rest of the night. The surge indicates a misalignment of the operational hours with actual occupancy, people are likely still present in the lift lobby after 7pm. During summer the measured temperature is stable around 25-26 ^{0}C, while the simulated temperature during occupied hours is 26 ^{0}C, out-of-hours temperature is significantly higher, at over 30 ^{0}C. Again indicating that internal gains or perhaps external gains in the simulated space are significantly different from the actual space, external, because the winter is exhibiting such low temperatures during the night. Although only two spaces are analysed, it gives a good reveals large differences between simulated and measured space temperatures, indicating that the underlying assumptions in the model are not in all facets representing the actual situation.

6.5.2 Uncertainty analysis

Total and predicted energy end-uses from 3000 simulations are compared to measured energy use in **Figure 6.56**. A large distribution exists for the predicted energy use as a total, inherited from the large variation of the servers, lighting and power energy use.

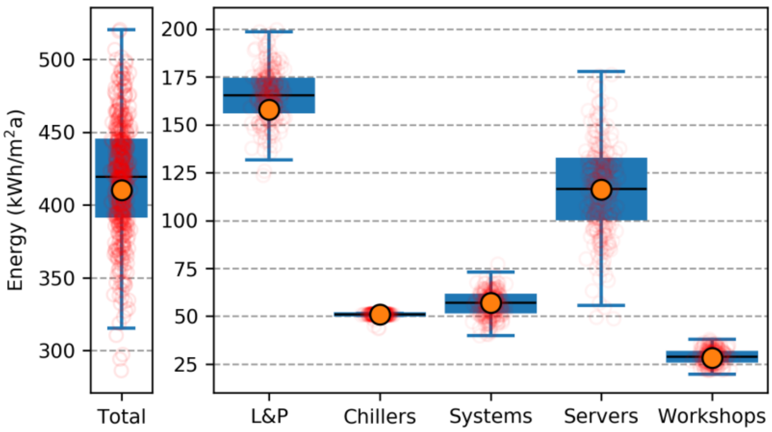


Figure 6.56: Total predicted (boxplots) and measured (orange dots) energy use for 3000 simulations.

Total predicted energy use for all runs is broken down per month to understand how seasonal variation is presented in the predictions, see **Figure 6.57**.

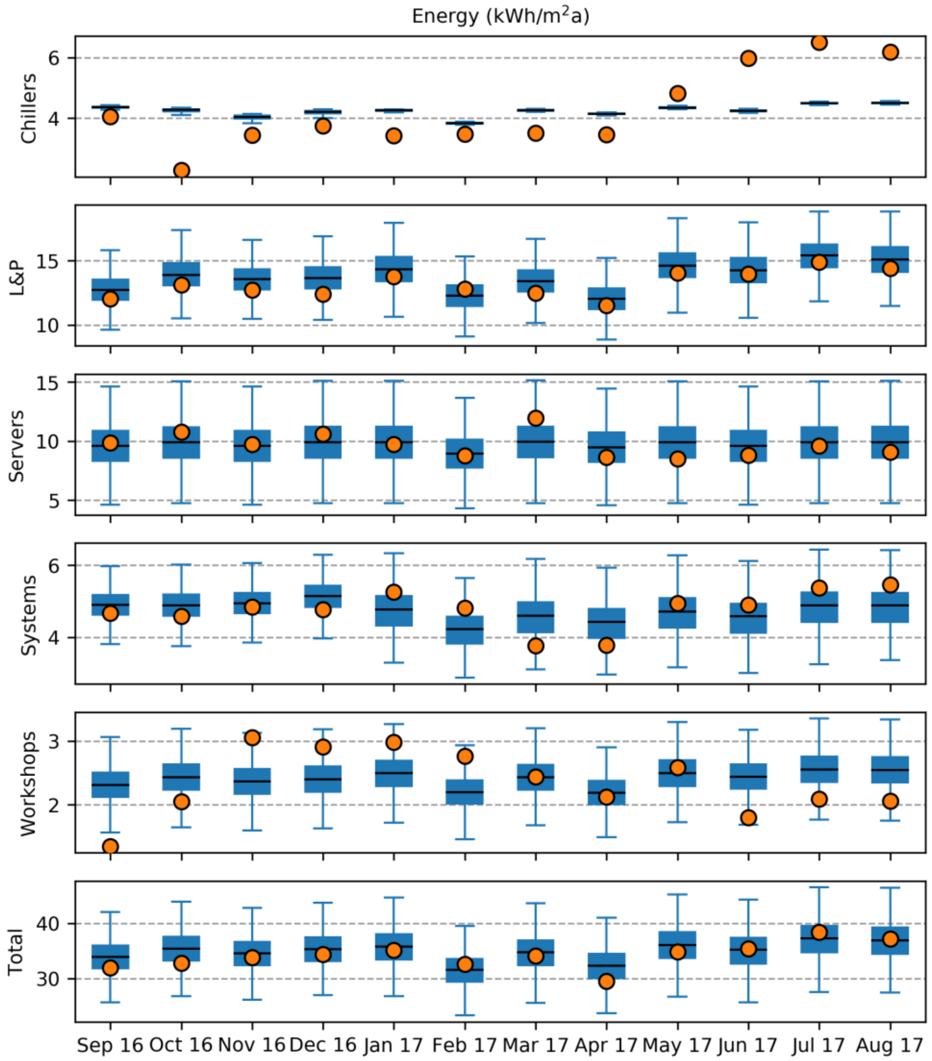


Figure 6.57: Monthly predicted (boxplots) and measured (orange dots) energy use for 3000 simulations.

For most months, predictions follow a similar pattern to measured energy use for the different end-uses, except for chiller energy use where measurement lie on the far sides of the spectrum. It was previously identified that although chiller electricity use was stable throughout the day, certain months showed a very sporadic profile, indicating that either more cooling was suddenly necessary or that some control measures were implemented to prevent the chiller from running 24/7. The graph does show that during the summer, chiller energy use goes up, but this behaviour was difficult to account for in the model.

6.5.3 Sensitivity analysis

The spearman correlation coefficients for 120 variable input parameters were computed. Most significant parameters, those with a coefficient of $\rho > 0.25$ and $\rho < -0.25$ are shown in the correlation matrix in **Figure 6.58**. Insignificant variables were:

- Equipment and lighting power densities in space types with relatively few spaces, lecture theatres, lift lobbies, meeting rooms, plant rooms, print rooms and reception, stairs and storage spaces.
- Natural ventilation $\rho < \pm 0.03$, its sensitivity is however dependent on three other variables, the temperature difference between inside and outside, the minimum indoor

and maximum outdoor temperatures. Which all three are varied, in contrast to CH, where these were assumed to be static.

- VRF heat pump CCOP and HCOP for the lift lobbies and server room G01.
- Seasonal weekend factor for all months ($\rho < \pm 0.05$), whereas the seasonal week factor ($\rho < \pm 0.25$) is much more significant as energy use during the week is much higher. These were varied for automated calibration purposes and adjust the monthly lighting, power and occupancy schedules.
- Activity level (i.e. metabolism of people) ($\rho = -0.02$ on chiller energy use), included to be variable.
- Cold water supply fixture flow rates ($\rho < \pm 0.1$), in contrast to hot water supply fixture flow rates, which have a significant impact on electricity use for domestic water heating ($0.62 < \rho < 0.75$).
- The weekend offset ($\rho < \pm 0.01$) on lighting and power schedules, which as was previously shown, has a very high baseload during the weekend throughout the year. Any horizontal change in the schedule will thus not affect energy use significantly. The weekday offset however is a bit stronger ($\rho < \pm 0.17$), but not as strong as in CH. This is because the profiles in MPEB have a very high baseload, so their relative effect on energy use is small.

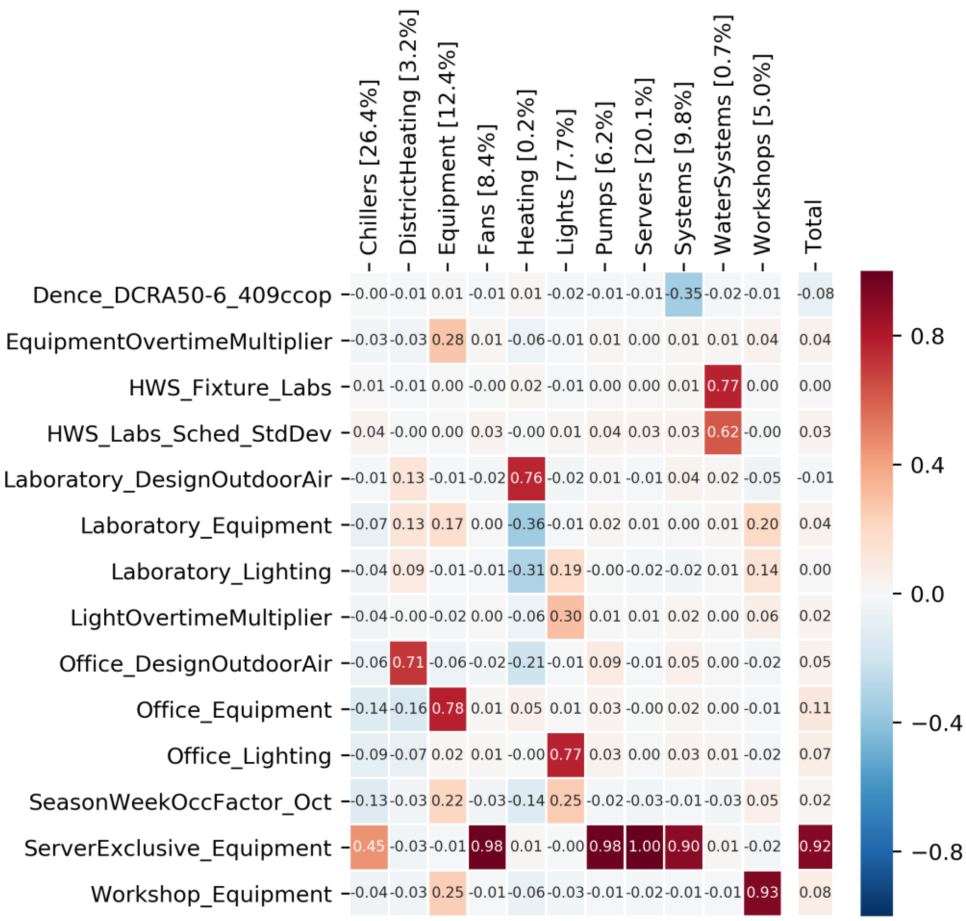


Figure 6.58: Spearman correlation coefficients per energy end-use for MPEB.

Significant parameters are mainly the equipment and lighting power densities in the offices, servers and laboratories. These in turn affect the coefficients of the VRF heat pump CCOP 'Dence_DCRA50' ($\rho = 0.35$ on systems energy use) conditioning the large server room and equipment ($\rho = 0.28$ on equipment energy use) and light (ρ = 0.30 on lights energy use) out-of-hours baseloads ('EquipmentOvertimeMultiplier' and 'LightOvertimeMultiplier' respectively). **Figure 6.59** shows the relation between the lights and equipment baseloads and lighting and equipment energy use respectively, varied at a standard deviation of 10% during parametric simulation.

Other significant parameters are the mechanical air flow rate in the offices and laboratory and hot water supply fixtures. Both of these and many of the other significant parameters are less important when considering the proportional effect on total energy use. The total sensitivity indices give a clear indication that energy use in the building is mainly driven by the large server rooms inside the building and indirectly their demands on the chillers, electric condensers and auxiliary equipment. Large energy savings could be made by reducing the server demand or efficiency of the system that conditioning the server spaces.

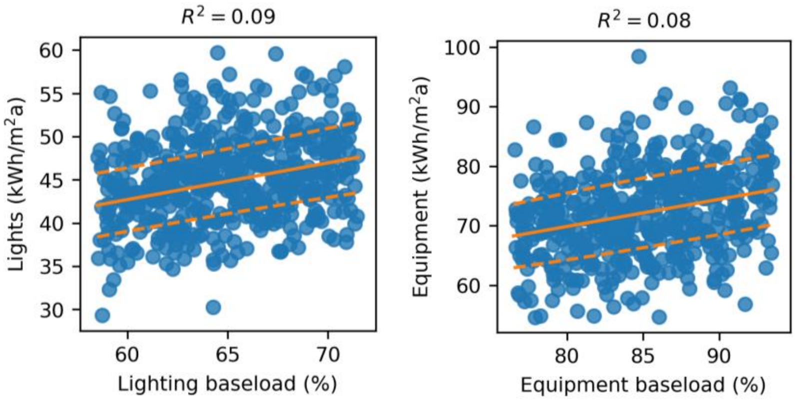


Figure 6.59: Correlation between lighting base load and lights energy use (left) and correlation between equipment base load and equipment energy use (right).

6.5.4 Impact of assumptions

The impact of typical (NCM) assumptions on the base case model have been assessed, the applied simplifications and their impact in percentage difference per energy end-use are given in **Table 6.9**.

Table 6.9: Effect of simplifications on the calibrated model as a percentage difference per yearly energy end-use.

	Simplification	Chillers	DH	Equipment	Fans	Heating	Lights	Pumps	Servers	WaterSystems	Workshops	Total
S0	No seasonality	-1	45	33	-3	-62	43	2	0	0	11	9
S1	Occupancy density	0	-1	0	0	-7	0	0	0	0	0	0
S2	Equipment power density	-6	14	-29	-44	29	0	-22	-50	1	-54	-25
S3	NCM schedule for equipment, lighting and occupancy	0	118	-33	-1	-2	-29	3	0	1	-37	-5
S4	NCM office H/C set points from 22/26°C to 22/24°C	0	91	0	-3	-16	0	3	0	0	-6	-1
S5	Infiltration (8 to 12 m³/m²h @ 50Pa)	0	6	0	0	5	0	0	0	0	0	0
S6	Combination S1 to S3	-6	2	-3	-46	-17	43	-21	-50	1	-44	-19
S7	Combination of S1 to S4	-7	141	-51	-45	13	-29	-18	-50	2	-69	-28
S8	Combination of S1 to S5	-9	133	-51	-39	3	-29	-15	-50	2	-69	-28
S9	H/C set-point server from 21/18°C to 23/20°C	0	95	0	-20	-16	0	0	0	0	-6	-3
S10	Server equipment power density from 1000 to 500 W/m²	-6	0	0	-44	0	0	-21	-50	0	0	-19
S11	Design weather	-1	116	0	-3	14	0	1	0	0	0	3

The impact of simplifications in annual energy use per floor area is given in **Figure 6.60**, where district heating (DH) is combined with electric heating under 'Heating'.

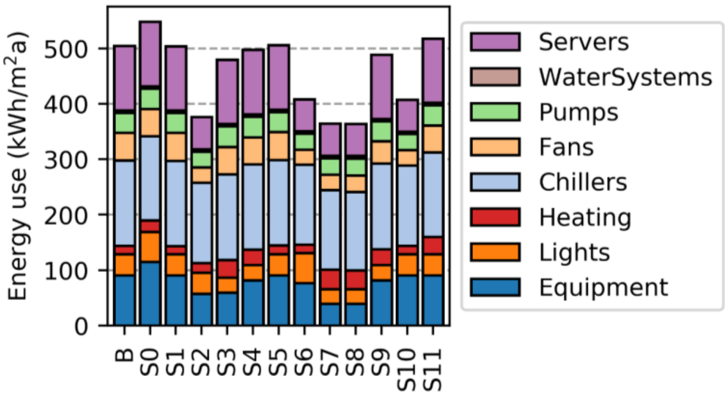


Figure 6.60: Energy use for the base case model with simplifications as numbered in **Table 6.9**

Similar to CH, a monthly seasonal variation was applied to the occupancy, equipment and lighting profiles to account for term-time variations in use. **Simplification 0** excludes this seasonal factor, which led to 9% increase in total energy, significantly smaller than the 21% in CH, this is because the computer clusters in MPEB are a dominant energy use, influencing also system energy consumption. This is evident from **Simplification 10,** where the base case model assumes an equipment power density of 1000 W/m², the NCM assumption of 500 W/m² impact not only reduces server energy use by 50%, but concurrently reduces auxiliary energy use and chiller energy use. **Simplification 2** reduces both the server load and takes into account NCM assumptions for equipment power density of the other space types, which are mostly smaller compared to base case assumptions based on an energy audit of these spaces. In particular, office spaces were 4 W/m² higher, workshops were assumed to be 55 W/m² instead of 6 W/m² and the library and lecture theatre were 8 W/m^2 and 10 W/m^2 respectively instead of 2 W/m^2. Although the high density IT spaces were determined to be somewhat lower than the 30 W/m^2 NCM assumption. These changes reduced total energy use by 25% only a 6% higher reduction compared to S10. **Simplification 3** is another indication that the servers have the most significant influence on energy use. Significant changes in the occupancy, equipment and lighting schedules reduced total energy use by only 5%. The profiles were applied to offices and similar space types as shown in **Figure 6.61**, whereas the servers are operated 24/7.

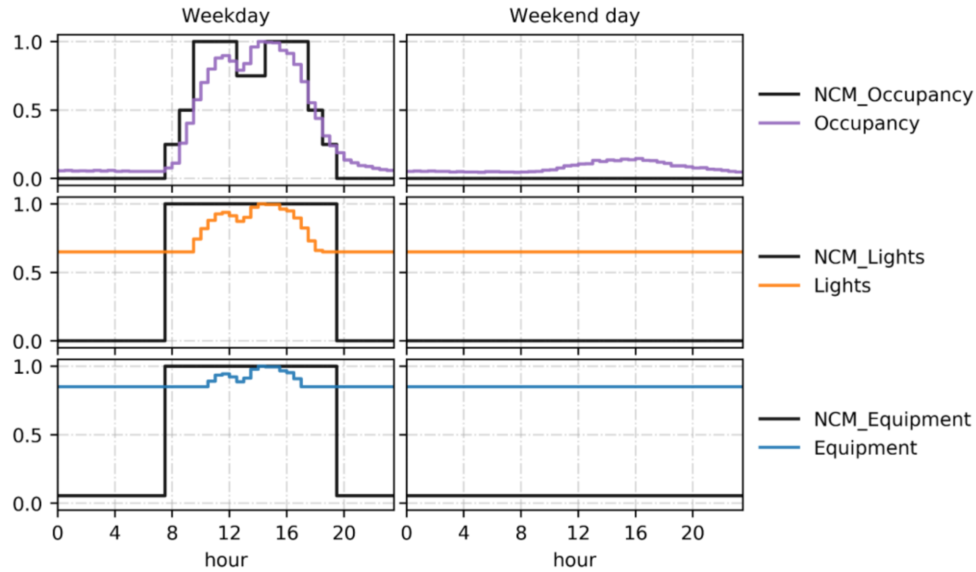


Figure 6.61: NCM schedules for occupancy, equipment and lighting and assumed base case profiles.

Simplification 9 changed the set-point temperature of the server rooms, from H: 21°C, C: 18°C to NCM assumption set-point temperatures H: 23°C, C: 21°C, which decreased energy use by 3%. **Simplification 11** compares the base case model using a weather file based on historical weather data and a design weather file, which increased total energy use by 3%. Simplifications 6, 7 and 8 look at combinations of the different simplifications, where S8, a combination of most NCM assumptions under predicts total energy use by 28%.

6.6 Summary

Manual calibration is an essential part of the research methodology. Most information on the physical aspects of a building can be found in the design specifications, typically collated in O&M manuals. Construction materials, systems and their operational strategy (although not always implemented), lighting- and other design specifications are described, but actual occupancy, the amount of small power equipment and when and how people interact with building systems and components retain much uncertainty in the modelling process. This was limited by utilising occupancy data from Wi-Fi and swipe card access, analysing operational performance and system characteristics and through building inspection, all of which are relatively time-consuming tasks. Furthermore, the use of manual calibration proved essential in understanding differences between predictions made by the model and measurements. Many solutions or configurations of input parameters or scheduling of operational processes can accurately fit the final measurements, but with an increasing level of data granularity, the wrong solutions become filtered out. The lack of measurement data for some of the end-uses in the case study buildings proved to limit the certainty of choices made in the calibration process. **Table 6.10** gives an overview of the observation and lessons learned, while Table 6.11 gives an overview of issues and limitations identified during development of the case study buildings and manual calibration that mitigated the discrepancy between predicted and measured energy use. Electrical hot water heating could not be distinguished from other power energy use, but is predicted to be a large contributor to energy use.

Table 6.10: Observations and lessons learned during manual calibration of the four case study buildings.

Office 17	Office 71	CH	MPEB
• Extensive amount of gas use compared to what was initially predicted due to night-time use. The FM put the boiler on a timer in later years. • High underestimation of equipment energy compared to NCM assumptions (54% lower than design stage assumption) • NCM simplifications introduced to the building energy model significantly underestimated measured energy use by 57%.	• Manual operation of system, such as turning off the boiler even though the model would still use some for radiator heating. • Replicating lighting and power schedules of use based on measured electricity gives a very close fit, closer than when using Wi-Fi data. • Weekend occupancy is very intermittent and difficult to take into account in the model, based on typical schedules for the whole year or per month. • Typical day trends and baseloads of power energy use, in contrast to lighting, was very different per floor. Indicating the importance of spatial granularity.	• Air handling unit observed to be out of order, which could have a large impact on energy use. • Continuous system energy use throughout the week from VRF heat pumps for air conditioning, indicating inefficient system. Replicating this behaviour required high heating set point temperatures and low cooling set point temperatures. • Holidays (i.e. days of vacancy and low energy use) in university buildings are based on semesters. • Operated 7 days a week, with minor differences in energy use between weekday and weekends	• High power and lighting base loads identified, server to count for largest proportion of energy use in the building. • Correctly establishing the H&C loads, uncertain due to lighting and power equipment assumptions is important as they affect other end-uses. Measured L&P electricity is essential, ideally on a per floor basis. • Holidays (i.e. days of vacancy and low energy use) in university buildings are based on semesters. • Operated 7 days a week, with minor differences in energy use between weekday and weekends.

Table 6.11: Identified issues and limitations during manual calibration of the four cast study buildings.

Office 17	Office 71	CH	MPEB
• No design specification available on material properties, thus based on previous assessment and audits.	• Systems energy use was excluded from the metering system for the VRF heat pumps. • No design specification available on material properties, thus based on previous assessment and audits.	• Heating and cooling energy use could not be separated as the VRF system provides both electrically, which is measured on a single meter. • Labels for L&P meters were unclear and often could not be distinguished.	• Server energy was connected to lighting and power bus bars on upper floors, which made accurately disaggregating these loads impossible. • District heating meter was broken during measurement period.

In all case study buildings, sub-metered data at a disaggregated level was essential to carry out the manual calibration process, but often the different meters were labelled incorrectly or insufficiently, did not cover certain important energy end-uses.

Collected operational data and energy audits of the existing buildings informed input assumptions of the building energy models, which were then iteratively adjusted to represent measured energy use. Final discrepancies for the manually calibrated models are shown in Table 6.12.

Table 6.12: Discrepancies between predictions and measurements

	Office 17	Office 71	CH	MPEB
NMBE (%)	-1.4	1.7	-2.9	2
$CV(RMSE)_{monthly}$ (%)	4.3	12.1	11	4.5
$CV(RMSE)_{hourly}$ (%)	67.8	66.8	28.8	11.2

The manually calibrated models were subsequently used to quantify the impact of underlying causes between predictions and measurement in particular concerning regulatory assumptions in building energy simulation. This was supported by assessing the sensitivity of input parameters on different energy end-uses and total energy use, a summary of the Spearman rank correlation coefficient on total energy use for the case study buildings is given in **Table 6.13**. In particular, equipment power density is a significant parameter to influence energy use, which is typical for non-domestic buildings, it is therefore of significant importance to make evidence based assumptions about these loads to make sure predictions are in line with measurements. For MPEB specifically, the large server rooms contribute to a significant proportion of total energy use and are therefore the main driver of energy use within the building as they indirectly affect system energy use. Other important factors are the heating and cooling set points, in particular in spaces with high internal gains, where temperatures can strongly fluctuate. Insignificant parameters to influence energy use are material properties, in contrast to domestic buildings, for non-domestic buildings there is a strong trade-off between the thermal performance of the envelope. An increase of the conductivity of the envelope can have a negative effect on energy use due to high internal gains in certain spaces, which inhibit heat loss to the outside and therefore have an increased cooling load. In high-density workspace in modern buildings, Passive house standard envelopes are unlikely to be a good design decision. Similarly, to the conductivity of materials, infiltration does not have a significant effect on energy use in any of the buildings, although there is a distinct difference between the seasons.

Table 6.13: Spearman correlation coefficients of the input parameters for the four case study buildings

	Office 17	Office 71	CH	MPEB
$\rho > \pm 0.75$	office equipment (W/m²)	office mechanical vent flow rate,	office lighting and equipment (W/m²), temperature difference in heating and cooling SP	server, office and workshop equipment (W/m²), office lighting
$\pm 0.75 > \rho$ $\rho > \pm 0.50$	heating SP	fixture flow rates for hot water, boiler hot water temperature, L&P (W/m²) in offices	flow rates of hot water fixtures, infiltration rate, server equipment, boiler efficiency	flow rates of hot water fixtures,
$\pm 0.50 > \rho$ $\rho > \pm 0.25$	natural ventilation air flow, lighting power density	boiler efficiencies, seasonal weekday factor	lighting and power offset, heating SP,	lab L&P, L&P base load modifiers, CCOP of server VRF
$p < \pm 0.25$	material properties, infiltration, system efficiencies,	L&P densities in space types with few spaces, VRF COPs, exhaust fan efficiency, natural ventilation rate	VRF CCOP and HCOP, L&P power densities in non-office space, natural ventilation air flow	L&P power densities in space types with only a few spaces, natural ventilation air flow, metabolism

For each case study buildings, several simplifications were applied to the manually calibrated models to quantify the effect of regulatory assumptions on predictions, as shown in **Table 6.14.** The typical NCM assumptions replaced calibrated input parameter values, in particular the significant parameters identified through sensitivity analysis. Certain parameters sets could not be directly included in the sensitivity analysis due to their non-stochastic representation during the Monte Carlo simulations. These are the seasonal variation factors, schedules (H/C, internal gains), and internal gain baseloads, which were also included as simplifications.

Table 6.14: Effect of simplifications on calibrated model in percentage difference on total energy use.

	Simplification	17	71	CH	MPEB
S0	No seasonality, the seasonal variation applied equipment, occupancy and lighting profiles for CH and MPEB are assumed to be unity for each month.	x	x	33.7	8.6
S1	Equipment power density for different space types is based on NCM values instead of based on observation.	-37.0	x	26.7	-25.4
S2	Combination of S0 and S1.	x	-5.8	58.0	-19.0
S3	Typical NCM occupancy, lighting and equipment profiles were used instead of those based on WiFi or sub-metering profiles.	-21.3	-19.9	-4.2	-4.9
S4	NCM heating and cooling profiles were used for the space types instead of those based on O&M manuals or where changed for calibration.	-10.0	-0.1	-13.7	-1.4
S5	Infitration was adjusted to more conservative values, increasing for CH and 71, and decreasing for MPEB as an airtighness test showed a value of 8 m^3/m^2h @ 50Pa.	x	0.1	-1.5	0.2
S6	Combination of S1 and S3	x	-23.1	18.5	-27.8
S7	Combination of S0, S3 and S4	-57.0	-23.1	27.5	-27.8
S8	Heating and cooling set-point temperatures in server room from (21, 18) to (23, 20).	x	x	x	-3.1
S9	Decreased power density in the servers from 1000 W/m^2 to 500 W/m^2, typical assumptions for computer clusters.	x	x	x	-19.2
S10	Used design weather file instead of weather file based actual weather data.	x	x	1	3

The simplifications concerned with a change of internal gains highlight the importance of accurately determining the assumptions for equipment and lighting power density in spaces, they have a large effect on total energy use, supported by the strong correlation coefficients for these parameters. In particular, server loads can be dominant in modern buildings, which was evident for MPEB, energy use in the building was primarily driven by the power use of the computer clusters, indirectly influencing systems energy use. Moreover, defining the right schedules for these loads is tantamount to establishing the right loads for spaces, as their effect on energy use was similar in the case study buildings, as shown by S1 and S3. As part of the manual calibration, it was found that high baseloads existed in the buildings, which had to be accounted for by adjusting the internal gains schedules. These were a large contributor to a discrepancy between regulatory predictions and measurements as there was a significant difference to the typically assumed NCM equipment power baseload. In the case study buildings, the baseloads for equipment in Office 17, 71, CH and MPEB were ~25%, ~20%, ~65%, ~85% respectively, compared to the NCM assumption of 5.3%. Besides internal gains, the heating and cooling temperatures in different space types can vary significantly from that initially assumed to that in operation, something that is difficult to replicate within a model when a variable strategy is in place. In CH, the operational set-point temperatures could be adjusted manually and this was difficult to replicate, especially since system energy use for conditioning was found to be constant in both CH from the VRF system and MPEB from the chillers. Replacing the calibrated set-point temperatures with NCM assumptions led to a significant decrease in energy use for Office 17 and CH, while for Office 71 the temperatures were similar to NCM assumptions. Finally, as simplification 10, the weather file based on historical weather data in London was replaced by a design weather file from Gatwick, which had a minor, but notable effect in increasing total energy use.

7 QUANTIFYING THE EFFECTS OF DATA GRANULARITY ON MODEL ACCURACY

Base case models for three of the four case study buildings were used to develop meta-models for automated calibration, trained on thousands of inputs and outputs generated using parametric simulations. These models were used to predict new sets of input parameters at different levels of data granularity, (M1) yearly energy end-uses, (M2) monthly energy use, (M3), monthly energy end-uses and (M4) monthly energy end-uses including a typical weekday and weekend day. Several machine-learning algorithms were tested and compared to achieve the highest level of accuracy, in particular partial least squares regression and artificial neural networks performed well and were used throughout the development. Meta-models that predicted monthly energy end-uses achieved r-squared values of 0.96, 0.97 and 0.84 for Office 71, CH and MPEB respectively. Meta-models were used during mathematical multi-objective optimisation to efficiently predict the objectives they were trained for (levels of data granularity) based on new sets of input parameters. A genetic algorithm (NSGA-II) was utilised to minimise differences between predictions and measurements. Predictions made by meta-models M1 and M2 and measurements were efficiently minimised for 5-12 different objectives, M3 and M4 however involved optimising 32 to 300 objectives and limitations in the optimisation process were identified. Optimised individuals (calibrated sets of parameters) were simulated using EnergyPlus and compared to the meta-models, base case and measurements. Model error arises due to the slight inaccuracy of the meta-models when translating inputs back to the building energy software. Meta-model multi-objective optimisation significantly reduced differences between predictions and measurements, however the model error when using optimised input parameters in EnergyPlus nullified these improvements.

7.1 Introduction

Automated calibration was utilised with the aim to improve the calibration accuracy of the base case models. However, as shown in the previous chapter, the base case models were already 'calibrated' using manual calibration, the iterative adjustment of the models based on observation and comparing predicted and measured energy use. In the calibration methodology, these activities are represented by the white and green shaded boxes shown in **Figure 7.1**.

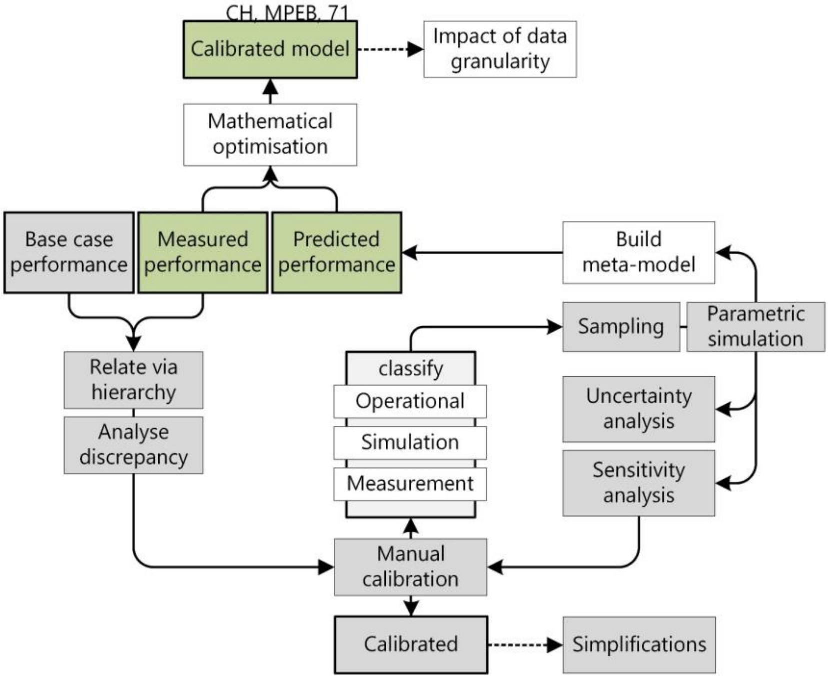


Figure 7.1: Calibration methodology, activities in grey are not discussed in this chapter.

The purpose of using automated calibration is to investigate its potential on improving the efficiency of the whole calibration process for full-scale building energy models, by alleviating some of the manual calibration tasks. In addition, it investigated and quantified the effectiveness in achieving increased model accuracy by decreasing the discrepancy between predicted and measured energy use, but found several limitations using automation at higher levels different levels of data granularity. Furthermore, several risks associated with using automated calibration at lower levels of granularity were identified and limitations concerning calibration of full scale models is discussed. Automated calibration has been applied at different levels of data granularity to the models of three case study buildings (Office 71, CH and MPEB). This approach was however not applied to Office 17, as it was initially a pilot study to inform the other studies and ultimately would unlikely have provided additional insights into the effects of data granularity on model calibration accuracy. The process includes the development of meta-models or surrogate models in order to enable optimisation, which would otherwise be limited due to computational intensity of full-scale first-principle building models.

Meta-models

Meta-models were created by training machine learning algorithms on varied inputs from parametric simulation of the building energy model and simulated outputs. The type of machine learning algorithm and amount of data for both the number of simulations performed and number of inputs and outputs determine the accuracy of the meta-model. More simulations create more training data, typically, this increases meta-model accuracy in predicting new values. However, an increase in the number of inputs and outputs increases complexity, making it more difficult for the learning algorithm to find relations within the data. There is thus a trade-off. Parametric simulation data (inputs & outputs) were split by 75% for training and 25% for testing, after training, the meta-model was used to predict new values and its accuracy is determined by comparing those predictions to the test data, which the model has not seen previously. The r-square, mean absolute error, average absolute percentage difference were used to quantify their accuracy. The main advantage of the meta-models is that they can quickly, less than a fraction of a second, make new predictions for what they were trained to predict.

Mathematical optimisation

The meta-models were subsequently deployed for automated calibration or mathematical optimisation, more specifically, using NSGA-II, a genetic algorithm for multi-objective optimisation. The objective here is to minimise the difference between predictions from the meta-models and measurement of energy use of existing buildings. The algorithm uses an evolutionary process that iteratively adjusts input parameters and finds those that improve the minimisation of objectives. It sorts a population (sets of input parameters) based on the performance of individuals (set of parameters in the population), an individual is calculated using a meta-model. Different meta-models are optimised, resulting in different calibrated models. The calibrated models are taken to be the best individuals (set of input parameters) per meta-model. The terms input space and solution space refer to the exhaustive set of input parameter combinations and their predicted outputs.

7.2 Office 71

7.2.1 Meta-model development

The meta-models were trained on the inputs and outputs from 1000 parametric simulations. In total, 115 input parameters were varied, including conductivity of building materials, system efficiencies, power densities for lighting and appliances and people density in different space types, parameters that determine when windows are opened, DHW hot water use, natural, mechanical and unwanted infiltration rates and a seasonal variation factor per month. Three meta-models were created by training them on the variable input parameters and energy use outputs. These three models predict energy use at different levels of data granularity as follows:

- Meta-model M1: Yearly end-uses (3 objectives: lights, power and gas).
- Meta-model M2: Monthly energy use (12 objectives: January 2014 to December 2014).
- Meta-model M3: Monthly energy end-uses (36 objectives: 3 end-uses per 12 months).

Several machine learning algorithms were tested and compared using the MAE and r-squared, a high r-squared value indicates a well fitted model. Different regression techniques; ridge (rr), Bayesian (BR), Lasso, linear (LR) and others perform similarly well. In **Figure 7.2** the accuracy in predicting monthly energy end-uses by four techniques is compared, Partial Least Squares regression (PLS), Ridge regression (RR), Lasso regression (Lasso) and the Artificial Neural Network (NN) perform similarly well, achieving near perfect r-square and considerably low errors. Other algorithms are available, such as Radial Basis Functions and Support Vector Regression, these have however not been explored further as a sufficient accuracy was obtained with employed techniques. Practically, only meta-model M3 needs to be trained as it is able to predict the highest level of data granularity, lower levels can then be inferred from this meta-model, however to show how a difference in the level of data granularity affects meta-model accuracy, the other meta-models were trained as well.

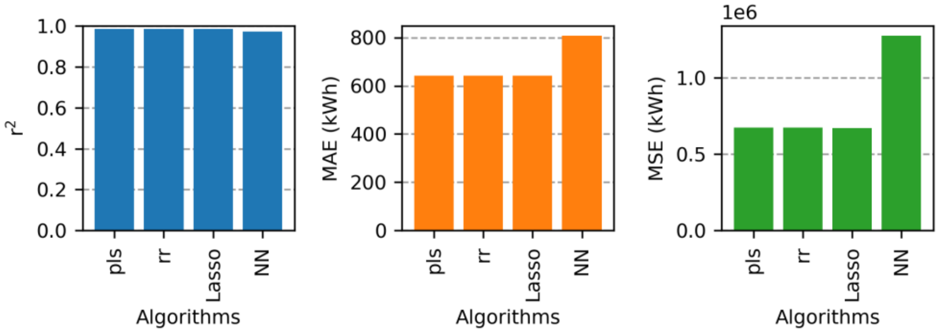


Figure 7.2: Machine learning algorithms and their respective scores; r-squared, mean absolute error (MAE) and mean square error (MSE) to evaluate their predictive accuracy in regards to training on monthly energy end-uses (M3).

In addition to their final accuracy, the accuracy of the meta-model is evaluated by looking at how they change when more model runs are included. For both methods, if the indicators are not changing or change is significantly small, then they are determined to be accurate and used in further analysis. The learning progression of training a meta-model that predicts monthly energy end-uses for Office 71 is shown in **Figure 7.3**.

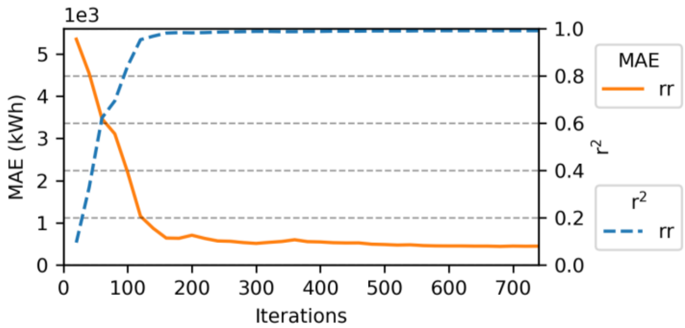


Figure 7.3: Learning progression of meta-model M3 using PLS, trained on monthly energy end-uses from 1000 simulations, showing statistical measures of the mean absolute error and r-squared.

A rapid increase in prediction accuracy is attained after only a few hundred simulations. These statistical measures indicate the accuracy of the meta-model predictions compared to what they are supposed to predict, the actual simulated data from the first-principle model. This comparison is visualised in **Figure 7.4**, where mean and standard deviation predicted monthly energy use from the meta-model is compared with that from simulations for a model trained on 75 (left) and 1000 (right) simulations. This visualises how effective the meta-model is in replicating the predictions of the first-principle model, it learns quickly because the data set is small, there are only 36 values (12 months x 3 end-uses) predicted.

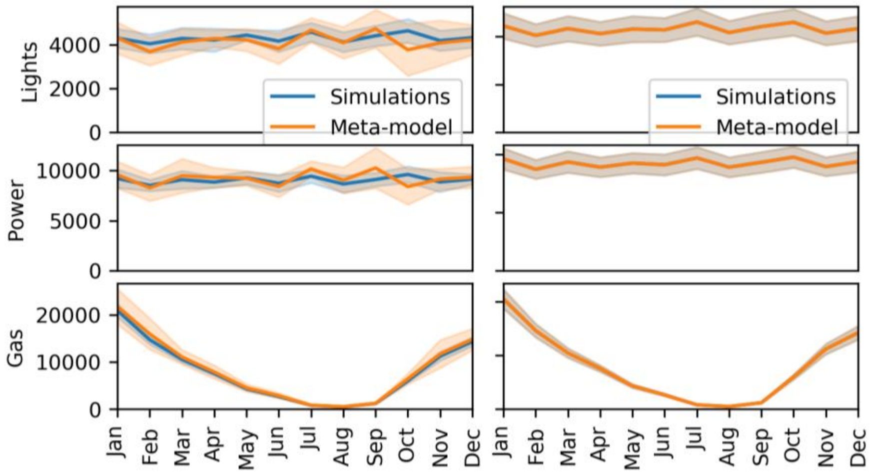


Figure 7.4: Simulation test data vs. meta-model predictions after 75 (left) and 1000 (right) simulations after training, showing the mean and standard deviation of the 25 and 500 predictions respectively from the test data.

The three constructed meta-models were used to run global sensitivity analysis, which needs many inputs and outputs variations. Subsequently, the meta-models were used for mathematical optimisation.

7.2.2 Mathematical optimisation

Optimisation is performed (i.e. automated calibration) by minimising the RMSE between predictions from the meta-model and measurements. **Figure 7.5** shows the minimisation of predicted yearly energy use for lights, power and gas. After only several generations, the genetic algorithm finds combinations of inputs that bring the predictions closer to measurements. After about 30 generations, the optimisation converged to a solution, i.e. it found an individual that predicts exactly the same amount of energy for the three different end-uses, in fact there will be many sets of input parameters (individuals) that will solve this particular problem.

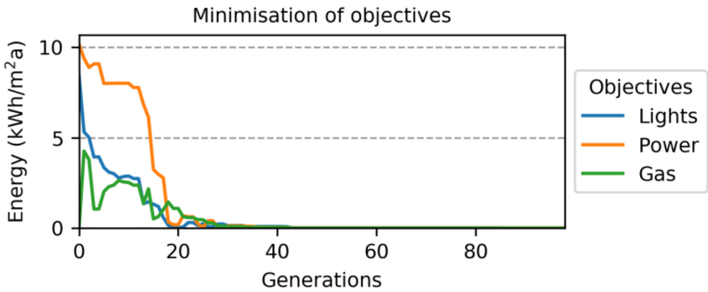


Figure 7.5: Minimisation of absolute difference between meta-model predicted and measured yearly energy end-uses per generation for Office 71.

However, finding solutions becomes more challenging with an increase in objectives. For example, optimising for total monthly energy use, shown in **Figure 7.6**, takes more generations, but is still manageable as the measurements fall closely within the solution space predicted by the first-principle model.

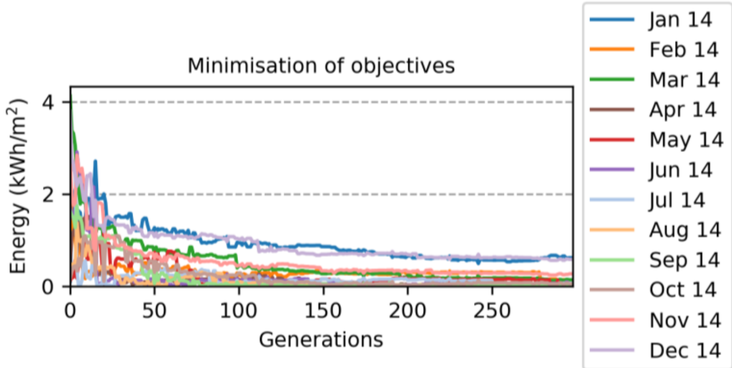


Figure 7.6: Minimisation of absolute difference between meta-model predicted and measured total monthly energy per generation for Office 71.

In contrast, when total monthly energy use is broken down, some of the measured data fall outside of this solution space, and the optimisation algorithm is unable to converge for each objective, as shown in **Figure 7.7**. The graph shows the absolute difference between predictions and measurements per energy end-use per floor per month, although converging towards zero, some of the objectives are unable to decrease after several generations or do not decrease at all. This is because the individuals or input parameter values are constrained to lower and upper limits. Without these constrains the algorithm would adjust certain variables to unrealistic values, for example, boiler efficiency (COP) could be set to 5, to more closely resemble systems energy use.

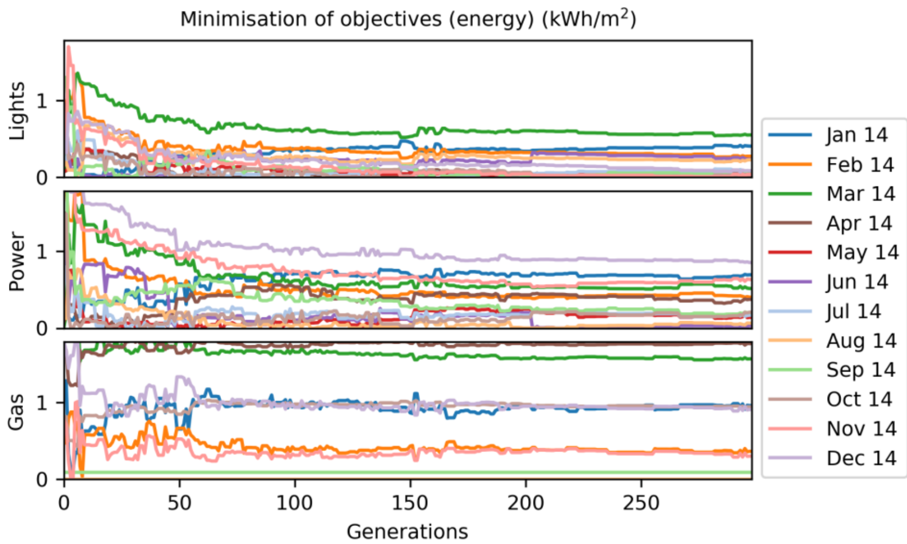
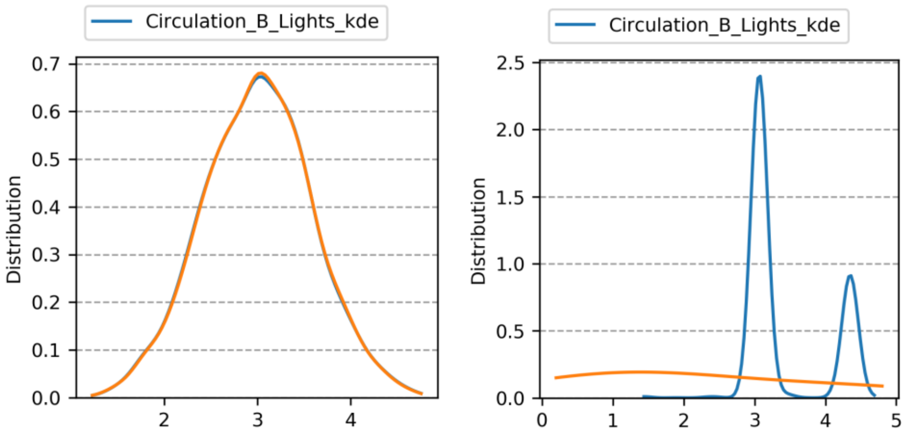


Figure 7.7: Minimisation of the absolute difference between predicted and measured monthly energy use per end-use for Office 71.

The meta-model alone predicts monthly energy use for a certain solution space, where many solutions fit the NMBE and CVRSME criteria of $\pm 5\%$ and $<30\%$ respectively. However, without optimisation, solutions that fit more stringent criteria are not found. The input parameters follow a normal distribution, as a result, so do the meta-model predictions, therefore it is unable to find combinations of parameters which lie at the ends of the normal distribution. This is why optimisation is used, it is able to find these solutions by iteratively selecting better fitting input parameters that converge towards possible solutions. The optimisation was able to find solutions that fit very strict criteria (NMBE $\pm 1\%$ and CV(RMSE) $< 1\%$). Filtering input variables that fit these criteria shows in what range the calibrated inputs are distributed. In **Figure 7.8**, the input distribution for an arbitrary input parameter, basement circulation lighting power density, is shown based on the actual search space (blue) and those put into the meta-model (orange) left, and the input parameter distribution after optimising the model on the right. The meta-model inputs follow a normal distribution, when filtering solution fit a criteria of CV(RMSE) <30%, it seems that for this particular variable any of the values within the normal distribution are possible solutions. If the criteria for filtering solutions is increased to CV(RMSE) <1%, then only the optimisation algorithm will be able to find solutions, for this particular input parameter it finds values at the high end of the distribution, at around 3 and 4.3 W/m^2. Due to the interrelation between input parameters on the output, different values are obtained when the range of other variables is adjusted or limited. For example, the basement lighting power density directly influences lighting energy use, when other input parameters that affect lighting energy use are limited to a smaller distribution space, the values for power density in the basement are likely to shift in contrast with the previous solution. It is therefore important to limit input parameters to a realistic range.



(meta-model in orange, search space in blue)

Figure 7.8: Meta-model predictions and single input (Basement circulation lighting power density) distribution that fit CV(RMSE) <30% (left) and CV(RMSE) <1% (right) using the optimisation algorithm for Office 71.

The optimisation algorithm looks through the search space to find sets of input parameters that minimise the objective function and finds many combinations that form calibrated models. **Table 7.1** shows the most significant (according to sensitivity analysis) parameters before and after automated calibrations for Office 71. Some of the input parameters have taken up much higher values than the default one determined for the base case. In particular, Office_Equipment power density is higher, even though the base case typical profile for the weekday and weekend day for power energy use was shown to be a good representation. The increase is however due to a decrease in the SeasonalWeekOccFactor, which directly influence the equipment schedules.

The higher levels of data granularity calibration (M2 and M3) used the seasonal factors to minimise the difference between monthly energy use and monthly energy end-uses respectively. Whereas M1 which does not concern itself with the monthly variation, instead it can adjust these parameters to justify an optimisation towards yearly energy use. Which in turn is likely to reduce the actual representation of reality even though the energy use predictions on a yearly basis converge towards the measurements.

Table 7.1: Most significant parameter values before (base case B) and after automated calibrations optimisation using meta-models (M1 to M3) for Office 71.

Parameter	B	M1	M2	M3
Office_1_Lights (W/m^2)	8	9.52	12.67	12.28
Office_2_Lights (W/m^2)	8	9.77	11.18	11.50
Office_3_Lights (W/m^2)	8	11.31	6.72	10.87
Office_1_Equipment (W/m^2)	18	29.49	21.96	19.15
Office_2_Equipment (W/m^2)	18	22.29	31.46	26.47
Office_3_Equipment (W/m^2)	14	14.87	24.17	26.23
Fixture_Kitchen (m^3/s)	5.56E-06	8.69E-06	4.22E-06	7.37E-06
HW_Boiler_1 (eff 0-1)	0.86	0.98	0.75	0.73
HW_Boiler_2 (eff 0-1)	0.86	0.75	0.751	0.81
Office_DesignOutdoorAir (l/s)	8	7.47	11.4	9.88
SeasonWeekOccFactor_Jan (0-1)	0.75	0.59	.517	0.85
SeasonWeekOccFactor_Feb (0-1)	0.75	0.61	0.71	0.68
SeasonWeekOccFactor_Nov (0-1)	0.75	0.97	0.97	0.99
SeasonWeekOccFactor_Dec (0-1)	0.75	0.50	0.94	0.76

Meta-models are limited to making predictions they were trained for, this means that the meta-models cannot predict hourly energy use variation, typical in building energy software. The calibrated models (M1 to M3), i.e. best performing sets of input parameters are saved and used in the first principle physics based software (EnergyPlus), using the base case model, but changing the optimised input parameters. These were then used to compare their predictions at an hourly level, these models are called with respect to their meta-models, C1, C2 and C3. Prediction results between the meta-models (M1 to M3) and building energy modelling predictions (C1 to C3) using the same input parameter are compared in **Table 7.2.**

Table 7.2: Model error showing difference in kWh/m^{2}a and % between meta-model and first-principle predictions for identical input parameter values.

	Total M2-C2		**Lighting M3-C3**		**Power M3-C3**		**Gas M3-C3**	
	Diff	%	Diff	%	Diff	%	Diff	%
Jan	1.56	-8.09	-0.04	1.29	-0.05	-0.64	-0.29	-2.52
Feb	0.05	-0.30	0.11	4.05	0.21	3.63	0.24	2.57
Mar	-0.17	0.95	0.29	12.77	0.57	10.78	0.10	1.31
Apr	0.86	-7.46	0.14	5.44	0.23	3.92	0.04	0.89
May	0.58	-7.11	0.18	6.48	0.34	5.80	0.00	0.00
Jun	0.42	-4.87	0.08	2.82	0.16	2.59	0.00	0.00
Jul	0.49	-5.43	0.12	3.97	0.16	2.53	0.00	0.00
Aug	0.53	-6.71	0.14	5.30	0.27	4.86	0.00	0.00
Sep	0.53	-6.14	0.14	5.00	0.20	3.22	0.00	0.00
Oct	0.72	-5.80	0.05	1.42	0.07	1.06	0.08	2.18
Nov	-0.51	2.92	-0.08	-2.37	-0.08	-1.19	0.08	1.25
Dec	-0.64	3.34	0.04	1.42	0.14	2.23	0.00	0.06
Sum	4.43	-2.80	1.17	3.35	2.23	2.99	0.26	0.50
M1 - C1			0.29	-0.83	0.42	-0.57	0.76	-1.48

Predictions are slightly different due to the meta-model training error (not giving exact predictions). Although the differences are marginal in most cases, they do affect the results, the optimised solutions found by the meta-model are less 'optimised' in the first-principle model. Their actual effect on results is analysed by comparing measured energy use compared with the different predictions from first-principle models (C1-C3) and the base case (B) predictions from the initial manual calibration.

7.2.3 Impact of data granularity

With two sets of models, the meta-models (M1 to M3), limited to monthly energy end-uses and first-principle model (C1 to C3), their performance can be compared to the initial base case and the measurements to see if the automated calibration process has found better calibrated models. In **Figure 7.9**, the hourly power energy use during October to December is shown for the calibrated model (C3), the calibrated model is significantly closer to the measured power energy use than the base case model. However, the calibration process is limited to changes in seasonal variation and power densities, which make up the power energy use. Therefore, the calibration will not be able minimise weekly differences, significant in December. Furthermore, in the case of Office 71, no horizontal offset is taken into account in the equipment schedules, thus the schedule can only be moved up and down the y-axis. Including weekly or even daily variability in the schedules could further improve accuracy. The automated calibration was able to improve energy use not only on a monthly basis, but also at an hourly granularity.

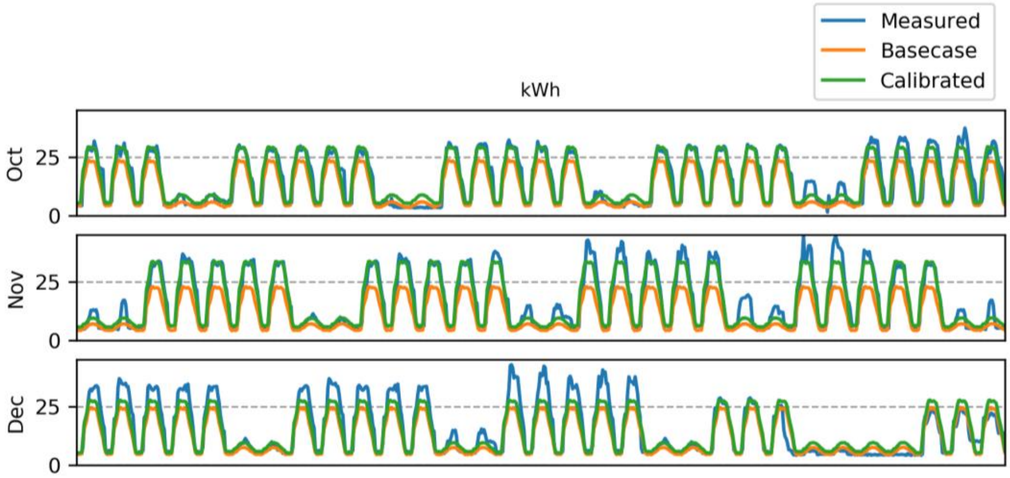


Figure 7.9: Hourly power energy use during Oct-Dec for the calibrated model (C3) and base case model (B) against measured power energy use in Office 71.

The calibrated models for the different meta-models and their first-principle model variants are shown in **Figure 7.10**, where the absolute difference in total energy use is compared with measurements. The graph shows that calibrated meta-model M1 is nearly 100% accurate, i.e. there is only a marginal difference between its predictions and the actual measurements. Meta-model (M3), which predicts monthly energy end-uses is considerably accurate and predicts within an NMBE of 2.5% for the yearly end-uses. Using the same calibrated inputs, the first-principle models (C1 to C3) were simulated, showing accuracy is somewhat reduced. C1 is now under predicting the measurements with NMBE 2.5%, whereas C2 and C3 show large margins of error. However, this does not necessarily mean that they are worse performing models, in fact, C2 and C3 are better at predicting monthly energy use and monthly energy end-uses respectively, which is evident from **Figure 7.10** and **Figure 7.11**.

Total monthly energy use predictions from the first-principle models (C1 to C3), meta-models (M2 and M3) and base case model are compared with measured data in **Figure 7.10**. The meta-model M2, specifically optimised to predict measured total monthly energy use shows the best fit, whereas C2 its first-principle variant is slightly less accurate. However, C1 in this case performs worse than the other models, as was hypothesised earlier. M3 and C3 perform similarly, with the base case monthly energy use predictions being closer to reality than both M3 and C3.

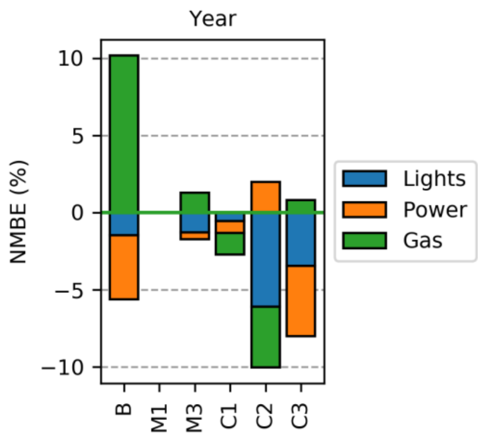


Figure 7.10: NMBE between total yearly energy end-uses from two meta-models (M1 and M3), first-principle models (C1 to C3), base case model (B) against measurements.

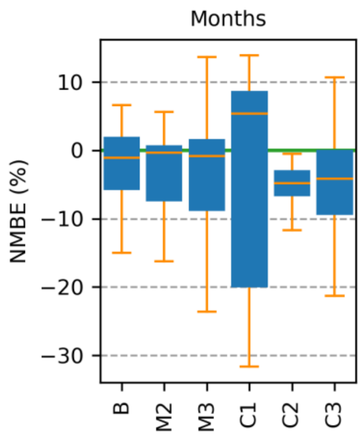


Figure 7.11: NMBE between monthly predictions from the two meta-models (M3 and M2), first-principle models (C1 to C3), base case model (B) against measurements, each boxplot representing 12 data points for the monthly CV(RMSE).

Finally, the normalised mean bias errors for three energy end-uses on a monthly basis were calculated, shown in **Figure 7.12.** Only M3 can be compared because it is specifically trained for predicting monthly energy end-uses.

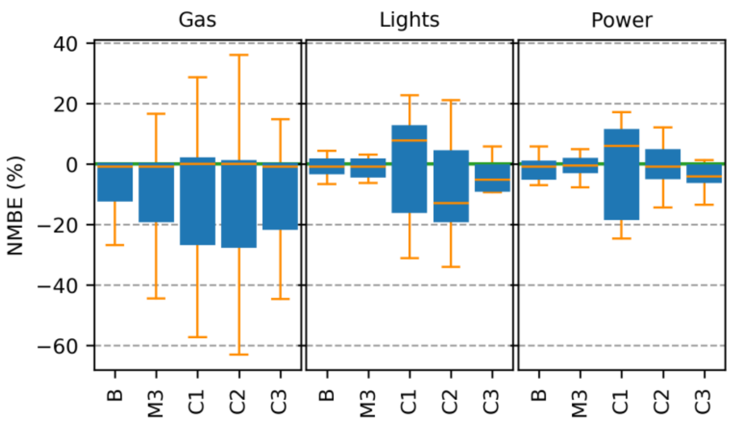


Figure 7.12: NMBE between monthly predictions from the meta-model (M3), first-principle models (C1-C3) and base case model (B) against measurements (zero-line), each boxplot representing 12 data points for the months.

Meta-model M3 performs well, it has small NMBE errors, however, the base case model is showing the same levels of error. C3, due to the model error, performs somewhat worse than M3, differences between monthly predictions and measurements is increased. The results also show how an increase in data granularity improves the accuracy of the model, this can be seen when comparing C3 with C1 and C2, which both have a larger prediction error as these models were not specifically calibrated at the same level of data granularity. The total monthly and hourly CV(RMSE) in addition to the NMBE were calculated to compare all models, shown in **Table 7.3.**

Table 7.3: Statistical measures for the performance between measurements and predictions of the base-case model, meta-model optimisations and first-principle models.

	B	M2	M3	C1	C2	C3
NMBE (%)	-1.7	-3.2	-3.3	-4.2	-5.8	-5.5
CV(RMSE) month (%)	12.1	6.5	10.8	17.3	7.0	12.6
CV(RMSE) hourly (%)	66.7			71.8	70.1	71.8

The meta-models M2 and M3 have the lowest error indicated by the CV(RMSE) for the months, the most important indicator for analysing the discrepancy, whereas the total NMBE is subject to cancellation errors (positive and negative differences cancel each other out). However, due to meta-model error, their first-principle models, C2 and C3 perform slightly worse. Because of this, C3 performs at a similar level as the base case model.

To see what how the optimisation has changed predictions on an hourly level, a comparison is shown in **Figure 7.13** between the models and measurements for typical weekday and weekend day for the different energy end-uses. Although a significant change in monthly energy use was apparent due to the variability allowed by the input parameters, this difference is reflected by minor changes in the typical daily profiles. Models C1 to C3 variate slightly and are either somewhat lower or higher than the base case model. No horizontal change is possible since the input parameters do not allow for this change to happen, the input parameters affect only the increase or decrease of energy use. However, the initial profiles for lighting, power and occupancy for the base case model were determined from actual energy use profiles and were therefore already significantly close the reality. The meta-models and automated calibration process would be more effective when there is a larger difference in energy use predicted by the base case model and measurements.

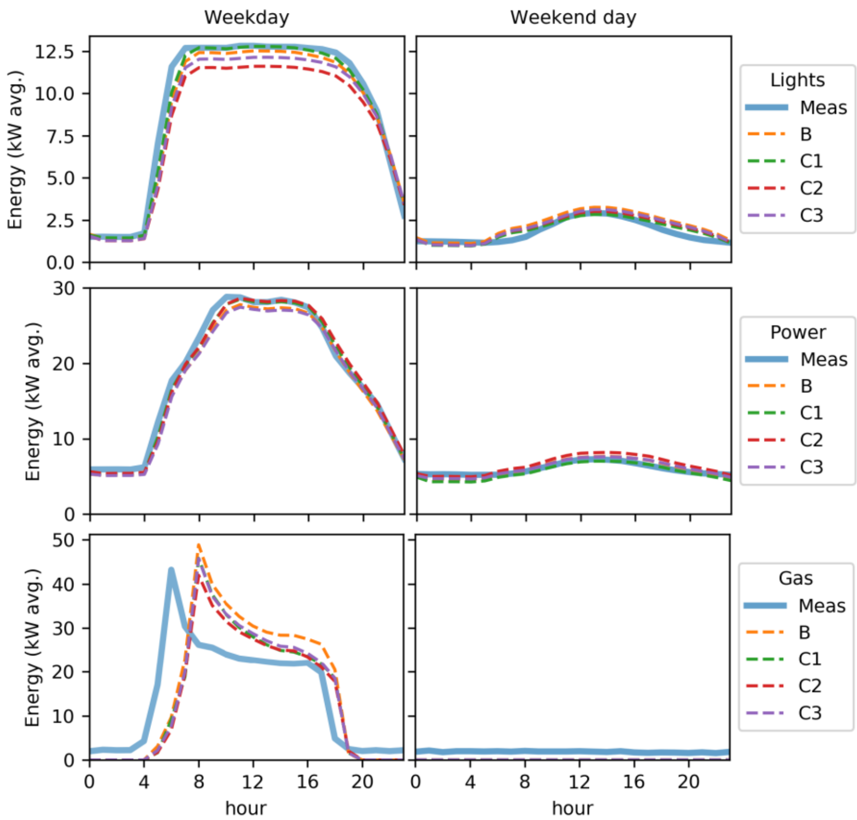


Figure 7.13: Electricity use for a typical weekday and weekend day for L&P and Systems, comparing measurements with predictions from the input parameters of the four calibrated meta-models at different levels of data granularity.

7.3 CH

7.3.1 Meta-model development

The meta-models were trained on the inputs and outputs from 3000 parametric simulations. In total, 85 input parameters were varied, including system efficiencies, power densities for lighting and appliances and people density in different space types, a horizontal offset to these profiles, parameters that determine when windows are opened, heating and cooling set-point temperatures, DHW hot water use, natural and unwanted infiltration rates, a seasonal variation factor per month. Four meta-models were created, by training them on the variable input parameters and energy use outputs. These meta-models predict energy use at different levels of data granularity as follows:

- M1: Yearly end-uses (3 objectives: L&P, systems, gas).
- M2: Monthly energy use (12 objectives: Sep '16 to Aug '17).
- M3: Monthly energy end-uses (36 objectives: 3 end-uses per 12 months).
- M4: Monthly energy end-uses plus typical weekday and weekend day per end-use (180 objectives: 12 months * 3 end uses + 3 * 24 hours * 2).

The meta-model developed for predicting energy end-uses on a monthly (and yearly) basis for Office 71 used partial least squares regression, to minimise differences between the training data (simulation results) and those predicted by the function. It did this effectively by achieving $r^2 > 0.98$ for both yearly and monthly energy end-uses. However, CH introduced additional complexity within the inputs and outputs by including additional parameters that take into account variations in the profiling, in addition to including systems energy use. Systems energy use is much more variable and less linearly correlated to changes in input parameters in comparison to lighting and power energy use. Due to this, similar levels of meta-model accuracy as seen in Office 71, could not be established in CH. Nevertheless, significantly high prediction accuracies were obtained for the four trained meta-models as shown in **Table 7.4.**

Table 7.4: Predictor scores for artificial neural network (NN) and partial least squares (PLS) used for multivariate regression at the different levels of granularity of data.

	R^2		MAE (kWh)		No. of objectives
	PLS	NN	PLS	NN	
M1	0.975	0.976	4094	4243	3
M2	0.959	0.958	1065	1137	12
M3	0.939	0.977	422	359	36
M4	0.956	0.978	117	113	180

The neural network outperforms partial least squares regression and the other linear regression techniques for all meta-models, see also **Figure 7.14** for M2 specifically. Although the differences are small, they are important to the final results as was shown in Office 71, where the translation of optimised meta-model input parameters into the building energy software introduced a small margin of error. The mean absolute error is significantly different between the different meta-models as they predict energy use at different levels of data granularity. More specifically, M1 has a high MAE because it indicates the model error in predicting yearly energy use, in comparison M3 predict monthly energy end-uses, MAE indicates the difference between predictions and measurements of 12 months per energy end-use.

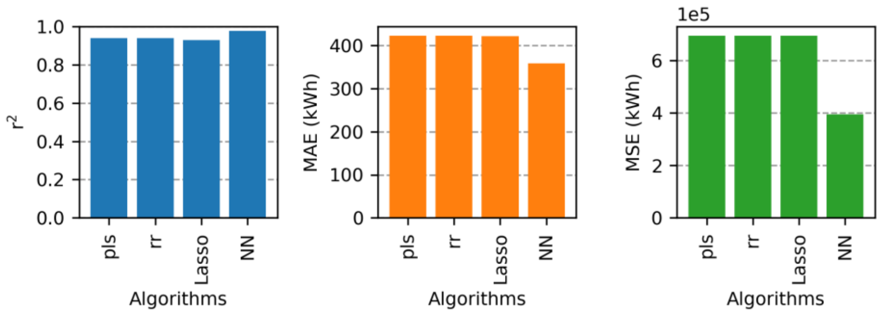


Figure 7.14: Machine learning algorithms (pls = partial least squares regression, NN = artificial neural network, rr = ridge regression and Lasso regression) and their respective scores to evaluate the predictive accuracy of meta-model M3.

A significant amount of data for training the regression algorithms is important when inputs and outputs are not so strongly linearly related, this is evident in **Figure 7.15**, where the increase in model accuracy is shown with an increase in the number of simulations. In Office 71, only 200 simulations achieved sufficient model accuracy, for CH, a similar accuracy is achieved after several thousand simulations.

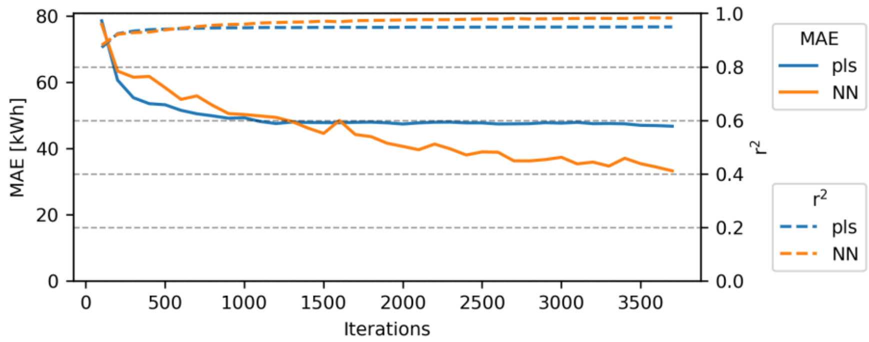


Figure 7.15: Learning progression of the meta-model trained on monthly energy use from 3750 simulations using partial least squares (pls) and an artificial neural network (NN). MAE and r^2 indicate their predictive accuracy with regards to monthly energy end-uses and typical weekday and weekend day consumption per month.

7.3.2 Mathematical optimisation

After training of the meta-models, they are employed for optimisation. **Figure 7.16** shows the minimisation of three objectives or yearly energy end-uses, L&P, systems and gas. After about 300 generations the objectives are minimised towards zero.

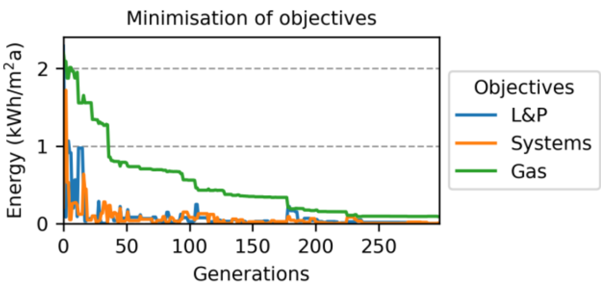


Figure 7.16: Minimisation of energy use between meta-model predictions and measurements for yearly energy end-uses (3 objectives).

Meta-model M2 was used to minimise differences between monthly energy use predictions and measurements as shown in **Figure 7.17**. Again, optimisation achieves differences of less than <1% CV(RMSE) monthly, significantly below typical calibration criteria.

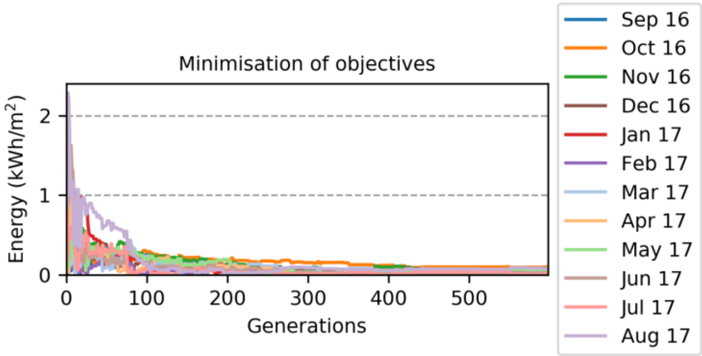


Figure 7.17: Minimisation of energy use between meta-model predictions and measurements for monthly energy use (12 objectives).

However, both M3 and M4 predict monthly energy uses, whereas M4 also predicts energy use for a typical weekday and weekend day for these end-uses, which are an additional 24 hours for two days for two end-uses. Minimisation of energy end-uses for both meta-models stabilise before reaching zero, results for M4 are shown in **Figure 7.18**, where M3 has similar optimisation results. Only L&P and some months for systems energy use were decreasing during the first few generations, but quickly stabilise after.

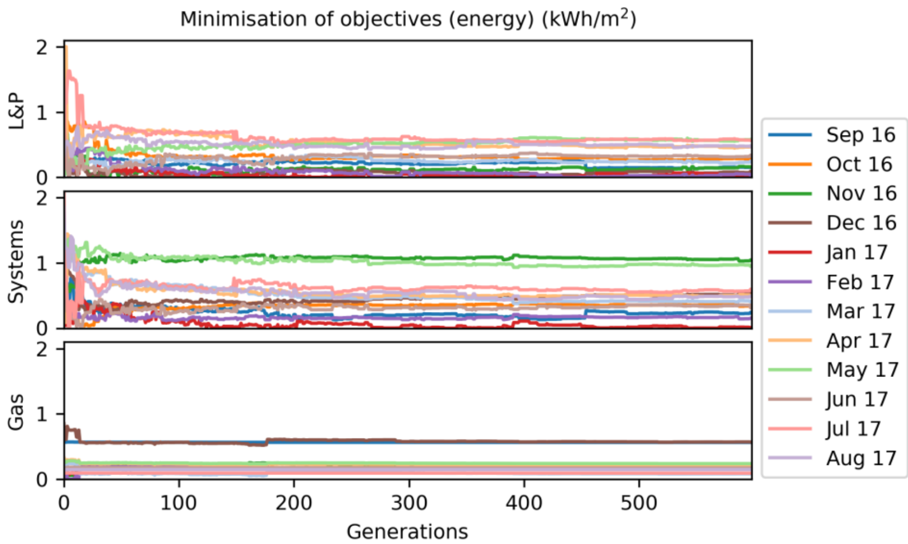


Figure 7.18: Minimisation of energy use between meta-model M4 predictions and measurements for monthly energy end-uses (36 objectives).

M4 also included the typical days as objectives for minimisation, but similarly these were not being minimised. Both systems and L&P energy use for a typical day fluctuate due to changing individuals (set of inputs parameters), but they do not minimise, as shown in **Figure 7.19**. Possibly, there are too many objectives to be minimised by the genetic algorithm and therefore it has difficulty in finding input parameter combinations that improve the fitness and / or the potential solutions available lie outside of the solutions space. However, when analysing the solution space created by the parametric simulations in the previous chapter (within uncertainty analysis), it becomes clear that most of the measurements fall within the uncertainty bands of the 3000 simulation run predictions. Nevertheless, a combination of input parameters might not exist that satisfies an exact representation of measured energy use. In addition, several of the monthly measured energy uses, in particular gas use during December and September 2016 are outside of their respective solution space, which can also be observed in the minimisation, where both of these do not decrease below ~0.65 kWh/m² in contrast to the other months of gas energy use.

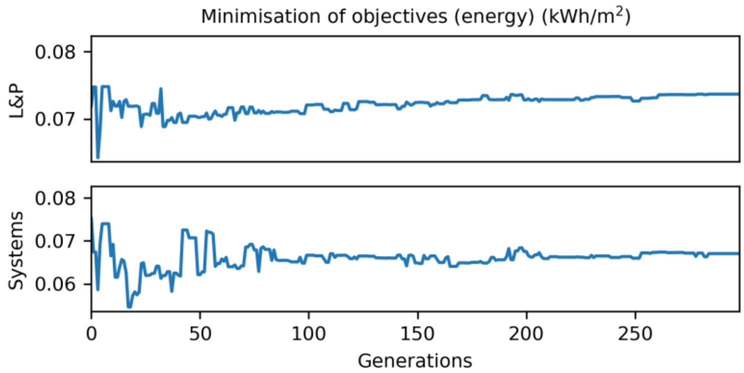


Figure 7.19: Minimisation of typical weekday energy use for meta-model M4, included as objectives.

Table 7.5 shows the most significant (according to sensitivity analysis) parameters before and after automated calibration for CH. In total, 85 input parameters were varied and most of them were adjusted during optimisation. If the optimisation includes end-uses, it will keep changing variables that affect a particular end-use to bring them closer to its respective measured energy end-use. Although all final parameters values seem very arbitrary, there are some fundamental differences between the possible parameter values due differences in the models and their objectives. Meta-model M1 has more solutions that are considered optimised than M3 or even M2. However, this is not clearly reflected in the final parameters.

Table 7.5: Most significant parameter values before (base case) and after automated calibrations optimisation using meta-models (M1 to M4) for CH.

Parameters	Limits	B	M1	M2	M3	M4
Boiler 1 (0-1)	(0.7, 0.93)	0.93	0.70	0.94	0.78	0.71
Circulation lighting (W/m²)	(3.2, 12.7)	8.0	4.1	5.4	5.0	6.4
Heating dead band (⁰C)	(0, 3)	3	2	1.5	2.5	2.5
Fixture female toilets (l/s)	(0.009, 0.036)	0.02	0.02	0.02	0.03	0.02
Fixture kitchenettes (l/s)	(0.008, 0.033	0.02	0.02	0.02	0.03	0.02
Fixture male toilets (l/s)	(0.009, 0.036)	0.02	0.03	0.02	0.02	0.02
Fixture showers (l/s)	(0.030, 0.121)	0.08	0.09	0.08	0.04	0.09
Infiltration (flow ext. surface) (m/h)	(1.238, 4.954)	3.09	4.86	2.66	3.24	3.09
Office equipment (W/m²)	(4.8, 19.2)	12	17	13	18	11
Office heating setpoint (⁰C)	(21, 24.5)	24	24	22.5	24	23.5
Office lighting (W/m²)	(4.06, 16.25)	10	12	10	11	12
Server equipment (W/m²)	(20, 80)	50	45	59	58	66
Weekday offset (30min)	(0, 4)	2	0	3	1	3

*based on the difference per yearly or monthly, monthly end-use prediction and measurement

7.3.3 Impact of data granularity

The different meta-models (M1 to M4) and their respective first-principle based model (C1 to C4) are compared for yearly energy use, monthly energy use and monthly energy end-uses in **Figure 7.20**, **Figure 7.21** and **Figure 7.22** respectively. Similar to Office 71, the meta-models optimised for their specific objectives perform best in predicting their objective energy use.

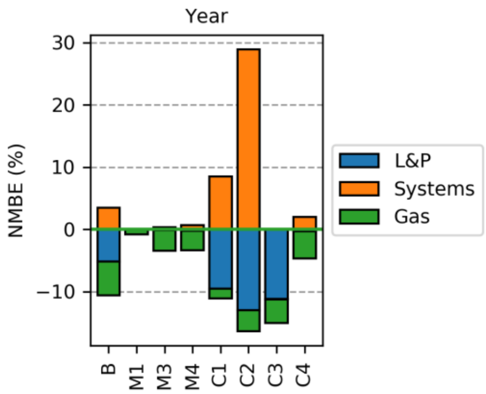


Figure 7.20: NMBE between total yearly energy end-uses for the two meta-models (M1 and M3), first-principle models (C1-C3) and base case model (B) and measurements.

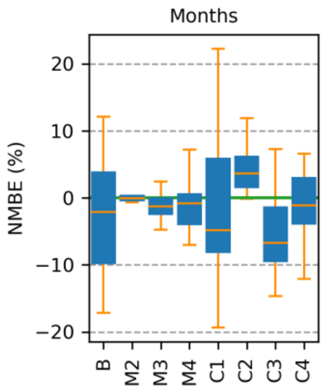


Figure 7.21: NMBE between monthly predictions from the two meta-models (M3 and M2), first-principle models (C1-C3) and base case model (B) against measurements, each boxplot representing 12 data points for the monthly NMBE.

It again highlights the importance of calibrating energy models at a high level of data granularity when employing automated calibration. With manual calibration, it is up to the modeller to understand how changes in inputs affect the output, however also for manual calibration it is necessary that results are analysed at a high level of data granularity, including end-uses and hourly variation. In contrast, the automated calibration procedure adjust whichever parameter minimises towards their objectives, therefore C1 is not well suited at predicting monthly energy use and vice-versa C2 is not well suited at predicting yearly energy end-uses. Their respective meta-models M1 and M2 are not even able to predict these values as they are trained to only predict their specific objectives. M3 and M4 on the other hand are more comprehensive, able to predict monthly energy end-uses, these can also predict yearly energy end-uses or monthly energy use by taking the sum. As such, M3 and M4 can be seen as multi-purpose meta-models, although their accuracy in predicting specific objectives of, for example M1 will be less accurate.

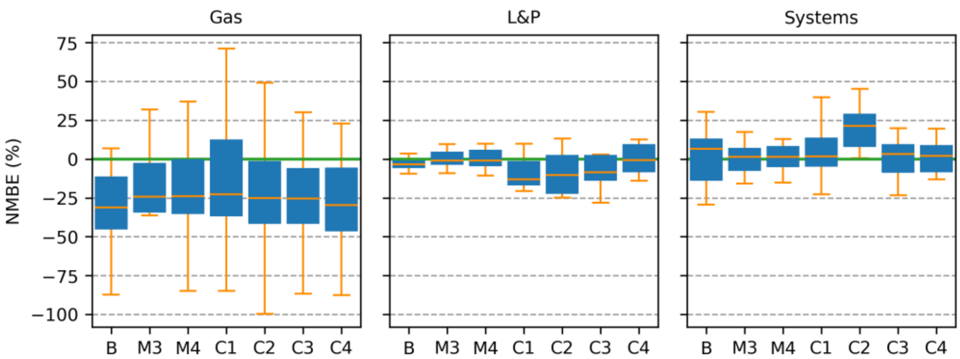


Figure 7.22: Normalised mean bias error between monthly predictions from meta-models M3 and M4, first-principle models (C1-C4) and base case model (B) against measurements, each boxplot representing 12 data points for the monthly mean bias error.

Table 7.6 shows the final total performance based on the NMBE and CV(RMSE) of the meta-models and their respective first-principle models with identical input parameters. The base case performed within the calibration criteria, the optimised meta-models M2 to M4 minimise the difference between predictions and measurements, but translation of the calibrated inputs into EnergyPlus again shows that it is affected by model errors. Finally, only C4 seems to perform significantly better than the base case, but the relevance of the increase in accuracy is arguable. During manual calibration it was identified that minor changes or previously unobserved changes in the building can significantly affect the difference between predictions and measurement, much more than the increase of accuracy that was attained through the automated calibration of the base case model for CH.

Table 7.6: Statistical measures for the performance between measurements and predictions of the base-case model, meta-model optimisations and first-principle models.

	B	M2	M3	M4	C1	C2	C3	C4
NMBE (%)	0.6	-0.1	-1.1	-0.9	-0.9	4.2	-5	-0.9
CV(RMSE) month (%)	8.4	0.4	4.2	3.9	12.4	5.5	8.4	4.9
CV(RMSE) hourly (%)*	26.6				32.9	32.8	30	29.2

*CV(RMSE) hourly is based on L&P and Systems hourly electricity use only

Finally, typical weekday and weekend day energy use for L&P and systems are compared, shown in **Figure 7.23**. Parametric simulation for CH included a horizontal offset of the lighting, power and occupancy profiles, which, as can be seen in the final calibrated input parameters, has also been adjusted during optimisation, different for the five models (B, C1 to C4). In contrast to Office 71, the profiles are now able to move horizontally, allowing automated adjustment of parameters that influence the width of the profiles instead of solely the height or magnitude.

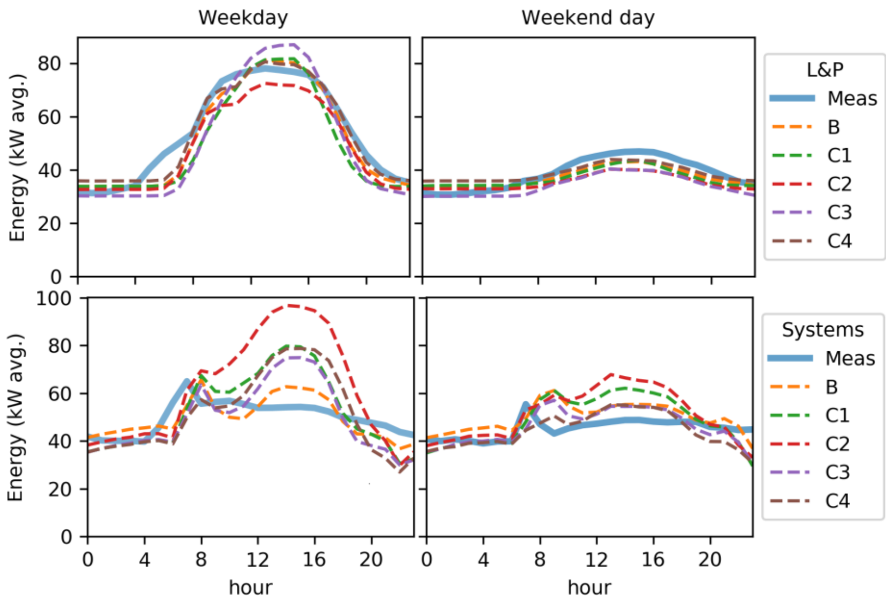


Figure 7.23: Electricity use for a typical weekday and weekend day for L&P and Systems, comparing measurements with predictions from the input parameters of the four calibrated meta-models at different levels of data granularity.

It seems however that the profiles, although adjusted, do not align more accurately with the measurements, at least not for the typical day profiles of energy use. As was evident in **Table 7.6**, the base case has the smallest error on an hourly basis. This was expected for the models C1 to C3, but not so, for C4, which included objectives for minimising the difference intended to lower the hourly error metric. The optimisation was over parameterised, i.e. there were many input parameters adjustable with too many objectives, but also limited to the variation in put parameters. The offset on the lighting and power profiles was not sufficient to allow for replicating measured energy use behaviour on an hourly level.

7.4 MPEB

7.4.1 Meta-model development

For MPEB four meta-models were trained on the inputs and outputs from 3000 parametric simulations. In total, 120 input parameters were varied, including system efficiencies, power densities for lighting and appliances and people density in different space types, a horizontal offset to these profiles, parameters that determine when windows are opened, DHW hot water use, natural, mechanical and unwanted infiltration rates and a seasonal variation factor per month. Four meta-models were created, by training them on the variable input parameters and energy use outputs. These models predict energy use at different levels of data granularity as follows:

- Meta-model M1: Yearly end-uses (5 objectives: L&P, chillers, systems, servers, workshops).
- Meta-model M2: Monthly energy use (12 objectives: Sep '16 to Aug '17).
- Meta-model M3: Monthly energy end-uses (60 objectives: 5 end-uses per 12 months).
- Meta-model M4: Monthly energy end-uses plus typical weekday and weekend day per end-use ((12 months * 5 end-uses + (5 end-uses * 24 hours * 2 days))

Table 7.7 shows a comparison of the prediction scores obtained for the trained meta-models for partial least squares regression and the use of artificial neural networks for regression. Accuracies obtained for the meta-models are different from previous two case studies. For Office 71, linear regression techniques were extremely accurate in learning the data patterns due to the linearity in the different end-uses, specifically lights, power and gas use were predicted. However, in CH, additional input parameters and the objective of systems energy use introduced complexity in the relationship between inputs and outputs, the neural network outperformed linear regression techniques in predicting three energy end-uses; system, gas and L&P. In MPEB however, the amount of objectives increased from 3 to 5 for yearly energy use and from 36 to 60 for monthly energy end-uses. In both cases, PLS and the neural network were unable to achieve similar levels of accuracy as those obtained for CH, r-squared scores of 0.84 are obtained for both. In contrast, for monthly energy use (no end-uses) an r-squared value of 0.99 is calculated, which is a near perfect fit. However, the RMSE for the neural network is about 1/3 higher than that predicted by the partial least squares regression, indicating that PLS is a better predictor.

Table 7.7: Predictor scores for artificial neural network (NN) and partial least squares (PLS) used for multivariate regression at the different levels of granularity of data.

	R^2		MAE (kWh)		RMSE (kWh)		no. of objectives
	PLS	NN	PLS	NN	PLS	NN	
M1	0.84	0.83	4648	6149	7612	9395	5
M2	0.99	0.99	1866	2669	2424	3427	12
M3	0.84	0.83	506	698	896	1119	60
M4	0.87	0.89	197	156	794	664	300

7.4.2 Mathematical optimisation

Optimisation towards the objectives using meta-models M1 and M2 achieved similar results as for Office 71 and CH. The minimisation converged to zero after several hundred iterations. Meta-model M3 and M4 also showed similar results to CH, in that they did not converge towards zero for all end-uses as can be seen for M3 for the monthly energy end-uses in **Figure 7.24** and meta-model M4, which includes the monthly energy end-uses and shows the typical weekdays for each end-use in **Figure 7.24**.

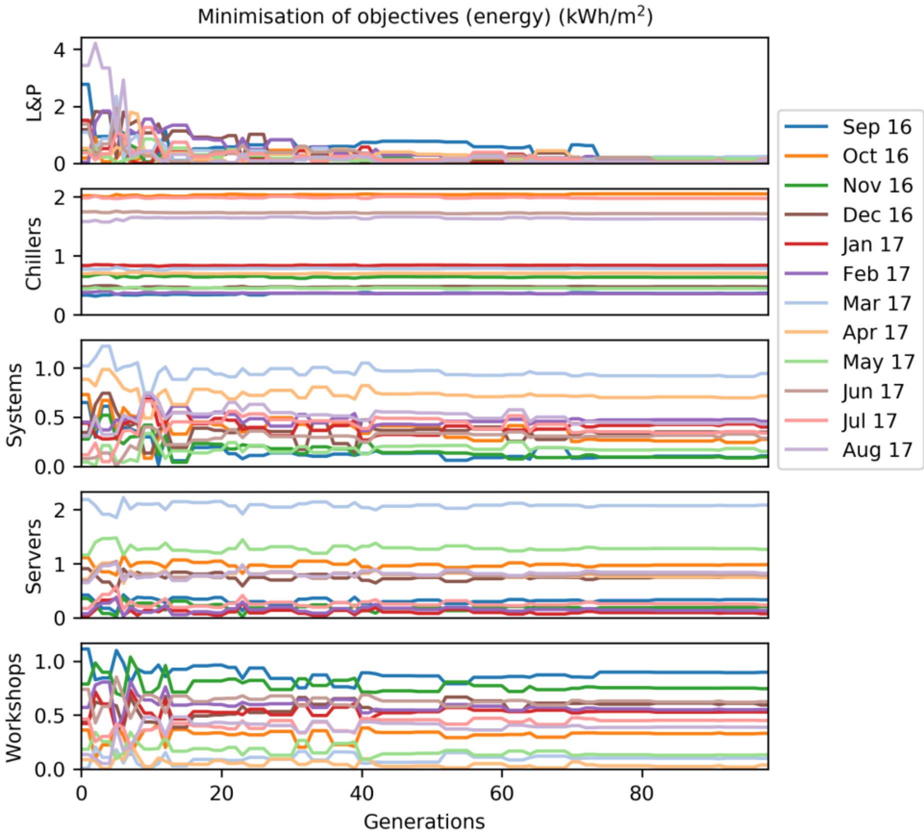


Figure 7.24: Minimisation of monthly energy end-uses between meta-model M3 predictions and measurements (60 objectives).

Although lighting and power is being minimised for both meta-models M3 and M4, the other end-uses remain relatively stable throughout the generations. The optimisation is not able to find better solutions and is limited by the input space.

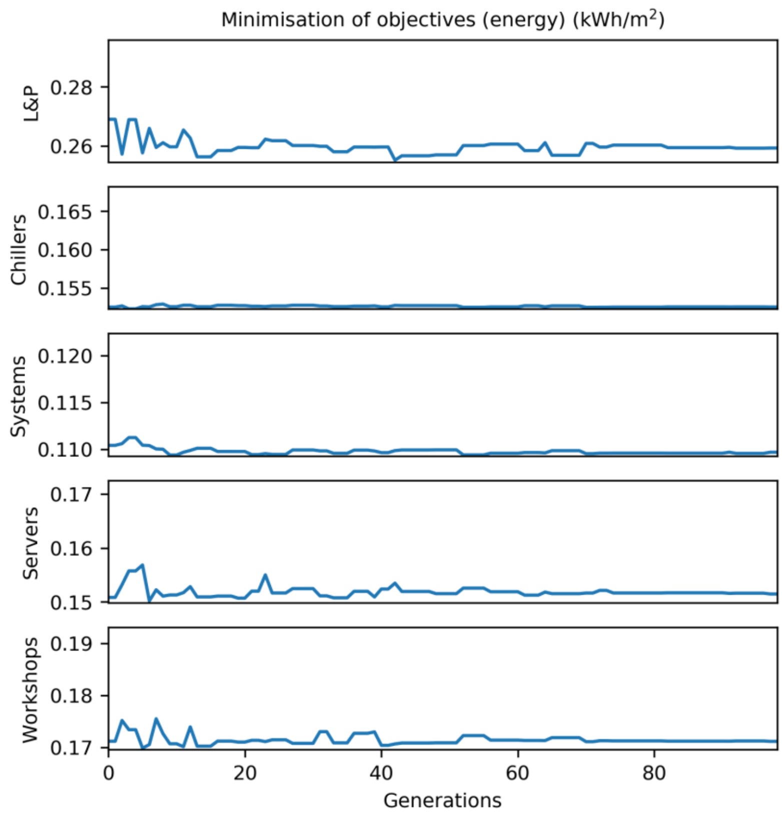


Figure 7.25: Minimisation of monthly energy end-uses between meta-model M4 predictions and measurements (300 objectives, including monthly energy end-uses).

The best individual for each optimisation using the different meta-models is shown in **Table 7.8.** Similar to the previous case studies, the change in input parameters seems to happen arbitrarily, but rather the optimisation process is trying to find the best set of input parameter combinations that fit the objectives (measured energy use targets). As such, many combinations will exist that minimise the difference, however with the amount of input parameter and objectives, local minimisation is very likely and difficult to avoid. This was clear when the optimisation was run for in particular the more complex meta-models M3 and M4, which when running the optimisation multiple times, in some cases stabilised earlier than others.

Table 7.8: Most significant parameter values before (base case) and after automated calibrations optimisation using meta-models (M1 to M4) for MPEB.

Parameters	Limits	B	M1	M2	M3	M4
Server 409 condenser CCOP	(1.16, 6.27)	3.38	2.7	4.6	3.2	4.1
Out-of-hours equipment baseload (%)	(60, 100)	85	95.9	93.9	74.6	98
HWS fixture labs (l/h)	(0.15, 0.61)	0.38	0.34	0.59	0.25	0.37
HWS labs schedule y-axis (0-1)	(0.1, 1)	0.50	0.24	0.39	0.37	0.63
Labs mechanical ventilation (l/s)	(1.6, 6.4)	4	4.4	3.5	4.1	3.7
Labs equipment (W/m^2)	(6.5, 26)	16	14.6	21.6	19.9	11.0
Labs lighting (W/m^2)	(8, 23)	16.26	15.1	16.8	21.0	11.4
Out-of-hours lighting baseload (%)	(30, 100)	65	70.0	80.1	68.2	73.7
Office mechanical ventilation (l/s)	(3.2, 12.8)	8	8.8	7.4	9.2	9.3
Office equipment (W/m^2)	(6.4, 25.6)	16	21.1	17.4	22.4	13.4
Office lighting (W/m^2)	(4.8, 19.2)	12.03	16.3	12.6	11.3	14.5
Workshop equipment (W/m^2)	(22, 88)	55	62.8	51.6	52.1	57.5

7.4.3 Impact of data granularity

The different meta-models (M1 to M4) and their respective first-principle based model (C1 to C4) are compared for yearly energy use, monthly energy use and monthly energy end-uses in **Figure 7.26**, **Figure 7.27** and **Figure 7.28** respectively. Similar to Office 71, the meta-models optimised for their specific objectives perform best in predicting their objective energy use.

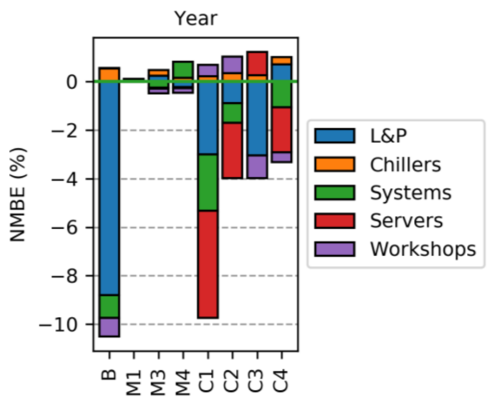


Figure 7.26: NMBE between total yearly energy end-uses for the two meta-models (M1 and M3), first-principle models (C1-C3) and base case model (B) and measurements.

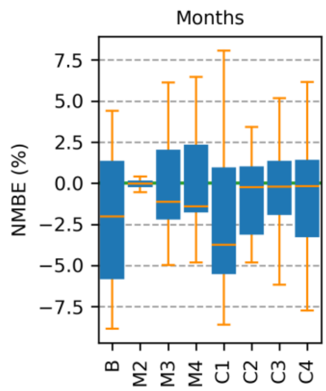


Figure 7.27: NMBE between monthly predictions from the two meta-models (M3 and M2), first-principle models (C1-C3) and base case model (B) against measurements, each boxplot representing 12 data points for the monthly NMBE.

The three figures show that the meta-model error propagates through to the first-principle models when these are simulated using the best optimised individuals (set of inputs). Comparing M1 to C1 in **Figure 7.26** shows a difference of nearly -10% NMBE, while smaller still significant differences for the models, M2 compared to C2 in **Figure 7.27** and M3 to C3 in **Figure 7.26.** The larger difference in the first meta-model M1 to C1 could be due to one or several input parameters that have a significant impact on energy use, possibly the server equipment power density input. In other words, the meta-model is trained based on inputs and outputs, but will not be 100% accurate in replicating the first-principle software. Thus, it may accurately represent the effect a certain parameter has on an output, which when computed in EnergyPlus will show a significant difference in the output from that predicted by a meta-model.

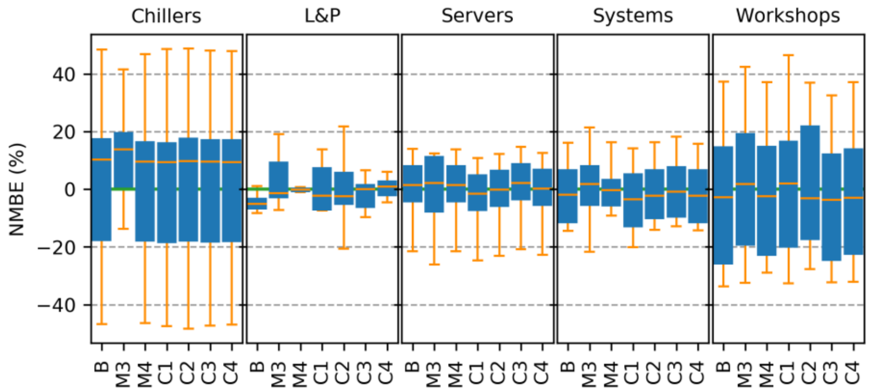


Figure 7.28: Normalised mean bias error between monthly predictions from the meta-model (M3), first-principle models (C1-C3) and base case model (B) against measurements, each boxplot representing 12 data points for the monthly mean bias error.

The base case, meta-model and first principle model predictions are compared with the measurements and their differences are quantified by three statistical measures, the NMBE and CV(RMSE) on a monthly and hourly basis as shown in **Table 7.9.** The meta-models established were reasonably accurate in predicting new sets of inputs, and used in the optimisation process establish significantly low CV(RMSE) values < 0.1 % for the monthly criteria. However, the meta-models are of limited use by themselves, as they are limited to only predict new solutions by changing parameters they are initially trained on. However, for predicting retrofit savings in existing buildings, or perhaps other purposes, a calibrated model will preferably predict a higher level of data granularity and also allow for more specific changes to be made within a model.

Table 7.9: Statistical measures for the performance between measurements and predictions of the base-case model, meta-model optimisations and first-principle models.

	B	M2	M3	M4	C1	C2	C3	C4
NMBE (%)	4.5	0.3	3.2	3.2	8.6	2.6	3.5	4.0
CV(RMSE) month (%)	-2.0	-0.1	0.0	0.1	-1.8	-0.6	-0.6	-0.5
CV(RMSE) hourly (%)*	9.8				12.3	8.8	9.2	9.7

Another limitation of the automated calibration approach taken is clarified when analysing differences at an hourly level. **Figure 7.29** shows the typical weekdays and weekend days for the different energy end-uses. The variable input parameters previously defined during parametric simulation mainly affect the magnitude of different components in the building energy simulation. For example, a change in lighting, equipment or a flow rate will directly increase or decrease energy use. Thus on a time scale, energy use will go up or down, but will not affect the typical profile or duration of energy use. These are typically static, limiting the measured energy use profile to be replicated through the automated calibration process. With this thought in mind, the approach taken here was to incorporate some variability into the schedules, in particular lighting, power and occupancy schedules were varied by introducing an offset. However, this offset was applied to the whole profile throughout a year, it changes a typical daily profile

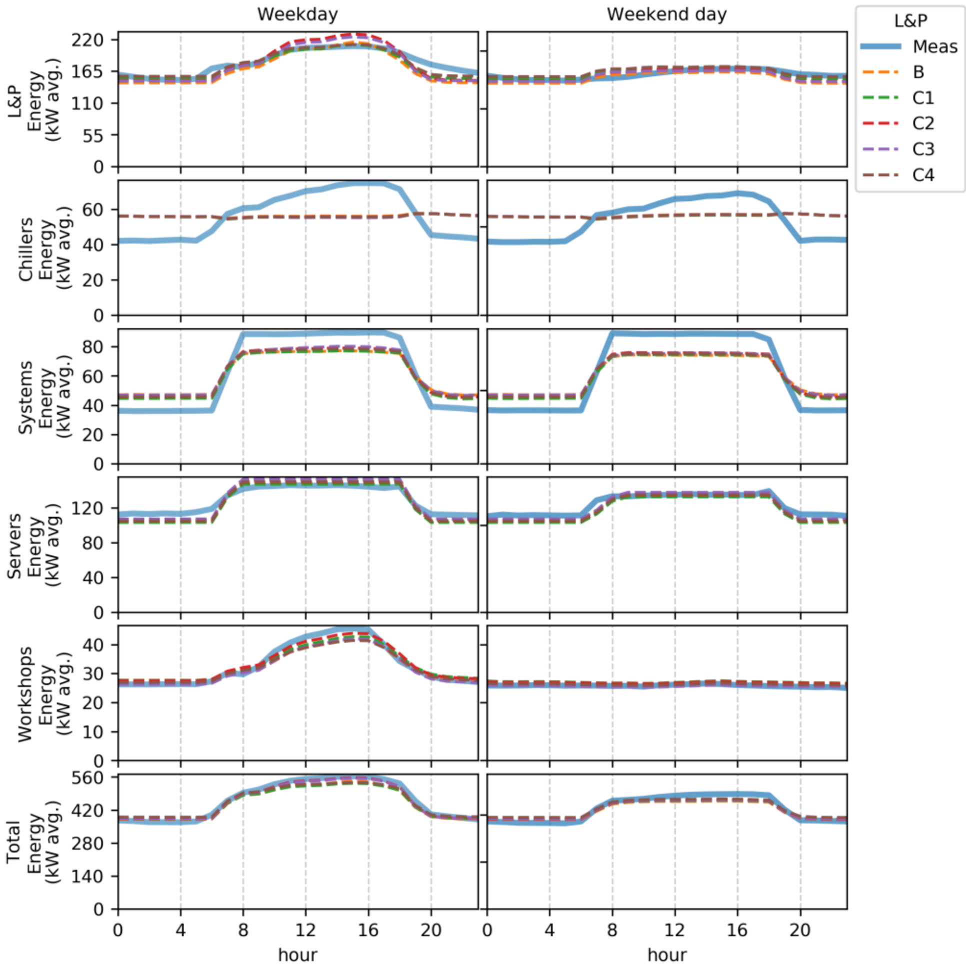


Figure 7.29: Electricity use for a typical weekday and weekend day for the different end-uses and total energy use, comparing measurements with predictions from the input parameters of the four calibrated meta-models at different levels of data granularity.

7.5 Summary

The automated calibration process sought to further reduce the discrepancy between predicted energy use from the manually calibrated models and measurements. It did this successfully at different levels of data granularity using the created meta-models and optimisation process. The automated calibration process extended the state of the art, by utilising meta-model based multi-objective optimisation to minimise energy use at a high level of data granularity. Typically, automated calibration is performed on yearly or monthly energy use, which as was found during the manual calibration procedure, masks energy use at a higher level of data granularity.

Meta-model development

This chapter investigated the application of developing meta-models as surrogates for the full-scale first-principle models. Several machine learning algorithms were tested to learn relations between variable input parameters and outputs (energy use). Multivariate regression models were developed using, in particular, partial least squares and artificial neural networks. Other algorithms, such as Radial Basis Functions, Support Vector Regression were not utilised, but have been reported to perform reliably in other research and should be further investigated concerning multivariate regression of building energy modelling data. Meta-models were built for different levels of data granularity, in particular for predicting yearly energy end-uses (M1), monthly energy use (M2), monthly energy end-uses (M3) and monthly energy end-uses including typical weekday and weekend days for these end-uses (M4). Fundamentally, only one meta-model would have been necessary, for example M4 predicts the outcomes of the other three meta-models (M1 to M4). However, the aim was to understand the limitations of meta-models and the impact of data granularity on model calibration accuracy. Their prediction scores for the different case studies are shown in **Table 7.10**.

Table 7.10: Machine learning (meta-) models best predictor scores, number of inputs and outputs.

	Office 71		CH		MPEB	
No. simulations	1000		3000		3000	
Inputs	115		85		120	
Outputs (M1 to M4)	(3, 12, 36)		(3, 12, 36, 180)		(5, 12, 60, 300)	
	R^2	MAE (kWh)	R^2	MAE (kWh)	R^2	MAE (kWh)
M1	0.992	450 (rr*)	0.975	4094 (PLS*)	0.84	4648 (PLS)
M2	0.986	107 (Lasso)	0.958	1065 (NN*)	0.99	1866 (PLS)
M3	0.985	73 (Lasso)	0.977	359 (PLS)	0.84	506 (PLS)
M4	-	-	0.978	113 (NN)	0.89	156 (NN)

*rr (ridge regression), PLS (partial least squares), NN (artificial neural network)

Meta-models developed with many inputs, but several outputs (M1 and M2) were trained relatively quickly and needed only several hundred simulations to achieve high prediction scores. Whereas more simulation runs were necessary with many more inputs (M3 and M4), in particular for CH and MPEB where the interactions between inputs and outputs were more complex. More specifically, for Office 71, only lights, power and gas use were trained on as energy end-uses which are mostly linearly related to the inputs. Which was not the case for heating, cooling and auxiliary energy use, such as fans and pumps, which fluctuate significantly and are much more dependent on the external environment. For CH and MPEB, the neural network outperformed the linear regression techniques significantly in learning non-linear relations. All four trained meta-models achieved high levels of accuracy, but with more complexity between inputs and outputs, more sophisticated machine learning techniques and data (number of simulations) were required. This was in particular evident in CH, where 3000 simulation runs achieved reasonable levels of accuracy, but where a 1000 simulations were shown to be 40% less accurate. In stark contrast to Office 71, were just 300 models were enough for predicting lights, power and gas energy use.

Mathematical optimisation

Optimisation (i.e. automated calibration) was performed using the developed meta-models. Significantly small margins of error were achieved for different levels of data granularity. However, optimisation at a higher level of hierarchical data granularity made it more difficult to find specific solutions to the objectives as they were not always within the solution space. In addition, multi-objective optimisation with many inputs and outputs introduced local minima where the optimisation converged to non-optimal solutions. Furthermore, the calibration approach included a certain level of uncertainty within the lighting, power and occupancy profiles, to allow for the automated calibration process to find better fitting profiles to represent measured energy use. However, the horizontal offset parameter still strongly constrained the optimisation algorithm, it allowed mostly only horizontal movement of the daily profile as a whole, in reality the profile varies on an hourly level. Therefore, the initial profile from the base case will retain the same shape and if this shape is not in line with measurements, it could not be replicated. Additional variable input parameters would have to be defined that allow changes on an hourly basis during parametric simulation, but this would mean hundreds of extra parameters if several different profiles were to be included. Such an increase in the number of parameters is likely to be too difficult as an optimisation problem, in addition, this would also affect the accuracy of the meta-model, and more simulation runs would be necessary.

Although high levels of prediction accuracy were obtained for the trained meta-models, some margin of error still existed and was propagated to the first-principle software when simulating the calibrated inputs from the optimisation process. In turn, this margin of model error reduced the effectiveness of the process, as these models did not improve the base case model. However, the use of meta-models was unavoidable, as the run time of full-scale models of existing buildings proved to be too time consuming for optimisation using genetic algorithms, which typically need thousands of simulations. Final results of the optimisations are shown as the difference between measurements and predictions by the base case model (B), meta-models (M1-M4) and first-principle model (C1-C4), which are the best individuals (calibrated sets of input parameters) simulated using EnergyPlus, see **Table 7.11**.

Table 7.11: Optimisation results, showing differences between measurements and predictions for the best individuals for the base case (B), meta-models (M1-M4) and first-principle models (C1-C4).

	CV(RMSE) %	B	M2	M3	M4	C1	C2	C3	C4
Office 71	month	12.11	6.54	10.78		17.34	7.04	12.63	
	hour	66.73				71.82	70.13	71.84	
CH	month	8.38	0.36	4.22	3.86	12.36	5.47	8.35	4.91
	hour	26.60				32.90	32.77	29.96	29.19
MPEB	month	-1.99	-0.05	0.00	0.07	-1.81	-0.59	-0.55	-0.46
	hour	9.81				12.31	8.84	9.22	9.70

Due to the limitations of automated calibration, manual calibration is a pre requisite in order to establish a base case model that represents the existing building operation. Although this was already evident in the manual calibration chapter, it is even more apparent during optimisation, where many parameters cannot be adjusted.

Impact of data granularity

The meta-models were used to minimise predicted and measured energy use at different levels of data granularity, purposefully to understand how automated calibration affects the accuracy of the model. When optimisation is performed with different objectives, or an increase in the number of objectives, the solution space will become smaller. A minimisation of yearly energy use will have many possible combinations of input parameters that will solve the minimisation problem, however with monthly energy end-uses the number of solutions will reduce. As such, after performing optimisation, at a lower level (yearly), this will not represent objectives at a higher level of data granularity, while this was not necessarily true when the optimisation is performed at a higher level of data granularity. In other words, minimisation of monthly energy end-uses (M3 objectives) performed well in predicting yearly energy end-uses (M1 objectives) or monthly energy use (M2 objectives). Thus, there is much risk involved in using automated calibration at lower levels of data granularity as they mask the underlying energy end-uses, and might even reduce the accuracy of a previously manually calibrated model. For example, for Office 71, C1 predicts monthly energy use significantly worse than C2 and C3, by about -20 and +10% NMBE, and about -20% and +20% NMBE for the monthly energy end-uses. Similarly, for CH, calibrated model C2 optimised towards monthly (total) energy use masked the energy end-uses on a yearly basis (C1) by -10 to 30% NMBE and for the monthly end-uses it was about 10 to 25% NMBE worse than C3. Therefore, when applying automated calibration, it is recommended to minimise differences on a monthly energy end-use basis (if such data is available), or that strict control of the input variables is maintained to prevent optimising towards inaccurate input parameters.

8 CASE RESEARCH DISCUSSION

A research methodology was developed in order to investigate and mitigate the discrepancy between predicted and measured energy use, this methodology was employed on four case study buildings. The methodology involved: (1) collecting design information and determining actual operational procedures through energy audits, O&M manuals, sub metering, building management system and other data sources. (2) Synthesising and ensuring quality of collected data to understand the building operation and its performance, including establishing a hierarchy to compare data at different levels (3). Modelling four case study buildings based on as-built collected information, which was then compared to measured performance to identify discrepancies. (4) Sensitivity and uncertainty analysis of parametric simulation results to understand how input parameters affect the outputs. (5) Manual calibration of the as-built models to measured performance of the existing building through iterative improvements identified through different data sources. (6) Quantifying the impact of underlying causes of the performance gap through implementing typical assumptions to the calibrated models. (7) Meta-model development and automated calibration at different levels of data granularity to determine the effect of data granularity on model calibration accuracy. Results from the case research are discussed in this chapter.

8.1 Utilising operational data to inform building performance simulation

Data collection was an essential part of this research, integral to the analysis and modelling of energy performance at a high level of data granularity. As-built information to inform the building model was collected by performing building audits, talking to facilities management, and searching for relevant information in O&M manuals. Quantitative energy performance data was collected from the sub-metering systems and additional short-term monitoring. Occupancy presence data was retrieved based on Wi-Fi connections and swipe card access data. Finally, system performance and environmental data were obtained through a building management system. This data was used to inform the building performance simulation process. In this research, Python programming language was used to integrate and analyse the different data sources; to compare like-for-like objects and evaluate the importance of data granularity on building model calibration. A data hierarchy was established to distinguish granularity at a hierarchical, spatial and temporal level.

Data collection

During the process of data collection, it became apparent that its quality was hard to ascertain. In the case of sub-metering, it was difficult and sometimes impossible to establish how building appliances were connected to their respective meters. Labelling was often incomprehensible and/or there were several electrical distribution schematics, from different contractors with different labels for the meters. Distribution schedules from O&Ms are rarely up to date, especially so for older buildings which have gone through several commissioning phases. In Office 71, a good level of disaggregation was established, but the large HVAC systems (VRF heat pump and AHU) were not measured, even though the VRF indoor units were. In contrast, for both MPEB and CH, many meters existed throughout the building, but some of the larger energy uses were not separated, while smaller ones were. In the end, the hierarchical granularity extended to separating lighting, power, system and gas energy use, where spatial granularity extended to all floors. In Office 17 and 71, lighting and power energy use was available for each floor separately, but this was not available from the sub-metering system in CH and MPEB. Although data was available on some floors, it was not for all floors, and additional sub-metering was necessary.

For the building management system, similar issues arose. Labelling of particular system components were difficult to distinguish, in particular when numerous components existed, such as with fan coil units. It was often unclear in what space a certain fan coil was located. Nevertheless, the BMS data proved helpful in tuning the performance of some of the HVAC systems, such as the air-handling units in MPEB, where the supply and return air temperatures were compared with those predicted. In addition, the internal space temperatures in MPEB were analysed to establish typical setpoint temperatures in different space types. This however, proved to be difficult as there was a lot of variability in between spaces and the control strategies were very dissimilar. In addition, the reported setpoint temperatures measured at the fan coil units were not always representative for the actual control in that space, internal temperatures often deviated from the setpoint temperatures.

Developing modelling assumptions

Collected operational data was used to inform as-built modelling assumptions, where design information from O&M manuals can describe operational strategies, measured data can verify these strategies and be used to develop specific assumptions. In particular, measured occupancy and electricity data was used to develop occupancy, lighting and equipment schedules. Occupancy presence from Wi-Fi and swipe-card access data informed when and where the building was occupied.

In Office 71, lighting and power were separated for each floor, measured electricity was used to create typical weekday and weekend day schedules for lighting and equipment per floor. Similarly, in MPEB and CH typical weekday and weekend day schedules and seasonal occupancy factors were applied to occupancy, equipment and lighting schedules to increase the accuracy of the model during manual calibration. However, at a yearly temporal granularity, an average typical profile masks the differences existent in the weeks and days. The level of detail could be increased by creating average schedules for each week or even implement actual daily data as schedules, but the model would then become fitted to the data. Instead typical daily schedules were created based on monthly data.

Lighting and equipment schedules are separate inputs into building energy software, with no direct relation to occupancy schedules, however, in reality occupancy presence has a large influence on both energy lighting and equipment energy use, and were therefore programmatically related in the case research. For MPEB and CH, this was accounted for by including variable parameters for the out-of-hours baseload and a horizontal offset. This allowed for representing any changes in occupancy affecting lighting and power energy end-uses. This approach was taken as occupancy data for CH and MPEB was available from Wi-Fi data, which showed a strong correlation between lighting and power energy use and occupants. Although Wi-Fi data and swipe card access can be a valuable data source for understanding occupancy presence, when such data is unavailable, detailed sub-metering data could be used as a proxy for occupancy and the creation of lighting and equipment schedules. However, the assumed baseload identified within the electricity profile needs to be correctly assumed to achieve an accurate representation.

Created schedules were transformed by applying a seasonal variation factor that was strongly present in both university building, applied on a monthly basis. The daily variability of occupancy was then included as uncertainty within the model. The occupancy schedules were then further adapted to create equipment and lighting schedules as these follow a very similar pattern, however the baseload of equipment and lighting is much higher than that of occupancy, (there are no occupants present during the night). This was accounted for when developing these schedules programmatically. In addition, a further variability in the scheduling was allowed for by quantifying the uncertainty of assumed baseloads and horizontal variability (i.e. time variability) of the schedules. Instead of using occupancy data, it was found that electricity can be used as a proxy for occupancy presence if detailed information on lighting and equipment electricity use is available. Occupancy schedules based on the Wi-Fi data were replicated through the manipulation of lighting and power electricity data. However, the assumed baseload identified within the electricity profile needs to be correctly assumed to achieve an accurate representation.

Space and set-point temperatures measured by the BMS were used to validate design assumptions from the O&M manuals. By analysing space temperatures in different space types and comparing them during occupied hours, it became apparent that a large variability in similar space types exists, but generally follow the design assumptions. Nevertheless, the manual control of space set-point temperatures is difficult to represent in a building energy model, specifically if temperatures can fluctuate significantly. In CH, the impact of space set-point temperatures was quantified by introducing a variability in set-point temperatures during parametric simulation. In MPEB, it was helpful to verify set-point temperature in the computer cluster rooms, which have a significant impact on energy use.

Besides space and set-point temperatures, the supply and return temperatures of the five air handling units in MPEB were analysed in order to validate if the design strategies in the O&M manuals were employed. The building model was then adjusted where necessary to represent the actual operation of these systems. The validation of design assumption and subsequent adjustment is actually part of the manual calibration process as it iteratively improves the accuracy of the model to the existing situation. This should be an integral part of the calibration process as incorrect assumptions can have a significant effect on model accuracy.

8.2 Quantifying the impact of underlying causes of a discrepancy

Model calibration of large existing buildings is a time-consuming process, although similar to performance modelling for building design, additional activities, such as the perusal of O&M manuals, energy audits and obtaining measurement data are necessary. This process would be much more time-efficient when previous design models are readily available, building information is well organised and complete, and sub-meters are properly commissioned and their data is collected in an easily accessible database. Fortunately, progression in software development and information technology has led to the development of such integration procedures, a major advancement has been building information modelling (BIM) which facilitates the exchange and interoperability of information in a digital form. Integrated energy modelling within BIM standardised software could drastically improve time-efficiency of typical modelling tasks and consequently the calibration of such models to existing building performance. Currently, model calibration is used for research purposes and Energy Performance Contracting, but may become more common when the development of calibrated models is a proven and less costly technique that has a significant benefit on improving and validating the performance of existing buildings, both through retrofitting and better operation. Until then, several issues concerning model calibration and necessary improvements highlighted throughout this research need to be alleviated.

Manual calibration process

Data collected in this research from sub-metering systems was directly compared with energy modelling outputs through customised meters in EnergyPlus. This process is however complicated when labelling of existing meters are incorrect or actual components are different from what is specified. In-depth analysis of all building components is a time-consuming process for large buildings. When a building has more than several hundred individual spaces, simulation runtimes can take up to more than 30 minutes on a high-end laptop, which makes manual calibration a time consuming process. A solution would be the simplification of building energy models (common practice among energy modellers), where similar spaces or floors are multiplied, however this introduces a margin of error. Avoided in this research, which aimed to capture all processes within a building and compare energy use data at a high level of data granularity at a hierarchical-, spatial- and temporal level.

This research focussed primarily on mitigating a discrepancy between predictions and measurements in energy use. However, environmental system variables can be considered to further improve calibration accuracy of a model. Variables such as space temperature are dependent on fundamental processes in the building, such as the operational strategy (system type, set-point temperatures and scheduling) and the heating and cooling loads (based on external environment and internal gains), and can therefore give a good indication of these processes, their actual settings and behaviour. Environmental data is however not always readily available, nor easily collected for a large number of spaces. Furthermore, with many space types and differences in spaces, it is difficult to find typical patterns and to represent these patterns with a model. Utilising space temperatures and system performance data for several spaces may be possible for manual calibration purposes, but when calibrating an existing building with numerous spaces, this becomes a complex process. A comprehensive framework for the automation of such tasks would need to be established, but this needs to take into account the fact that each building is very different. Furthermore, to benefit from such a framework, rigorous and extensive data collection and data management are essential and need to vast improvement over what is typically available in existing buildings. The immediate benefit of such a calibration framework for the sole purpose of predicting energy saving measures is untenable as the time-intensiveness of setting up such a comprehensive model is likely too time consuming. Rather, a real-time calibrated model or platform would be more constructive if it were used continuously. Another potential application would be the use of machine learning to learn and forecast the behaviour of a building, based solely on measurements of space temperatures, occupancy and external and internal environmental parameters, instead of relying on a virtual model.

Predicted and measured energy use was mitigated through manual calibration, i.e. the iterative manual adjustment of a building energy model to reduce differences between performance predictions and measurements. The reliability of this process is strongly dependent on the availability of data, both design and measured data, which is the evidence to support any iterative changes. If there is a lack of such data, any iterative changes will be arbitrary, even though a discrepancy will be mitigated. Under this rationale it becomes clear that a higher level of data granularity supports the development of a more accurate model, but that a lack of information can mask the real situation. As such, the utilisation of operational data to inform building performance simulation assumptions is essential in model calibration. Furthermore, O&M manuals need to be reviewed to understand the design and intended operation of a building, which subsequently have to be validated through energy audits. Especially the latter is important, in the existing buildings it was clear that the intended design and design strategies differ from that observed, which in many cases had a significant effect on building performance. For example, in Office 17, the boiler radiator heating was manually turned off during the summer, in CH an AHU was out of operation throughout the measurement period, in MPEB some of the server load was connected to the L&P bus-bars as communicated by the facilities manager. Each of these observations had to be accounted for in the models to represent the actual situation. Calibration is an underdetermined process, where many configurations of input parameters can accurately fit the solution. A higher level of data granularity and availability of evidence, can filter out the wrong solutions and increase the accuracy of the model. In this research, a significant amount of data was collected and utilised to inform the calibration models, but in some buildings the lack of certain end-uses introduced significant uncertainty into these models, as these could not be directly compared. In MPEB, a faulty heat meter was not logging district heating data during the measurement period, and in Office 71 the air handlers were not connected to the sub metering system, both which were likely to have a substantial effect on the model accuracy, even though this could not be directly validated. The impact and significance of uncertainty in input parameters was assessed through uncertainty and sensitivity analysis.

Uncertainty and sensitivity analysis

Uncertainty assigned to input parameters results in a spread in predictions from parametric simulation. The spread signifies the different combinations of inputs and outputs, where the input parameters follow a certain distribution (normal, uniform, triangular). Uncertainty in the predicted end-uses was analysed using the coefficient of variation, which quantifies the variability that is predicted by the numerous simulations. Uncertainty in outputs was compared to the measured data, which indicated that measured data points did not always fall within the uncertainty of the predictions. This is an important observation, as subsequent automated calibration will not be able converge towards non-existing solutions, further manual calibration was necessary in these cases to be able to effectively apply automated calibration.

The influence of parameters on energy use can be understood based on the fundamental principles of building performance simulation and its underlying equations. Quantitatively, sensitivity analysis can compute their significance for detailed building energy models and system analysis where the relationship between inputs and outputs is not so easily understood. In the case study buildings, the highest energy end-uses are typically lighting, equipment, systems and or gas energy use, the input parameters that affects these energy uses directly and most significantly are often the lighting and equipment power density and efficiency coefficients respectively. Although this depends on the uncertainty assigned to these parameters. Smaller input uncertainty will lead to a smaller significance coefficient, in case of the boiler coefficients and VRF heat pump COPs, the change from a 20% variability to 2-5% variability was significant. The uncertainty in equipment and lighting loads is likely much larger as it is much more dependent on the number of people present and equipment in the spaces, and type of lighting. In an office building, lighting loads were therefore being assigned a lower uncertainty as they are likely much more static than equipment loads. Again, people presence is typically disconnected in building energy modelling, i.e. lighting and equipment loads have their own schedule, in this case, the uncertainty should actually be larger than when the occupancy schedule directly influences these load schedules.

Impact of typical assumptions

The manually calibrated models were used to compare how typical design assumptions influence energy use. Typical design assumptions are those defined under the National Calculation Methodology, in the UK, which pre-scribe the inputs for specific space types in order to determine the minimum performance requirements of a building for Building Regulations (i.e. compliance modelling). Compliance modelling should not be used as a design tool by informing building efficiency improvements. In practice, however, this occurs as increasingly stringent building efficiency targets set by the government or local councils are not being met. Further refinement of the design is necessary to achieve these targets, tested through the compliance model. Inevitably, this model is then used to test the necessary efficiency measures and will be used to support design changes, even though the compliance model is not an actual representation of the to be build building. As such, efficiency measures can have a significant impact on the compliance model, but may have a less significant or even adverse effect on the carbon emissions or energy use of the actual building. A stronger emphasis on the use of performance modelling is needed in order to drive design decisions that will effectively mitigate energy use and the energy performance gap that arises.

The impact of typical assumptions on energy use highlighted the importance of using performance modelling over compliance modelling. It identified the significant differences that exist between typical assumptions and the actual operation of a building, giving a better understanding of how and what assumptions should be made when using performance modelling. The impact of these assumptions, or 'simplifications' were applied to the calibrated model, in particular significant input parameters, as determined through sensitivity analysis, were analysed. The findings emphasise the need to confirm, most importantly; future equipment loads, equipment and lighting (or occupancy) schedules, seasonality of use, and heating and cooling strategies of a building. In addition, typical lighting and equipment baseloads under compliance modelling are a gross under prediction of actual baseloads measured in the four case study buildings. Some input parameters, such as future occupancy presence, are difficult to determine and may need to be based on realistic use profiles, preferably based on those in similar existing buildings, rather than the simplified assumptions under NCM.

8.3 Quantifying the effects of data granularity on model accuracy

Model calibration is a complex and time-consuming process. Automating calibration can alleviate some of the time-intense tasks. As such, automated calibration was employed based on a meta-model of the first-principle model using a genetic algorithm for mathematical optimisation. It proved to introduce an extra level of complexity, but reduced the time needed to run simulations. This is particularly useful for the larger models which took more than half an hour to run for a full-year simulation in EnergyPlus. In terms of effectiveness, the automated calibration process was found to have some limitations; convergence is not always guaranteed, analysing retrofit options with the meta-model is limited to the initial variable parameters and the meta-model introduces model error when feeding back calibrated parameters to the first-principle modelling software.

Automated calibration was used to find calibrated models within the initial search (possible input ranges) and solution space (possible outputs), computed during parametric simulation. When measured data is 'out of bounds' of the predicted solutions space, the optimisation will not be able to converge (as long as input parameters values are within the same bounds as during parametric simulation). Convergence is dependent on the base case model used during parametric simulation and the variability of its input parameters, an increase in variability increases the number of potential solutions. Preferably, the uncertainty in variables is limited, a lower uncertainty implies that the modeller has a higher level of confidence of the value of occurrence. This means that the initial manual calibration of the base case model determines the effectiveness of the automated calibration process to find calibrated solutions. In the case research this became more apparent with an increasing level of data granularity, when calibrating to just monthly energy use, the mathematical optimisation would be able to find many solutions that fit the measured data extremely well (<1% CV(RMSE)). However, when the meta-model tried to calibrate between predicted and measured monthly energy end-uses, it often could not converge, as the measured data was simply out of bounds. To increase the likelihood of convergence, seasonal variation factors were introduced (based on the seasonal occupancy variation), that adjusted the lighting and power schedules on a monthly basis, but even then, convergence was not always ascertained.

The meta-models were trained on inputs and outputs from the parametric simulations using the first-principle model (EnergyPlus), after training, the meta-models can predict new sets of inputs with remarkable accuracy. However, because the meta-model is based on an initial set of parameters, it is unable to compute new parameters unknown to the meta-model, which for the purposes of assessing efficiency measures for retrofitting, is a limitation. Nevertheless, it can be used for analysing a large amount of input parameters, such as coefficients of performance of systems, material properties, equipment power density, and many others. Alternatively, the calibrated set of parameters from the meta-model can be used in the first-principle software for further analysis. However, it was found that feeding back the calibrated inputs into the first-principle model introduced a model error (<1% CV(RMSE) for monthly energy use). In other words, when the calibrated input parameters from the meta-model were simulated with the first-principle software, it showed slight variations in predicted monthly energy use to that predicted by the meta-model. Although small, the calibrated inputs were then less accurate in predicting measured energy use. This model error occurs because the trained meta-models were not exact in their predictions.

The accuracy of an increase in data granularity was assessed using meta-model optimisation. The increase in accuracy of using manual calibration at a higher level of data granularity is difficult to determine without bias, therefore automated calibration clarified how different convergence criteria in data granularity will mask some of the energy end-uses and can lead to inaccurate calibration results. Optimisation towards lower levels of data granularity finds more solutions that are considered calibrated, but were more likely to mask the deeper levels of data granularity. For example, when calibrating towards yearly end-uses, the calibration accuracy at that level was easily obtained. After several evolutionary generations tight calibration criteria of CV(RMSE) $<0.01\%$ between predicted and measured energy use were achieved. However, using the calibrated input parameters in the first-principle model showed that monthly energy end-uses were not necessarily accurate. The highest levels of accuracy were obtained when calibrating for monthly energy end-uses, however this in turn meant that sometimes the yearly end-uses were not necessarily as accurate, as convergence at the monthly level was not always ascertained.

Although only included in the manual calibration process, automated model calibration towards environmental parameters and system performance in addition to energy use is likely to improve prediction accuracy, but with an increase in complexity. Research into the feasibility and importance of calibrating towards different building performance aspects would be beneficial. Especially for future integration of building energy models in existing building and real-time performance forecasting, in addition to retrofit engineering to understand which calibration performance criteria are important to replicate a real situation, while taking into account complexity and feasibility.

9 CONCLUSION

The aim of this research was to quantify and mitigate the energy performance gap and its underlying causes, achieved by fulfilling several interrelated objectives. Existing knowledge on the energy performance gap was reviewed, a classification of different gaps was identified, typical underlying causes were summarised and several suggestions were made in regards to reducing the discrepancy between predicted and measured energy use. In addition, it identified model calibration as a way to investigate predicted and measured energy use in order to further understand why and how the performance gap can be mitigated, it also realised there was a need to further improve model calibration capabilities, in particular concerning its complexity and accuracy. One of the main findings from reviewing existing literature was that differences between predicted and measured energy use are caused by factors existent in different stages of the building life cycle, related to activities performed by different stakeholders in the construction industry. As such, an exploratory study of industry perspectives on delivering reliable building performance identified common barriers, suggested how such barriers can be overcome and in what way stakeholders need to be engaged. Instead of focussing on energy, it looked at generic building performance, including any stakeholder incentives for building procurement. The remaining objectives in this engineering doctorate involved conducting case research into four existing non-domestic buildings. A methodology was developed to quantify the discrepancy between predicted and measured energy use, utilising model calibration techniques and extensive data collection. This methodology was applied to the case study buildings in order to investigate discrepancies and their underlying causes to build on the existing body of knowledge. Finally, it determined the importance of collecting a high level of operational data granularity to ensure accuracy of building model calibration.

9.1 Main findings

9.1.1 Literature review

Predicted and measured energy use has been shown to deviate significantly, also termed 'the performance gap'. This gap can be classified as a difference between compliance and measured energy use (the regulatory gap), but also as a gap performance modelling and measured energy use (the static performance gap) or calibrated prediction with measured energy use with longitudinal perspective (the dynamic performance gap). Literary sources were reviewed that quantify the magnitude of the regulatory performance gap to be +35% with a standard deviation of 55%, based on 62 buildings. The main underlying causes for the performance gap were related to be from specification uncertainty in building modelling, occupant behaviour and poor practice in operation, with an estimated effect of 20-60%, 10-80% and 15-80% on energy use respectively.

The literature review identified different types of performance gaps. It explains how the regulatory performance gap is a complicated issue where many facets of the building procurement and delivery process have an impact. To reduce this gap, key measures for further work and research by the building industry needs to be established:

- Accessible energy data is required for a continued gathering of evidence on the energy performance gap, this can be established through collaborative data gathering platforms.
- Legislative frameworks set limits for predicted performance and penalize buildings for high operational energy use. More effectively, however governments should relate predicted to measured performance through predictive modelling and in-use regulation. Furthermore, it should consider mandating the disclosure of design stage calculations and assumptions as well as operational energy use outcomes in Building Regulations.
- Monitoring and data analysis of operational building performance is imperative to driving change and management in operation. While well-defined assumptions need to be established through detailed calibration studies identifying the driving factors of energy use in buildings.

Understanding and mitigating differences between predicted and measured energy use requires an expansion of research efforts and focus on its underlying causes. Detailed energy audits and model calibration are invaluable techniques in order to quantify these causes. Furthermore, tools are necessary to support intuitive visualisation and data disaggregation to display energy uses at detailed levels and for different time granularity, comparing predicted and measured energy use taking a longitudinal approach. This would help in better understanding typical profiles of use and discrepancies at a higher level of granularity, whereas most current building simulation software solely provide an overview of energy use at a monthly level or as a time series (half-hourly/hourly) without the ability to further analyse the statistics of the information, such as typical profiles. Currently, such analysis would have to be done separate to the building simulation software.

9.1.2 Industry perspective of building performance

An exploratory study was carried out in partnership with the UK Green Building Council in order to answer the question, "How can the construction industry deliver better building performance and more reliable outcomes?" A group of industry experts was brought together to seek out and highlight process improvements that design, construction and property communities, as well as occupiers, might adopt deliver building which perform more predictably in operation. Through semi-structured interviews and round-table discussions with industry experts, supported by desk-based research, behaviour and process were examined across the built environment that affect building performance.

Building performance was defined as the incentives for industry stakeholders procure buildings, which for a capital provider could mean a return on investment and yield, for a designer it is the provision of a safe, resilient and sustainable building with great aesthetics, and for facilities management this is the assurance of an operable building in which occupants can be comfortable. Economic, environmental and social aspects in building procurement can complement each other, creating a business case for delivering high building performance. However, the construction industry is fragmented and incentives are rarely aligned, demonstrating and communicating the business case is necessary to transition the market to delivering better buildings.

The study identified many different barriers and gaps that need to be overcome and structured necessary changes around five key success factors that need to be driven through the different stages of the building life cycle in order to deliver better building performance. (1) Aspiration, set targets for final outcomes of the project. Expectations and requirements need to be clear from the start and can then be driven throughout the project delivery supply chain, aiming to tie in different stakeholders intended to limit fragmentation. (2) Control, procurement methods have been found to considerably affect inter-relationships between stakeholders and the delivery of building performance, the construction industry has started moving away from traditional methods (design-bid-build) to design-build and on towards more collaborative frameworks. One of those is collaborative procurement/contracting, which promotes teams to 'work together' and aims to find the best solution that creates and shares values for all parties involved. (3) Design for performance, with a target set out, there is a need to understand what is necessary to achieve it. This requires going beyond compliance. Performance needs to be incorporated in early design and actual use needs to be taken into account to alleviate later changes in the design process, but more importantly, the building needs to operate according predictions. (4) Feedback, reciprocal links need to be established between stakeholders, facilities management and design, but also to building occupiers. Feedback of performance is essential and needs to be measured, verified and demonstrated. Raising awareness, understanding of operational features, improving commissioning processes, and informing new designs are essential to reliable building performance. (5) Knowledge, and skills need to be improved in many facets of the construction industry. Enhancing understanding of the role that each part of the supply chain plays, not only within stakeholder groups, but also across them, will facilitate more meaningful conversations on how to achieve reliable building performance. Developers/owners need to be engaged to adopt new method for capturing a property's value and understand life-cycle benefits of a highly performing asset. Operational staff need to be given resources to carry out the necessary activities to attain a well-performing building, be it training or actual investments for effective operations. Tenants need to be educated in using the building and its systems to satisfy their own needs of health, comfort and safety. On-site workmanship needs to be trained to increasing levels of complexity in building construction. Finally, information needs to be in the right language, several people that were interviewed indicated that those engaged in the building performance debate tend to speak to each other rather than the whole industry.

As a continuation of the UKGBC work, the case research; (1) provided comprehensive examples of measured energy performance of existing buildings compared to detailed performance modelling (i.e. demonstrating performance and measurement and verification); (2) investigated the effect of simplifications of modelling assumptions on energy performance predictions and explains the limitations of regulations in the context of the energy performance gap; (3) explored the use of detailed measurements to identify patterns in energy use and to inform design assumptions (i.e. feedback). It demonstrates how energy performance can be better predicted during the design stage by proposing techniques for incorporating uncertainty in assumptions and more representative assumptions of actual use.

9.1.3 Case research

A methodology was developed and applied to compare and mitigate predicted and measured energy performance at a high level of data granularity in four non-domestic buildings. The methodology involved the following activities; collection of operational data and building information through building audits, synthesis of information and development of a data hierarchy, utilising operational data to inform building performance simulation assumptions, sensitivity and uncertainty analysis to understand how input parameters affect the outputs, mitigating differences between predicted and measured energy use through manual calibration, quantifying the impact of underlying causes of the regulatory performance gap, utilising meta-model based optimisation to quantify the impact of data granularity on model calibration accuracy. The methodology is based upon previous research and extended it by introducing several processes and techniques within the model calibration process. First, the parametric modelling process uses variable scheduling of occupancy presence, equipment- and lighting loads, to allow for automated adjustment of schedules to improve the accuracy of the calibration model. Second, the methodology introduced seasonal variability in profiling and uncertainty in heating and cooling setpoints in addition to uncertainty in typical static parameters. This allowed for replicating seasonal trends of an existing building, important for certain building types (e.g. university buildings). Finally, the methodology focussed on achieving a high level of data granularity and describes a hierarchy for comparing performance data.

Up to date O&M manuals that include as-built design and commissioning data supported by operational data of building performance are essential to improve the accuracy of building calibration models. It was found that as-built design information needs to be verified through energy audits; systems are often not functioning as expected, the metering strategy in schematics deviates due to mislabelling and intermittent adjustments, where space layouts may have changed and system components (FCUs, lighting, etc.) may have been added or replaced. Especially in older buildings design data is unreliable, where newer buildings have more up-to-date drawings and schematics, with only slight deviations from the original design. The collection of O&M design data and carrying out of an energy audit are a necessity in order to ensure a good level of understanding of the building, its processes and operation. This is strengthened by the analysis of operational data, collected at a high level of data granularity from; sub-metering systems, building management system and data from other sources, such as environmental sensors, Wi-Fi routers and swipe-card access data to quantify occupancy presence in a building. In the case research, data on these aspects was collected, analysed, and utilised to inform building performance simulation assumptions and verification of design strategies. Both electricity use and Wi-Fi data were used to understand the occupancy presence in the buildings to inform occupancy-, equipment- and lighting schedules.

Occupancy presence data, when contrasted with energy use, showed that occupants have a strong influence on energy use. In particular, on lighting and power (r^2 = 0.86 and 0.82 for CH and MPEB respectively) and less so on systems (r^2 = 0.22 and 0.43 for CH and MPEB respectively). Typical weekday and weekend day schedules were created based on occupancy presence and then manipulated using electricity baseloads for lighting and power to create separate schedules. Additionally, the process incorporated a horizontal variability to account for daily variations and differences between occupancy and equipment and lighting loads. The typical weekday occupancy profiles that were created based on the data were nearly identical for CH and MPEB, indicating that perhaps they may prove to be very similar in other buildings as well. It is suggested that a database is established based on this type of occupancy data for different buildings and building types, this can then in the future prove an essential source of information for both the calibration of models, but also for performance modelling purposes in the design stage of a building. Finally, system operation and performance was validated through analysing BMS data, such as supply-and return air temperatures of AHUs and operational setpoints on FCUs. Furthermore, space temperatures were analysed and compared to set-point temperatures to inform the heating- and cooling strategy.

The case study buildings were compared to typical, albeit outdated, benchmarks, and did not provide a meaningful comparison. It highlights the need for more comprehensive benchmarking techniques. Understanding if buildings fall within typical or good benchmarks as presented is not straightforward due to internal differences such as additional services; large server rooms in MPEB or laboratory type spaces within a building. The office buildings were more similar to the benchmarks (as the benchmarks are also intended for offices), but determining their efficiency in operation or performance against the stock was difficult. As such, their electricity patterns were also compared at a higher level of data granularity by computing their representative load pattern and other load shape metrics. These proved to be more effective in identifying operational differences within the buildings. Total energy use baseloads were significantly different, from high to low, MPEB, CH and Office 17 and 71, with baseloads of about 60%, 45%, 30% and 10% respectively. In addition, on-hour duration for MPEB was determined to be around 13 hours, while the other buildings fluctuate around 16-18 hours. This was mainly due to the steep difference in energy use between the day and night in MPEB, whereas the profiles for the other buildings are much smoother and therefore calculate longer on-hour durations.

Manual calibration was an essential part of the research methodology, as it gave an understanding of common differences between predicted and measured energy use and its underlying causes. It was evident that many solutions or configurations of input parameters or scheduling of operational processes can deem a model to be calibrated. However, this was narrowed down through collecting data at a high level of granularity so as to filter out incorrect solutions. Through the calibration process, it became clear that the reliability of a model is strongly dependent on the availability and accuracy of both design and measured data, which is the evidence to support any iterative changes. A higher level of data granularity proved essential in understanding and implementing correct operational schedules and HVAC strategies.

Parametric simulations were employed in combination with uncertainty and sensitivity analysis to assess the sensitivity of input parameters on different energy end-uses and total energy use. It became clear that in all buildings, equipment power density was one of the more significant parameters to influence energy use, typical for non-domestic buildings. It is therefore important to make evidence based assumptions about these loads to make sure predictions are in line with measurements. For MPEB specifically, the large server rooms contribute to a significant proportion of total energy use and are the main driver of energy use within the building as they also affect system energy use. Other important factors are the heating and cooling set points, in particular in spaces with high internal gains. Insignificant parameters to influence energy use are material properties, in contrast to domestic buildings. For non-domestic buildings there is an unintuitive trade-off between the thermal performance of the envelope and energy use. An improvement of the conductivity (decrease in U-value) of the envelope can have a negative effect on energy use due to high internal gains in certain spaces, which inhibit heat loss to the outside and therefore have an increased cooling load. Similar to the conductivity of materials, infiltration did not show to have a significant effect on energy use in any of the buildings, although there is a distinct difference between the seasons.

The manually calibrated models were used to quantify the impact of simplifications (i.e. typical NCM assumptions) on the regulatory performance gap. The simplifications concerned with a change of internal gains highlight the importance of accurately determining the assumptions for equipment and lighting power density in spaces, as they have a large effect on total energy use. In particular, server loads can be dominant in modern buildings, in MPEB energy use was strongly affected by the power use of the computer clusters, indirectly influencing systems energy use. Defining the right use-schedules for these loads is tantamount to establishing the right load assumptions for these spaces. As part of the manual calibration process, it was found that high baseloads exist in the buildings, which had to be accounted for by adjusting the internal gains schedules. These were a large contributor to a discrepancy between regulatory predictions and measurements as there was a significant difference to the typically assumed NCM equipment power baseload. In the case study buildings, the baseloads for equipment in Office 17, 71, CH and MPEB were ~25%, ~20%, ~65%, ~85% respectively, compared to the NCM assumption of 5.3%. Besides internal gains, the heating and cooling temperatures in different space types can vary significantly from that initially assumed to that in operation, without detailed temperature or BMS data, this would be difficult to replicate within a model. Moreover, in CH the operational set-point are manually adjustable, making it even more difficult to implement static strategies within a model. Replacing the calibrated set-point temperatures with NCM assumptions led to a significant decrease in energy use for Office 17 and CH, while for Office 71 the temperatures were similar to NCM assumptions. Finally, the weather file based on historical weather data in London was replaced by a design weather file from Gatwick (simplification #10), which had a minor, but notable effect in increasing total energy use.

Manual calibration achieved the statistical criteria that according to available guidelines deem a model to be calibrated, where automated calibration further mitigated differences between predictions and measurements to CV(RMSE) < 1%. Automated calibration at different levels of data granularity showed that calibrating to monthly energy end-use predictions compared to yearly end-uses and monthly energy use increased model accuracy by 23.5% and 11% respectively. Nevertheless, energy model calibration is a complex and time-consuming process. Automation alleviated some of the time-intense tasks, but also added another level of complexity. In terms of effectiveness, the automated calibration process was found to have some limitations; convergence is not always guaranteed, analysing retrofit options with the meta-model is limited to initial variable parameters, a meta-model introduces model error when feeding back calibrated parameters to the first-principle physics based software. Automated calibration can only find calibrated models within the initial search (possible input ranges) and solution space (possible outputs), computed during parametric simulation. The meta-model used for automated calibration is based on an initial set of parameters, it is therefore unable to compute new parameters unknown to the model, which for the purposes of retrofit is a limitation. Feeding back calibrated inputs into the first-principle physics based model introduced a model error <1% CV(RMSE) for monthly energy use. Finally, tight control and understanding of input parameters remains essential, even though calibration might obtain values and models that accurately fit measured data, it does not necessarily mean they are representing the real situation. Nevertheless, this was alleviated by increasing the level of data granularity, which in turn increases the accuracy of input parameters.

Building performance modelling is a time-intense process, further protracted by the exhaustive nature of building model calibration, the verification of design data and analysis of operational data is helpful in improving model calibration accuracy, but at a cost. Automating some of the underlying processes can improve its time efficiency, further improved by prioritising tasks in order of significance on increasing model accuracy. Through the process of calibrating four full-scale building energy models to measured existing performance, the author found that the accuracy of models is strongly dependent on the granularity of energy data collected. At a hierarchical level, calibrating towards total energy use masks the underlying end-uses significantly, while at a temporal level, sub-hourly data can inform on typical schedules of use to represent the actual processes more accurately than when only monthly data is available.

Ultimately, the necessary level of detail of building calibration models is dependent on the purpose of calibration. For example, (1) if the model is needed to make recommendations of the indoor environment of a museum exhibition area for a close-control strategy, model calibration should focus on calibrating to space temperatures and humidity. While assumptions concerning material properties, control systems and people presence in this space need to be captured at a high level of detail. However, at this level of detail, such information is typically not directly available from built-in metering systems and would need additional sensors. In contrast, (2) if the purpose is to understand the impact of energy use by changing from a natural ventilation to comfort cooling strategy, then calibration towards energy uses should be at a high granularity.

9.1.4 Contributions to knowledge

This thesis and the author made three contributions to knowledge.

First, in partnership with the UKGBC an investigation was carried out to understand how high building performance can be delivered throughout the building life cycle. It did this through qualitative data collection in the form of 15 semi-structured interviews with building construction industry experts, including architects, developers, investors, project managers, contractors, owner/occupiers and those in facilities management. In addition, round-table discussions were held with other industry experts, which formed a 'task-group' that aimed to explore different aspects in the building life cycle with regards to building performance, by discussing results from interviewees, desk-based research and their expert knowledge. Common barriers to delivering building performance were identified. It developed five key factors which should be adhered to in order to deliver reliable building performance. This research was published in an industry-focussed report by the UKGBC.

Second, case research was conducted to mitigate differences between predicted and measured energy use in four existing buildings, while quantifying the impact of typical underlying causes of the regulatory performance. The case research highlighted common issues in data collection and performance prediction. Operational data was collected and used to inform building performance simulation, where sensitivity analysis and manual calibration mitigated differences between predictions and measurements. It furthermore explored the effects of building parameter assumptions and uncertainties on predictions and indirectly the energy performance gap. This contribution was support by developing a calibration methodology that compared predicted and measured energy use in existing buildings, for the application of both manual and automated calibration to mitigate discrepancies. It introduced several parametric techniques that improved the calibration process over previous research.

Third, the effects of data granularity on model calibration accuracy was quantified through the application of meta-model based optimisation in three of the four case study buildings. Building on previous research in the area of model calibration, it sought out to understand how the data granularity affects model accuracy. Such knowledge is useful as it determines the relationship between model accuracy and need for quantitative data and information, establishing a trade-off between time consumption, cost and accuracy. Subsequent to the second contribution, which used manual calibration processes, meta-models were created based on the relationships between inputs and outputs created through parametric simulation. Higher levels of accuracy were achieved through using these meta-models for optimisation (i.e. automated calibration), but this was found to introduce additional complexity.

9.2 Further work

The main research findings alleviate some of the limitations found in model calibration, while reiterating others that require further research for improvement. The author recognises that further improvements can be made to increase the calibration accuracy of building energy models and recommends that further research is beneficial in order to; effectively utilise calibrated models in retrofitting, advance automation in developing calibrated models, improve automated calibration techniques, and refine the use of machine learning in supporting calibrated models. In particular, the following areas for further work were identified:

- There is a need for further development of more stringent and specific statistical measures for the acceptance criteria of calibrated models set out in different standards. In particular, they should account for higher levels of hierarchical data granularity, as a minimum, separating lighting, power and system energy use. Garret and New (2016) provide some background into the suitability of these criteria. Although, the availability of data at a high level of granularity is difficult to obtain in many existing buildings, recent regulatory changes require building energy data to be collected and disaggregated.
- Model calibration in this research has mainly focussed on energy use, but can be extended to systems, see (Yin, et al., 2016) and environmental performance, see (Royapoor & Roskilly, 2015). This is another level of model validation, making them better at representing reality. In addition, when this level of data granularity is not accurately represented in the model, it is unlikely to accurate predict conservation measures that are related to specific system changes or changes in their operational control. Research into the importance of these different data sources and their effect on model accuracy are necessary to understand how they increase model calibration accuracy.
- Full-scale building models were constructed in this research to perform an in-depth comparison of energy use at a high level of data granularity. However, in practice, such time-consuming work is infeasible as it takes weeks or even months to collect all the necessary data and build detailed energy models. Efforts at reducing the workload could look into the simplification of this process depending on the purpose of the calibrated model, several suggestions are made:
    - Research into the effect of simplifying the model by replicating similar zones and floors on accuracy of the model.
    - Software development to include features that support parametric simulation and automated calibration, which in many cases are already under way.
    - Building Information Modelling (BIM) could potentially make a positive difference in this area, as it aims to integrate information flows between different disciplines through all stages of the building life cycle, increasing efficiency and accuracy of model development.
- Linking calibrated models to operational buildings for real-time forecasting of performance, and analysis of control strategies. Some research was done by O'Neill et al. (2014) and Pang et al. (2012), however research in this field is still in its early stages and is primarily limited to the availability and quality of data. Which is likely to change in the near future due to the internet of things and need for improved energy efficiency, health and wellbeing in buildings.

- The use of machine learning algorithms in building performance forecasting. In contrast to linking calibrated models with existing building for real-time forecasting, they can be used to learn and forecast the behaviour in the buildings solely based on measured building performance, including system performance, occupancy presence, weather data and energy performance. This avoids the need for building energy modelling, but will need high quality data at a high level of granularity.

10 REFERENCES

[1] Ahmed, M. & Culp, C., 2006. Uncalibrated building energy simulation modeling results. *HVAC&R Research,* 12(4), pp. 1141-1155.

[2] Ahuja, S. & Peles, S., 2013. *Building energy models: Quantifying uncertainties due to stochastic processes.* Florence, Italy, IEEE Conference on Decision and Control.

[3] ASHRAE, 2002. *Measurement of Energy and Demand Savings,* Atlanta, U.S.: American Society of Heating, Refrigeration and Air-Conditioning Engineers.

[4] ASHRAE, 2004. *ASHRAE Standard 90.1: Energy standard for buildings except low-rise residential buildings,* Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc..

[5] ASHRAE, 2013. *ASHRAE Fundamentals.* Atlanta, U.S.: ASHRAE.

[6] Atink, R., Garrigan, C., Bonetti, M. & Westaway, R., 2014. *Greening the building supply chain,* Nairobi, Kenya: UNEP, Sustainable Buildings and Climate Initiative.

[7] Austin, B., 2013. *The performance gap - causes and solutions,* London, UK: Green Construction Board - Buildings Working Group.

[8] Azar, E. & Menassa, C., 2012. A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings. *Energy and Buildings,* Volume 55, pp. 841-853.

[9] Bertagnolio, S., Randaxhe, F. & Lemort, V., 2012. *Evidence-based calibration of a building energy simulation model: Application to an office building in Belgium.* Manchester, UK, International Conference for Enhanced Building Operations ICEBO 2012.

[10] Bhandari, M., Shrestha, S. & New, J., 2012. Evaluation of weather datasets for building energy simulation. *Energy and Buildings,* Volume 49, pp. 109-118.

[11] Bloomfied, D., 1988. *An investigation into analytical and empricial validation techniques for dynamic thermal models of buildings,* Watford, UK: BRE (Building Research Establishment, U.K.).

[12] Bordass, B., Leaman, A. & Ruyssevelt, P., 2001. Assessing building performance in use 5: conclusions and implications. *Building Research & Information,* 29(2), pp. 144-157.

[13] Bou-Saada, T. & Haberl, J., 1995. *An improved procedure for developing calibrated hourly simulation models.* Montreal, Canada, IBPSA Proceedings.

[14] Brown, N., Wright, A., Shukla, A. & Stuart, G., 2010. Longitudinal analysis of energy metering data from non-domestic buildings. *Building Research & Information,* 38(1), pp. 80-91.

[15] Bucking, S., Zmeureanu, R. & Athienitis, A., 2014. A methodology for identifying the influence of design variations on building energy performance. *Journal of Building Performance Simulation,* 7(6), pp. 411-426.

[16] Bunn, R. & Burman, E., 2015. *S-Curves to model and visualise the energy performance gap between design and reality - first steps to a practical tool.* London, UK, University College London.

[17] Burhenne, S., Jacob, D. & Henze, G., 2011. *Sampling based on Sobol' sequences for Monte Carlo techniques applied to building simulations.* Sydney, Australia, 12th Conference of International Building Performance Simulation Association.

[18] Burman, E., 2016. *Assessing the operational performance of educational buildings against design expectations - a case study approach,* London, UK: University College London.

[19] Burman, E., Mumovic, D. & Kimpian, J., 2014. Towards measurement and verification of energy performance under the framework of the European directive for energy performance of buildings. *Energy,* 77(1), pp. 153-163.

[20] Burman, E., Rigamonti, D., Kimpain, J. & Mumovic, D., 2012. *Performance gap and thermal modelling: a comparison of simulation results and actual energy performance for an academy in north west england.* Loughborough, UK, BSO12.

[21] Calderone, A., 2011. *Simulation versus reality - 4 case studies.* Sydney, Australia, IBPSA Proceedings.

[22] Carbon Trust, 2012. *Closing the gap - Lesson learned on realising the potential of low carbon building design,* s.l.: s.n.

[23] CarbonBuzz, 2015. [Online]
Available at: www.carbonbuzz.org
[Accessed May 2015].

[24] Chaudhary, G. et al., 2016. Evaluation of "Autotune" calibration against manual calibration of building energy models. *Applied Energy,* Volume 182, pp. 115-134.

[25] Chegut, A., Eichholtz, P. & Kok, N., 2013. Supply, Demand and the Value of Green Buildings. *UrbanStudies,* pp. 1-22.

[26] CIBSE, 2005. *Natural ventilation in non-domestic buildings, CIBSE Applications Manual AM10.* London: The Chartered Institute of Building Services Engineers.

[27] CIBSE, 2006. *TM22: Energy Assessment and Reporting Methodology,* London, UK: Chartered Institution of Building Services Engineers.

[28] CIBSE, 2007. *Environmental Design, CIBSE Guide A.* 7th edition ed. London, U.K.: The Chartered Institute of Building Services Engineers.

[29] CIBSE, 2013. *CIBSE TM54: Evaluating operational energy performance of buildings at the design stage,* London, U.K.: The Chartered Institute of Building Services Engineers.

[30] CIBSE, 2015. *CIBSE TM57: Integrated school design,* London, U.K.: The Chartered Institute of Building Services Engineers.

[31] Clarke, J., 2001. *Energy Simulation in Building Design.* 2nd ed. s.l.:Butteworth Heinemann.

[32] Clements-Croome, D., 2014. *Intelligent Buildings: An Introduction.* Abingdon, UK: Routledge.

[33] Clevenger, C. & Haymaker, J., 2006. *The impact of the building occupant on energy modelling simulations.* Montreal, Canada, Joint International Conference on Computing and Decision Making in Civil and Building Engineering.

[34] Coakley, D., Raftery, P. & Keane, M., 2014. A review of methods to match building energy simulation models to measured data. *Renewable and Sustainable Energy Reviews,* Volume 37, pp. 123-141.

[35] Coakley, D., Raftery, P. & Molloy, P., 2012. *Calibration of whole building energy simulation models: detailed case study of a naturally ventilated building using hourly measured data.* Loughborough, UK, BSO12.

[36] Coakley, D., Raftery, P., Molloy, P. & White, G., 2011. *Calibration of a detailed BES model to measured data using an evidnce-based analytical optimisation approach,* Galway, Ireland: 12th Conference of International Building performance Simulation Association.

[37] Cohen, R., Standeven, M., Bordass, B. & Leaman, A., 2001. Assessing building performance in use 1: the Probe process. *Building Research & Information,* 29(2), pp. 85-102.

[38] Colorni, A., Dorigo, M. & Maniezzo, V., 1999. *Genetic algorithms and highly constrained problems: the time-table case.* Dormund, Proceedings of the first international workshop on parallel problem solving from nature..

[39] Commission, E., 2011. *Energy Roadmap 2050 - Communication from the commission to the european parliament, the council, the european economic and social comittee and the committee of the regions,* Brussels: European Commission.

[40] Construction Task Force, 1998. *Rethinking Construction,* London, U.K.: Department of the Environment, Transport and the Regions.

[41] Cox, A. & Townsend, M., 1997. Latham as half way house: A relational competence approach to better practice in construction procurement. *Engineering, Construction and Architectural Management,* 4(2), pp. 143-158.

[42] Crawley, D., Hand, J., Kummert, M. & Griffith, B., 2008. Contrasting the capabilities of building energy performance simulation programs. *Building and Environment,* 43(4), pp. 661-673.

[43] Daly, D., Cooper, P. & Ma, Z., 2014. Understanding the risks and uncertainties introduced by common assumptions in energy simulations for Australian commercial buildings. *Energy and Buildings,* Volume 75, pp. 382-393.

[44] Davis III, J. & Nutter, D., 2010. Occupancy diversity factors for common university building types. *Energy and Buildings,* Volume 42, pp. 1543-1551.

[45] de Wilde, P., 2014. The gap between predicted and measured energy performance of buildings: a framework for investigation. *Automation in Construction,* Volume 41, pp. 40-49.

[46] de Wilde, P., Tian, W. & Augenbroe, G., 2011. Longitudinal prediction of the operational energy use of buildings. *Building and Environment,* Volume 46, pp. 1670-1680.

[47] de Wit, S., 2001. *Uncertainty in predictions of thermal comfort in buildings,* Delft, the Netherlands: Delft University of Technology.

[48] de Wit, S. & Augenbroe, G., 2002. Analysis of uncertainty in building design evaluations and its implications. *Energy and Buildings,* Volume 34, p. 9510958.

[49] Deb, K., 1999. Multi-objective genetic algorithms: problem difficulties and construction of test problems. *Evolutionary Computation,* 7(3), pp. 205-230.

[50] DECC, 2013. *Energy Consumption in the UK,* s.l.: Department of Energy & Climate Change.

[51] Department of the Environment, 2003. *Energy efficiency in offices Energy Consumption Guide ECG019,* London, UK: Energy Efficiency Best Practice programme.

[52] Diamond, R. et al., 2006. Evaluating the energy performance of the first generation of LEED-certified commerical buildings. *Proceedings of the 2006 American Council for an Energy-Efficient Economy Summer Study on Energy Efficiency in Buildings*.

[53] Eisenhower, B., O'Neill, Z., Fonoberov, V. & Mezic, I., 2011. Uncertainty and sensitivity decomposition of building energy models. *Journal of Building Performance Simulation,* 5(3), pp. 171-184.

[54] Eisenhower, B. et al., 2012. A methodology for meta-model based optimization in building energy models. *Energy and Buildings,* Volume 47, pp. 292-301.

[55] Ellison, L. & Sayce, S., 2007. Assessing sustainability in the existing commerical property stock: Establishing sustainability criteria relevant for the commerical property investment sector. *Property Management,* 25(3), pp. 287-304.

[56] Evins, R., 2013. A review of computational optimisation methods applied to sustainable building design. *Renewable and Sustainable Energy Reviews,* Volume 22, pp. 230-245.

[57] EVO, 2009. *International Performance Measurement & Verification Protocol; Concepts and options for determing energy and water savings Volume I,* Oak Ridge, U.S.A.: The Efficiency Valuation Organization.

[58] Fabrizio, E. & Monetti, V., 2015. Methodologies and Advancements in the Calibration of Building Energy Models. *Energies,* Volume 8, pp. 2548-2574.

[59] Fedoruk, L., Cole, R., Robinson, J. & Cayuela, A., 2015. Learning from failure: understanding the anticipated-achieved building energy performance gap. *Building Research & Information.*

[60] Fellows, R. & Liu, A., 2012. Managing organizational interfaces in engineering construction projects: addressing fragmentation and boundary issues across multiple interfaces. *Construction Management and Economics,* 30(8), pp. 653-671.

[61] Finch, E. & Zhang, X., 2013. Facilities Management. In: *Design and Management of Sustainable Built Environment.* London, UK: Springer-Verlag, pp. 305-326.

[62] Fortin, F. et al., 2012. DEAP: Evoluationary Algorithms Made Easy. *Journal of Machine Learning Research,* Volume 13, pp. 2171-2175.

[63] Fowler, K., Rauch, E., Henderson, J. & Kora, A., 2010. *Re-assessing green building performance: a post occupancy evaluation of 22 GSA buildings,* Richland, Washington, U.S.: Pacific Northwest National Laboratory.

[64] Francis, J., 1976. *Supervision and Examination of Higher Degree Students,* London, UK: University of London.

[65] Frankel, M., Edelson, J. & Colker, R., 2015. *Getting to Outcome-Based Building Performance,* s.l.: s.n.

[66] Fuerst, F., 2015. *The financial rewards of sustainability: A globl performance study of real estate investment trusts,* Cambridge, UK: University of Cambridge.
[67] Garrett, A. & New, J., 2016. *Suitability of ASHRAE Guideline 14: Metric for Calibration.* Orlando, USA, ASHRAE.
[68] GCB, 2014. *Mapping the Real Estate Lifecycle for Effective Policy Interventions,* s.l.: s.n.
[69] George, C. & Garrod, E., 2017. *Getting to zero - Developing policy recommendations for a zero emission capital,* London, UK: Elementa consulting.
[70] Guest, G., Bunce, A. & Johnson, L., 2006. How many interviews are enough? An experiment with data saturation and variability. *Field Methods,* 18(1), pp. 59-82.
[71] Guyon, G., 1997. *Role of the model user in results obtained from simulation software program.* IBPSA1997, IBPSA Proceedings.
[72] Haberl, J. & Culp, C., 2005. *Review of methods for measuring and verifying savings from energy conservation retrofits to existing buildings,* Texas, USA: Energy Systems Laboratory.
[73] Hawkins, D., Mumovic, D. & Hanna, S., 2012. *Appying articial neural networks to estimate building energy use in the early stages of architectural design.* London, UK, CIBSE ASHRAE Technical Symposium 2012.
[74] HEEPI, SUST and Thirdwave, 2008. *High Performance Buildings 1. The Business Case for Universities and Colleges,* London, UK: HEEPI and SUST.
[75] HEEPI, SUST and Thirdwave, 2008. *High Performance Buildings 2. The process of Delivery for Universities and Colleges,* London, UK: Higher Education Environmental Performance Improvement, The Lighthouse on Sustainability, Thirdwave Ltd..
[76] Heidarinejad, M. et al., 2013. *Building classification based on simulated annual results: towards realistic building performance expectations.* Chambéry, France, IBPSA 2013.
[77] Henderson, V., 2011. *An evaluation of Display Energy Certificates as a policy initiative to improve energy-efficiency within public buildings in Cornwall,* Leicester: De Montfort University.
[78] Heo, Y. et al., 2015. Scalable methodology for large scale building energy improvement: Relevance of calibration in model-based retrofit analysis. *Building and Environment,* Volume 87, pp. 342-350.
[79] Heo, Y., Choudhary, R. & Augenbroe, G., 2012. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings,* April, Volume 47, pp. 550-560.
[80] Herman, J., Kollat, J., Reed, P. & Wagener, T., 2013. Technical Note: Method of Morris effectively reduces the computational demands of global sensitivity analysis distributed watershed models. *Hydrology and Earth System Sciences,* Volume 17, pp. 2893-2903.
[81] Herman, J. & Usher, W., 2017. SALib: An open-source Python library for Sensitivity Analysis. *The Journal of Open Source Software,* 2(9).
[82] HM Goverment, 2008. *Climate Change Act 2008,* London, UK: The Stationary Office Limited.
[83] HM Goverment, 2017. *The non-domestic Prvate Rented Property minimu mstandard - landlord guidance,* London, UK: Department fo Business, Energy & Industrial Strategy.
[84] HM Government, 2010. *Low Carbon Construction, Innovation & Growth Team,* London, U.K.: Department for Business, Innovation & Skills.
[85] HM Government, 2013. *The Building Regulations 2013 - L2A Conservation of fuel and power in new buildings other than dwellings,* London, U.K.: HM Government.
[86] Hogg, N. & Botten, C., 2012. *A Tale of Two Buildings - Are EPCs a true indicator of energy efficiency?,* London: Jones Lang LaSalle & BBP.
[87] Hong, S. & Steadman, P., 2013. *An analysis of Display Energy Certificates for Public Buildings, 2008 to 2012,* London: Energy Institute, University College London.
[88] Hong, T., Chou, S. & Bong, T., 2000. Building simulation: an overview of developments and information sources. *Building and Environment,* Volume 35, pp. 347-361.
[89] Hong, T. & Lin, H., 2013. *Occupant behavior: impact on energy use of private offices,* Berkeley, California, U.S.: Ernest Orlando Lawrence Berkeley National Laboratory.
[90] Hopfe, C. & Hensen, J., 2011. Uncertainty analysis in building performance simulation for design support. *Energy and Buildings,* Volume 43, pp. 2798-2805.

[91] House of Commons, 2008. *Construction matters - Ninth Report of Session 2007-08 Volume 1,* London, UK: The Stationery Office Limited.

[92] Hsu, D., 2014. How much information disclosure of building energy performance is necessary?. *Energy Policy,* Volume 64, pp. 263-272.

[93] IEA, 2010. *World Energy Outlook 2010.* s.l.:International Energy Agency.

[94] Iooss, B. & Lemaitre, P., 2015. A review on global sensitivity methods. In: G. Dellino & C. Meloni, eds. *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications.* Boston, M.A., U.S.: Springer, pp. 101-122.

[95] ISO, 2008. *ISO 13790:2008 - Energy performance of building -- Calculation of energy use for space heating and cooling,* s.l.: International Organization for Standardization.

[96] ISO, 2017. *ISO 52016-1:2107 - Energy performance of buildings -- Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads -- Part 1: Calculation procedures,* s.l.: International Organization for Standardization.

[97] Izenman, A., 2008. *Modern Multivariate Statistical Techniques - Regression, classification and manifold learning.* 1st ed. Philadelphia, USA: Springer.

[98] Jones, S. et al., 2016. *World Green Building Trends 2016: Developing Markets Accelerate Global Green Growth SmartMarket Report,* s.l.: Dodge Data & Analytics.

[99] Jonlin, D., 2014. *Bridging the energy performance gap: real-world tools.* Pacific Grove, C.A., U.S., ACEEE Summer Study on Energy Efficiency in Buildings.

[100] Judkoff, R., Wortman, D., O'Doherty, B. & Burch, J., 2008. *A methodology for validating building energy analysis simulations,* Golden, Colorado, US: National Renewable Energy Laboratory.

[101] Karlsson, F., Rohdin, P. & Persson, M.-L., 2007. Measured and predicted energy demand of a low energy building: important aspects when using building energy simulation. *Building Services Engineering Research and Technology,* 28(3), pp. 223-235.

[102] Kawamoto, K., Shimoda, Y. & Mizuno, M., 2004. Energy saving potential of office equipment power management. *Energy and Buildings,* Volume 36, pp. 915-923.

[103] Kemna, R., van Elburg, M., Li, W. & van Holsteijn, R., 2007. *Eco-design of CH-Boilers - Definitions, test standards, current legislation & measures,* Brussels, Belgium: European Commission.

[104] Kimpian, J. et al., 2014. *Getting real about energy use in non-domestic buildings.* Dublin, Ireland, CIBSE ASHRAE Technical Symposium.

[105] Kimpian, J. et al., 2014. *Getting real about energy use in non-domestic buildings.* Dublin, Ireland, s.n.

[106] Kim, S. & Augenbroe, G., 2013. Uncertainty in developing supervisory demand-side controls in buildings: A framework and guidance. *Automation in Construction,* Volume 35, pp. 28-43.

[107] Kim, Y., Heidarinejad, M., Dahlhausen, M. & Srebric, J., 2017. Building energy model calibration with schedules derived from electricity use data. *Applied Energy,* Volume 190, pp. 997-1007.

[108] Kim, Y. & Park, C., 2016. Stepwise deterministic and stochastic calibration of an energy simulation model for an existing building. *Energy and Buildings,* Volume 133, pp. 455-468.

[109] Kleber, M. & Wagner, A., 2007. *Results of monitoring a naturally ventilated and passively cooled office building in Frankfurt a.M., Germany.* Helsinki, Finland, Proceedings of Clima 2007 Wellbeing Indoors.

[110] Knight, I., Stravoravdis, S. & Lasvaux, S., 2008. Predicting operational energy consumption profiles - finding from detailed surveys and modelling in a UK educational building compared to measured consumption. *International Journal of Ventilation,* 7(1).

[111] Korjenic, A. & Bednar, T., 2012. Validation and evaluation of total energy use in office buildings: A case study. *Automation in Construction,* Volume 23, pp. 64-70.

[112] Korkmaz, S., Horman, M., Molenaar, K. & Gransberg, D., 2010. *Influence of Project Delivery Methods on Achieving Sustainable High Performance Buildings,* Michigan, USA: Charles Pankow Foundation.

[113] Li, B. & Lim, D., 2013. Occupant behavior and building performance. In: *Design and Management of Sustainable Built Environments.* London, UK: Springer-Verlag, pp. 279-304.

[114] Lim, C. & Mohamed, M., 1999. Criteria of poject success: and exploratory re-examination. *International Journal of Project Management,* 17(4), pp. 243-248.

[115] Luo, X., Hong, T., Chen, Y. & Piette, M., 2017. Electric load shape benchmarking for small- and medium-sized commercial buildings. *Applied Energy,* Volume 204, pp. 715-725.

[116] Macdonald, I., 2002. *Quantifying the Effects of Uncertainty in Building Simulation,* Glasgow, U.K.: University of Strathclyde.

[117] Mahdavi, A. & Pröglhöf, C., 2009. *User behavior and energy performance in buildings.* Vienna, Austria, IEWT.

[118] Maile, T., 2010. *Comparing measured and simulated building energy performance data,* Stanford, CA, U.S.: Standford University.

[119] Maile, T., Bazjanac, V. & Fischer, M., 2012. A method to compare simulated and measured data to assess building energy performance. *Building and Environment,* October, Volume 56, pp. 241-251.

[120] Maile, T., Fischer, M. & Bazjanac, V., 2010. *Formalizing assumptions to document limitations of building performance measurement systems,* Standford, CA, U.S.: CIFE Working Paper.

[121] Manfren, M., Aste, N. & Moshksar, R., 2013. Calibration and uncertainty analysis for computer models - A meta-model based approach for integrated building energy simulation. *Applied Energy,* Volume 103, pp. 627-641.

[122] Mara, T. & Tarantola, S., 2008. Application of global sensitvity analysis of model output to building thermal simulations. *Building Simulation,* Volume 1, pp. 390-302.

[123] Martani, C. et al., 2012. ENERNET: Studying the dynamic relationship between building occupancy and energy consumption. *Energy and Buildings,* Volume 47, pp. 584-591.

[124] Masoso, O. & Grobler, L., 2010. The dark side of occupants' behaviour on building energy use. *Energy and Buildings,* Volume 42, pp. 173-177.

[125] Menezes, A., 2013. *Improving predictions of operational energy performance through better estimates of small power consumption,* London: Loughborough University.

[126] Menezes, A., Cripps, A., Bouchlaghem, D. & Buswell, R., 2012. Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. *Applied Energy,* Volume 97, pp. 355-364.

[127] Mills, E., 2011. Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions in the United States. *Energy Efficiency,* 4(2), pp. 145-173.

[128] Moezzi, M., Hammer, C., Goins, J. & Meier, A., 2013. *Behavioral strategies to bridge the gap between potential and actual savings in commercial buildings,* s.l.: Air Resources Board.

[129] Morant, M., 2012. *The performance gap - non domestic building: final report (CEW1005),* Cardiff: AECOM.

[130] Mulville, M., Jones, K. & Heubner, G., 2014. The potential for energy reduction in UK commercial offices through effective management and behaviour change. *Architectural Engineering and Design Management,* 10(1-2), pp. 79-90.

[131] Murphy, E. & Castleton, H., 2014. *Predicting the energy performance of a new low carbon office building: a case study approach.* London, UK, Sheffield University.

[132] NCM, 2013. *National Calculation Methodology (NCM) modelling guide (for buildings other than dwellings in England),* London, U.K.: Department for Commnities and local Government.

[133] Neymark, J. et al., 2002. Applying the building energy simulation test (BESTEST) diagnostic method to verification of space conditioning equipment models used in whole-building energy simulation programs. *Energy and Buildings,* Volume 34, pp. 917-931.

[134] Nguyen, A., Reiter, S. & Rigo, P., 2014. A review on simulation-based optimization methods applied to building performance analysis. *Applied Energy,* Volume 113, pp. 1043-1058.

[135] Norford, L., Socolow, R., Hsieh, E. & Spadaro, G., 1994. Two-to-one discrepancy between measured and predicted performance of a 'low-energy' office building: insights from a reconciliation based on the DOE-2 model. *Energy and Buildings,* Volume 21, pp. 121-131.

[136] Oberkampf, W. & Roy, C., 2010. *Verification and validation in scientific computing.* s.l.:Cambridge University Press.

[137] OECD, 2012. *OECD Green Growth Studies - Energy,* Paris, France: OECD Publishing.

[138] Olivier, D., 2001. *Building in ignorance - demolising complacency: improving the energy performance of 21st century homes,* s.l.: s.n.

[139] O'Mara, M. & Bates, S., 2012. *Why invest in high-performance green buildings? Utilizing smart building strategies to reduce environmental impact, ensure occupant satisfaction, drive better business outcomes, and maximize asset value,* London, U.K.: Schneider Electric.

[140] O'Neill, Z. & Eisenhower, B., 2013. Leveraging the analysis of parametric uncertainty for building energy model calibration. *Building Simulation,* 6(4), pp. 365-377.

[141] O'Neill, Z. et al., 2014. Model-based real-time whole building energy performance monitoring and diagnostics. *Journal of Building Performance Simulation,* 7(2), pp. 83-99.

[142] Oreszczyn, T. & Lowe, R., 2009. Challenges for energy and buildings research: objectives, methods and funding mechanisms. *Building Research & Information,* 38(1), pp. 107-122.

[143] Palmer, J., Terry, N. & Armitage, P., 2016. *Building Performance Evaluation Programme: Findings from non-domestic projects - Getting the best from buildings,* London, U.K.: Innovate UK.

[144] Pang, X., Wetter, M., Bhattacharya, P. & Haves, P., 2012. A framework for simulation-based real-time whole building performance assessment. *Building and Environment,* Volume 54, pp. 100-018.

[145] Pan, Y., Huang, Z. & Wu, G., 2007. Calibrated building energy simulation and its application in a high-rise commercial building in Shanghai. *Energy and Buildings,* 39(6), pp. 651-657.

[146] Parys, W., Saelens, D. & Hens, H., 2010. *Implementing realistic occupant behavior in building energy simulatios - the effect on the results of an optimization of office buildings.* Antalya, Turkey, Proceedings of the 10th REHVA World Congress" Sustainable Energy use in Buildings.

[147] Pedregosa, F. et al., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research,* Volume 12, pp. 2825-2830.

[148] Pegg, I., Cripps, A. & Kolokotroni, A., 2007. Post-occupancy performance of five low-energy schools in the UK. *ASHRAE Transactions.*

[149] Peles, S., Ahuja, S. & Narayanan, S., 2012. *Uncertainty quantification in energy efficient building performance simulations.* Purdue, s.n.

[150] Phillips, E. & Pugh, D., 1994. *How to get a PhD - a handbook for students and their supervisors,* London, UK.: Open University Press.

[151] Pianosi, F. et al., 2016. Sensivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software,* Volume 79, pp. 214-232.

[152] Piette, M., Nordman, B., deBuen, O. & Diamond, R., 1994. *Over the energy edge: results from a seven year new commercial buildings research and demonstration project.* Berkeley, CA, USA, American Council for an Energy Efficient Economy (ACEEE) Summer Study.

[153] Pless, S. et al., 2012. *How-To Guide for Energy-Performance-Based Procurement,* Golden, CO, U.S.: National Renewable Energy Laboratory.

[154] Preiser, W. & Vischer, J., 2005. *Assessing Building Performance.* Burlington, MA: Elsevier.

[155] pyDOE, 2017. *pyDOE Design of Experiments for Python.* [Online]
Available at: https://pythonhosted.org/pyDOE/
[Accessed 2017].

[156] Raftery, P., Keane, M. & O'Donnell, J., 2011. Calibrating whole building energy models: An evidence-based methodology. *Energy and Buildings,* pp. 2356-2364.

[157] Raslan, R. & Davies, M., 2010. Results variability in accredited building energy performance compliance demonstration software in the UK: an inter-model comparative study. *Journal of Building Performance Simulation,* 3(1), pp. 63-85.

[158] Reddy, T., Maor, I. & Panjapornpon, C., 2007. Calibrating detailed building energy simulation programs with measured data-Part 1: General methodology (RP-1051). *HVAC&R Research,* February, 13(2), pp. 221-241.

[159] Rencher, A., 2002. *Method of Multivariate Analysis.* 2nd ed. Utah, USA: Wiley-Interscience.

[160] RIBA, 2013. *RIBA Plan of Work 2013.* [Online]
Available at: http://www.ribaplanofwork.com/
[Accessed June 2015].

[161] Roberti, F., Oberegger, U. & Gasparella, A., 2015. Calibrating historic building energy models to hourly indoor air and surface temperatures: Methodology and case study. *Energy and Buildings,* Volume 108, pp. 236-243.

[162] Robertson, C. & Mumovic, D., 2013. *Assessing building performance: designed and actual performance in context of industry pressures.* Coventry, U.K., SB13 Sustainabile Building and Construciton Conference.

[163] Royapoor, M. & Roskilly, T., 2015. Building model calibration using energy and environmental data. *Energy and Buildings,* Volume 94, pp. 109-120.

[164] Ruyssevelt, P., 2014. *Analysis of CarbonBuzz Data,* London, U.K.: UCL Energy Institute.

[165] Rysanek, A. & Choudhary, R., 2014. DELORES - a tool for stochastic prediction of occupant services demand. *Journal of Building Performance Simulation,* pp. 1-22.

[166] Salehi, M. et al., 2013. *Improving the performance of a whole-building energy modeling tool by using post-occupancy measured data.* Chambéry, France, Building Simulation Optimisation .

[167] Schwartz, Y. & Raslan, R., 2013. Variations in results of building energy simulation tools, and their impact on BREEAM and LEED ratings: a case study. *Energy and Buildings,* Volume 62, pp. 350-359.

[168] Scott, D., 1992. *Multivariate Density Estimation: Theory, Practice and Visualization.* New York: John Wiley & Sons.

[169] Sobol, I., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics in Computer in Simulation,* Volume 55, pp. 271-280.

[170] Sun, K., Hong, T., Taylor-Lange, S. & Piette, M., 2016. A pattern-based automated approach to building energy model calibration. *Applied Energy,* Volume 165, pp. 214-224.

[171] Sun, T. & Reddy, A., 2006. Calibration of building energy simulation programs using the analytic optimization approach (RP-1051). *HVAC&R Research,* 12(1), pp. 177-196.

[172] Sun, Y. et al., 2014. Uncertainty quantification of microclimate variables in building energy models. *Journal of Building Performance Simulation,* 7(1), pp. 17-32.

[173] Swarup, L., Korkmaz, S. & Riley, D., 2011. Project Delivery Metrics for Sustainable High-Performance Buildings. *Journal of Construction Engineering and Management,* 137(12), pp. 1043-1051.

[174] ten Broeke, G., van Voorn, G. & Ligtenberg, A., 2016. Which sensitivity analysis method should I use for my agent-based model?. *Journal of Artificial Societies and Social Simulation,* 19(5).

[175] Torcellini, P. et al., 2006. *Lessons learned from field evaluation of six high-performance buildings,* Pacific Grove, CA, U.S.: ACEEE Summer Study on Energy Efficiency in Buildings.

[176] Turner, C. & Frankel, M., 2008. *Energy Performance of LEED for New Construction Buildings,* WA, USA: New Buildings Institute.

[177] UKGBC, 2007. *Carbon reductions in new non-domestic buildings,* London: Queen's Printer.

[178] UKGBC, 2016. *Delivering Building Performance,* London: UK Green Building Council.

[179] van Dronkelaar, C., Dowson, M., Spataru, C. & Mumovic, D., 2016. A review of the regulatory performance gap and its underlying causes in non-domestic buildings. *Frontiers in Mechanical Engineering,* 1(17).

[180] Wang, L., Mathew, P. & Pang, X., 2012. Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building. *Energy and Buildings,* Volume 53, pp. 152-158.

[181] Wang, S., Yan, C. & Xiao, F., 2012. Quantitative energy performance assessment methods for existing buildings. *Energy and Buildings,* Volume 55, pp. 873-888.

[182] Way, M., Bordass, B., Leaman, A. & Bunn, R., 2014. *The Soft Landing Framework - for better briefing, design, handover and building performance in-use,* Bracknell, U.K.: BSRIA.

[183] Wetter, M., 2011. A view on future building system modeling and simulation. In: J. Hensen & R. Lamberts, eds. *Building Performance for Design and Operation.* Berkeley, USA: Spon Press, pp. 482-504.

[184] WGBC, 2013. *The business case for green buildings - A Review of the Costs and Benefits for developers, investors and occupants,* New York, USA: The World Green Building Council.

[185] WGBC, 2014. *Health, Wellbeing & Productivity in Offices - the next chapter for green building,* New York, USA: The World Green Building Council.

[186] WGBC, 2016. *Health, Wellbeing & Productivity in Retail: The impact of green buildings on people and profit,* London, UK: The World Green Building Council.

[187] Williamson, B., 2012. *The gap between design and build: construction compliance towards 2020 in Scotland, Edinburgh,* Edinburgh, UK: CICStart.

[188] Yang, Z. & Becerik-Gerber, B., 2015. A model calibration framework for simultaneous multi-level building energy simulation. *Applied Energy,* Volume 149, pp. 415-431.

[189] Yin, R., Kiliccote, S. & Piette, M., 2016. Linking measurements and models in commercial buildings: A case study for model calibration and demand response strategy evaluation. *Energy and Buildings,* Volume 124, pp. 222-235.

[190] Yun, G., Kim, H. & Kim, J., 2012. Effects of ocupancy and lighting use patterns on lighting energy consumption. *Energy and Buildings,* Volume 46, pp. 152-158.

[191] ZCH, 2010. *A review of the modelling tools and assumptions: Topic 4, Closing the gap between designed and built performance,* London, UK: Zero Carbon Hub.

[192] ZCH, 2014. *Closing the gap between design and as-built performance - end of term report,* London, UK: Zero Carbon Hub.

[193] ZCH, 2014. *Closing the gap between design and as-built performance - evidence review report,* London, U.K.: Zero Carbon Hub.

[194] ZCH & NHBC Foundation, 2010. *Carbon compliance for tomorrow's new homes - a review of the modelling tool and assumptions - Topic 4 - Closing the gap between design and built performance,* London, U.K.: Zero Carbon Hub, NHBC.

[195] Zhang, T., Siebers, P. & Aickelin, U., 2011. Modelling electricity consumption in office buildings: An agent based approach. *Energy and Buildings,* Volume 43, pp. 2882-2892.

Table A 1: Predicted and measured data taken from literary sources

Source	Type	Base + small power (kWh/m²/y)	Base (kwh/m²a)	Actual (kWh/m²/y)	Percentage difference	Year	Prediction type	Area (m²)	Tool
Norford et al. (1994)	Office	125		325	160%	1994	B	12000	DOE-2.1C
Piette et al. (1994)	Office		79	108	36%	1994	A	2304	DOE 2.1
Piette et al. (1994)	Office		80	118	48%	1994	A	1505	DOE 2.1
Piette et al. (1994)	Office		84	172	104%	1994	A	72734	DOE 2.1
Piette et al. (1994)	Office		89	118	33%	1994	A	288	DOE 2.1
Piette et al. (1994)	Office		91	226	148%	1994	A	2332	DOE 2.1
Piette et al. (1994)	Multipurpose		91	205	124%	1994	A	557	DOE 2.1
Piette et al. (1994)	Office		92	140	52%	1994	A	492	DOE 2.1
Piette et al. (1994)	Office		93	86	-7%	1994	A	279	DOE 2.1
Piette et al. (1994)	Office		99	108	9%	1994	A	195	DOE 2.1
Piette et al. (1994)	Office		106	140	32%	1994	A	790	DOE 2.1
Piette et al. (1994)	School		111	108	-3%	1994	A	2917	DOE 2.1
Piette et al. (1994)	Office		111	129	16%	1994	A	7404	DOE 2.1
Piette et al. (1994)	Multipurpose		123	258	110%	1994	A	307	DOE 2.1
Piette et al. (1994)	Multipurpose		125	258	106%	1994	A	1236	DOE 2.1
Piette et al. (1994)	Office		144	151	5%	1994	A	1245	DOE 2.1
Piette et al. (1994)	School		149	140	-6%	1994	A	5175	DOE 2.1
Piette et al. (1994)	Office		158	108	-32%	1994	A	399	DOE 2.1
Piette et al. (1994)	Multipurpose		191	215	13%	1994	A	4366	DOE 2.1
Piette et al. (1994)	Multipurpose		200	484	142%	1994	A	1171	DOE 2.1
Piette et al. (1994)	Office		215	269	25%	1994	A	10986	DOE 2.1
Piette et al. (1994)	Office		226	237	5%	1994	A	36139	DOE 2.1
Piette et al. (1994)	Office		276	215	-22%	1994	A	8482	DOE 2.1
Piette et al. (1994)	Restaurant	166		237	43%	1994	B	1960	DOE 2.1
Piette et al. (1994)	Restaurant	384		657	71%	1994	B	232	DOE 2.1
Piette et al. (1994)	Restaurant	1147		1399	22%	1994	B	251	DOE 2.1
Piette et al. (1994)	Restaurant	1658		1442	-13%	1994	B	381	DOE 2.1
Piette et al. (1994)	Supermarket	775		581	-25%	1994	B	307	DOE 2.1
Piette et al. (1994)	Supermarket	472		495	5%	1994	B	3865	DOE 2.1
Diamond et al. (2006)	Library	240		164	-32%	2006	B (ASHRAE)	38275	
Diamond et al. (2006)	Multipurpose	218		151	-30%	2006	B (ASHRAE)	20476	
Diamond et al. (2006)	Office	199		148	-25%	2006	B (ASHRAE)	1672	
Diamond et al. (2006)	Multipurpose	177		139	-21%	2006	B (ASHRAE)	5946	
Diamond et al. (2006)	Office	164		151	-8%	2006	B (ASHRAE)	1689	
Diamond et al. (2006)	Office	215		246	15%	2006	B (ASHRAE)	567	
Diamond et al. (2006)	Library	208		309	48%	2006	B (ASHRAE)	2044	
Diamond et al. (2006)	Office	110		192	74%	2006	B (ASHRAE)	6503	
Ahmed & Culp (2006)	Multipurpose	247		214	-13%	2006	B (ASHRAE)	11516	DOE-2.1E
Ahmed & Culp (2006)	Multipurpose	500		438	-12	2006	B (ASHRAE)	17837	DOE-2.1E
Ahmed & Culp (2006)	Multipurpose	423		413	-2	2006	B (ASHRAE)	12156	DOE-2.1E
Ahmed & Culp (2006)	Multipurpose	698		1030	48	2006	B (ASHRAE)	16450	DOE-2.1E
Diamond et al. (2006)	Multipurpose		486	151	-69	2006	A (ASHRAE)	2992	
Diamond et al. (2006)	Office	530		498	-6	2006	B (ASHRAE)	779	
Diamond et al. (2006)	Office	218		218	0	2006	B (ASHRAE)	11664	
Diamond et al. (2006)	Office	249		230	-8	2006	B (ASHRAE)	4853	
Diamond et al. (2006)	Office	394		69	-82	2006	B (ASHRAE)	6689	
Diamond et al. (2006)	Office	836		404	-52	2006	B (ASHRAE)	4840	

Source	Type	Base + small power (kWh/m²/y)	Base (kwh/m²a)	Actual (kWh/m²/y)	Percentage difference	Year	Prediction type	Area (m²)	Tool
Diamond et al. (2006)	Laboratory	1114		1126	1	2006	B (ASHRAE)	6544	
Diamond et al. (2006)	Laboratory	842		855	1	2006	B (ASHRAE)	6685	
Diamond et al. (2006)	Laboratory	470		915	95	2006	B (ASHRAE)	31178	
Diamond et al. (2006)	University	85		104	22	2006	B (ASHRAE)	7322	
Diamond et al. (2006)	Multipurpose	145		325	124	2006	B	34221	
Pegg et al. (2007)	School	103		219	112	2007	D	10627	
Pegg et al. (2007)	School	100		228	127	2007	D	10529	
Pegg et al. (2007)	School	99		209	111	2007	D	13000	
Knight et al. (2008)	University	286		310	8	2008	B (SBEM)	11150	SBEM
Calderone (2011)	Office	67		76	13	2011	B (NABERS)	14500	IES VE
Calderone (2011)	Office	83		61	-26	2011	B (NABERS)	34000	IES VE
Calderone (2011)	Office	168		122	-27	2011	B (NABERS)	14500	IES VE
Calderone (2011)	Office	89		76	-15	2011	B (NABERS)	5900	IES VE
Menezes et al. (2012)	Office	60	30	128	113	2012	E	2000	TM22
Menezes et al. (2012)	Office	60	30	103	72	2012	E	2000	TM22
Korjenic & Bednar (2012)	Office	82		115	40	2012	H	4811	BuildOpt
Bertagnolio et al. (2012)	Office	153		170	11	2012	C	4432	ISO 13790
Austin (2013)	Office	240		353	47	2013	B (NCM)	9144	VIS DOE 3.0
Salehi et al. (2013)	Multipurpose	35		59	69	2013	B (NECCB)	5700	IES VE
Daly et al. (2014)	Office	68		66	-2	2014	B (ASHRAE)	7500	ECOTECT
Kimpian et al. (2014)	School	160		280	75	2014	F	10490	DSM / TM22
Kimpian et al. (2014)	School	190		245	29	2014	F	10172	DSM / TM22
Kimpian et al. (2014)	School	111		212	91	2014	F	10418	DSM / TM22
Kimpian et al. (2014)	School	75		168	124	2014	G	14610	DSM / TM22
Kimpian et al. (2014)	School	92		125	36	2014	F	2834	DSM / TM22
Kimpian et al. (2014)	Office	142		100	-30	2014	F	2310	DSM / TM22
Murphy & Castleton (2014)	Office	168		140	-17	2014	B (SBEM)	9996	SBEM
Ruyssevelt (2014)	Office	138		225	63	2014	unclear		
Ruyssevelt (2014)	School	119		163	37	2014	unclear		
Ruyssevelt (2014)	University	124		317	156	2014	unclear		
Ruyssevelt (2014)	Retail	174		195	12	2014	unclear		
Torcellini et al. (2006)	Multipurpose	0		52		2006	B (ASHRAE)	1265	DOE 2.2
Torcellini et al. (2006)	Retail	48		78	62	2006	B (ASHRAE)	1076	DOE 2.1E
Torcellini et al. (2006)	Office	86		114	31	2006	B (ASHRAE)	3205	DOE 2.2
Torcellini et al. (2006)	Office	73		126	73	2006	B (ASHRAE)	2900	EnergyPlus
Torcellini et al. (2006)	Multipurpose	39		90	132	2006	A (ASHRAE)	930	DOE 2.1E
Torcellini et al. (2006)	Multipurpose	64		124	95	2006	B (ASHRAE)	3940	DOE 2.1E

Table A 2: Prediction methods, see **Table A1.**

	Type of prediction
A	Design stage calculation, excluding unregulated loads
B	Design stage calculation, including equipment loads, standard operation
C	Design stage calculation, including equipment loads and detailed operation
D	CIBSE building energy code 1 (1998) using monthly average temperatures and included unregulated loads, no thermal modelling, similar to CIBSE TM22 bottom-up approach
E	CIBSE TM22 Bottom-up approach
F	NCM in thermal modelling + equipment, external lighting and lift using TM22
G	NCM in thermal modelling + benchmarking for DHW and Auxiliary loads +equipment, external lighting and lift using TM22
H	Monthly balance method (EN-ISO 13790) + equipment loads
I	Quasi steady-state hourly simulation relying on simple normative models (EN-ISO 13790)

Table A 3: Underlying causes identified in literature case studies

Source	Effect on energy use	Underlying issue	Context
Austin (2013)	44%	Poor practice	Fans, pumps and cooling showed greatest divergence between modelling and metered, underestimated caused by the way the plant is controlled, different from assumptions
Austin (2013)	25%	Specification uncertainty	Lighting and small power overestimated
Austin (2013)	30%	Calculation methodologies	Main server room and its air conditioning not included in original model, a substantial load of the total electricity use
Bhandari et al. (2012)	7%	Scenario uncertainty	Predicted energy use can vary up to 7% as a function of the provided location's weather data
Bloomfield (1988)	-46% to 106%	Heuristic uncertainty	BRE observed 25 users giving different predicted energy consumption of a large complex building
Brown et al. (2010)	9%	Energy use variability in operation	Present a longitudinal analysis of 25 buildings in the UK and found an increase of 9% in energy use on average per year, with a standard deviation of 18%
Burman et al. (2012)	10%	Modelling uncertainty	Found that pumps' auxiliary power could not be modelled for compliance purposes and had to apply default values based on HVAC system type the case study building, this amounts to approximately 9.3 $kWh/m^2/yr$ error in estimating the actual energy usage of pumps (10% of total energy use), whereas a simple back of envelope calculation based on the pumps' ratings would yield a result close to the sub-metered data
Burman et al. (2012)		Scenario uncertainty	Wrong setpoints, too high estimation in DHW use
Burman et al. (2014)	25% of gap	Poor commissioning	Poor implementation of control strategy for mechanical ventilation led to failure of demand-controlled ventilation (fans run at full capacity regardless of actual demand), led to excessive auxiliary and heating energy use. Poor actuator control at the sliding header interface between GSHP and gas-fired boilers, which led to low contribution of the heat pumps to heating, almost half of design intent.
Burman et al. (2014)	75% of gap	Poor practice	Schedules of operation of heating and ventilation systems were not controlled
Calderone (2011)	+16%	Specification / Scenario uncertainty	Higher ventilation rates than design and lower internal heat loads resulted in an increase in 40% increase in gas use
Calderone (2011)	-5%	Specification uncertainty	HVAC pumps using 40% more electricity than predicted due to optimistic profile use for variable speed pumping strategy
Calderone (2011)	+5%	Specification / Scenario uncertainty	Heat pump VRF system is using 6 times more electricity than predicted due to an increase in longer than expected operating hours and higher infiltration rates
Calderone (2011)		Modelling uncertainty	Modelling software cannot model all control scenarios installed within a building
Calderone (2011)		Poor practice	Building management can have a large impact on overall energy use based on their understanding and operation of systems
Calderone (2011)		Scenario uncertainty	Weather variation adversely impact energy consumption of HVAC system, especially noticeable for extreme events which are generally not found in test reference year data
de Wilde (2014)	5%	Energy use variability in operation	Heat pump coefficient of performance high effect on prediction, degradation of heat pump due to being used at partial load
Fedoruk et al. (2015)		Institutional issues	Primary impediments were institutional in the sense that they arose from the way the various stages of the building life cycle were specified, contracted and implemented. Problems had to do with a lack of useful information, interpretation, communication, feedback and integration than with the expense and technical difficulty involved in implementing design goals
Fedoruk et al. (2015)		Measurement system limitations	Significant fault in meters that did not recognize flow directionality, were mislabelled, incorrectly installed and calibrated, left out of the energy understanding boundary, or which used differing convention for various system components.
Guyon (1997)	±40%	Heuristic uncertainty	Investigated influence of 12 energy modellers on predicted energy use of residential house
Karlsson et al. (2007)	2%	Inter-model variability	Differences between three difference dynamic simulation tools (domestic building)

Source	Effect on energy use	Underlying issue	Context
Karlsson et al. (2007)	7%	Occupant behaviour	Difference in internal gains due to occupant habits
Karlsson et al. (2007)	10%	Poor practice	Airflow control
Karlsson et al. (2007)	20%	Specification uncertainty	Heat exchanger efficiency
Kimpian et al. (2014)		Poor practice	Excessive heating consumption due to BMS shortfalls and lack of seasonal commissioning
Kimpian et al. (2014)		Scenario uncertainty	Setpoints and operational schedules often higher and longer than assumed
Kimpian et al. (2014)		Changes after design	Automatic doors had sensors that were difficult to adjust to achieve smaller openings during winter time and security camera links that would have enabled the doors to stay shut longer were value-engineered out.
Kimpian et al. (2014)		Degradation of materials	Window seals were found to degrade within the first year of a school's operation
Kimpian et al. (2014)		Poor practice	Thermal setpoints for plenum ventilation flaps were not enabled in the BMS, resulting in complaints about overheating
Kimpian et al. (2014)		Specification uncertainty	Energy use predictions routinely under-estimate occupier electrical loads arising from appliance and IT loads
Kimpian et al. (2014)		Specification uncertainty / poor practice	Full load specific fan powers (SFP) achieved at the commissioning stage were higher than the design intent in all cases. Dirty and clogged air filters found in the buildings further aggravated these.
Kimpian et al. (2014)		Energy performance target	The lack of energy-focused commissioning had a major impact on performance outcomes
Kimpian et al. (2014)		Changes after design	Users found actuators for automated windows disruptively noisy at these buildings, motors specified were replaced with a cheaper model. These resulted in excessive heat consumption during winter and discomfort from overheating during the summer.
Kleber & Wagner (2007)		Poor practice	Monitored and office building and found that failures in operating the building's facilities caused higher energy consumption, they underline the importance of continuous commissioning
Kawamoto et al. (2004)		Occupant behaviour	82-97% power down overnight
Maile (2010)		Poor commissioning	Uncalibrated sensors on electrical sub-meters were off by a factor eight
Maile (2010)		Poor practice	Wrongly assumed control strategies
Martani et al. (2012)	63-69%	Occupant behaviour	Analysed two building and revealed significant correlation between electricity consumption and occupancy
Masoso & Grobler (2010)		Occupant behaviour	Night-time energy use (leaving office equipment on)
Moezzi et al. (2013)		Poor practice	Building operations can help achieve higher energy savings through better operation, and indicate that assumption may not be met.
Mulville et al. (2014)		Occupant behaviour	70-72% power down overnight
Murphy & Castleton (2014)		Changes after design	Report in their case study that the roll-out of unspecified low-energy equipment not taken into account in the predicted model affected final unregulated loads, influencing both unregulated energy use and cooling energy due to lower internal gains.
Murphy & Castleton (2014)		Changes after design / Specification uncertainty	Equipment loads 50% lower due to "thin-client" pc terminals
Murphy & Castleton (2014)		Heuristic uncertainty	Reviewed a compliance model, which they wished to compare with measured energy use. Although not having access to the model, they realised that the compliance model output suggested a total floor area twice the size of the actual floor area, likely due to the model taking the ceiling void into account as occupied spaces
Murphy & Castleton (2014)		Specification uncertainty	Lower lighting energy use than predicted
Neymark et al. (2002)	4-40%	Inter-model variability	A comparative study of 7 different tools
Neymark et al. (2002)		Numerical uncertainty	Faulty algorithms caused errors of up to 20-45% in predicting COP values

Source	Effect on energy use	Underlying issue	Context
Norford et al. (1994)	24%	Scenario uncertainty	Extended HVAC operation
Norford et al. (1994)	64%	Specification uncertainty	Lighting and office equipment powers in design 10.8 and 5.4 and practice 16.2 and 15.1
Norford et al. (1994)	12%	Specification uncertainty	HVAC equipment not operating to specs (assumed)
Pang et al. (2012)		Poor practice	Poor operational practice, malfunctioning equipment, incorrectly configured control systems and lack of continuous commissioning are often reported as main drivers for underperformance in operation
Parys et al. (2010)	±10%	Occupant behaviour	Reported a standard deviation of up to 10% on energy use to be related to occupant behaviour
Pegg et al. (2007)	30-45% heating	Poor practice	Heating energy consumption underestimated 30-45% due to mechanical ventilation on outside occupied hours
Pegg et al. (2007)	3x Operational hours	Scenario uncertainty	Hours of operation underestimated by a factor of 3 although design lighting load correct
Pegg et al. (2007)	57-75% small power energy	Specification uncertainty	57% to 70% underestimation of small power energy use, different load and operation
Piette et al. (1994)	6%	Energy use variability in operation	Variability in use from 3th to 4th year by 6%
Raslan & Davies (2010)		Modelling uncertainty	Steady state models have a difficulty to model complex systems while dynamic simulation systems were capable
Salehi et al. (2013)		Specification uncertainty	Underlying assumption for plug loads and lighting significantly underestimated actual energy use
Schwartz & Raslan (2013)	35%	Inter-model variability	Differences between three difference dynamic simulation tools
Torcellini et al. (2006)		Changes after design	Important changes to the design occurring late in the design development were not updated in the energy models
Torcellini et al. (2006)		Changes after design	Omission of a lighting display area
Torcellini et al. (2006)		Changes after design	Electrical lighting circuiting not installed as designed, limiting daylight control strategy and giving uneven light distribution.
Torcellini et al. (2006)		Changes after design	Not installed specified window and door frame thermal breaks
Torcellini et al. (2006)		Energy performance target	Measurable performance goals translate into efficient building performance
Torcellini et al. (2006)		Poor practice and malfunctioning equipment	Energy-saving technologies such as lighting controls, CO2 sensors an desiccant heat recovery are not used or faulty
Torcellini et al. (2006)		Procurement issue	Foundation missing specified perimeter insulation resulting in a thermal bridge and affecting occupant comfort and energy performance
Torcellini et al. (2006)		Scenario uncertainty	Simulation create idealistic controls, actual performance showed different setpoints and less setup and setback of space temperatures
Torcellini et al. (2006)	Twice those assumed	Specification uncertainty	Plug loads were often greater than design predictions, and loads underestimated for server room, exterior lighting and mechanical accessories
Torcellini et al. (2006)		Specification uncertainty	Insulation values are optimistic compared to actual construction techniques
Torcellini et al. (2006)		Specification uncertainty and Occupant behaviour	Designers were too optimistic about the behaviour of occupants and their acceptance of systems
Wang et al. (2012)		Poor practice	Uncertainty due to operational practices poor practice increase in energy use by 49-79% while good practice reduces consumption by 15-29%
Wang et al. (2012)	-4% to 6%	Scenario uncertainty	Year-to-year weather fluctuation on energy use -4% to 6%
Wetter (2011)		Modelling uncertainty	Mechanical systems often simplified and do not capture their dynamic behaviour and part-load operation
Zhang et al. (2011)		Occupant behaviour	Surveyed 143 staff and found that 60% never turn down computer, while 31% did so occasionally

APPENDIX B. BACKGROUND OF QUOTED INTERVIEWEES

The qualitative research carried out through the use of interviews and round-table discussions with industry experts resulting in a variety of quotes that were used. Although all interviewees are anonymised, **Table B1** provides some context to the quotes given in chapter 3.

Table B1: Role and background of quoted interviewees in chapter 3.

Code	Professional role	Background
INT1	Head of sales at a proprietary software developer	Wealth of experience in multi-disciplinary, collaborative, building design and specific expertise in energy modelling, building physics and building performance. Previously led Building Physics team and large global engineering consultancy.
INT2	Head of energy & sustainability at multinational defence technology company in the UK	Experience in energy and carbon, sustainability in the built environment, economic decision making, statistics and financial appraisal techniques. Previous experience as a sustainability and environmental operations manager for a large project management and construction services company for the UK government.
INT3	Head of sustainability at large property investment and development business	Responsible for developing and delivering the sustainability agenda for a large property investment and development business headquartered in London. Previous experience as a sustainability and environmental design manager.
INT4	Sustainability manager at one of the largest property managers in the UK	Environmental and sustainability professional with experience in the property industry for the management and delivery of complex schemes at asset and portfolio level. Experience in negotiating and facilitating discussions with key stakeholder groups within the built environment throughout the RIBA stages.
INT5	Head of Sustainability at large global asset management company	Previous experience at global engineering consulting companies, heading the sustainability teams therein. 25 years' experience in the sustainability and carbon management sector.
INT6	Head of innovation and property solutions at a large commercial property investment company	Responsible for creating and implementing long-term engineering, environmental, smart procurement, intelligent building and product innovation strategy. With previous experience in heading up the design and engineering departments at a large property investment company
INT7	Portfolio manager at a multinational technology company	Strong technical engineering and environmental background and broad range of experience across areas including strategic planning, operations and change management. Previous experience as sustainability program manager and consultant.
INT8	Head of sustainable development at international professional engineering association	With 10 years of experience in a large consulting firm, specialises in low-carbon buildings, health and wellbeing and a variety of built environment projects from early master planning through to post-occupancy evaluation and policy work.

APPENDIX C. DESCRIPTION OF CASE STUDY BUILDINGS

Office 17

Office 17 is a pre-1930's office building with a total gross floor area of ~1924 m^2, occupied by about 200 people. It partially adjoins other office buildings and provides daylight through openable windows and roof lights. All floors, from basement to second floor are open plan office with typically 1 or 2 meeting rooms. In addition, the basement contains a large print room, server room and canteen. The ground floor reception is the main entrance to the building and directly connects to the open plan office. The front façade and an aerial view of the building are shown in **Figure C1** and **Figure C2** respectively.

Figure C1: Front façade of the building

Figure C2: Aerial view of the building

Building elements

The building fabric consists of solid walls with no insulation and single glazing in aluminium frames. Two stairwells connect the basement, ground- and first floor, in addition a lift connects all floors. In the basement, a canteen and print room are located with open connection to the open plan office space.

Heating and cooling

Heating is provided by two gas-condensing boilers installed on the first floor, whereas domestic hot water is provided by decentralised electric heaters. The print room, canteen and office spaces in the basement and the server are air-conditioned, as well as most of the meeting rooms, while all other floors are naturally ventilated with fans located on ceilings to provide extra air movement during the summer. Local desk fans are provided during the summer months.

Lighting and power

T-8 high frequency fluorescents provide electric lighting in the office spaces, with LED lights in the reception and meeting rooms, additional desk lamps are present in some cases. The office spaces are densely populated. Desks typically have two screens and a laptop or in some cases a desktop computer to run heavier engineering software. The building is equipped with a sub-metering system connected to all floors and is used to provide half-hourly data used in this study.

Office 71

Office 71 is a 6-storey mid-terrace office building. The building has been substantially refurbished both internally and externally in 1997. Consequently, the building can be considered to be in good condition commensurate with its age (thought to be late 1950s/early 1960s).

Figure C3: Office 71 located in dense urban area.

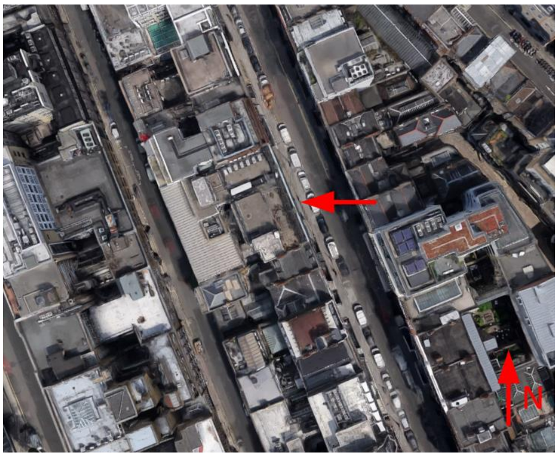


Figure C4: Aerial view of Office 71

Building elements

Building structure includes a reinforced concrete frame with solid concrete basement floor, in situ cast concrete upper floors and roof decks with asphalt finished flat roofs. Assumed cavity brick and blockwork to the rear elevation, with reinforced concrete cladding to the front elevation. Double-glazed aluminium tilt and turn windows are provided throughout, with secondary glazing to the ground floor. External areas are minimal, with a small loading bay to the rear of the building.

Heating and cooling

Building services were designed and installed in 1997, since, a heat pump installation (VRV) was added in 2000. Heating for the building is provided by a low pressure low temperature hot water (LPHW) heating system, generated by 2 gas fired cast iron boilers each rated at 147kW, arranged in parallel and controlled by a sequence controller. Circulation through the primary heating circuit is maintained by duplicate direct drive centrifugal pumps installed in the primary circuit flow, arranged in parallel to serves the boiler primary heating circuit and air-handling unit heater batteries. Radiators are located around the external elevation beneath the windows on each floor. In addition, an air-handling unit provides supplementary heating and cooling for the office floors, supported by an air-cooled modular multi-zone 2-pipe VRV system. The air-cooled units are located within the roof plant area and condensing units are located on the external area of roof. Fan coil units serve the offices and meeting rooms, this system operates independently of the LPHW heating and ventilation plant. Domestic hot water is generated via local electric point of use water heaters located adjacent to tea points and toilets. Domestic hot water for the kitchen is generated via a 550-litre capacity unvented direct storage calorifier.

Ventilation

An air-handling unit distributes tempered fresh air to each floor via a row of side wall extract grilles running at high level along the south elevation of each floor. The return air is distributed via an extract riser separate from the air handler. Additionally, mechanical fans located on the roof, extract air from the kitchen, print room and toilet spaces. An overview of the HVAC systems in Office 71 is shown in **Figure C5**.

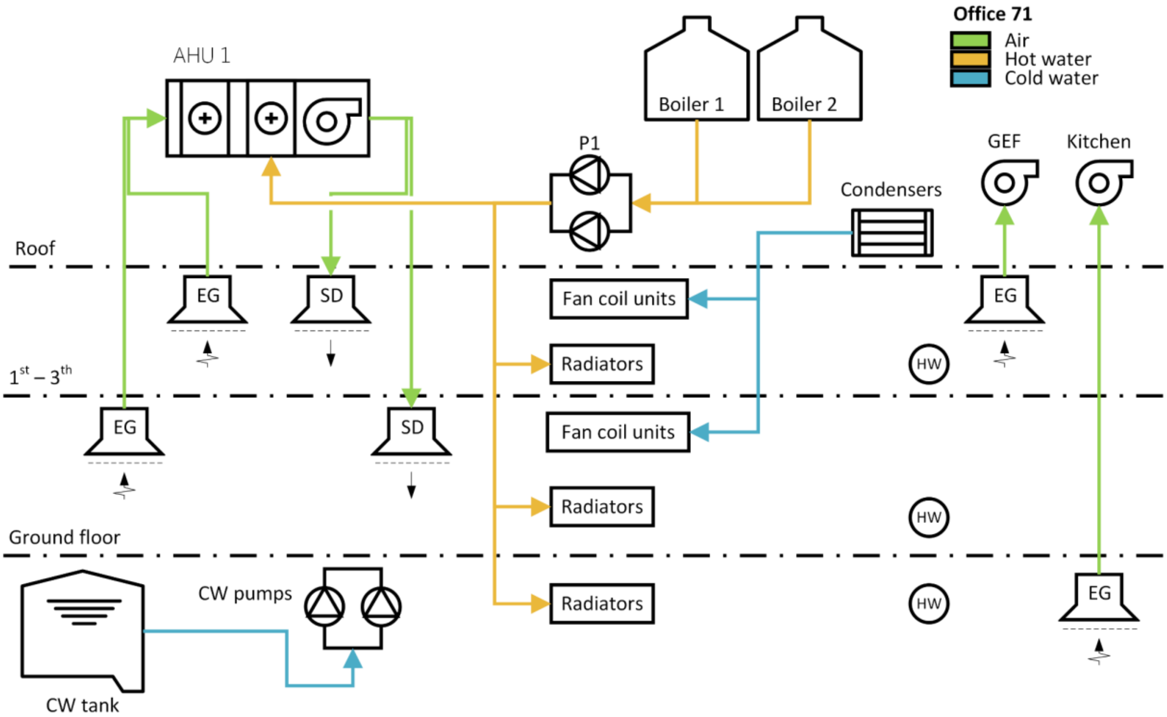


Figure C5: Office 71 HVAC schematic, arrows representing supply to sub-system components.

Lighting and power

Compact fluorescent lighting exists throughout the building, typically fittings were used with lamps in two varieties, a 26W and 58W fitting. Equipment power use in the building is mainly through computers/laptops and screens, where some ancillary equipment can be found in both the office and meeting spaces, such as projector, printers and TV's.

CH

CH hosts the University College London departments of the Bartlett School of Environment, Energy and Resources, Communications & Marketing, the Development & Alumni Relations Office (DARO). UCL acquired the majority of CH in mid-2009. Originally built around the early 1900's, it had a major refurbishment on all seven floors in 2010. CH provides 4585 m^2 of office accommodation, library and other facilities over basement, ground floor and six upper floors.

Figure C6: CH front entrance

Figure C7: Aerial view of CH

The building consists primarily of higher education office (44%). There are student work areas in the basement, which are more like meeting rooms and several IT intensive office spaces (2%), which provide hot desks for students. On the ground floor, there is a lecture theatre that is regularly used in the evenings for lectures or gatherings and during the day by students. The refurbishment created space for a new library (300 m^2) to provide both storage for literary works and working space for students.

Building elements

The building structure consists of a combination of cast concrete floors and brick walls. External walls have been insulated on the inside with a thermal lining composite board, fixed to the brickwork. In addition, secondary glazing has been added behind the single glazing at every window, where double glazing is fitted to the ground floor skylights.

Heating and Cooling

Heating and cooling to all offices, meeting rooms, comms rooms and learning spaces is provided by a VRV multi split system with heat recovery. 19 outdoor condensing units are located on the roof and connect to a multitude of indoor evaporators of differing capacities. The system allows for simultaneous heating and cooling in different zones. Central control panels allow for adjusting and resetting the setpoint temperatures of the units. Hot water is provided to the toilet areas, kitchenette areas and cleaners' cupboard through electric hot water heaters. Radiator heating to the stair spaces is provided by two condensing boilers located within the basement boiler house. An overview of the HVAC systems in is given in Figure C8.

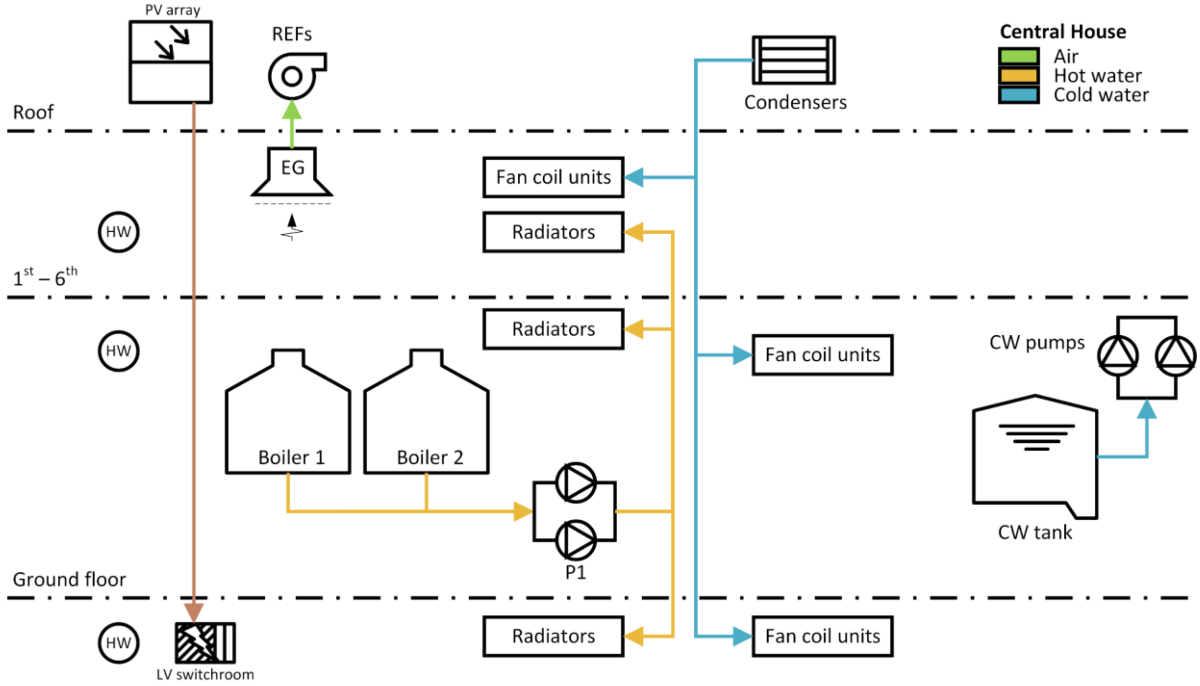


Figure C8: CH HVAC schematic.

Ventilation

Nearly all perimeter spaces have openable windows to allow for natural ventilation, spaces without windows have air provided through transfer grills from other spaces. For some spaces in the basement and ground floor, an air-handling unit provides fresh air. These come with complete DX coils, cooling and heating the air as required. Ducted extract systems serve the toilets and shower areas on all floors.

Lighting and Power

In the lavatories and circulation spaces, lighting is activated through occupancy detection, for all other areas lighting is turned on manually, but will switch off when people leave the space. Lighting in offices and meeting rooms is typically provided by recessed 600x600mm modular luminaires (3x24W T16) with dimming. In the circulation spaces there are single downlights luminaires (TC-TEL 32W). The reception and library have generally lower wattage luminaires (18W LEDs) installed. Equipment power is similar a typical office building, but follows a different pattern of use as it serves as a university building, where occupancy fluctuates heavily throughout the year.

Renewables

Electricity is generated by 18 photovoltaic panels each 190W, able to generate 3.42 kWp, which is estimated to produce approximately 3,000 kWh per year. The PV system is connected to a distribution board, separately monitored to determine the exact amount of electricity generation.

MPEB

MPEB was completed in 2005 and is home to a number of UCL Engineering departments as well as the Institute of Making. The main facade is divided into three sections, covered with glass and terracotta.

Figure C9: Front entrance of MPEB on UCL campus

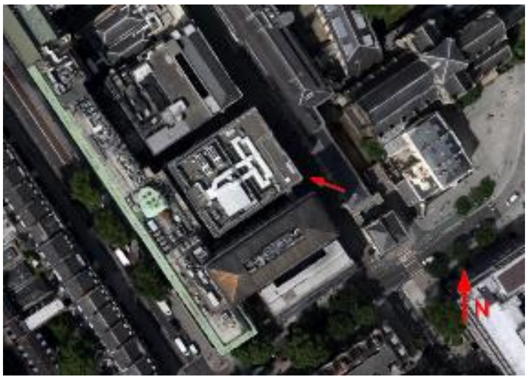


Figure C10: Aerial view of MPEB

The UCL Department of Computer Science, Medical Physics & Biomedical Engineering and Mechanical Engineering, occupy most of the building. The basement is split into three main functional areas - engineering workshop, technology laboratory, and a thermodynamics laboratory. The ground floor comprises of an entrance foyer/reception, loading bay, computer cluster, fuel store, and plant room. The upper floors are predominately used for research. The main teaching facilities are located on the first floor, with cellular offices and labs on second to fourth floors, and a combination of open-plan and cellular office accommodation on fifth to eighth floors. The building is located within the UCL campus premises and closely surrounded by other UCL estates owned buildings. The building is multi-functional and contains offices, laboratory, workshops, large server space, computer labs, meeting rooms and a small number of kitchen and reception areas. The upper levels consist of office space, whereas the basement, first floor and ground floor mostly contain workshops, seminar rooms and laboratory spaces.

Building elements

The building structure is made of a nine storey steel frame (1 basement and 9 floors in the superstructure). Some of the steel columns extend down to basement level for founding. The superstructure is generally pre-case concrete planks with a concrete box basement. The cladding finish to North, South & West elevations is profiled PVF2 coated galvanised steel sheeting with ribbon windows on the 1st to 8th floors. The East elevation finish is curtain walling with terracotta tile panels. The front of the building has a glazed curtain wall, which encloses the entrance foyer, and passenger lift lobbies on all above ground floors.

Heating and cooling

MPEB is connected to the UCL district heat network. This serves the heating system with LTHW to the air handling units, re-heater heating coils, radiators and local duct mounted heater batteries (fan-coil units and over door curtain heater). Radiators provide heating primarily to office spaces at the perimeter of the building, about half of the office spaces are provided with fan coil units sometimes in combination with radiators. Chilled water is generated by two packaged air-cooled water chillers located on the roof. Chilled water is distributed to serve the cooling coils in the air handling units, remote duct mounted cooling coils and fan coil units throughout the building. Basement laboratory and workshop areas are provided with electric storage water heaters. domestic hot water is serviced from the existing heating network from the building next door, providing hot water through various risers to the male and female lavatories.

Ventilation

Plenum fresh air is provided to the 2nd to 8th floor general areas (circulation and several open plan offices) from AHU5. The lift lobby at the front of the building at every floor and the lavatories are provided with fresh air from AHU4, and separately extracted through independent twin toilet extract fans, located at roof level. The first floor areas including the lecture theatres are provided with treated, heated and cooled supply air and extract ventilation from AHU1, located in the ground floor plant room. Each of these rooms is also provided with a duct mounted LPHW re-heater coil and chilled water cooling coil to adjust the supply condition to suit the various zone load conditions. Ground floor areas are provided with treated, heated and cooled supply air and extract ventilation, served by AHU3 located on the ground floor with separate extract. Most of the basement areas are served by AHU2 with separate extract, where additional ventilation is and extract are provided to the engine dynamometer cells through louvres and high duty extract fans. Fume extract is provided to several of the workshop and separate Institute of Making. Along the perimeter, offices on the first floor and up have openable windows for natural ventilation. An overview of the HVAC systems in MPEB is shown in **Figure C10**.

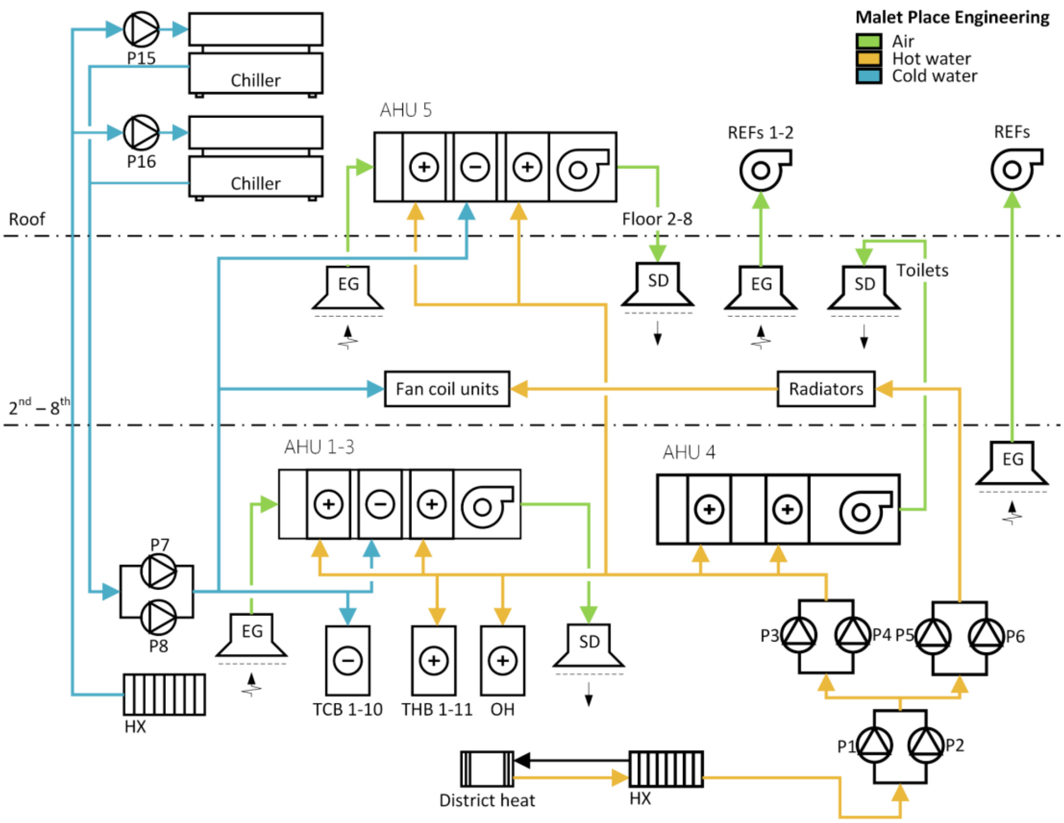


Figure C10: MPEB HVAC schematic, arrows representing exclusively supply to sub-systems.

Lighting and power

From the ground floor to the 8th floor, lighting comprises of recessed- and compact fluorescent luminaries controlled via a lighting management system at each floor, movement sensors are provided throughout the general use areas. Lecture theatre and meeting rooms are provided with a dimming interface to allow the lighting to be controlled. Local switching is provided in spaces such as storage, cupboards, dark rooms, laboratories and plant rooms. Equipment loads in the spaces are somewhat higher than those found in CH, although most of the spaces in MPEB are offices, there are multiple spaces such as laboratory, workshops spaces and smaller server rooms that contain a large amount of additional equipment.

APPENDIX D. BUILDING MODELLING INPUTS

In **Table D1** an overview is given of the space types present in the four buildings, the number of spaces assigned to them and their total floor area.

Table D1: Space types in the case study buildings,
the number of spaces assigned, and floor area in their respective models.

Space type	Office 17		Office 71		CH		MPEB	
	no.	m^2	no.	m^2	no.	m^2	no.	m^2
Office	4	1258	3	1042	69	1999	114	2588
Circulation	1	55	28	489	46	895	95	1698
Storage	3	54	9	137	28	111	33	231
Lavatory	3	45	6	143	24	246	28	345
Meeting	4	120	6	217	24	235		
Lift	1				16	111	27	130
Kitchen			1	41	11	171	2	34
Server	1	54			9	62	2	168
Shower					6	16	1	8
Lecture Theatre					4	172	5	281
Plant	1		1	80	3	77	9	320
Reception	1	155	1	121	2	116	2	35
Computer Cluster					2	76	6	492
Print Room	1	115			1	9	3	28
Library					2	303		
Workshop							9	714
Laboratory								
Lift Lobby							10	524
Canteen	1	68	1	284				
All spaces		1924	56	2554	225	4598	467	8451

Materials

In Office 17, different material properties were included as uncertain parameters and it was evident that mainly the conductivity of materials was important. Macdonald (2002) quantified the uncertainties in material properties for conductivity, solar absorptance and emissivity to be 5%, 1% and 12.25% respectively. These uncertainties were taking into account for the materials in Office 17, however only the conductivity in the remaining three case studies was taken into account. See Appendix E for properties of the building materials used and their build-up.

People density and metabolic rate

People have a large influence on the energy use of a building, they are the ones that use the equipment, lighting and for whom generally a conditioned space is provided. In building energy modelling the amount of people in a building or zone is defined by the people density. The number of people in a building or zone can be difficult to determine. For CH and MPEB, use is made of Wi-Fi and swipe card access data to establish when and how many people are present in the building. Base people density values were taken from the National Calculation Methodology (NCM) modelling guide for buildings other than dwellings in England (NCM, 2013). These were combined with the by the peak number of occupants to establish a specific people density for different space types, shown in **Table D2**.

Table D2: People density (m^2/p) per space type

Space type	Office 17	Office 71	CH	MPEB
Office	5	9	7	11
Circulation	15	9	16	25
Storage	20	40	16	25
Lavatory	9	9	7	11

Space type	Office 17	Office 71	CH	MPEB
Meeting		9	7	11
Lift	20	20	16	25.2
Kitchen			3	5
Server			7	11
Shower			7	11
Lecture Theatre			3	5
Plant		9	20	11
Reception	5	10	10	11
Computer Cluster			6	5
Print Room	5		7	11
Library			7	11
Workshop				11
Laboratory				11
Lift Lobby				5
Canteen	5	3.5		
Peak no. of people			600	800

The number of people can directly affect parameters such as natural and mechanical ventilation flow rates. Additionally, people dissipate heat, the amount depends on the type of activity they perform. Typical values for the heat generation of people for different types of activities are assigned to different space types as shown in **Table D3.**

Table D3: Typical metabolic rate and heat generation per unit area of body surface for various activities

Activity		Metabolic rate (M_a)	Heat generation (W/m^2)	Model input (W)	Space type
Resting:			CIBSE / AHSRAE		
-	sleeping	0.7	41 / 40	74	
-	seated, quiet	1.0	58 / 60	**104**	Residential, Lavatory
-	standing, relaxed	1.2	70 / 70	**126**	Lift
Office work:					
-	reading, seated	1	58 / 55	104	
-	typing	1	58 / 65	**104**	Office, Meeting
Walking					
-	3.2 km/h (0.9 m/s)	2	116 / 115	**209**	Circulation
-	4.3 km/h (1.2 m/s)	2.6	151 / 150	272	
Occupational					
-	light machine work	3	93 – 116 / 115 - 140	**188**	Plant, Workshop, Laboratory

Note 1: Average surface area of an adult human is about $1.8m^2$, used for the W calculation. 1 M_a is 58.1 W/m^2
Note 2: Figures from CIBSE (CIBSE, 2007) and ASHRAE (ASHRAE, 2013)

There is however a large uncertainty in measuring metabolic rates and in defining the tasks, it is therefore reasonable to assume a ±20% variability for engineering purposes for well-defined activities with $M_a < 1.5$ and ±50% for poorly defined activities with $M_a > 3.0$ (CIBSE, 2007). Other space types not mentioned in the table have the same value as that for the Office space type.

Lighting loads

Lighting loads are based on the type of lighting used in a space and the number of fixtures. This is determined from available drawings in O&M manuals and verified during building audits. For each space type the lighting load in all spaces belonging to that space type were aggregated and divided by floor area. The lighting load (W/m^2) is then determined per space type, the total lighting load is then based on the schedule of operation, i.e. when lighting is on/off in spaces. Spaces are assigned to space types, however they are not identical, a standard deviation was therefore calculated where multiple spaces assigned to the same space type, shown in **Table D4**. Lighting loads were varied at 20% standard deviation and varied according to the noted standard deviations in the table for the 'evidenced' ranges.

Table D4: Lighting load (W/m²) per space type with their respective standard deviation

Space type	Office 17	Office 71	CH	MPEB
Office	15	8	10 (3)	12 (4)
Circulation	10	3	8 (4)	9 (6)
Storage	12	8	14 (4)	11 (7)
Lavatory	8	7	13 (3)	19 (9)
Meeting	10	8	19 (4)	12 (5)
Lift			7 (3)	5 (3)
Kitchen			11 (4)	11 (5)
Server	12		15 (4)	18 (6)
Shower			14 (5)	15 (1)
Lecture Theatre			12 (3)	10 (1)
Plant	7	7	14 (2)	7.67 (10)
Reception	20	10	12 (3)	15 (3)
Computer Cluster			16 (3)	11 (1)
Print Room	15		16 (0)	12 (2)
Library			8 (0)	10 (0)
Workshop				9 (4)
Laboratory				16 (9)
Lift Lobby				8 (3)
Canteen	15			

Plug loads

Similar to the lighting loads, plug loads are defined by the appliances used by people, however, such data is not available on drawings as these items are typically not permanent components in a building. Instead, an approximation is made based on building audits and counting the pieces of equipment used in different space types. Because access to all spaces was not always possible, benchmarks were established for different spaces types, by counting the number of laptops/computers, screens and other type of small equipment. The benchmark wattage is used to determine the total plug load in a space. This is done for several spaces and then the average is taken for each space type and used as a modelling input, in **Table D5.**

Plug loads have a larger uncertainty than lighting loads, due to the method used for data collection and larger variability in loads as the actual load has a higher dependency on the amount of people present. Theoretically one person present in a space requires the lighting for a whole space or large part of a space to be on, whereas one person only uses one or two pieces of equipment in a space where normally perhaps 5 to 6 people are present. Plug loads were varied at 20% standard deviation.

Table D5: Plug load (W/m²) per space type

Space type	Office 17	Office 71	CH	MPEB
Office	30	18	16	25
Circulation	1.5	1.5	2	2
Storage	0	0	0	0
Lavatory	2	5	2	2
Meeting	12	12	12	12
Lift	0	2	2	2
Kitchen			12	12
Server	50		50	700
Shower	3		3	3
Lecture Theatre			10	18
Plant	50	50	50	50
Reception	10	8	16	12
Computer Cluster			16	30
Print Room	30		16	20
Library			10	12

Space type	Office 17	Office 71	CH	MPEB
Workshop				50
Laboratory				35
Lift Lobby				10
Canteen	25	25		

Building systems and control

Most spaces in the building are conditioned, heated (H) through radiators or and/or cooled (AC) through more active systems, such as air handling units or fan coil units, a few spaces are mechanically ventilated (mV), but most have a provision for natural ventilation (nV) through window openings, while lavatory spaces have only extract ventilation (eV). An overview of the type of conditioning per space type is shown in **Table D6.**

Table D6: Zone conditioning by space type

Space type	Office 17				Office 71				CH				MPEB			
Office	H	mV	nV		H	AC	mV	nV	H	AC	nV		H	AC	mV	nV
Circulation	H	nV			H				H				AC	mV		
Storage																
Lavatory	eV				eV				mV				H	mV	eV	
Meeting	H	nV			H	mV	nV		H	AC	nV		H	AC	mV	nV
Kitchen									H	AC	nV		H	nV		
Server	H												H	AC		
Shower	H	eV			H	eV			H	nV			H	mV	eV	
Lecture									H	AC	nV		H	AC	mV	nV
Plant																
Reception	H	AC	mV		H	AC			H	AC	nV		H	AC	mV	
Computer Cluster									H	AC	mV	nV	H	AC	mV	
Print Room	H	AC							H	AC	nV					
Library									H	AC	nV		H	AC	mV	nV
Workshop													H	AC	mV	
Laboratory													H	AC	mV	
Canteen	H	AC	mV		H	mV										
	*mV only for basement								*Heating in staircases							

Information regarding system capacities and performance coefficients is available from O&M manuals, but actual figures might retain a certain level of uncertainty. As suggested by Heo et al. (2015), uncertainty in actual system efficiency is quantified by industry standards for different system types. Uncertainties of 2% for boiler efficiency (Kemna, et al., 2007, p. 27) and 5% for air conditioners and heat pump efficiency were taken into account.

Mechanical ventilation

Mechanical ventilation is several spaces is available in MPEB, Office 71 and partially in Office 17 in the basement, assumed provision of tempered fresh air to different space types are given in **Table D7.** All design air flow rates were varied at 20% standard deviation according to a normal distribution. CH does have an air handling unit located outside of the building, but was out of operation during the building audits. According to the O&M manual it is supposed to provide fresh tempered air to several spaces in the basement, spaces without operable windows.

Table D7: Design airflow (ltr/s) rate for mechanical ventilation.

Parameter	Office 17	Office 71	MPEB
Office	8 (basement)	8	8
Circulation		0	4
Lavatory		10	10
LiftLobby, Workshop			8

Parameter	Office 17	Office 71	MPEB
Computer Cluster			6
Meeting		8	4
PrintRoom	8		4
Lecture Theatre, Laboratory, Kitchen, Library			4
Reception		8	4
Shower			10
Canteen	8	8	

Natural ventilation and infiltration

The principle role of ventilation is to provide an appropriate level of indoor air quality (IAQ) by removing and diluting airborne contaminants (CIBSE, 2005). This can be accomplished by using extract ventilation, whole building ventilation (supply and extract) or purge ventilation (used for activities that have intermittent high concentration release of pollutants, such as in a workshop or laboratory. In the building models purposeful (natural) ventilation is represented by allowing windows to be opened to provide a certain flow rate to a space.

For Office 17 the "AirflowNetwork" in EnergyPlus was utilised which allows for more detailed configuration of natural ventilation into the building. The amount of natural ventilation is determined window dimensions and criteria for opening the window. In particular, the discharge coefficient indicates the fractional effectiveness for air flow through a window or door a certain opening factor, where the opening factor is determined based on the dimensions and type of the window (pivoting horizontal/vertical). In addition, a minimum window opening factor is applied that determines that a window is at least open a certain percentage (minimum window open factor) as long as temperature and enthalpy indoor and outdoor conditions are met. Uncertainty in the discharge coefficient and minimum window open factor are taken into account as shown in **Table D8.**

Table D8: Natural ventilation uncertainty parameters for Office 17.

Parameter	Mean (μ)	Standard dev. (σ)
Discharge coefficient (DischargeC)	0.68	15%
Minimum window open factor (WinOpen)	0.3	15%

Note: both upper and lower limits are set to 3 standard deviations of the nominal value.

Implementing the AirflowNetwork is however considerably complex in larger buildings with numerous windows and spaces, natural ventilation was therefore simplified in Office 71, CH and MPEB. Instead, windows are opened when the following temperature criteria for inside (T_{in}) and outside (T_{out}) are met;

$$\text{IF } T_{in} > 23 \text{ AND } T_{out} < 30 \text{ AND } T_{in} - T_{out} < \Delta T \text{ THEN open window.} \quad (12)$$

The delta temperature is the temperature difference between T_{in} and T_{out} at which ventilation is shut off. A negative value allows ventilation to occur even if the outdoor temperature is above the indoor temperature, which is typically the case. Input parameters for the natural ventilation objects in the models are shown in **Table D9.**

Table D9: Natural ventilation model parameters.

Parameter	Mean (μ)	Standard dev. (σ)
Flow rate per person (m^3/s)	0.008	20%
Min indoor temperature (^0C)	23	1
Delta temperature (^0C)	-2	0.5
Min outdoor temperature (^0C)	18	1
Max outdoor temperature (^0C)	30	1

Note: both upper and lower limits are set to 3 standard deviations of the nominal value.

Infiltration in buildings can have a large influence on the heating and cooling demands of a building. Over the years, the air tightness of buildings has improved as its significance in reducing energy use and improving comfort was acknowledged. Since Part L 2002 an air permeability of 10 m^3/m^2h at 50Pa is required for commercial buildings in the United Kingdom. However, three of the case study buildings are built before 2002, in fact they are significantly older. CIBSE Guide A provides empirical values for the air infiltration rate for different office types. In general terms, a leaky building falls within the range of 20 m^3/h-m^2 @ 50Pa, while 10 m^3/h-m^2 @ 50Pa complies with Part L 2002 regulations, 7 m^3/h-m^2 @ 50Pa complies with 2005 regulations, 5 m^3/h-m^2 @ 50Pa is deemed a tight building, and 3 m^3/h-m^2 @ 50Pa a very tight building (CIBSE, 2007).

For CH and Office 71 infiltration values have been estimated by taking into account the age of the building, height, building type (naturally ventilated) and any recent refurbishments. Infiltration values for MPEB are based on actual test data performed after completion of the building. For Office 17, the AirflowNetwork in EnergyPlus has been utilised for ventilation and infiltration, therefore infiltration is calculated differently from the other buildings. More specifically, it is calculated using:

$$Q = (Crack\ factor) \cdot \left[\frac{\rho_o}{\rho}\right]^{n-1} \left[\frac{v_o}{v}\right]^{2n-1} \cdot C_Q \cdot \Delta P^n \qquad (13)$$

where, Q = air mass flow (kg/s), C_Q= air mass flow coefficient, ΔP = pressure difference across crack (Pa), n = air flow exponent, ρ and v are the air density and kinetic viscosity of air respectively. The air flow exponent in this equation has been varied to quantify the uncertainty in the air infiltration for Office 17. Air permeability inputs are shown in **Table D10.**

Table D10: Air permeability parameters

Parameter	Office 17	Office 71	CH	MPEB
Air permeability (m^3/h-m^2 @ 50Pa)	-	12	16	8
Flow per exterior surface area (m^3/h-m^2 @ 4Pa)	-	0.00065 (σ = 20%)	0.00075 (σ = 20%)	0.00043 (σ = 10%)
Flow exponent	0.65 (σ = 0.05)	0.65	0.65	0.65
Source	Estimate	Estimate	Estimate	Based on test

Note: Flow per exterior surface area is obtained by multiplying the air permeability by 40.65 / 500.65, using a flow exponent of 0.65.

In EnergyPlus the flow per exterior surface area is used which requires the air permeability values to be converted from 50Pa (used in airtightness tests) to a reference value of 4Pa, the flow is then applied to all external surfaces. In addition, the DOE-2 air change method is used to represent a change in summer and winter conditions.

Hot water use

Energy use from hot water services is a factor with a large uncertainty. Used for showering, washing hands and drinking (typically from zip taps), the amount depends on the volume to be heated, peak loads and duration of use. Both ASHRAE and CIBSE provide peak hourly and daily demands for various categories of buildings, whereas ASHRAE provides data on hot water demand per fixture for different building types, as shown in **Table D11.**

Table D11: Hot water demand for office buildings.

Hot water daily use (l/person)	CIBSE[1] / ASHRAE[2]
Min daily average	- / 1.5
Daily average	14 / 3.8
Max daily average	26 / 7.6
Fixtures (l/hr)	
Basin, public lavatory	7.6
Shower	114
Kitchen sink	76

Average daily hot water demands were used to determine the use per fixture as input in the building energy model. Electric water heaters provide hot water to fixtures in the showers, offices, lavatories and kitchens. In the building energy model, the peak flow rate (m^3/s) and their schedules were defined and determine the hot water use load per fixture. Energy use for hot water heating is based on the efficiency of the hot water heating system. There is large difference between the figures found in the ASHRAE and CIBSE documentation, therefore a large uncertainty is taken into account for the hot water use loads in the buildings. Final calculated loads for the buildings are given in **Table D12.**

Table D12: Hot water use input parameters and variability (σ in %) on peak flow rate

No. of Fixtures	Office 17	Office 71	CH	MPEB
Office	3	4		
Canteen	1	1		
Showers	3	2	8	
Lavatory sinks	9		57	
Kitchen sink / zip tap	2		10	
Cleaners' tap			7	
Laboratory				10
Workshop				10
Peak flow rate (l/hr)	140	329 (±20%)	774 (±20%)	684 (±20%)

Weather data

The building model was calibrated towards measured energy use, to accurately perform such a calibration it is essential that historical meteorological factors were used. Weather files were obtained covering the measurement periods. The weather files are based on data collected in the City of London, with the following latitude and longitude coordinates; 51.5168° N, 0.0987° W. This location is in close vicinity (within a 4 km radius) to all case study buildings.

APPENDIX E. BUILDING MATERIALS

Table E1: Building constructions and opaque materials properties in CH.

Material (out > in)	Thickness (mm)	Conductivity (W/mK)	Density (kg/m³)	Specific heat (J/kgK)	Thermal resistance (m²K/W)
External walls					
Brick	200	0.9	1845	840	
Air cavity	50				0.15
Brick	100	0.9	1845	840	
Air cavity	10				0.15
Insulation Celotex	40	0.025	50.4	840	
Plasterboard	12.5	0.17	670	800	
MDF board	10	0.3	750	1700	
Exposed floor					
Cast Concrete	300	0.2	1900	930	
Concrete deck	50	0.2	1900	930	
External roof					
Roof Membrane	10	0.16	1121	1460	
Cast Concrete	300	0.2	1900	930	
Insulation Celotex	40	0.025	50.4	840	
Plasterboard	12.5	0.17	670	800	
Internal partition					
Plasterboard	12.5	0.17	670	800	
Air cavity	10				0.15
Insulation partition	25	0.035	33	840	
Air cavity	10				0.15
Plasterboard	12.5	0.17	670	800	
Internal ceiling					
Carpet					0.1
Plywood	40	0.15	608	1630	
Air cavity (floor)	130				0.18
Cast Concrete	300	0.2	1900	930	
Air cavity (ceiling)	350				0.18
Plasterboard	12.5	0.17	670	800	

Table E2: Window materials in CH.

External windows	Thickness (mm)	Conductivity (W/mK)	Slat width (mm)	Slate separation (mm)	Visible transmitta nce	Solar heat gain coefficient	U-factor (W/m²K)
Secondary glazing	6 / 6				0.71	0.54	2.91
Venetian blinds	1	220	0.025	0.01875			

Table E3: Building constructions and opaque materials properties in Office 71.

Material (out > in)	Thickness (mm)	Conductivity (W/mK)	Density (kg/m³)	Specific heat (J/kgK)	Thermal resistance (m²K/W)
External walls					
Brick	90	1.5	2083	921	
Air cavity	12				0.18
HW Concrete block	90	1.5	2234	837	
Gypsum plasterboard	19.1	0.15	801	837	
Exposed floor					
London clay	750	1.41	1900	1000	
Brickwork	250	0.84	1700	800	
Cast concrete	100	0.64	2000	1000	

Material (out > in)	Thickness (mm)	Conductivity (W/mK)	Density (kg/m³)	Specific heat (J/kgK)	Thermal resistance (m²K/W)
EPS insulation	63.5	0.025	30	1400	
Synthetic carpet	10	0.06	160	2500	
External roof					
Stone chippings	10	0.96	1800	1000	
Cast concrete	350	0.64	2000	1000	
Internal partition					
Plasterboard	12.5	0.16	600	800	
Brickwork	100	0.62	1700	800	
Plasterboard	12.5	0.16	600	800	
Internal ceiling					
Carpet					0.1
Cast Concrete	200	1.4	2100	840	

Table E4: Window materials in Office 71.

External windows	Thickness (mm)	Conductivity (W/mK)	Visible transmittance	Solar heat gain coefficient
Pilkington	6	1.06	0.71	0.69
Argon	12			
Pilkington		1.06	0.71	0.69
Internal shading	0.5	220		

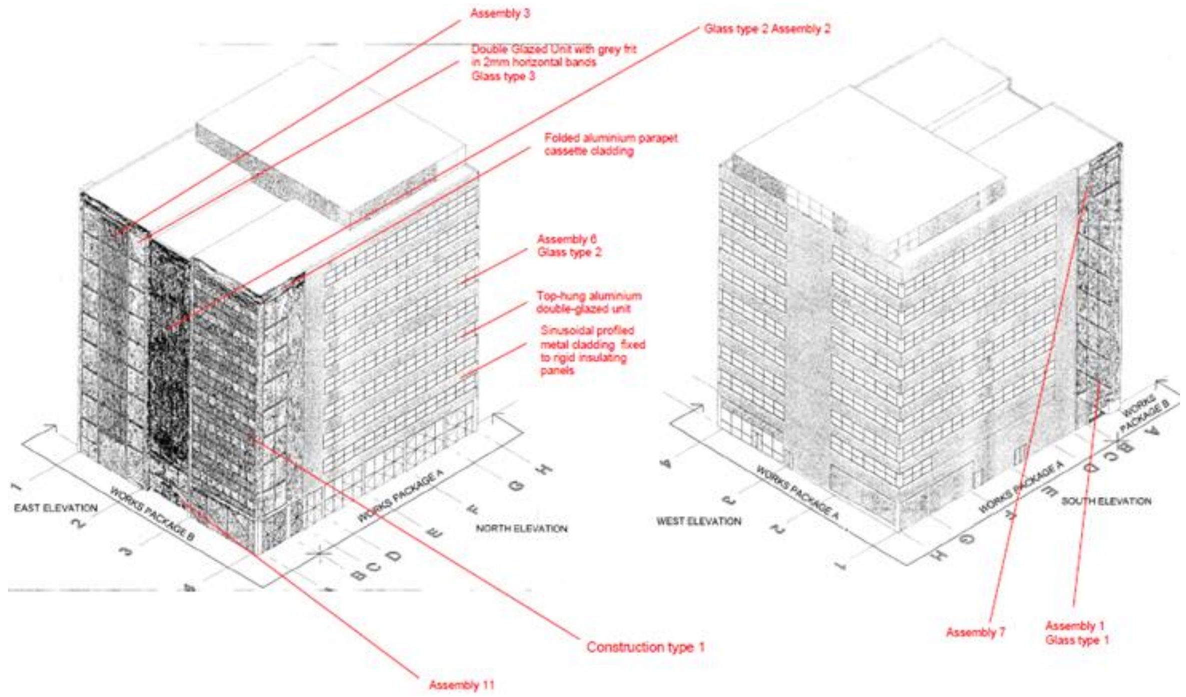


Figure D1: Location of different façade materials in MPEB.

Table E5: Building constructions and opaque materials properties in MPEB.

Material (out > in)	Thickness (mm)	Conductivity (W/mK)	Density (kg/m³)	Specific heat (J/kgK)	Thermal resistance (m²K/W)
External walls					
Sinusoidal Corrugated Sheet	0.7	50	7850	490	

Material (out > in)	Thickness (mm)	Conductivity (W/mK)	Density (kg/m³)	Specific heat (J/kgK)	Thermal resistance (m²K/W)
Air cavity	25				0.15
Mineral fibre RW2 Rockwool	150	0.035	33	840	
Sinusoidal Corrugated Sheet	0.7	50	7850	490	
Plasterboard	12	0.17	670	800	
External walls (ground)					
Concrete construction	600	2.15	1900	930	
Mineral fibre RW2 Rockwool	60	0.035	33	840	
Exposed floor					
Mineral fibre RW2 Rockwool	60	0.035	33	840	
Concrete construction	500	2.15	1900	930	
External roof					
Roof Membrane	10	0.16	1121	1460	
Mineral fibre RW2 Rockwool	120	0.035	33	840	
Plywood	25	0.15	608	1630	
Sinusoidal Corrugated Sheet	0.7	50	7850	490	
Air cavity	350				0.18
Plasterboard	12	0.17	670	800	
Internal partition					
Plasterboard	30	0.17	670	800	
Insulation partition	70	0.035	33	840	
Plasterboard	30	0.17	670	800	
Internal ceiling					
Carpet					0.1
Plywood	40	0.15	608	1630	
Air cavity (floor)	130				0.18
PreCast Concrete	150	0.2	1900	930	
Air cavity (ceiling)	350				0.18
Plasterboard	12.5	0.25	670	800	

Table E6: Window materials in MPEB.

External windows	Visible transmittance	Solar heat gain coefficient	U-factor (W/m²K)
Type A SingleClear StairCase	0.95	0.7	5.5
Type B LowE Double	0.52	0.5	1.5
Type C/D LowE Double N/W/S	0.71	0.5	2.2
Typce C/D LowE Double Main Entrance	0.54	0.7	1.9
Type D Low E Double Lift CurtainWall	0.71	0.5	1.8
Type E Double North CurtainWall	0.61	0.8	1.9

APPENDIX F. ENERGY USE OVERVIEW

For all case study buildings, there is some electric sub-hourly metering data available. Buildings are serviced with distribution boards (DB's) to provide electricity to lighting and appliances in spaces and to HVAC equipment. Typically, several distribution boards are located on each floor and are connected either directly to a low-voltage power distribution (LV-panel) or indirectly via bus-bars. The LV-panel typically has one or several main electrical incomers, which should aggregate all electricity use in a building. Lighting and appliances in spaces are typically connected to the same distribution board, while in some cases these are configured to be on separate meters, it is otherwise difficult to correctly separate their electricity use. For predicted data, this can be more easily accomplished, and possibly energy use for each individual component in a building could be separated. For the metered data however, some assumptions in regards to the ratio of electricity use from power and lighting had to be established in order to make a comparison viable. Not all distribution boards have a meter installed, to fully capture and understand all energy end-uses in the building it was deemed necessary to install additional short-term monitoring. In **Figure F1** and **Figure F2** two distribution boards are shown that are sub-metered and short-term monitored respectively.

Figure F1: Distribution board with permanent meter located next to it.

Figure F2: Distribution board with short-term monitoring equipment.

Measured energy use can identify typical patterns of use, which will feed back into the building energy simulations and the underlying assumptions made. Typical profiles for different energy end-uses were identified, such as when the lighting is on, when small power equipment is being used and when systems are running. This is not always as straightforward when working with many meters. The aim was to capture the highest level of data granularity in the sub-metered data, this meant keeping the distribution boards disaggregated where possible, while at the same time making sure that different boards capture similar energy end-uses as those that will be predicted by the model. However, most electricity meters capture multiple building components, which can make it difficult or impossible to establish all energy end-uses as could be calculated by a building energy model.

Office 17

Gas and energy data from April 2012 to October 2017 were available, electricity is disaggregated into equipment, lights, lift, AC, server, canteen and print room. The data contained erroneous (peaks) and missing data points, corrected for by replacing them with the average half-hourly energy use over their respective year at a similar time and day. Gas energy use is missing for the winter of 2014/2015 (due to a faulty meter), but is available for a full year in 2013, the year used for comparison to predictions. Monthly energy use over the measurement period is shown in **Figure F3.** The dominant energy uses in 2013 were from gas (~37%), equipment (~35%), lighting (~10%) and server (~10%).

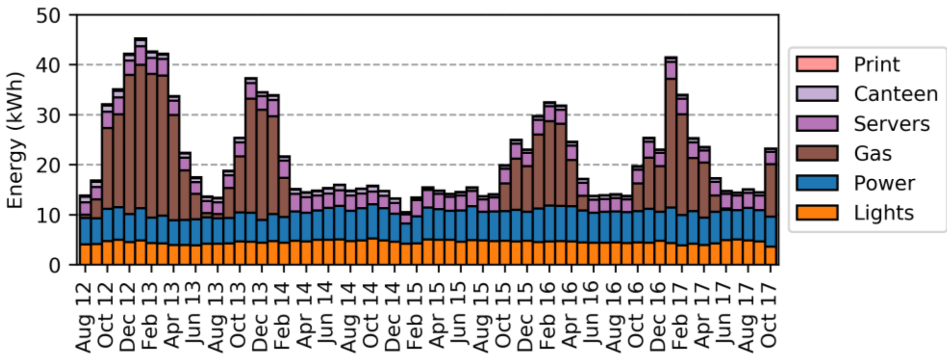


Figure F3: Monthly energy use in Office 17.

There is a large difference between weekday and weekend day energy use as shown in **Figure F4.** Equipment energy use high during the night and weekend. It is likely that a lot of equipment is being left on, whereas the server is run continuously at a constant load. The baseload during the night and weekend is found to be around 35% of the peak load during the day, excluding the gas use.

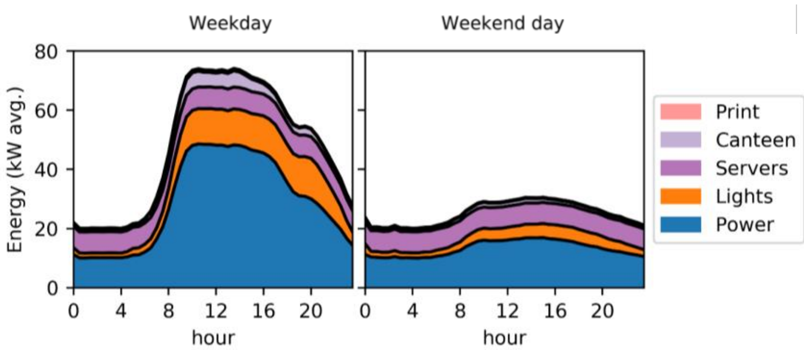


Figure F4: Energy use for a typical weekday and weekend day in Office 17, based on 2013 excl. gas use.

The typical weekday and weekend day profiles are based on a whole year of data, averaging out gas use, which should be proportionality higher during the winter than the summer. Lighting and power fluctuate during the day and night, while the server is operated 24/7. Two boilers use gas mainly during the winter months and is therefore separated from the other end-uses, shown separately in **Figure F5.** Gas use made up 50% of total energy use in 2013 due to the heating system being on constantly, this was notified to the facilities manager. During subsequent years, the meter turned out to be faulty and no data was collected for the winter of 2014 and 2015, however the meter was fixed during 2016 and coincidentally the boiler was replaced with a more efficient one, while the heating system was adjusted to operate on a time-basis, saving a significant amount of gas use.

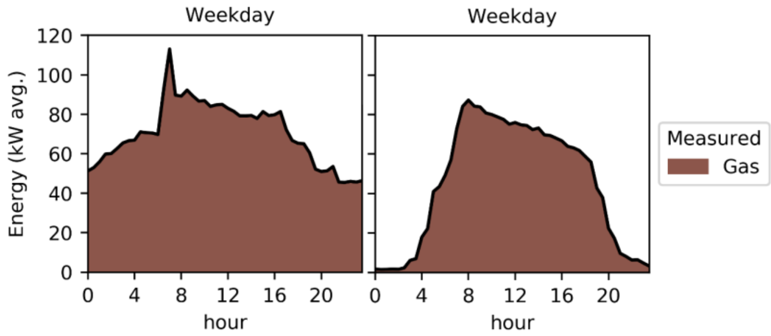


Figure F5: Gas energy use for an average weekday during January to April 2013 (left) and 2016 (right).

Office 71

For Office 71 a sub-metering system is in place that disaggregated energy use into lighting, power, air conditioning (AC), lifts, canteen and gas use. Lighting, power and AC meters are disaggregated per floor. Unexpectedly, the metering system did not extend to the outdoor VRF system and air-handling unit, and these are therefore could not be analysed. Energy use data was available from April 2012 to October 2017, monthly energy use for this period is shown in **Figure F6.**

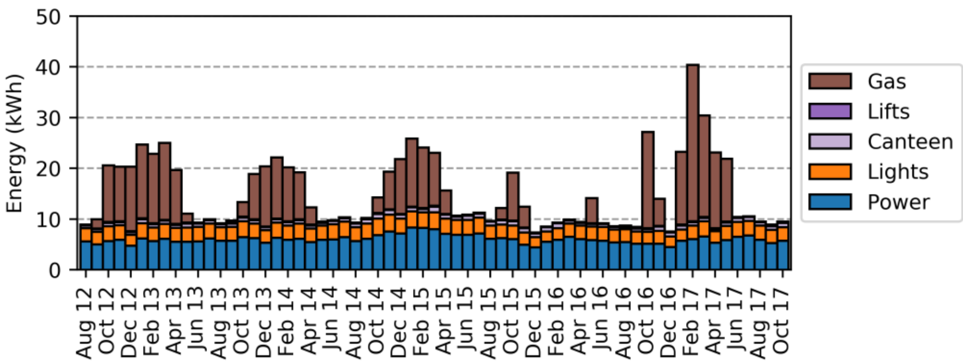


Figure F6: Monthly measured energy use in Office 71.

Electricity use throughout the year is fairly stable, gas use however fluctuates significantly, being more than 50% of monthly energy use during the winter months. Dominant energy uses are from gas (~38%), equipment (~36%) and lighting (~18%). Air-conditioning (mainly the FCUs in the spaces) use a negligible amount of energy use, even though these systems are part of the HVAC strategy. It is likely that the AHU on the roof provides the necessary cooling conditioning to the spaces, whereas the radiators provide the heating. However, the mechanical plant room on the roof, nor the VRV heat pumps were measured separately, so actual building energy use cannot be fully established without these components. Therefore, the model should separate these energy end-uses, and calibration will only focus on the energy use that is measured. Energy use for a typical weekday and weekend day provide more insights into the patterns of use, shown in **Figure F7** for electricity use. Equipment is on until late, indicating that some people are still present during the evening. Lighting energy use is relatively low during the night, but turned on early in the morning, due to cleaning, and turned off late.

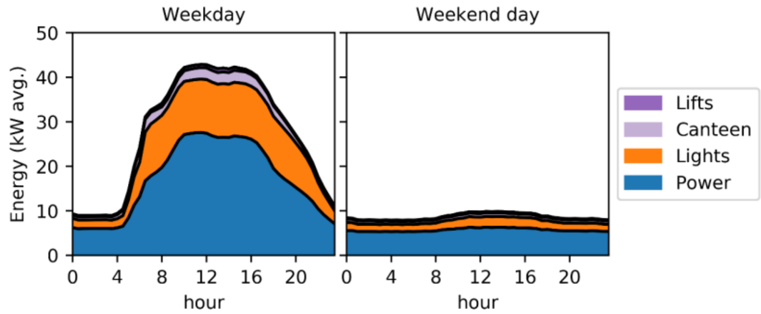


Figure F7: Energy use for a typical weekday and weekend day, based on 2014.

CH

The distribution schematic for CH is shown in **Figure F8**, most of the meters in the building are sub-metered (green), additional short-term monitoring (orange) was deemed necessary to capture all electrical energy use on the different floors. Meters not monitored, measure negligible loads or are logged by a parent meter. On some of the office floors, the lighting and power are segregated, used for separate analysis of their loads.

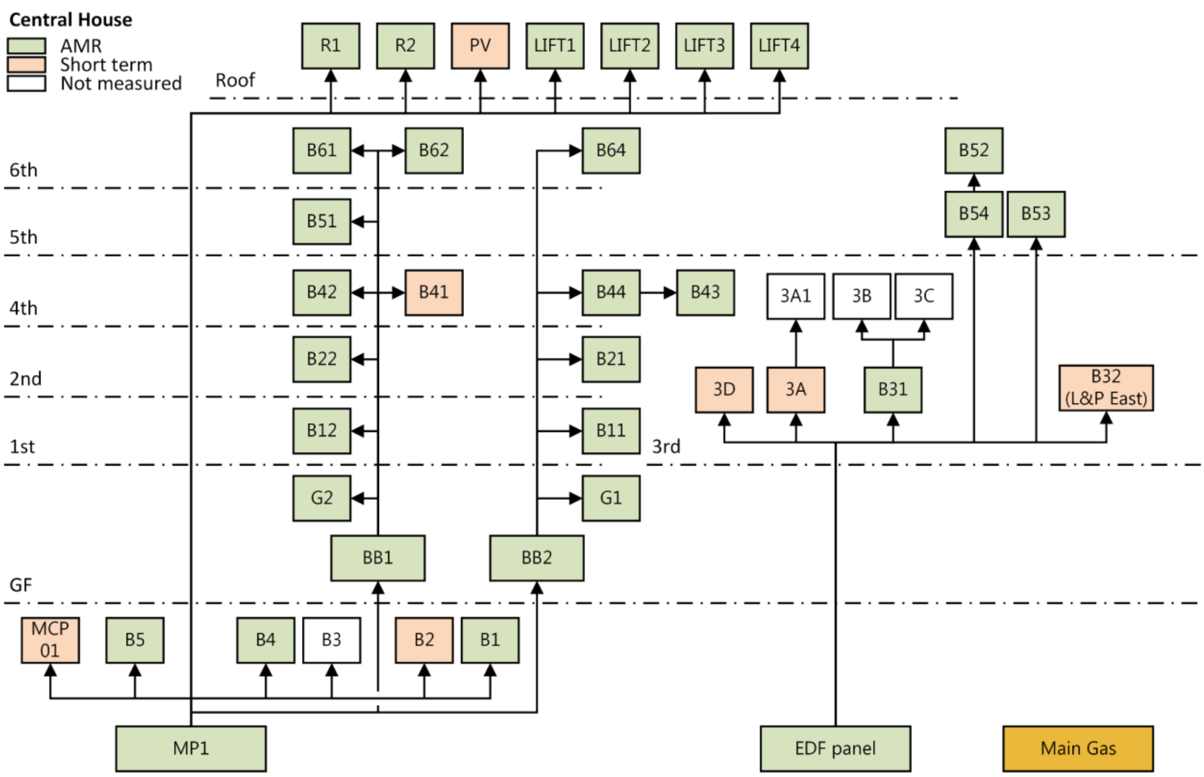


Figure F8: Electrical distribution schematic for CH.

In CH, the level of disaggregation is not exhaustive of all the energy using components in the building. Many components are not monitored separately, some of which would be useful in order to understand how they were used. Distribution boards serve different parts of the building, which parts of the building they serve and the type of components they serve need to be determined in order to make predicted and measured energy use comparable. Lighting and equipment are in some cases separated by floor, their boards being connected to the BB1 and BB2 bus bars, however these often also provide electricity to other types of equipment, such as FCUs and electric water heaters, ideally separated, as they are large energy users. Another example is the systems, in CH conditioning is primarily provided by VRF heat pumps for heating and cooling. The heat pumps are connected to R1 and R2, but heating and cooling cannot be separated, as is typically done in building energy model. For several meters lighting, power and systems are separately measured, lifts are connected to separate distribution boards. Finally, energy use was aggregated in; lighting and power (available per floor), systems, server, lifts and gas use, shown in **Figure F9**.

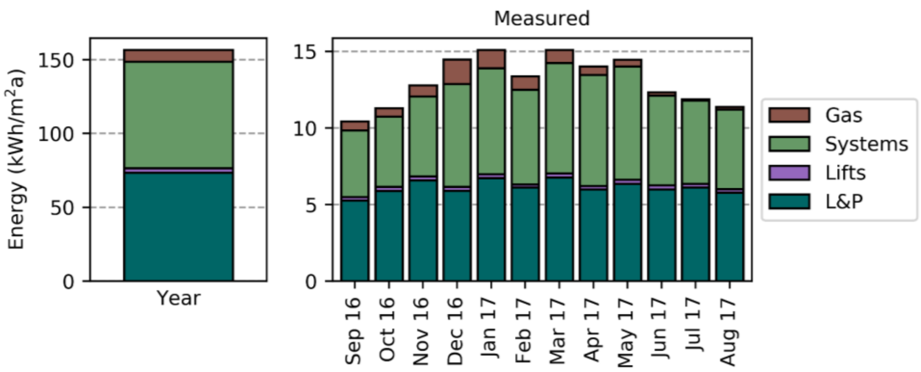


Figure F9: Monthly measured energy use in CH.

A clear distinction can be made between the size of different energy end-uses, systems and lighting and power are dominant in CH, as they are in most buildings. Although lighting and power were measured separately on some floors, an accurate distinction at the building level could not be made and for calibration purposes both were combined as L&P. Nevertheless, some further analysis into the separate lighting and power meters proved helpful in determining their typical profiles of use. Energy use from gas, the server and lifts were relatively marginal, but could be clearly distinguished. Another observation is the large difference in energy use throughout the months, gas use is slightly higher during the winter (only serves the radiators).

Aggregated electricity energy use is visualised at a higher temporal granularity in **Figure F10**, plotted as a heat map, electricity use during January and February 2017 is represented by hourly blocks, and is aggregated into daily totals shown in the bottom graph. Electricity use during the night and on weekdays are similar. Furthermore, there is a clear distinction between occupied and unoccupied hours, where night time use is about 3/5$^{\text{ths}}$ of electricity use during day time. People seem to be mainly present from 9am to 6pm, with a few coming in earlier (from 6am) and leaving later (until 9pm).

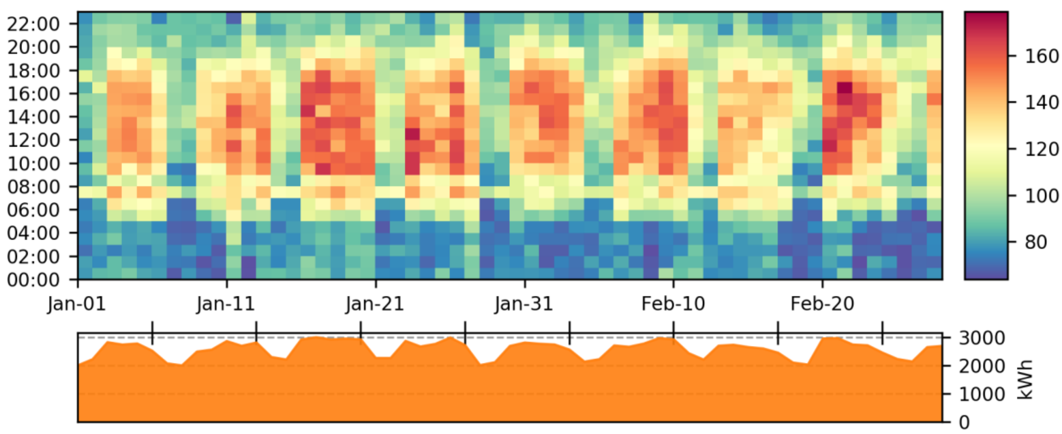


Figure F10: Heat map showing total energy use (kWh) for a month of August for CH.

Figure F11 shows the variation between day and night-time electricity use, typically in compliance modelling the difference between day and night-time use be significantly larger as they assume more ideal (less energy consuming) profiles of operation. Evidently, systems energy use (VRF system) and lighting and power energy use are significant during the night and the weekend, this will need to be accounted for in the energy model.

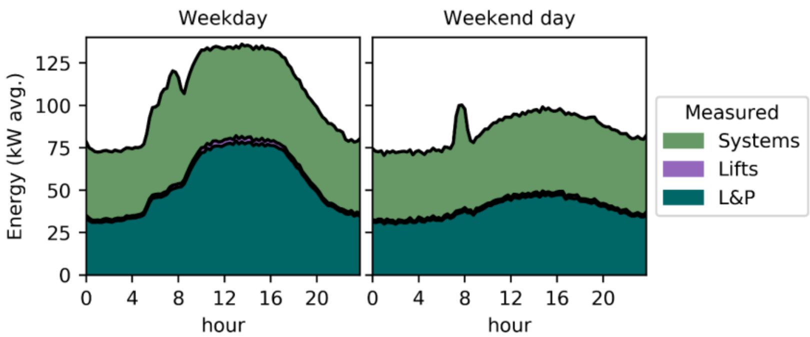


Figure F11: Typical weekday and weekend day for CH.

On the roof of CH, several PV panels generate renewable electricity for the building. Generated electricity from PV is however negligible (<0.5% of total) compared to total energy use and has therefore been neglected in further analysis.

MPEB

The distribution schematic for MPEB is shown in **Figure F12**, most of the meters in the building are sub-metered, and additional short-term monitoring was used to retrieve data on lighting and power data on several representative floors. The building electricity separates into two main meters, LV1 and LV2, LV1 serves the workshops in the basement, chiller 1, lifts and mechanical plant room (MCCB03) on the ground floor. LV2 serves two bus bars; BB1 and BB2, in addition to chiller 2, mechanical roof plant room (MCCB01) and the server room DB409, which connect to it directly. The bus bars in turn serve lighting and power on separate floors, but none of the meters were logged initially. Due to difficulties with the installation of the heat meter, no data has been collected on the district heating system, therefore this case study focusses solely on the electricity use in the building.

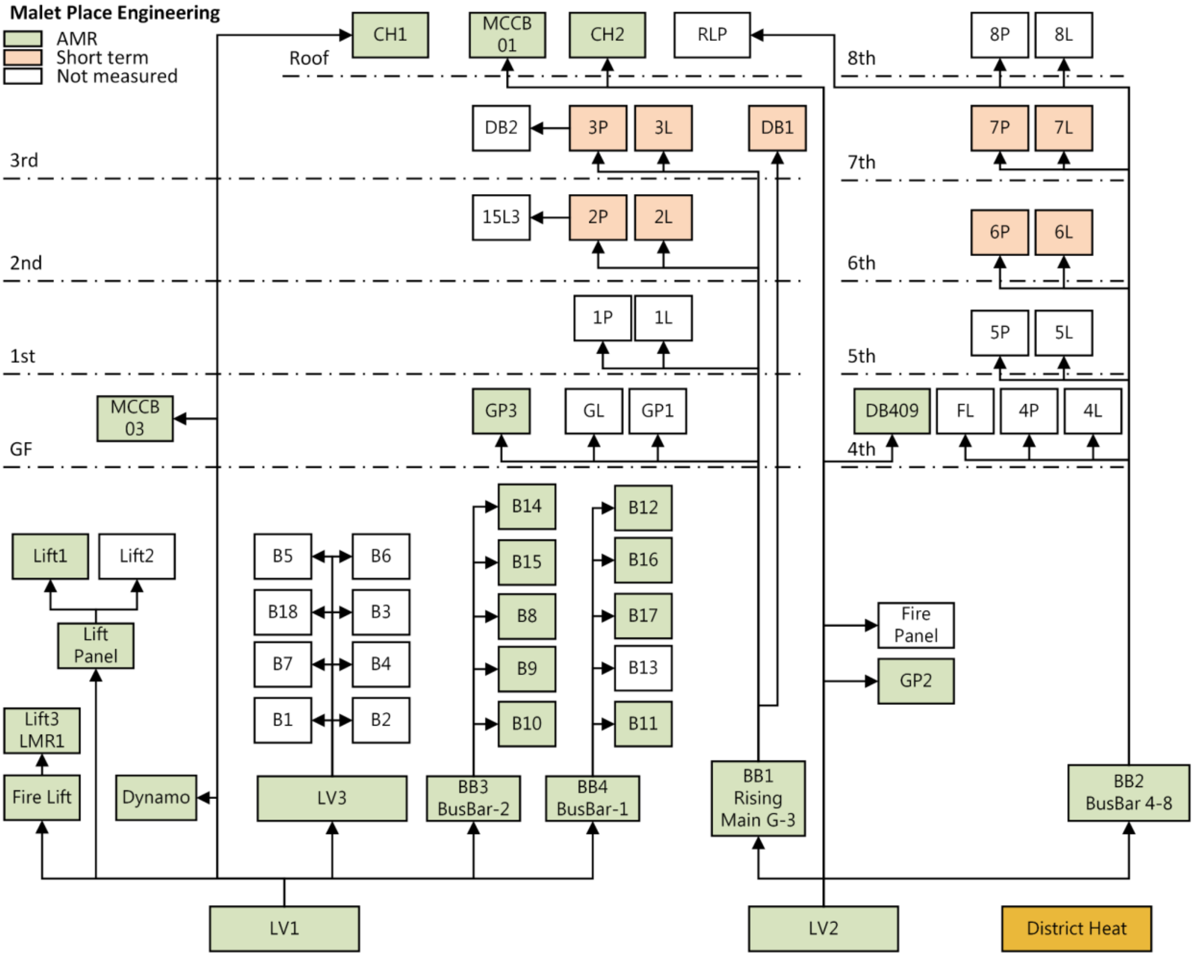


Figure F12: Electrical distribution schematic for MPEB, each rectangle corresponds to a distribution board, bus bar or low-voltage mains.

Systems energy consists of two chillers, two mechanical plant rooms and the dynamometer (uses minimal amount of energy, but is considered to be part of the systems), the chillers are considered separately. All the distribution boards in the basement, apart from the dynamometer, lifts and GP2 were taken together and are defined as workshops. DB409 and GP3 serve two large servers, in particular DB409 is a significant energy user and additional FCUs besides the computer cluster are connected to it. GP3 has a somewhat smaller computer cluster, but serves additional spaces such as the VR room, entrance space with over door heater and an additional office on the ground floor. Measured energy use is aggregated from the sub-metered and short-term monitoring data in; lighting and power, chillers, systems (mechanical plant rooms), workshops, servers (DB409 and GP3) and lifts, collected from September 2016 until April 2017, shown per month in **Figure F13**.

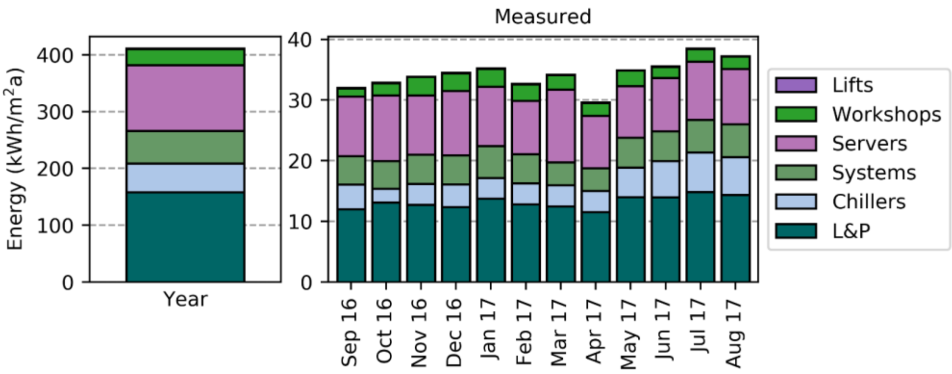


Figure F13: Monthly measured energy use in MPEB.

Nearly half the electricity in the building is consumed by lighting and power, system energy use including the chillers make up only about one quarter of the total, where servers consume only marginally less. There is similar difference between the total amount of energy used monthly when compared to CH, but it is less apparent in MPEB. **Figure F14** shows the electricity use during a typical weekday and weekend day for the year, indicating a time operation between around 7am and 7pm, in particular so the systems. Furthermore, it is clear that MPEB operates 7 days a week as there is no difference between weekdays and weekend days. Other end-uses remain relatively stable throughout the day and night.

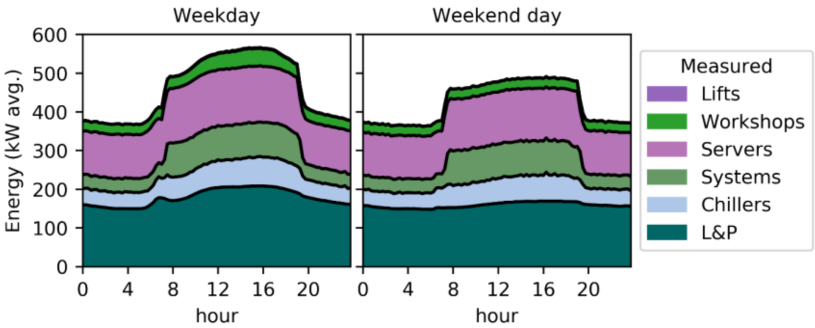


Figure F14: Electricity use for a typical weekday and weekend day.

Systems located in the mechanical plant rooms are the air handling units and auxiliary equipment, conditioning the occupied spaces. Whereas, servers conditioned 24/7 to manage high internal gains from the computer clusters, chillers therefore run continuously, as shown in **Figure F15**. However, if AHUs are time controlled, a difference in cooling load would be expected during occupied and unoccupied hours, only marginally evident in the graph.

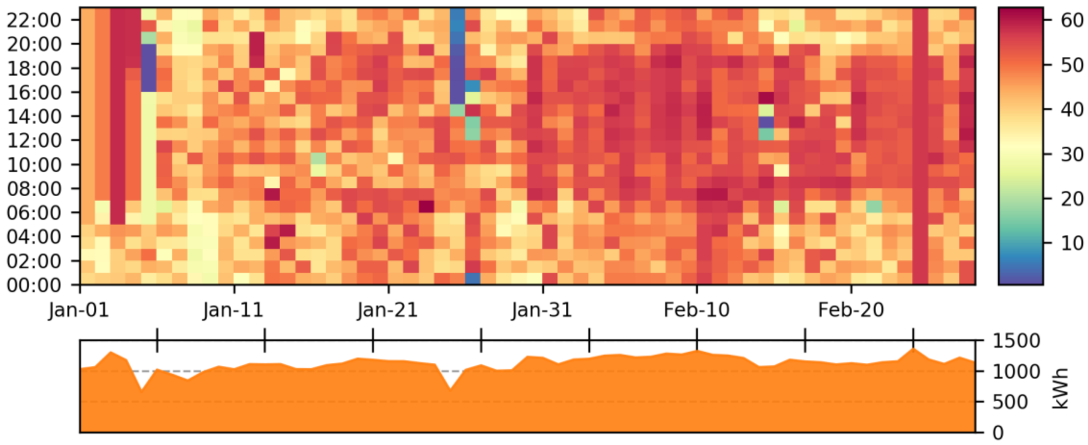


Figure F15: Chiller electricity consumption (CH1 and CH2)

Lighting and power were not sub-metered separately, therefore additional short term monitoring was put in place to retrieve information on their typical energy use. Four floors were measured; the 6th and 7th floor were deemed representative of other office floors, and the 2nd and 3rd floors contain computer labs and laboratory spaces, likely to be more energy intensive due to higher equipment loads.

Finally, lift energy use (from three lifts) is analysed, daily energy use during the measurement period is shown in **Figure F16.** There is some fluctuation throughout the seasons, where considerably less use is made of the lifts during March and April 2017. The baseload or standby energy use is also quite significant, during unoccupied hours the lifts still use about 30% of their daily peak loads. Relatively however, lift energy is negligible, accounting for less than 0.5% of total energy use.

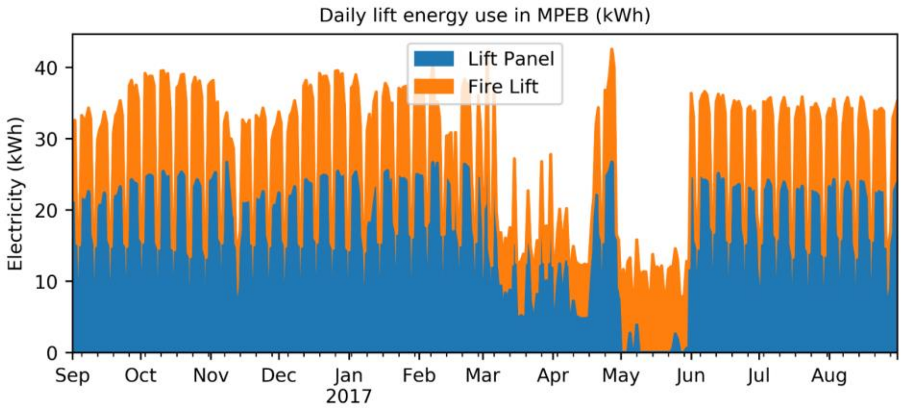


Figure F15: Daily energy use (kWh per day) from lifts in MPEB.