Reducing the performance gap using calibrated simulation models

NISHESH JAIN

UCL Institute for Environmental Design and Engineering
Bartlett School of Environment, Energy and Resources
University College London

October 2020

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Supervisors:
Professor Dejan Mumovic
Dr Esfand Burman
Professor Michael Davies

Declaration

I, Nishesh Jain, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my research supervisors, Prof. Dejan Mumovic, Dr Esfand Burman and Prof. Michael Davies, for providing invaluable guidance throughout the research. Their vision, sincerity and motivation for the subject have deeply inspired me. Their invaluable insights have informed my work and it was truly a privilege and honour to work under their guidance. I would also like to thank Dr Andy Tindale, my industrial supervisor, and the entire DesignBuilder Software team at the Stroud office for the technical insights and software guidance in completing various simulation tasks.

This thesis was closely aligned with The 'Total Performance' of Low Carbon Buildings in China and the UK ('TOP') project funded by The Engineering and Physical Sciences Research Council (EPSRC) (Grant Reference Code: EP/N009703/1). Most of the case study building monitoring data was obtained as a part of the 'TOP' project and therefore I would like to acknowledge the tireless contribution of Dr Esfand Burman and Dr Samuel Stamp, along with all others involved in that project, in collecting all the necessary data and information from the case study buildings. I would also like to thank the facility managers and staff of the case studies who helped in the understanding of the building and its systems.

I would like to gratefully acknowledge the financial support I received in completing this research from UCL Overseas Research Scholarships (UCL-ORS), DesignBuilder Software Limited. I extend my heartfelt thanks to their continued support. This thesis could not be completed without their unwavering commitment to my research. UCL Centre in Virtual Environments, Interaction and Visualisation (VEIV) also provided support during the initial stages of research.

Special thanks go to all the colleagues at UCL IEDE whose support acted as fuel to my research journey and made it such an enjoyable learning experience.

Finally, I must thank my parents for their love, prayers and constant support for my education and future. I can say, without a doubt, that I will forever owe all my successes and achievements to them.

ABSTRACT

Buildings have a significant impact on the environment. Construction of buildings and their operation accounts for 36% of global final energy use and 40% of energy-related carbon dioxide (CO_2) emissions. Also, as per the 2019 International Energy Agency (IEA) and United Nations Environment Programme (UNEP) report, the building sector has a strong potential to provide long-term energy and greenhouse gas emission savings without high financial costs.

Building performance simulation tools, ranging from steady-state calculations to dynamic simulation methods, can calculate the anticipated energy consumption of a building with adequate levels of accuracy. However, there is considerable evidence to suggest that buildings underperform post-completion when compared against the expected performance prediction during the design-stage. The difference between the actual operation and the design intent is termed the 'performance gap'. While the energy performance gap in buildings is a well-known phenomenon, its in-use interpretation is quite vague. It is important to understand the basis of assumptions and protocols used in design-stage performance calculations to assess the causes of the performance gap.

In the context of the performance gap, energy performance is generally the most emphasised. The gap, however, is not limited to energy – it also applies to indoor environmental quality (IEQ) parameters, such as temperature and air quality. Moreover, the pursuit of energy efficiency may have the unintended consequence of compromising IEQ, thereby requiring a comprehensive approach to performance assessment. It is therefore important to consider energy and IEQ performance issues together.

This thesis contributes to an improved understanding, quantification and resolution of performance gap related issues by using a novel simulation-based approach that enables systematic identification and classification of the root causes of the performance gap. A new measurement and verification (M&V) framework that is underpinned by building performance simulation and calibration is proposed.

A key aspect of this new methodological framework is the identification and separation of the three types of performance gaps because of:

1. Use of inappropriate design-stage calculation methods (such as those used for regulatory compliance),
2. Technical issues with the building, its systems and their operations, and
3. Operational changes that the building has gone through to meet its functional requirements.

For the first type of performance gap, CIBSE TM54 (CIBSE, 2013a) already provides guidance to reduce the perceived gap and enable improved estimates of building performance during the ***design-stage***. This thesis focuses on the understanding of ***operational-stage*** issues and their detailed causes, related to the second and third types of the performance gap.

This thesis is the first study that systematically defines, identifies and separates,

- the technical issues that cause the performance gap between design intent and actual operation, and
- the deviations of operating conditions from the design that are driven by the building's function and occupancy.

This is achieved by integrating the conventional post-occupancy performance assessment approach with building performance modelling and evidence-based model calibration. Another addition to the conventional approach, explored in this study, is the incorporation of IEQ. The issue of IEQ is addressed in two ways: first, by using zonal temperatures for calibration cross-validation, and second, by assessing the energy-related unintended consequences of IEQ underperformance which may happen during building operations.

The calibrated simulation models are *operationally accurate virtual representation* of the actual building and can help to isolate the performance issues and validate the findings. The new framework proposed in this thesis is better suited than conventional M&V protocols such as ASHRAE (American Society of Heating Refrigerating and Air Conditioning Engineers) Guideline 14 and IPMVP (International Performance Measurement and Verification Protocol). These conventional M&V protocols also propose a calibration-based approach, but they generally focus on broad statistical requirements and are not tied to a framework for a procedural verification of all the most important issues that can cause the performance gap. It is likely that using these conventional protocols will identify some key issues during investigations while leaving other potential issues hidden.

The guidance on calibration and validation provided in conventional M&V protocols is commonly used for all model calibration exercises. However, the conventional protocols were developed for calculating energy savings in retrofit applications, and the calibration criteria defined in them are mainly for checking the accuracy of building-level energy use totals. The calibration criteria do not check for the uncertainty or the accuracy of dependent parameters, such as zone temperatures and other environmental outputs, which could cross-validate the model. Mathematically, meeting just the statistical criteria for building-level energy use totals in a highly parameterized model and an under-determined search space can lead to unrealistic solutions also being validated.

To better support the calibration accuracy with the new proposed M&V framework, advanced model validation criteria have also been developed. New multi-level calibration criteria are proposed, which factors in data quantity, quality and granularity. In this new advanced validation criteria, the current industry standard of monthly energy use checks is the lowest level of calibration, with higher levels requiring detailed checks, using granular and disaggregated energy use. However, all levels of calibration require minimum dependent parameter checks, such as IEQ checks for typical zone temperatures. Dependent parameter checks are desirable in model calibration; however, current statistical criteria used for calibration are not suitable for these checks. Revised and new metrics and thresholds are proposed and explored in this thesis for use in advanced calibrated model validation checks.

Beyond the use of IEQ parameters (e.g. zone temperatures) in model calibration, another area of focus of this thesis is the unintended IEQ underperformance captured during the monitoring. The scope of this assessment is limited to underperformance in IEQ parameters linked to achieving high energy efficiency objectives, thermal comfort and indoor air quality (IAQ). Amongst the various IEQ parameters, thermal comfort and IAQ have complex and dynamic interactions with buildings energy end-uses. Comprising of multiple factors, which are both subjective and empirical, thermal comfort and IAQ performances have a high interrelation with the energy performance objectives. Therefore, along with conventionally tracked parameters of temperature and CO_2, additional IAQ parameters (not used during the calibration process), such as NO_2, $PM_{2.5}$ and PM_{10}, are analysed to enhance the understanding of unintended energy-related IEQ underperformance.

The new methodology proposed in this study is applied to five case study buildings across four building sectors – offices, schools, hospitals and apartment blocks. These buildings represent a large cross-section of the UK building stock and, therefore, can provide useful insights into the issues in the construction sector that drive the performance gap. While detailed performance assessment and advanced validation is done for all five case study buildings using the proposed framework, in one case study building the multi-level calibration checking criteria is also fully explored.

Using this methodology on the various case study buildings, cross-sectoral lessons, related to root causes of the energy performance gap and applicable in the wider industry context, are uncovered. Linking to the three types of performance gaps mentioned earlier, analysis of the results shows that, in most of the case studies, some of the energy performance gap is the ***perceived gap*** (related to point 1: use of inappropriate design-stage calculation methods) or is because of ***operational changes*** (related to point 3: changes that the building has gone through to meet its functional requirements). However, the most critical cause of the gap is due to ***technical issues*** (related to point 2: issues with the building, its systems and their operations) identified across the case studies. These issues were either design errors, improper construction and installation, poor commissioning or shortcomings in building systems and the use of new low-carbon technologies. It was observed that long-term involvement (with responsibility for the operational performance) of the design and construction teams are effective in lowering performance gaps.

Issues related to IEQ were also observed across the case studies, such as overheating risks and poor IAQ. These added to the existing knowledge of energy-related IEQ issues and highlighted the need to address IEQ simultaneously with energy through better design, advanced operational controls and by incorporating regular IEQ measurements as part of operations and maintenance protocols.

The novel approach presented here builds a case to move building performance calculations towards an operational context, where design projections are done using advanced simulation and with a view of tracking the projections through to operation using measurable performance outcomes. Overall, the study shows the importance of the early involvement of all stakeholders and their accountability to minimise performance issues. Integrating the findings from the case studies, a case could be built for having IEQ performance objectives in energy performance contracts. This can mitigate the trade-offs of IEQ against energy performance that leads to unintended health consequences for occupants. Further, this work promotes a way of integrating dynamic thermal simulation in regular post-occupancy checking and management of buildings.

Impact Statement

This research deepens the understanding of the root causes of the performance gap across the building sector and advances the current practices followed during the energy modelling of buildings. The work proposes a new framework for assessing performance issues in buildings. The framework defines the performance gaps based on the calculation protocols used to generate the baseline (simulated) performance - the reference point of ascertaining the magnitude of the gap from actual performance. The work pushes for an evidence-based calibration approach for the identification of building performance issues. During the process a multi-level calibration checking framework is also proposed to improve quality of calibrated models. The advancements proposed to current calibration practices, such as improved validation methods and minimum IEQ verification, enables development of more robust and useful calibrated models. Overarching all the energy related analysis, this work expands the domain of performance assessment to encompass aspects of IEQ so that the trade-offs of IEQ against energy performance that lead to unintended health consequences for occupants can be mitigated.

The new framework and proposed methods in this thesis have informed an industry guidance document, CIBSE TM63: Operational performance: Modelling for evaluation of energy in-use. Additionally, this work has informed the development of new software tools by DesignBuilder software to enable the industry to deliver on operational performance assessments. In a wider context, the research pushes the purview of energy-centred performance goals to a holistic balancing of achieving energy efficiency whilst maintaining acceptable levels of IEQ. This work has contributed to the development of 'CIBSE TM63: Operational performance: Building performance modelling and calibration for evaluation of energy in-use' and 'CIBSE TM61: Operational performance of buildings'.

List of Publications

Journal

Jain, N.; Burman, E.; Stamp, S.; Shrubsole, C.; Bunn, R.; Oberman, T.; Barrett, E.; Aletta, F.; Kang, J.; Raynham, P.; Mumovic, D.; Davies, M. (2021) Building Performance Evaluation of a New Hospital Building in the UK: Balancing Indoor Environmental Quality and Energy Performance. *Atmosphere*, 12(1), 115. https://doi.org/ 10.3390/atmos12010115

Jain, N., Burman, E., Stamp, S., Mumovic, D., & Davies, M. (2020). Cross-sectoral assessment of the performance gap using calibrated building energy performance simulation. *Energy and Buildings*, *224*, 110271. https://doi.org/10.1016/j.enbuild.2020.110271

Jain, N., Burman, E., Robertson, C., Stamp, S., Shrubsole, C., Aletta, F., Barrett, E., Oberman, T., Kang, J., Raynham, P., Mumovic, D., & Davies, M. (2019). Building performance evaluation: Balancing energy and indoor environmental quality in a UK school building. *Building Services Engineering Research and Technology*, *41*(3), 343–360. https://doi.org/10.1177/0143624419897397

Papachristos, G., Jain, N., Burman, E., Zimmermann, N., Mumovic, D., Davies, M., & Edkins, A. (2020). Low carbon building performance in the construction industry: A multi-method approach of project management operations and building energy use applied in a UK public office building. *Energy and Buildings*, *206*, 109609. https://doi.org/10.1016/j.enbuild.2019.109609

Papachristos, G., Jain, N., Burman, E., Zimmermann, N., Wu, X., Liu, P., Mumovic, D., Lin, B., Davies, M., & Edkins, A. (2020). Low carbon building performance in the construction industry: a multi-method approach of system dynamics and building performance modelling. *Construction Management and Economics*, *38*(9), 856–876. https://doi.org/10.1080/01446193.2020.1748212

Conference

Jain, N., Burman, E., Mumovic, D., & Davies, M. (2020). Using model calibration to develop a measurement and verification framework for managing the energy performance gap in buildings. Glasgow, UK: CIBSE Technical Symposium 2020.

Jain, N., Burman, E., Mumovic, D., & Davies, M. (2020). Improving model calibration: A case study application of incorporating IEQ with Energy. Chicago, USA: 2020 BPAC and SimBuild.

Jain, N., Burman, E., Mumovic, D., Davies, M., & Tindale, A. (2019). Calibrating Energy Performance Model of a Hospital Building: Dealing with Practical Issues of Data Availability and Granularity in a Case Study Building in the UK. Rome, Italy: 16th IBPSA Conference Building Simulation 2019.

Jain, N., Burman, E., Robertson, C., Mumovic, D., & Davies, M. (2019). Moving beyond energy, integrating it with Indoor Environment Quality: UK school building case study. Sheffield, UK: CIBSE Technical Symposium 2019.

Jain, N., Burman, E., Mumovic, D., Davies, M., & Tindale, A. (2018). Comparative analysis of protocols used in measurement and verification of energy performance: Dealing with practical issues of data availability and granularity in a UK school building. Cambridge, UK: 4th IBPSA-England conference Building Simulation and Optimization (BSO) 2018.

Jain, N., Burman, E., Mumovic, D., Davies, M., & Tindale, A. (2017). Improving the Energy Performance Contracting Process using Building Performance Simulation: Lessons Learnt from a Post Occupancy Investigation of a Case Study in the UK. San Francisco, USA: 15th Conference of International Building Performance Simulation Association (IBPSA).

Papachristos, G., Jain, N., Burman, E., Zimmerman, N., Mumovic, D., & Davies, M. (2018). A multi-method approach to low carbon building performance: Coupling project management to building performance in a UK public building case study. London, UK: CIBSE Technical Symposium 2018.

Papachristos, G., Jain, N., Burman, E., Zimmerman, N., Mumovic, D., & Davies, M. (2018). Low carbon performance in the construction industry: A multi-method approach of project management operations and building energy use applied in a UK public office building. Reykjavik, Iceland: 36th International Conference of the System Dynamics Society.

Papachristos, G., Jain, N., Burman, E., Zimmerman, N., Mumovic, D., & Davies, M. (2018). Towards the low carbon transition in the construction industry: A multi-method framework of project management operations and total building performance. Reykjavik, Iceland: 36th International Conference of the System Dynamics Society.

Burman, E., Stamp, S., Jain, N., Shrubsole, C., Mumovic, D., & Davies, M. (2018). Total energy and environmental performance of low carbon buildings: a cross sectoral study. London, UK: CIBSE Technical Symposium 2018.

Industry Guidance

Jain, N., Burman, E., Mumovic, D., & Davies, M. (2020). TM63 Operational performance: Building performance modelling and calibration for evaluation of energy in-use. London, UK: The Chartered Institution of Building Services Engineers.

Aletta F, Barrett E, Bunn R, Burman E, Cappuccio V, Chatzidiakou L, Van Creveld K, Davies M, Dowson M, Van Dronkelaar C, et al. (2020). TM61 Operational performance of buildings. London, UK: The Chartered Institution of Building Services Engineers.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF TERMS

ANOVA: Analysis of Variance

ASHRAE: American Society of Heating Refrigerating and Air Conditioning Engineers

BMS: Building Management System

CFL: Compact Fluorescent Lamp

CHP: Combined Heat and Power

CO_2: Carbon dioxide

C_V(RMSE): Coefficient of Variation of the Root Mean Square Error

DEC: Display Energy Certificate

DSM: Dynamic Simulation Method

DX chillers: Direct Expansion chillers

ECM: Energy Conservation Measures

EER: Energy Efficiency Ratio

EPBD: Energy Performance of Buildings Directive

EPC: Energy Performance Certificate

FAST: Fourier Amplitude Sensitivity Testing

FEMP: Federal Energy Management Program

FM: Facilities Management

HVAC: Heating Ventilation and Air Conditioning

IAQ: Indoor Air Quality

ICT: Information Communications Technology

IEQ: Indoor Environment Quality

IPMVP: International Performance Measurement and Verification Protocol

IWEC: International Weather for Energy Calculation

LED: Light Emitting Diode

M&V: Measurement and Verification

MAE: Mean Absolute Error

MVHR: Mechanical Ventilation with Heat Recovery

NCM: National Calculation Methodology

NMBE: Normalised Mean Bias Error

NO_2: Nitrogen dioxide

O&M: Operation and Maintenance

PIR: Passive Infra-Red

PM: Particulate Matter

PMV: Predicted Mean Vote

POE: Post Occupancy Evaluation

PPD: Percentage of People Dissatisfied

PROBE: Post-Occupancy Review of Buildings and their Engineering

RH: Relative Humidity

RMSE: Root Mean Square Error

SBEM: Simplified Building Energy Model

UN SDG: United Nations Sustainable Developments Goals

VOC: Volatile Organic Compound

WHO: World Health Organisation

WMO: World Meteorological Organization

1 INTRODUCTION

This chapter describes the context, aim and key objectives of the research on how calibrated simulation models can be used in a procedural assessment and in closing the performance gap across the UK building sector. The building and construction sector accounted for 36% of global final energy use and 39% of energy and process-related carbon dioxide (CO_2) emissions in 2018 and this sector offers the most cost-effective options to achieve the goals of the Paris agreement commitment and the United Nations Sustainable Developments Goals (UN SDGs) (IEA & UNEP, 2019). In the UK, the latest amendment to the Climate Change Act 2008 requires the UK to ensure net-zero carbon emissions by the year 2050 (HM Government, 2019). To meet the environmental performance related UN SDGs and the UK's net-zero carbon emissions targets, various schemes have been implemented in the UK building sector that focuses on improving energy efficiency and quantification of performance at both the design stage and during operations. While UK Building Regulations (Part L) and asset ratings (Energy Performance Certificate - EPCs) focus on design-stage quantification, the operational rating Display Energy Certificate (DEC) scheme focuses on operational-stage performance.

The 'Post-Occupancy Review of Buildings and their Engineering - PROBE' studies (Bordass, et al., 2001a) and more recent Innovate UK building performance evaluations (Palmer, et al., 2016a) show that buildings underperform post-completion when compared against the performance predicted during the design stage. This underperformance is seen in terms of building fabric, services and technologies, as well as user satisfaction and wellbeing, leading to a *performance gap*. The performance gap is the difference between the actual performance of the building and the intended one, often calculated using dynamic thermal simulation tools (Carbon Trust, 2012; Burman, et al., 2014; de Wilde, 2014; van Dronkelaar, et al., 2016; de Wilde, 2018).

Building performance simulation tools are used to calculate thermal loads and resulting energy use, along with related metrics for occupant comfort and indoor environment quality (IEQ). In the context of simulation, in this work where a model is an object, simulation is a process, tool is a vehicle and performance is a result, models are simulated in tools to assess performance. A building performance simulation model is considered as a good model if it delivers performance predictions with adequate levels of accuracy, factoring in most of the complex physical interactions and interrelations. Due to the assumptions used in defining the inputs, some variability in simulation outputs is to be expected; however, the scale of the discrepancy is often too wide to be acceptable and inadvertently reduces confidence levels in the results of simulation calculations (de Wilde, 2018).

1.1 BACKGROUND

Evaluating performance of buildings after occupancy requires understanding of how well a building meets the needs of clients and building occupants. It has been a focus in building standards, such as Government Soft Landings (Philp, et al., 2019), British Standard 8536-1:2015, WELL Building Standard and Passivhaus (RIBA and Hay, et al., 2016). However, a recent report identifies that the delivery of post occupancy evaluations, in general, is patchy, and the lessons learnt have not been consistently embedded into the knowledge and processes of the construction industry (RIBA, 2019). It highlights that the performance gap in new and existing buildings is a key issue that must be urgently addressed in order for the UK to meet UN SDGs and its own net-zero carbon emissions targets (RIBA, 2019).

The term 'performance gap' is commonly used to define the discrepancy between predicted and measured performance in the context of building energy use (de Wilde, 2018). However, its practical definition is quite vague – depending on the baseline* chosen or the calculation protocols used, the magnitude and cause of the gap can vary (Burman, 2016; van Dronkelaar, et al., 2016). For example, under the Energy Performance of Buildings Directive (EPBD), National Calculation Methodology (NCM) was developed for the UK non-domestic sector to facilitate simplified and dynamic calculation tools (Simplified Building Energy Model (SBEM) and Dynamic Simulation Method (DSM)) for regulatory compliance calculations. These methods calculate energy use under standardised operating conditions, comparing the calculated energy performance of a building to the energy performance of notional or reference buildings, depending on the assessment type. Despite the *relativist* nature of these NCM based regulatory calculations, they are sometimes used to project the *absolute* energy performance of buildings (Burman, et al., 2014). These regulatory compliance calculations are suitable for policy implementation purposes, where calculations need to be simple, replicable, verifiable and, most importantly, suitable for performance estimations and comparisons across the building stock. However, due to a lack of understanding of intentions, limitations and finer details of the calculations used, there is a prevalence of interchangeable and contentious acceptance of the outcomes of these regulatory compliance calculations as an individual building's projected design-stage performance. To address this issue, CIBSE TM54 (CIBSE, 2013a) was developed for estimating operational energy use at the design-stage.

The modelling approach in CIBSE TM54 accounts for all end uses in the building alongside realistic operating conditions and occupant behaviour. The estimates of energy use calculated at the design-stage using the framework provided in CIBSE TM54 can be used to check the performance of the building in operation. In line with this, while CIBSE TM54 advocates post-occupancy evaluation of the

* reference point for calculating the performance gap

building performance, the scope of guidance presented there is tailored to improve design-stage calculations. Once the building is operational, a separate process is needed for performance evaluation by using a new measurement and verification (M&V) framework to check and improve the performance of the existing buildings. In these operational assessments, it is necessary to identify performance deviations that are caused due to changes in building usage due to functional requirements and other issues that are due to other (technical) shortcomings. Calibrated simulation models can be used for M&V and to help identify differences between how a building was designed to perform and how it is functioning. M&V protocols such as American Society of Heating Refrigerating and Air Conditioning Engineers (ASHRAE) Guideline 14 (ASHRAE, 2014), Federal Energy Management Program (FEMP) (Webster, et al., 2015) and International Performance Measurement and Verification Protocol (IPMVP) (EVO, 2016) suggest the use of calibrated computer simulation modelling during M&V to assess energy performance and check potential issues in the building's operations. These protocols provide acceptable calibration tolerances and requirements to create a calibrated model that is fit for applications such as energy conservation measures (ECM) evaluation, optimisation of building system controls and identification of underlying performance gap issues. However, the calibration criteria defined in these protocols are limited and do not provide a detailed framework for developing and cross-validating these calibrated models.

In the context of the performance gap, energy performance is generally the most emphasised. However, the gap between actual and expected performance is not limited to energy; it may also be identified for IEQ parameters, such as temperature, relative humidity, air quality (pollutants, CO_2), noise and lighting (Tuohy & Murphy, 2015; Fabbri & Tronchin, 2015; Phillips & Levin, 2015). The direct relationship of occupant wellbeing, comfort and productivity to IEQ in various building types is well measured and documented (Wyon & Wargocki, 2013; Chatzidiakou, et al., 2014; Al Horr, et al., 2016). Moreover, the ways to achieve better IEQ and building user satisfaction might contradict measures to achieve better energy performance. Therefore, if the focus is only on energy or carbon emissions, this can lead to the unintended consequence of poor IEQ in buildings. Energy use reduction is not enough unless it allows buildings to perform their desired functions, i.e. to be healthy, comfortable and productive places in which to live and work. To avoid the unintended consequence of compromising IEQ, a holistic approach to performance assessment that includes IEQ is needed (Shrubsole, et al., 2018).

This thesis focuses on extending the use of building performance simulation in the domain of operational performance. It addresses the challenge of making energy models of existing buildings more accurate by reliably calibrating them with monitored data, so that they can be used for identification and mitigation of performance gap issues. Using a case-study approach, the work focuses on performance issues prevalent across the UK building sector and identifies improvements in the energy model calibration process. In this process, IEQ in the case study buildings is also monitored and analysed to identify any energy-related unintended underperformance; also, typical zone temperatures are used for cross-validation of the calibrated models.

1.2 AIM AND OBJECTIVES

The research aims to **develop a simulation-based framework for a procedural examination of the root causes of the performance gap in buildings**, focusing on energy use and energy-related unintended IEQ underperformance**.** Using calibrated simulation, in this operational performance assessment, an effective model calibration and validation protocol can be created for effective implementation in M&V to manage the performance gap. The study has the following objectives:

1. Analyse the operational performance of buildings for their energy use and the performance of energy-related IEQ parameters.
2. Develop a systematic and replicable framework for M&V, using calibrated building performance simulation models to identify performance issues.
3. Explore model calibration and improved validation approaches for implementation within M&V, when used in the context of performance issues identification.
4. Examine the findings from different types of buildings and determine the cross-sectoral performance issues and their underlying root causes.
5. Assess the unintended IEQ underperformance that might occur when the focus of the design is primarily to meet increasingly stringent objectives of energy efficiency.

The main research method used to achieve these objectives is a case study analysis of five buildings from different building sectors in the UK (two offices, a school, an apartment block and a hospital). The justification of a case study approach is discussed in detail in Section 3.4.2.

The method used to develop the protocol was based on a desk-based study of existing peer-reviewed journals and other industry specialist and professional body literature. Then, for each of the case studies, site-based data collection was undertaken. Data was collected through design documentation review, metering and monitoring, on-site observations and feedback from stakeholders. This was followed by a simulation-based study, where analysis of the collected data was done to identify issues in case study buildings using dynamic thermal simulation in DesignBuilder Software (DesignBuilder Software Ltd., 2019a). Finally, findings from case study level analysis were reviewed and conclusions were drawn in a desk-based study; also, linking them with findings from existing literature. These are explained in detail in Chapter 3.

1.3 CONTRIBUTION TO KNOWLEDGE

A calibrated computer model, which is an *operationally accurate virtual representation* of the actual building, can be used to investigate, and potentially fix performance issues. Undertaking a cross-sectoral examination of the performance gap and identification of its root causes for various building categories provides insights into endemic issues within the industry.

Extending the existing body of work, this thesis proposes a new framework of analysis and assesses its implementation through case studies. The work addresses the shortcomings of existing procedures and provides lessons both at the case study level and for process improvements at the industry level. The key contributions of knowledge of this research are:

1. **Development of a new calibration-based methodology for M&V, which is procedural (step-by-step) and replicable.**

   The research proposes a new framework for assessing performance issues in buildings. The framework defines the performance gaps based on the calculation protocols used to generate the baseline (simulated) performance which is used as the reference point when ascertaining the gap from actual performance. It also identifies and separates the underlying causes of the deviations between actual and simulated performance. The M&V method proposed in this research is able to identify and separate performance issues because of:

   - Use of inappropriate design-stage calculation methods,
   - Technical issues with the building, its systems and their operations, and
   - Operational changes that the building has gone through to meet its functional requirements.

   This methodology forms the basis of industry guidance document CIBSE TM63: Operational performance: Modelling for evaluation of energy in-use (elaborated in Section 1.4)

2. **Cross-sectoral assessment of performance gap issues across the UK building sector using a new calibration-based methodology, which is procedural (step-by-step) and replicable.**

   The research assesses performance issues across the building sector with a detailed analysis of four types of buildings (office, school, hospital and apartment block), using an evidence-based model calibration with enhanced validation. This provides a robust assessment of performance issues and corroboration of their causes, with many common themes and lessons identified. While drawn from individual case studies, the issues identified have applicability in other buildings and across the construction sector.

3. **Improving the quality of model calibration through a multi-level (tiered) framework for checking calibration accuracy by incorporating the advanced cross-validation methods with existing protocols.**

   This research explores several ways in which the quality of a calibrated model can be improved and suggests a multi-level calibration checking framework. The improvements considered include the use of disaggregated data, integration of zone temperature cross-validation and using uncertainty-based hybrid validation methods. The current industry standard of total building energy use at monthly timescales is determined to be the lowest ‘acceptable’ level of calibration. Higher levels of calibration would require validation of total energy use along with disaggregated end-uses whilst meeting the acceptance criteria and minimum validation requirement for zone temperatures. During the calibration process, the impact of data availability and granularity on calibration level and accuracy of the calibrated results were also examined.

4. **Identification of unintended consequences of IEQ underperformance that are a result of focus on energy efficiency due to sometimes conflicting strategies to achieve both better energy performance and better IEQ performance.**

   The focus on energy efficiency alone as the performance objective in buildings built with low energy use intentions can lead to unintended IEQ concerns, such as poor air quality or overheating. Issues across the case studies show the need to address IEQ simultaneously with energy through better design, advanced operational controls and, most importantly, incorporating regular IEQ measurements. The study indicates the importance of contractual accountability to minimise performance issues, building a case for including IEQ in energy performance contracts to mitigate the trade-offs of IEQ against energy performance that lead to unintended health consequences for occupants.

1.4 Contribution to Practice Statement

This research deepens the understanding of the root causes of the performance gap across the building sector and advances the current practices followed during the energy modelling of buildings. It pushes for an evidence-based calibration approach for the identification of building performance issues. The advancements proposed to current calibration practices, such as improved validation methods and minimum IEQ verification, would enable the development of more robust and useful calibrated models. The new framework and proposed methods have informed an industry guidance document, CIBSE TM63: Operational performance: Modelling for evaluation of energy in-use. Additionally, this work has informed the development of new software tools by DesignBuilder software to enable the industry to deliver on operational performance assessments. In a wider context, the research pushes the purview of energy-centred performance goals to a holistic balancing of achieving energy efficiency whilst maintaining acceptable levels of IEQ. This work has contributed to the development of 'CIBSE TM63: Operational performance: Building performance modelling and calibration for evaluation of energy in-use' and 'CIBSE TM61: Operational performance of buildings'.

1.5 STRUCTURE OF THE THESIS

Figure 1.1 explains the conceptual flow of the research undertaken and the structure of the thesis. The different colours indicate how the flow of research links with the overall structure of this document. The study is conceptually divided into four parts: context, focus areas, processes and outcomes. Each part is covered in the thesis sequentially through various chapters. In this section, each part is explained, with links to the different chapters and a brief outline of their content. Further, to show how each of the objectives of the thesis are met, Figure 1.2 maps the objectives with the various sections and chapters.

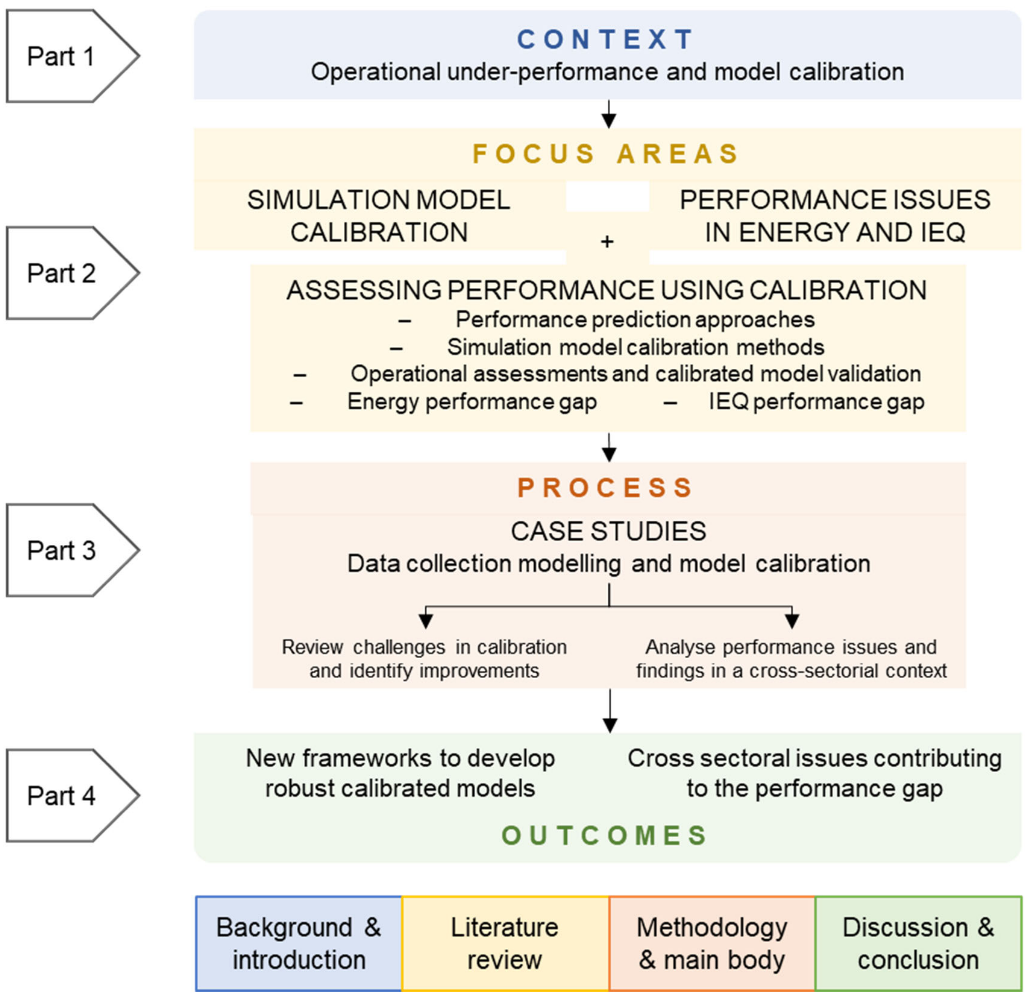


Figure 1.1: Study design of the thesis

1.5.1 Part 1: Context

This section sets out the larger framework within which the current study is undertaken. These include the policy, standards and protocols related to energy performance in new and existing buildings; the importance of ensuring better operational performance of buildings and minimising the performance gap; and the role simulation can play in predicting and managing performance in buildings. Context is set up in Chapter 1 (*Introduction*).

Chapter 1: Introduction

This chapter introduces the issue of the performance gap in energy use in buildings and energy-related unintended underperformance in IEQ. It also looks at how calibrated energy models can be used in identification, validation and correction of performance gap issues, and explains how these are connected to the aim and objectives of this research. This is followed by the design and structure of the whole document.

1.5.2 Part 2: Focus areas

Focus areas of the thesis are explored in detail through a literature review in Chapter 2. The study explores two areas: first, the creation of calibrated models that can accurately represent the operational performance of buildings, and second, the identification of root causes of the performance gap that are prevalent across the UK building sector. Linking the two areas together, the study, assessing performance issues using calibration, focuses on five key aspects.

Chapter 2: Literature review

This chapter reviews building performance model calibration and validation, finding key determinants in calibration method selection, challenges for various calibration approaches and improvements required in validation protocols. The chapter also reviews the performance gap, its magnitude, underlying causes, and approaches for mitigation. In addition, it also explores the underperformance issues in IEQ which are an unintended consequence of focusing only on energy in the context of building performance.

1.5.3 Part 3: Process

The process is covered in the methodology (Chapter 3), which defines the procedural workflow on using calibration for performance gap assessment. In the main body (Chapters 4-7) the application of the methodology on the case study buildings is shown. In these sections, both the focus areas are continually explored for lessons, challenges faced, and necessary improvements related to the calibration process. In this part, specific lessons related to performance issues in the case studies are also discussed, along with their larger applicability within the larger construction sector.

Chapter 3: Methodology

This chapter describes the framework developed for undertaking model calibration and validation for systematically assessing the root causes of the performance gap. It presents a step by step workflow of data collection, model fine-tuning, calibration checks and analysis of root causes of underperformance in energy use, including unintended impacts on IEQ. This chapter forms the basis of the next four chapters (Chapters 4 to 7), where performance analysis and calibration of case study buildings are undertaken and lessons learnt for the case study building are discussed along with their larger cross-sectoral impacts.

Chapter 4: Case study buildings and operational performance

This chapter introduces the case study buildings and provides a detailed description of the design-stage (construction and operational) characteristics of each of the five buildings, which are typical examples of new buildings in their respective sectors in the UK. The chapter also presents the design-stage calculation results of the case study buildings and compares them with the actual energy use to quantify the performance gap. IEQ performance parameters for temperature, relative humidity (RH), CO_2, nitrogen dioxide (NO_2), particulate matter ($PM_{2.5}$ and PM_{10}) are also presented, and their underperformance, if any, is highlighted. All the metering and monitoring data for energy and IEQ parameters reported here were collected during a related project on the broader operational performance of buildings.

Chapter 5: Operational analysis and model calibration (Monthly)

The previous chapter shows the design-stage and actual performance of each building and quantifies the magnitude of the performance gap. This chapter shows the development of calibrated models for each of the case study buildings to identify the case-specific causes of the performance gap. The chapter first describes the development of the calibration-baseline (as-design) model based on design-stage data. Then, after fine-tuning of the as-design model (based on on-site observations and monitoring), this chapter presents monthly energy use calibration results for all the five case study buildings. These calibrated models are cross-validated against monitored zone temperatures. The chapter summarises the deviations between the calibrated model and the design-stage model which are verified through calibration and are the most likely causes of performance gap in the respective case study buildings.

Chapter 6: Detailed model calibration (Hourly) and validation

In the previous chapter all building models were calibrated for monthly energy use to identify the causes of the performance gap identified in Chapter 4. However, for one of the case study buildings, highly granular and disaggregated data was available, which provided an opportunity to take the monthly calibration model further. This chapter describes the detailed model calibration under a multi-level calibration framework for that case study building. The detailed calibration is at hourly temporal granularity and is for disaggregated (spatial and end-use) sub-metered data using additional monitoring information (such as using IEQ data and occupancy data to create typical profiles). To check the robustness of the developed model, like with the monthly calibration, cross-validation is done through zone temperature checks. Additional uncertainty-based probabilistic checking of model outputs is also done. The chapter reflects upon the added data requirements and the applicability of models calibrated to hourly (compared to monthly) resolution. The chapter exploring advanced validation also assesses the metrics and thresholds used during calibration checks.

Chapter 7: ***Cross-sectoral performance gap issues***

Identification of the causes of the performance gap in Chapter 5 was undertaken at a building-level. This chapter explores the issues identified in a more generic context. First, the issues identified for each case study building are categorised into either technical issues or operational changes due to occupant behaviour issues or needed for the building to perform its function in practice. Then, drawn from the analysis of common themes and lessons identified at the building-level among the individual case studies, this chapter examines the likely applicability of the findings in other similar buildings and across the construction sector. This examination focused on energy, also looks at the causes of unintended IEQ underperformance regarding thermal comfort and indoor air quality (IAQ).

1.5.4 Part 4: Outcomes

The outcomes of the study are summarised in the *Synthesis of knowledge* (Chapter 8) and *Conclusion* (Chapter 9). This final part of the research compiles the lessons learned about the two focus areas described in *Part 2* (Section 1.5.2), the new calibration assisted framework and the cross-sectoral performance gap issues.

Chapter 8: **Synthesis of knowledge**

This chapter reflects on the main outcomes of the study for the two focus areas, model calibration process improvement and cross-sectoral performance issues related to not just energy but IEQ as well. The model calibration issue is covered in Chapters 5 and 6, whereas the performance issue is explored in Chapters 4, 5 and 7. The chapter is presented as a reflection on each of the study objectives described in Section 1.2.

Chapter 9: **Conclusion**

In this chapter, the main findings of the research on improving the energy model calibration process in the context of the performance gap assessment using energy and environmental data are presented. Additional research gaps and improvement of certain techniques are discussed as areas for future work.

<table>
<tr><th colspan="4">AIM
Develop a simulation-based framework for a robust and procedural examination of the root causes of the performance gap, focusing on energy use and energy related unintended IEQ underperformance</th></tr>
<tr><th>OBJECTIVE</th><th>LITERATURE</th><th>METHODOLOGY</th><th>RESULT CHAPTERS</th></tr>
<tr><td>1. Analyse the operational performance of buildings for their energy use and performance of energy related IEQ parameters.</td><td>Part 1: Performance prediction approaches</td><td>Step 1: Data Collection

Step 2: Comparison (Energy and IEQ)</td><td>Chapter 4: Case study buildings and operational performance</td></tr>
<tr><td rowspan="2">2. Develop a systematic, robust, and replicable framework for M&V, using calibrated building performance simulation models to identify performance issues.

3. Explore model calibration and improved validation approaches for implementation within M&V, when used in the context of performance issues identification.</td><td rowspan="2">Part 2: Simulation model calibration

Part 3: Operational assessments and calibrated model validation</td><td rowspan="2">Step 3: Modelling
- Baseline
- Monthly calibration
- Hourly calibration
- Advanced validation</td><td>Chapter 5: Operational analysis and model calibration (Monthly)</td></tr>
<tr><td>Chapter 6: Detailed model calibration (hourly) and validation</td></tr>
<tr><td>4. Examine the findings from different types of buildings and determine the cross sectoral performance issues and their underlying root causes.</td><td>Part 4: Energy performance gap and its causes</td><td rowspan="2">Step 4: Analysis and Lessons</td><td rowspan="2">Chapter 7: Cross sectoral performance gap issues</td></tr>
<tr><td>5. Assess the unintended IEQ under-performance, which might occur when the focus of design is primarily to meet increasingly stringent energy efficiency objectives.</td><td>Part 5: IEQ performance gap and its causes</td></tr>
</table>

Figure 1.2: Relationship of the study objectives with various sections and subsections of the thesis

2 LITERATURE REVIEW

Operational stage performance assessments in existing buildings are done for applications such as evaluation of ECMs in building retrofits, identification of performance issues, retro-commissioning, building diagnostics and optimising the operation of building services. M&V processes are structured operational stage performance assessments. In M&V, data planning, measuring, collection and analysis are undertaken for verifying and reporting performance. There are many M&V methods, such as Post Occupancy Evaluation (POE), which is a broad review of building performance and user satisfaction, and IPMVP/ASHRAE Guideline 14, which are more focused on operational verification in specific applications, such as evaluation of ECMs. Findings from existing studies on operational stage building performance assessment show that there are significant shortcomings in energy and IEQ performance and that there is little concurrence between the design assumptions and actual on-site observations.

One of the ways to identify and validate performance issues in buildings is to use building performance simulation tools, developing a calibrated simulation model of the actual building. Calibration in the context of building performance simulation is the process of fine-tuning the input parameters of a simulation model to create a digital equivalent of a real building. A review of the literature has been undertaken in this section to understand how calibrated simulation models can provide valuable insights into operational stage performance issues with a high level of confidence. The review extends the understanding of the prevalent issues in the construction sector. Covering the two broad themes of model calibration and performance issues, this literature review chapter is divided into five sections. Figure 2.1 shows the links between the literature review and the overall study objectives.

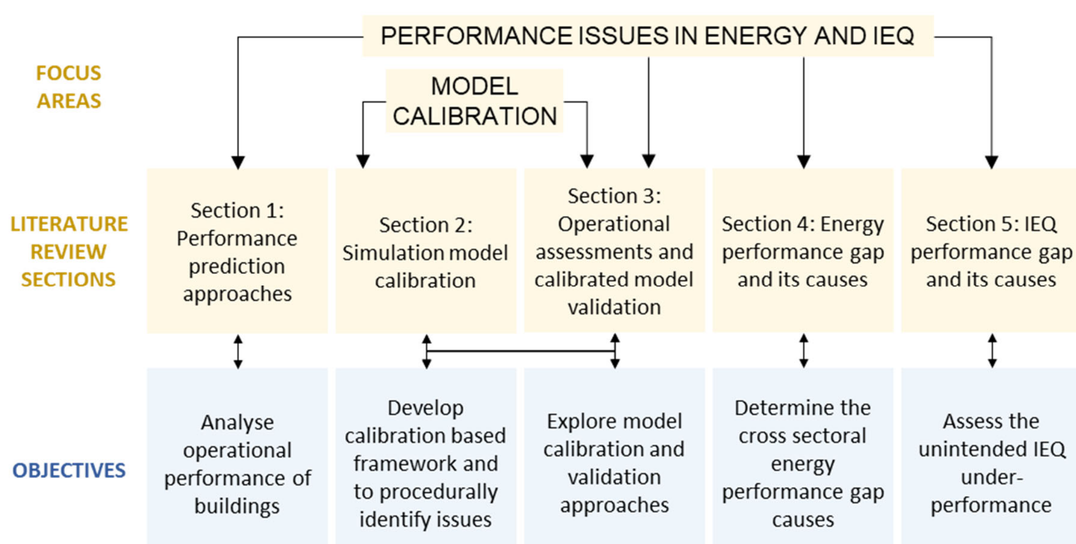


Figure 2.1: Links of literature review with study objectives

In the first section, the chapter reviews the approaches used to calculate building energy performance and differences between their predictions. In the second section, the process and methods of model calibration are reviewed. The key determinants in method selection and challenges for various calibration approaches are identified. In the context of the use of calibration in post-occupancy assessments and M&V, commonly used protocols and advanced model validation techniques are discussed in Section 3. In the fourth section, findings from existing studies regarding the energy performance gap across the construction industry are discussed, highlighting, and analysing the magnitude, root causes and mitigation measures. Further, as the energy performance gap alone does not capture the full impact of buildings on occupants and the wider environment, the performance gap issues that impact occupant wellbeing and IEQ are also reviewed in the last section.

2.1 PERFORMANCE PREDICTION APPROACHES

As policies increasingly focus on creating low-energy buildings and benchmarking their performance, the use of computer modelling has become essential to assessing and improving performance throughout a building's life (design, construction, operation and retrofit). There are two key determinants of the scope and complexity of the calculations used for performance assessments. One is the project procurement stage at which the calculation is done, and the other is the purpose of the calculations. Based on these factors, calculations can be categorised into:

1. Compliance modelling, done during design and construction stages for regulatory compliance and comparative benchmarking,
2. Performance modelling, done during design and construction stages for building-specific estimates and performance prediction, and
3. Actual performance modelling, done during the operational stage of the building for evaluating current performance and the potential for new retrofits.

In this section, these modelling approaches are explained in more detail, and differences between them and their results are highlighted.

2.1.1 Compliance modelling

In most countries modelling is primarily undertaken for regulatory compliance; this type of modelling can be termed *compliance modelling* (van Dronkelaar, et al., 2016). Under the EU EPBD, in the UK's regulatory compliance systems, SBEM or DSM compliance models are created to assess energy performance and generate EPCs for comparative assessments (DCLG, 2017). To calculate energy use, the compliance modelling approach followed in the UK uses standardised operating conditions, as per the UK NCM (DCLG, 2017). These models often do not accurately reflect the building's actual operating conditions, such as occupancy, temperature set points and schedules of operation of heating, ventilation and air conditioning (HVAC) systems. Furthermore, compliance modelling calculations in the UK do not report energy use related to equipment (plug loads) (DCLG, 2017).

Compliance modelling are yes or no calculations, suitable for policy application where there is a need for simplicity, replicability, verifiability and, most importantly, applicability for the entire building stock, and where 'relative' performance is important (e.g. comparing the energy performance of a building against a reference building using the same method). However, these compliance calculation methods are sometimes inappropriately used to evaluate the energy use performance of buildings as discussed in CIBSE TM54 (CIBSE, 2013a) and analysed in Burman (2016). This happens due to a lack of understanding about the intentions, limitations and finer details of the calculations used, often resulting in misinterpreted predictions of actual performance.

Comparing the actual energy use of a building with compliance modelling results, as per current UK Building Regulations which do not necessarily use real operation settings and do not report non-regulated energy (such as small power equipment) in building energy use totals, can inflate the performance gap (difference between actual energy use and design intended projections) and lead to a *perceived gap* (Burman, et al., 2014), (van Dronkelaar, et al., 2016).

2.1.2 Performance modelling

To address the issue of misinterpretation of compliance calculation results as predicted performance, a new industry guidance document was created in the UK, CIBSE TM54 (CIBSE, 2013a). CIBSE TM54 sets out a framework, using dynamic simulation models, to provide estimates of the likely operational energy performance of buildings at the design stage. It allows designers to tailor the operating conditions as per the project brief and the predicted performance accounts for all end uses including equipment loads. CIBSE TM54 is intended to be used during design and construction stages, and the tailoring of the calculations for actual operating conditions creates a more realistic design stage projection of building performance. This calculation result can also be used as a more appropriate baseline[*] for estimating the magnitude of the energy performance gap. CIBSE TM54 methodology is not meant for benchmarking; therefore, in contrast with *compliance modelling* this approach is termed *performance modelling* (van Dronkelaar, et al., 2016).

One aspect of CIBSE TM54's calculation is that its underlying model to predict HVAC loads is suggested to be typically based on a simplified approach, owing to the cost and time associated with detailed HVAC system design and simulation. In the context of improvements in performance accuracy, learning from the NABERS scheme in Australia (OEH, 2019), the pilot study for the Design for Performance program (Cohen, et al., 2018) highlighted the lack of detail of the HVAC system modelling as a reason for the unreliability of the predictions, even with the use of TM54 modelling protocols. By contrast, undertaking uncalibrated simulations for four case study buildings, Ahmad and Culp (2006) suggest that, due to the added uncertainties, there are high discrepancies between the simulations and the measured data even with detailed system modelling. Ahmad and Culp concluded that noticeable improvements were not obtained with the added effort over the simpler modelling effort.

2.1.3 Actual performance and its modelling

Actual performance is the measured and monitored performance of a building when it is in a steady mode of operation. This is the performance that modellers aim to calculate during the design stages, to meet energy performance objectives and minimise the difference between their predictions and

[*] reference point for calculating the performance gap

actual energy use (performance gap). Figure 2.2 describes the various baseline* calculations and their relation to the different performance gaps. The difference between the compliance model and the actual energy use is the 'inflated' perceived gap. However, the gap between performance modelling calculation result and actual energy use is the *actual gap*.

Compliance calculation from compliance modelling: *Results of regulatory compliance calculations that are often misinterpreted as a prediction of the actual performance.*

Design performance from performance modelling: *Results of calculations done at the design stage that accounts for all end-uses and uses building specific inputs as per CIBSE TM54 or equivalent protocols.*

Perceived Gap: *The gap between the actual energy use and the often underestimated performance when compliance calculation results are misinterpreted as the predicted performance.*

Actual Gap: *Performance gap that is due to changes over time and the technical issues identified in the building and its systems.*

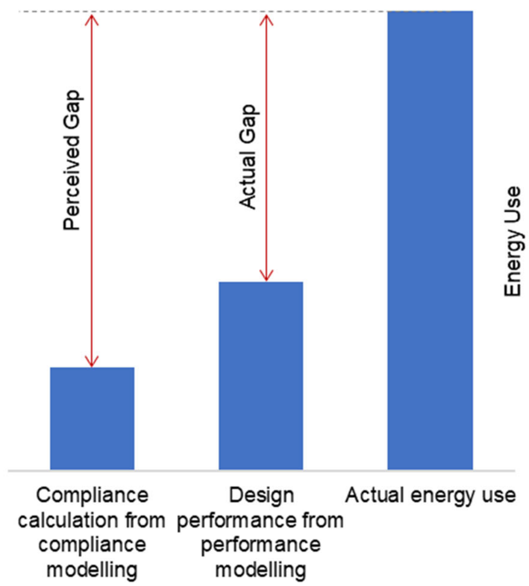


Figure 2.2: Performance calculations and associated gap categorisations

CIBSE TM54 calculations are designed to deal with design stage issues. Problems during construction and operation stages, such as technical issues arising from poor workmanship or maintenance, or changes in building functions or occupancy trends over time, can lead to underperformance but are not reflected in the model created using CIBSE TM54. To deliver a model that can accurately predict the actual performance once a building is in the operational stage, a calibrated simulation model can be developed. Calibration is a process of fine-tuning the input parameters of an initial calibration-baseline model (generally a performance model, such as that developed as per CIBSE TM54) so that its results match the actual performance. Model calibration is reviewed in Section 2.2.

2.1.4 Benchmarking of energy performance

Benchmarking of building energy performance is a process comparing the actual energy consumption of buildings against other similar buildings and/or under standardised protocols and targets. As benchmarking is either a peer-to-peer comparison or aimed at achieving specific targets, it can be an effective strategy to quantitatively track and manage energy use in buildings, for example, ensuring

* reference point for calculating the performance gap

that your building is performance better than other similar buildings. Additionally when employed to large number of buildings, benchmarking can also enable policy led management of energy performance at building stock levels.

In the UK, and asset ratings (EPCs) that focus on design-stage quantification and the operational rating DEC scheme that focuses on operational-stage performance both rely on benchmarking to drive stick level energy performance improvements. In an EPC, during the design stage estimation of energy performance for compliance purposes (See Section 2.1.1) the performance of the proposed building is also rated in comparison to a reference building. This reference building is an equivalent building (i.e. a building of the same size, shape and use as the actual building) constructed to a notional building designed to a specified standard (DCLG, 2017). Similarly for operational stage DEC ratings (DCLG, 2015), actual energy performance is compared against a benchmark, which is the energy consumption of the median building of the same type and are published in CIBSE TM46 (CIBSE, 2008).

CIBSE Guide F (CIBSE, 2012) describes all the known UK Energy and component benchmarks, building up on the statutory building energy benchmarks set out in CIBSE TM46. Table 2.1 provides a summary of the good practice benchmarks as per CIBSE Guide F for the four case-study building categories.

Table 2.1: CIBSE Guide F good practice benchmarks for various case-study building categories

Sector	Benchmark (kWh/m^2/yr.)		CIBSE Guide F category
	Gas (fossil)	Electricity	
Office	79.0	54.0	Offices: naturally ventilated, open plan
School	108.0	25.0	Education (schools): secondary
Hospital	422.3	74.4	Hospitals: acute and maternity
Apartment	247.0	44.0	Residential and nursing homes

The benchmarks in CIBSE Guide F are the most detailed benchmarks available for UK Building stock. While some of the data contained in CIBSE Guide F, published in 2012, is updated by CIBSE TM46, some of it is still based on the data is from the 1998 edition of the Guide with some of the primary source publications are no longer available. Studies exploring the robustness of the benchmarks found that there were noticeable changing trends towards higher electricity consumption and lower fossil thermal energy consumption in many benchmark categories, suggesting that the benchmarks were no longer accurately depicting the current pattern of energy use in the stock (Hong, et al., 2013). Work is underway to gradually update and replace the energy benchmarks in CIBSE Guide F (CIBSE, 2020).

2.2 Simulation model calibration

Building simulations tools use detailed computational methods to predict a building's performance, for both new constructions and retrofits to assist designers in decision making around "what if"-type questions, e.g. "What would happen if something is changed in the design or this retrofit strategy is used?" (de Wit, 2004).

Calibration is a process in which simulation models are corrected (by modifying input assumptions) so that the results more closely match the actual performance. A model is said to be calibrated when the difference between simulated results and actual measurements is less than predefined thresholds, known as the validation criteria. Use of calibrated computer simulation models during M&V to model energy performance and check potential issues in the building's operations has been described in protocols, such as ASHRAE Guideline 14, FEMP and IPMVP. In this section, we look at existing studies related to calibration, its process, methods and data requirements.

2.2.1 Calibration process

Model calibration is typically an iterative process of adjusting model inputs and comparing the results to measured data. However, model calibration is an over-specified problem (i.e. too many inputs to calibrate and too little of metered or measured information), and there are practical and financial limitations to the amount of monitored data available (Coakley, et al., 2014). A typical calibration process requires two data sets: one is the simulation data, which is often based on the as-design values and operational assumptions, and the other is the metered data from monitoring the real building. The former is composed of many inputs. Therefore, depending on the objective of the calibration, the most influential parameters affecting the specific outputs concerned are selected for fine-tuning to create a match between the two data sets. Figure 2.3 shows the basic outline of the overall calibration process, which is divided into three parts, data collection, modelling and fine-tuning. The model is considered calibrated if it meets some pre-defined validation criteria for accuracy.

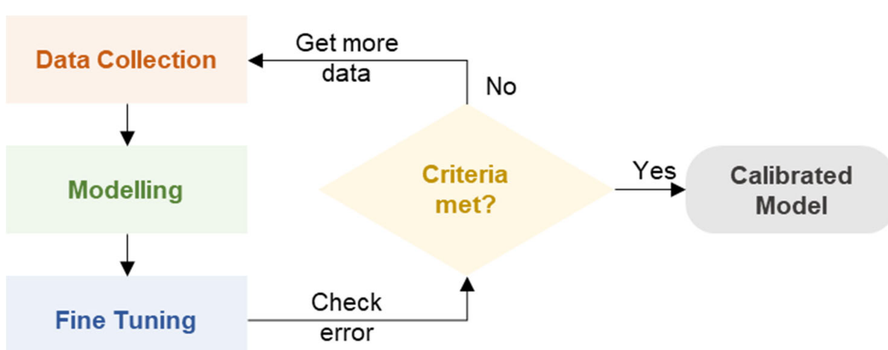


Figure 2.3: Simplified graphic of the calibration process

Data collection

Data collection is the first step in model calibration. The objective of data collection is to develop a detailed understanding of the physical and operational characteristics of the building. Information needed for modelling and calibration can be captured in many ways, such as analysing design-stage documents, operation and maintenance manuals, building audits, spot-measurements, longitudinal metering and monitoring, and stakeholder feedback. Procedures for a detailed audit of buildings have been defined by ASHRAE (2011). IPMVP recommends following the data collection protocols set out in ASHRAE Guideline 14 for creating calibrated energy performance models.

Data gathered about the building should capture design-stage assumptions, intents and targets as well as operation-stage performance and modifications. Data collected include architectural design information (dimensions and properties of building surfaces), monthly and hourly whole-building utility data, design, commissioning and operation information of HVAC and other building system components, building operating schedules and weather data. Required data typically include the following:

- Utility meter readings for at least one year at monthly, hourly, or sub-hourly intervals
- Architectural, mechanical and electrical drawings
- Building performance targets, regulatory compliance benchmarks and intended operating conditions
- HVAC system equipment and their control details (design and actual)
- Lighting system, equipment and process loads details and controls (design and actual)
- Occupant density, occupation and operation schedules
- Detailed, disaggregated monitoring of energy end uses and indoor environment performance characteristics
- Actual equipment operating power for installed lighting, plug load and HVAC equipment
- Weather data for the monitoring period
- Interviews with facilities management (FM) teams

For the operational data collection, the intrusiveness of measurement procedures and the duration of data can determine its usefulness. Short-term monitoring (i.e. >2 weeks per season) can be used to identify typical energy end-use profiles and/or base-loads (Lunneberg, 1999; Wassmer, 2013). In intrusive testing, run over short periods of 3-5 days, the building is conditioned in a controlled manner (Subbarao, 1988) and/or on-off tests (Soebarto, 1997) are performed. This helps in accurately determining building parameters like heat-loss coefficient or lighting and plug loads. One of the most useful types of data that can be collected is high-quality (i.e. disaggregated) and high-resolution (i.e.

at hourly or sub-hourly timescales) metering and monitoring information for an entire year (Clarke, et al., 1993; Raftery, et al., 2011). As opposed to utilising daily load profiles or monthly utility bill data, high-resolution data helps in more effective trend identification and fine-tuning because it eliminates the averaging ("compensation error effect") that happens in lower resolutions.

In the context of in-use building performance, structured evolution approaches, such as POE, also collect useful information that can be used to calibrate the models. POE is a systematic and rigorous process of building evaluation by measuring factors such as building use, energy consumption, operations and management, and user satisfaction. POE helps in understanding building operations, highlights any building issues and provides lessons for improvement. CIBSE's PROBE studies describe the practical evaluation of several buildings (Bordass, et al., 2001a). Additionally, a RIBA report, Building Knowledge: Pathways to Post Occupancy Evaluation, provides practical guidance for undertaking POE (RIBA and Hay, et al., 2016).

Modelling

Once the design and operational data are collected, the next step is to develop a simulation model. To create a calibrated model, first, an initial calibration-baseline model is needed. A calibration-baseline model is the uncalibrated simulation model that uses preliminary (often design-stage) data regarding building and systems' design and operation assumptions. This calibration-baseline model can be the performance model as per CIBSE TM54 (as described in Section 2.1.2), where the simulation results closely represent the design projected or intended performance of the building. This model represents building characteristics, system and equipment specifications, and operation and control mechanisms as they were intended. Key data required to create a calibration-baseline model are physical properties of the facility, equipment and system types and efficiencies, appropriate weather data and control sequences.

Fine-tuning

After creating the initial calibration-baseline model, the fine-tuning of the model is done. Fine-tuning is the iterative adjusting of model inputs in the initial calibration-baseline model by modifying them according to the actual operating conditions and incorporating all the findings identified as deviations from the design intent during the data collection, metering and monitoring. The calibration should be done for a period of one year with the minimum granularity of a month for building-level analysis. But higher granularity (hourly or daily) is preferred, especially in cases where the calibrated model is used for a system and a sub-system level investigation (EVO, 2016). Accuracy of energy model calibration is checked by comparing simulation results with actual monitored whole-building and/or end-use energy data. Model input parameters require fine-tuning until the statistical indices set as validation criteria have been met. Methods used to undertake this fine-tuning are described in Section 2.2.2 below.

2.2.2 Calibration methods review

There are four main categories of calibration methodology identified by Clarke, et al. (1993) and later adopted by Reddy and Maor (2006):

1. Manual calibration methods based on an iterative approach,
2. Graphical-based calibration methods,
3. Calibration based on special tests and analysis procedures, and
4. Automated techniques for calibration, based on analytical and mathematical approaches.

From the four main categories above, different methods and combinations can be used during the same calibration process. In a review of various calibration techniques by Coakley et al. (2014), these were further simplified as manual approaches and automated approaches. This classification was much more in line with the current trends and direction that is seen in recent calibration studies. Both manual and automated techniques may employ procedures and analytical tools to assist calibration. However, automated processes use automated (i.e., not user-driven), mathematical and statistical techniques to reach their goal. A review of various calibration techniques used in existing research studies, classified as per the categories above, is presented in detail by Coakley et al. (2014) and Fabrizio and Monetti (2015).

New advancements in calibration techniques focus on three key areas of improvements: replicability, simplification and automation. How the calibration techniques used in recent studies link to one or more of these areas is listed in Table 2.2 and further elaborated in subsections below.

Table 2.2: Focus areas of calibration techniques in recent studies

Focus Area	Calibration Technique	Studies
Replicability	Use an evidence-based procedural calibration process to improve the reliability and reproducibility of calibration.	(Yoon, et al., 2003; Monfet, et al., 2009; Raftery, et al., 2011; Coakley, et al., 2012; Bertagnolio, et al., 2012; Mustafaraj, et al., 2014; Ji & Xu, 2015)
Simplification	Reduce the dimensionality of the problem by performing a sensitivity analysis of uncertain variables to identify parameters of greatest influence.	(Bertagnolio, et al., 2012; Heo, et al., 2012; Coakley, et al., 2012; Manfren, et al., 2013; Mustafaraj, et al., 2014)
Automation	Decrease the iterative, ad hoc, and experience-based manual user interventions, and time-taken for the calibration process by using automated techniques. Automated techniques include optimisation, Bayesian approach, artificial neural networks, and meta-modelling.	*Optimisation*: (Taheri, et al., 2013; O'Neill & Eisenhower, 2013) *Bayesian approach*: (Heo, et al., 2012; Chong & Lam, 2015; Manfren, et al., 2013), *Artificial neural networks*: (Neto & Fiorelli, 2008; Edwards, et al., 2012) *Meta-modelling*: (Eisenhower, et al., 2012; Manfren, et al., 2013)

Evidence-based procedural calibration (replicability)

Evidence-based calibration requires a structured approach to the calibration process, with the main objective of ensuring reliability and reproducibility. Therefore, it uses a procedural method for model development. In this process, changes are made according to source evidence and clearly defined priorities rather than ad-hoc manual pragmatic interventions that rely on user knowledge, experience, statistical expertise, engineering judgement and an abundance of trial and error. Figure 2.4 shows a typical evidence-based calibration process. The main components of a typical evidence-based method will be (Raftery, et al., 2011; Coakley, et al., 2012; Bertagnolio, et al., 2012; Mustafaraj, et al., 2014):

1. Building a simulation model as per design data,
2. Using a defined hierarchy of the most influential parameters,
3. Iteratively modifying each parameter and check for calibration criteria at each step.

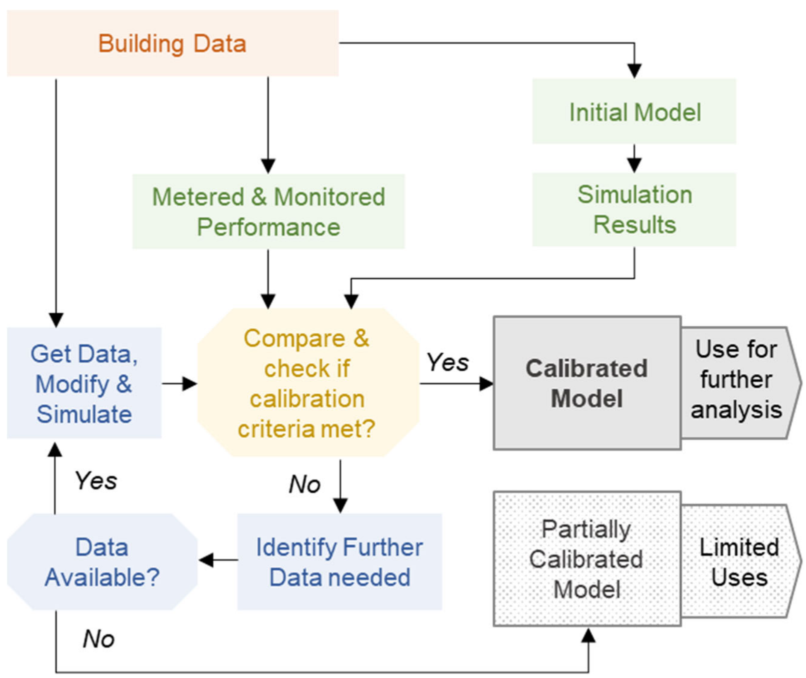


Figure 2.4: A typical evidence-based calibration process

While model evidence-based calibration can provide reasonable results, this technique can be integrated with other techniques mentioned in Table 2.2 to improve/fasten/automate the calibration.

Simplification of problem

Calibration is an over-specified and underdetermined mathematical problem. Many approaches to simplifications have been used to achieve a 'locally calibrated model' (model calibrated for certain objectives and a certain purpose) instead of a 'globally calibrated model'.

Normative models: Normative energy models are lightweight, quasi-steady state formulations of heat balance equations designed to calculate the energy consumption by main end-uses in a building. They approximate energy flows in a building at the macro level with a simplified description of the building. They are appropriate for modelling large sets of buildings in a fraction of the time compared to dynamic simulation (Heo, et al., 2012). However, these models are based on macro-level parameters and may not adequately capture interactive, cumulative effects of real building components. This makes the reproducibility of these normative models difficult.

Day-typing and Zone-typing: This reduces the number of parameters as per the statistical characterisation of inputs. Using zone-typing (i.e., grouping similar zones together) or day-typing (i.e., defining typical days like weekdays, weekends and holidays) to characterise building performance enables condensing a large quantity of complex monitored building data into relatively few input points or schedules (Raftery, et al., 2011). This simplification needs to be analysed on a case-by-case basis as each building is unique. Moreover, this approach makes the use of high-resolution zone-specific inputs and outputs difficult.

Sensitivity Analysis: Sensitivity analysis orders, by importance, the strength and relevance of the inputs in determining the value of the output. As sensitivity analysis of simulation outputs can help to identify the most influential simulation input variables, it can be used to determine the sequence of input data for iterative changes either at evidence-based or experience-based stages of fine-tuning (Bertagnolio, et al., 2012). Sensitivity-analysis-based model calibration does not aim to find a unique and best calibrated solution but rather a small set of the most plausible solutions (Reddy, et al., 2006). Sensitivity analysis, closely linked with uncertainties in modelling and measurements, is further discussed in detail in Section 2.2.5.

Automated calibration techniques

These techniques use mathematical and statistical formulations to find the values of inputs that meet the calibration criteria. These can be used in conjunction with previously defined methods.

Objective-based optimisation: This technique employs some form of optimisation objective function to minimise the difference (Normalised Mean Bias Error - NMBE and Coefficient of Variation of the Root Mean Square Error - C_V(RMSE)) between measured and simulated data (Sun & Reddy, 2006). Several input parameters of the model are systematically varied within specified ranges, to minimise the objective function, using algorithms like GenOpt (Taheri, et al., 2013; O'Neill & Eisenhower, 2013).

Bayesian Approach: Bayesian analysis is a statistical method that employs probability theory to compute a posterior distribution for unknown parameters given the observed data. The procedure begins with the quantification of uncertainty in model parameters using prior probability distributions. This is followed by an application of Bayes's Theorem to produce an updated posterior probability distribution for each parameter through a process of random walks. This updating process is driven by the log-likelihood of the model parameters, given the observations. Although Bayesian calibration has been shown to be effective, it is currently not suitable for use with building performance simulation programs due to their large number of input parameters (Heo, et al., 2012; Chong & Lam, 2015; Manfren, et al., 2013).

Artificial Neural Networks and Meta-models: A meta-model is a simpler and computationally faster version of the simulation model (model of a model) created by reducing the model complexity (Eisenhower, et al., 2012; Manfren, et al., 2013). These models reduce the main detailed energy models to key input and output variables, whilst retaining the complex relationships between those inputs and outputs (Neto & Fiorelli, 2008; Edwards, et al., 2012). These virtual models learn to emulate the performance prediction from their parent models and can then be used with optimisation algorithms to speed up the entire calibration process.

2.2.3 Key determinants in method selection

Selecting the appropriate technique for developing the calibrated model has two key determinants.

1. *Intended use of the calibrated model*: The intended use of a calibrated model can range from identification of ECMs to real-time building management system (BMS) control. It is therefore important to undertake a thorough investigation and analysis of literature that can establish the appropriateness of calibration methods for the intended use, especially to close the performance gap and optimise building systems operations.
2. *Issue of data granularity:* Calibration of models can be done at monthly, daily, hourly, or sub-hourly time steps. As the resolution of data increases, it becomes increasingly practical for calibrated models to be used for detailed analysis and real-time decision making. High-resolution data helps in more effective trend identification and fine-tuning because it eliminates the averaging ("compensation error effect") that happens in lower resolutions. However, as increasing the granularity comes at a cost of accuracy and time, a clear identification of the minimum data (in terms of data streams and granularity) that can achieve the desired level of accuracy in calibration is essential.

2.2.4 Uncertainties in modelling

Building simulations are widely used to estimate the performance of the building design for both new constructions and post-occupancy assessments for retrofits (de Wit, 2004). While performing energy simulations, detailed information about the building envelope, building systems, occupancy and usage, weather conditions, etc. are used as model inputs. Data used as inputs in these models is often estimated or is a deterministic, best-guess value. This data can contain both aleatory uncertainty (probabilistic uncertainty within the actual measured values) and epistemic uncertainty (uncertainty due to a lack of extensive appropriate data).

Many studies have been directed towards an investigation into the input uncertainties that influence the results of building simulations (MacDonald, 2002; Hopfe & Hensen, 2011; Tian & de Wilde, 2011; Heo, et al., 2012; Wang, et al., 2012). Factors such as building operations (occupancy, systems, loads, etc.) estimation, weather conditions, envelope properties and internal modelling algorithm inadequacies affect modelling predictions, whereas measurement errors and metered data inaccuracies affect the quality of operational data collected.

Measurement data errors can be found in the equipment manufacturer's literature or by conducting equipment calibration. Additionally, some data quality issues can arise because of incorrectly located sensors. This can be mitigated by having multiple sensors in a space and taking an average.

To summarise, the main areas that cause uncertainties in simulation models are the following:

1. The accuracy simulation tool and approximations in their computational formulas.
2. Uncertainty of input parameters used in the models to describe the design conditions of the building's envelope and its systems.
3. Weather data used for the simulations.
4. Actual building occupant behaviour and usage.
5. Observation errors in the measured data that affect data quality.

2.2.5 Sensitivity analysis

An understanding of the level of uncertainties is required to make the modelling predictions more relevant and thus ensure a better comparison of building energy simulation results. The implication of inputs uncertainties on the calculated results can be assessed using sensitivity analysis.

There are two approaches for sensitivity analysis, local and global (Campolongo, et al., 2011). Local methods change one input parameter at a time and overlook interactions between different model inputs, whereas a global approach varies multiple parameters at the same time. Methods to undertake sensitivity analysis can be divided into four types (Fabrizio & Monetti, 2015):

1. *Screening-Based Methods:* This is a local analysis usually aimed at identifying the most important or influential parameters to be considered in further global sensitivity analysis. It includes techniques such as sensitivity index calculation (Saltelli, et al., 2007), differential

sensitivity analysis (Heiselberg, et al., 2009) and the Morris method (Morris, 1991; Campolongo, et al., 2011).

2. *Regression Analysis:* This method estimates the relationships between different variables in a model by investigating how an output (dependent) variable changes based on the variation of model input (independent) variable. Through a standardised regression coefficient, this global sensitivity analysis ranks the most important influential input parameters for the output (Saltelli, et al., 2004).
3. *Variance-Based Methods:* These methods can deal with the interdependencies of input variables by measuring sensitivity in two parts, first order and total order. Analysis of variance (ANOVA), Fourier amplitude sensitivity testing (FAST) and Sobol methods are commonly used variance-based sensitivity methods (Saltelli, et al., 2004).
4. *Monte Carlo Method:* This is one of the most common global approaches, where each input is varied within a defined probability distribution and range to quantify overall uncertainty in the model predictions based on the uncertainties in the input parameters. Various sampling methods, such as Random, Stratified and Latin-Hypercube, can be used to choose the values of input variables (Sun & Reddy, 2006).

2.2.6 Challenges in model calibration

Detailed investigation and data collection that is necessary to assess the performance gap can be hindered by practical issues such as cost, stakeholder reputational concerns and liability (Robertson & Mumovic, 2014). However, besides this, there are multiple technical challenges that model calibration faces (Tupper, et al., 2011; RMI, 2011):

1. Building energy modelling being an over-specified problem. It is challenging to identify the few key variables that affect quality calibration.
2. It is hard to get useful data on key inputs, such as occupancy schedules.
3. The expense and time needed to obtain the required hourly sub-metered data is high. Benchmark data does not include sufficiently granular information on end-use breakdowns.
4. There are no standardized procedures for calibration. Depending on the calibration accuracy and objectives there is a possibility masking of modelling inaccuracies at aggregated levels.
5. The lack of integrated tools and automated methods that could assist calibration.

2.3 Operational assessments and calibrated model validation

Statistical indices are generally used to check whether a model should be considered calibrated. These checks determine how well simulated energy consumption matches with the measured data. In the context of using calibrated energy simulation for evaluating building performance, model validation approaches and statistical requirements for calibration are provided in the M&V protocols. These M&V protocols generally focus on quantitative requirements for calibration-baseline model creation and goodness-of-fit of the simulation results. In this section, we review the various M&V protocols used and the calibrated model validation criteria proposed in them. We also look at other validation methods that can be used to check the accuracy of the calibrated model.

2.3.1 Performance evaluations and M&V

Assessment of building operations and performance after occupancy is necessary to understand if there are any issues with a building and identify possible solutions. M&V is the process in which planning, measuring, collecting, and analysing data is undertaken for verifying and reporting a building's performance.

POE

POE is a process of revisiting and learning from in-use building performance, understanding how well a building meets the needs of clients and building occupants. While the scope of POE can be much wider, covering all building performance aspects, in the context of energy, it is a systematic and rigorous process of building evaluation by measuring factors such as building use, energy consumption, operations and management and user satisfaction. PROBE studies, one of the early energy-focused POE studies, reports the results of investigations on several buildings (Bordass, et al., 2001a). A more recent POE study, Building Performance Evaluation Programme (BPE), assessed buildings' fabric, systems and occupant satisfaction with an aim to identify best strategies and pitfalls causing performance gap (Palmer, et al., 2016a; Palmer, et al., 2016b). A RIBA report, Building Knowledge: Pathways to Post Occupancy Evaluation, provides practical guidance for undertaking POE (RIBA and Hay, et al., 2016).

A POE process comprises an evidence-based investigation of the current performance of the building, which can identify issues causing underperformance and analyse possible causes. However, POE often does not go into identifying the root causes of the issues in a procedural and quantifiable manner.

NCM framework for energy efficiency finance

This framework (previously known as 'The Green Deal' for non-domestic buildings) is an approach to model and analyse the in-use performance of buildings with an intention of evaluating the impact of

ECMs (BRE, 2012). In this framework, an initial model is first developed using SBEM or DSM, as done in regulatory compliance calculations in the UK. Then this model is updated by changing the standardised operating conditions assumed under the NCM, as per actual operations, for a closer definition of the building (e.g. occupant density and schedule, equipment gains, temperature set points and ventilation rates). Further, several multipliers are derived to adjust the outcomes of the models to accommodate the effect of building management and maintenance. Finally, actual data on fuel wise annual consumption and the reliability of the data are used to calculate a normalisation factor that will be applied to the modelling results.

The simplicity of methods used in NCM calculations limits the detail and accuracy of the estimation of the performance of many processes. Also, there are no criteria set for the accuracy of thermal models in relation to actual energy consumption. Moreover, an unbounded and unexplained normalisation factor provides an avenue for having unsubstantiated factors to make the predicted results more closely match actual performance (Burman, et al., 2014).

IPMVP and ASHRAE Guideline-14 M&V protocols

IPMVP (EVO, 2016) and ASHRAE Guideline 14 (ASHRAE, 2014) are complementary frameworks created for measuring and ascertaining performance using best practice techniques in the implementation of ECMs. Another measurement and verification guideline, FEMP (Webster, et al., 2015), is also related to IPMVP and ASHRAE Guideline -14, with cross-referenced methods and protocols. These protocols are structured around calculating savings due to ECMs and provide guidance to reduce uncertainty in assessing performance; however, the procedures recommended in them can be followed for establishing an energy use monitoring system or assessing the opportunities for improving energy performance.

IPMVP and ASHRAE Guideline 14 provides guidance on measurement boundary, measurement period, ways to calculate impact (savings) and undertaking operational verification. Depending on the type of ECM and its relationship with other building performance input or output parameters, IPMVP provides four options for calculating ECM impact.

1. Option A: Retrofit isolation, Key Parameter Measurement
2. Option B: Retrofit isolation, All Parameter Measurement
3. Option C: Whole-Facility
4. Options D: Calibrated Simulation

While the first two options look at isolating the assessment and analysis to one or a few building systems that are affected by the ECM, the other two are done at a whole building-level. For each of the options, IPMVP explains the data required and the monitoring and measurement protocols. It also suggests ways to undertake accurate calculations, validate results, and link to other relevant standards and protocols.

IPMVP Option D: Calibrated Simulation suggests using building performance modelling tools for energy consumption and demand simulation, calibrated with hourly or monthly utility billing data. The IPMVP framework provides a step by step method to fine-tune the model to reflect the building and its operating conditions accurately. It is underpinned by ASHRAE Guideline 14, Measurement of Energy, Demand and Water Savings, for criteria to check calibration accuracy either at hourly or monthly intervals (discussed in Section 2.3.2). In these protocols, detailed operational information needs to be collected during site surveys and by measurements to calibrate the model.

While IPMVP and ASHRAE Guideline 14 provide thresholds for statistical indices to measure the goodness-of-fit of the building simulation model to the real data, they do not provide a detailed methodology for undertaking model calibration. The primary focus of these protocols is on quantitative requirements and guidance given in these protocols is not tied in a framework for a procedural verification of all issues. Therefore, it is not certain that the technical issues uncovered in a building through, e.g., on-site investigations, reflect all or most of the key causes of the performance gap. It is likely that some key issues are identified during the investigations, whilst other potential issues are not uncovered. Also, if the operation-stage information is limited, it is not possible to procedurally evaluate, with high confidence, the value of inputs to progress with calibration. At that a stage modeller needs to rely on estimations to meet the criteria.

BS EN 15603 approach

BS EN 15603: Energy performance of buildings (BSI, 2008) standard also includes a procedure for validation of the building calculation models. This validation process enables the attainment of a higher level of confidence in the building simulation model and input data, compared to the abovementioned methods, by probabilistically comparing calculated data with actual energy use. The main difference is that the method allows for determining the confidence interval arising from uncertainties for all input data. The input data that cannot be assessed is to be taken from inference rules, national references, or standards. Under this procedure, validation is carried out based on annual energy performance. However, in BS EN 15603, no specific criteria are provided to define reasonable consistency. Therefore, under this approach accuracy of a simulation model should be determined on a case-by-case basis, depending on the application it is being used for.

2.3.2 Validation criteria

Calibrated model validation criteria confirm whether the model is a reasonable representation of the actual building. In most studies and M&V protocols, NMBE and C_V(RMSE) are the two statistical indices that are calculated to check for validation. M&V protocols define the acceptable error thresholds for these matrices. Table 2.3 lists the thresholds for NMBE and C_V(RMSE) as defined in ASHRAE Guideline 14, IPMVP and FEMP for monthly and hourly data, with ASHRAE Guideline 14 thresholds being the most frequently used.

Table 2.3: Calibration criteria defined in measurement and verification protocols

Index	Guideline 14	IPMVP	FEMP
$NMBE_{monthly}$ (%)	±5	±20	±5
$C_V(RMSE)_{monthly}$ (%)	15		15
$NMBE_{hourly}$ (%)	±10	±5	±10
$C_V(RMSE)_{hourly}$ (%)	30	20	30

Using statistical criteria as a calibration check, within the intended use of ECM evaluation in these M&V protocols, is expected to provide suitable results. However, reviewing these tolerances, Royapoor and Roskilly (2015) suggest that when hourly annual data is available, the ASHRAE Guideline 14 requirements of NMBE and C_V(RMSE) can be brought down to +/- 5% and 20% respectively. When hourly data is not available, the ASHRAE Guideline 14 limits are realistic.

ASHRAE Guideline 14 focuses on two levels of calibration based on the temporal granularity of the available data, monthly and hourly. For detailed performance assessments, using hourly data calibration is time consuming, but monthly data might be too coarse. Due to the underdetermined nature of model calibration, the current criteria defined for calibration for monthly data is suitable; however, if a more detailed analysis is required at hourly resolution, further cross-checking and validation (in addition to NMBE and C_V(RMSE)) is also needed (Ruiz & Lemort, 2014). For high-resolution data, Ruiz and Lemort (2014) highlight that NMBE and C_V(RMSE) criteria are necessary but not sufficient because it is not possible to accurately match the highly dynamic behaviour of both simulated and actual results at small time-steps. They suggest that along with the above-mentioned statistical indices, bin analysis could be used.

Limitations of using these metrics have been discussed in Garrett and New (2016) and Ruiz and Bandera (2017). As the criteria are a deterministic statistical index, it fails to capture that multiple solutions may exist that might meet the criteria, but some of the solutions may not necessarily reflect the actual performance. For example, the gap in heating demand can be closed by increasing either indoor set point temperatures (observed by IEQ measurements of some typical zones) or mechanical ventilation supply (based on BMS observations during site visits), or both. Therefore, based on limited data, even when ASHRAE Guideline 14 criteria are met, it is not possible to deterministically identify the exact deviations in both these areas.

Moreover, some of the solutions identified might be mathematically correct but physically impossible to achieve. If the operation stage information is limited, it is not possible to procedurally estimate, with certainty, the exact value of inputs to progress with calibration. At that stage, the modeller needs to rely on estimation to deliver the calibrated model. Defining weekly and daily level checks and cross-validation using secondary data streams, such as disaggregated end uses, loads checks and zone set points (EVO, 2016), can help fix some of the unrealistic solutions issues. Further, the incorporation of an uncertainty based probabilistic approach can be useful in determining confidence levels in the validated model (BSI, 2008).

2.3.3 Advanced calibration checks

To mitigate the limitations of just using NMBE and C_V(RMSE) thresholds as calibration checks, more data streams and analysis methods can be used. Annex C of ASHRAE Guideline 14 discusses graphical data techniques, such as 24-hour profile plots, box-and-whisker-plots and 3-D plots. These could be used to visually evaluate the accuracy of the calibrated model. Also, IPMVP recommends checking of building loads and energy use patterns and comparing measured and simulated data in form of bar charts; monthly percent difference time-series graphs and scatter plots. These are good techniques that can be used for detailed evaluations; however, neither of the documents provide specific implementation guidelines of these techniques. Therefore, practical use of these techniques is limited, and it is up to the modeller's judgment and ability to analyse the accuracy. Addressing the issue of the need for calibration cross-checking that can be evaluated objectively, this section reviews three advanced calibration checks, highlights their uses and defines their requirements for proper application.

Additional statistical and graphical checks

Coefficient of determination (R^2): This is a statistical index that also measures deviation. R^2 is a recommended metric in ASHRAE Guideline 14 but does not form a part of the formal checks within which its model validation criteria are defined. The R^2 measures how close simulated values are to the regression line of the measured values. Between 0.00 and 1.00, the higher the value, the better the predictive capability of the model. R^2 value is linked to the slope of the best fit line in the actual vs predicted performance graph. Higher slopes tend to show a higher value of R^2 even though the variance measured by C_V(RMSE) is the same. Figure 2.5 shows two different datasets with the same C_V(RMSE) but vastly different R^2 values. The lesser the slope of the line, the more difficult achieving high R^2 value becomes.

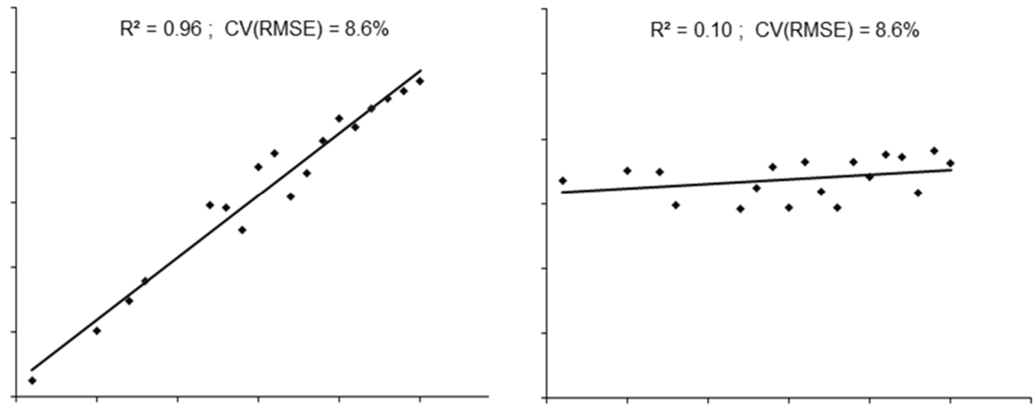


Figure 2.5: R^2 values of datasets with same C_V(RMSE) (Vesma, 2016)

ASHRAE Guideline 14 suggests an R^2 value of 0.75 as the target value for energy use parameters. However, statistically, R^2 values of 0.3, 0.5, and 0.7 are interpreted to show weak, moderate and strong effects (Moore, et al., 2015). Achieving an R^2 of 0.75 for energy data, as per ASHRAE Guideline 14, is appropriate, as the data spread is like the left graph in Figure 2.5. However, for data that have clustering like the right graph in Figure 2.5, such as zone temperatures, achieving an R^2 of 0.5 (i.e. moderate effect) can be considered suitable and within an acceptable range of variation.

Other statistical indices: To check for model calibration, indices of NMBE and C_V(RMSE) are used. These indices are scale-dependent and therefore can lead to high variation if the same tolerance limits as energy performance are applied for other variables which are in other units, such as zone temperatures. Statistical indices that address absolute errors such as mean absolute error (MAE) and root mean square error (RMSE) are used in climate research (Chai & Draxler, 2014) and also used by Roberti, et al. (2015) for temperature calibration checks in energy simulation models. Indices not used by ASHRAE Guidelines, FEMP and IPMVP, but which have been found to be useful for calibration purposes in other studies are the goodness-of-fit (Ruiz, et al., 2016), Range Normalised Root Mean Squared Error (Chakraborty & Elzarka, 2018), and Unscaled Mean Bounded Relative Absolute Error (Chen, et al., 2017).

Graphical techniques: Along with these statistical checks, some complementary graphical tests mentioned earlier, can be used to assess the deviations, such as bin analysis which is the most used graphical method. Bin analysis provides a probabilistic distribution of data, and data can be presented as relative frequency and cumulated relative frequency for each of the bins. These bin analyses can be done for absolute values as well as residual error values (which is linked to R^2), and mean and standard deviation of the errors can be calculated. Figure 2.6 and Figure 2.7 show examples of these graphical plots sourced from calibration studies using them.

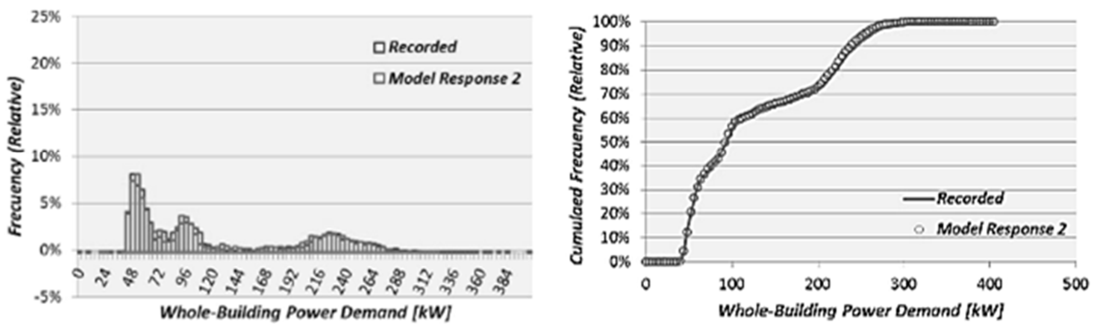


Figure 2.6: Relative and cumulated frequency comparison between recorded and predicted data (Ruiz & Lemort, 2014)

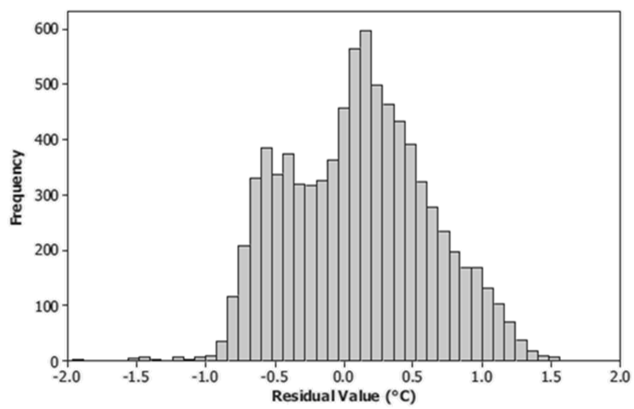


Figure 2.7: Histogram of hourly temperature residuals (Royapoor & Roskilly, 2015)

Calibration for disaggregated data

The minimum output data needed for comparison in any calibration is facility-level energy use, for all fuels. However, calibrating for more data streams through disaggregated energy use can further improve the accuracy of, and confidence in, the calibrated model (Soebarto, 1997; Reddy, 2006; Raftery, et al., 2011). This information is increasingly available and can be taken during audits and short-term and long-term end-use metering of energy data (Reddy & Maor, 2006; Penna, et al., 2015).

Disaggregation compartmentalises energy use, thereby reducing the chances of cross-compensation. Disaggregation can be done for different end-uses or spatially using energy sub-meters. Sub-metering of energy use based on end-uses can help in isolating some inter-dependent uses and undertaking a more granular analysis, where a dominant highly influential parameter for an end-use does not end up masking the influential parameters for other energy end-uses. At a minimum, such disaggregation at the end-use level should separate weather-dependent loads, such as heating and cooling, from other occupant-driven and non-weather-dependent ones, such as small power and lighting (Soebarto, 1997).

The disaggregation of sub-meters spatially separates areas that have distinct operating conditions. These disaggregated meters can provide insights into usage patterns and operations at a highly granular level. Disaggregation of hourly energy use data in this manner can help in trend analysis, thereby giving an opportunity to create typical profiles. Spatial disaggregation should isolate different floors, zone types and tenancy (Coakley, et al., 2014). Building codes and standards such as Part L in the UK (DCLG, 2013) and Title 24 in California, US (CEC, 2018), among others, now have minimum requirements for sub-metering.

Integrating spatial disaggregation with end-use disaggregation, typical recommendations for sub-metering should include sub-meters for all individual fuels which are separated for all energy end-uses. These sub-meters should have larger end uses separated by floor, area, type, or tenancy. Zones or zone clusters more than 500 m^2 in area and having a load of more than 25 kW should be considered to have separate meters (CEC, 2018).

For disaggregated meters, it is better to have full-year high-resolution data that coincides with the overall calibration period, for better use in fine-tuning of the model. However, if that is not possible due to practical limitations, then short-term intensive monitoring should be carried out where granular (hourly or finer) data is collected for typical weeks in different seasons to generate the typical profiles (Penna, et al., 2015). The criteria used for assessing calibration for these short-term periods can be based on the same statistical parameters but with finer acceptability ranges.

Cross-validation of other dependent results

The quality of calibrated models can be improved by cross-validating the simulation results with other dependent parameters. Zone temperatures and other IEQ parameters are easily available and can be a reliable data source. Monitoring of IEQ data streams can also provide evidence for detailed building operational profiles. Temperature data can help in finding the set-points used, and CO_2 and $PM_{2.5}$ concentrations can help to check occupancy, ventilation and infiltration rates (Kapalo, 2013; Parsons, 2014; Batterman, 2017).

IPMVP suggests verification of system loads and zone level set points (e.g. temperature and humidity) when validating a calibration. Besides this, IEQ parameters, such as CO_2 concentration and peak and off-peak load profiles, could also be used. However, IPMVP does not provide definitive criteria for undertaking these checks. Structured incorporation of cross-validation checks in model calibration needs requirements to be defined, such as the selection of the parameters and their measurement frequency and duration.

Minimum cross-validation checks in calibrated models should be done for zone air temperatures, as IEQ assessments are often largely based on room temperature distributions (Royapoor & Roskilly, 2015; Roberti, et al., 2015). These checks should be done for representative rooms for all zone types, orientations and operational conditions, covering at least 10% of the floor area. Guidance regarding these is given in BS EN 12599 (BSI, 2012) and BS EN 16798 (BSI, 2019). The checks need to be done hourly, and, if not for the entire year, they should cover at least two consecutive weeks for summer, winter and any other seasons.

As discussed earlier, for temperature checks, the NMBE and C_V(RMSE) criteria for energy use calibration validation in ASHRAE Guideline 14 might not be suitable. Indoor temperatures vary within a small range, and ASHRAE Guideline 14 acceptance limits can lead to an inclusion of large deviations from the comfort bands. There are no standards or formal guidelines that define acceptance criteria for zone air-temperature calibration. Being relative statistical indices, NMBE and C_V(RMSE) are scale-dependent and can lead to high variation in temperatures if the same limits applied to energy performance are used to check model calibration for temperatures. Statistical indices that address absolute errors such as MAE and RMSE were used by Roberti, et al. (2015). In most studies, acceptable absolute error recorded was in the range of ±1-2°C for most of the temperature data points (Booten & Tabares-Velasco, 2012; Royapoor & Roskilly, 2015; Ruiz, et al., 2016; Beizaee, 2016). Therefore, MAE of 1°C and RMSE of 1.5°C can be used as reasonable targets for air temperature calibration checks.

Probabilistic calibration results

Due to the uncertainty in input parameters and also when there is a lack of regularly monitored data through the calibration period, there can be numerous possible permutations that can create a validated ASHRAE Guideline 14 model. A probabilistic approach can be more suitable to represent and explore these cases. The BS EN 15603 (BSI, 2008) approach, discussed earlier, suggests including the input uncertainty in energy simulations and providing energy use results probabilistically. Similar to that approach, in model calibration, the observed deviations in measured data that are based on trends can be used to define the upper and lower range within which the uncertain input parameters can lie. The results of the simulation model can then be presented with confidence bands that will represent uncertainty around the most likely value of the data points. When this approach is used in model calibration, the measured results will be plotted against the corresponding data point of simulation results with its uncertainty range.

The aim of calibration in this representation is then to ensure that the measured value is within the upper and lower ranges of uncertainty. If required, the model can be fine-tuned to reduce the output uncertainty band. Figure 2.8 shows a probabilistic box plot of predicted monthly energy consumption with the measured data. Using a box plot provides a graphical display of the predictions' symmetry, skewness and spread at a glance.

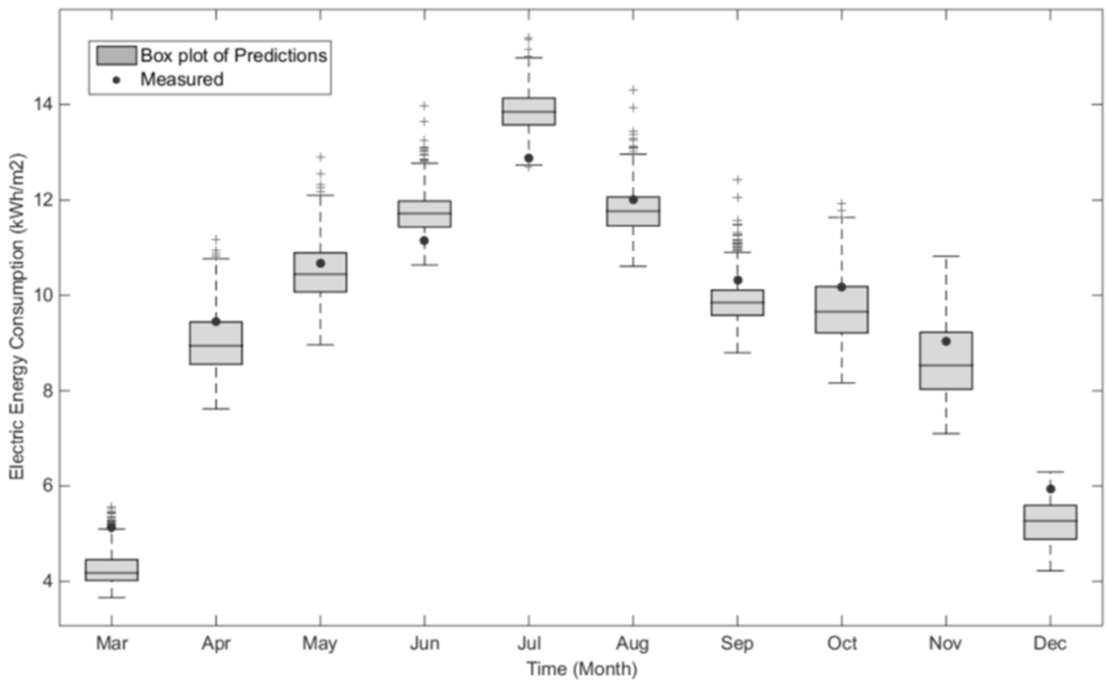


Figure 2.8: Probabilistic calibration results (Chong, et al., 2015)

2.4 ENERGY PERFORMANCE GAP AND ITS CAUSES

The phenomenon in which the estimated (simulated) performance of the building shows a significant discrepancy to the actual (measured) performance during occupation is referred to "the performance gap" (Carbon Trust, 2012; Menezes, et al., 2012; Burman, et al., 2014; de Wilde, 2018). The most visible and researched 'performance gap' is seen in energy performance predictions. Initially highlighted by The PROBE studies (Bordass, et al., 2001a) and later supported by numerous other studies, actual energy performance can be 2.5 times greater than predicted (Menezes, et al., 2012). In this section, we look at the current definition and classification of the energy performance gap, its evidence across the building sector and the main causes of the gap.

2.4.1 Definition of the performance gap

The performance gap is a widely known term; however, its in-use definition is quite vague. As discussed in Section 2.1, depending on the baseline* selected, the magnitude and scope of the performance gap can vary; for example, the use of compliance modelling results instead of performance modelling results can lead to a perceived gap which can have artificial inflation added to the actual performance gap.

Comparing design estimates from performance modelling with the actual performance gives the actual gap (See Figure 2.2). This actual gap represents the correct gap, as the baseline* used to ascertain the performance gap is based on design intents at the design or handover stage of building procurement. However, building performance and its issues can be analysed throughout its lifecycle, and calibrated simulation models can be used to assess the performance gap with a longitudinal perspective. From a longitudinal perspective, Bunn and Burman (2015), classify the actual performance gap during the operational stage into two parts, static and dynamic performance gaps.

The static performance gap is the gap between the actual performance of the building attains when it is in a steady mode of operation during a fixed period, such as after one-year post-occupancy. The dynamic performance gap is the gap between design intent and ongoing performance, which will include subsequent deviations in building operations that may increase or decrease the gap, as the building carries on with its regular operations. Technical issues are major contributors to static gaps and operational changes are to the dynamic gaps.

Issues that cause the performance gap are discussed and qualitatively analysed in the next section to understand their role in various stages of the building lifecycle (design, construction, commissioning, handover and operations).

* reference point for calculating the performance gap

2.4.2 Causes of the energy performance gap

A key finding of the PROBE studies was that there was very little connection between design assumptions and actual on-site observations. The actual energy use of most buildings was almost twice the design estimates (Bordass, et al., 2001a). Evidence from other countries corroborates the findings of PROBE. A two-to-one discrepancy between actual and predicted energy performance in office buildings was identified and investigated in the US by Norfold et al. (1994). Studies on LEED-certified buildings also revealed significant deviations in measured performance from their design projections (Turner & Frankel, 2008; Samuelson, et al., 2014; Burman, 2016). Table 2.4 provides a summary of the design and actual energy performance in the UK reported across the four case-study building sectors. Being voluntarily provided data, design calculations are not all necessarily calculated in the same way, with some being calculated with proper design assumptions and others based on regulatory compliance calculations. However, because of the significant sample size, this data can be used to highlight the performance gaps indicatively.

Table 2.4: Design and actual energy performance for various building sectors in the UK

Sector	Design* ($kWh/m^2/yr.$)		Actual ($kWh/m^2/yr.$)		Source
	Gas	Elec.	Gas	Elec.	
Office	46	71	73	121	CarbonBuzz (Kimpian & Chisholm, 2011) 80 design prediction & 113 actual performance samples
School	57	56	84	106	CarbonBuzz (Kimpian & Chisholm, 2011) 133 design prediction & 203 actual performance samples
Hospital	317	122	373	143	(Morgensterna, et al., 2016) About 150 general acute hospital samples.
Apartment	29	15	73	38	(Palmer, et al., 2016b) 76 homes. Data in $kgCO_2/m^2/yr$. Carbon intensity ($kgCO_2/kWh$) Gas 0.194; electricity 0.55

** Sourced from voluntary platforms, some of the design energy performance within the sample can be based on compliance modelling.*

The performance gap can be due to many factors during building design, construction and operation (de Wilde, 2014), and understanding the causes is essential to increasing, and more importantly maintaining, confidence in the performance estimates and the tools used for them. A review of 28 case studies in the UK found that 75% of buildings underperformed due to serious shortcomings in construction practices, control strategies, commissioning, building fine-tuning in the early stages of post-occupancy, user training, building management and maintenance (Carbon Trust, 2012; Shrubsole, et al., 2018). Another cross-sectoral study documented problems associated with building fabric, control strategies, commissioning, installed metering strategies and inadequate provision of training (Palmer, et al., 2016a). Similarly, a study analysing buildings built under the Australian Building Greenhouse Rating scheme uncovered examples of poor controls design, construction and commissioning issues along with problems related to building maintenance and operation (Bannister, 2009). Some of the main causes of the performance gap are.

1. **Limitations and complexity of the design decisions:** Calculations and decisions done at the building design stage are done with a lot of assumptions and limited knowledge of the impact of early design decisions. The complexity of the design leading to subsequent changes, such as value engineering, adds to the uncertainty and deviations from the original design intent.
2. **Modelling and calculation mistakes:** Besides design decisions, calculation methods, assumptions and tools used to ascertain performance are also sources of uncertainty. This is directly related to uncertainties in building energy simulation, especially model simplification, user error and errors introduced in the discretisation of the model results.
3. **Poor construction and commissioning:** During the construction and handover stages of a project, on-site workmanship issues such as discontinuous insulation or punctured airtight barriers can occur. Also, poor system balancing and checking during commissioning results in reduced system efficiency and can compromise the airtightness and ventilation strategies.
4. **Operational inefficiencies:** Occupants and building managers have a substantial influence on the energy performance of a building as they affect operations and control requirements of various systems, which may deviate from assumed schedules. Also, building conditions may change over-time, such as building occupancy, space usage and deterioration of physical elements. These factors can also lead to deviation from the intended performance.

Table 2.5 summarises the major performance gap factors reported in the existing literature and classifies them to various construction stages of a project.

Table 2.5: Energy performance gap factors for various building stages

Building stage	Performance Gap Factors	Source
Design	**Design targets:** Energy targets may be poorly specified in briefing and design criteria and can conflict with other targets (e.g. energy efficiency vs. minimum air change rates).	(Bordass, et al., 2001a; Bannister, 2009)
	Design issues: Recurring issues include specification of centralised systems for local loads, HVAC system oversizing and poor zoning, and poor control interface between low or zero-carbon technologies and backup systems.	(Bordass, et al., 2001a; Bannister, 2009; Burman, 2016)
	Modelling issues: The method used for energy performance calculation, modelling uncertainty, software variability and user variability can have a significant impact on the accuracy of such calculations.	(Guyon, 1997; Ahmad & Culp, 2006; Wang, et al., 2012a; Neymark, et al., 2002)
Construction	**Poor build quality:** Recurring issues include poor airtightness and gaps in fabric insulation, thermal bridging at construction junctions, and poor installation of building services, especially the air distribution systems in mechanically ventilated buildings.	(Bordass, et al., 2001b; Petersen & Hviid, 2012; Wingfield, et al., 2013; Bordass, et al., 2004)

	Value engineering process: If key determinants of energy performance are not protected, the value engineering process may compromise the original design intent.	
Commissioning	**Basic commissioning:** Major commissioning flaws can compromise energy efficiency (e.g. incorrect equipment installation and poor commissioning of controls).	(Piette, et al., 1994; Bannister, 2009; Pang, et al., 2012; ZCH, 2014)
	Lack of seasonal commissioning: This could be critical, especially in complex buildings (e.g. advanced naturally ventilated buildings with different settings for the heating season and summer performance).	(Burman, 2016)
Handover	**Inadequate training:** Effective training of building users including key personnel such as facility managers is essential.	(Carbon Trust, 2012)
	Incomplete documentation: (e.g. building logbooks).	(Palmer, et al., 2016a)
Operation	**Lack of building fine-tuning:** Post-occupancy evaluation can help identify and address performance issues.	(Menezes, et al., 2012)
	Occupant behaviour: This is a source of uncertainty at the design stages but also affects actual building operation.	(Azar & Menassa, 2012; Martani, et al., 2012)
	Poor maintenance: Longitudinal performance of the building including quality of maintenance will affect the performance gap.	(Bannister, 2009; de Wilde & Jones, 2014)

2.5 IEQ PERFORMANCE GAP AND ITS CAUSES

There are many parameters which can fall within the remit of building performance. These include but are not limited to energy, thermal comfort, IAQ, lighting and acoustics. In the context of the performance gap, as mentioned earlier, energy performance is the most emphasised. However, depending on the scope of performance assessment, the definition and purview of the 'performance gap', its magnitude, underlying causes and mitigation measures can vary significantly. IEQ is closely linked to and inter-dependent with energy performance and achieving high performance in IEQ as well as energy use in buildings can be convergent or divergent objectives. This section discusses various IEQ parameters, their performance issues and reasons for underperformance.

2.5.1 Changing focus towards IEQ

The importance of the indoor environment in buildings (primarily dwellings) and the impact of shortcomings in the indoor environment on health and associated mitigation measures have been documented and laid down by the World Health Organisation (WHO) as early as 1990 (WHO, 1990). The direct relationship between occupant well-being, comfort and productivity with IEQ in various building types was also measured and documented in later research (Wyon & Wargocki, 2013; Chatzidiakou, et al., 2014; Al Horr, et al., 2016).

The policy-driven focus of lowering carbon emissions has made energy efficiency a key objective for new buildings and major renovations. Also, the debate about the performance gap has raised awareness about the need to meet energy performance goals in practice. With this focus on energy efficiency and the 'energy' performance gap, buildings are generally designed to achieve a specific energy use and carbon emissions target. Also, improved metering with readings at hourly and half-hourly intervals makes the 'energy' performance gap easily identifiable and disproportionately more visible (de Wilde, 2014).

However, this is not necessarily the case for IEQ parameters. While adhering to IEQ performance standards is an essential aspect at the design development stage, the energy and carbon emissions reductions are the primary, and often the only, performance objective. IEQ has only come to prominence relatively recently. This can be seen in the changes to the revised version of the EPBD (EPBD 2018/844). The original EPBD (EPBD 2010/31/EU) focused on reducing energy use. The revised EPBD highlights the importance of health, comfort, IAQ and indoor climate. However, details about their practical implementation are yet to be addressed (Council Directive, 2018).

2.5.2 IEQ performance evaluation factors

The key constituents of IEQ that are linked to occupant health, comfort, productivity and overall wellbeing are thermal comfort, IAQ, lighting and acoustics. Incorporation of good IEQ in buildings and their assessment is done objectively; however, these are experienced by building occupants in a subjective manner. This makes the analysis and evaluation of IEQ performance difficult.

Thermal comfort

Thermal comfort in buildings is a combination of personal parameters, such as activity (metabolic rate) and clothing and environmental parameters, such as air temperature, mean radiant temperature, air speed and humidity (ASHRAE, 2017). The most commonly used thermal comfort indices are predicted mean vote (PMV) / percentage of people dissatisfied (PPD) (Fanger, 1970) and adaptive thermal comfort (de Dear & Brager, 1998).

PMV/PPD: PMV is the 'predicted mean vote' on a seven-point scale. The scale is from -3 (cold) to +3 (hot), with 0 as neutral/comfortable. PPD is the ‘percentage of people dissatisfied’ at each PMV. As PMV deviates from zero towards +3 or -3, PPD increases. PMV index is based on the combination and interdependencies of metabolic activity, clothing insulation, air temperature, mean radiant temperature, air movement and humidity. The PPD index quantifies dissatisfaction with the environment by relating it to the PMV. At PMV = 0, PPD is the lowest, at 5%, and at the extreme PMVs, PPD increases towards 100%.

Adaptive thermal comfort: Adaptive thermal comfort is an index based on the principle that building occupants can adapt to and accept different internal temperatures based on outdoor climatic conditions. Cultural and technical practices and past thermal history influence the thermal expectations and preferences of the occupants. Types of adaptive actions that can be taken are modifying (de Dear & Brager, 1998):

- Internal heat generation, such as by shivering or siesta.
- Rate of body heat loss, such as by sweating or wearing more clothes.
- Thermal environment, such as by opening windows.
- Selecting a different environment, such as by moving closer to the fire or window.

ASHRAE Standard 55 (ASHRAE, 2017) provides an adaptive comfort model for determining acceptable thermal conditions in naturally conditioned spaces. In naturally conditioned buildings, thermal neutrality for operative comfort can be calculated based on mean monthly outdoor air temperature. When internal operative temperatures are within ±2.5°C and ±3.5°C of the comfort operative temperature, then building complies with ASHRAE Standard 55’s thresholds of 90% acceptability and

80% acceptably respectively. The 80% acceptable limit is used generally, and 90% acceptability is used when a higher thermal comfort standard is required.

<u>Overheating:</u> Overheating is a phenomenon of more than acceptable indoor temperatures, often seen in spaces that do not have active cooling systems. Due to local and regional climatic and cultural differences, the threshold of overheating and acceptable indoor thermal environment varies. Generally, indoor air temperature is used to assess overheating; however, other thermal comfort factors, such as relative humidity, radiant heat and air movement determine thermal comfort. Additionally, the duration and intensity of heat exposure can also influence thermal sensation (BRE, 2016).

In the UK, CIBSE TM52 (CIBSE, 2013b) and CIBSE TM59 (CIBSE, 2017) provide guidance for calculating the risk of overheating in non-domestic and residential buildings, respectively. CIBSE TM52 sets out two comfort temperature thresholds based on BS EN 15251 (BSI, 2007) (now replaced by BS EN 16798-1 (BSI, 2019)), maximum acceptable temperature (T_{max}) and upper limit temperature (T_{upp}). CIBSE TM52 recommends a T_{max} of 3°C and T_{upp} of 7°C above the comfort temperature. Comfort temperature, based on outdoor temperature, is calculated as per exponentially weighted running mean of the daily mean outdoor air temperature when the building is in a free-running mode.

For mechanically conditioned spaces, CIBSE TM52 recommends maximum space temperatures. When the clothing is assumed to be 1.0 clo in winter and 0.5 clo in summer, temperatures should not exceed 24°C in winders and 26°C in summers. In residential spaces, this threshold is 25°C for winters.

Indoor air quality

IAQ is constituted by the individual and cumulative effects of multiple pollutants in indoor air. These pollutants are generated from indoor sources that include, cooking, cleaning and emissions from construction materials and space furnishing. Additionally, in dense urban environments, the IAQ is heavily affected by outdoor sources such as road traffic, industry and construction seeping through the building envelope (façade, windows and doors).

CO_2 has been widely accepted in the industry as an indicator of IAQ because its main generation source is occupancy, so levels of fresh air can be determined easily. However, traffic-related external pollutants such as $PM_{2.5}$, PM_{10} and NO_2 are linked to adverse health impacts as well (AQEG, 2004; AQEG, 2005). This is important in buildings with natural ventilation, where air exchange between the indoor and outdoor environment occurs without any filtration. In such circumstances, the use of CO_2 as the only metric used as a proxy for IAQ is questionable. While CO_2 levels provide the first indication of exposure to poor air quality, indoor levels of traffic-related pollutants need to be considered separately (Chatzidiakou, et al., 2014).

A comprehensive IAQ assessment would include the evaluation of concentrations of all indoor air pollutants (including the numerous volatile organic compound - VOCs). However, practical limitations of measurements and their accuracy, along with the need to integrate them in system controls, makes the use of proxies unavoidable. Using CO_2, $PM_{2.5}$ and NO_2 as proxies of indoor and outdoor contaminants and traffic pollution could provide a holistic way to control ventilation strategies for optimised IAQ.

CIBSE Guide A (CIBSE, 2015) provides guidance on recommended CO_2 levels for various categories of IAQ, Class I is the most stringent category and Class IV is the most relaxed, minimum acceptable, category. Indoor CO_2 concentrations of 400 ppm above the outdoor CO_2 concentration level is the limit for Class I category. And indoor CO_2 concentrations of 1000 ppm above the outdoor CO_2 concentration level is the limit for the Class IV category. Regarding $PM_{2.5}$, PM_{10} and NO_2, thresholds for exposure have been defined by the WHO (2018). Guideline values for fine particulate matter ($PM_{2.5}$) are defined as 10 $\mu g/m^3$ annual mean and 25 $\mu g/m^3$ 24-hour mean. For coarse particulate matter (PM_{10}), guideline values have been defined as 20 $\mu g/m^3$ annual mean and 50 $\mu g/m^3$ 24-hour mean. The current WHO guideline value for NO_2 is set as 40 $\mu g/m^3$ annual mean and short-term concentration exceeding 200 $\mu g/m^3$ is deemed to be toxic.

Lighting and acoustics

The functional requirement of lighting in a building is to ensure visual comfort while the tasks that require good light quality can be carried out with ease and proficiency. Lighting has two components daylight and artificial lighting. Daylight is a natural source of light and must meet the functional demand of light, but at the same time deal with glare and heat gain along with reducing energy consumption by offsetting the need for artificial lighting. Daylight needs to create an environment where occupant's health and wellbeing can benefit from it. Artificial lighting is installed to supplement the lighting in the absence of adequate daylight. Key factors that can quantifiably determine the quality of lighting for visual performance are luminance and illuminance levels and their distribution, colour appearance, and colour rendering.

Acoustics in buildings require management of sound with respect to noise in and around buildings and from the quality of internal sound for functions that are particularly relevant for acoustics (e.g. learning and teaching spaces). In most spaces, a background ambient noise level of 65 dB is a threshold for acoustic comfort. Along with basic background noise levels measurement, BB93 (Department for Education, 2003) suggests acoustic measurements of reverberation time to check for speech intelligibility. CIBSE TM40: Health and wellbeing in building services (CIBSE, 2020) provides detailed guidance on regulatory framework and performance criteria for lighting and acoustics.

IEQ performance requirements

Performance requirements for each of the IEQ parameters have been discussed in their respective sections. In addition to those, BS ISO 17772-1 (BSI, 2017) "Energy performance of buildings. Indoor Environmental Quality. Indoor environmental input parameters for the design and assessment of energy performance of buildings" is a complete standard for IEQ assessment. This standard specifies requirements for temperature, IAQ, lighting and acoustics. It includes IEQ criteria for local thermal comfort factors, draught, ventilation rates, radiant temperature asymmetry, vertical air temperature differences and floor surface temperatures. It also specifies occupancy schedules to be used in standard energy calculations and defines four levels of IEQ depending on the level of expectations occupants might have.

2.5.3 Causes of the IEQ performance gap

PROBE studies investigated thermal comfort, acoustic performance, perceived control, and the misfit between building performance and user expectations in the case studies and found that buildings were underperforming (Bordass, et al., 2001b). In current practice, while adhering to IEQ standards is an essential aspect during design, the actual performance post-construction is not usually evaluated. Similar to the performance gap in energy use, the underperformance is also seen in IEQ parameters such as temperature, air quality (pollutants, CO_2), noise and lighting (Tuohy & Murphy, 2015; Fabbri & Tronchin, 2015; Phillips & Levin, 2015). Table 2.6 describes the IEQ performance gap factors documented in various studies across the stages of building procurement.

Table 2.6: IEQ performance gap factors for various building stages

Building stage	Performance Gap Factors	Source
Design	**Design targets and specifications:** Lack of pre-defined IEQ goals in briefing and design criteria. Poor choice of construction material (e.g. non-low-emitting materials or materials with high emissions at high temperatures).	(Bordass, et al., 2001a; Bannister, 2009; Chatzidiakou, et al., 2014; Poppendieck, et al., 2016)
	Design complexities: Not factoring in all sound sources for managing acoustic performance conflicts because of space planning and design (e.g. quiet areas near noisy spaces or poor daylight design and poor control interface between natural and artificial lighting systems). Issues with designing HVAC system control.	
	Modelling issues: Modelling uncertainty, software variability, lack of integrated design software, and user variability can have a significant impact on the accuracy of such calculations.	(Guyon, 1997; Ahmad & Culp, 2006; Neymark, et al., 2002; Won, et al., 2005)
Construction	**Poor build quality:** Recurring issues include poor airtightness and gaps in fabric insulation, acoustic leaks, thermal bridging at	(Bordass, et al., 2001b; Wingfield, et al., 2013;

	construction junctions, and poor installation of building services, especially pipe and ductwork insulation.	Bordass, et al., 2004; Yu & Crump, 2010)
	Changes after design: Changes during construction can affect IEQ parameters significantly, as addressing the resultant IEQ issues might require significant alterations, even structural changes.	
Commissioning and Handover	**Systems' commissioning:** Major commissioning flaws can compromise the functioning of lighting and ventilation systems (e.g. incorrect equipment installation, faulty sensors, and poor commissioning of controls).	(Gupta & Kapsali, 2015)
	Inadequate training: Effective training of building users including key personnel such as facility managers is essential.	(Burman, et al., 2018)
	Incomplete documentation: (e.g. building logbooks, operation, and maintenance manuals).	(Palmer, et al., 2016a)
Operation	**Controls:** Design of controls strategies, user training, and fine-tuning following user feedback are likely to influence how the users interact with the building and the IEQ performance gap.	(Menezes, et al., 2012)
	Occupant behaviour: This is a source for uncertainty and affects actual building operation because of their role in controlling noise levels, lighting levels, and most times ventilation rates (window openings) and space conditioning set-points.	(Mumovic, et al., 2009; Tanner, et al., 2013)
	Poor maintenance: Quality of maintenance will affect performance. (e.g. poor light output from dirty luminaries or if filters are not cleaned regularly, then they can affect IAQ).	
	Longitudinal performance of the building: Changes over the life of the building can include changes in space use and occupancy levels, adding more equipment, repainting of surfaces and replacing of materials used, etc. Also, as the building ages, the efficacy of equipment (e.g. light sources) deteriorates.	(Mavrogianni & Mumovic, 2010)

2.5.4 Unintended energy-related IEQ underperformance

Ways to achieve better IEQ and building user satisfaction might contradict with measures to achieve better energy performance. Therefore, if the focus is only on energy or carbon emissions, this can lead to the unintended consequence of poor IEQ in buildings. Energy use reduction is not enough unless it allows buildings to perform their desired functions, i.e. to be healthy, comfortable and productive places to live and work (Jain, et al., 2017). This section explores some of the conflicts between various building objectives.

Energy vs thermal comfort and IAQ

Indoor thermal conditions expected by building occupants are critical for energy performance. High temperature set points in winter and low cooling set points in summers, where air conditioning is installed, can compromise energy performance. There is also a risk of conflict between the operation of heating and cooling systems if an effective dead band between the respective set points is not specified. It is important to verify that the installed building services can provide an acceptable level of thermal comfort.

An example of these conflicts is the overheating and air quality issues that are uncovered in some new buildings which are constructed to higher energy standards, with high insulation and airtightness (Logue, et al., 2011; Larsen, et al., 2012; Maivel, et al., 2015; Abadie & Wargocki, 2016). High ventilation rates are required to avoid overheating in buildings without air conditioning; these are often higher than fresh air requirements for IAQ. Therefore, it is important that the choice and the management of the installed ventilation system, which is closely related to energy performance, can provide and maintain acceptable air changes.

Energy vs IAQ

CO_2 concentrations are often used as a proxy for IAQ. However, in urban areas, traffic-related external pollutants, such as fine particulate matter ($PM_{2.5}$) and NO_2, are linked to adverse health impacts and can potentially compromise IAQ (AQEG, 2004; AQEG, 2005). This may have significant implications where ventilation strategies are adopted that improve energy efficiency but do not necessarily address outdoor pollution.

Energy vs lighting

Lighting control is critical to saving electricity. The daylight factor is a measure of how much natural light a space receives. Zones, within rooms, which are closer to natural light sources should have separate manual and automated lighting control to help save energy. Zoning of lighting sensors, the sensitivity of daylight sensors, and sensitivity and time offs of presence/absence detection sensors have an impact on energy performance.

Energy vs acoustics, thermal comfort, IAQ and lighting

Protecting an indoor environment from excessive ambient noise levels may lead to a requirement of mechanical ventilation instead of operable windows, which is a major risk factor for energy performance if not managed well. On the other hand, using thermal mass as a passive measure to moderate indoor temperatures may lead to excessive reverberation in the absence of acoustic absorption rafts. Therefore, assessment of internal noise and reverberation times are closely linked to the assessment of ventilation strategy, thermal comfort and energy performance.

Another perspective on the provision of operable windows for natural ventilation and comfort cooling in summer is to address outdoor noise ingress issues, and windows also need to be integrated within lighting comfort requirements. Outdoor noise ingress through openings can be an issue, along with the use of blinds for glare prevention. Internal blinds can conflict with airflow from open windows because of rattling. To address this, an integrated design solution for the façade balancing all requirements is required.

2.6 SUMMARY

The definition of performance gap varies depending on the baseline[*] selected (perceived gap vs actual gap) and the operational stage at which performance is evaluated (static gap vs longitudinal gap) (van Dronkelaar, et al., 2016). Once the performance gap has been suitably defined, it can be reduced by improving the predictive capabilities of the simulation tools and/or by reducing the deviation of actual operations from the initially intended operations. Primarily targeting the former, efforts have been made to close energy performance gaps by using more sophisticated simulation tools, developing a better understanding of building-occupant interactions and calibrating the models using monitored data (Andre, et al., 2008; Haldi & Robinson, 2008; van Dronkelaar, et al., 2016; de Wilde, 2018).

Calibration is one of the methods used to increase accuracy in simulation results and to more confidently forecast and understand energy use in buildings during operational stages. Calibration methods mentioned in Reddy and Maor (2006), Raftery, et al. (2011), Coakley, et al. (2012), Bertagnolio, et al. (2012), Mustafaraj, et al. (2014) and Ji and Xu (2015) provide promising, procedural and replicable methodologies for calibrating energy models. These are similar evidence-based iterative methods applied to solitary case study buildings for varying calibration objectives. They show promise in terms of flexibility and reproducibility by incorporating some automated procedural extensions and can be used for generic building types along with more advanced and varying calibration objectives.

It is evident from the literature that all methods of calibration require some degree of approximation and simplification that can reduce the dimensionality of the problem. It is difficult (virtually impossible) to calibrate a building performance simulation model for all scales and types of outputs. Based on the needs of each specific calibration problem and the intended use of the calibrated model, specific techniques need to be adopted. There is no one-size-fits-all solution to create a calibrated model. Also, the calibration process requires an in-depth review, ideally over a prolonged period, of the intended (design stage) and actual (operational stage) working of the building to fine-tune the energy model. Therefore, a calibrated energy model can be used to understand the performance gap. Calibration can give insights into the operational inefficiencies and pinpoint underlying causes for the performance gap (Burman, 2016).

The actual energy use of most buildings is almost twice the design estimates, as reported in studies globally. The causes of the deviation can be mapped to various stages of building procurement, during design, construction and operation. The main causes of underperformance of buildings, reported in

[*] reference point for calculating the performance gap

existing studies, is due to shortcomings in construction practices, building fabric, control strategies, commissioning, building fine-tuning in early stages of post-occupancy, user training, building management and maintenance (Carbon Trust, 2012; Shrubsole, et al., 2018; Palmer, et al., 2016a).

The IEQ performance gap, often under-reported and under-analysed, can be seen in operational performance achieved for thermal comfort, IAQ, lighting and acoustics. The direct relationship between occupant well-being, comfort and productivity and IEQ in various building types has been widely documented (Wyon & Wargocki, 2013; Chatzidiakou, et al., 2014; Al Horr, et al., 2016) and has recently come to prominence as a key performance objective. Similar to the performance gap in energy use, the underperformance seen in IEQ parameters such as temperature, air quality (pollutants, CO_2), noise and lighting (Tuohy & Murphy, 2015; Fabbri & Tronchin, 2015; Phillips & Levin, 2015) can be linked to the various stages of building procurement. These factors include poor design stage specifications, construction practices, building fabric, commissioning issues and operation and maintenance shortcomings.

IEQ has close linkage and inter-dependency to energy performance and achieving high performance in IEQ as well as energy use in buildings can be convergent or divergent objectives. Overall, a holistic energy and environmental performance approach is necessary to understand the intricate interrelationship between these performance aspects to avoid unintended consequences and address gaps in performance.

Analysing all the existing research on performance gap studies on energy and IEQ and the use of calibration in operational performance assessment reveals many gaps in knowledge, which can be addressed through this work. Table 2.7 lists the key gap areas, mapped to the various literature review sections and study objectives.

Table 2.7: Gaps in knowledge, identified during the literature review.

Study Objectives	**Literature Review Sections**	**Knowledge Gaps**
Analyse operational performance of buildings	**1: Performance prediction approaches**	Most design energy calculations are conducted in the context of compliance modelling. Performance modelling using the CIBSE TM54 methodology makes simulation results more representative of actual performance at the design stage. However, as it is not focused on operational assessment, if applied to buildings in-use during M&V, it does not provide a method to procedurally identify and classify causes of the gap.

Develop calibration-based framework and to procedurally identify issues	**2: Simulation model calibration**	Using energy model calibration as a tool for performance gap assessment is a useful approach to identify, quantify and validate the performance issues in buildings. However, calibrating building simulation models to monitor data for this application is a new trend. M&V protocols such as ASHRAE Guideline 14 and IPMVP advocate the use of model calibration but focus on quantitative requirements and are not tied in a framework for procedural verification of all issues. It is likely that some key issues are identified during the investigations whilst other potential issues are not uncovered. Therefore, the development and use of systemised and replicable methods is highly recommended as a potential research area.
Explore model calibration and validation approaches	**3: Operational assessments and Calibrated model validation**	With the variability in the practical availability of operational data and its quality, depending on the data available and the use of the calibrated model, there is a need for a more robust model calibration and validation framework, that goes beyond the whole building statistical metrics of NMBE and C_V(RMSE), recommend by current M&V protocols. Calibration with high-resolution disaggregated data, supported with using IEQ data for cross-validation, is under-explored.
Determine the cross sectoral energy performance gap causes	**4: Energy performance gap and its causes**	The current performance gap and calibration studies are focused mostly on office buildings as case study examples. While these studies identified various issues, depending on the building type the relative importance of performance parameters varies. The issues highlighted in these studies may be because of the function of the building's, rather than an endemic issue in the construction industry. Therefore, analysing the root causes of the performance gap in various building types can provide an interesting insight into the common issues across the building stock.
Assess the unintended IEQ under-performance	**5: IEQ performance gap and its causes**	Performance gaps in IEQ in low energy low carbon design buildings have not been well-established. Sometimes, divergent ways of achieving low-energy and high-IEQ objectives, subsequent focus on energy only can lead to the unintended consequence of poor IEQ.

3 METHODOLOGY

This chapter reviews the research approach and the methods used in this thesis. In the context of building performance evaluations, it is evident from the earlier chapters that the current processes used for operational-stage assessments (such as M&V protocols) are not suitable for a systematic evaluation of performance gaps in a procedural and replicable manner. The existing processes focus on quantitative requirements, and they are not bound in a framework that can procedurally verify of all issues. Therefore, it is not certain that the technical issues uncovered in a building through, e.g., on-site investigations, reflect all or most of the key causes of the performance gap. It is likely that some key issues are identified during the investigations whilst other potential issues are not.

This chapter describes the development of a systematic calibration-based methodology to identify, quantify and validate performance gap issues. Within this evidence-based calibration process, advancements to calibrated model validation are assessed. To analyse the effectiveness of this systematic calibration-assisted method, a case study approach is used. Also, applying this approach to multiple typologies of case study buildings, lessons that may also be widely applicable, across the sector are identified.

This chapter first explains the development of the calibration-based methodology. The presentation of the method is divided into four key components that are linked to the various stages in the process proposed: data collection, comparison, modelling and analysis and lessons. This methodological overview is diagrammatically presented in Figure 3.1 and mapped to the study objectives in Figure 3.2. After reviewing each of the four parts of the methodology, the last part of the chapter discusses the application, benefits and limitations of assessing this method via a case study approach, creating a basis for the application of the findings from individual buildings in a cross-sectoral context.

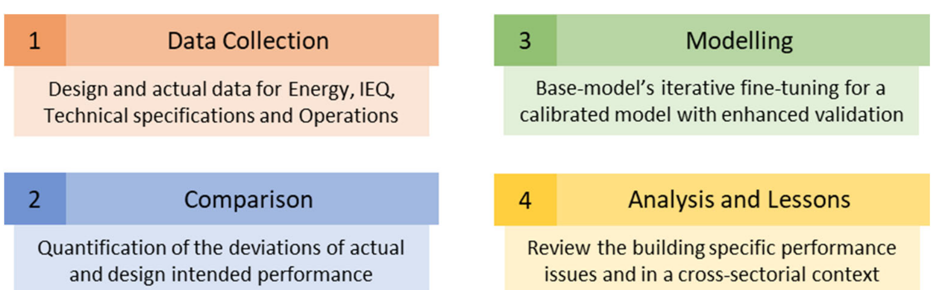


Figure 3.1: Methodological overview

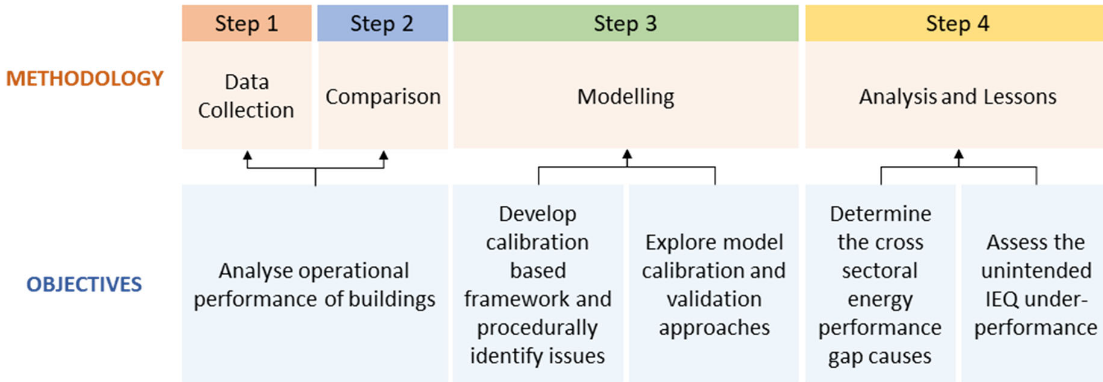


Figure 3.2: Links between the methodology components and the study objectives

3.1 Data collection

3.1.1 Data collection approach

Performance, in any system, is based on inputs, processes, outputs and outcomes, and the same is true in cases where the goal is to have high energy and environmental performance of buildings. During the design stage, information regarding each of these four components is based on assumptions and intentions, and as the project moves towards handover and operational stages, the assumptions and intentions become specific and real. Operational-stage performance assessments require a comparison of two stages: one is what the intention was (design data from the design stage) and the other is what happened in reality (actual data from the operational stage). Figure 3.3, adapted from Brown (2006), describes the design and actual data/information used/available under the four headings when the goal is the high energy and environmental performance of a building.

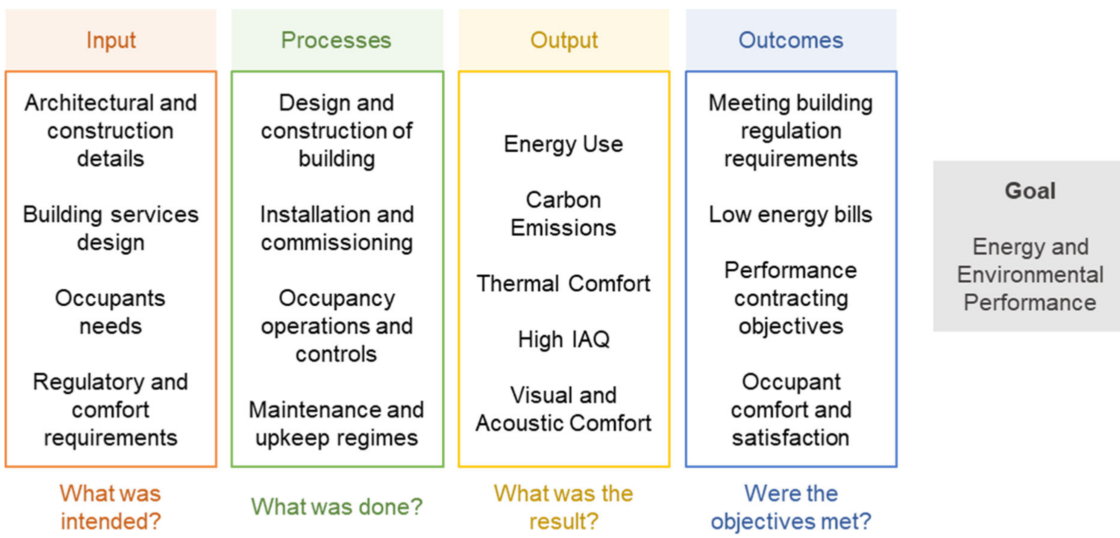


Figure 3.3: Inputs, processes, outputs and outcomes for energy and environmental performance of buildings

The data was collected under two headings, design data and actual data and is based on two factors:

1. *Scope of operational-stage performance assessment:* The objective of this exercise is to explore broad systemic issues and their root causes in building design, construction and operation, rather than in-depth, system-level investigation, fault detection and diagnostics. For energy use investigations, full building and sub-metered data is collected and for IEQ investigations, data collected is only for the parameters that are related to energy. Amongst the various IEQ parameters, thermal comfort and IAQ have complex and dynamic interactions with buildings energy end-uses and their performances have a high interrelation

with the energy performance objectives. Therefore, IEQ data collection is limited to thermal comfort and IAQ, as these two aspects are directly linked to energy modelling, occupant requirements and monitoring strategy.

2. *Availability and quality of information:* The level of information is determined according to the level of detail at which data is available from the building both at the design and operational stages. Design-stage information quality is determined by available documentation and access to stakeholders. Operational-stage information quality is determined by the level of metering and monitoring and access to a building and information from its facility managers.

This research work was aligned with a bigger project focusing on the total operational performance of buildings, EPSRC project (EP/N009703/1), The 'Total Performance' of Low Carbon Buildings in China and the UK ('TOP'). Therefore, all physical data collection for all the case studies was done under that project.

Following the scope and availability factors discussed above, the collected data has been processed and used for this work. Table 3.1 shows the key data that was collected for all the case studies and their sources for design data and actual data.

Table 3.1: Data collection overview

Data type	Use	Design Data	Actual Data
Architectural details and specifications	To understand building geometry, fabric and construction details for creating the simulation model.	Design drawings & documentation	Site visits, On-site Inspections
Building services details and specifications	To understand design and operational parameters for HVAC systems, lighting, electrical and other loads.	Design drawings & documentation	Operation and maintenance (O&M) manuals, building mangers, BMS
Occupancy details	To ascertain the occupant density and occupancy levels for learning operation schedules during the evaluation period.	Design teams and calculation reports	Site visits and discussions with building managers
Performance targets	To identify the building energy and IEQ performance targets and regulatory compliance benchmarks used.	Design stage calculation reports	NA
Hourly weather data	To accurately estimate the performance in the calibrated model.	NA	Actual weather data from the nearest station
Energy use	To estimate the actual energy use over one year at various levels of spatial and temporal disaggregation.	NA	Meter readings and Utility supplier
Indoor environment performance	To determine the quality of the indoor environment for thermal comfort and IAQ and for performance cross-validation.	NA	Regular IEQ monitors

Information gathered from the design data is necessary for the creation of the as-built model and understating of the design intentions. The actual data is used to calibrate the as-built model and to identify patterns of use and the performance of the building and its systems. The next two sections explain the design and actual data collected for the case studies, highlighting the secondary information used to infer the key data in case of missing information.

3.1.2 Design data

Design-stage information was collated from the design and construction documentation to understand the building procurement, design, construction, commissioning and handover process. Data was collected from the sources listed in Table 3.1. Architectural and building services drawings, sub-metering schematics, equipment inventory lists, commissioning data, energy and IEQ targets and standards, technical specifications of building systems and building operational parameters were collected.

However, in cases where initial estimate data was not available in detail, missing information was updated using measurements, surveys with stakeholders, or otherwise using building standards/codes or typical benchmarks. In cases where an atypical approach was used for calibration-baseline modelling, data is presented after modifications to ensure consistency across the case studies. These atypical instances and any modifications in data presentation are highlighted where relevant.

Design data collected for each of the case study buildings is described in Chapter 4 in Sections 4.1.1, 4.2.1, 4.3.1, 4.4.1, and 4.5.1, titled 'Design-stage details'.

3.1.3 Operational data

Regular measurements, observations and discussions with the facility managers at monthly or bi-monthly intervals over a period of one year were used to collect post-occupancy data and information. Actual data collected during site-visits recorded metered energy use and documented building technical and operational details. Actual IEQ performance was captured through regular monitoring of typical zones for temperature, RH, CO_2 concentrations (a proxy for fresh air), NO_2 (primarily a traffic-driven pollutant) and $PM_{2.5}$.

For all the buildings, energy use (gas and electricity) data was taken from utility meters and other meters linked to the BMS system. The data collected was disaggregated for different energy end uses. IEQ data for temperature, RH and CO_2 concentration were monitored in representative zones, covering 5-10% of the floor area with a frequency of at least 10 minutes for one year, in accordance with BS EN 15251:2007 (BSI, 2007), (measurement accuracies: T: ± 0.4°C, RH: ± 4.5 %, CO_2: ±75 ppm). Also, short-term intensive monitoring of parameters of thermal comfort and various air pollutants such as CO_2, $PM_{2.5}$, PM_{10} and NO_2 were also done, where data was recorded every minute by data loggers and sensors in the monitored spaces. Within the limitations of this study, data collected from IEQ parameters was either used for model cross-validation checks or to highlight energy-related unintended IEQ underperformance.

Obtaining disaggregated, regular and highly granular data for any building can be a particularly challenging task. While the coordination of extensive monitoring is always a factor, practical issues during site visits, such as access, regularity and data quality, are also encountered. The level of access and data availability was slightly different in each case study. Where there were shortcomings in directly monitored data, statistical post-processing was used to fill the gaps in the collected data using trends from site-level half-hourly energy and IEQ data. For example, to create monthly energy use profiles for monitored data at irregular intervals, proportional redistribution methods were used, adjusting for weekdays, weekends and holidays. These details are explained in detail in the relevant sections of the case study chapter for each of the buildings.

While monthly data has been used for calibrating energy models for all buildings and to assess the cross-sectoral impact of performance issues, it is recognised that monthly data can conceal inaccuracies at a higher temporal resolution, such as hourly data. For this reason, one of the buildings has been calibrated at an hourly level, and its data collection is at high granularity.

Operational data collected for each of the case study buildings is described in Chapter 4. Sections, 4.1.2, 4.2.2, 4.3.2, 4.4.2, and 4.5.2, titled Building energy performance contain the energy related data and IEQ related data is explained in Sections, 4.1.3, 4.2.3, 4.3.3, 4.4.3, and 4.5.3, titled Building IEQ performance.

3.2 COMPARISON

3.2.1 Data comparison approach

In the next stage of operational assessment, a comparison of collected data is undertaken for each of the four stages shown in Figure 3.4. The input and processes data collected in the context of operational performance assessments are individual data points and can be compared directly against each other. This comparison helps to highlight individual *deviations*. On the other hand, output and outcomes are computed information and therefore must be compared numerically for *quantification* of the differences.

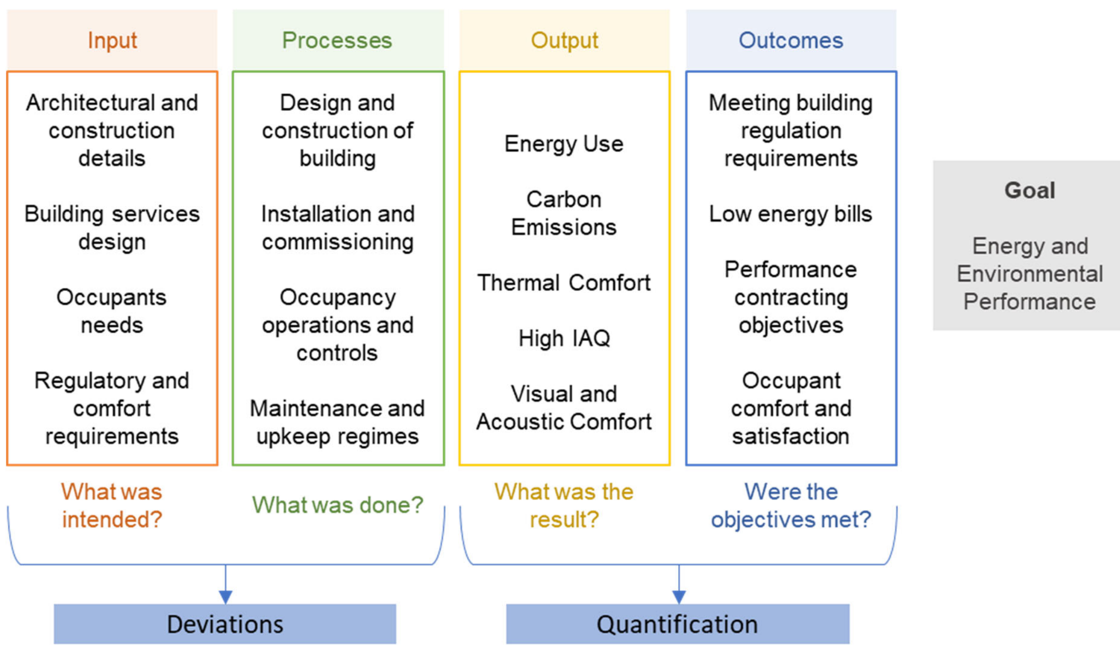


Figure 3.4: Comparing inputs vs processes and outputs vs outcomes

Comparing the technical specifications and operations of the building and systems against the intended ones gives insight into possible operational-stage deviations to help in the calibration process. Also, by applying this approach to the data collected for design and actual metered energy performance for each of the buildings, the comparison helps to quantify the magnitude of the performance gap. Figure 3.5 illustrates the comparison between design and actual data; the deviation and quantification thus generated are discussed in the next section.

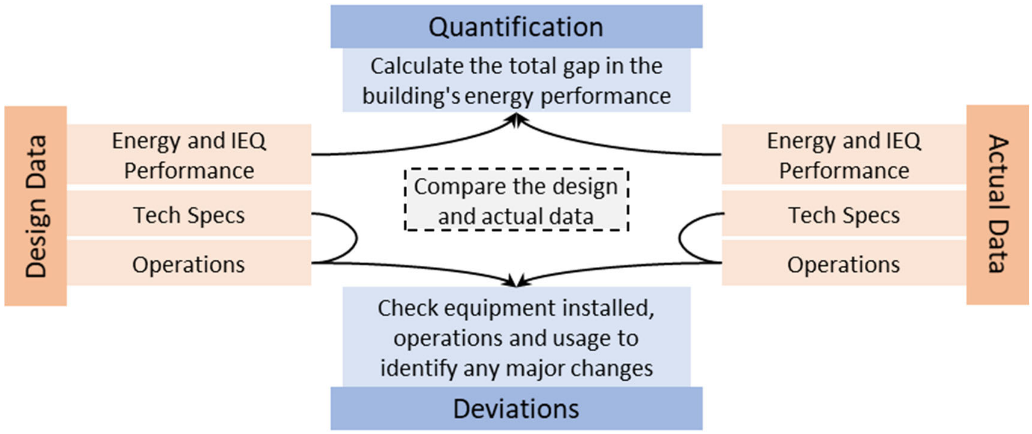


Figure 3.5: Use of data collected for comparative assessments

3.2.2 Deviations

In this comparison, technical specifications and operational parameters in design and actual data are checked against each other. Table 3.3 shows the comparison categories and relevant ways the deviations were checked.

Table 3.2: Identification of deviation in design and actual data

Category	Area	Method
Tech Specs	Building fabric (Walls, windows, roofs, floors, infiltration)	Review design documentation against any operation stage data for deviation, including non-intrusive on-site observations Thermographic surveys of buildings and in one of the building where U-value performance was very poor, advanced heat flux-based U-Value measurement was done for the walls.
	Lighting system	On-site verification of installed equipment against design specifications (non-intrusive, verification of models, nameplates and equipment rating).
	HVAC system	
	Installed equipment	
	Natural ventilation	On-site verification of the natural ventilation sources such as window openings and vents and checking of the (CO_2 based) automated vent controls.
Operations	Occupancy and schedule	Match the design intended operations against metered and monitored data, sub-metered load profiles, and data-trends from IEQ data and on-site observations. Also, identify events that are outside typical trends.
	Lighting and its controls	
	HVAC and its controls	
	Equipment and its controls	

Deviations identified in each of the case study buildings is provisionally described in Chapter 4, in the Sections, 4.1.4, 4.2.4, 4.3.4, 4.4.4, and 4.5.4, titled Observed deviation areas and issues and explained in detail in Chapter 5, Sections 5.1.2, 5.2.2, 5.3.2, 5.4.2, and 5.5.2, where they get confirmed during the calibration process (explained in Section 3.3.2)

3.2.3 Quantification

For each of the case study buildings, in Sections, 4.1.2, 4.2.2, 4.3.2, 4.4.2, and 4.5.2 the energy performance gap has been quantified. Disaggregated energy end-use estimated at the design stage has been compared with the current operational performance and good practice benchmarks as per CIBSE Guide F (CIBSE, 2012) at an annual level. For consistency in data presentation, the disaggregation of energy end-use has been first presented in six categories (heating, hot water, lighting, cooling, small power, auxiliaries). These comparisons are captured in Table 4.3, Table 4.5, Table 4.7, Table 4.9 and Table 4.11. The design stage energy projection in these tables are the ones which the design teams understood to be the projected energy use during operations. For most of the case studies it was the compliance modelling calculation results (see section 2.1.1) except for the public office and university office buildings. Public office building had design performance calculated by a performance modelling approach (see Section 2.1.2) whereas university office did not have that data available so quantification of the gap could not be done.

Monthly fuel-wise actual consumption has been presented in Figure 4.2, Figure 4.9, Figure 4.15, Figure 4.22, and Figure 4.27 for each of the case study buildings. Further disaggregation, where available, has also been presented separately.

3.2.4 IEQ comparison

Besides energy use, a comparison of IEQ parameters is undertaken for thermal comfort parameters (temperature and RH) and IAQ parameters (CO_2, $PM_{2.5}$, PM_{10} and NO_2). This is explained in Sections, 4.1.3, 4.2.3, 4.3.3, 4.4.3, and 4.5.3, titled Building IEQ performance. Table 3.3 lists the major requirements set out for the IEQ of various building categories against which the case study buildings' performance is being evaluated.

Table 3.3: IEQ requirements and their sources applicable to the case studies

<table>
<tr><th>Building category</th><th>Thermal Comfort</th><th>CO_2</th><th>$PM_{2.5}$</th><th>PM_{10}</th><th>NO_2</th></tr>
<tr><td>Office</td><td>CIBSE TM52 overheating criteria</td><td>1200 ppm
BS EN 16798
(BSI, 2019)</td><td rowspan="4">WHO annual mean limit
10 µg/m^3
(WHO, 2005)</td><td rowspan="4">WHO annual mean limit
20 µg/m^3
(WHO, 2005)</td><td rowspan="4">WHO annual mean limit
40 µg/m^3
(WHO, 2005)</td></tr>
<tr><td>School</td><td rowspan="3">Comfort temperature thresholds as described in BS EN 16798 (BSI, 2019)</td><td>1500 ppm
BB 101 Limit
(DfE, 2006)</td></tr>
<tr><td>Hospital</td><td>950 ppm
BS EN 16798
(BSI, 2019)</td></tr>
<tr><td>Apartment Block</td><td>1200 ppm
BS EN 16798
(BSI, 2019)</td></tr>
</table>

3.3 MODELLING

3.3.1 Modelling approach and scope

The scope of modelling proposed in this methodology is to create an *operationally accurate virtual representation* of the building intended at the design stage and the building under current operations. Therefore, the scope can be broadly divided into two parts

1. **Calibration-baseline model:** The dynamic thermal simulation model is built with design-stage assumptions, whose results represent the design stage's intended performance.
2. **Model calibration:** Fine-tuning of the building model inputs based on deviations observed in the data comparison stage, so that the end-state model outputs match the actual building performance, as per the validation criteria defined.

The buildings selected in this study had their own design stage calculations done by designers using dynamic thermal models, created either for initial design-stage projections or for building regulation compliance (the results of which were used in the quantified comparison whose methodology was described in the last section, Section 0). These models could have been further used as the calibration-baseline (as-design) models for subsequent calibration. However, there are two factors that suggest developing a new model as the best practice approach for the modelling part of this framework. First is the bias or mistakes which a modeller can bring to their models due to the approach used (Guyon, 1997), and second is the variability and uncertainty in calculation between different simulation tools (Schwartz & Raslan, 2013). Comparing results from models created with different approaches or software is therefore not justified. A new calibration-baseline (as-design) model is created using DesignBuilder V6.0 (DesignBuilder Software Ltd., 2019a), which is a graphical interface for the EnergyPlus simulation engine (DOE, 2019). As mentioned earlier, the scope of this exercise is to explore broad issues rather than in-depth system-level investigations; therefore, detailed component-level HVAC system modelling is not undertaken.

The model calibration method used in the framework is developed based on earlier work and is further extended to fit the specific research objectives of this study. As this study requires an in-depth review of root causes of the performance gap, an 'evidence-based' methodology was used (Raftery, et al., 2011; Coakley, et al., 2012; Bertagnolio, et al., 2012; Mustafaraj, et al., 2014). Other calibration approaches are not suitable for this application. A normative model approach, used in large stock models, is generic and does not provide enough modelling detail to capture building-specific interactions and issues. Also, automated approaches change many variables, and as the modeller does not know what change brought what impact, the approach can often mask the more crucial factors.

Evidence-based calibration clarifies the importance of different data sources as it supports the need for evidence when iterative changes are made to calibrate the model.

The last chapter proved that the currently used calibrated model validation procedures are not suitable. Developed for evaluating model calibration in the context of calculating savings due to ECM, the criteria are deterministic, failing to capture that some of the solutions that meet the criteria may not necessarily reflect the real performance. Improvements in the existing protocols require added cross-validation of other dependent outputs. An advanced model validation approach is also formulated as part of this framework. The manual evidence-based calibration, which is systematic, replicable and reliable in effectively identifying and estimating the impact of the underlying causes of discrepancies, was improved with advanced validation checks. In this section, we describe these aspects in detail.

3.3.2 Calibration-baseline modelling and model calibration

Error! Reference source not found. shows the process used to create the calibration-baseline (as-design) model and various steps proposed to develop the calibrated model. The first part of the modelling process is to use the collected building information to set up the inputs in the calibration-baseline (as-design) model, a virtual representation of the design intended building. An initial as-design model, with all design stage input parameter values, was created using DesignBuilder V6.0 (DesignBuilder Software Ltd., 2019a). For each of the case study building this as-design model creation is described in Sections 5.1.1, 5.2.1, 5.3.1, 5.4.1 and 5.5.1, titled Calibration-baseline (as-design) model, in Chapter 5.

Actual weather data, as per official station and satellite measurements for the relevant geographical location for the calibration period, was obtained from the DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b). This weather data was used in the simulations to create a weather corrected model. Using weather data from the nearest weather station for the same calendar year as the energy-use meter readings is important to ensure that there is a like for like comparison of simulated vs actual performance. At the end of this process, the model results are the intended design-stage performance of the building for the time under investigation.

Calibration is a process in which simulation models are corrected (by changing input assumptions) so that the results more closely match the actual performance. A typical calibration process compares simulation data against the metered data from the monitoring of the real building. As simulation data is composed of many inputs, therefore, depending on the objective of the calibration, the most influential input parameters affecting the specific output are fine-tuned. After the weather corrected model is complete, specifications of installed systems and equipment, controls, actual operations and occupancy, which are identified as modified in the initial comparison (as discussed in Section 3.2) are corrected to create an intermediate model. These deviations are explained in detail in Chapter 5, Sections 5.1.2, 5.2.2, 5.3.2, 5.4.2, and 5.5.2.

After that, an initial comparison of disaggregated metered and monitored monthly/hourly energy data is done against the corresponding simulation results of the intermediate model. This comparison finds whether the modifications explain the quantified performance gap. Basic calibration checks are undertaken using the ASHRAE Guideline 14 Calibration criteria check of NMBE and C_V(RMSE).

The next step is to improve the model calibration by iteratively reducing the deviation of simulated vs metered data and assessing which possible further measurements can be obtained. Analysis of disaggregated (energy end-use or spatial disaggregation) data trends at both system and subsystem levels is done. Fine-tuning is done iteratively until the calibration criteria are met. Calibration results after advanced validation (explained in next section, Section 3.3.3) for the case study buildings are shown in Section 5.1.3, 5.2.3, 5.3.3, 5.4.3 and 5.5.3.

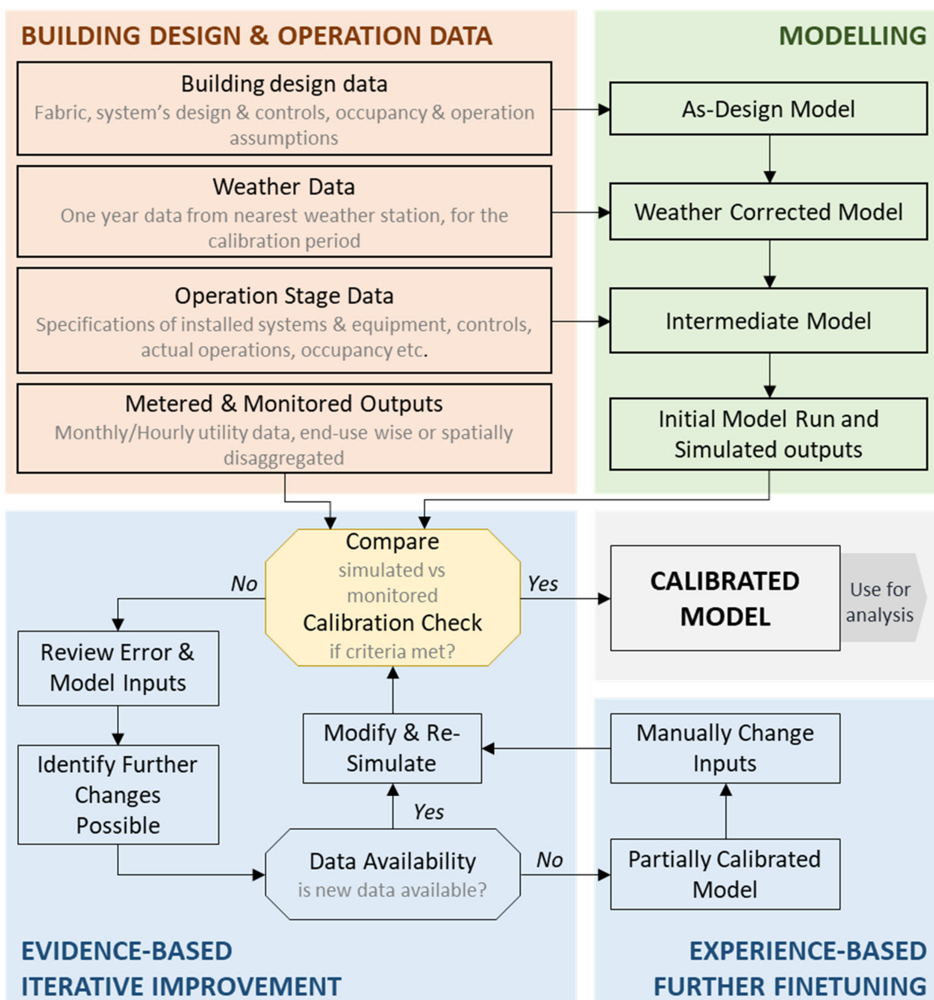


Figure 3.6: Method to undertake calibration

3.3.3 Advanced model validation

For all case study models cross-validation of model output is done by comparing air temperatures for typical weeks against measured values in a sample space. Additionally, for one building (University Office), a more detailed calibration framework is followed. The overall validation framework followed for all cases is based on these principles:

1. Statistical tests (NMBE, C_V(RMSE)) provide the basis of mathematical accuracy.
2. Cross-validation of dependent variables with a statistical and graphical approach ensures confidence in the correctness of underlying calculations and simulation input parameters.
3. Disaggregated checking reduces cross-compensation.
4. Probabilistic validation ascertains whether uncertain and yet unknown parameters can explain the residual gap between measured and predicted values.

Incorporating all the aspects mentioned above, a three-tier calibration framework is used in the study depending on data availability. The first tier looks at the spatial granularity of metered data, the second tier looks at the temporal granularity and the third tier assesses the level of cross-validation. Table 3.4 describes the framework in detail and Table 3.5 lists the statistical validation criteria for energy and temperature checks.

Table 3.4: Multi-tier calibration framework

Tier	Details			
Sub-metering granularity	**Level 1**		**Level 2**	
	Building-level totals checks: Annual calibration as per existing protocols for building-level totals for individual fuels.		End-use and spatially disaggregated data checks: Annual calibration for all major disaggregated energy end uses, especially separating the weather-dependent space-conditioning loads from other loads.	
Temporal granularity	**A**	**B**	**C**	**D**
	Monthly	Weekly	Daily	Hourly
Cross-validation detail	**NIL**	**SIM**	**INT**	**ADV**
	No cross-validation data were available.	Dependent variables typical weeks' checks.	Dependent variables check for typical weeks across seasons and different zone types.	Dependent variable checks for zone level IEQ parameters. Annual or peak season calibration for zone level space temperatures covering all zones with a different activity or HVAC system.

Table 3.5: Validation criteria for various checks

Energy calibration validation criteria	C_V(RMSE) (%)	NMBE (%)	Temp. calibration validation criteria	MAE	RMSE
A: Monthly	15%	±5%	**Hourly (full year)**	1°C	1.5°C
B: Weekly	15%	±5%			
C: Daily	18%	±6%			
D: Hourly	30%	±10%			

Sub-metering granularity

Level 1: This is the base level when only building-level energy data is available. Similar to the current standards, this level requires separate calibration for individual fuels. Calibration should be done for the building-level energy use for a period of one year.

Level 2: Increasing the requirements in Level 2, this calibration can be done where detailed metered data is available, disaggregated according to end uses and spatially. To comply with this calibration level, calibration should be done separately for weather-dependent space-conditioning loads (heating and cooling) and other loads. Also, as zone clusters of more than 500 m^2 in area and 25 kW loads could have separate meters (CEC, 2018), each of these sub-meters should be calibrated separately.

Temporal granularity and statistical compliance

To meet the calibration levels, for Level 1 and Level 2, the current standard criteria, which is based on statistical parameters of NMBE and C_V(RMSE), is needed to be complied with for each of the data streams. Derived from existing M&V protocols, Table 3.5 describes the limits for NMBE and C_V(RMSE) at A: Monthly, B: Weekly, C: Daily and D: Hourly temporal resolutions for energy use.

Calibrations to the hourly and daily billing data generally produced more accurate results than calibrations to monthly data (Robertson, et al., 2013). Monthly data analyses can easily miss significant errors at a daily or hourly resolution; however, obtaining high-resolution hourly data is not always possible (Raftery, et al., 2011). Therefore, intermediary resolution calibration at weekly and daily levels can provide useful insights and a better calibrated model than monthly checks. The limits for NMBE and C_V(RMSE) in Table 3.5 have been obtained by interpolating monthly and hourly criteria.

Cross-validation detail

This is the final tier of validation where dependent variable data is available, such as zone level temperatures (IEQ parameters) and loads. To meet this tier of compliance, cross-validation checks should be done for at least zone level space temperatures. The zones selected for these checks should cover all spaces with a different activity or HVAC system; these should account for a total of 10% of the building's floor area (BSI, 2012; BSI, 2019). These monitoring tests need to be at hourly timescales (ADV) for a full year. However, if a full year is not possible, the checks should cover at least two consecutive weeks in the year (SIM), or two weeks in summer, winter and other transitional seasons (INT). The limits for MBE and RMSE for hourly temperature checks are described in Table 3.5.

Application in the case studies

Depending on the application of the calibrated model and monitoring data availability, a particular level of the framework and compliance criteria can be used (the lowest and highest level will be Level 1A-NIL and Level 2D-ADV respectively). For all case study buildings, a Level 1A-SIM model is created. In the case of the university office, the Level 1A-SIM model initially created is further calibrated to the most detailed calibration level, Level 2D-ADV.

Additionally, to determine the confidence in the calibrated model or if the statistical validation criteria cannot be met, a probabilistic validation method can be used. This has been done for the university office building in the form of scenario analysis. Probabilistic validation can be done in cases where the uncertainties of key parameters that explain the deviations have been clearly defined and refined. For the university office building, to create the scenarios for probabilistic validation, standard uncertainty ranges outlined in existing literature (MacDonald, 2002; BSI, 2008) have been used as the baseline with modifications as per on-site observations and information gathered from stakeholders. Using these uncertainty ranges, high, medium and low scenarios are calculated and analysed.

Through this multi-level advanced validation process across the case studies, with the objective of having systematic model calibration checks, improvements in the criteria and lessons relating to model calibration and validation were identified.

3.4 ANALYSIS AND LESSONS

The identified changes between design-stage and actual data in the building, its systems and overall operations in Section 3.2 (comparison stage) were verified during the site investigations. Subsequently, those deviations that had led to the creation of the calibrated model in Section 3.3 (modelling stage) are assumed to be the key factors causing the gap. Table 5.1, Table 5.2, Table 5.3, Table 5.4, and Table 5.5 in Section 5.6.2: Summary of deviations (issues) list the key areas that were the causes of the performance gap, confirmed through the calibration process. Figure 3.7 shows how the methodology has evolved until now, the data collection stage feeding into the comparison stage feeding into the modelling stage. The next stage is to analyse this process on the case study buildings and deduce the lessons related to the specific study objectives.

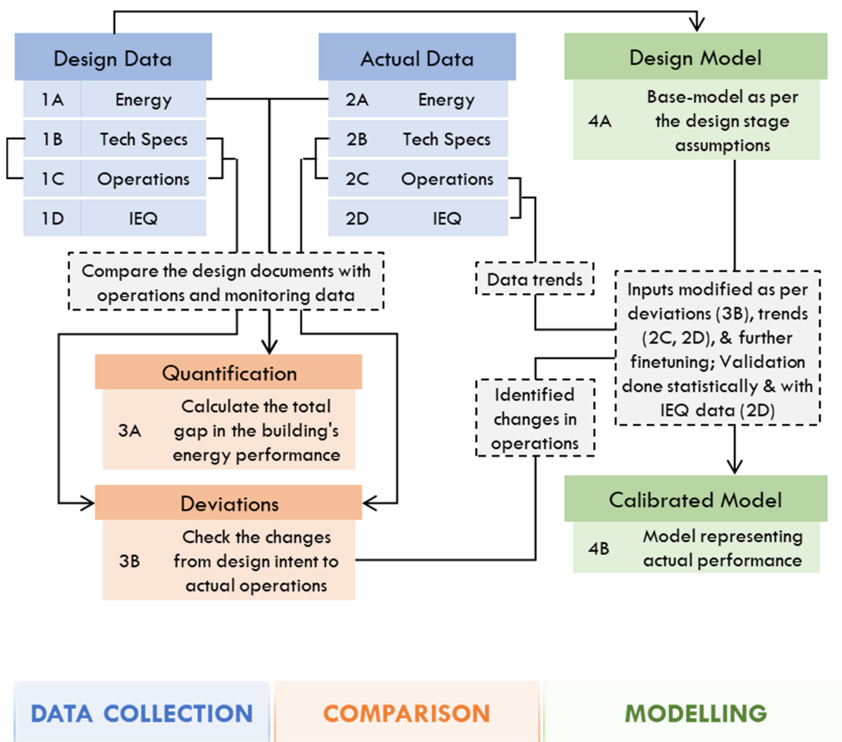


Figure 3.7: Relationship of the first three components of the methodology, data collection, comparison and modelling

3.4.1 Analysis

In the final analysis stage, to understand the perceived gap, first, the classification shown in Figure 2.2 is used to assess each case study building's performance. This compares the results of compliance calculation from compliance models from those of the design performance from the performance model. Second, an operational-baseline is created and the gap due to technical issues is calculated by reverting the technical changes in the calibrated model to their design intents. Figure 3.7 explains the *operational-baseline* and the two types of gaps that constitute the *actual gap*.

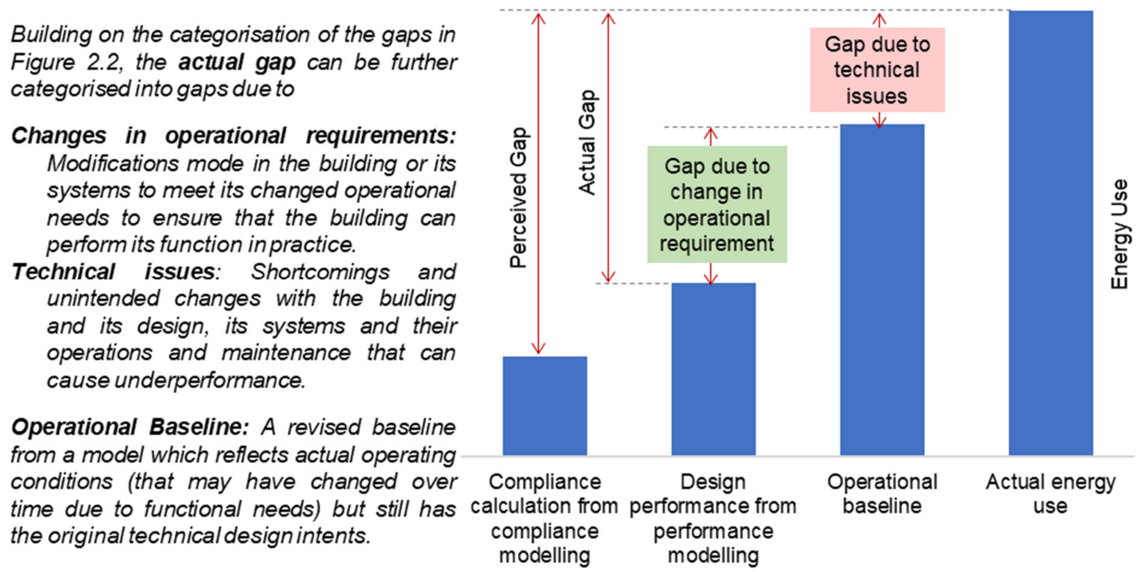


Figure 3.8: Building on Figure 2.2, new operational-baseline based on the actual gap categorisations

The calibrated model created earlier closely represents the actual building. Therefore, to ascertain the gap due to changes in operational requirements and the gap due to technical issues, the following steps were taken:

1. All the deviations (and their potential causes) concerning the building's use, operation settings, systems' functioning and controls were divided into discrepancies that were due to change in operations and those due to technical issues.
2. Deviations relating to technical issues were reverted to their design intents. This new simulation model provides the current operational-baseline, thereby quantifying the performance gap due to technical issues. This revised baseline[*] reflects the actual operating conditions, such as space-time use of the building and its systems, currently followed by the building to perform its functions.
3. Comparing the actual energy use with this operational-baseline determines the actual energy performance gap arising from technical issues, whereas comparing the operational-baseline with the original simulation results shows the effect of the changes in the operational requirements of the building.

During the performance gap analysis it is possible that the deviations identified can be related to a combination of both, operational changes and technical issues. In those instances the deviation should be proportionally attributed to both of them (based on engineering judgement and discussions with

[*] reference point for calculating the performance gap

the stakeholders such as building managers). An example of this can be instances where changes in operations of a system, which already has technical issues. Another case of such straddling of issues might occur due to system performance degradation over time coupled with technical problems. This is typically a case in older buildings and in those instances the proportional attribution would determine the longitudinal degradation as an operation change and rest of the deviation as technical issue because the operational baseline is the most optimum (technical) performance given the current state of the building, its systems and its operational requirements.

Section 7.1: Categorisation of the causes of the performance gap separates the technical issues from operational changes in the case study buildings and 7.2 Operational-baseline and the actual gap explains the various performance gaps based on their root causes in these case study buildings. Applying this approach to all the case study buildings, common themes for performance issues were used to determine cross-sectoral problems (described in Section 7.3 for the cases studied here). Also, findings from IEQ data were compared with intended performance and building standards and potential underperformance issues and conflicts related to energy objectives were identified (described in Section 7.4 for the cases studied here). This way this method applies to the case studies is explained in Section 3.4.2 below.

3.4.2 Lessons learnt

A case study approach is used in this research to meet its key objectives, which are to validate the effectiveness and robustness of the proposed framework and then use the findings from multiple cases to infer and develop generalised theories. In this section, the applicability of the case study approach is first discussed; then, in the second part, inferences that could be drawn after the application of this framework on multiple case studies are explained.

Justification of a case study approach

A case study approach is often stereotyped as a weak, imprecise and insufficiently rigorous method (Yin, 2014). However, the approach is extensively used in various disciplines to analyse a phenomenon, generate hypotheses or validate a method (Teegavarapu & Summers, 2008), including building performance evaluations. It is one of the most used methods in the research on building energy efficiency assessments published in leading building energy journals (Addy, et al., 2014). The reason for this is explained by Yin's (2014) definition of the case study approach as: "*an empirical method that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident*". In operational-stage building performance assessments, the buildings under investigation are well integrated into real-life operations. Evaluating their performance holistically (the phenomenon) requires them to continue to operate in the complex and intertwined processes (the context).

Other competing/complementary approaches for operational-stage building performance assessments that are often considered more structured and rigorous are controlled experiments and surveys. However, by themselves, these approaches are insufficient in building performance assessments. In building performance assessments, it is not possible to control all behavioural events, making controlled experiments difficult. Also, these assessments require answering, among others, 'why' types of questions, which are difficult to do through surveys. The case study approach can deal with these situations where the variables under analysis are more than the data points, by giving the observer an opportunity to collect data using different techniques (triangulation), such as a survey, an interview, or an experiment, all under a single study. A robust structured plan incorporating the various techniques can address a key issue that case study researchers have to deal with, which is that case studies do not have a strict and systematic approach. Case studies can lack academic rigor and can be influenced by the subjectivity of the researcher (Yin, 2014).

Another concern around the case study approach, which is valid in the context of building performance assessments, is that one cannot generalise from a single or a set of cases (Yin, 2014). However, case studies are analytical investigations, similar to scientific experiments, and do not aim for statistical generalisation. The main aim of a case study is not to propose a theory that is universally valid but to arrive at a theory that is valid for a specific context. Case studies, being in-depth assessments, are not required to follow sample logic. There could be a single case from which a generalisation is made, the main factor being the appropriateness of the case and the way the case is chosen (George & Bennett, 2005). Multiple cases could be analysed under a single case study, referred to as a multiple-case study, in which selection of cases is based on repetition logic, rather than sample logic (Yin, 2014). In view of this, the five case study buildings selected in this cross-sectoral examination are from four distinct categories: offices, schools, hospitals and apartment blocks. Offices, schools and hospitals account for around 46% of the England and Wales non-domestic building stock, and 37% of energy consumption (BEIS, 2016). Apartments account for 21% of the England and Wales household spaces (ONS, 2011). These buildings represent a large proportion of the UK building stock. Within each category, the buildings selected were newly built and had good accessibility and data availability, in line with the aims and objectives of this research.

The case study approach uses multiple methods in a small number of samples for analytic generalisation within the context of the problem, with minimal intrusion. The limitations of the case study approach, such as a lack of representativeness and generalisability (discussed earlier) and inability to quantify the average effects or to eliminate selection bias, can be addressed by meeting the criteria of general soundness, which includes construct validity, internal validity, external validity and reliability (Yin, 2014). In the current study,

- Construct validity is ensured by the use of multiple case studies, getting structured information from multiple sources and following evidence-based processes.
- Internal validity is ensured by analytic generalisation across multiple cases considering all evidence, addressing rival explanations and cross-validating the findings from various sources.
- External validity of the theory is checked by referring to other similar research conducted in the field as well as replication logic across cases.
- Reliability is ensured by using the same framework across multiple studies.

Drawing inferences from case studies

Extending the process followed in Figure 3.7 and Figure 3.8 for individual buildings across the case studies, lessons could be drawn that relate to the main focus areas of this research. Figure 3.9 shows how various stages of the proposed methodology, from Data Collection, Comparison and Modelling to Analysis, link to the key lessons.

1. **5A: Calibration Process Lessons** – Reviewing the challenges in undertaking conventional as well as advanced model calibration and validation, lessons for better calibration are concluded (Summarised in Section 8.4).
2. **6D: Cross-sectoral Lessons** – Reviewing the performance issues found across the case studies, common themes and lessons that may apply to the wider industry are collated (Summarised in Sections 8.2 and 8.5).
3. **7A: Unintended IEQ Issues** – Through the assessment of the design targets and actual IEQ performance, underperformance issues are reported and root causes that can be linked to energy efficiency objectives are identified across the case studies (Summarised in Sections 8.3 and 8.6).

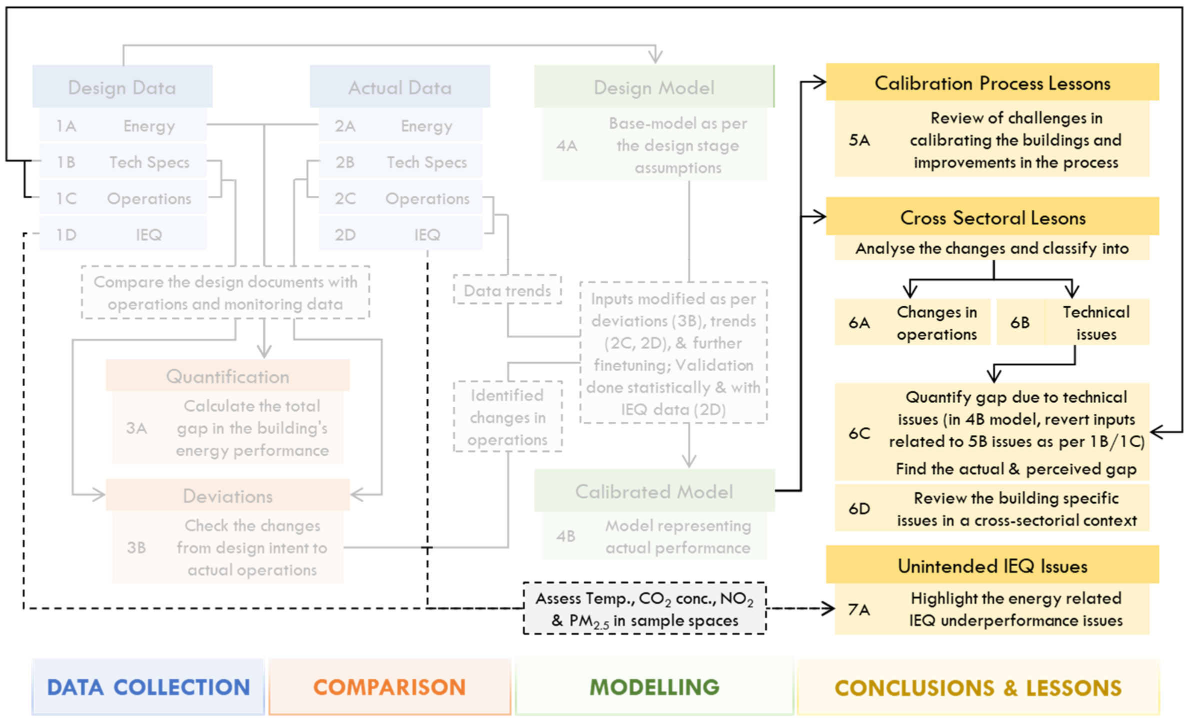


Figure 3.9: How the components of the methodology link with the key conclusions and lessons

3.5 SUMMARY

This study uses a systematic calibration-assisted method to find and validate building-level performance issues, finding and separating deviations related to operating conditions and technical issues. Applying this process across the case study buildings provides lessons that may be applicable across the building sector, along with an understanding of model calibration issues and unintended consequences on IEQ. The overall method is diagrammatically presented in Figure 3.10.

First, the building's design intents and actual data were collected. Design data was collected from the design documents having details about the energy (1A) and IEQ (1D) targets, technical specifications of the building systems (1B) and building operational parameters (1C). Actual data was collected during site visits, recording metered energy use (2A) and documenting building technical and operational details (2B/2C). The actual IEQ performance (2D) was captured through regular monitoring of typical zones for temperature, CO_2 concentrations (a proxy for fresh air), NO_2 (primarily a traffic-driven pollutant) and $PM_{2.5}$.

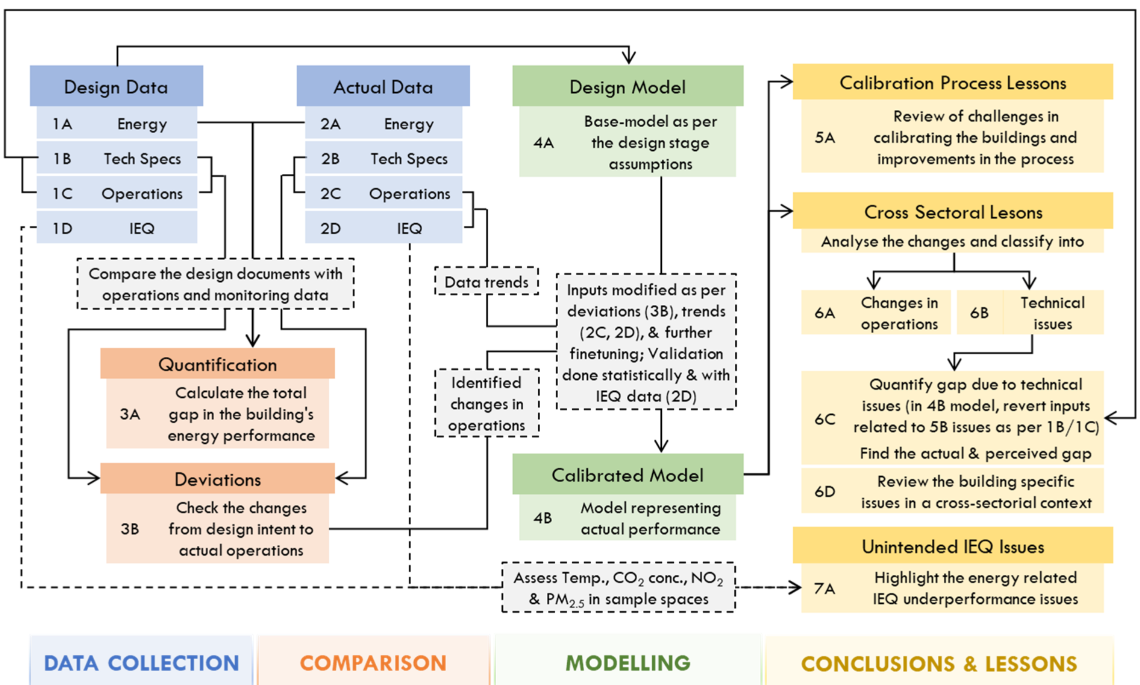


Figure 3.10: Summary of the method for performance gap assessment using calibrated energy simulation

Then, design-intended energy performance was compared with actual measured monthly energy use to quantify the magnitude of the performance gap (3A). Also, comparing the technical specifications and operations of the building and systems gives insight into possible operational-stage deviations (3B).

In the next stage, an initial design model (4A) was created using DesignBuilder V6.0 (DesignBuilder Software Ltd., 2019a), based on design data. This design model was then changed per actual operations (2C) and IEQ measurements (2D), and operational-stage deviations (3B). Further model calibration was undertaken by evidence-based manual fine-tuning of various input parameters. Actual weather data as per official station and satellite measurements for relevant geographical locations for the calibration period were obtained from the DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b) and were used in the simulations. The resultant calibrated model (4B) was validated using the monthly calibration criteria of ASHRAE Guideline 14 and IPMVP (C_V(RMSE) < 15%; NMBE < ±5%). Also, temperatures for typical weeks were compared against measured values (2D) in a sample space. The classification shown in Figure 2.2 is used to assess each of the case study buildings' energy performance issues.

Once a calibrated model was created, general lessons relating to model calibration and validation of results were found (5A). Regarding building performance issues, all the deviations (and their potential causes) in relation to building use, operation settings, systems' functioning and controls were divided into discrepancies that were due to a change in *operational requirement* (6A) and those due to *technical issues* (6B).

Subsequently, to quantify the performance *gap due to technical issues* (6C), deviations relating to *technical issues* (6B) were reverted to their design intents (1B/1C) to create an *operational-baseline*. This revised baseline* reflected the actual operating conditions (such as space-time use of the building and its systems) that are needed for the building to perform its functions in practice but assumed that the original design-intended technical specification are met. Comparing the *actual energy use* with this baseline* determines the *actual energy performance gap arising from technical issues*.

After analysing all four case study buildings, common performance issues were used to suggest some remedies that can affect a large cross-section of the building stock (6D). Also, findings from IEQ data showing potential underperformance issues and conflicts related to energy objectives were reported (7A).

* reference point for calculating the performance gap

4 CASE STUDY BUILDINGS AND OPERATIONAL PERFORMANCE

This chapter describes the five case study buildings in four categories: offices, schools, hospitals and apartment blocks. The buildings include two offices and one recently built building from each of the other categories. These buildings represent a large proportion of the UK building stock (as discussed earlier) and can provide useful insights into the endemic issues in the construction sector that drive the performance gap.

First, Table 4.1 describes key information about these buildings; subsequent sections describe the detailed building characteristics, design vs actual energy use, and the performance gaps in Energy and IEQ. To systematically understand the case study buildings, their details have been presented in four sections:

1. Design-stage details: Explains the building's architectural design features, fabric, HVAC system, lighting, operations and metering.
2. Energy performance: Details the current energy performance of the building and compares it with design data to quantify the magnitude of the energy performance gap.
3. IEQ performance: Examines the temperature and IAQ parameter performance in typical spaces and identifies any unintended energy-related unintended consequences.
4. Performance deviation areas: In the energy and IEQ performance section, areas of deviation between design intent and actual operations seen during data collection have been discussed. Issues found in the case study building, during the operational performance assessment stage have been separately listed. In Chapter 5 these issues inform model calibration and, if applicable, they are validated as some of the likely causes of the performance gap.

Table 4.1: Key information about the case studies

Building	Location	Useful Floor Area (m^2)	Date Completed	Remarks
Public Office	Somerset (Keynsham)	6,363	2014	• Open-plan office with meeting rooms; four floors with narrow floor plates which are inter-connected by atriums and cut-outs • Project procured under an Energy Performance Contract and within a Soft-Landings framework (Way, et al., 2009)
School	London (Wandsworth)	21,405	2014	• Secondary school and sixth form with academy status • Eight campus buildings, four floors each • Performance analysis at the campus level with a detailed assessment of one building

Hospital	Bristol (City Centre)	16,122	2015	• Patient ward building within a large hospital campus; ‘Acute’ type hospital (as defined in ECG72 (BRECSU, 1996)) • Nine floors; include wards, consulting rooms, offices, diagnostics, operating theatres, a canteen, and the usual amenity rooms.
Apartment Block	London (Tower Hamlets)	7,940	2015	• Block with 98 flats and maisonettes across two buildings. • Performance analysis at the block level with a detailed assessment of a typical flat
University Office	London (Camden)	4,500	1900’s (Refurbished in 2010)	• University campus building with a basement, ground, and six upper floors • Primarily consists of open plan and cellular office spaces, meeting rooms, a library and computer clusters

4.1 PUBLIC OFFICE

The four-storey office building is approximately 6,500 m^2 and is located in Keynsham. It is a new building completed in 2014, designed primarily to house open-plan offices with meeting rooms. To ensure that the project teams are responsible for the operational performance of the building, it was built under an energy performance contract. Initially, a conceptual design was developed for the tender stage. The main contractor was then appointed to develop a detailed design and execute as per the concept, while ensuring the energy performance targets were met. The contractor and their designers developed the detailed design and the concept design team was engaged to review the process.

Figure 4.1: External view of the public office building

4.1.1 Design-stage details

Architectural design

The fabric of the building was designed to be highly insulated, and the architectural design promoted passive design. Narrow floor plates, connected by atriums and cut-outs, created an interconnected and open environment, provided deep natural light penetration and enhanced natural ventilation by creating a stack effect. Fabric U-values (W/m^2K) were as follows: Wall: 0.20; Window: 1.40; Roof: 0.15; Ground: 0.15; and design airtightness was 5 $m^3/hr/m^2$ @ 50 Pa.

HVAC systems design

Heating was designed to be primarily provided by heat pumps that can produce simultaneous heating and cooling and was distributed to the building via heating and cooling buffer vessels. The heat pumps were also designed to satisfy the cooling needs of the IT server room. But the heat pumps only operated if there was a heating demand in the building. In the absence of heat demand, a free cooling chiller satisfied the cooling needs, via a chilled water buffer vessel. When the amount of heat produced by the heat pumps was insufficient for the building heat load, in the event of low external temperatures, then the additional heat needed was to be provided by modular condensing gas-fired boilers which were designed to meet peak loads and give full back-up. When working in heat transfer mode, the heat pumps had a maximum combined COP (40% cooling / 60% heating) of 6.5. The energy efficiency ratio (EER) (cooling mode, full load) was 2.75 and Heating COP (full load) was 2.31. The boiler seasonal efficiency was 95.6%.

There were two heating loops planned, one constant temperature loop that ran at a fixed flow temperature of 45°C and weather-compensated variable temperature loop that ran at a maximum of 65°C during boost time. Underfloor radiant heating was provided in circulation and communal areas and trench heaters supplied heating to all offices and meeting rooms. Cooling for space conditioning, only provided for meeting rooms, was supplied by chilled beams.

Natural ventilation was the primary form of ventilation and was provided by vents which were controlled via a BMS, based on CO_2 concentration and air temperature. A night-cooling strategy was specified to keep the open-plan offices cool in summer. Manually openable vents were also provided. Toilets and other enclosed occupied spaces had dedicated mechanical exhausts.

Lighting and electrical systems design

The building, designed to be daylit, used low-energy artificial lights. In open-plan office areas, general background lighting was provided via energy-efficient light emitting diode (LED), and T16 lamps defined the circulation routes. Additional freestanding up-and-down-lighters with compact fluorescent lamps (CFLs) were provided for the desks. Lights were dimmable and were controlled by passive infra-red (PIR) and daylight sensors. The building was planned to have low-energy equipment and thin client computers. Among other loads, the building was to have a constant server load and loads for catering, lifts, actuators and CCTV.

Building and services operations

The building was assumed to be occupied during the weekdays, and all services were intended to operate during occupied times. A centralised BMS was to control all operations, and manual overrides and controls were also provided. Table 4.2 lists the design intended characteristics of the building for occupancy, services and operations.

Table 4.2: Public office occupancy, services, operations and loads details at the design stage

Category	Details
Occupancy	Number: 455 persons Weekdays: 0700 -1900 with diversity and some out-of-hours occupancy; Weekends: zero.
Heating	Operation: 07:00 to 20:00 (warmup at 06:00) during weekdays; Off during weekends Set point: 19°C
Cooling	Operation: 10:00 to 17:00 during weekdays; Off during weekends Set point: 23°C
Auxiliaries; Mech. Vent.	Load: 0.5 W/m^2 (Pumps: 2.5 kW; Fans: 0.7 kW) Operation: 09:00 - 19:00 during weekdays; Off during weekends
Lighting	Loads: 5 W/m^2 – daylight integrated Operation: 4 hours/day during occupied times during weekdays; Off during weekends
Small Power	Loads: Server Peak 29 kW; Standby 15 kW; Other small power 20 W/m^2 Operations: Server Peak: 09:00-19:00 during weekdays; Standby during weekends; All other loads follow occupancy.

Metering strategy

There were separate meters for all systems and end-uses to record the disaggregated energy use in high-resolution. Separate meters were provided for heating, cooling, hot water, lighting, small power, servers (electrical and cooling), pumps, vents and lifts. All uses were broken down per floor and per zone except where the total end-use is less than 0.5 kW. The meters were designed to be integrated to the BMS.

Other design information

As a unique case of performance contracting, the building was designed to achieve a DEC-A rating and net annual emissions of less than 18.78 $kgCO_2/m^2$. To achieve DEC-A, a building's emissions need to be 75% less than the DEC typical office (DCLG, 2015). The DEC typical office's net annual emissions were 75.12 $kgCO_2/m^2$. To ensure a DEC-A rating and to aid a smooth transition from design to construction, a risk matrix was created at the concept stage. The main risks identified were value engineering, controls' optimisation, user behaviour and small power loads. Technical compliance parameters for mechanical and electrical systems were also pre-defined at the concept design stage, to be followed during detailed design development. Performance contracting integrated within initiative Soft Landings Framework (Way, et al., 2009) ensured the extended involvement of the design and contracting team (up to three years), so that the performance objectives were met.

Unlike energy, there is no evidence of *specific* operational IEQ performance objectives in the performance contract. However, design and construction documents (like other projects) do show the requirement to achieve an acceptable indoor environment as per BS EN 15251:2007 (BSI, 2007), along with Part L (DCLG, 2013) and Part F (DCLG, 2010) of UK Building Regulations, thereby providing design stage overheating, ventilation, lighting and acoustic targets.

4.1.2 Building energy performance

The actual performance from November 2015 to October 2016, two years after building handover and following initial fine-tuning, although encouraging, showed that the building was not at that point operating at a DEC-A level. The contractor and the FM were still working to optimise the performance under the Soft Landings Framework. Table 4.3 shows the design and actual energy use disaggregated at the annual level, and Figure 4.2 shows the actual monthly gas and electricity use patterns. The design calculations were done during concept design to determine the performance targets and went beyond the restrictions of regulatory compliance calculations. The actual data was collected from the sub-meters installed on-site. Actual gas use for heating and hot water was metered together, and actual cooling was reported in the small power electricity meter.

Table 4.3 Annual design vs actual energy use comparison for the public office

Annual design-vs-actual energy use				**Good Practice Benchmark****
Fuel	**End Use**	**Design**	**Actual**	
Gas	Heating	13.9	28.9	
	Hot Water	5.0		
	Total	**18.9**	**28.9**	**79.0**
Electricity	Lighting	6.1	11.1	
	Cooling	0.2	31.2	
	Small Power	22.6		
	Server	26.4	15.2	
	Fans/Pumps	1.7	10.0	
	Total	**57.0**	**67.5**	**54.0**

*All values are in kWh/m^2/yr.; ** CIBSE Guide F benchmark for Offices: naturally ventilated, open plan (CIBSE, 2012)

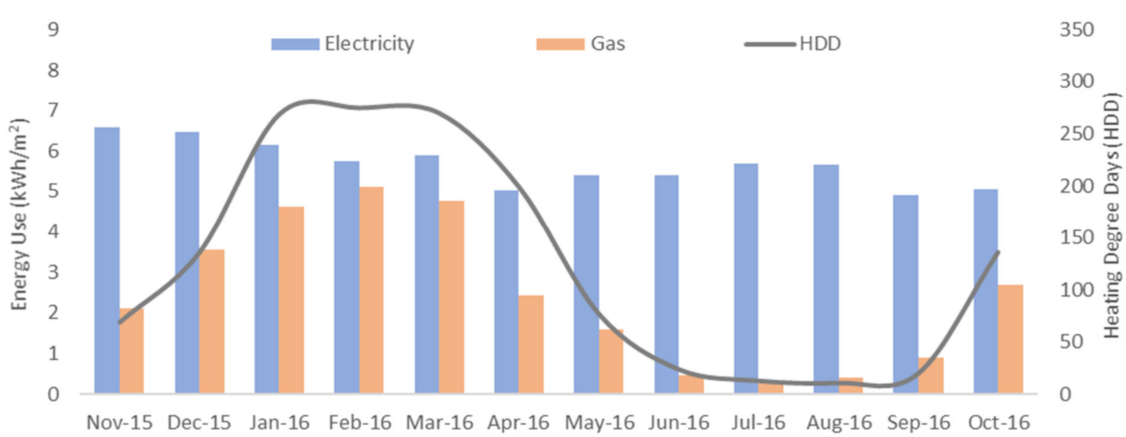


Figure 4.2: Monthly electricity use and gas use for the public office building

Table 4.3 shows that there are discrepancies in designed and actual performance for almost all end-uses. Initial observations found that the 50% increase in gas use was because the heat pumps, designed to use rejected heat from the servers to heat the building, had technical issues which caused them to malfunction. As a result, the gas-fired boilers supplied all the heating. Design stage under-predictions were also seen in fans, lighting and small power use, which can be due to a 20-30% increased occupancy observed in the building. However, the actual electricity used by the servers was less than the design estimate due to the overestimation of the server load at the design stage.

Based on the actual performance from November 2015 to October 2016, the building was achieving a DEC-B rating. The net carbon emissions were at 25.12 $kgCO_2/m^2$ which is 50% more than the DEC-A target of 18.78 $kgCO_2/m^2$. However, it should be noted that comparing the current performance with the very stringent target shows a magnified level of underperformance for an otherwise, contextually, well-performing building. Due to the stringency of the targets and the use of performance contracting integrated within the Soft Landings Framework, along with committed engagement of the design team, the contractors and the client, since the project's inception, ensured that performance targets were kept in sight. The building's actual performance compared favourably against the good practice benchmark (CIBSE, 2012) for naturally ventilated public office buildings in the UK.

4.1.3 Building IEQ performance

In the office building, the parameters for thermal comfort (Temperature and RH) and various air pollutants (CO_2, $PM_{2.5}$, PM_{10} and NO_2) were recorded in four typical locations: open offices on the first and third floors, meeting room on the second floor, and external space (near the building close to the main road). This section reports the findings along with the relevant limits for comfort and wellbeing. The results are presented in box plots for each parameter, with more detailed graphs where unique trends or serious issues are seen.

Thermal comfort

The building has space heating provision but no comfort cooling; in a naturally ventilated building without comfort cooling, summer overheating is a concern. Figure 4.3 shows that there were no significant thermal comfort issues in the building. The building mostly maintained a comfortable indoor temperature and RH in most spaces during both non-heating and heating seasons. Indoor temperatures in the monitored spaces were kept between 21°C and 24°C during the non-heating season, and about 2°C higher during the heating season. However, there were a few instances where indoor temperatures were recorded in excess of 28°C; subsequently, a detailed analysis was done to understand the trends. Figure 4.4 and Figure 4.5 show air temperatures during a typical summer week in August and during a hot spell in July, respectively.

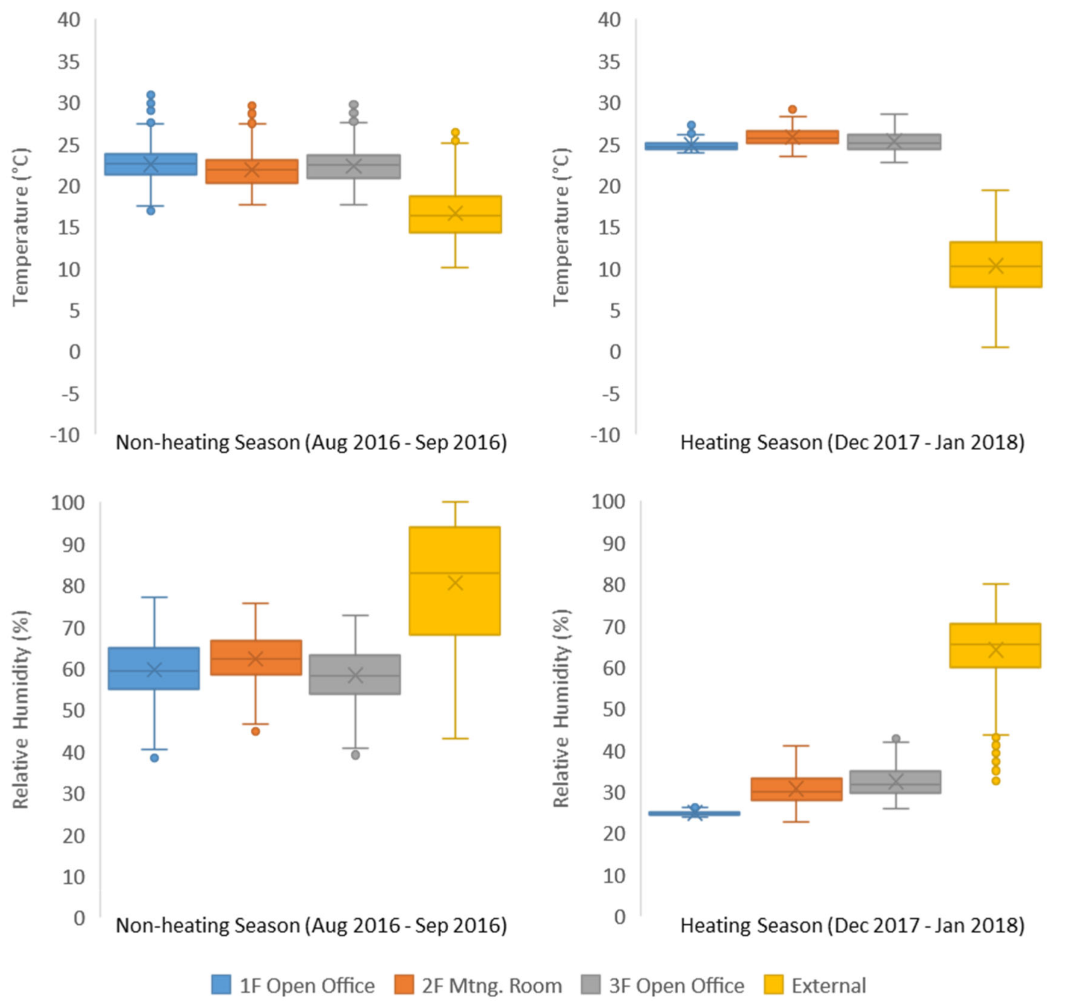


Figure 4.3: Box plots showing the spread of Temperature and RH during non-heating and heating seasons in the public office building

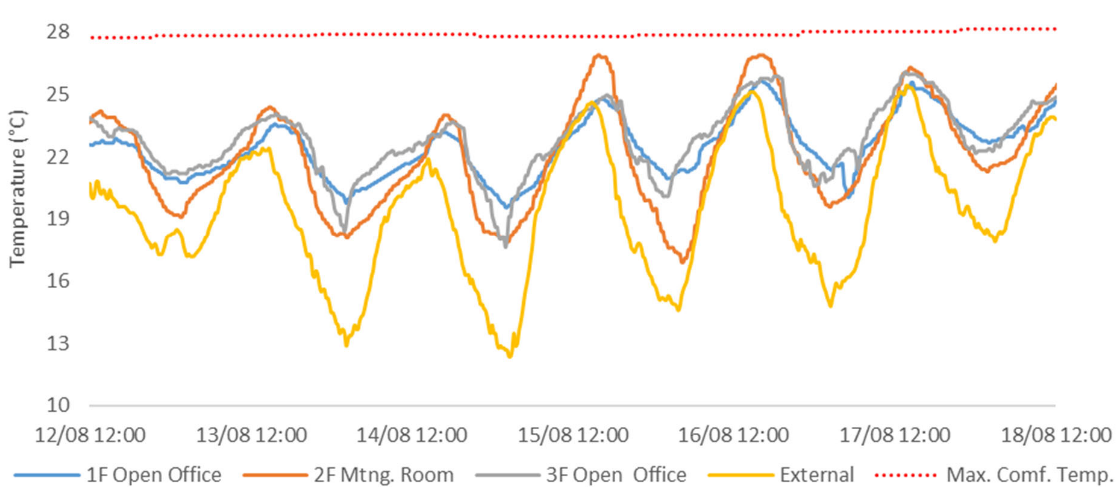


Figure 4.4: Indoor temperatures in a typical summer week in the public office building

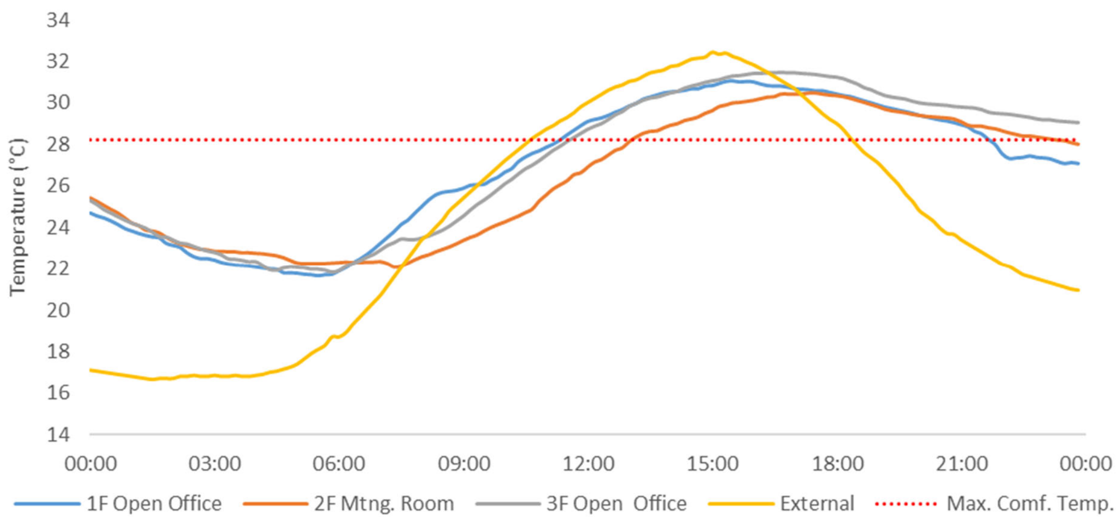


Figure 4.5: Indoor temperatures on a hot summer day in the public office building

In the typical summer week (Figure 4.4), space temperatures were within the comfort levels (as per TM52). However, temperatures did go beyond 25°C, due to high external temperature. Also on a particularly hot day (Figure 4.5), when the peak outside air temperature was 33.2°C, the indoor air temperature exceeded the comfort temperature significantly, suggesting some overheating issues.

As per the design strategy, night cooling could have kept the building comfortable in summers by maintaining a temperature of 2-3°C less than peak outdoor temperatures. However, due to the limitations of present controls, some operational issues occurred which could have been responsible for the high indoor temperatures. To avoid overcooling to the extent that heating was required when the office opened in the morning, the vents were only open at nights, when the indoor temperature was above 19°C. This made the night cooling strategy ineffective.

Indoor air quality

As the building was naturally ventilated, fresh air was provided via automatic vents controlled by the BMS as well as additional manually operable vents. Figure 4.6 shows that IAQ in the spaces monitored was under their respective thresholds. Seasonal variation is seen in both NO_2 and mean CO_2 levels. The concentration levels of NO_2 are significantly higher during summer months, when vents and windows were open. Whilst the current concentration levels are lower than the WHO limit, the numbers show the potential risk of control strategies that are solely based on CO_2 concentration for IAQ control.

Also, it is notable that CO_2 levels in the open-plan office on the third floor are higher than the lower floors. There are episodes where CO_2 concentration levels reach to 2500 ppm. Figure 4.7, reproduced from the BMS data, shows the CO_2 concentrations for the third-floor open plan office on a typical winter's day. The average CO_2 concentration during the workday was >1900 ppm. For, more than 90% of the working hours it was above 1500 ppm, reaching up to 2500 ppm.

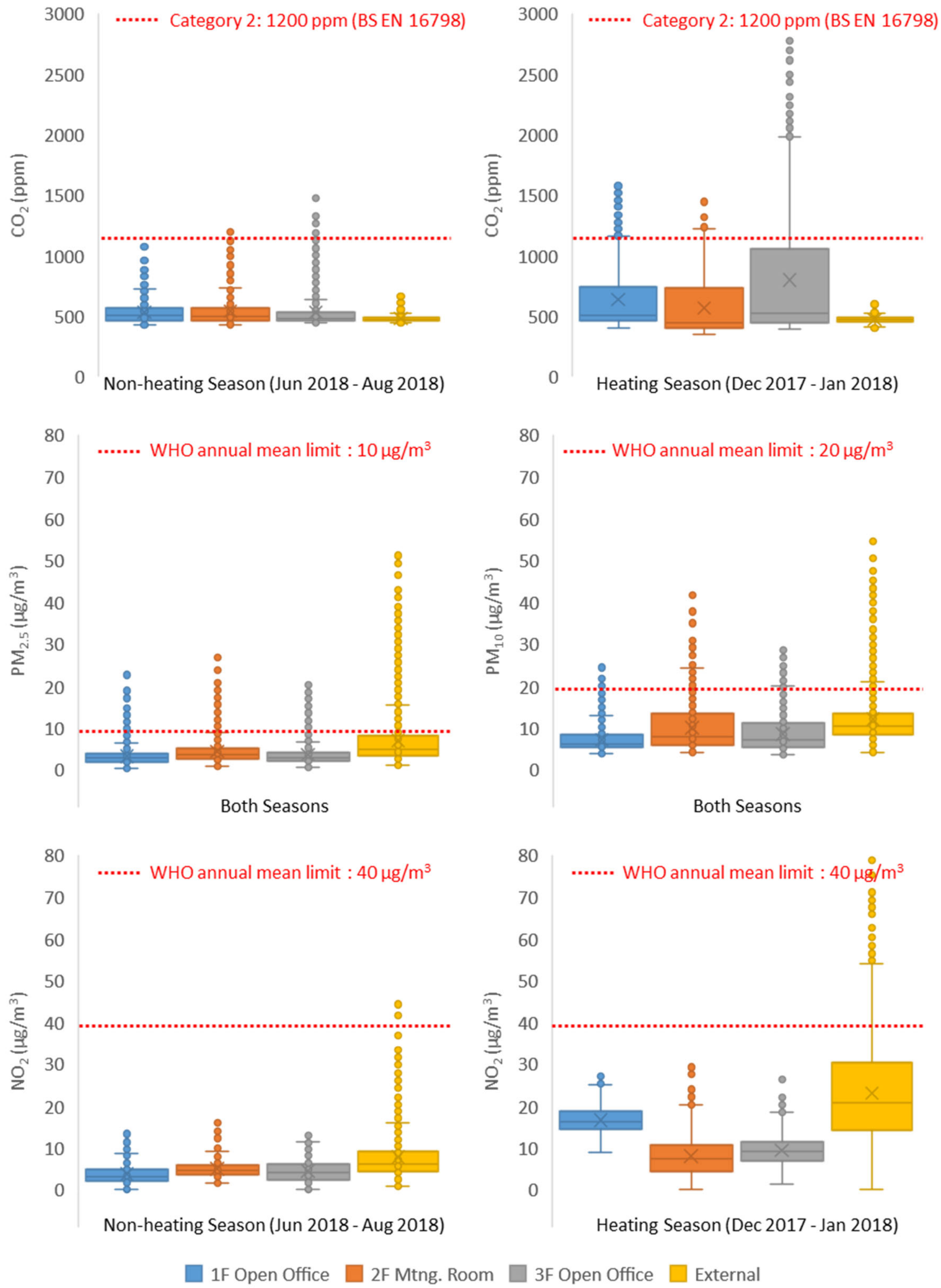


Figure 4.6: Box plots showing the spread of IAQ parameters in the public office building

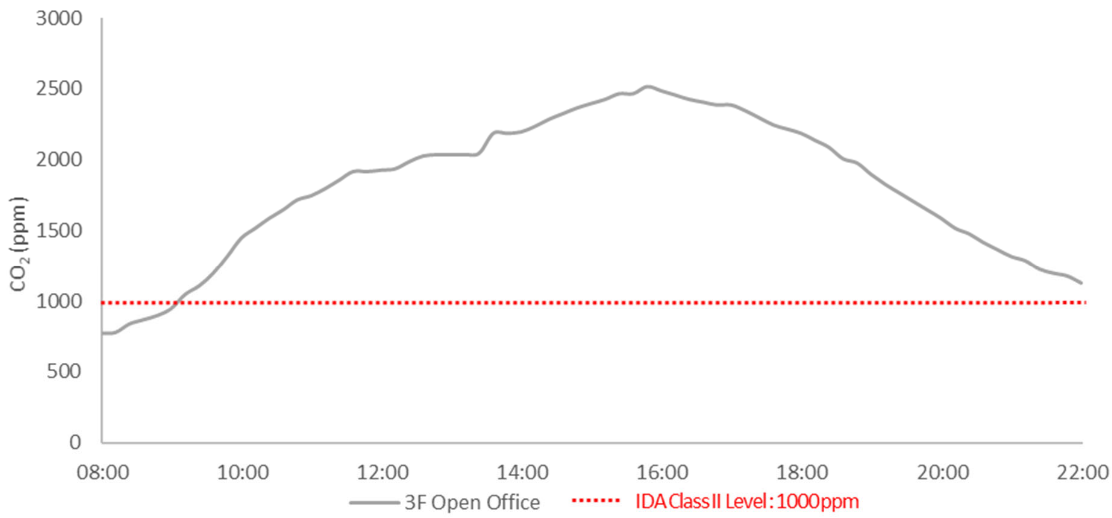


Figure 4.7: CO_2 concentration on a typical winter's day in the third-floor open-office in the public office building

Apart from the effect of stack ventilation that may cause concentration levels slightly higher than lower floors, on-site observations found a malfunctioning CO_2 sensor that led to the underestimation of concentration levels on the third floor, meaning the motorised vents were not open when required.

4.1.4 Observed deviation areas and issues

During the operational performance assessment of the building, many deviation areas related to operations and technical issues were found. These issues were discussed in the last two sections and are summarised here:

- The total occupancy of the building was about 25-30% higher than design-stage estimates, leading to more workstations being added and the use of small power and lighting was for longer durations.
- Technical issues with heat exchangers and the flow rate specification caused the heat pumps to malfunction. As a result, the gas-fired boilers supplied all the heating.
- Server loads were overestimated in design. This adversely impacted the heating system efficiency, as there was less free heat.
- Heating set-point temperatures maintained in the building were about 2°C higher than the design intention, driven primarily by occupant needs.
- Some of the automated ventilation control sensors malfunctioned.

The issues listed here inform further data analysis and the model calibration, which is presented in the next chapter, along with a more detailed evaluation and identification of these and other deviation areas.

4.2 SCHOOL

The secondary school and sixth form, with academy status, is located in inner London, England. The school was redeveloped in 2014 with the construction of six new buildings (including teaching spaces, a sports hall, and a performance arts and dining hall) and the refurbishment of several existing ones (swimming pool and gymnasium building, and assembly hall). The buildings are generally four stories high, with a total useful floor area of 21,405m^2. The project was required to implement on-site renewable energy technologies to meet the local council's planning conditions. To satisfy this, a biomass boiler utilising solely wood pellets and solar thermal collectors was implemented. Along with an overall campus-level assessment, detailed issues were assessed in one of the teaching buildings.

Figure 4.8: External view of the school building

4.2.1 Design-stage details

Architectural design

The envelope was made of prefabricated concrete panels that were assembled at the site. The building, designed for high energy efficiency, had low fabric U-values, and had an emphasis on avoiding thermal bridging. Fabric U-values (W/m^{2}K) were as follows: Wall: 0.17; Window: 1.63; Roof: 0.20; Ground: 0.15; and design airtightness was 5 m^3/hr/m^2 @ 50 Pa. Spaces were designed to have large windows for daylighting and were also designed to be partially operable for natural ventilation and free cooling in summer.

HVAC systems design

Heating was to be provided through a centralised plant for the entire campus via a pressurised low-temperature hot water system. A biomass boiler (heating seasonal efficiency: 0.75), sized to meet annual domestic hot water demand, and two gas-fired boilers (heating seasonal efficiency: 0.84) were installed to provide heat in the building. Variable refrigerant flow systems were installed in rooms with high information communications technology (ICT) loads and server rooms (heating/cooling seasonal energy efficiency ratio: 1.47/3.80). There was no provision for comfort cooling to any other space.

A mechanical ventilation system with heat recovery (efficiency: 0.75) via a centralised roof-mounted AHU was designed to provide fresh air in the buildings, distributed through wall-mounted diffusers/grills. BMS-controlled ventilation was provided in the spaces based on the installed CO_2 sensors in each room.

Lighting and electrical systems design

The lighting system was designed to be energy efficient, with T5 florescent lamps in classrooms and offices. The rooms were also equipped with daylight and PIR sensors to provide automated lighting control. Besides lighting, there was a constant server load in addition to other loads, including small power, lifts, pumps and fans.

Building and services operations

The building was designed to be occupied during the weekdays as per normal school hours, with half days on Saturdays. All services were intended to run during occupied times. Table 4.4 lists the design intended characteristics of the building for occupancy, services and operations.

Table 4.4: Building occupancy and services operations and loads details at the design stage

Category	Details
Occupancy	Number: 2250 (2000 pupils and 250 staff) Weekdays: 0830 -1500 with diversity Saturdays: 0900 -1300; Sundays & holidays: zero.
Heating	Operation: Setpoints maintained follow occupancy. Set point: 20°C
Cooling	Operation: Setpoints maintained follow occupancy. Set point: 23°C
Auxiliaries; Mech. Vent.	Mechanical ventilation rate: 5 to 12 l/s/person; Specific fan power: 1.8 W/l/s. Operation: follows occupancy
Lighting	Loads: 12 W/m^2 – daylight integrated. Operation: follows occupancy
Small Power	Loads: Classroom 20 W/m^2; Offices 10 W/m^2; ICT Areas: 50 W/m^2. Operations: loads follow occupancy.

Metering strategy

Gas use in the facility, metered at the site level, was recorded in utility bills monthly. Each new building had its respective heat meter which was linked to the BMS. The mains electricity meter recorded half-hourly electricity use at the site level, which was available from the utility supplier. At the building-level, disaggregated energy use for lights, small power, lifts, server, pumps and fans could be read through the BMS.

4.2.2 Building energy performance

The design-stage projections were done at the design development stage as a part of UK Building Regulations (Part L) compliance calculations. The calculation, conducted for the whole facility, reported annual energy use projections for each building separately. The actual performance of the teaching building was monitored from October 2016 to September 2017. The gas used in the teaching building, supplied through a district heating network, was recorded by the building-level heat meter. The metered readings, combined for heating and hot water use, have been reported here, after adjusting them for the efficiency of the campus boiler. Also, three spaces (ICT rooms and server room) which have cooling and heating provided by VRF systems had their conditioning energy use accounted for in the small power electricity.

Table 4.5 shows a comparison between annual design and actual performance of one of the teaching buildings, and Figure 4.9 shows the actual monthly gas and electricity use patterns. For this school building, all data was not recorded directly. While disaggregated electricity use readings were available from the BMS, they could only be recorded during site visits, which happened at varying (one to two month) time intervals. The high-granularity electricity use data from the utility supplier was at the facility-level with no disaggregation. This data was used to create monthly energy use profiles for the building. Trends for facility-level half-hour electricity use data trends were used to proportionally redistribute the building energy use data recovered from the BMS. This way the BMS readings were appropriately distributed, accounting for weekdays, weekends, term times and holidays. This method assumed that the whole facility's energy use trends across the buildings are similar.

Heating energy use, supplied by gas-fired boilers, also did not have high granularity in the data for generating detailed use profiles. Utility supplier data for gas use at the facility-level was available at irregular intervals spanning one to two months. Also, building-level heat use was recorded by a heat meter during site visits at varying time intervals. Therefore, to create a monthly heat demand profile for the building, the heat demand recorded during site visits was proportionally distributed. The distribution was based on daily heating degree days and, further, redistributed for weekdays, weekends, term times and holidays (using the trends from electricity use data).

Table 4.5 Annual design vs actual energy use comparison for the school

Annual design-vs-actual energy use				Good Practice Benchmark**
Fuel	**End Use**	**Design**	**Actual**	
Gas	Heating	4.9	168.3	
	Hot Water	14.5		
	Total	**19.4**	**168.3**	**108.0**
Electricity	Lighting	9.3	16.2	
	Cooling	1.2	15.4	
	Small Power	18.0***		
	Server		22.7	
	Fans/Pumps	8.5	34.2	
	Total	**20.0**	**88.5**	**25.0**

*All values are in kWh/m²/yr.; ** CIBSE Guide F benchmark for Education (schools): secondary (CIBSE, 2012); ***Not reported in totals as the calculations were done for regulatory compliance which factor in small power during calculations but do not report in the totals.

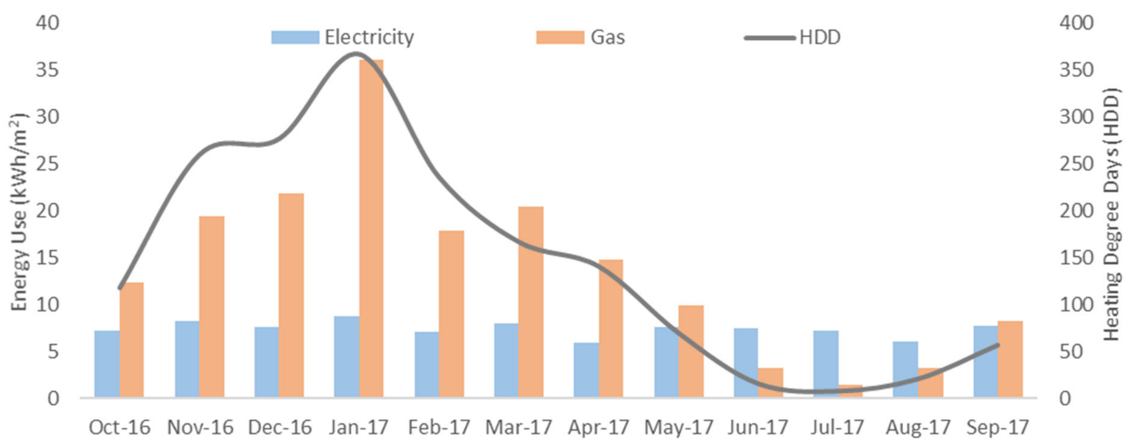


Figure 4.9: Monthly electricity use and gas use for the school building

Analysing the data, it can be seen that there is a significant underestimation of energy use in the design-stage calculations. One of the reasons for this is the interpretation of regulatory calculations at the design stage as the predicted performance. Regulatory calculations in the UK use standardised UK-NCM input parameters, such as default power densities for pumps. Also, some of the key energy use areas such as equipment, server and lift, are not included in the totals when reporting design projections. Moreover, the occupancy and operational profile of the building is also calculated based on standard schedules, which, in real scenarios, can be significantly different, such as the assumption that all systems are shut down during half-term breaks and school holidays. Despite being a new building, gas and electricity use was 1.5 times and 2.5 times more, respectively, than the good practice benchmarks (CIBSE, 2012). Some of the deviations can be explained by the fact that the biomass boiler installed was not used in practice due to logistical issues of managing the biomass. Also, site observations suggested that the building was occupied after school hours and during half-term breaks and school holidays.

4.2.3 Building IEQ performance

In the school building, the parameters for thermal comfort (Temperature and RH) and various air pollutants (CO_2, $PM_{2.5}$, PM_{10} and NO_2) were recorded in four typical locations: library (Ground Floor, West), classroom (Building 3 First Floor, North), science lab (Building 4 First Floor, East) and external space (within the campus, close to the main road). This section reports the findings along with the relevant limits for comfort or wellbeing. The results are presented in box plots for each parameter, with more detailed graphs where unique trends or major issues are seen.

Thermal comfort

The building has space heating provision but comfort cooling is only in high ICT spaces; therefore, in naturally ventilated spaces without comfort cooling, summer overheating is a concern. Figure 4.10 shows that the winter temperature in the library and the lab are towards the higher end or above the recommended range. This may be indicative of higher set points used in these spaces, which can be a contributing factor to the high heating consumption of the school.

Analysing the heating season temperature trends in more detail (Figure 4.11), it is observed that during occupied times the indoor temperatures in all sampled rooms were maintained above 23°C. Also, during holidays, temperatures were recorded much higher than the outdoor levels for the lab and the library, while the classroom temperature profile was similar to the external measurements.

During the non-heating season, while most of the spaces did not suffer from overheating, rooms on the south façade, lacking solar controls (blinds/shades) had high heat gains. They were susceptible to overheating risks in 'hot' summers. Figure 4.12 shows indoor temperatures in a south-facing classroom on the second floor during a hot spell in June. To evaluate the overheating risk of mechanically ventilated buildings, a threshold of 26°C is specified by BS EN 15251 (BSI, 2007); this was seen to be exceeded as external temperatures remained high. The standard, valid when monitoring was done, has been superseded by BS EN 16798-1:2019 (BSI, 2019). However, for all sampled spaces the BB101:2006 (DfE, 2006) overheating criteria, which was the basis of the design, was met. The temperatures of the sample spaces were within the BB101 criteria for hours-of-exceedance-limits, daily-weighted-exceedance-limits and upper-limit-temperature. This suggests that while the current overheating risk is not high, for longer periods, in the context of changing climate and increasing temperatures, the building will need to adapt by using strategies such as night purge ventilation.

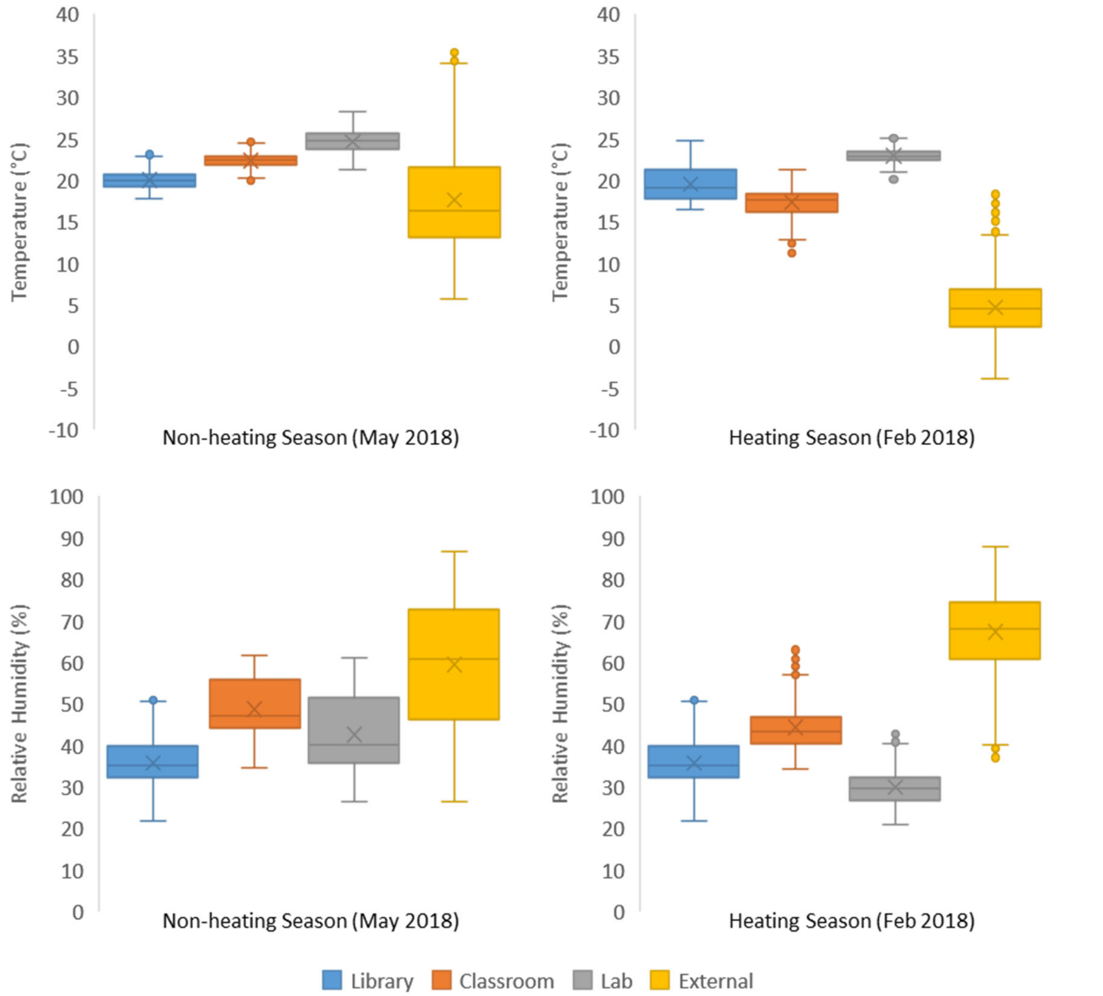


Figure 4.10: Box plots showing the spread of Temperature and RH during non-heating and heating seasons in the school

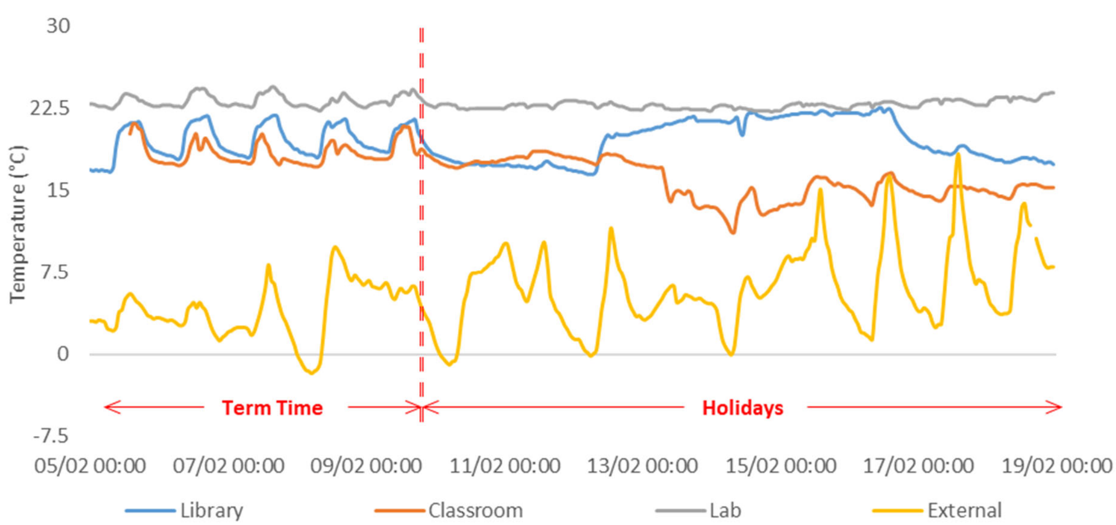


Figure 4.11: Temperature variations during term time and holidays in the heating season

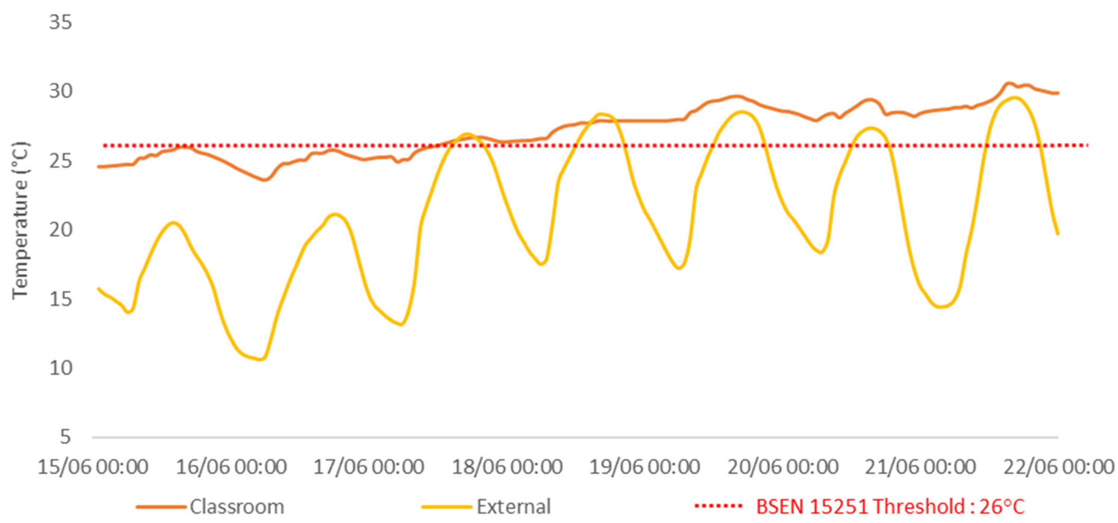


Figure 4.12: Indoor monitored temperatures in a classroom during a hot summer week

Indoor air quality

As buildings are mechanically ventilated in this school, most spaces had adequate fresh air supply during occupancy hours. Figure 4.13 shows that IAQ in the spaces monitored was generally under their respective thresholds.

The daily averaged value of 1500 ppm, recommended by BB101:2006, was the basis of the design (DfE, 2006). CO_2 levels during the heating and non-heating seasons in the monitored space were always under 1500 ppm, except for the classroom where there was a failure of a supply fan in the respective air handling unit.

NO_2 concentration levels in the library which is located next to the main road are higher than the other monitored zones, although the median value is still lower than the WHO recommended annual mean. Concentration levels of fine particulate matter ($PM_{2.5}$) are significantly lower than the WHO limit, as a result of air filtration in air handling units. Also due to mechanical ventilation and good airtightness, the fresh air intake is controlled and filtered. $PM_{2.5}$ and PM_{10} concentrations in the monitored spaces were always below external values and significantly below the WHO 24-hour mean threshold of 25 μg/m^3 and 50 μg/m^3 for $PM_{2.5}$ and PM_{10} respectively.

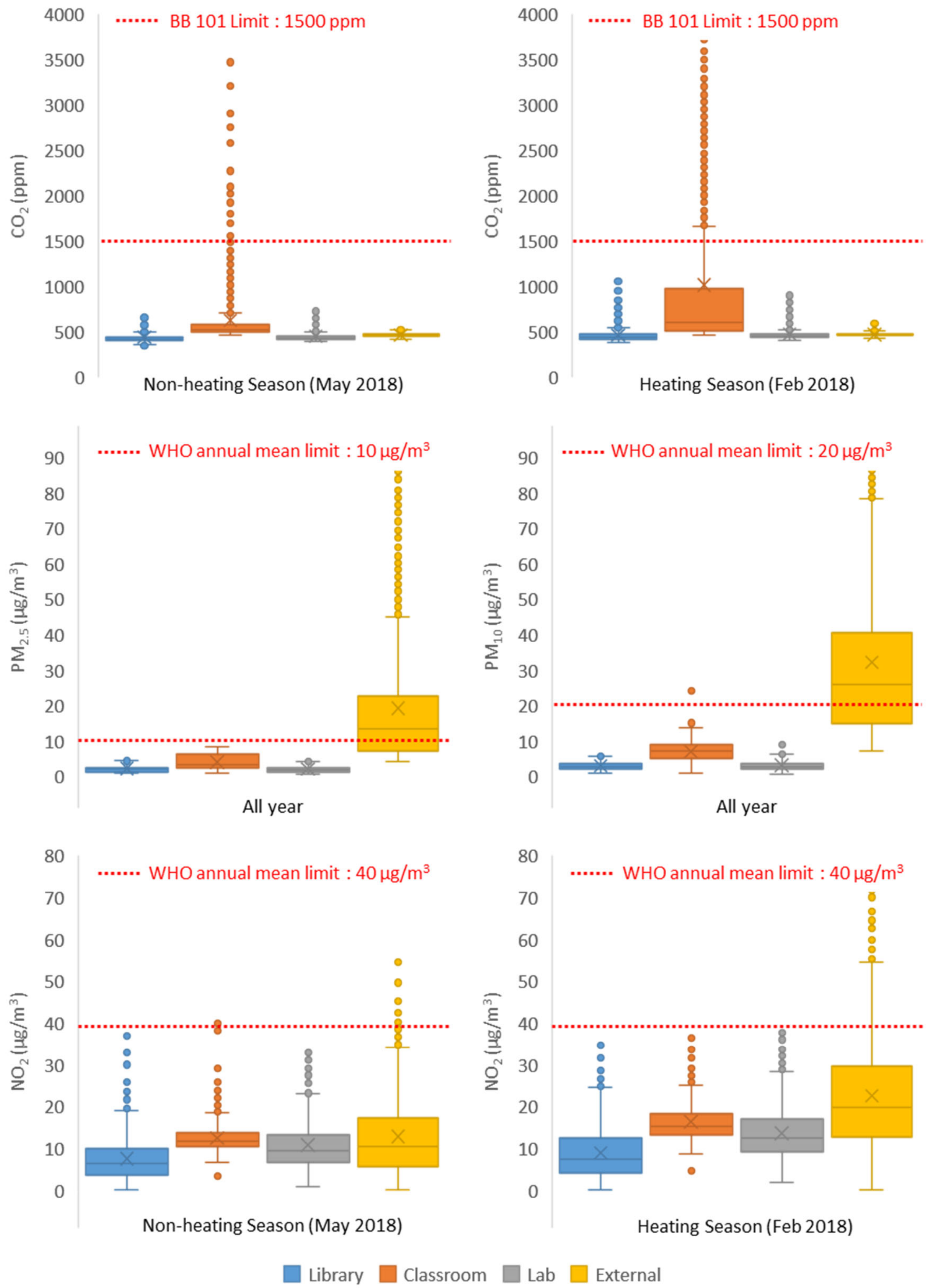


Figure 4.13: Box plots showing the spread of IAQ in the school

4.2.4 Observed deviation areas and issues

During the operational performance assessment of the building, many deviation areas of operations and technical issues were found. These issues were discussed in the last two sections and are summarised here:

- Not all spaces were occupied throughout the day. Any given space was occupied only 60-70% of the full working day. Also, during academic breaks, the school is not completely shut. There are extra-curricular activities and events, especially during the summer holidays.
- The biomass boiler, installed for decarbonising energy use, was never used, citing practical and logistical issues of using biomass as fuel.
- The actual boiler efficiency (tested) was 88%, compared to the design intent of 96%.
- The building heating and ventilation systems were working during unoccupied hours and indoor temperatures monitored during the winter season were about 2°C higher than the set-point temperatures, and sometimes even more.

The issues listed here inform further data analysis and the model calibration, which is presented in the next chapter, along with a more detailed evaluation and identification of these and other deviation areas.

4.3 HOSPITAL

This building is a new ward complex built within an existing hospital campus in 2015 and is classified as an 'acute hospital', as defined in ECG72 (BRECSU, 1996). The nine-floor building is in Bristol, in southwest England, with a useful floor area of ~14,700m^2. Typical spaces include wards, consulting rooms, offices, diagnostic spaces, operating theatres, a canteen and the usual building amenity rooms.

Figure 4.14: External view of the building

4.3.1 Design-stage details

Architectural design

The building was designed as a curtain wall construction with concrete lattice floor slabs. It was planned to be highly insulated and heavyweight in terms of thermal mass. The building design had a 23% window to wall ratio and solar control glazing. There were no external shadings besides the surrounding buildings. Fabric U-values (W/m^{2}K) were as follows: Wall: 0.22; Window: 1.60; Roof: 0.23; Ground: 0.25; and design airtightness was 5 m^3/hr/m^2 @ 50 Pa.

HVAC systems design

Most of the building is designed to be heated and cooled using an all air supply/extract ventilation system with ventilation terminal heating coils. The heat in the building was to be provided via a low efficiency (~70%) old campus-wide steam-based central heating network. Mechanical ventilation with heat recovery and cooling was to be provided through rooftop mounted air handling units and chillers. The mechanical ventilation system (along with heating and cooling) was to be controlled via BMS, along with individual rooms having analog manual overrides.

Lighting and electrical systems design

The building was designed to have low energy artificial lights, primarily T5 & CFL lamps. The building also had low-energy equipment and thin client computers along with other loads, including hospital medical equipment, small power, IT, pumps and fans.

Building and services operations

Most of the building spaces were planned to be occupied 24 hours on all days apart from the consulting rooms and offices, which were to be occupied only on weekdays from 8 am to 6 pm. Details about the occupancy were as per typical occupant patterns; however, the irregular nature of processes made it difficult to plan for typical events and average patterns with high certainty. A centralised BMS was to be installed to control all operations, along with manual overrides and controls. Table 4.6 lists the design-intended characteristics of the building for occupancy, services and operations.

Table 4.6: Building occupancy and services operations and loads details at the design stage

Category	Details
Occupancy	Number: 200 Beds (Typical Floor: 28 beds and 22 staff) Hospital Wards: 24 hours occupied; Office spaces: Weekdays 0800-1800, Weekends: Zero.
Heating	Operation: 24-hours Set point: 22°C
Cooling	Operation: 24-hours Set point: 24°C
Auxiliaries; Mech. Vent.	Auxiliaries Load: 2.7 W/m^2; Mechanical ventilation, Wards: 17.4 l/s/person, Food areas: 25 l/s/person, Operation theatres: 129 l/s/person, Other spaces: 10 to 12 l/s/person Specific Fan Power: 1.7 W/l/s; Heat Recovery efficiency: 0.75; Operation: 24-hours.
Lighting	Loads: 15 W/m^2 in wards but varies as per space. Operation: 24-hours
Small Power	Loads: 12.5 W/m^2 in wards but varies as per space. Operations: 24-hours

Metering strategy

Gas use in the facility was to be metered at the site level, with each building having its own heat meter which would provide building-wise heat demand. The local (building-level) electricity meter that could be accessed via BMS was enabled to record hourly electricity use. Disaggregated energy use for lights, small power, IT, pumps and fans, and cooling was also available. All meters were designed to be integrated to the BMS.

4.3.2 Building energy performance

The design-stage projection of energy performance, done as a part of UK Building Regulations (Part L) compliance calculations, was interpreted as the intended performance. The actual performance monitoring was done from November 2015 to October 2016. Table 4.7 shows the design and actual energy use disaggregated at annual level, and Figure 4.15 shows the actual monthly gas and electricity use patterns. The disaggregated electricity use of the building was obtained from the BMS. Heat for both space heating and domestic hot water, supplied through a district network, was recorded from a building-level heat meter. The gas use reported was adjusted for the efficacy of the district network.

Table 4.7 Annual design vs actual energy use comparison for the hospital

Annual design-vs-actual energy use				Good Practice Benchmark**
Fuel	**End Use**	**Design**	**Actual**	
Gas	Heating	83.8	303.5	
	Hot Water			
	Total	**83.8**	**303.5**	**422.3**
Electricity	Lighting	17.7	33.9	
	Cooling	3.3	14.6	
	Small Power	43.4***	32.0	
	Server	N.A.	N.A.	
	Fans/Pumps	14.0	33.3	
	Total	**35.0**	**113.8**	**74.4**

*All values are in kWh/m^2/yr.; ** CIBSE Guide F benchmark for Hospitals: acute and maternity (CIBSE, 2012); ***Not reported in totals in calculations done for regulatory compliance which factor in small power during calculations but do not report in the totals.

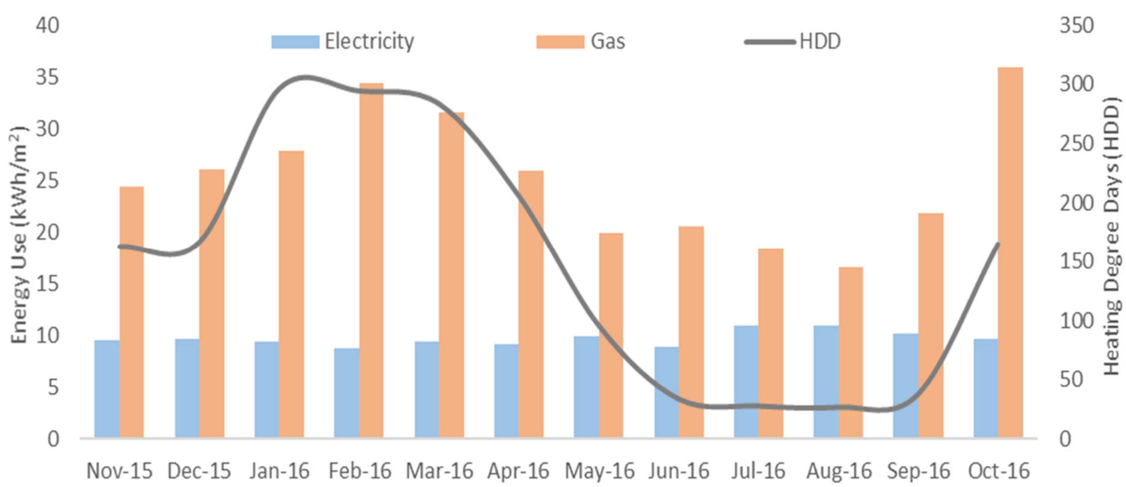


Figure 4.15: Monthly electricity use and gas use for the hospital building

Hospitals are energy-intensive buildings with complex and varying specifications for their functions and are operational 24/7. The energy use projections of various end-uses, done as part of the design-stage Part L calculations, differ from actual use and were severely underestimated. The major reason for this is that these calculations, primarily aimed at benchmarking, exclude key energy use areas, such as plug-in equipment, in the total projections. Moreover, the occupancy and operational profiles are calculated based on UK NCM defaults which, in real scenarios, can be significantly different.

The high proportion of medical equipment energy (~30%) is unique to hospitals and has a specific load and usage. This explains some of the overestimation for this end-use in the design projections.

One of the big reasons for underestimation of the gas use was the ongoing use of the low-efficiency district heating system, which, as per design intent, was to be replaced by a new high-efficiency combined heat and power (CHP) system. The new CHP plant was to be installed following a major renovation to maximise the efficiency savings across the facility rather than as a separate system for the new building only. As this had not happened when the monitoring was done, the thermal performance of the building was much worse than what would be expected from a new building.

For this hospital, hourly disaggregated electricity use data was available from the BMS. However, as a central steam network was used for space heating and domestic hot water, only monthly spot measurement of the heat meter was available. Therefore, to create a monthly heating demand profile for the building, the heat demand recorded in the spot measurement was proportionally distributed. The distribution was based on daily heating degree days and, further, redistributed for weekdays and weekends (using the trends from electricity use data).

Also, the monthly breakup of actual performance shows significant gas use during non-heating periods (Jun-Sep). This base demand is primarily for domestic hot water use, which in hospitals in the UK has been reported to be very building-specific and constitutes a significant proportion of total energy use (DECC, 2018). It was not possible to monitor disaggregated heat demand for space heating and domestic hot water. However, as hot water demand in hospitals is largely independent of external weather conditions, the approximate baseline demand for it can be calculated using Heating Degree Days (HDD). Figure 4.16 plots the heat demand against HDD. The intercept of the best fit line provides the approximate monthly hot water heat demand.

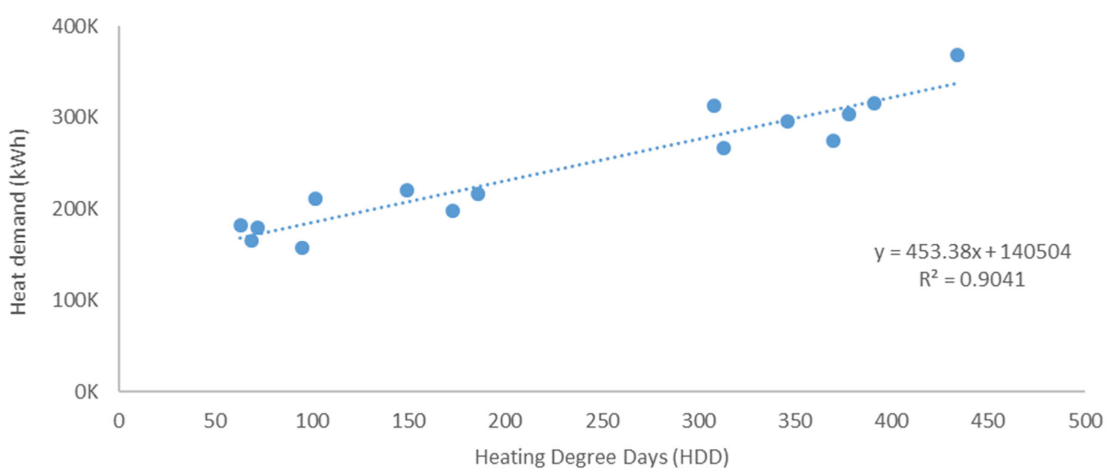


Figure 4.16: HDD analysis to estimate the non-space-heating portion of monthly gas demand (Data from 2015 to 2016)

4.3.3 Building IEQ Performance

In the hospital building, the parameters for thermal comfort (Temperature and RH) and various air pollutants (CO_2, $PM_{2.5}$, PM_{10} and NO_2) were recorded at four locations, three patient wards (on the third, fourth and seventh floors) and outside (near the main road). This section reports the finding, along with the relevant limits for comfort or wellbeing. The results are presented in box plots for each parameter, with more detailed graphs where unique trends or major issues are seen.

Thermal comfort

As the building operated with close control over the indoor environment, thermal comfort issues were significantly lower, as seen in Figure 4.17. The temperatures were between 20-24°C and were about 1-2°C higher in summers than in winters. Temperatures in the fourth-floor ward, which were occasionally higher than the recommended range during the summer period, were more of an occupant control issue rather than a system malfunction issue.

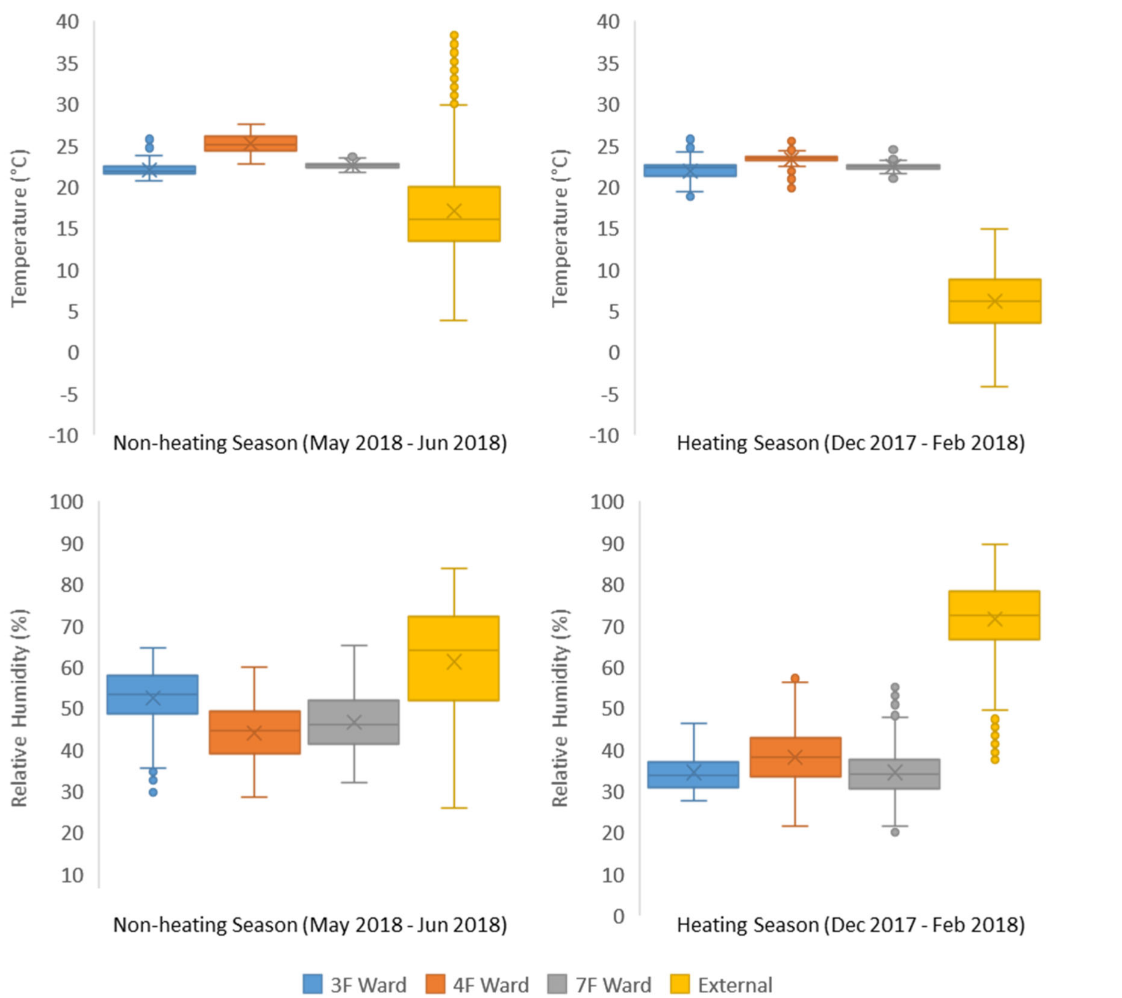


Figure 4.17: Box plots showing the spread of Temperature and RH during non-heating and heating seasons in the hospital

Indoor air quality

As the building is mechanically ventilated, filtered fresh air is provided. Figure 4.18 shows that due to a high level of control, IAQ in the spaces monitored was generally under their respective thresholds.

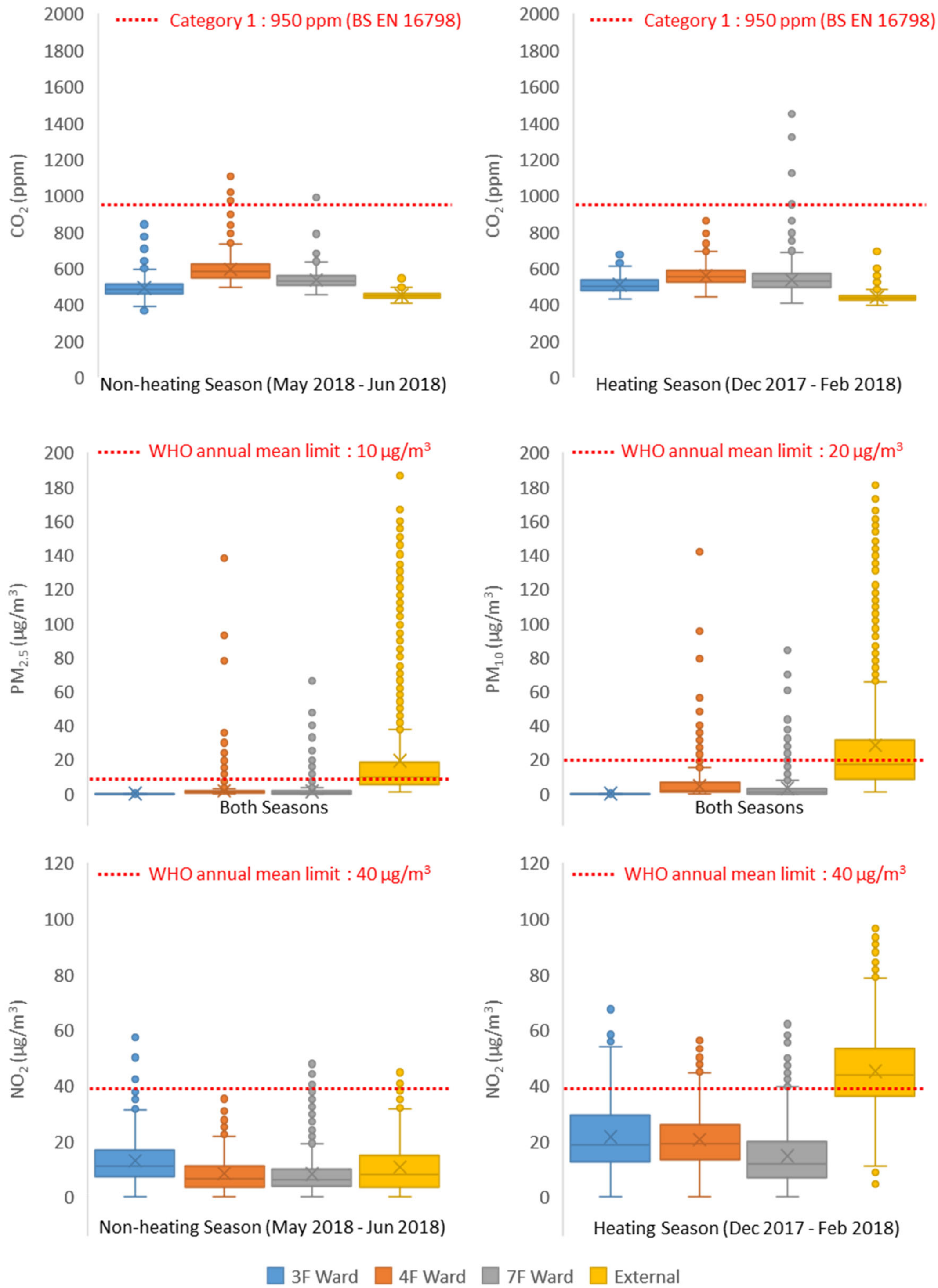


Figure 4.18: Box plots showing the spread of IAQ parameters in the hospital

CO_2 levels in all monitored wards remained below 950 ppm, due to effective mechanical ventilation and high air change rates. This is in accordance with the class 1 requirement of ≤ 550 ppm above outdoor as per BS EN 16798 (BSI, 2019). To meet these requirements, the mechanical ventilation system provided 10 ach to most medical spaces and 6 ach to examination and measurement rooms. $PM_{2.5}$ levels were also significantly lower than external levels and were less than the WHO 24-hour mean threshold of 25 μg/m³ (WHO, 2005). This shows that the mechanical ventilation air filters effectively controlled the ingress of micro particles. It should be noted that maintaining high IEQ through high air changes and filtration comes at an energy expense. Consequently, fans and pumps used to provide this close control use around 30% of the total electricity (see Table 4.7).

A key finding in this building, located in a congested urban area, was the lack of measures against ambient NO_2. Indoor NO_2 levels recorded in the hospital wards very closely followed the external levels (Figure 4.19). These recordings were below the WHO daily mean threshold of 200 μg/m³ but suggest a potential risk of exposure to more than the WHO recommended levels if the external air becomes more polluted for prolonged periods. Advanced chemical filtration and controls that consider the balance between the requirement for fresh air and protection from outdoor sources of pollution could provide a healthier environment and at the same time save energy.

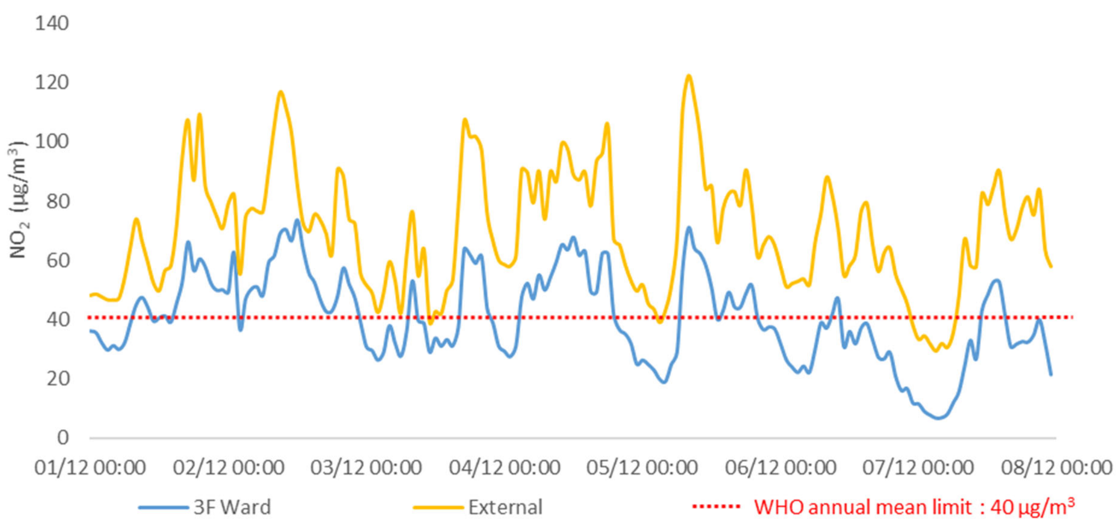


Figure 4.19: NO_2 measurements in a ward

4.3.4 Observed deviation areas and issues

During the operational performance assessment of the building, many deviation areas of operations and technical issues were found. These issues were discussed in the last two sections and are summarised here:

- Different clinical processes had unique needs. It was difficult to generalise typical operational trends of various spaces, their equipment loads; and set-point temperatures.
- An increased number of beds were seen in the patient wards during the site visits.
- A low efficiency, old, steam-based central heating network serves the building. A new CHP plant was to be installed for the entire facility rather than just for the new building, to maximise savings. However, this has not happened yet.
- Hospital hot water energy use was high due to clinical requirements.

The issues listed here inform further data analysis and the model calibration, which is presented in the next chapter, along with a more detailed evaluation and identification of these and other deviation areas.

4.4 APARTMENT BLOCK

The apartment block, completed in 2015, is located in East London. The block consists of 98 flats across two towers, as well as a community centre. The two towers are fourteen floors and ten floors high, with 48 apartments in the taller one and 34 apartments as well as 16 maisonettes in the other.

Figure 4.20: External view of the building

Performance issues were assessed for the entire apartment block, with a detailed investigation of a typical flat on the eighth floor. Figure 4.21 shows the spread of energy use in all the flats. The typical flat selected for detailed analysis has energy use close to the average value across the apartment block.

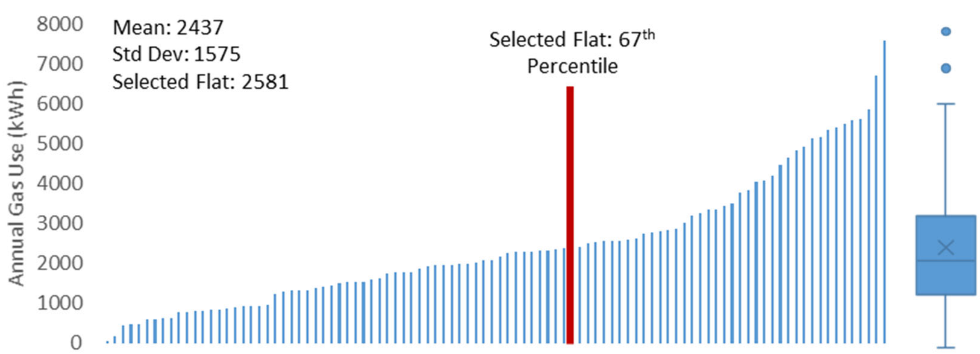


Figure 4.21: Spread of actual gas use of all flats in the apartment block. (Red line: typical flat used for detailed assessment.)

4.4.1 Design-stage details

Architectural design

Various sustainability strategies were integrated such as low U-values, good daylighting through full-height windows and high thermal mass to reduce overheating risk. Fabric U-values (W/m^2K) were: Wall: 0.18; Window: 0.92; Roof: 0.13; Ground: 0.12; and design airtightness was 5 $m^3/hr/m^2$ @ 50 Pa.

HVAC systems design

Heating and hot water in the apartment block were planned to be provided by a centralised boiler system, as part of a district heating network. Radiators were planned to be installed in the apartments to deliver space heating. There was no provision for mechanical cooling in the apartments. All the apartments had a dual mixed-mode ventilation strategy that used both mechanical ventilation with heat recovery (MVHR) and natural ventilation. Each flat had its own MVHR unit to supply the minimum background ventilation necessary for the health and wellbeing of occupants. All windows were to be operable to supplement natural ventilation for ventilation and space cooling.

Lighting and electrical systems design

The building, designed to be largely daylit, used low-energy artificial lights (CFLs). Other loads were assumed to be primarily for electricity consumed by appliances and house equipment. These loads are variable and would be largely dependent on the occupants.

Building and services operations

The apartments were assumed to be occupied as typical residences, with complete occupant control over all aspects of building operations. Table 4.8 lists the design intended characteristics of the apartment for occupancy, services and operations.

Table 4.8: Apartment occupancy and services operations and loads details at the design stage

Category	Details
Occupancy	Number: 5 persons Typical domestic patterns as per UK-NCM.
Heating	Operation: On-demand, linked with occupancy Set point: 21°C
Cooling	N. A.
Auxiliaries; Mech. Vent.	Mechanical ventilation air change rate: 0.5 ach; Specific Fan Power: 0.44 W/l/s; Heat Recovery efficiency: 0.75; Operation: On demand, linked with occupancy
Lighting	Loads: 15 W/m^2. Operation: On-demand, linked with occupancy
Small Power	Loads: Typical domestic loads as per UK-NCM (3.58 W/m^2 – 30.28 W/m^2) Operations: On-demand, linked with occupancy

Metering strategy

Heating and electricity use for all apartments were to be metered separately by the utility supplier. Readings (sometimes partly interpolated/extrapolated for electricity) would be available at monthly and daily levels for the apartment block.

4.4.2 Building energy performance

Energy performance was calculated for Building Regulation Compliance using a Standard Assessment Procedure (SAP) and was interpreted as the design-stage projection. Like the EPC, SAP is a method for calculating performance in domestic buildings in the UK. The actual performance was monitored from January 2017 to December 2017. Table 4.9 shows the design and actual energy use disaggregated at annual level, and Figure 4.2 shows the actual monthly gas and electricity use patterns. Monthly gas and electricity use were obtained from the utility bills, and end-use disaggregation was not possible. Also, heat supplied as part of a district heat network was adjusted for the district system's efficacy.

Table 4.9 Annual design vs actual energy use comparison for the apartment

Annual design-vs-actual energy use				**Good Practice Benchmark****
Fuel	**End Use**	**Design**	**Actual**	
Gas	Heating	11.5	113.4	
	Hot Water	29.3		
	Total	**40.8**	**113.4**	**247.0**
Electricity	Lighting	4.0	35.5	
	Cooling	N.A.		
	Small Power	29.4***		
	Server	N.A.		
	Fans/Pumps	2.2		
	Total	**6.2**	**35.5**	**44.0**

*All values are in kWh/m²/yr.; ** CIBSE Guide F benchmark for Residential and nursing homes (CIBSE, 2012); *** Not reported in totals in regulatory compliance calculations which factor in small power during calculations but do not report in the totals.

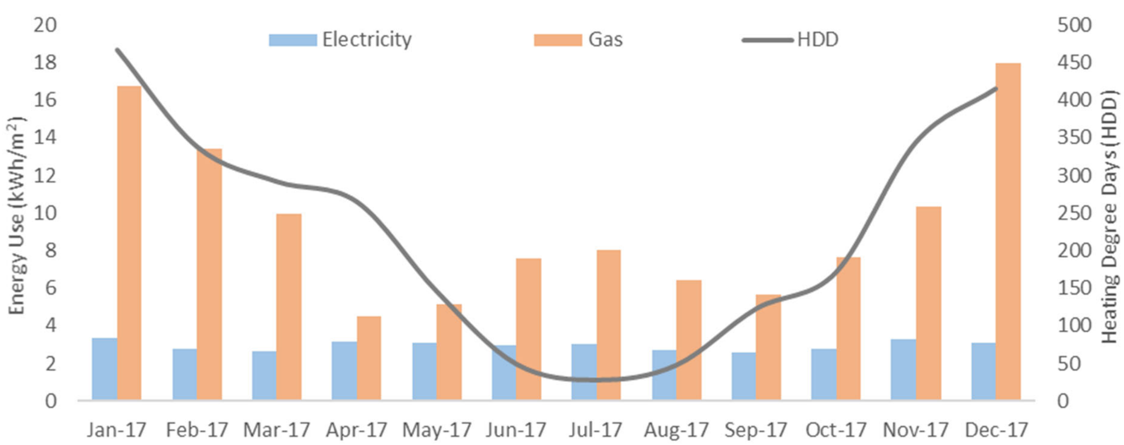


Figure 4.22: Monthly electricity use and gas use for the apartment

Both heating and electricity were underestimated in the SAP calculations. While actual gas use was marginally more than the design estimates, actual electricity use was much more. This is because small power use was not reported in the total use in the SAP calculations. In the actual monthly energy use, the dip in the heat demand in April and May, and the rise in June and July, can be attributed to changes in occupancy (such as holidays, guests) and other epistemic uncertainties of occupant behaviour and opening of windows in summer. These can be explored further and validated using calibrated models.

4.4.3 Building IEQ performance

In the apartment block, the parameters for thermal comfort and various air pollutants (CO_2, $PM_{2.5}$, PM_{10} and NO_2) were recorded in the bedroom, living room, kitchen and external space (near the main road on-site). This section reports the findings. The results are presented in box plots for each parameter, with more detailed graphs where unique trends or major issues are seen.

Thermal comfort

The building has space heating provision; however, as no comfort cooling is provided, summer overheating can be a concern. Figure 4.23 shows that the apartment was often outside the range for comfortable indoor temperature and RH in both heating and non-heating seasons.

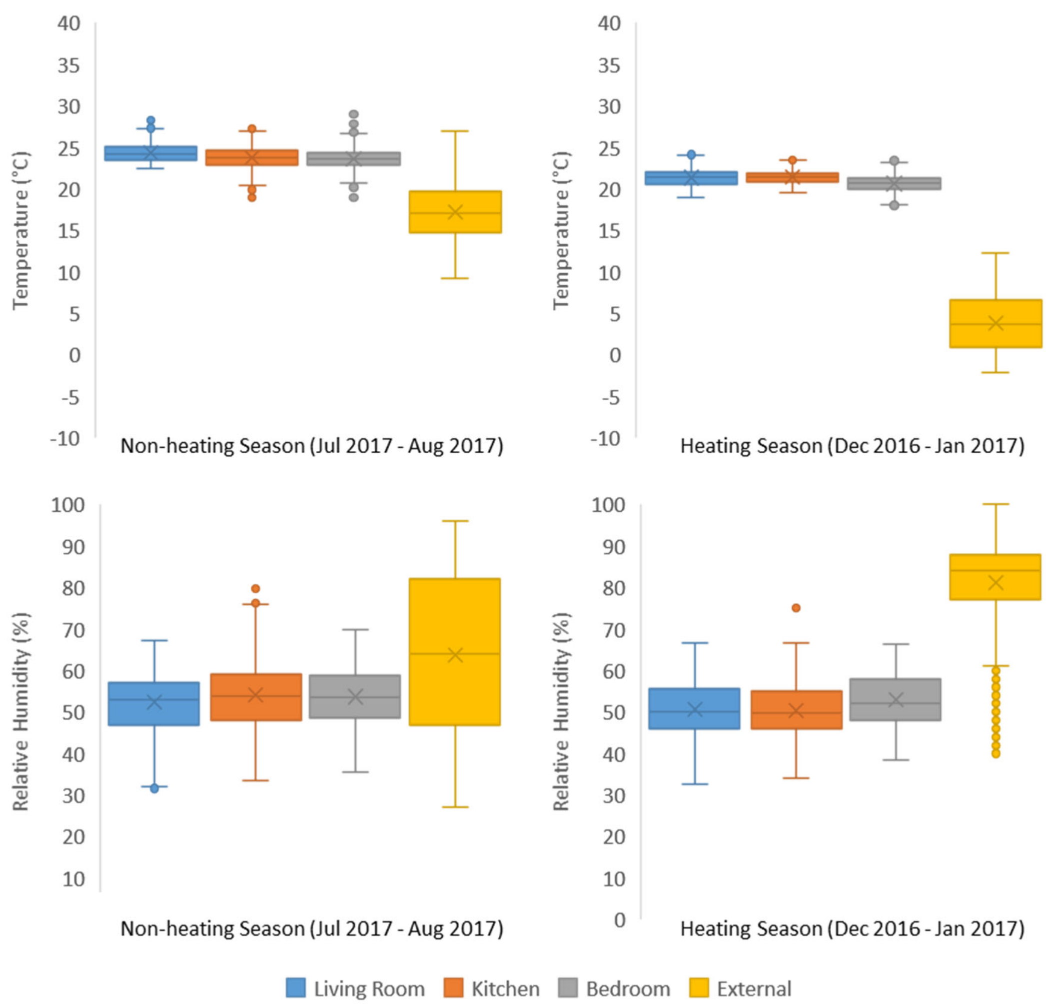


Figure 4.23: Box plots showing Temperature and RH during non-heating and heating seasons in the apartment

Indoor temperatures in the monitored spaces, were, on average, 24°C and 21°C during non-heating and heating seasons, respectively. Part of this can be due to occupant comfort preferences and their behaviour, as there was no significant issue raised by them on the issue of thermal comfort.

Indoor air quality

Being mechanically ventilated, fresh air is mainly provided via a MVHR system controlled by the BMS along with manually operable windows. Figure 4.24 shows the IAQ in the monitored spaces.

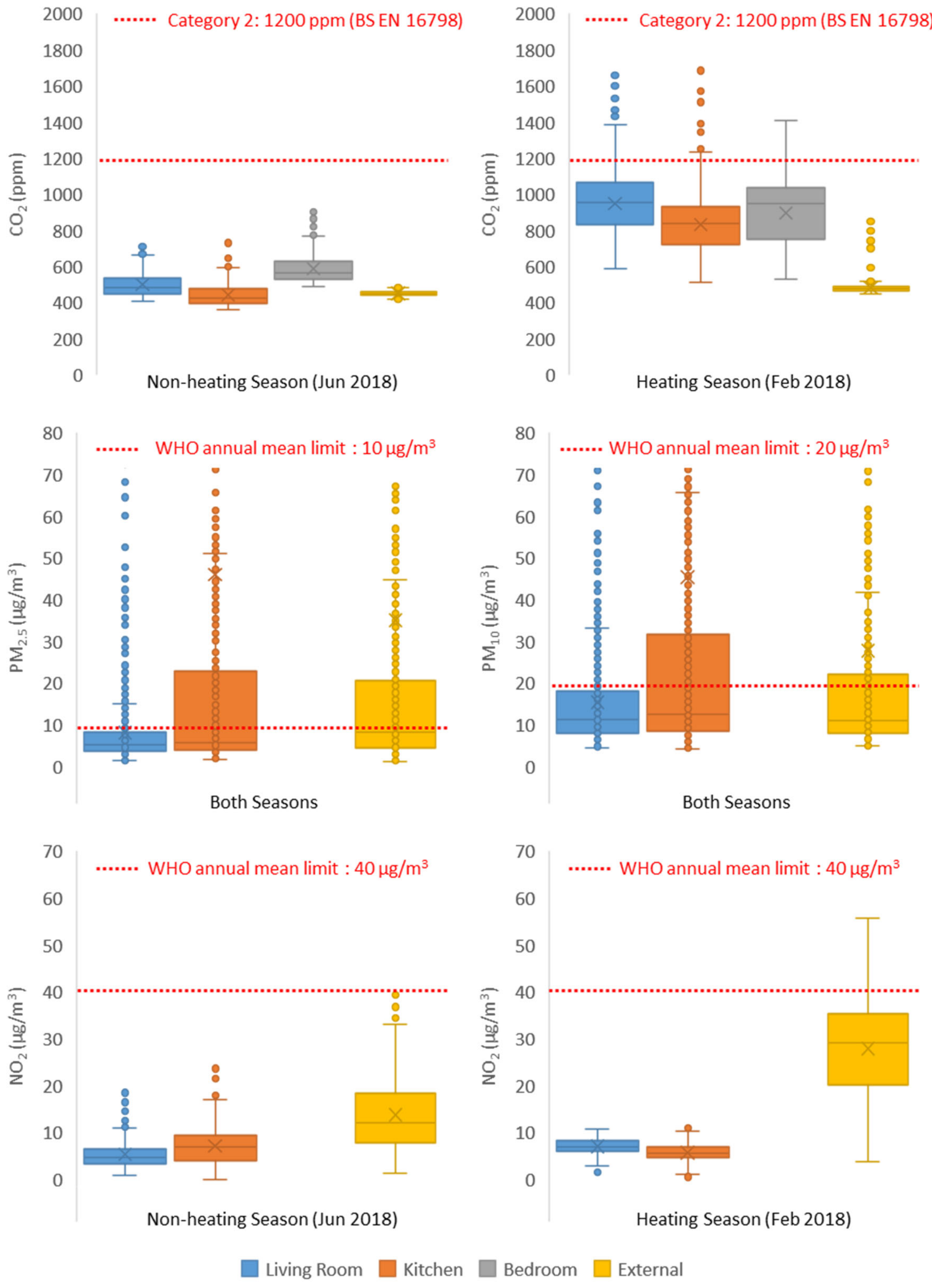


Figure 4.24: Box plots showing the spread of IAQ parameters in the apartment

Concentration levels of particulate matter are kept below the guideline limits most of the time. Whilst outdoor concentration levels of NO_2 are often higher than the WHO annual mean (40 µg/m³, 21 ppb), indoor NO_2 levels are generally lower than this limit. The intakes for the MVHR system are far from the main road and in most cases much higher than the ground level. In winter, when most windows are closed, the ventilation strategy is mainly dependent on the MVHR system (i.e. there is less air exchange between inside and outside). In addition to the effect of outdoor sources, PM and NO_2 levels in apartments could be increased by internal sources, especially in the kitchen.

Air quality issues were seen in the mechanically ventilated apartment block. $PM_{2.5}$ levels in the flat closely followed external levels and were frequently higher than the WHO daily average chronic limit of 25 µg/m³ (Figure 4.25). As ambient $PM_{2.5}$ levels are typical for London, the use of high-grade filtration, while slightly increasing electricity use, will help reduce $PM_{2.5}$ levels.

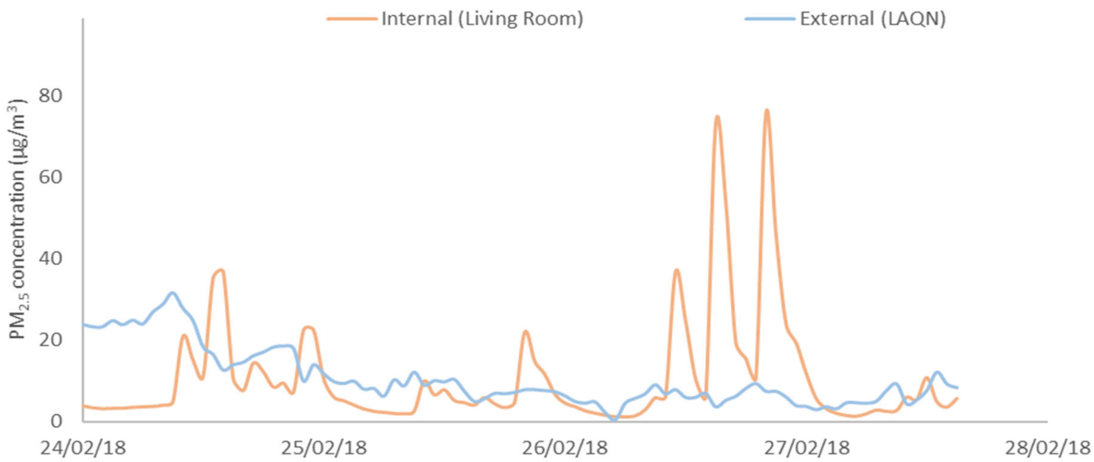


Figure 4.25: $PM_{2.5}$ measurements in the living room

4.4.4 Observed deviation areas and issues

During the operational performance assessment of the building, many deviation areas of operations and technical issues were found. These issues were discussed in the last two sections and are summarised here:

- The occupancy pattern of the individual flat is very uncertain to determine, occupants had full control over the heating system and MVHR, determining their set-points and operations.
- There are instances where, during summer, the windows were opened for enhanced ventilation while the heating system was still operational.
- Maintaining of filters in the de-centralised MVHR systems was left to the occupants and therefore there were maintenance issues (e.g. dirty filters) in most flats.

The issues listed here inform further analysis and the model calibration, which is presented in the next chapter, along with a more detailed evaluation and identification of these and other deviation areas.

4.5 University office

The ~4,500 m^2 university office building is located in central London, England. Built in the 1900s, the building recently underwent a major refurbishment, which was completed in 2010. Having a basement, ground and six upper floors, the building primarily consists of open-plan and cellular office spaces, meeting rooms, a library and computer clusters, among other facilities.

Figure 4.26: External view of the university office building

4.5.1 Design-stage details

Architectural design

The cast concrete and brick wall structure was made energy-efficient during the refurbishment. The external brick walls were fitted with composite board insulation and secondary double glazing was installed on the inside of existing single glazed windows. Fabric U-values (W/m^2K) were as follows: Wall: 0.20; Window: 1.40; Roof: 0.15; Ground: 0.15; and design airtightness was 5 $m^3/hr/m^2$ @ 50 Pa.

HVAC systems design

Heating and cooling were primarily supplied by a VRF multi-split system with heat recovery. The system allows for simultaneous heating and cooling in different zones. There were roof-mounted outdoor condensing units, and indoor spaces were provided with evaporators of differing capacities, with central control panels for adjusting and resetting the set-point temperatures. Hot water was to be provided through electric water heaters. Additionally, heating in the circulation area was provided by two condensing boilers and supplied in the spaces through hot water radiators. However, the energy used by the boilers was a very small proportion of the total building energy use. Also, since granular metered data is not available for gas use, boiler energy use has not been assessed in the study.

The building was primarily naturally ventilated through operable windows all around the perimeter, and spaces with no windows had fresh air provided through transfer grills from other spaces. However, some areas in the basement and ground floor were served with preconditioned air through an air-handling unit. Ducted extract systems served the toilets and shower areas on all floors.

Lighting and electrical systems design

Equipment load was planned to be like a typical office building but following a more diversified pattern of use, as it serves as a university building, where occupancy is related to academic term times. For office space lighting, recessed modular luminaires with T16 lamps were primarily used, controlled by absence detection sensors. Other than that, LEDs and CFLs were used in circulation areas, controlled by occupancy sensors.

Building and services operations

The building was designed to be occupied during the weekdays and all services were intended to operate during occupied times. However, the building has a different pattern of use compared to a typical office, as it serves as a university building, where occupancy fluctuates throughout the year. Table 4.10 lists the design intended characteristics of the building for occupancy, services and operations.

Table 4.10: Building occupancy and services operations and loads details at the design stage

Category	Details
Occupancy	Number: 600 persons (peak) Weekdays: 0900 -1800 with diversity; Low occupancy from in the morning, late evenings and on the weekends. Shutdown from 2300 to 0500.
Heating and Cooling	Operation: Linked to occupancy, setback running during unoccupied hours; manual control with occupants in each space. Set point: Varies by space and time of the year between 21-24°C; Setback: 12°C/30°C
Mech. Vent.	Not Applicable (air handling unit for basement and other inner zones was not operational during monitored time periods)
Lighting	Loads: Varies by space between 7-20 W/m^2. Operation: Linked to occupancy; Off during shutdown times
Small Power	Loads: Server/Plant 50 W/m^2; Other small power: 10-16 W/m^2 Operations: Linked to occupancy; Off during shutdown times.

Metering strategy

Most of the energy used in the building was electricity, with minimal gas used for heating in circulation areas. Mains energy use was metered at building-level meters and by the utility supplier at half-hourly intervals. Besides this, parts of the building had separate sub-meters, recording both spatial and end-use disaggregation, with an hourly reporting frequency. Space conditioning (heating + cooling) energy use at the building-level is recorded separately and floor level sub-meters record lighting and small power (including a small amount of hot water) energy use. The server and lifts were metered separately. The meters were designed to be integrated to a BMS. It was not possible to separate heating and cooling, and is it also not possible to separate lighting from small power; however, some short-term meters can provide typical trends for both to help in some trend estimations.

4.5.2 Building energy performance

Extensive metering and monitoring for the building was undertaken from August 2016 to July 2017. Table 4.11 shows the design and actual energy use disaggregated at the annual level, and Figure 4.27 shows the actual monthly gas and electricity use patterns. Being a refurbishment project, design-stage assumptions and calculations were not available for the building, and the actual data was collected from the sub-meters installed on-site. For this building, as noted earlier, most meters captured multiple building components that were difficult to separate from each other, such as lighting and small power on each floor. Another example is the inseparability of heating and cooling end-uses due to the use of a VRF system. These issues make it difficult to disaggregate different end-uses.

Table 4.11 Annual design vs actual energy use comparison for the university office building

Annual design-vs-actual energy use				Good Practice Benchmark**
Fuel	**End Use**	**Design**	**Actual**	
Electricity	Cooling	Design Calcs are not available	88.6	
	Heating			
	Hot Water		89.0	
	Lighting			
	Small Power			
	Server		5.4	
	Fans/Pumps		N.A.	
	Total	**N.A.**	**183.0**	**133.0**

*All values are in kWh/m^2/yr.; ** CIBSE Guide F benchmark for Offices: naturally ventilated, open plan has been added for fossil fuel and electricity (CIBSE, 2012)

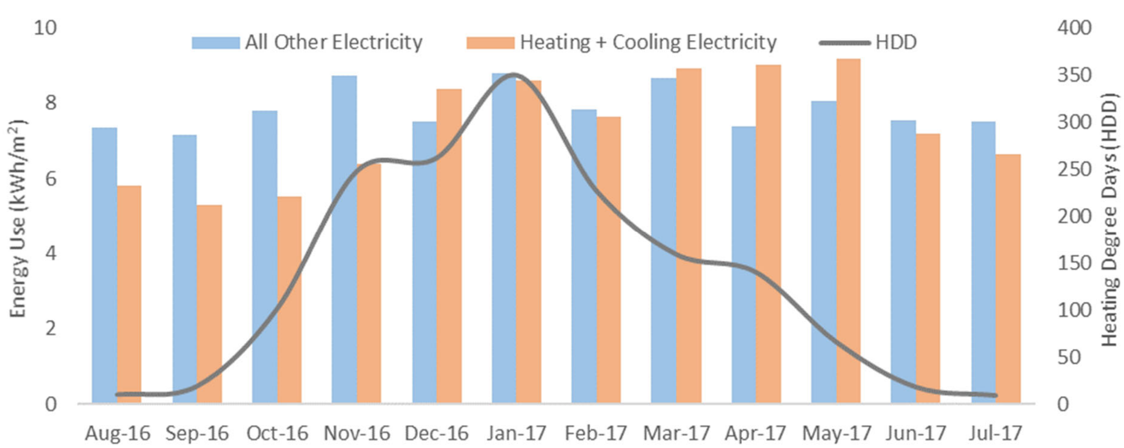


Figure 4.27: Monthly electricity use and gas use for the university office building

Heating and cooling, and lighting and small power are the dominant energy end-uses in the building. While lighting and power were measured separately on some floors, an accurate distinction at the building-level could not be made.

Initial observation, however, showed that occupancy was not consistent with a typical office, varying monthly and seasonally (based on term times), and there was higher out of hours use, including weekends. While the AHU in the basement was not working, all the other machinal systems operated throughout the day and night, even during unoccupied times. This could explain the significantly high (50% more) building energy use when compared to the good practice benchmarks for a similar type of building.

The trends and patterns of actual use could be clearer with granular data which is needed for a higher resolution calibration, as typical weekday and weekend patterns can be detected from this data. In the next chapter, detailed monitoring data is analysed to assist in calibration. Also in Chapter 6, which describes hourly calibration, energy performance issues have been elaborated in more detail.

4.5.3 Building IEQ performance

The case study building intended to explore a high granularity calibration process; therefore, the extent of IEQ monitoring was limited to Temperature and RH and indoor CO_2 concentrations, as they would help in the calibration process. This section reports the results, presented in box plots for each parameter, for two open-plan offices, one cellular office and the external space.

Thermal comfort

The building has a VRF based space conditioning system, and therefore can be heated, or cooled as needed. Figure 4.28 shows that there were no significant issues with thermal comfort in the building. Temperatures were within 20°C and 25°C and RH between 30% and 60%. The recorded temperatures were therefore 2-3°C higher than typically assumed set points.

Indoor air quality

As the building is naturally ventilated, fresh air is predominantly provided via the perimeter windows. Figure 4.29 shows that CO_2 levels in the spaces monitored were significantly lower than the Category 2 requirement. Some seasonal variation is seen in CO_2 levels, as the concentration levels are higher during winter months, when windows are kept closed.

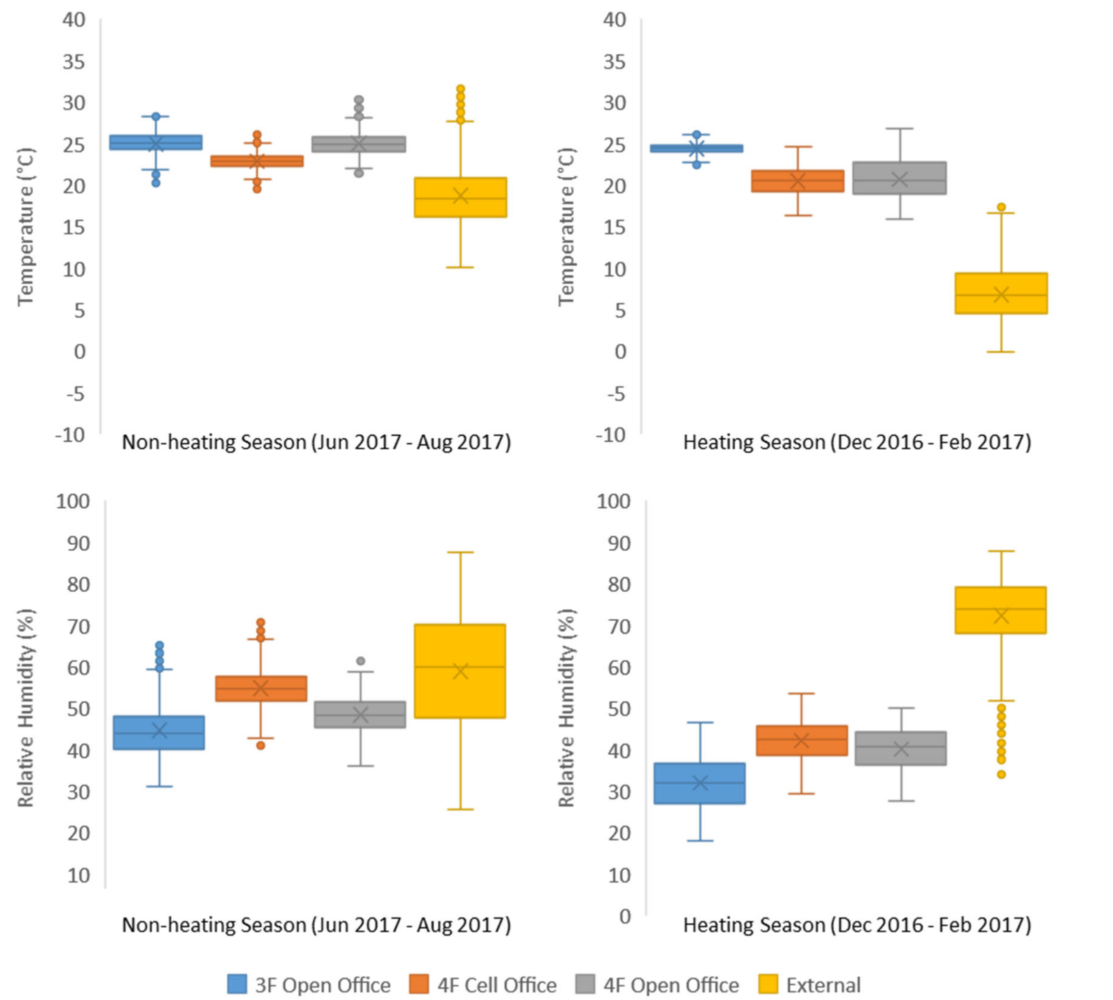


Figure 4.28: Box plots showing the spread of Temperature and RH during non-heating and heating seasons in the university office building

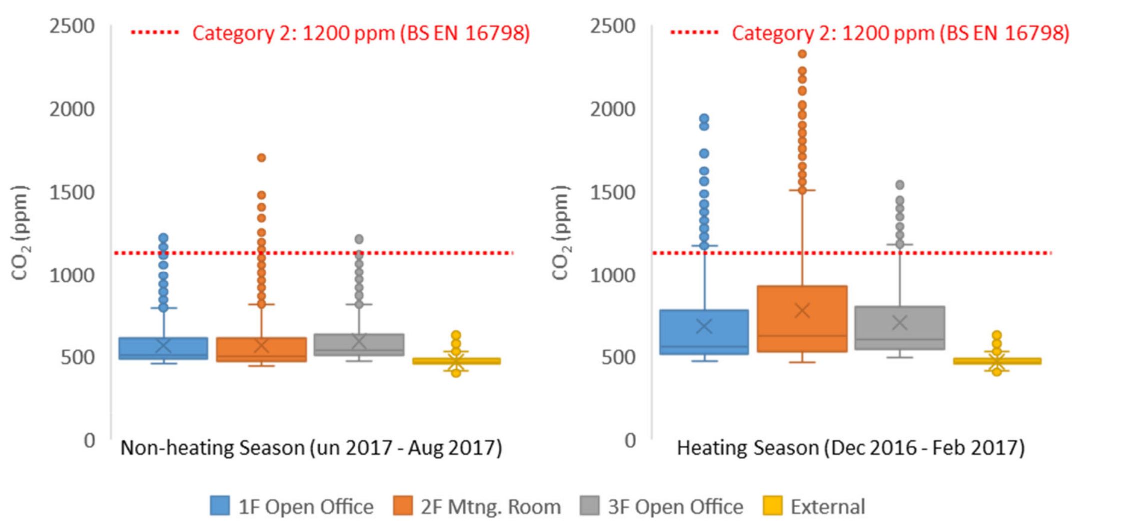


Figure 4.29: Box plots showing the spread of CO_2 conc. during non-heating and heating seasons in the university office building

4.5.4 Observed deviation areas and issues

During the operational performance assessment of the building, many deviation areas of operation and technical issues were found. These issues were discussed in the last two sections and are summarised here:

- Occupancy varies monthly and seasonally (based on term times).
- Higher out of hours use including on weekends.
- The AHU in the basement was not working.
- Systems operated throughout the day and night, even during unoccupied times.
- 2-3°C higher than typical (20-22°C) set-point temperatures were maintained in the building.

The issues listed here inform further data analysis and the model calibration, which is presented in the next two chapters (monthly calibration in Chapter 5 and hourly calibration in Chapter 6), along with a more detailed evaluation and identification of these and other deviation areas.

4.6 Summary

The case studies consist of five buildings in four categories: two offices, one school, one hospital and one apartment block. Analysing issues and improvements in these buildings relating to design, construction, operations and management can provide significant cross-sectoral learnings.

Data for design and operation was collected for the case studies. Design-stage information was collated from the design and construction documentation to understand the building procurement, design, construction, commissioning and handover process. Regular measurements, observations and feedback from the facility managers, at monthly or bi-monthly intervals over a period of one year, were used to collect post-occupancy data and information. Actual data collected during site visits recorded metered energy use and documented building technical and operational details. Actual IEQ performance was captured through regular monitoring of typical zones for Temperature, CO_2 concentrations (a proxy for fresh air), NO_2 (primarily a traffic-driven pollutant) and $PM_{2.5}$. This detailed building information and data collection allows for creating exhaustive building simulation models and model calibration.

The actual performance of all the buildings reported for one calendar year shows deviations from the intended design levels. The use of regulatory compliance calculation results as the design-stage predicted performance for performance gap comparison is not correct. When compared against actual energy use, the resultant gap is often inflated due to the underestimation of energy results of regulatory calculation as opposed to realistic design stage calculations. This was seen in the school, the hospital and the apartment block. These calculations were based on simplified methods and did not factor in building specific operations and all energy end-uses. As a result, the actual energy use was more than 2-4 times what was predicted using the compliance calculations. While there are some under-performance issues in the public office building, it is by far a very high-performing building in its category. Part of this is due to the performance contracting procurement process followed in the project. This process guaranteed the involvement of the contractor and designers after the handover to fine-tune the building and ensure low carbon objectives were met. Performance contracting also integrated a Soft Landings Framework, which promotes extended involvement of the design and contracting team, helping to ensure that these objectives are met.

In all the case study buildings, energy performance was the only specific quantified performance objective. There was no metering, monitoring and reporting strategy for any of the IEQ parameters planned for regular operational stage evaluations. However, as discussed earlier, lower energy and better IEQ can be divergent objectives. The temperature, CO_2, NO_2 and $PM_{2.5}$ measurements done in this study provided insights into the IEQ performance and suggested their potential energy-related implications. Conflicts within IEQ objectives for air quality and between energy and IEQ for thermal comfort, especially in new low-energy buildings constructed in polluted urban areas, were observed.

Table 4.12 and Table 4.13 summarise the energy and IEQ performance respectively for all the case study buildings. Table 4.14 lists all the performance deviations seen during initial observations and data collection. These deviations will form the basis of detailed deviations assessments and model calibration in the next chapter.

Table 4.12: Energy performance summary of all case studies

Building	Fuel	Design*	Actual*	Benchmark*	Remarks
Public Office	Gas	18.9	28.9	79.0	• The design calculations were done to determine the performance targets and went beyond the restrictions of regulatory compliance calculations. • Due to the use of performance contracting integrated within the Soft Landings Framework, performance targets were kept in sight and then compared favourably against the benchmark.
	Electricity	57.0	67.5	54.0	
School	Gas	19.4	168.3	108.0	• The design stage projections were done as part of Building Regulations compliance calculations and severely underestimated the energy use. • Despite a new building, gas and electricity are 1.5 and 2.5 times more compared to the benchmarks.
	Electricity	20.0	88.5	25.0	
Hospital	Gas	83.8	303.5	422.3	• The design stage projections were done as part of Building Regulations compliance calculations and severely underestimated the energy use. • Hospital's high energy use intensity is primarily for functional needs, so benchmark comparisons are more academic than realistic.
	Electricity	35.0	113.8	74.4	
Apartment Block	Gas	40.8	113.4	247.0	• The design stage projections were done as part of regulatory compliance SAP calculations and underestimated the energy use. • Actual performance, which depends a lot on occupants in apartments, is lower than the benchmarks.
	Electricity	6.2	35.5	44.0	
University Office	Electricity	N.A.	183.0	133.0	• Being a refurbishment project, design stage assumptions and calculations were not available for the building. • Actual performance is 1.5 times the benchmark. One of the reasons is the building's atypical usage pattern when compared to other offices.

	Deviation is less than benchmark					Deviation is more than benchmark			
Legend:	>75%	50% - 75%	25% - 50%	Up to 25%	Benchmark	Up to 25%	25% - 50%	50% - 75%	>75%

*All values are in kWh/m^2/yr.

Table 4.13: IEQ performance summary of all case studies

Building	Thermal Comfort	Indoor Air Quality
Public Office	• Summer overheating is a concern because of no space cooling. There were few instances of high indoor temperatures during hot summer spells.	• IAQ in the spaces monitored was under their respective thresholds; however, there were episodes of high CO_2 levels in one of the offices due to malfunctioning vents.
School	• Summer overheating risk is a concern in rooms on south facing facades due to a lack of solar controls.	• IAQ in the spaces monitored was under their respective thresholds however there were some issues. • CO_2 level in one of the classrooms was above 1500 ppm because of the failure of the supply fan in the AHU. • NO_2 levels in the library, located next to the main road, although lower than the WHO recommended annual mean, are higher than the other zones.
Hospital	• The building operated with close control over indoor environment and had no significant thermal comfort issues.	• Filtered fresh air is provided and due to a high level of control, IAQ was generally under their respective thresholds. • There was a lack of measures against ambient NO_2. Indoor NO_2 levels recorded, while not exceeding WHO limits, very closely followed the external levels.
Apartment Block	• As there is no comfort cooling, summer overheating can be a concern. Indoor temperatures were often recorded to be outside comfort ranges in both summers and winters, but that can be a user comfort preference.	• Concentration levels of IAQ parameters are kept below the guideline limits mostly. • However, there were some filtration issues with the MVHR system. $PM_{2.5}$ levels in the flat closely follow external levels and were often higher than WHO limits.
University Office	• Due to the VRF based space conditioning system, building can be heated, or cooled as per need. There were no significant thermal comfort issues in the building.	• Fresh air, provided via the perimeter windows, keep the CO_2 levels in the monitored spaces generally lower than the thresholds.

Table 4.14: Identified deviations in the case study buildings

Building	Issues
Public office	• The total occupancy of the building was about 25-30% higher, leading to more workstations being added and longer use of small power and lighting. • Technical issues with heat exchangers and the flow rates specification caused the heat pumps to malfunction. As a result, the gas-fired boilers supplied all the heating. • Server loads were overestimated in design. This adversely impacted the heating system efficiency as there was less free heat. • Heating set-point temperatures maintained were about 2°C higher than the design intention, driven primarily by occupant needs. • Some of the automated ventilation control sensors malfunctioned.
School	• All spaces were not occupied throughout the day. Any given space is occupied only 60-70% of the full working day. Also, during academic breaks, the school is not completely shut. There are extra-curricular activities and events, especially during the summer holidays. • Biomass boiler, installed for decarbonising energy use, was never used, citing practical and logistic issues of using biomass as fuel. • Actual boiler efficiency (tested) was 88% compared to the design intent of 96%.

	• Building heating and ventilation systems were working during unoccupied hours and indoor temperatures monitored during the winter season were about 2°C higher than the set-point temperatures and sometimes even more.
Hospital	• Different clinical processes had unique needs. It was difficult to generalise typical operational trends of various spaces; their equipment loads; and set-point temperatures. • Increased number of beds were seen in the patient wards during the site visits. • A low efficiency, old, steam-based central heating network serves the building. A new CHP plant was to be installed for the entire facility rather than just for the new building to maximise savings. However, this has not happened yet. • Hospital hot water energy use was high due to clinical requirements.
Apartment block	• Occupancy of the individual flat is very uncertain to determine, occupants had full control over the heating system and MVHR, determining their set-points and operations. • There are instances where during summer, the windows were opened for enhanced ventilation, while the heating system was still operational. • Maintaining of filters in the de-centralised MVHR systems was left to the occupants and therefore there were maintenance issues (e.g. dirty filters) in most flats.
University Office	• Occupancy varies monthly and seasonally (based on term times) • Higher out-of-hours use including on the weekends. • The AHU in the basement was not working. • Systems operated throughout the day and night, even during unoccupied times. • 2-3°C higher than typical (20-22°C) set point temperatures were maintained in the building.

5 Operational analysis and model calibration (monthly)

This chapter identifies the areas of deviations that happened at the operational-stage, observed during the data comparison stage, which were fine-tuned during the calibration process. The design intent of the projects was captured in design documentation and obtained during data collection. This information was used to develop the calibration-baseline (as-design) model. Analysing the deviation of the model input parameters in this calibrated model against the calibration-baseline (as-design) model revealed key performance gap issues. The specific potential reasons for deviation from the design assumptions were catalogued through on-site observations, analysis of metered and monitored data for energy and IEQ, and feedback from stakeholders.

Once the model is calibrated, the changes made in the calibrated model account for most of the performance gap. Therefore, model calibration confirms the issues as the most important and likely reasons for the gap. For each case study building, this chapter first describes the development of the calibration-baseline (as-design) model. Then the deviations/shortcomings are observed, along with the evidence found during investigations and monitoring that could contribute to the underperformance. Finally, monthly calibration results, as per the evidence-based step by step method defined in Section 3.3.2, are presented.

5.1 PUBLIC OFFICE

5.1.1 Calibration-baseline (as-design) model

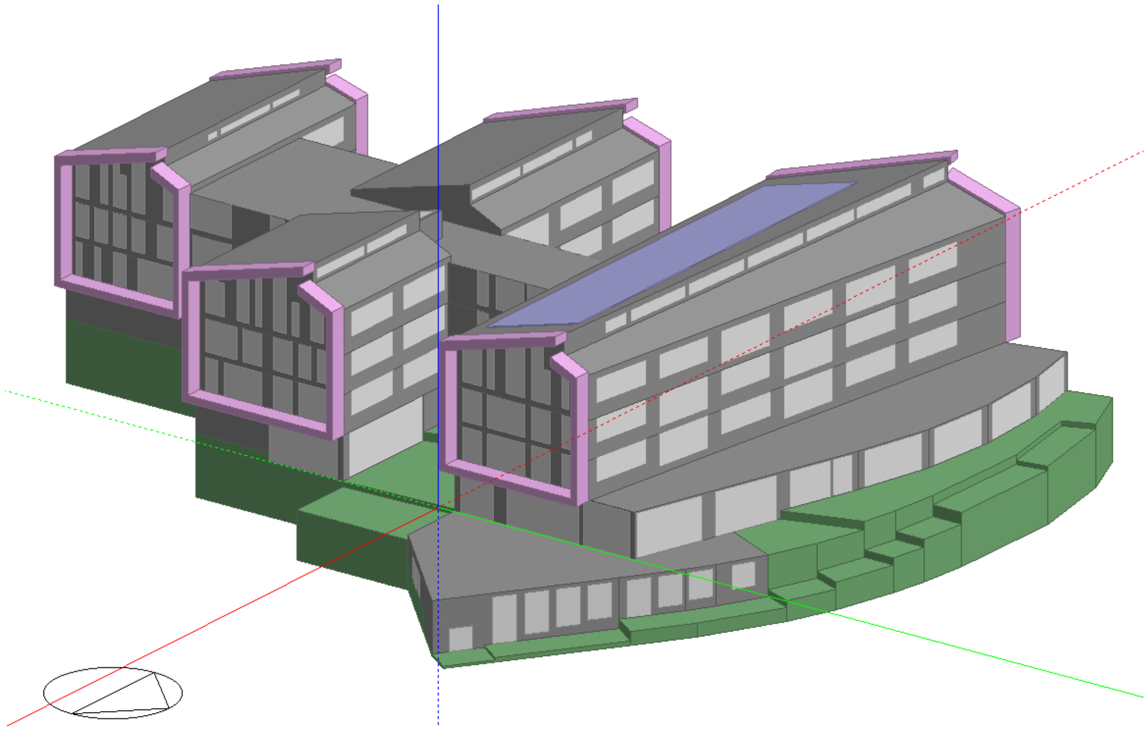


Figure 5.1: Model visualisation of the public office building

As a first step towards calibration, a calibration-baseline (as-design) model (Figure 5.1) was created as per the details obtained from the design data in the last chapter, described in Section 4.1.1. Most of the design information and design intent was well documented; therefore, a model could be developed with that information. Details for the architectural design, fabric, occupancy, loads and HVAC system were taken from design-stage documentation and calculations. Modelling of the HVAC system was done as a simplified system that calculates heating and cooling loads as per EnergyPlus ZoneHVAC:IdealLoadsAirSystem and factors in all operation control and overall efficiency. Design weather data, from International Weather for Energy Calculation (IWEC), for Birmingham (World Meteorological Organization - WMO: 035340) was used as the nearest weather station to the site.

5.1.2 Deviations

Analysing the information gathered from operational stage investigations into the building's performance, this section explains the deviations of building design and operations from the initial design intent.

Simulation weather data

Actual weather as per station and satellite measurements for Bristol WEA Centre (WMO: 37260) for the calibration period was obtained from DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b) and was used in the simulations.

Building design and construction

The overall building design and construction specifications were as per expectations for a newly built building. However, thermal imaging of the envelope (Figure 5.2) revealed leakage through poorly sealed vents and at construction junctions. The observed heat loss was higher than anticipated, especially from the doors installed behind natural ventilation louvers.

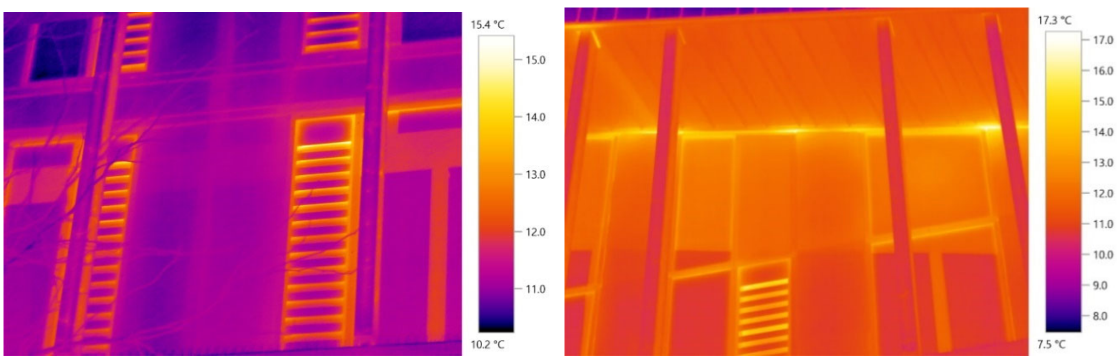


Figure 5.2: Heat loss from the doors installed behind natural ventilation louvers in a closed position (left), and at construction junctions (right)

Occupancy and occupant behaviour

The total occupancy of the building was about 25-30% higher than panned, leading to more workstations being added. Initially designed for 455 occupants, the building was being used by 550-600 people. Additionally, there were longer occupancy hours in the building; than were assumed in the design calculations. This partially contributed to the higher small power and lighting energy use. Figure 5.3 shows the net electric demand of the building. The dip during the day in electric demand is when the Solar PV system is meeting part of the load. Therefore, start and end times can be better inferred during winter months, where the building loads start increasing around 6 am and get back to base levels around 8 pm to 9 pm. This is consistent with the on-site observations.

The implications of this were also affecting the contracted performance targets. Occupancy was a critical risk in the pre-tender risk register. Use of a smart card to activate thin client computers or people counter cameras were considered to record occupancy if extended hours were to be used in DEC assessment. However, subsequent documents do not show these strategies being taken forward, thus making it difficult to create a more accurate benchmark for DEC calculation.

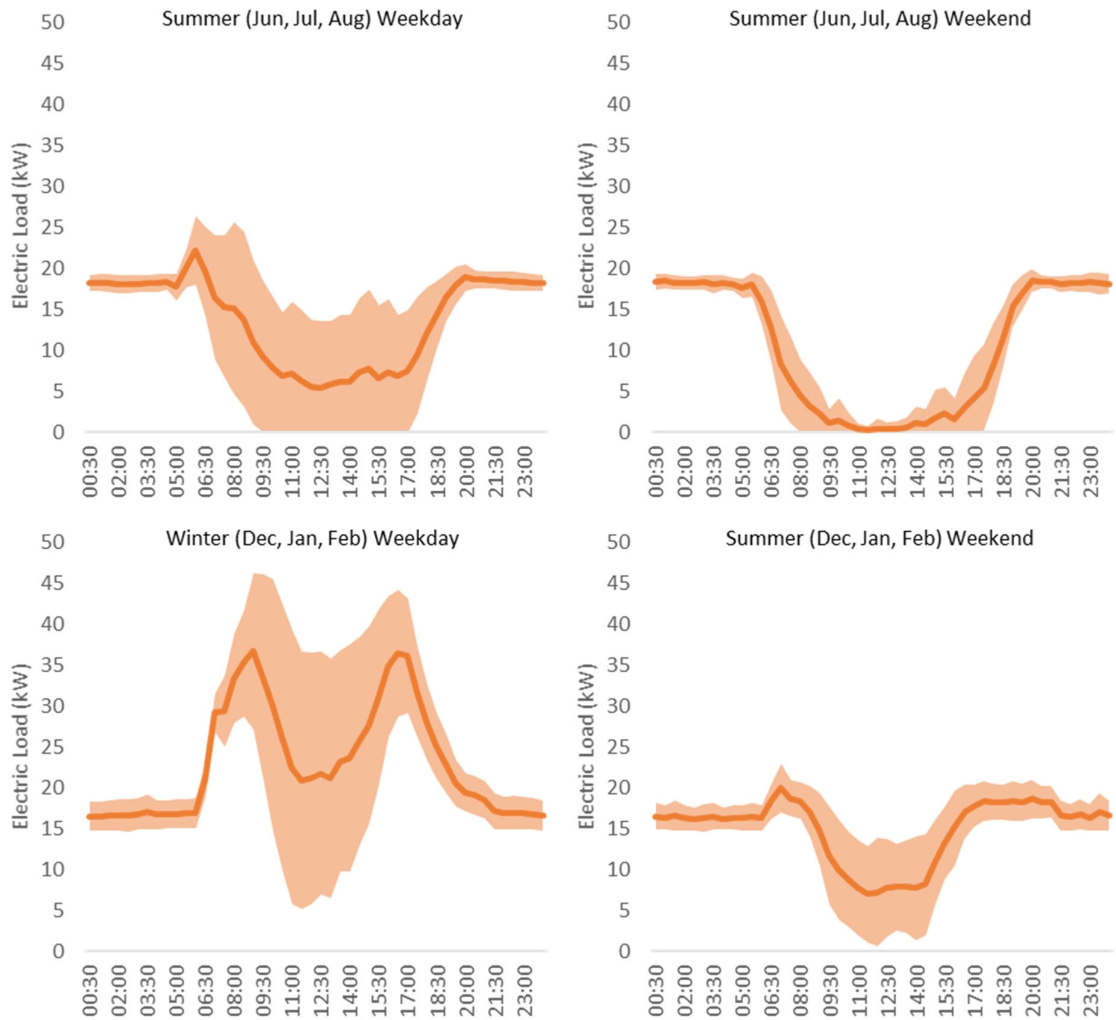


Figure 5.3: Average electrical demand and its standard deviation for the office building, for summer and winter months

Hot desking, using a thin-client IT system, was initially planned for optimum space-time use of the building by using flexible workstations and hydraulic isolation of heating and cooling systems in unoccupied zones. However, this was not followed in practice. The departmental structure of the occupant organisation restricted movement and led to inefficiencies in building services operation during out-of-hours building use.

Equipment, lighting and other loads and their operation

High parasitic loads were identified when the building was unoccupied. Figure 5.3 shows the high level of baseload in the building, during unoccupied times. However, this issue remained hidden for a long time as the BMS integration of the meters was not undertaken correctly, making it difficult to identify parasitic loads by monitoring the disaggregated energy use.

The original design concept allowed for a server cooling load of 29 kW. Whilst the larger load may occur in the future, the current connected load is only 15 kW. Therefore, the electricity used for servers is significantly less than the design estimates. However, if heat pumps provided cooling (which is not the case, as explained in the next part), there would be an adverse impact on the efficiency of the heat pumps as there would be significantly less free heat available.

HVAC system design and operations

The building has a complex interdependent heating and cooling strategy. The evidence suggests that the heat pumps have not been highly effective as a primary source for heating and cooling because of technical issues with the buffer vessel's heat exchangers and the flow rates. Consequently, the entire heating load of the building has been shifted to the gas boilers, and there is no active comfort cooling being provided. Server cooling is also provided by two backup unitary direct expansion chillers (DX chillers). A review of the technical specification of the heating terminals suggests that the sizing of these terminals is not consistent with the low-temperature heating flow required for energy efficient operation of the heating system.

On the ventilation front, there is some overriding of automatic vent opening to avoid thermal discomfort due to drafts around the floor cut-outs. Some of the ventilation control sensors also malfunctioned and required a subsequent modification to the control strategy (by increasing the set-point temperatures) to overcome the system shortcomings.

On the operation front, heating set-point temperatures were about 2°C higher than the design intention of 19°C, driven primarily by occupant needs. This can be seen in the monitored temperature graphs (Figure 5.4) and also in the BMS snapshot (Figure 5.5) taken during the heating season.

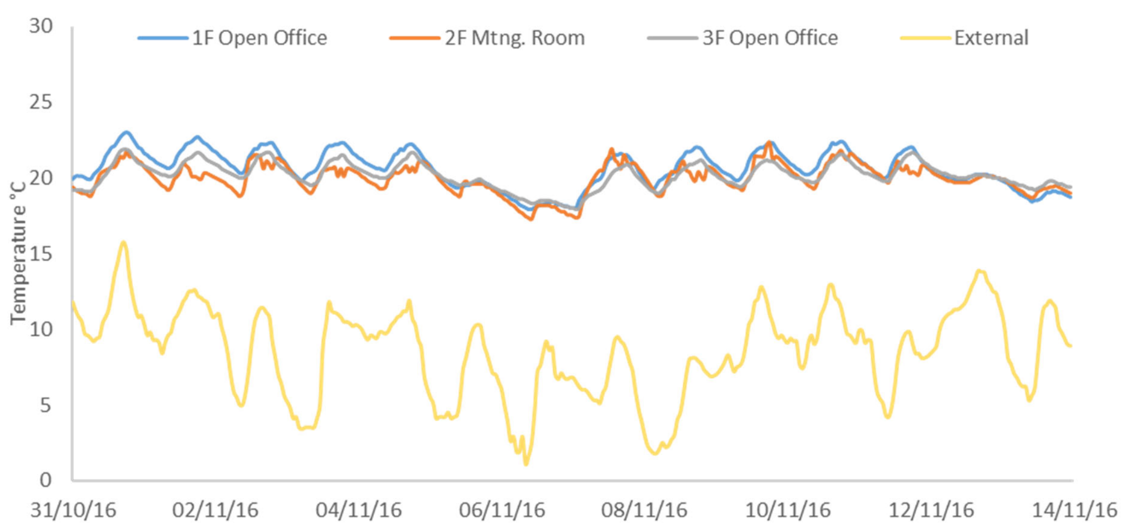


Figure 5.4: Monitored space temperature during the heating season in the public office building

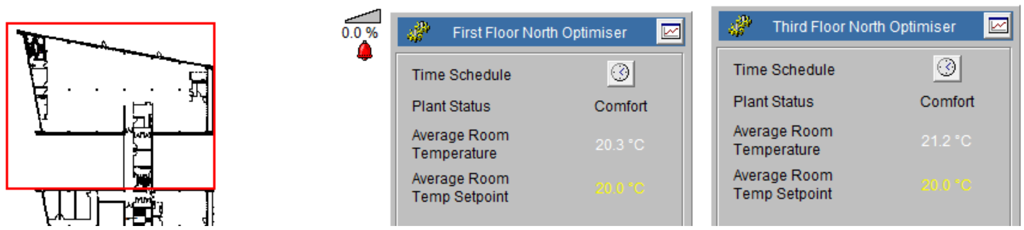


Figure 5.5: BMS screenshots from the public office building showing the average temperatures and set-points in the first- and third-floor open offices

5.1.3 Model calibration results

Figure 5.6 compares the monthly simulation results against the actual measured data for electricity and gas use for the building and reports the key statistical error values. By using the correct weather data, operational assumptions, loads and system configuration, this building's monthly energy use profile was estimated with an acceptable accuracy as per ASHRAE Guideline 14, i.e. C_V(RMSE) < 15% and NMBE < ±5%.

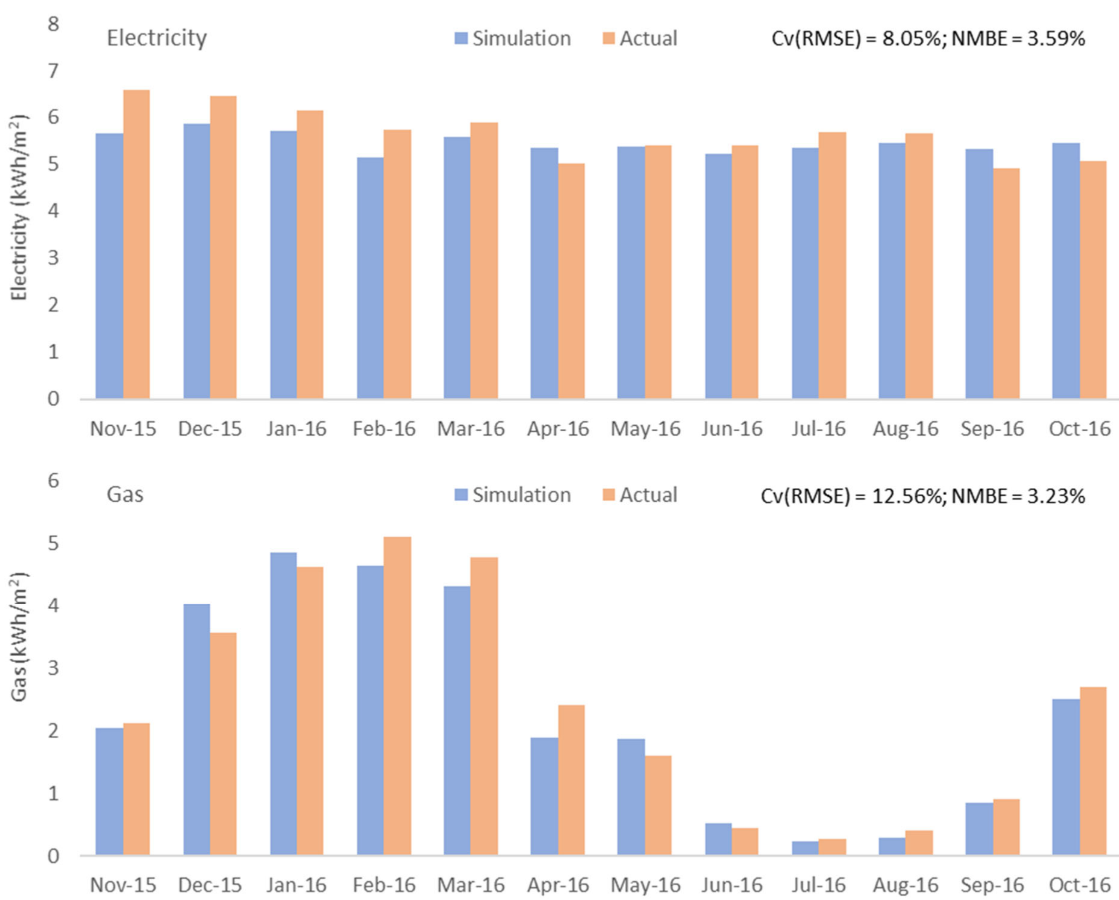


Figure 5.6: Monthly calibration results for electricity and gas use for the office building

Further, to ensure that the calibrated model reasonably reflected the actual building's operation, Figure 5.7 shows the temperature trends for typical days for a sample space in the office building. The simulated temperatures closely follow the measured air temperature, except for a few deviations. For example, there is a dip in simulated temperatures during the weekends. Also, the modelled radiant temperature is sometimes less than the modelled air temperature due to the poor thermal envelope and air leakage through the vents and heat losses at construction junctions (Figure 5.2).

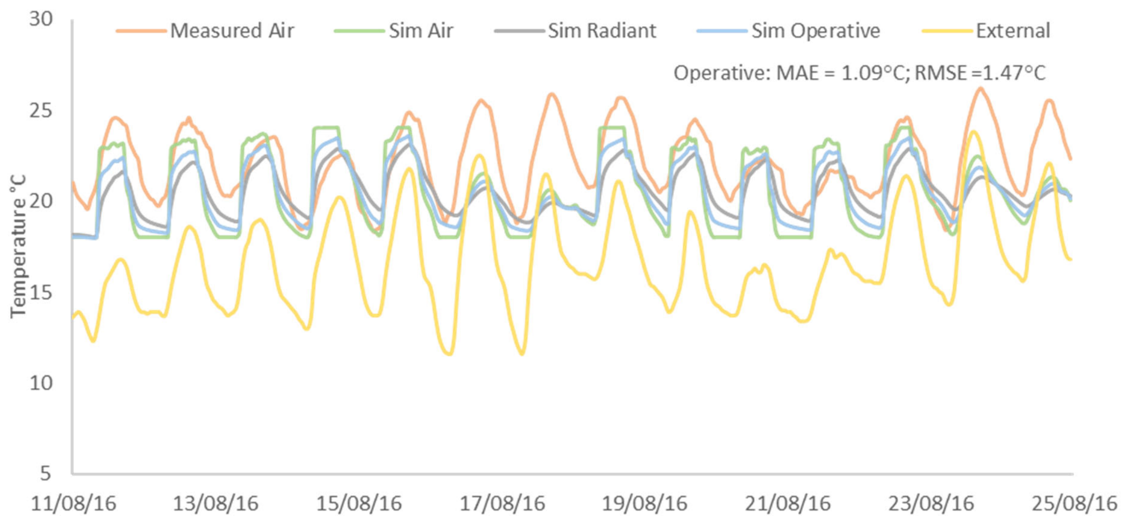


Figure 5.7: Hourly simulated and measured temperature profiles for typical days for a sample space in the office building

5.2 SCHOOL

5.2.1 Calibration-baseline (as-design) model

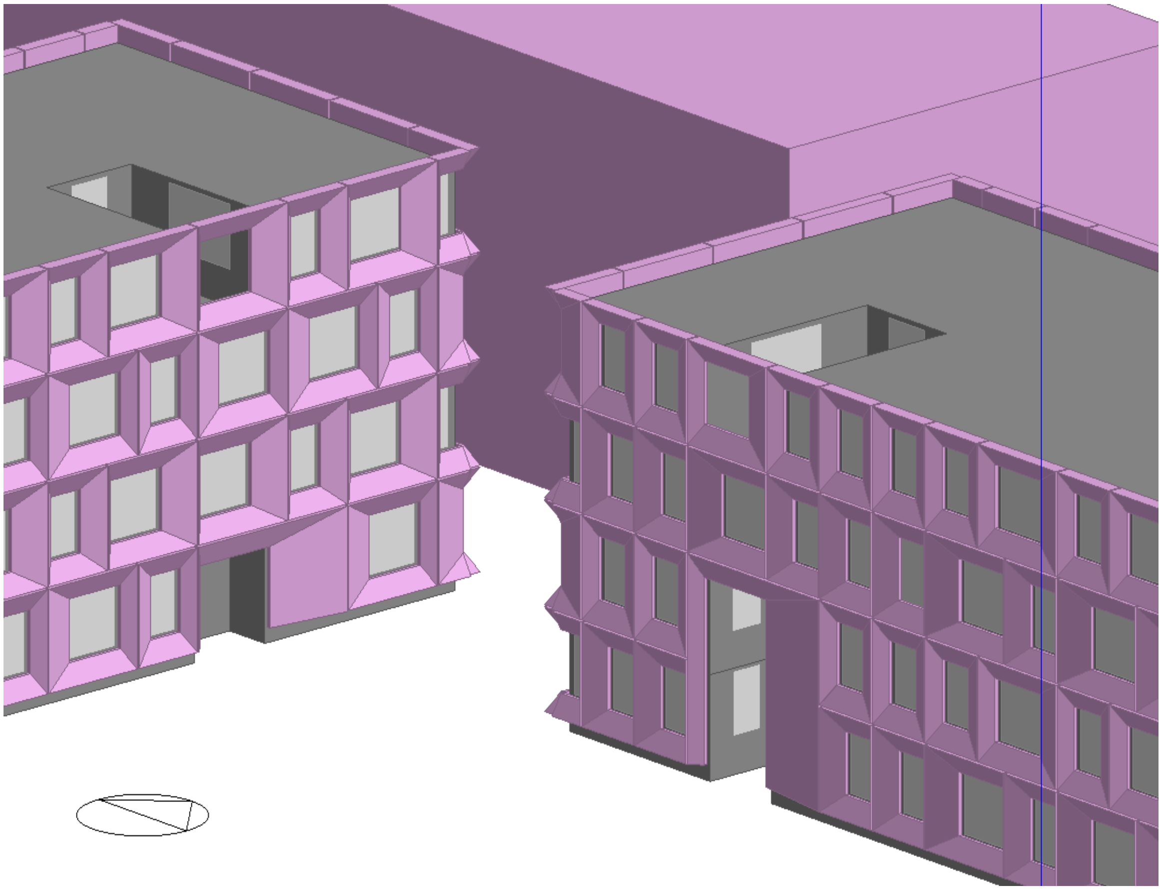


Figure 5.8: Model visualisation of the school building

As a first step towards calibration, a calibration-baseline (as-design) model (Figure 5.8) was created as per the design-data described in the last chapter, described in detail in Section 4.2.1. Most of the design information and design intent was well documented and was also available from the design calculations that used UK NCM defaults. The information that was not directly available was inferred from other data or what was happening on-site. Design data collected covered details for architectural design, fabric, occupancy, loads and HVAC system. Modelling of the HVAC system was done as a simplified system that calculates heating and cooling loads as per EnergyPlus ZoneHVAC:IdealLoadsAirSystem. The modelling factors in all operation control and overall system efficiency. Weather data, from IWEC, for London Gatwick (WMO: 37760) was used as the nearest weather station to the site.

5.2.2 Deviations

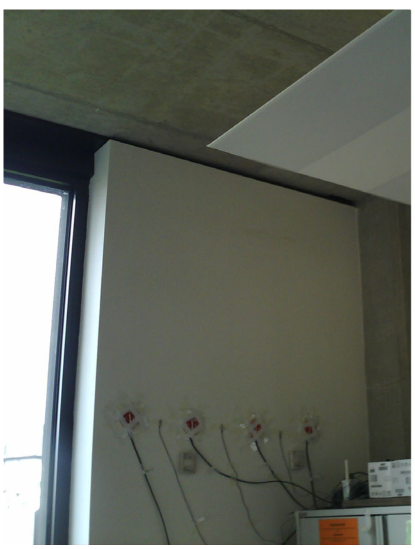
Analysing the information gathered from operational stage investigations into the building's performance, this section explains the deviations of building design and operations from the original design intent.

Simulation weather data

Actual weather as per station and satellite measurements for Kew-in-London (WMO: 37750) for the calibration period was obtained from DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b) and was used in the simulations.

Building design and construction

The overall building design and construction specifications were as per expectations for a newly built building. However, significant thermal bypasses around insulation were observed as the tested U-values were much higher than the design. In-situ U-value measurements on a single north-facing section of the façade in the school building (Figure 5.9) found U-values between 0.72-0.78 W/m²K, a significant increase on design values of 0.17 W/m²K.



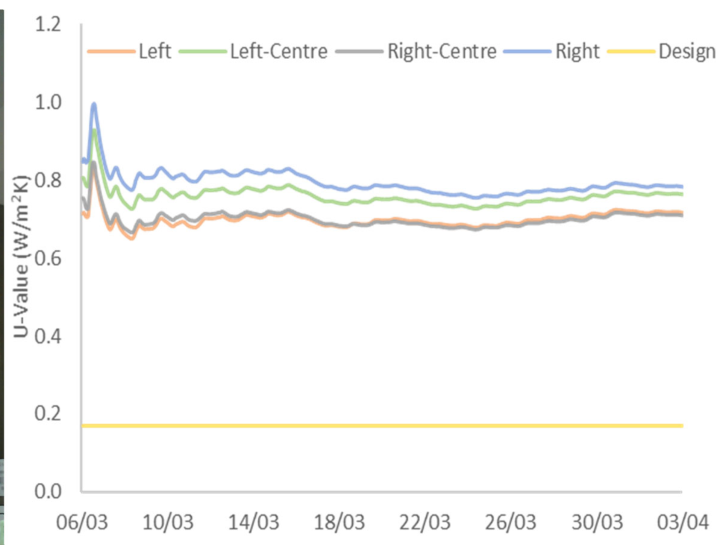


Figure 5.9: In-situ U-value measurements on a single north-facing section of the façade in the school building

This was also seen in thermal images (Figure 5.10) taken during winter, when heat losses through the envelope of precast concrete panels and the frame were much more than through the windows.

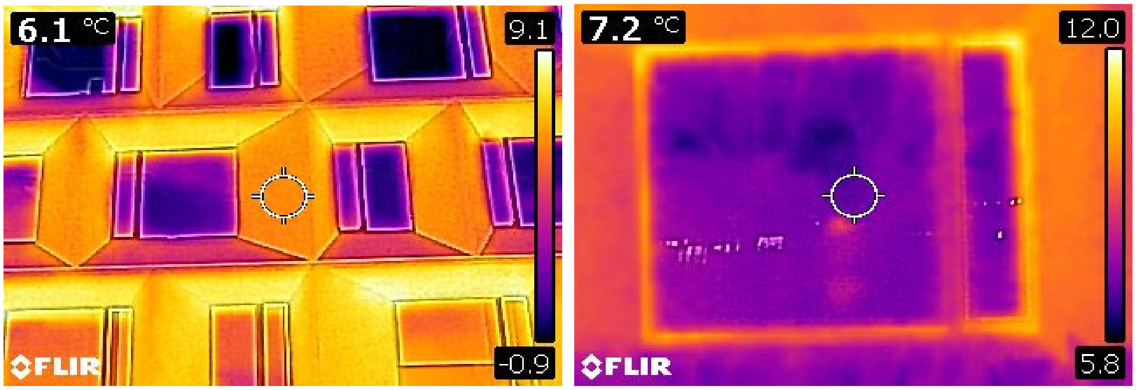


Figure 5.10: Heat loss from the precast concrete planes in the façade (left), and at window frames (right).

Occupancy and occupant behaviour

Occupancy in classrooms during term time is as per class timetables. It was observed that not all spaces were occupied throughout the day. Any given space was occupied only 60-70% of the working day. Therefore, this variability in occupancy and the subsequent operation profile for internal loads need to be accounted for. During term breaks the school was not completely shut; extra-curricular activities and events took place, especially during the summer holidays. Figure 5.11 shows the increase in electric load during the daytime in August, when the school had summer holidays. The site-level daytime electrical demand in August reached around 100 kW, when the baseload was 50 kW. This was verified by on-site observations. Monitored CO_2 concentrations (Figure 5.12) also show increased levels beyond the regular school hours, sometimes beyond 6 pm on weekdays, indicating that some evening classes were occurring. Also, some occupancy was seen on Saturdays as well. This is similar to the low-level occupancy trends seen during term breaks.

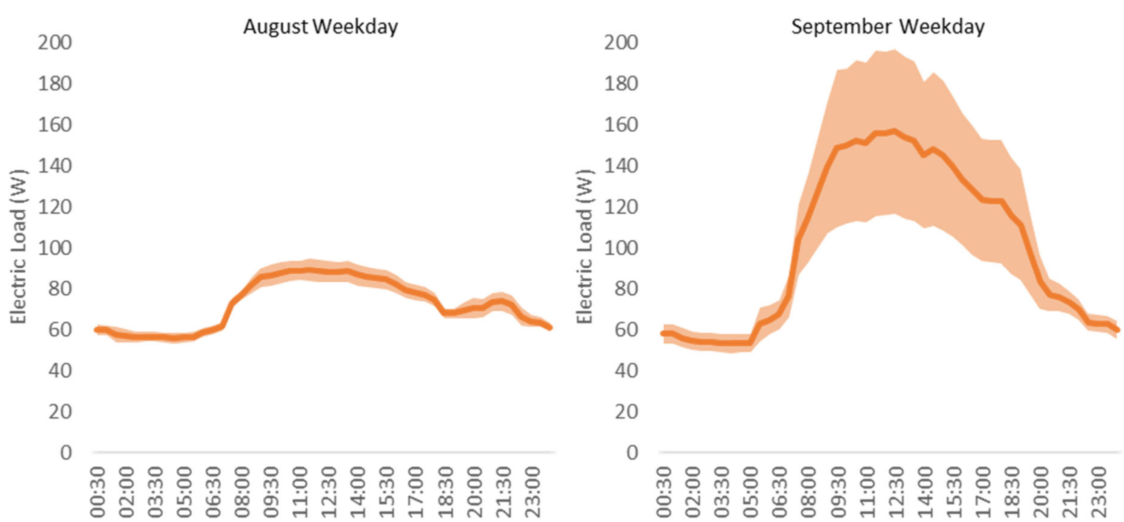


Figure 5.11: Average electric load and its standard deviation for August (holiday time) and September (term time)

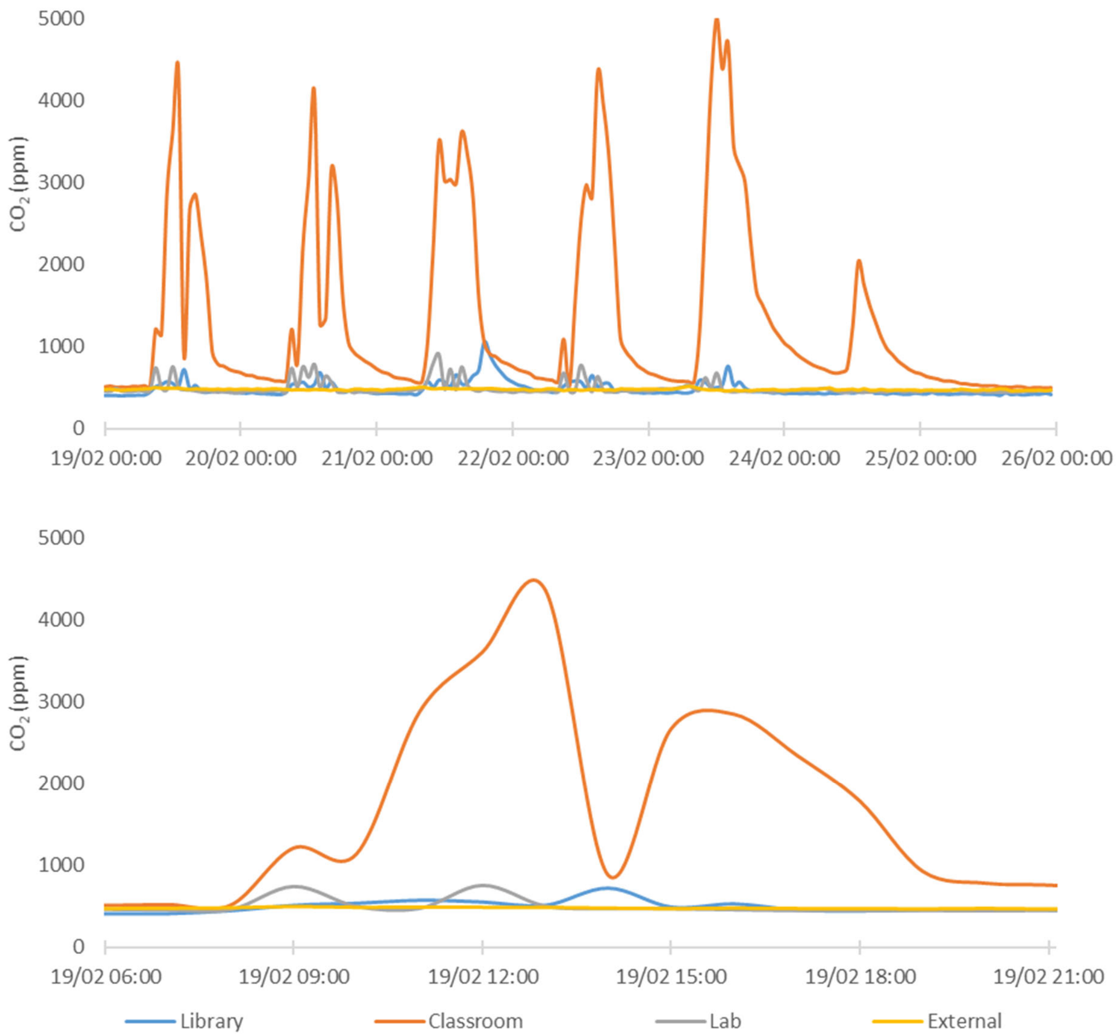


Figure 5.12: CO_2 levels in school spaces during a typical week (above) and typical day (below)

Equipment, lighting and other loads and their operation

It was seen during site visits that lights in the circulation areas were switched on even after the end of the classes. Daylight sensors for automatic dimming of lights were poorly commissioned and PIR sensors had very long delay times. Also, many computers were found to be 'on' when computer rooms were not occupied.

HVAC system design and operations

A biomass boiler was designed and installed to provide maximum 50% of the total DHW demand, with the intent of decarbonizing energy use, as a policy measure by the local council. However, the boiler was never used by the facility managers, citing practical and logistical issues of using biomass as fuel. This, while not impacting net energy use by the building, led to an increase in net CO_2 emissions due to the use of gas-fired boilers, as the CO_2 emissions conversion factor of biomass (0.025 kg CO_2/kWh) is significantly lower than that of natural gas (0.19 kg CO_2/kWh). Also, the gas-fired boiler's actual efficiency (tested) was 88%, compared to the design intent of 96%.

The design intent of the boiler being turned on three hours before classrooms were occupied until the end of classes was not happening in the building. During the site visits, it was noted that the system was on even after the end of classes. The indoor monitored temperature data (Figure 5.13) shows that the building heating system and mechanical ventilation systems were functional during unoccupied hours as well, both during term time and during holidays. Also, it can be seen that during the winter season the indoor temperatures throughout the day were in the range of 21-22°C (about 2°C higher than the design set-point temperatures) and sometimes a little higher.

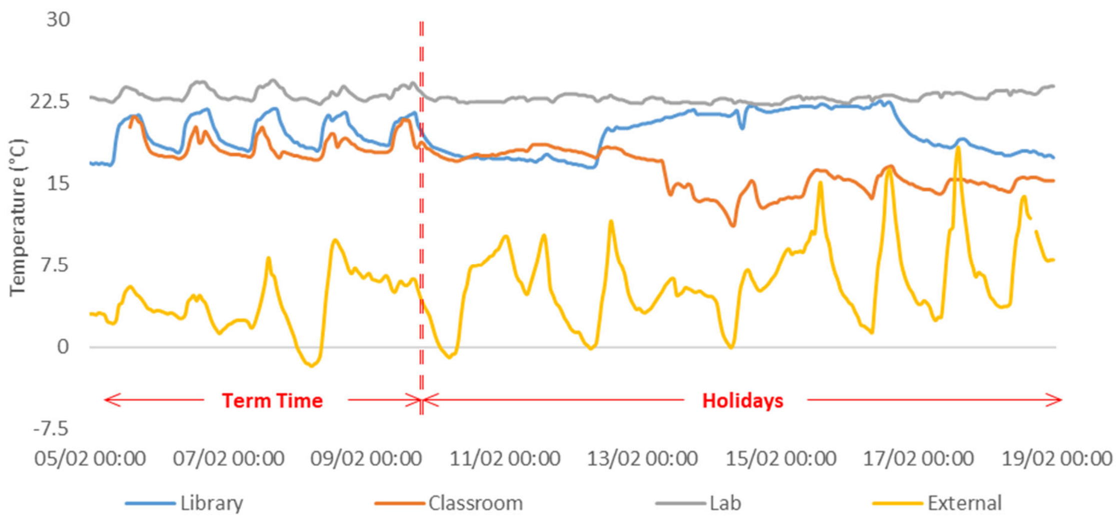


Figure 5.13: Temperature variations during term time and holidays in the heating season

Optimum space-time utilisation of educational facilities is key to optimise energy performance. For example, only a few thermal zones of a school often need to be conditioned for extra-curricular activities or night classes. Centralised system design (one control and sensor for multiple zones) led to inefficient operations during out-of-hours use. Conditioned fresh air was provided in the building by a centralised air handling unit. During out-of-regular-hours and half-term breaks, when there is very low occupancy, conditioned fresh air and heating is served to multiple zones.

The Specific Fan Power in the AHU commissioning sheets was 66% higher than the values used in design stage estimations. Also, during unoccupied times the supply fan was running inefficiently, operating at 30% to 40% of its nominal speed.

5.2.3 Model calibration results

Figure 5.14 compares the monthly simulation results against the actual measured data for electricity and gas use for the building. By using the correct weather data, operational assumptions, loads and system configuration, this building's monthly energy use profile was calculated with an acceptable accuracy as per ASHRAE Guideline 14, i.e. C_V(RMSE) < 15% and NMBE < ±5%.

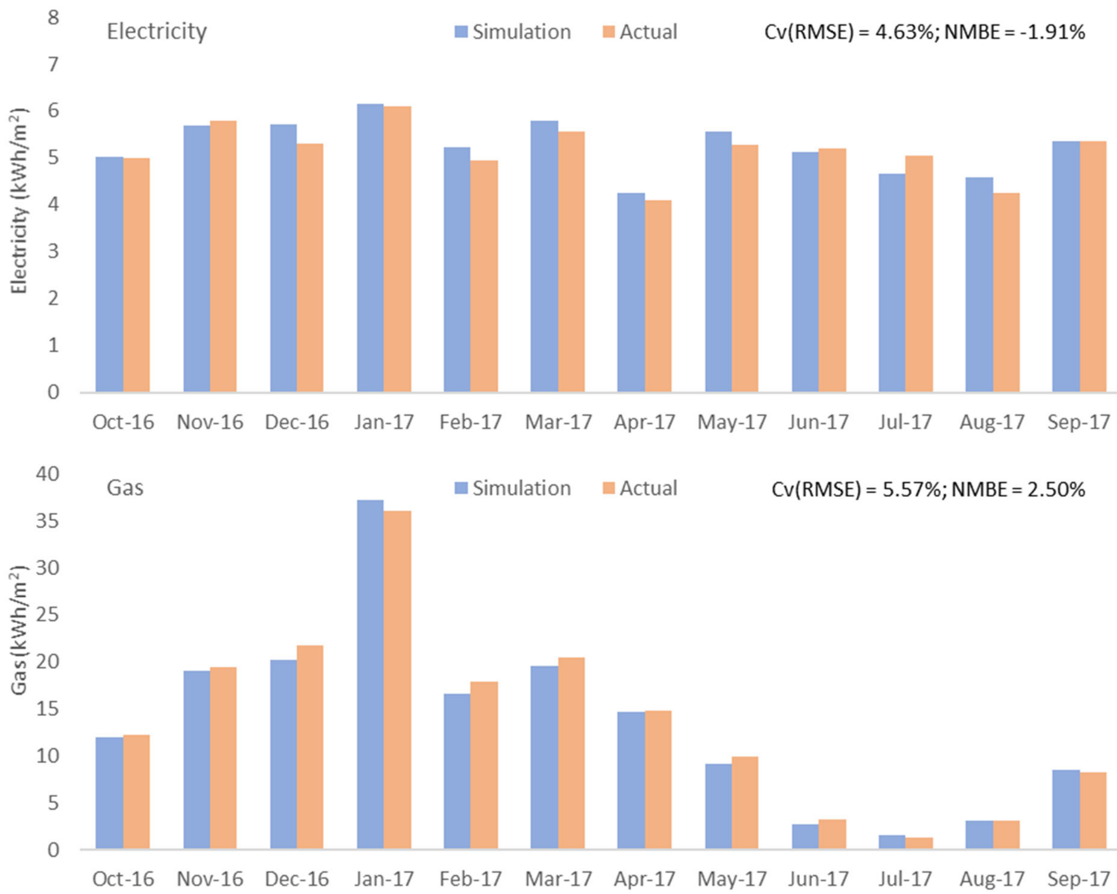


Figure 5.14: Monthly calibration results for electricity and gas use for the school building

Further, to ensure that the calibrated model reasonably reflected the actual building's operation, Figure 5.15 shows the temperature trends for typical days for a sample space in the school building. The simulated temperatures closely follow the measured air temperature, except for the few instances where there are some deviations. For example, overnight simulated temperatures over the weekends in the school are lower than the monitored values. Also, the modelled radiant temperature is sometimes less than the modelled air temperature due to the poor thermal envelope caused by thermal bypasses around insulation, observed and verified with U-value measurements (Figure 5.9), and the heat losses observed in thermal imaging (Figure 5.10).

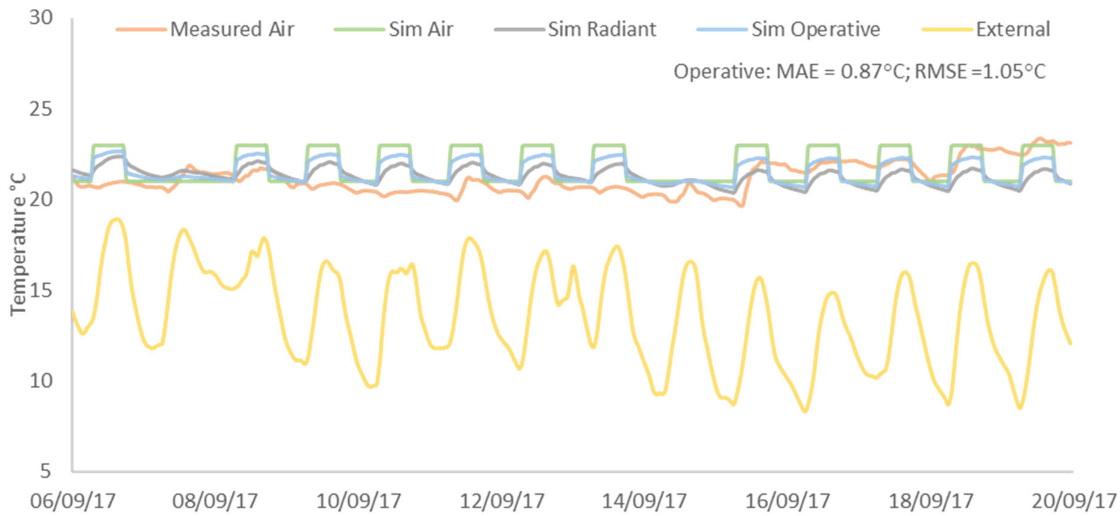


Figure 5.15: Hourly simulated and measured temperature profiles for typical days for a sample space in the school building

5.3 Hospital

5.3.1 Calibration-baseline (as-design) model

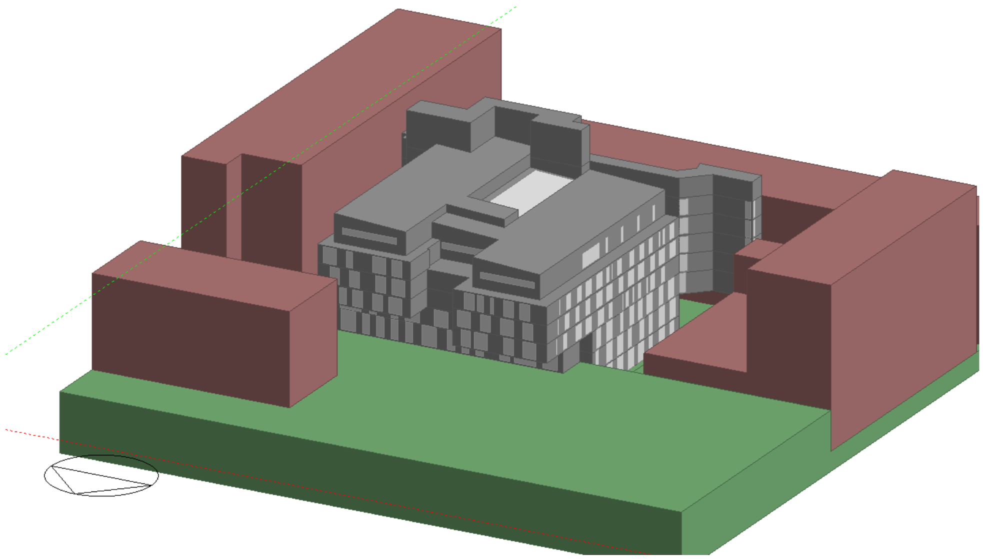


Figure 5.16: Model visualisation of the hospital building

As a first step towards calibration, a calibration-baseline (as-design) model (Figure 5.16) was created as per the details obtained from the design data described in the last chapter, in Section 4.3.1. Most of the design information and design intent was well documented; therefore, the model could be developed based on that information. Details for architectural design, fabric, occupancy, loads, and HVAC system were taken from the design stage data collected and design calculations that used UK NCM defaults. Modelling of the HVAC system was done as a simplified system that calculates heating and cooling loads as per EnergyPlus ZoneHVAC:IdealLoadsAirSystem and factors in all operation control and overall efficiency. Weather data, from IWEC, for Birmingham (WMO: 035340) was used as the nearest weather station to the site.

5.3.2 Deviations

Analysing the information gathered from operational-stage investigations into the building's performance, this section explains the deviations of building design and operations from the initial design intent. Information about some of the inputs were known more precisely than others. While the building fabric and occupancy profiles of 24/7 use were certain, data for inputs such as exact operational trends of various spaces, their equipment loads and equipment operation, as well as set-point temperatures, were highly uncertain. The uncertainty in these inputs arose because the irregular nature of processes in hospitals makes it difficult to identify typical events and average durations of use. The transient nature of the building occupants and their needs further added to this uncertainty. To develop the calibrated model, best estimated values for inputs were used, catalogued through on-site observations, walkthrough audits, analysing metered and monitored data for energy and IEQ, and stakeholder interviews.

Simulation weather data

Actual weather as per station and satellite measurements for Bristol WEA Centre (WMO: 37260) for the calibration period was obtained from DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b) and was used in the simulations.

Occupancy and occupant behaviour

The building is occupied 24/7; however, different clinical processes have different needs, and the irregular nature of the processes makes it difficult to describe typical events and average durations of use for various functions. Additionally, hospitals continuously evolve to meet changing needs. This may result in changes to space use, equipment inventory or building services, which might not be updated in the facility managers' records.

The transient nature of occupancy (patients and some staff) results in limited knowledge of local customs, which could be collected during site visits. While most occupancies and operations are standard, an increased number of beds was observed in some of the patient wards during the site visits.

Equipment, lighting and other loads and their operation

This is one of the most uncertain areas for energy use estimation in hospitals. Different parts of the facility have differing energy intensities of clinical processes and specialist medical equipment and of building service requirements. Depending on the type of department, these loads can vary significantly. For example, for consulting areas, field studies have shown loads to be between 3-7 W/m^2 (Sheppy, et al., 2014), whereas UK NCM uses 27.31 W/m^2. Therefore, it is difficult to generalise typical operational trends of various spaces and their equipment loads. Additionally, a detailed audit is difficult due to access restrictions. Fine-tuning of these loads and operations after initial assumptions must be based on the modeller's judgement.

HVAC system design and operations

Underestimation of fossil-thermal energy is due to the low efficiency of the steam-based central heating network that serves the building. A new, efficient, combined heat and power (CHP) plant was used in calculations instead of the existing network, which was conditionally allowed. The new CHP plant was to be installed following a major renovation to maximise the efficiency savings across the facility rather than as a separate system for the new building only. As this has not happened yet, the present thermal performance of the building is much worse than expected from a new building.

Hospital hot water energy use is reported to vary widely, ranging from 10% of heating energy to up to 40%. The actual daily demand in the hospital was four times the UK NCM assumption of 3.4 l/m^2.

To maintain appropriate IEQ, through high air changes and filtration, energy used auxiliary systems to supply that fresh air was more than the design assumptions. Fans and pumps provided this close control and used around 30% of total electricity, which, as per the design, could have been more effective through proper demand control ventilation.

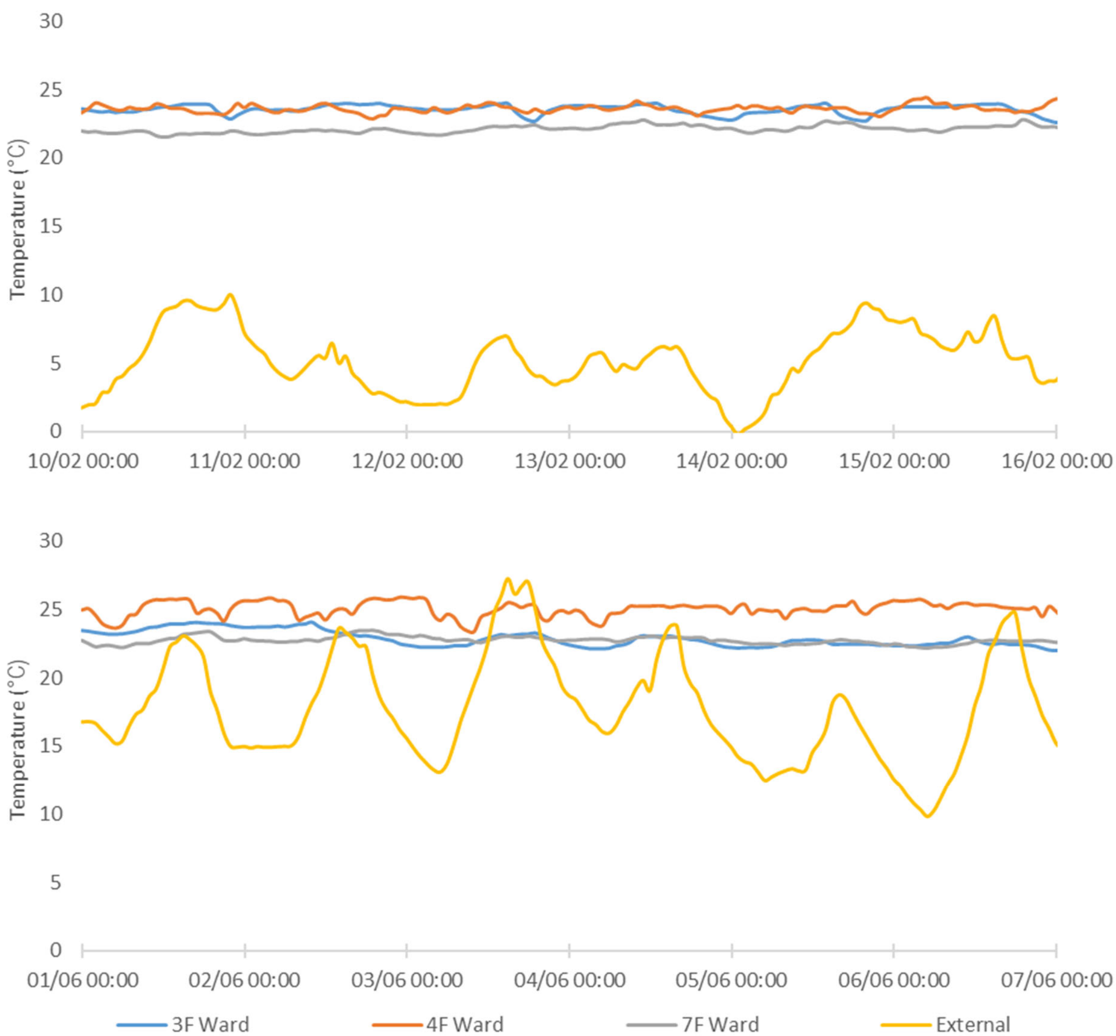


Figure 5.17: Monitored temperatures in sample spaces during the heating season (top) and non-heating season (bottom)

Controls for IEQ are typically decentralised, by having overrides to modulate local temperature and sometimes air flow rates. The design stage data, as per UK NCM used 22°C and 24°C as heating and cooling set-points for the wards throughout the year. However, it was observed in sample rooms that actual temperatures were around 20-23°C in the heating season and 22-25°C in the non-heating season (Figure 5.17). Spot measurements done on the BMS system, which observes all the spaces, showed that room set-points were anywhere from 18°C to 24°C depending on local requirements.

5.3.3 Model calibration results

Figure 5.18 compares the monthly simulation results against the actual measured data for electricity and gas use for the building and also reports the key statistical error values. By using the correct weather data, operational assumptions, loads and system configuration, this building's monthly energy use profile was estimated with an acceptable accuracy as per ASHRAE Guideline 14, i.e. C_V(RMSE) < 15% and NMBE < ±5%.

Further, to ensure that the calibrated model reasonably reflected the actual building's operation, Figure 5.19 shows the temperature trends for typical days for a sample space in the hospital building. The simulated temperatures closely follow the measured air temperature, except for the few instances where there are some deviations, which decline over the longer term.

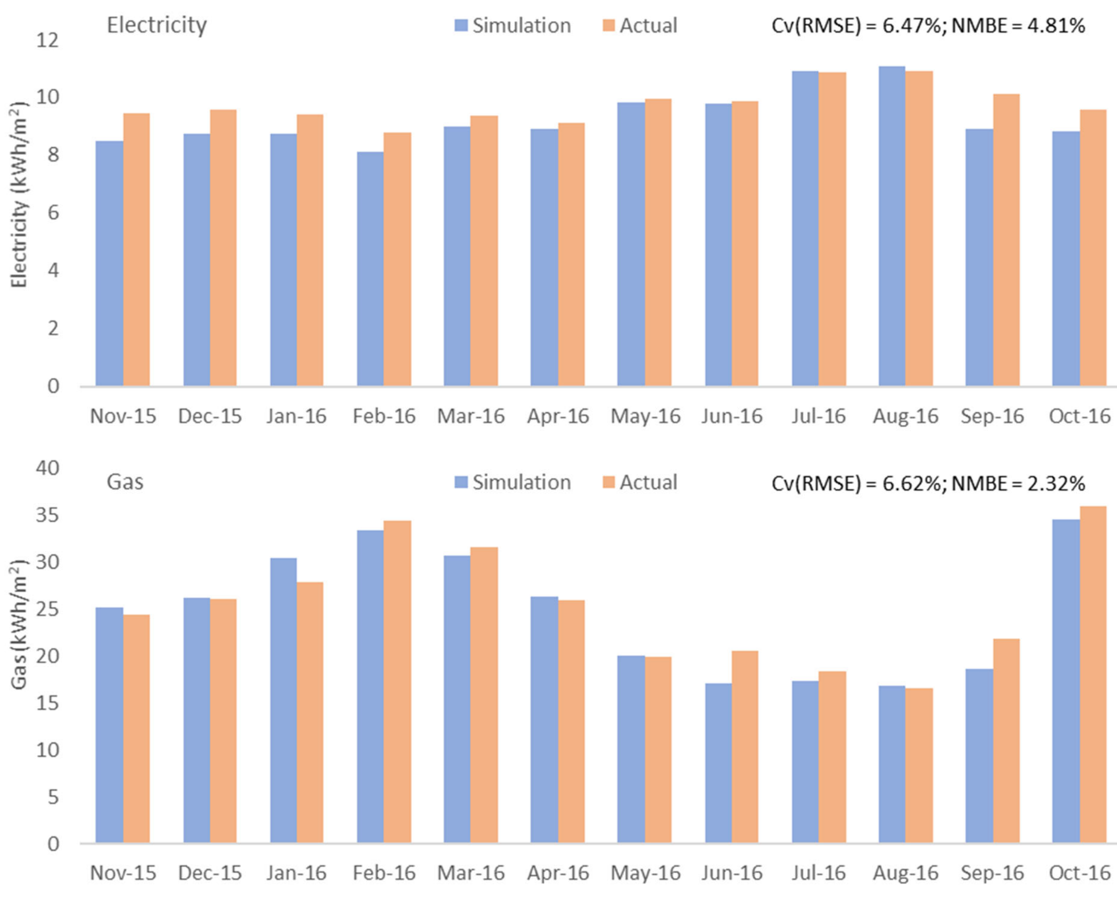


Figure 5.18: Monthly calibration results for electricity and gas use for the hospital building

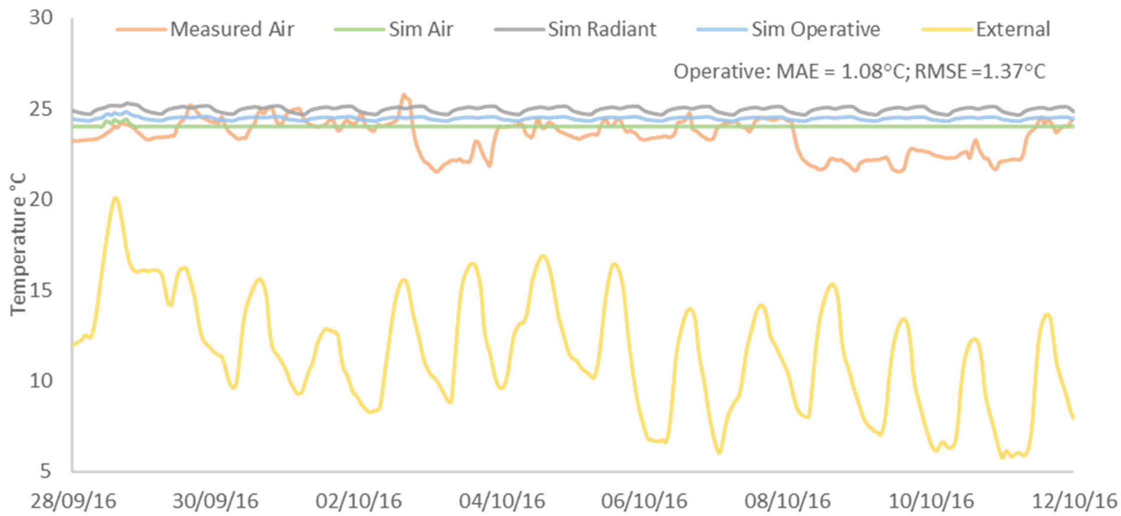


Figure 5.19: Hourly simulated and measured temperature profiles for typical days for a sample space in the hospital building

5.4 Apartment

5.4.1 Calibration-baseline (as-design) model

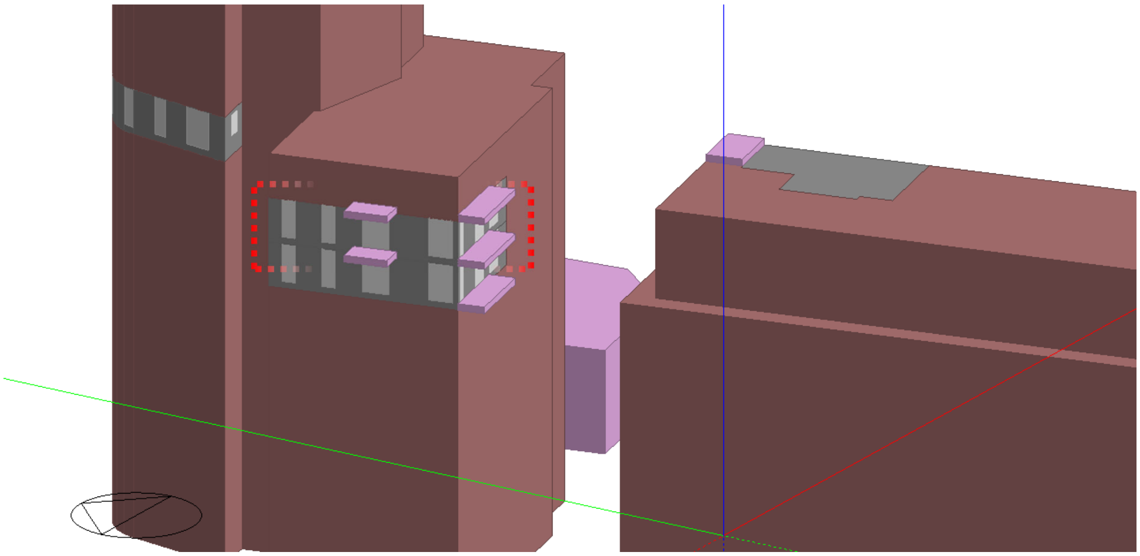


Figure 5.20: Model visualisation of the apartment building

As a first step towards calibration, a calibration-baseline (as-design) model (Figure 5.20) was created as per the details obtained from the design data described in the last chapter, in Section 4.4.1. As mentioned earlier, a model calibration exercise has been done for a typical flat in the apartment block, located on the eighth floor. Details for architectural design, fabric, occupancy, loads, and HVAC system were taken from design-stage documentation and calculation using standard NCM assumptions. Modelling of the HVAC system was done as a simplified system that calculates heating and cooling loads as per EnergyPlus ZoneHVAC:IdealLoadsAirSystem and factors in all operation control and overall efficiency. Weather data, from IWEC, for London Gatwick (WMO: 37760) was used as the nearest weather station to the site.

5.4.2 Deviations

Analysing the information gathered from operational-stage investigations into the building's performance, this section explains the deviations of building design and operations from the initial design intent.

Simulation weather data

Actual weather as per station and satellite measurements for Blackwall (WMO: 37820) for the calibration period was obtained from DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b) and was used in the simulations.

Occupancy, occupant behaviour, equipment, lighting and other loads and their operation

Apartments are most susceptible to the impact of occupant behaviour on energy use, as occupants have control over most of the operational aspects of the spaces, including added small power appliances and their use. Occupancy of an individual flat is very uncertain, especially due to the high degree of variability in occupant behaviours and high sensitivity of occupancy to events such as holidays or guests. CO_2 monitoring can show a possible lack of occupancy, sometimes for extended periods (Figure 5.21).

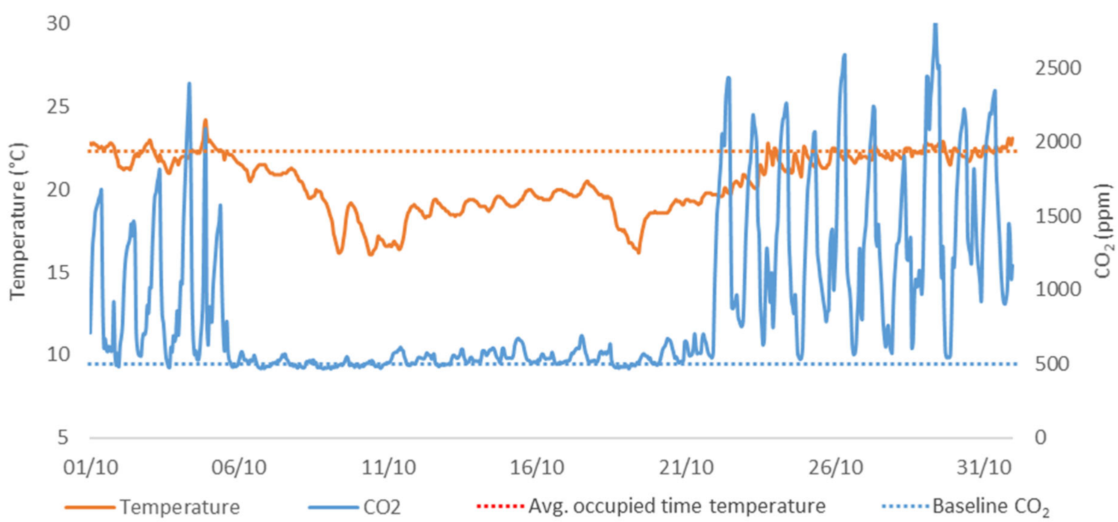


Figure 5.21: Monitored temperature and CO_2 trends in the bedroom show a two-week period when the flat was unoccupied (representative data from a neighbouring flat on the eighth floor)

Additionally, there are instances where it is evident that during summer, windows were opened for enhanced ventilation while the heating system was still operational (Figure 5.22).

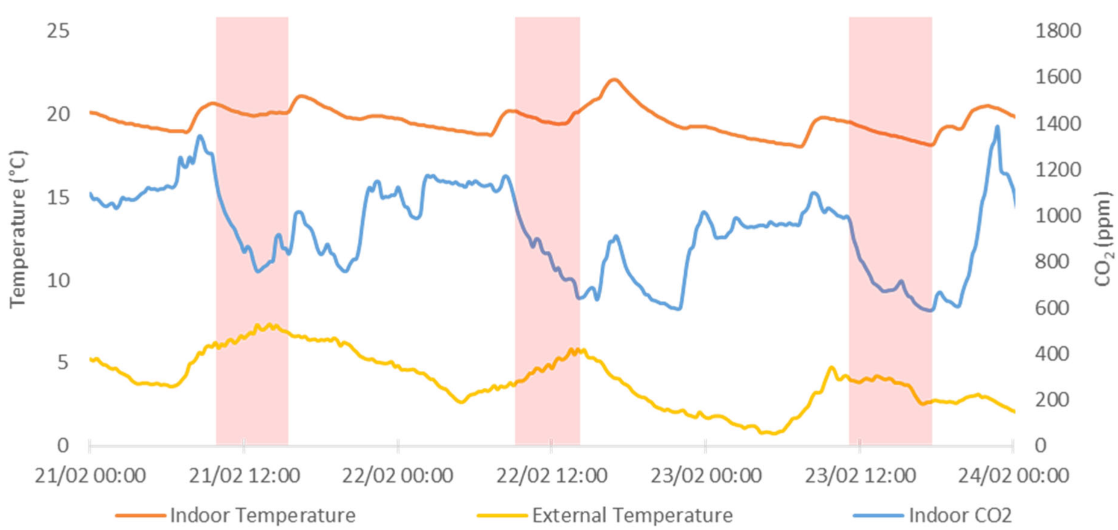


Figure 5.22: CO_2 reductions in the living room in the late morning are likely due to open windows, which is happening whilst the heating system is also running as indoor temperatures don't drop significantly

HVAC system design and operations

Maintaining of filters in the de-centralised MVHR systems was left to the occupants. However, most residents were not aware of the maintenance requirements of the installed MVHR systems. MVHR filters had not been replaced after three years of operation and were rarely cleaned, contributing to poor air quality and efficiency losses.

The annual gas use data for heat delivered to the apartments show an average annual heating efficiency of 50%, significantly lower than the design specification of 87%. This worsens the performance gap between actual operation and design intent.

Occupants had full control over the systems, determining their set-points and operations. Figure 4.23 shows that indoor temperatures were significantly higher than the comfort ranges recommended for the heating season (17-19°C in bedroom and kitchen, 22-23°C in living room, as per CIBSE Guide A).

5.4.3 Model calibration results

Figure 5.23 compares the monthly simulation results against the actual measured data for electricity and gas use for the building and also reports the key statistical error values.

Figure 5.23: Monthly calibration results for electricity and gas use for the flat

By using the correct weather data, operational assumptions, loads and system configuration, this apartment's monthly energy use profile was estimated with acceptable accuracy as per ASHRAE Guideline 14. Further, to ensure that the calibrated model reasonably reflected the actual building's operation, Figure 5.24 shows the temperature trends for typical days for a sample space in the flat. The simulated temperatures closely follow the measured air temperature.

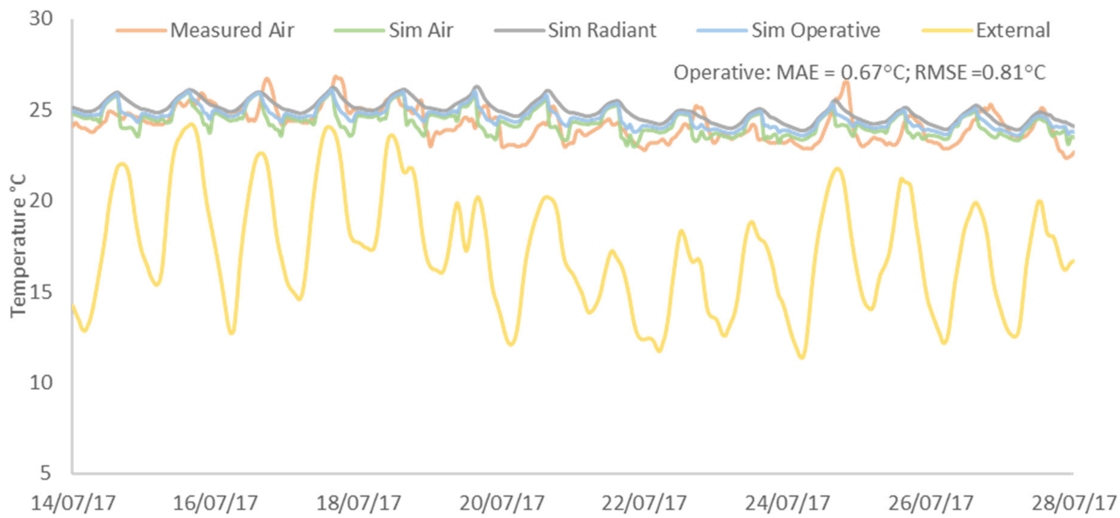


Figure 5.24: Hourly simulated and measured temperature profiles for typical days for a sample space in the flat

5.5 UNIVERSITY OFFICE

5.5.1 Calibration-baseline (as-design) model

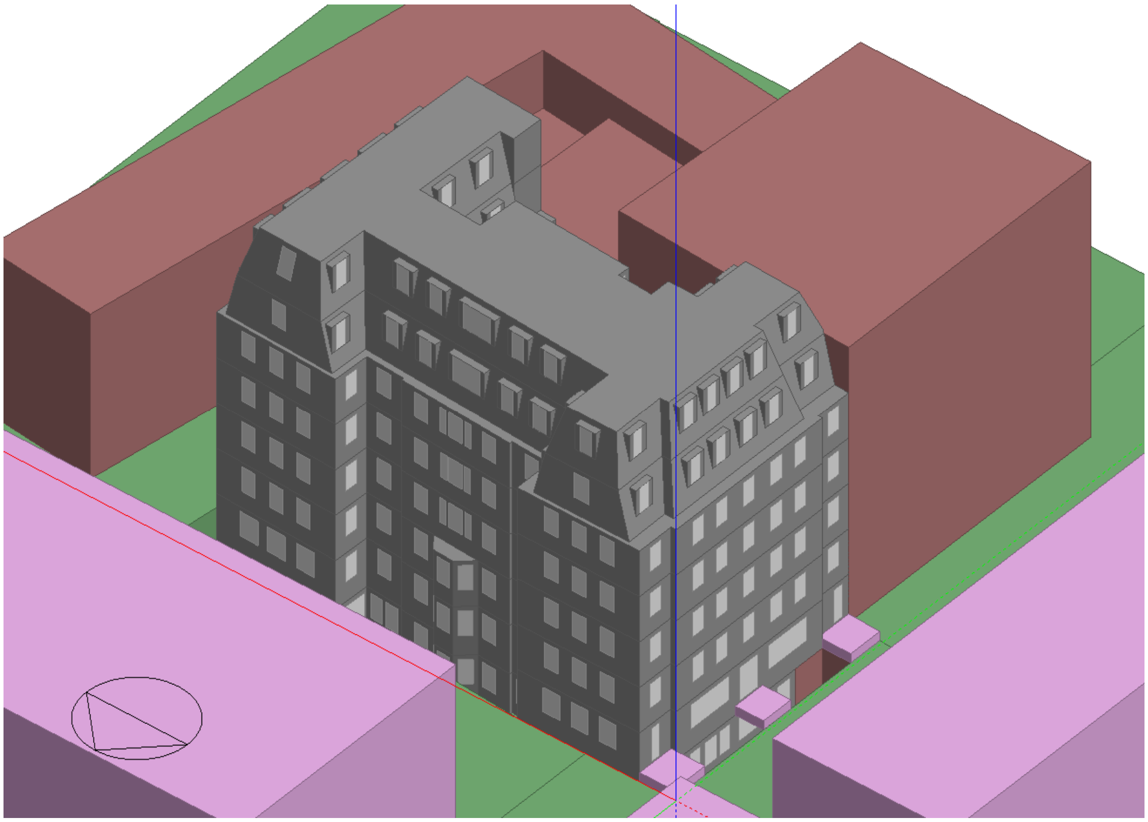


Figure 5.25: Model visualisation of the university office building

As a first step towards calibration, a calibration-baseline (as-design) model (Figure 5.25) was created as per the details obtained from the design data described in the last chapter, in Section 4.5.1. While some of the design information was available from design-stage drawings, there were no design-stage performance calculations available. The rest of the design data was supplemented through operation and maintenance (O&M) manuals, on-site observations and stakeholder discussions. Modelling of the HVAC system was done as a simplified system that calculates heating and cooling loads as per EnergyPlus ZoneHVAC:IdealLoadsAirSystem and factors in all operation control and overall efficiency. Weather data, from IWEC, for London Gatwick (WMO: 37760) was used as the nearest weather station to the site.

5.5.2 Deviations

Analysing the information gathered from operational-stage investigations into the building's performance, this section explains the deviations of building design and operations from the initial estimates used to create the calibration-baseline model.

Simulation weather data

Actual weather as per station and satellite measurements for London WEA Centre (WMO: 37780) for the calibration period was obtained from DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b) and was used in the simulations.

Occupancy and occupant behaviour

Being a university building, occupied by staff and students, monthly variation was seen in occupancy and loads, which can be linked to term times and the academic calendar. Additionally, unlike a typical office, occupancy was generally higher than assumed for out-of-hours use and weekends. This effect was not measured directly but could be inferred from detailed load profiles for typical months.

Figure 5.26 shows the electrical load for summer and winter months. The loads increase between 6 am and 8pm on weekdays. They revert to baseload levels very slowly, between 5 pm and 10 pm.

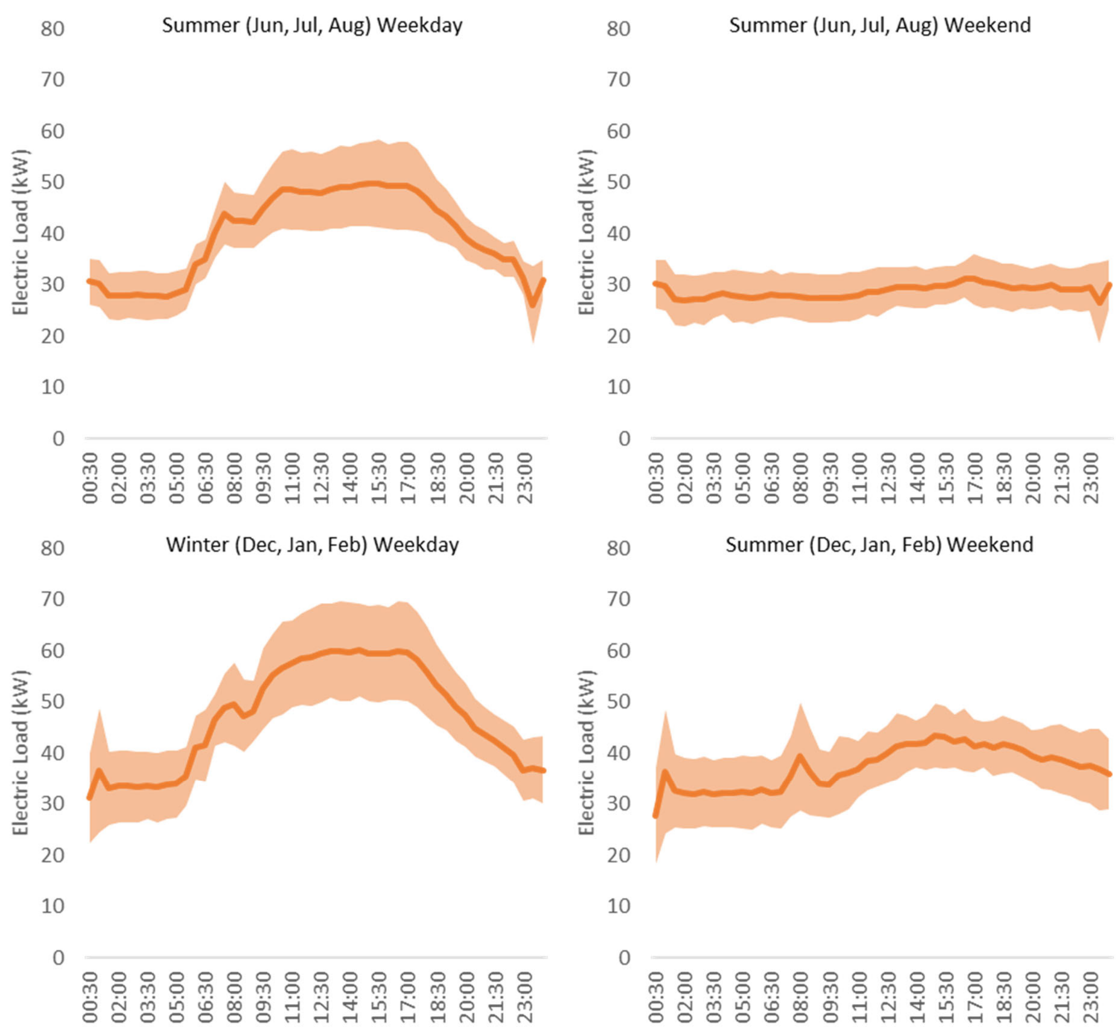


Figure 5.26: Average electrical demand and its standard deviation, for summer and winter months

This shows significant out-of-hours use of the building. An increase in loads on the weekends in winter months also points towards partial occupancy during term times, when students are more likely to use the buildings regularly, compared to summer (off-term time), when there are fewer academic activities on campus.

Equipment, lighting and other loads and their operation

During unoccupied times, there is very high lighting and power baseload (Figure 5.26) because some of the parasitic loads and also because some of the equipment (including servers) is either on or in standby mode. The lighting and power load during unoccupied time is about 50%. To account for these factors, deviations in baseline loads for lighting were assumed to be around 60%, and equipment around 35%. Additionally, following the occupancy data, adjustments for peak load and diversity were made to address the seasonal variation between term time and off-term time (August vs March in Figure 5.27).

HVAC system design and operations

Systems were operated throughout the day and night, even during unoccupied times. The distinction between occupied and unoccupied times for the heating + cooling heat maps in Figure 5.27 is not as clear as it is in the lighting + power heat maps. This shows that the systems were operating continuously.

The continuous operations of the HVAC system is also seen in the space temperature graphs in Figure 5.28, where the average occupied time temperature was 23°C, and unoccupied time temperature was 21°C, which are both significantly above the external temperatures. Also, temperature trends show a more distinct increase in zone temperature during occupied times in the open-plan spaces as compared to the cellular office, which, having its individual thermostat, remains at a steady temperature.

Another issue with the mechanical systems was that the AHU on the ground floor was observed to be out of order, which led to more opening of windows in the basement.

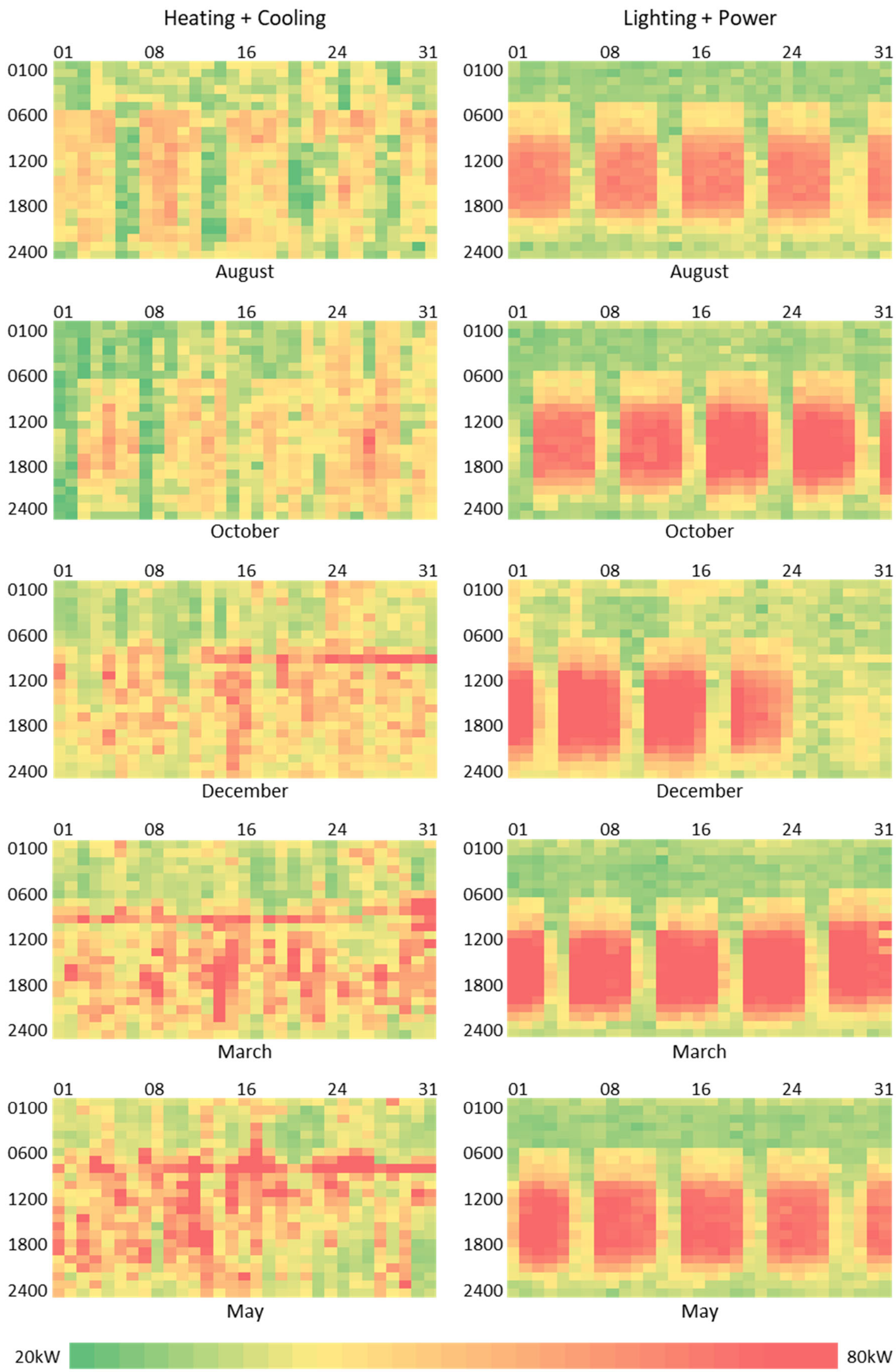


Figure 5.27: Heat map of selected months showing hourly loads for 'heating + cooling' and 'lighting + power' in the university office building.

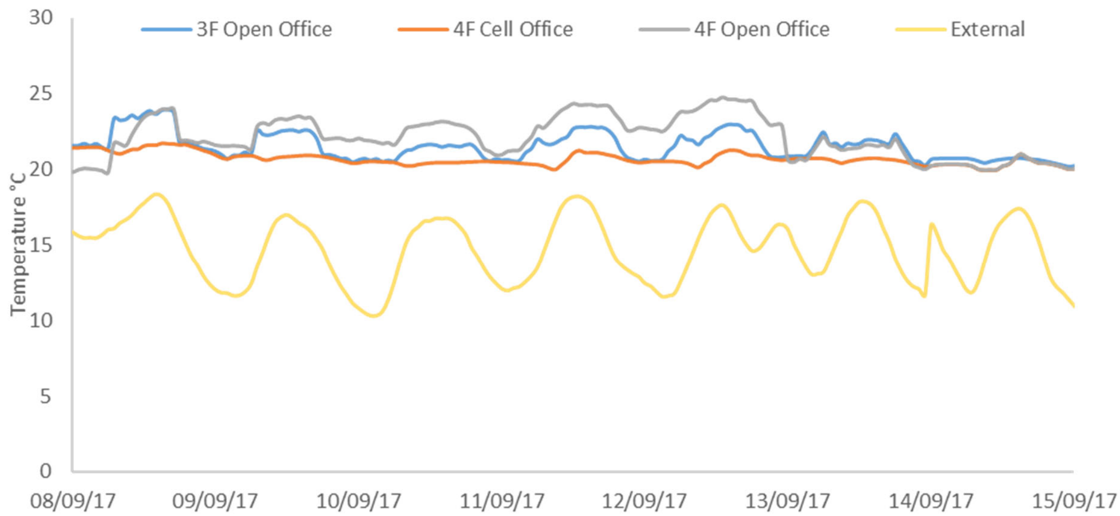


Figure 5.28: Monitored space temperature in sample spaces in a typical week in the university office building

5.5.3 Model calibration results

Electricity is the primary fuel used in the building for all end-uses (heating + cooling and lighting + power). Figure 5.29 compares the monthly simulation results against the actual measured data for electricity use for the building and also reports the key statistical error values. By using the correct weather data, operational assumptions, loads and system configuration, this building's monthly energy use profile was estimated with acceptable accuracy as per ASHRAE Guideline 14, i.e. C_V(RMSE) < 15% and NMBE < ±5%.

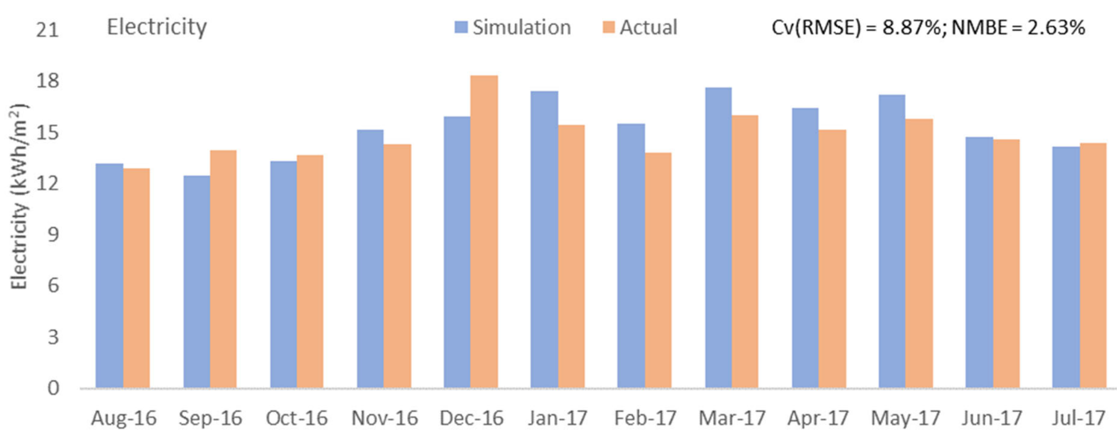


Figure 5.29: Monthly calibration results for electricity use for the university office building

Further, to ensure that the calibrated model reasonably reflected the actual building's operation, Figure 5.30 shows the temperature trends for typical days for a sample space in the university office building. The simulated temperatures follow the measured air temperature except for some deviations. There is a dip in simulated data on the weekends and overnight. However, over a longer period, these get close to the measured values.

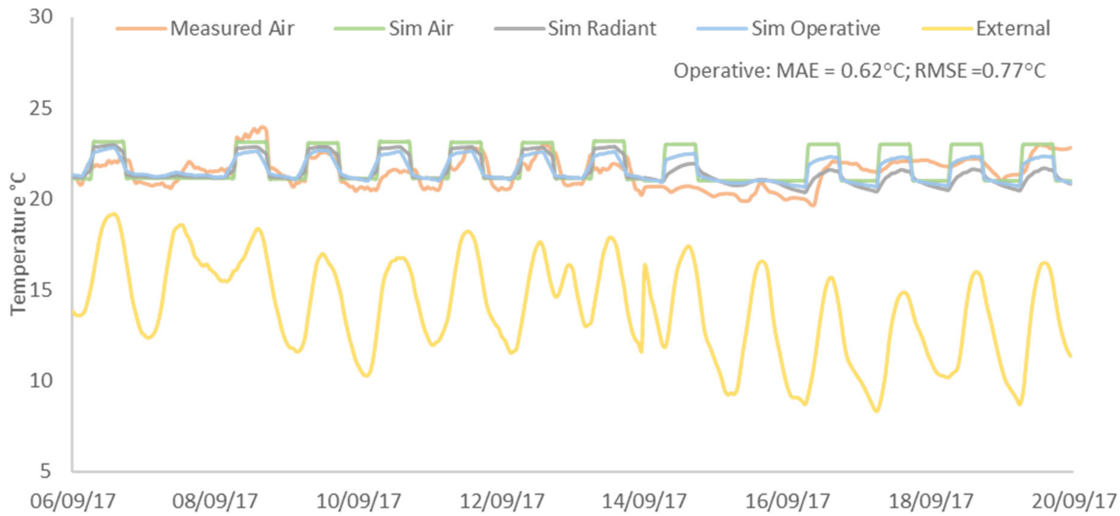


Figure 5.30: Hourly simulated and measured temperature profiles for typical days for a sample space in the university office building

5.6 SUMMARY

5.6.1 Modelling and Calibration

An initial design model (calibration-baseline) was created using DesignBuilder V6.0 (DesignBuilder Software Ltd., 2019a) based on the design data. This design model was then modified in accordance with actual operations and IEQ measurements, and operational-stage deviations. Further calibration was undertaken by evidence-based manual fine-tuning of various input parameters. Actual weather as per station and satellite measurements for the relevant location during the calibration period was obtained from DesignBuilder Climate Analytics tool (DesignBuilder Software Ltd., 2019b) and was used in the simulations. The calibrated model was validated using the monthly calibration criteria of ASHRAE Guideline 14 (C_V(RMSE) < 15%; NMBE < $\pm$5%). Also, hourly temperatures for typical weeks were compared against measured values in a sample space (Criteria used: MAE < 1°C; RMSE < 1.5°C). The intention of this stage is that once the calibrated model is created and validated, then the identified deviations are considered to be verified.

5.6.2 Summary of deviations (issues)

To identify the causes of the performance gap, evidence-based manual calibration enables a procedural investigation of issues and the certainty of capturing most of them. Fine-tuning of all case study buildings' calibration-baseline models was as per the evidence collected from the site or inferred from metering and monitoring data. Apart from the weather data used in the simulation, which will always be uncertain during the design-stage, the available evidence shows that most buildings had issues in similar areas. Issues identified include poor construction, changed space use and building operation trends, incorrect/changed specification of HVAC system components, modifications to the control strategy, inefficient operations, issues with commissioning and sometimes optimistic design-stage assumptions. These issues are key factors contributing to the gap. The following tables in this section summarise the causes for the energy performance gap in the case study buildings and model changes that were done during calibration.

Table 5.1: List of deviations (issues) in the public office building, fine-tuned for monthly calibration

Issue	Project stage of occurrence	Explanation/model changes
Heat loss due to leakage through poorly sealed vents and at construction junctions.	Construction Operation	It is difficult to obtain high airtightness in a naturally ventilated building. This increases the inflation rates.
The total occupancy of the building was about 25-30% higher, leading to more workstations being added	Design	In a new building, this could have been determined in the design-stage or handover-stage. This has a knock-on effect by having longer operation schedules and increased loads for all energy end uses.
Extended operation hours led to longer use of small power and lighting.		

Hot desking, using flexible workstations was not followed in practice. Hydraulic isolation of HVAC systems in unoccupied zones during out-of-hours use could not be followed.	Design Operation	These strategies require buy-in by the occupants during design and then follow-up during the operation-stage. As this was not practical for the building occupants, this led to inefficient operation of the building.
Parasitic loads were seen when the building was unoccupied.	Operation	Better management practices are necessary to minimise these. These lead to higher baseloads in equipment operation schedules.
Server loads were overestimated in design. This adversely impacted the heating system efficiency as there was less free heat.	Design	The primary system design relied on heat pumps and on the free heat from server rooms but as they had design issues, they were non-functional, and gas-fired boilers provided the heating.
Technical issues with heat exchangers and the flow rate specification caused the heat pumps to malfunction.		
The heating terminal's sizing was not consistent with the low-temperature flow required for energy-efficient operation of the system.	Design Commissioning	High set-point temperatures were used to compensate.
Some of the ventilation control sensors malfunctioned and required a later modification to the control strategy to overcome the system shortcomings.	Operation	Less fresh air was supplied to the zones as some vents remained closed.
Overriding of automatic vent opening to avoid thermal discomfort due to drafts around the floor cut-outs.		
Heating set-point temperatures maintained were about 2°C higher than the design intention, driven primarily by occupant needs	Operation	Higher set-points in temperature schedule.

Table 5.2: List of deviations (issues) in the school building, fine-tuned for monthly calibration

Issue	**Project stage of occurrence**	**Explanation/model changes**
Significant thermal bypasses around insulation were seen. The tested U-values were higher than designed.	Construction	Improper construction leading to higher U-value for the walls.
All spaces were not occupied throughout the day. Any given space is occupied only 60-70% of the full working day.	Design	These are typical operational patterns of schools these days and could have been incorporated during designing. This has a knock-on effect on the design of services. Overall, this leads to longer operation schedules and increased loads for all energy end uses.
There are extra-curricular activities and events, especially during the summer holidays.		
Centralised system design (one control and sensor for multiple zones) led to inefficient operations during out-of-hours use.		
Daylight sensors for automatic dimming of lights were poorly commissioned and PIR sensors had exceptionally long delay times.	Commissioning	Longer operation schedule for lights in affected spaces.
The biomass boiler installed, was never used because of logistic issues of using biomass.	Operation	Higher primary energy use and emissions.
Actual boiler efficiency (tested) was 88% compared to the design intent of 96%.	Design	Design values were optimistic. Lower COP for the boilers is used. And fan power is increased.

The Specific Fan Power in the AHU commissioning sheets was 66% higher than the values used in design-stage estimations.		
The supply fan during unoccupied times was running inefficiently, operating at 30% to 40% of its nominal speed.	Commissioning Operation	The base-load operation schedule of the fan increased
Building heating and ventilation systems were running during unoccupied hours.	Operation	Longer operation schedule for the HVAC system.
Indoor temperatures monitored during the winter were about 2°C higher than the design set-points and sometimes even more.	Operation	Higher set-points in temperature schedule.

Table 5.3: List of deviations (issues) in the hospital building, fine-tuned for monthly calibration

Issue	Project stage of occurrence	Explanation/model changes
Different clinical processes had unique needs. It was difficult to generalise typical operational trends of various spaces; their equipment loads; and set-point temperatures.	Operation	These changes are needed for the building to function in practice. Separate schedules and loads were used based on-site visit information and average values were determined from other monitored data.
An increased number of beds were observed in the patient wards during the site visits.	Operation	
Hospitals had a high changeover rate of patients. Therefore, uncertainties in set-point temperatures and operations of hospital equipment were quite high.	Operation	
A low efficiency, old, steam-based central heating network serves the building. A new CHP plant was to be installed for the entire facility rather than just for the new building to maximise savings. However, this has not happened yet.	Design	Higher primary energy use and emissions.
Hospital hot water energy use was high due to the clinical requirements.	Operation	Increased hot water demand.
Fans and pumps are used to provide close control to the indoor environment. Demand control ventilation could have been used more effectively and heating energy use could be reduced whilst keeping comfortable indoor temperatures.	Operation	Increased baseload and operating schedule for system auxiliaries.

Table 5.4: List of deviations (issues) in the apartment block, fine-tuned for monthly calibration

Issue	Project stage of occurrence	Explanation/model changes
Occupancy of the individual flat is very uncertain to determine, especially holidays.	Operation	Changes to schedules for occupancy and other loads.
CO_2 monitoring show a possible lack of occupancy during school holiday times.		

There are instances where it is evident that during summer, the windows were opened for enhanced ventilation, while the heating system was still operational.	Operation	Additional window opening.
Maintaining filters in the de-centralised MVHR systems was left to the occupants and therefore there were maintenance issues (e.g. dirty filters) in most flats.	Operation	Lower efficiency of fan.
Efficiency of the district heating network, estimated at 91% in design calculations, was around 50% in practice.	Operation	Higher primary energy use and emissions.
Occupants had full control over heating system & MVHR, determining their set-points and operations.	Operation	Set-points determined based on average temperature measurements.

Table 5.5: List of deviations (issues) in the university office building, fine-tuned for monthly calibration

Issue	**Project stage of occurrence**	**Explanation/model changes**
Measured occupancy varied monthly and seasonally. It was higher than that assumed for out-of-hours use and weekends.	Operation	Changes to schedules for occupancy and other loads.
Baseline energy use for lighting and power consumption is high, at about 50% lighting and power load during the unoccupied time.	Operation	Increased lighting and equipment baseloads.
Systems were operated throughout the day and night, even during unoccupied times.	Operation	System operational schedules changed.
High set point temperature maintained during occupied (23°C) and unoccupied (21°C) times.	Operation	Higher set-points in temperature schedule.
AHU in the basement not working and subsequent more opening of windows in the basement.	Operation	Modified system and window operations.

6 DETAILED MODEL CALIBRATION (HOURLY) AND VALIDATION

Model calibration done for assessment of ECMs and identification of underperformance issues should be done for one year, with minimum granularity of a month. But higher granularity data (hourly or daily), while being more onerous, is preferred in cases where the calibrated model is used for a system or a sub-system level analysis (EVO, 2016). The calibration undertaken in the last chapter focused on monthly energy use and was effective for the identification and validation of the key performance issues in the case study buildings. However, for one of the case study buildings, the university office building, more granular and disaggregated data was available. This allowed for a more thorough understanding of the building's performance, as detailed model calibration for disaggregated sub-metered data, using advanced validation checks, could be attempted.

The hourly calibration for the university office building was used to analyse the demands and usefulness of such calibration. This analysis helps to show that weather hourly data calibration is worth the difficulty of obtaining and analysing it, in other words, a trade-off between effort and accuracy, of more granular model calibration. Based on the lessons from this detailed calibration, the chapter also reflects upon the amount of additional data required and the applicability of hourly calibration models compared to monthly calibration models. During this process, advanced model validation techniques are also explored.

In this chapter, the monthly calibration results are first compared against the hourly calibration results. Then, the hourly calibration model is further fine-tuned by undertaking calibration for sub-metered data, available with spatial and end-use disaggregation, cross-validated by checking the air temperature in different space types. Finally, to present the calibration results in context, uncertainty-based validation, by defining the remaining uncertainty for key inputs, was done to establish the robustness of the model.

6.1 INTRODUCTION

The calibration model for the university office building, developed in the last chapter, is taken further by following a new multi-level framework for validating a calibrated model. This framework, explained in Section 3.3.3 is based on the following principles:

- Statistical tests provide the basis for accuracy;
- Disaggregated checks reduce cross compensation;
- Cross-validation of dependent variables ensures the correctness of inputs and the calculations; and
- Probabilistic validation ascertains that uncertain, yet unknown, parameters can explain the residual gaps.

As created in Section 3.3.3, Level 1 calibration is done at the building-level meters, and Level 2 calibration is done for disaggregated sub-meters. Results for these can be assessed with different temporal granularity, A: Monthly, B: Weekly, C: Daily, and D: Hourly. For each of these, three levels of cross-validation can be attempted, simple (SIM), intermediate (INT) and advanced (ADV). For this university office building, Level 1A-SIM (i.e. building-level, monthly with simple cross-validation) calibration results were shown in the previous chapter. In this chapter, following further analysis and calibration, results are presented as follows:

- Weekly, daily and hourly calibration results of the (monthly) Level 1A-SIM model are analysed.
- Further calibration of the model to the most detailed Level 2D-ADV is attempted with advanced levels of validation and uncertainty quantification.

Comparing the Level 1 model calibration results with advanced Level 2 results across various temporal resolutions gives insights into issues such as data granularity vs accuracy of calibrated models. Level 1 model calibration is suitable for finding and validating areas of underperformance, but it may not be suitable for a detailed assessment of system issues. As developing a highly granular and detailed calibrated model takes time and is computationally expensive, exploring this in detail in one model provides an understanding of which type of model calibration will be sufficient for which purposes.

6.2 MONTHLY VS HOURLY (LEVEL 1 MODELS)

Section 5.5 explains the fine-tuning of the calibration-baseline model of the university office building for monthly electricity use with simple cross-validation (Level 1A-SIM). The model was fine-tuned for issues such as system faults and modifications to typical operations as per site information and interactions with facility managers (Section 5.5.2). As that model was calibrated for building-level for monthly data points, acceptable criteria were achieved even though at a more granular level there was some degree of cross-compensation. Table 6.1 shows the calibration statistics (C_V(RMSE) and NMBE) of the Level 1A-SIM calculated for monthly, weekly, daily and hourly granularity. Most of the calibration statistics at finer resolutions are within the thresholds described in Table 3.5. $C_V(RMSE)_{hourly}$ does not meet that threshold, and the increased $C_V(RMSE)_{hourly}$ with low NMBE can be seen in the scatter plot in Figure 6.1, where the best fit line is significantly away from the ideal fit line for hourly data. This provides an initial indication that the Level 1A-SIM model is not a good predictor of performance at higher granularity. We will explain this in more detail.

Table 6.1: C_V(RMSE) and NMBE at monthly, weekly, daily and hourly granularity for the Level 1A-SIM calibrated model

Index	**Monthly**	**Weekly**	**Daily**	**Hourly**
NMBE (%)	-2.63	-2.63	-2.63	-2.63
C_V(RMSE) (%)	8.44	10.49	14.14	33.72

Green: Table 3.5 thresholds met; Red: Thresholds not met.

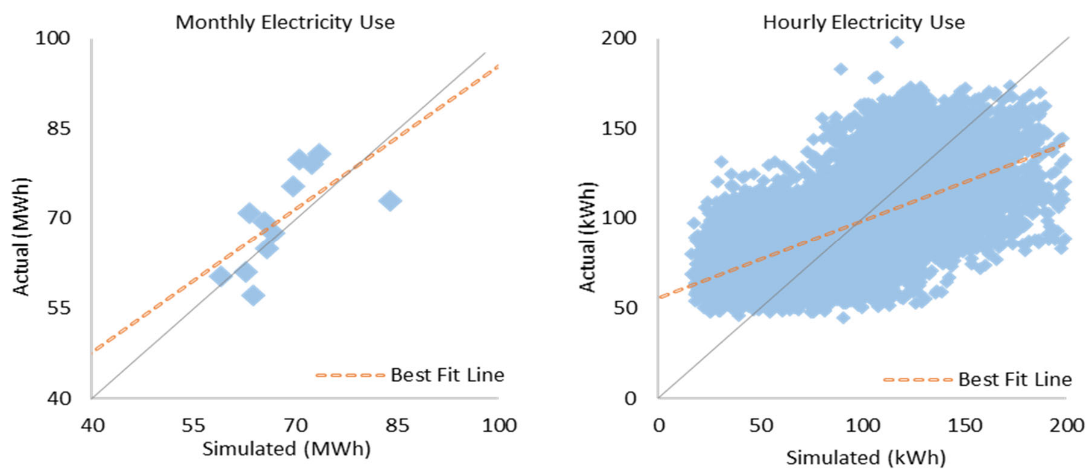


Figure 6.1: Scatter plot for the building's total electricity use at monthly and hourly calibration granularity for the Level 1A-SIM calibrated model

Figure 6.2 shows actual monthly electricity use and the Level 1A-SIM calibrated model's monthly simulated results (top graph) compared against the monthly box plots of hourly actual electricity use and hourly simulated results of the same model (bottom graph). While the median and mean values in the box plots are close for each month, interquartile range and minimum and maximum values for simulated results are much wider than for actual monitored values. This suggests that, while there is no compensation between months, there is a certain level of cross-compensation for the hourly time-steps within a month. This cross-compensation is more clearly seen in Figure 6.3, which shows the NMBE and C_V(RMSE) for hourly electricity use calculated per month shows that the hourly calibration thresholds that in many months are not met.

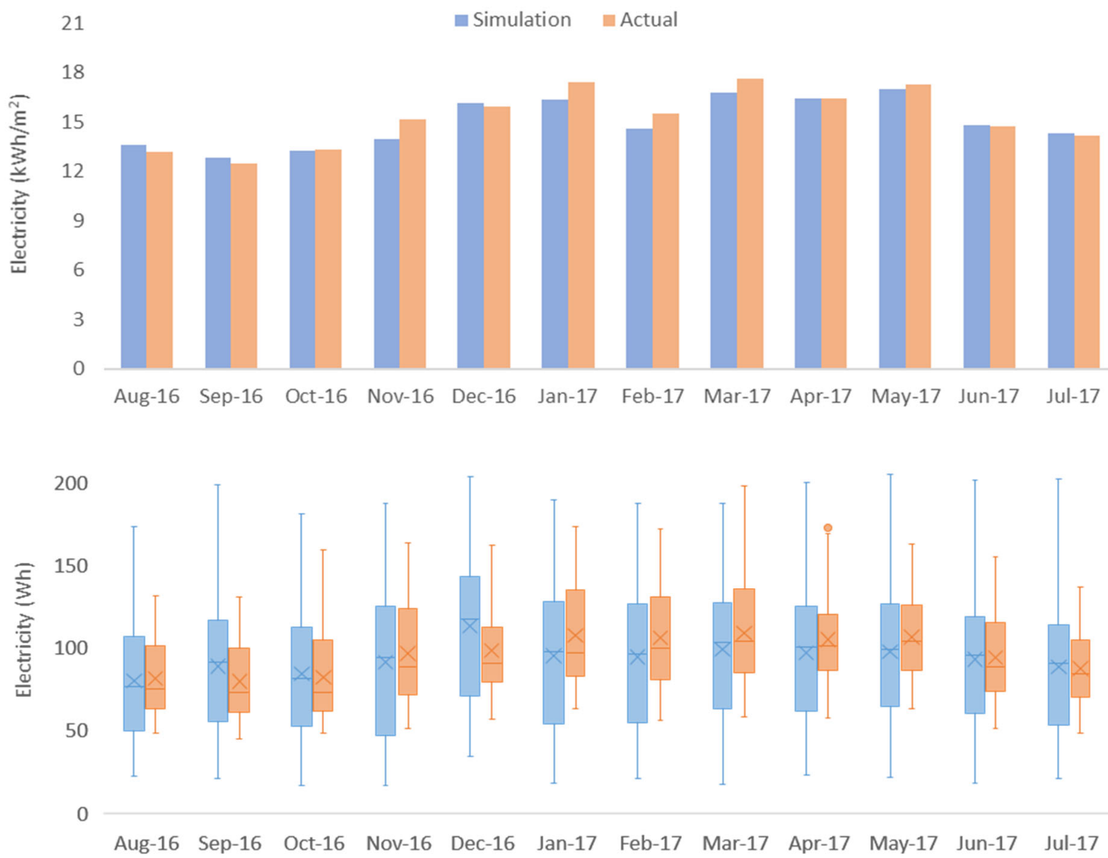
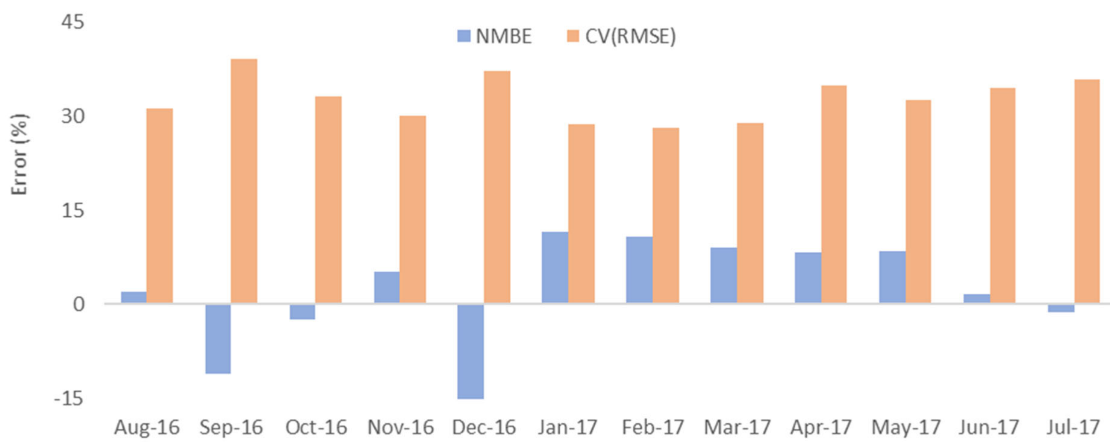


Figure 6.2: (Top) Monthly simulation vs actual electricity use of a Level 1A-SIM calibration model; (Bottom) Hourly simulated vs actual electricity use spread per month for the same Level 1A-SIM calibration model



	Aug-16	Sep-16	Oct-16	Nov-16	Dec-16	Jan-17	Feb-17	Mar-17	Apr-17	May-17	Jun-17	Jul-17
NMBE (%)	1.89	-11.23	-2.50	5.23	-15.25	11.59	10.84	9.05	8.21	8.45	1.56	-1.38
CVRMSE (%)	31.09	39.00	33.03	29.99	37.17	28.60	28.09	28.95	34.85	32.59	34.45	35.71

Figure 6.3: NMBE and C_V(RMSE) based on hourly electricity use data calculated per month for the Level 1A-SIM calibrated model (Green: Table 3.5 thresholds met; Red: Thresholds not met)

Analysing NMBE and C_V(RMSE) based on monthly and hourly electricity use data calculated for various sub-meters in the building reveals a similar cross-compensation between sub-meters. Table 6.2 shows that there is a lot of cross-compensation between floors for lighting + power end-uses, with some meters having a deviation in excess of 100%.

The heating + cooling sub-meter has better monthly calibration than other meters, partly because of the simple cross-validation checks done for various spaces. However, the large C_V(RMSE) at hourly resolution suggests larger seasonal variation, which needs to be addressed in Level 2 calibrations.

Table 6.2: NMBE and C_V(RMSE) for sub-metered monthly and hourly electricity use for the Level 1A-SIM calibrated model

End-use	Sub-meter	Monthly		Hourly	
		NMBE (%)	**C_V(RMSE) (%)**	**NMBE (%)**	**C_V(RMSE) (%)**
Heating + Cooling	Heating + Cooling	-1.59	12.37	-1.59	62.21
Lighting + Power + Hot Water + Lift	Basement	-5.86	17.49	-5.86	99.13
	Ground Floor	-3.64	13.77	-3.64	44.48
	First Floor	33.58	34.89	33.58	47.05
	Second Floor	-9.60	12.31	-9.60	39
	Third Floor	97.17	98.05	97.17	71.01
	Fourth Floor	-49.76	50.24	-49.76	116.47
	Fifth Floor	-2.85	15.37	-2.85	49.34
	Sixth Floor	10.10	14.51	10.10	62.24
Server	Server	0.00	1.75	0.00	12.81
Building Total		**-2.63**	**8.44**	**-2.63**	**33.72**

Green: Table 3.5 thresholds met; Red: thresholds not met.

6.3 SUB-METERED DATA CALIBRATION (LEVEL 2)

The level 1A-SIM calibrated model, while being suitable for monthly calibration and understanding general building operable performance trends was less accurate at high granularity. A Level 1D-SIM or Level 1D-ADV model for hourly building results could have been created by fine-tuning inputs at a higher granularity. However, it might still not have addressed the cross-compensation between various sub-meters. Despite the limitations that heating + cooling and lighting + power were not metered separately, Level 2 calibration with advanced validation for sub-metered data with spatial and end-use disaggregation was done for this building.

6.3.1 Further fine-tuning

This section describes the further fine-tuning that was done to the previous Level 1A-SIM model. The fine-tuning was based on detailed assessments of the disaggregated sub-metered energy data and of the IEQ measurements.

Occupancy and operations schedule

Floor wise lighting + power electricity use data helped to identify typical occupancy for various zone types. There were significant variations seen in schedules and hours of use depending on the occupant profile. The first floor was used by administrative staff and had a more regular occupancy compared to the floors used by research students, which had a diversified usage.

Figure 6.4 shows typical weekday and weekend electric loads for the first and fourth floors. On the first floor, during the weekdays the load profile was similar to that of a typical office. The electric load increases around 9 am and starts reducing after 5 pm. Weekend loads on the first floor show a consistent baseload with a very low standard deviation across the year, indicating that there is no occupancy. This is very different to the loads on the fourth floor, where the load profile is highly diversified across the day. Even on the weekends, the standard deviation of the fourth-floor loads is quite high, indicating partial occupancy year-round for that floor.

Further fine-tuning was also done for low occupancy spaces during the off-term time (mid-June to mid-September), Christmas and Easter closures, and for bank holidays. The heat map in Figure 6.5 shows a distinctly lower electric load during these days.

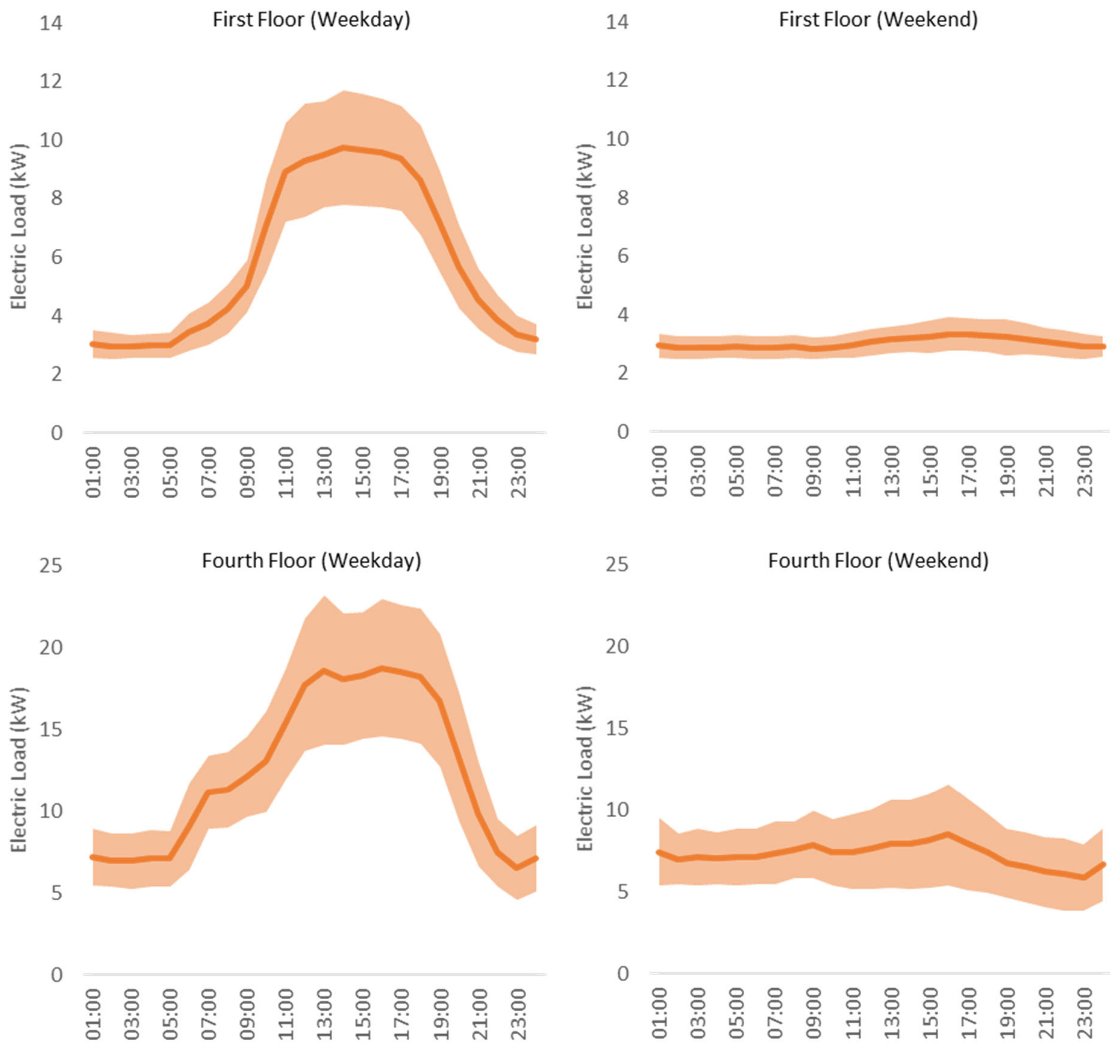


Figure 6.4: Average electric load trends and their standard deviation for first and fourth floor for weekdays and weekends

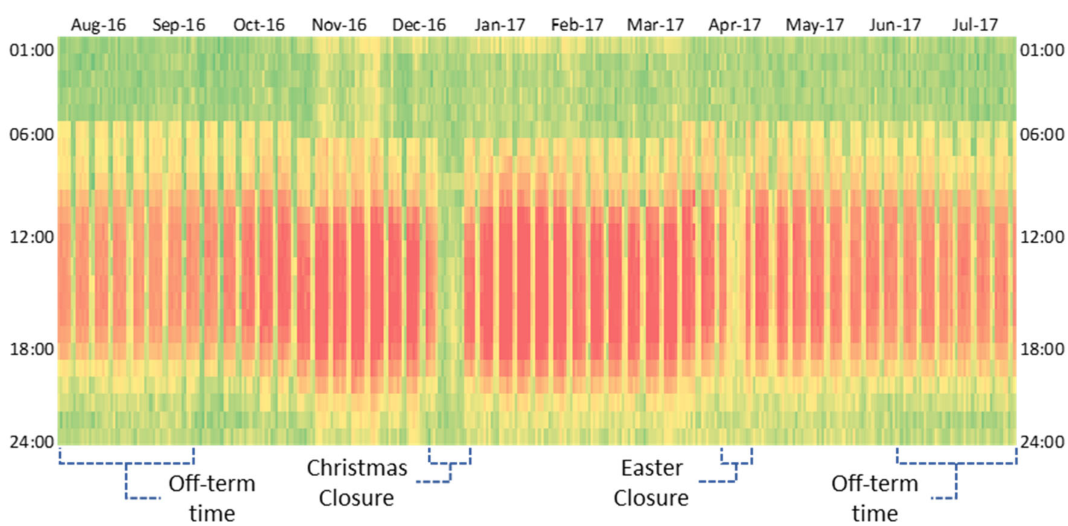


Figure 6.5: Heat map showing electricity use trends throughout the year

Small power, lighting load and schedule

Figure 6.4 shows that the baseload for fourth-floor lighting + power is twice that of the first floor. Redistribution of lighting + power loads across the floors was done proportionally as per the variation in baseload. Additionally, schedules for light and power use, which are linked with space occupancy were changed based on average monitored electric load patterns on different floors.

HVAC system design and operations

Figure 6.6 shows that the system runs continuously throughout the day and even on weekends. There is an increased start-up load for a couple of hours in the morning.

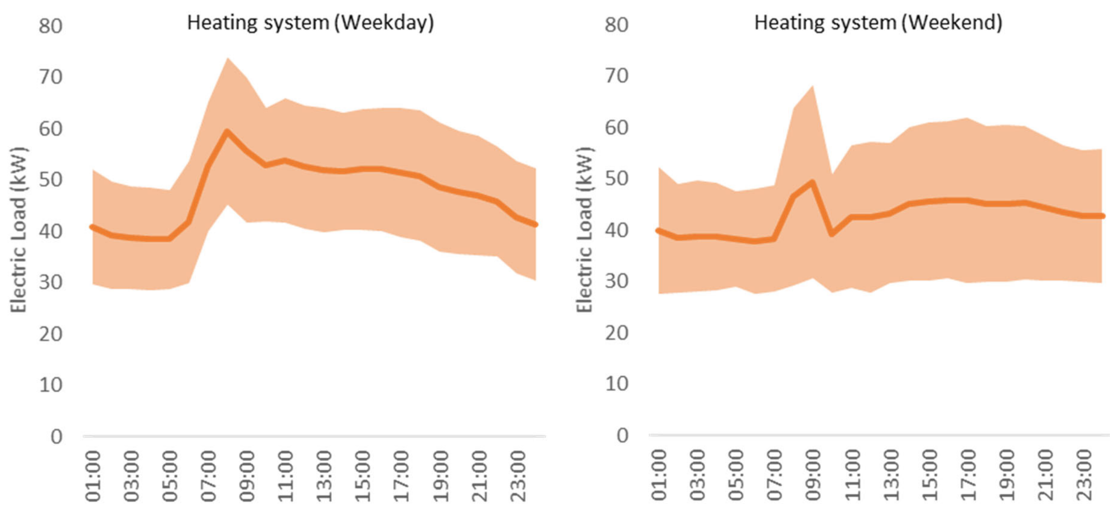


Figure 6.6: Heating system's average electric load trends and their standard deviation for weekdays and weekends

Figure 6.7 shows the average temperature in a couple of open-plan offices for summer and winter, on weekdays and weekends. Space temperatures maintained varied seasonally; they were recorded between 21°C and 24°C during winter and around 24°C and 25°C during summer. However, the temperature difference between the occupied and unoccupied and times was 1-2°C, which is further evidence of continuous HVAC system operations. This was also confirmed from the BMS logs, and these averages were used to define the space set-point schedules at higher granularity for calibration.

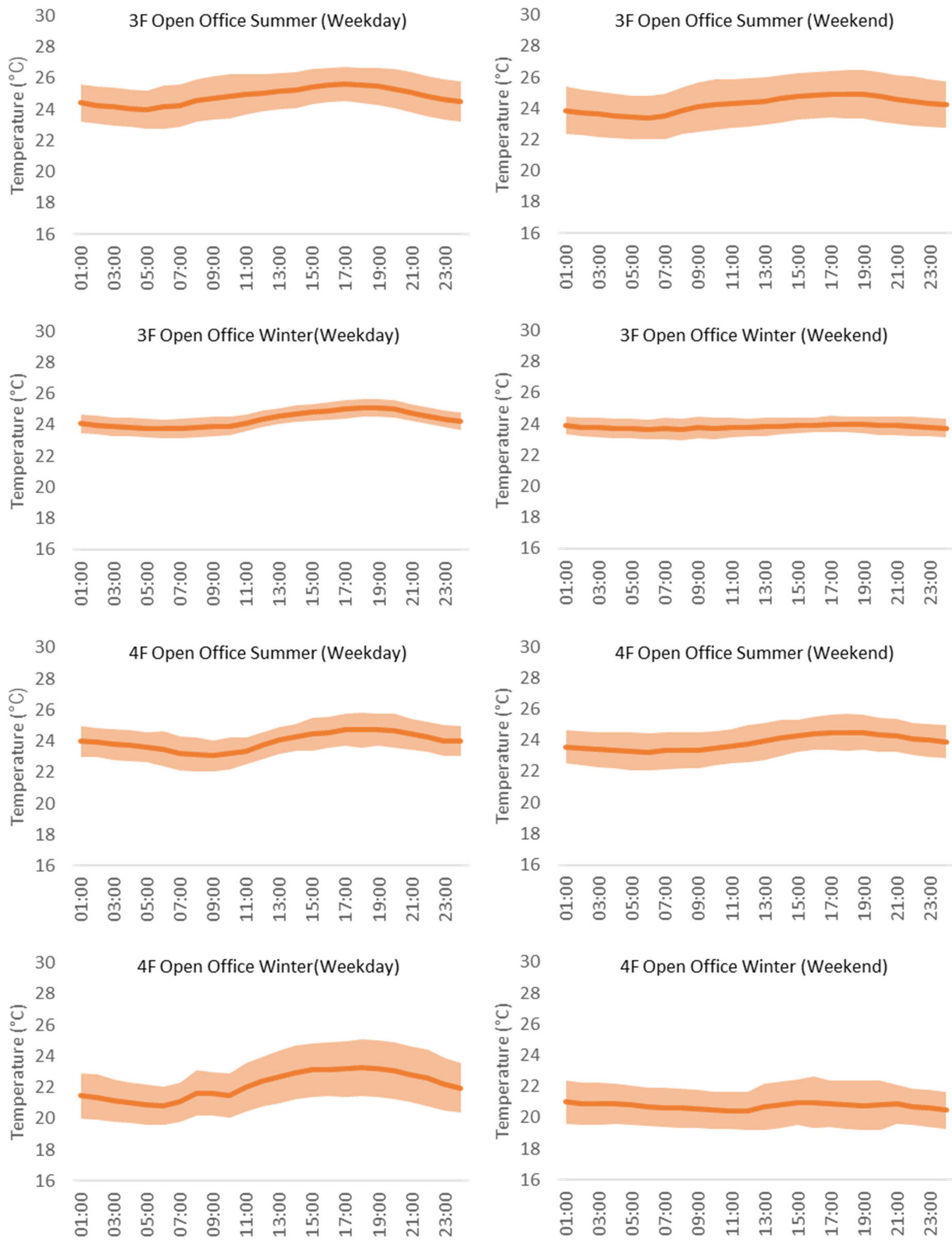


Figure 6.7: Average temperature and its standard deviation for a couple of spaces during summer and winter season for weekdays and weekends

6.3.2 Calibration results

The level 1A-SIM calibrated model was further fine-tuned for occupancy, operations, load profiles and system controls based on the granular trends explained in the earlier section. This enabled us to undertake Level 2D calibration for end-use and spatially disaggregated metered data with Advanced (ADV) cross-validation. Table 6.3 shows the calibration statistics (C_V(RMSE) and NMBE) for the building-level energy use of the Level 2D-ADV (hourly) calibrated model, with additional error calculation for monthly, weekly and daily granularity. Figure 6.8 shows the actual vs simulated scatter plot for building-level energy use for monthly and hourly data.

Table 6.3: C_V(RMSE) and NMBE for the Level 2D-ADV calibrated model along with monthly, weekly, daily and hourly granularity

Index	Monthly	Weekly	Daily	Hourly
NMBE (%)	1.52	1.52	1.52	1.52
C_V(RMSE) (%)	3.98	8.36	10.99	16.37

Green: Table 3.5 thresholds met; Red: Thresholds not met.

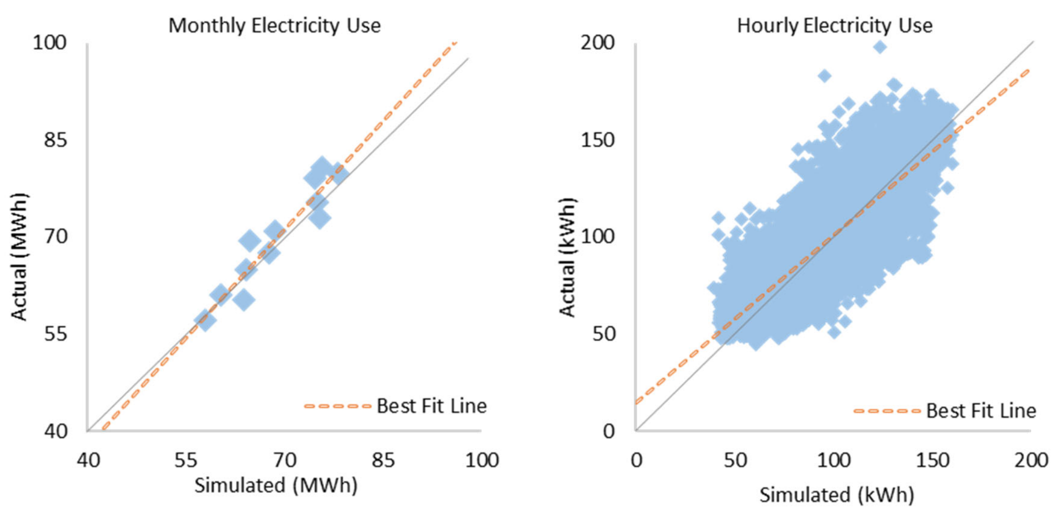


Figure 6.8: Scatter plot for the building's total electricity use at monthly and hourly calibration granularity for the Level 2D-ADV calibrated model

Compared with the calibration results in Table 6.3 from the Level 1A-SIM model (Table 6.1 and Figure 6.1), all the calibration statistics at different resolutions are within the thresholds (as per the criteria defined in Table 3.5). This can also be seen in Figure 6.8, where there is tight clustering of the scatter and the best fit line is close to the ideal fit line. Further, Figure 6.9 shows monthly electricity use and month-wise box plots of hourly electricity use, showing significantly reduced deviations and variability of simulation results. Reduction in cross-compensation is also seen in Figure 6.10, as hourly calibration thresholds are met in all the months.

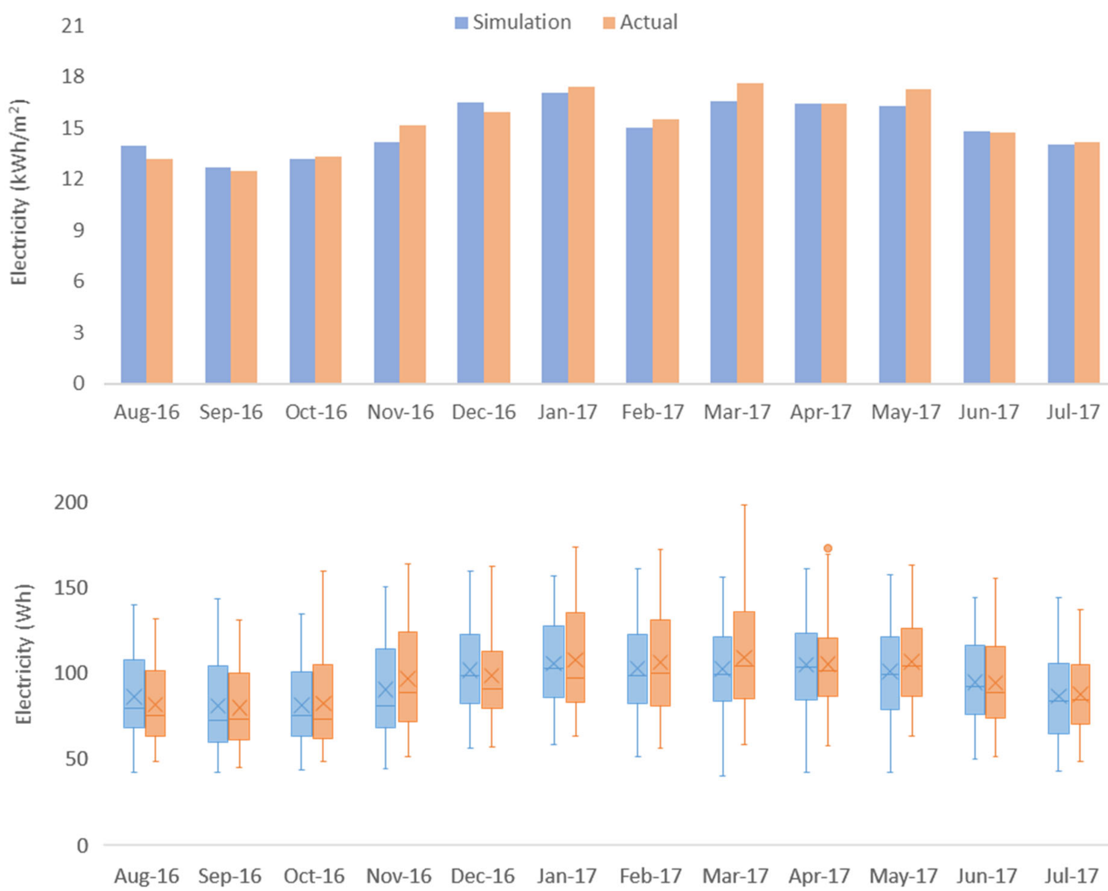
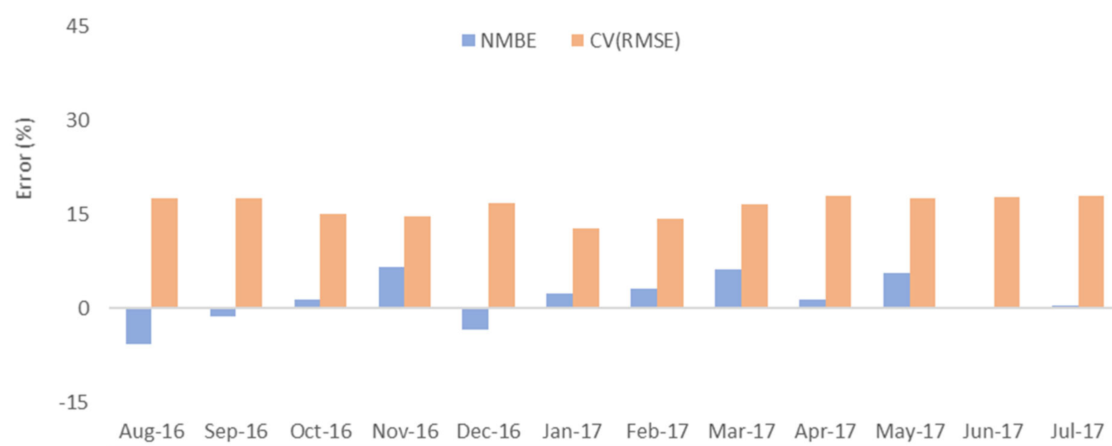


Figure 6.9: (Top) Monthly simulation vs actual electricity use of a Level 2D-ADV calibration model; (Bottom) Hourly simulated vs actual electricity use spread per month for the same Level 2D-ADV calibration model



	Aug-16	Sep-16	Oct-16	Nov-16	Dec-16	Jan-17	Feb-17	Mar-17	Apr-17	May-17	Jun-17	Jul-17
NMBE (%)	-5.71	-1.29	1.40	6.54	-3.47	2.23	2.99	6.13	1.31	5.50	-0.14	0.48
CVRMSE (%)	17.43	17.44	15.01	14.54	16.58	12.62	14.18	16.48	17.94	17.60	17.79	17.97

Figure 6.10: NMBE and C_V(RMSE) based on hourly electricity use data calculated per month for the Level 2D-ADV calibrated model (Green: Table 3.5 thresholds met; Red: Thresholds not met)

The key emphasis of Level 2D-ADV calibration is to undertake fine-tuning of end-use and spatially disaggregated metering data. Table 6.4 lists the NMBE and C_V(RMSE) for monthly and hourly electricity use calibration, calculated for various sub-meters in the building. It can be seen that all the threshold criteria are met.

Table 6.4: NMBE and C_V(RMSE) for sub-metered monthly and hourly electricity use for the Level 2D-ADV calibrated model

End-use	Sub-meter	Monthly		Hourly	
		NMBE (%)	C_V(RMSE) (%)	NMBE (%)	C_V(RMSE) (%)
Heating + Cooling	Heating + Cooling	0.31	4.13	0.31	29.15
Lighting + Power + Hot Water + Lift	Basement	3.82	7.37	3.82	24.31
	Ground Floor	5.44	12.53	5.44	27.24
	First Floor	3.83	9.36	3.83	24.84
	Second Floor	4.05	7.34	4.05	21.45
	Third Floor	4.19	8.53	4.19	29.45
	Fourth Floor	1.22	6.95	1.22	27.30
	Fifth Floor	3.65	6.33	3.65	29.92
	Sixth Floor	-0.31	6.57	-0.31	29.91
Server	Server	4.75	5.06	4.75	13.05
Building Total		**-2.63**	**1.52**	**1.52**	**16.37**

Green: Table 3.5 thresholds met; Red: thresholds not met.

6.3.3 Advanced validation

Advanced cross-validation was done for Level 2D, which includes two parts:

1. Checking the air temperature in different space types for a year. Statistical metrics of MAE and RMSE (as defined in Table 3.5) are used to ascertain if the deviation is within acceptable limits.
2. Uncertainty-based validation to establish the robustness of the model by identifying the range of calibration errors by developing the best and worst scenarios.

Temperature checking

Temperature checks across different space types showed variations in temperature in the monitored and the simulated data, especially, on holidays and weekends. Use of granular set-point temperature schedules for the monitored spaces, varying the schedules seasonally and by day of the week, the temperature variations between the simulated and the monitored data were reduced. The simulated temperatures followed the monitored data well, except for some weekend and overnight variations. These could have been further fine-tuned with finer control over set-points. However, the deviations were not huge and fluctuated at normal levels over longer periods. Table 6.5 gives the temperature calibration statistics over the year for some of the spaces, and the variation across the sample spaces is captured in Figure 6.11, which shows the spread of the error.

Table 6.5: MAE and RMSE of temperature calibration check for the Level 2D-ADV calibrated model

Zone	MAE (°C)	RMSE (°C)
Open Office 3F	1.29	1.27
Open Office 4F	0.76	1.45
Cellular Office 4F	1.09	1.61

Green: Table 3.5 thresholds met; Red: thresholds not met.

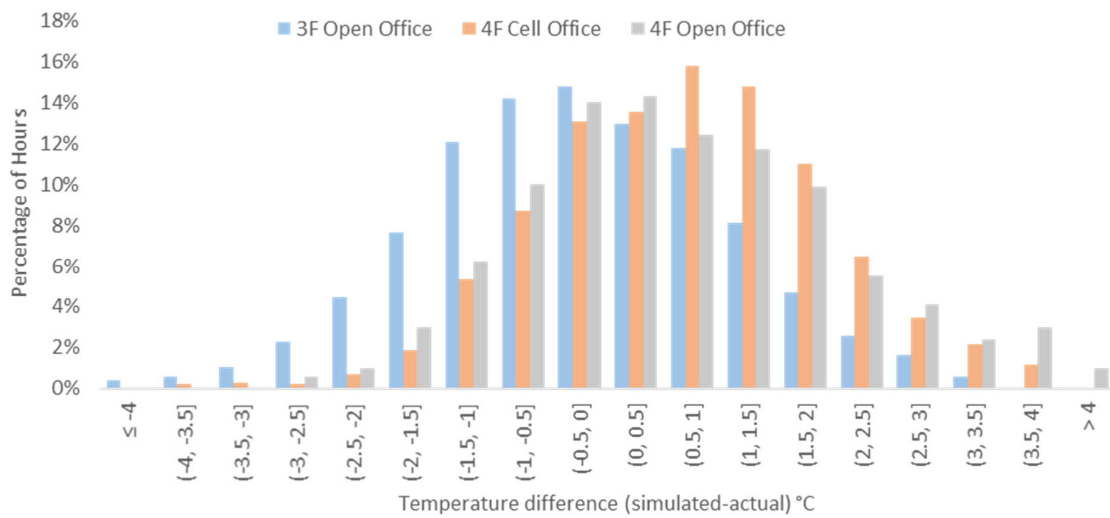


Figure 6.11: Histogram of hourly temperature difference between simulated and actual data for typical spaces

Uncertainty checks

Current building use patterns, controls and system configurations (such as user control of space temperatures) makes defining operations in a calibrated model uncertain, beyond typical averages. The current level of disaggregation of total electricity use in heating + cooling and lighting + power, being combined, is not detailed enough to help understand the underlying behaviour of the building. The model could mask some of the effects, such as levels of partial occupancy, when certain spaces are being heated or cooled and the contribution of domestic hot water in energy totals. Due to the aggregation of measurements for systems, mitigating this difference is difficult, as the actual operation cannot be analysed. Completely eliminating this uncertainty is not possible. In this case, study, to analyse the effect of the uncertainty, the best-estimated trends can be used. For example, in the case of temperature checks, some of the simulated temperatures are 0.5°C outside the threshold limits. This is because the current calibrated model input settings are based on average trends of energy demand and space temperatures, which have been created as typical patterns for weekdays and weekends and for different times of the year (seasons, holidays, etc.). These inputs could be even more granular to further fine-tune the further, which, however, will not be based solely on monitored evidence but also on assumptions and anecdotal observations.

To address this, uncertainty-based validation establishes the robustness of the model. Uncertainly was established as the standard deviation in the average trends in earlier graphs, such as in Figure 6.4, Figure 6.6 and Figure 6.7. The standard deviations in those graphs are the typical ranges within which the values for these trends lie. Along with this, other uncertain inputs, relating to fabric performance, were also defined for the typical variation to create the best and worst scenarios. Figure 6.12 shows the spread of the high and low (best and worst) scenarios and their calibration criteria mapped with the actual data. Table 6.6 lists the uncertain inputs and the reason for their deviation and the ranges. The deviations observed in the building-level totals, even for the extreme range scenarios, are within two times the acceptable range of calibration, which is reasonable.

Figure 6.12: Calibrated model uncertainty checks of the Level 2D-ADV model. (Top) Monthly data; (Bottom) Histograms for hourly data for each month

Table 6.6: Uncertain inputs and their deviation in the Level 2D-ADV model

Input	Deviation
Occupancy	Up to ±20% variation in occupancy schedule and occupancy density per space type and time of the day (determined from lighting + power electric load trends)
Light and Power	Up to ±20% variation in lighting + power schedule and loads per space type and time of the day (determined from lighting + power electric load trends)
Temperature Set-points	Up to ±1°C in winter and ±2°C in summers based on deviations seen in zone temperature measurements (maintaining heating vs cooling dead band of 2°C)
System baseload	Up to ±20% variation per space type and time of the day (determined from heating + cooling electric load trends)
U-Value	±10% due to poor documentation of changes during refurbishment and could not be verified on-site

Detailed monitoring evidence can be obtained to reduce the deviations. Also, the probability across the range values will more likely have a normal distribution rather than a uniform one, i.e. the upper and lower extremes of the range are less likely to occur than the central values. A probabilistic uncertainty assessment could be done to determine the more likely scenarios instead of extreme ones, but that lies outside the scope of this thesis.

6.4 SUMMARY

This chapter provides an understanding of how the temporal granularity of data can affect the accuracy of a calibrated model. The university office building model in Chapter 5 (Level 1A-SIM) was calibrated for monthly energy use with simple cross-validation. That model masked some of the underlying data at an hourly level. Also, as the model was calibrated for building-level total electricity use, significant cross-compensation was seen between end-uses and floor-wise lighting + power sub-meters.

In this chapter, the Level 1A-SIM model for the university office building was calibrated for higher granularity, separately for end-uses and other sub-meters, with advanced validation (Level 2D-ADV). This Level 2D-ADV model provides another level of detail to enable better calibration. Further fine-tuning was done in the Level 1A-SIM model to achieve a Level 2D-ADV calibrated model, and the modified parameters are summarised in Table 6.7.

Table 6.7: Further fine-tuning for Level 2 calibration

Input	Deviation
Occupancy	Using average floor level lighting + power electricity use data, bespoke schedules for each floor were created for different times of the year for weekdays, weekends, and holidays.
Small power and lighting load and schedule	Analysing the floor level lighting + power electricity use data, baseline loads were redistributed for each floor. Also, adjustments for peak load and diversity were done to the schedules to address the monthly variation and holidays.
Heating system and temperature set-points	Based on system sub-meter data, an increased start-up load for a couple of hours in the morning was modelled. Also, based on temperature monitoring of typical spaces in the building, set-point schedules were changed as per average trends for different times of the year for weekdays, weekends, and holidays.

Advanced cross-validation done for this model verified that the system, running overnight, had high set points even during unoccupied times. However, there were some limitations. As the heating and cooling were provided by a VRF system, it was not possible to isolate the space heating from cooling energy. This is important in London's climate where summers are mild and cooling load can be driven by internal gains, especially in an office building. This can often lead to instances of simultaneous heating and cooling in different zones. Also, the building envelope construction was not documented well during the refurbishment; therefore, the issues and uncertainties about material properties could not be assessed. However, in the Level 2D-ADV advanced model, the residual uncertainty in the overall results was not very high.

Uncertainty-based calibration can be used to reinforce confidence in a calibrated model, ensuring that even with input uncertainty the measured data lies within a reasonable output range. A probabilistic approach can be used to resolve the limitation of using just the statistical criteria, as they fail to address the fact that multiple solutions may exist that do not necessarily reflect the real performance.

Monthly calibration with simple validation can provide sufficient accuracy for identifying main performance gap issues. However, its predictive accuracy reduces gradually at higher granularity as cross-compensation of finer data occurs at aggregated levels. Onerous and highly-granular submeter-level advanced-calibration is required here higher accuracy is needed (such as in system and sub-system level assessments). Therefore, the level of calibration can be chosen depending on the application and resources available.

7 CROSS-SECTORAL PERFORMANCE GAP ISSUES

Depending on the application of the calibration model the level of accuracy required can be different. In applications such as assessment causes of performance gap, the intent of the calibrated model is to ensure that the major trends in the building behaviour are captured. Calibrating energy models at monthly temporal resolution ensure that the key trends and seasonal variations are accurately matched, thereby providing the conformation for identified performance gap issues. However as seen in the Chapter 6, the same model may not be suitable for applications where more accurate fine-tuning is required such as detailed system analysis or occupant behaviour analysis. Therefore for the application of identifying performance issues in buildings, the monthly calibrated model created in Chapter 5 can be considered to reasonably represents the actual building. The identified changes in the building, its systems and overall operations that were verified during site investigations and led to the creation of the calibrated model are assumed to be the key factors causing the performance gap. In this final stage (i.e. in this chapter), following the process in Section 3.4.1, an operational-baseline is created and the gap due to technical issues is calculated for four case study buildings: the public office, school, hospital and apartment block. These buildings are examples of the four building sectors under consideration in this study. These four buildings, newly built, had an advanced level of design-stage information available for them. Therefore, they were suitable for this assessment. The fifth case study, the university office building, being a refurbishment project, had limited design-stage information and calculations for detailed assessment of causes of the gap.

The performance issues identified in the individual case studies in Chapter 5 are analysed in a more generic context. For example, the lessons from the school building might be applicable to other schools as well, and the root causes of the issues might be applicable to other buildings as well. Therefore, to understand the root causes of the gap in the case studies in detail and to identify the cross-sectoral issues, the following is presented in this chapter:

- First, all the deviations identified for the buildings in Chapter 5 and validated by calibration (summarised in Table 5.1, Table 5.2, Table 5.3, Table 5.4 and Table 5.5) are divided into those that were due to changes that happened for the building to perform its functions in practice and those that were due to technical issues and shortcomings.
- Then, in the calibrated model, reverting the technical issues to their design intents, a new operational-baseline is created, thereby quantifying the performance gap due to the technical issues and due to the effect of changes in operational requirements.

- Further, analysing all the case study buildings, common themes for performance issues that may also be applicable to other buildings, are used to determine cross-sectoral problems.
- Also, findings from IEQ data are analysed and potential underperformance issues and conflicts related to energy objectives are identified.

In this chapter, first, the classification of operational and technical factors responsible for building underperformance is shown. Then, new operational-baselines developed are compared against the design predictions and the actual gap is ascertained. The last two parts of the chapter present common themes and lessons identified by analysing the buildings' performance that can impact a large cross-section of the building stock.

7.1 Categorisation of the Causes of the Performance Gap

The calibration process validated the identified issues in the case study buildings by ensuring that the changes made in the calibrated model can account for most of the performance gap. Based on the analysis of the deviation of the model input parameters in the Level 1A-SIM calibrated models against the as-design model, the key performance gap issues are listed in Table 5.1 Table 5.2, Table 5.3, Table 5.4 and Table 5.5 for each of the case studies. To assess the issues in detail, based on engineering judgement, these issues are categorised as per their root causes into either those related to changes in operational requirements or those arising from technical issues. Table 7.1 lists all the factors responsible for building underperformance, classified into changes in operational requirements and technical issues.

Table 7.1: Operational and technical factors responsible for building underperformance

Building	Changes in operational requirement	Technical Issues
Public office	• The total occupancy of the building was about 25-30% higher, leading to more workstations being added. • Extended operation hours led to longer use of small power and lighting. • Hot desking, using a thin-client IT system, by using flexible workstations was not followed in practice. The departmental structure of the occupant organisation restricted movement. Therefore, hydraulic isolation of heating and cooling systems in unoccupied zones during out-of-hours use of the building could not be followed. • Heating set-point temperatures maintained were about 2°C higher than the design intention, driven primarily by occupant needs.	• Heat loss due to leakage through poorly sealed vents and at construction junctions. • Parasitic loads were identified when the building was unoccupied. • Server loads were overestimated in design. This adversely impacted the heating system efficiency as there was less free heat. • Technical issues with heat exchangers and the flow rate specification caused the heat pumps to malfunction. • The heating terminal's sizing was not consistent with the low-temperature flow required for energy-efficient operation of the system. • Some of the ventilation control sensors malfunctioned and required a subsequent modification to the control strategy to overcome the system shortcomings. • Overriding of automatic vent opening to avoid thermal discomfort due to drafts around the floor cut-outs.
School	• All spaces were not occupied throughout the day. Any given space is occupied only 60-70% of the full working day. • During academic breaks, the school is not completely shut. There are extra-curricular activities and events, especially during the summer holidays. • Indoor temperatures monitored during the winter season were about 2°C higher than the set-point temperatures and sometimes even more.	• Significant thermal bypasses around insulation were observed as the tested U-values were much higher than the design. • Daylight sensors for automatic dimming of lights were poorly commissioned and PIR sensors had very long delay times. • The biomass boiler, installed for decarbonising energy use, was never used, citing practical and logistic issues of using biomass as fuel. • Actual boiler efficiency (tested) was 88% compared to the design intent of 96%.

		• Centralised system design (one control and sensor for multiple zones) led to inefficient operations during out-of-hours use. • The Specific Fan Power in the AHU commissioning sheets was 66% higher than the values used in design-stage estimations. • The supply fan during unoccupied times was running inefficiently, operating at 30% to 40% of its nominal speed. • Building heating and ventilation systems were operating during unoccupied hours.
Hospital	• Different clinical processes had different needs. It was difficult to generalise typical operational trends of various spaces; their equipment loads; and set-point temperatures. • An increased number of beds were observed in the patient wards during the site visits. • Hospitals had a high changeover rate of patients. Therefore, epistemic uncertainties in set-point temperatures and operations of hospital equipment were quite high. • Hospital hot water energy use was high due to the clinical requirements.	• A low efficiency, old, steam-based central heating network serves the building. A new CHP plant was to be installed for the entire facility rather than just for the new building to maximise savings. However, this has not happened yet. • Fans and pumps are used to provide close control to the indoor environment. Demand control ventilation could have been used more effectively and heating energy use could be reduced whilst maintaining comfortable indoor temperatures.
Apartment block	• Occupancy of the individual flat is very uncertain to determine, especially holidays. • CO_2 monitoring show a possible lack of occupancy during school holiday times. • There are instances where it is evident that during summer, the windows were opened for enhanced ventilation, while the heating system was still operational. • Occupants had full control over the heating system & Mechanical Ventilation with Heat Recovery (MVHR), determining their set-points and operations.	• Maintaining filters in the de-centralised MVHR systems was left to the occupants and therefore there were maintenance issues (e.g. dirty filters) in most flats. • The efficiency of the district heating network, estimated at 91% in design calculations, was around 50% in practice.
University office	• Measured occupancy varied monthly and seasonally. It was generally higher than that assumed for out-of-hours use and weekends. • High set point temperature maintained during occupied times (23°C) and unoccupied times (21°C).	• Baseline energy use for lighting and power consumption is very high at about 50% lighting and power load during the unoccupied time. • Systems were operated throughout the day and night, even during unoccupied times • AHU in the basement not working and subsequent more opening of windows in the basement.

Some of the deviation identified in the buildings might be due to a combination of operational changes and technical issues, such as in the case of university office where there was a deviation with its heating system. Unlike the school building where there was a same deviation due to operational changes, the increased heating energy use in the university building was due to a combination of increased heating setpoint (operational change) and longer duration of system running (due to technical issue). In such cases the issues are split or proportionally attributed to each factor, based on engineering judgement and discussions with stakeholders.

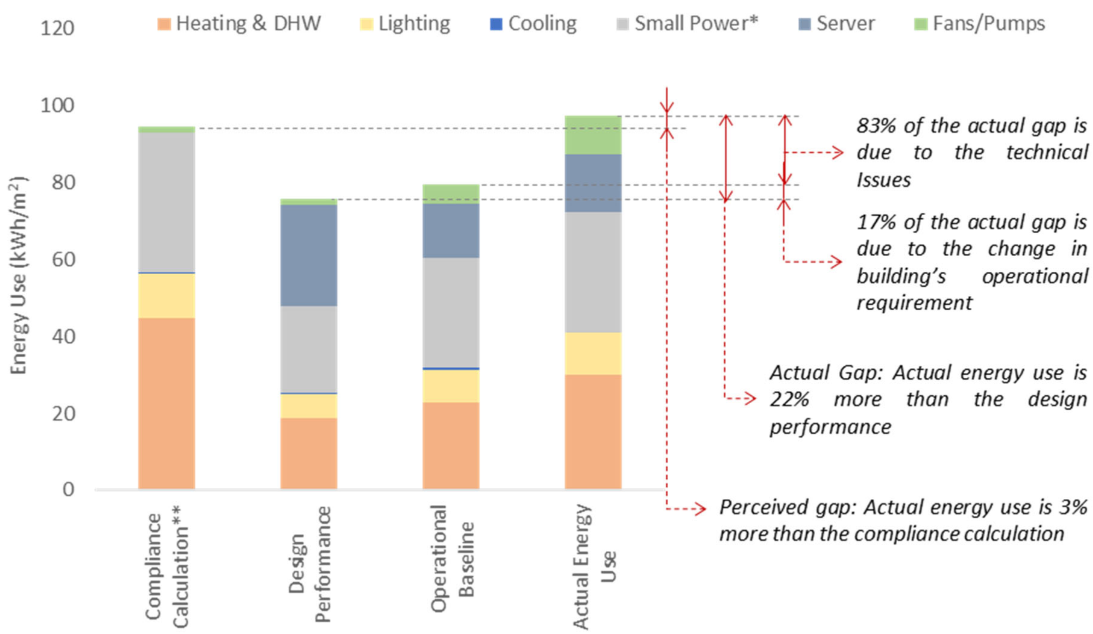
7.2 Operational-baseline and the actual gap

To quantify the performance gap due to technical issues, all the technical issues identified in Table 7.1 were reverted to their design intents according to what was captured during data collection. In other words, in the calibrated model, the technical characteristics of the areas with technical issues in the building were converted back to their values and settings intended in the design stage. For example, in the school building, the U value measured as 0.78 W/m^2K and used to calibrate the model, was reverted to 0.17 W/m^2K as used in the design specification and the design model. The model thus created, by reverting the technical issues to design intent, is the new operational-baseline which the building should be running at to meet its functional requirements without any technical issues or shortcomings. In this section, for the four case studies the new operational-baseline is compared against compliance calculation results, design performance projections (i.e. CIBSE TM54 based performance projections at the design stage) and actual energy use.

7.2.1 Public office

Figure 7.1 shows the comparison of the operational-baseline against the design and actual performance of the public office building and associated gaps between the projections. The perceived gap is lower than the actual gap in this building because the design performance calculations were lower than the estimate of compliance calculations of energy use (with equipment energy use). The lower design performance is because the building was designed to achieve a very high energy rating and had strict performance targets. Being procured under an energy performance contract, along with the compliance calculations (which were done as per regulatory requirements), detailed design performance estimates were separately calculated for this building. Additionally, performance contracting ensured significant aftercare and fine-tuning of the building after handover; therefore, the actual gap is low.

With 83% of the actual gap due to technical shortcomings, the root causes of the performance gap in this building are the technical issues identified in Table 7.1. Even with the current operational changes in the building done to meet the functional requirements, such as additional occupancy and operating set-point temperatures, actual performance can be brought closer to design performance if technical issues with the system, building operations and maintenance are resolved and parasitic loads are reduced. Some of these issues were being fine-tuned under the performance contracting, such as those related to efficiency and operations improvements. However, as the building is now running in a steady mode of operation, fixing other issues that require significant modifications to the installed systems might not be practically feasible at this stage.



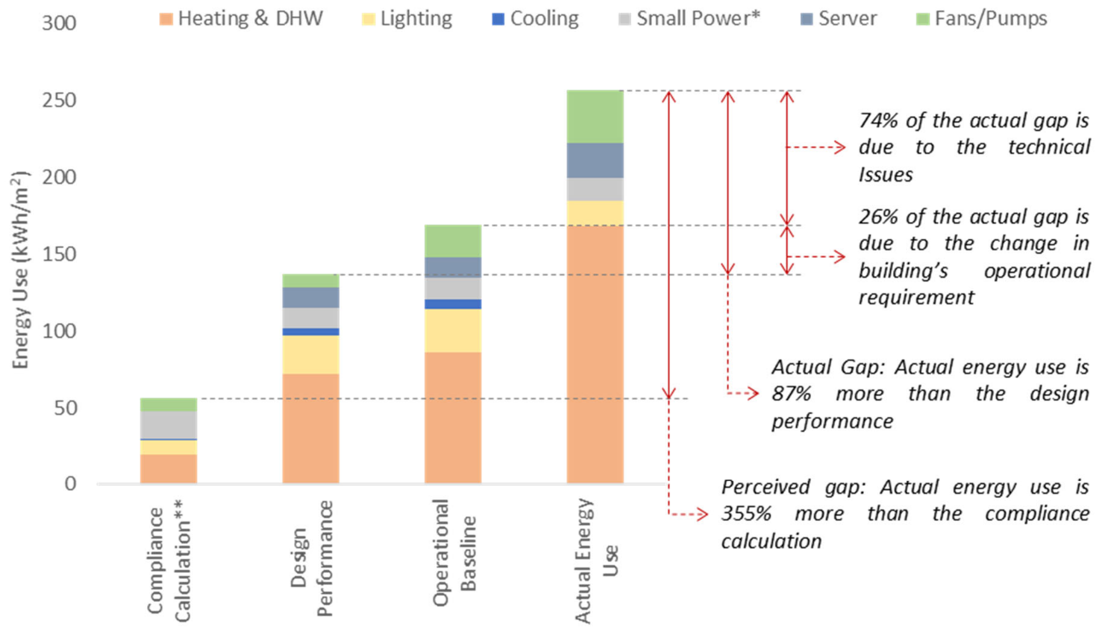
* Small Power includes Cooling energy use in Actual Energy Use; ** Compliance calculation were not used for design stage projections, and in compliance calculation in the UK, small power use is calculations but is not reported in totals, however they have been reported here for a meaningful comparison against other results.

Figure 7.1: Performance calculations and associated gaps for the public office building

7.2.2 School

Figure 7.2 shows the comparison of the operational-baseline against the design and actual performance of the school building, and associated gaps between the projections. In the school, compliance calculation results, created from modelling done as part of Building Regulations compliance, were misinterpreted as a prediction of actual performance, contributing to the 355% perceived gap. CIBSE TM54 design performance calculations were done for the calibration-baseline (as-design) model in Section 5.2.1 and compared against actual energy use to ascertain the actual gap of 87%. Part of the actual gap (26%) is due to necessary operational changes, such as evening classes and extra-curricular activities.

The remaining actual gap (74%), the difference between the operational-baseline and actual energy use, is the performance gap due to the technical issues identified in Table 7.1. Changing trends in the use of school buildings, such as partial occupancy during half-term breaks and occupancy after normal operating hours for extracurricular activities and evening classes, have increased space-time utilisation. However, systems not designed for this flexibility, which subsequently lead to inefficient operations and maintenance; highlight the technical problems in the building. Resolving the identified technical issues would address and improve the actual performance.

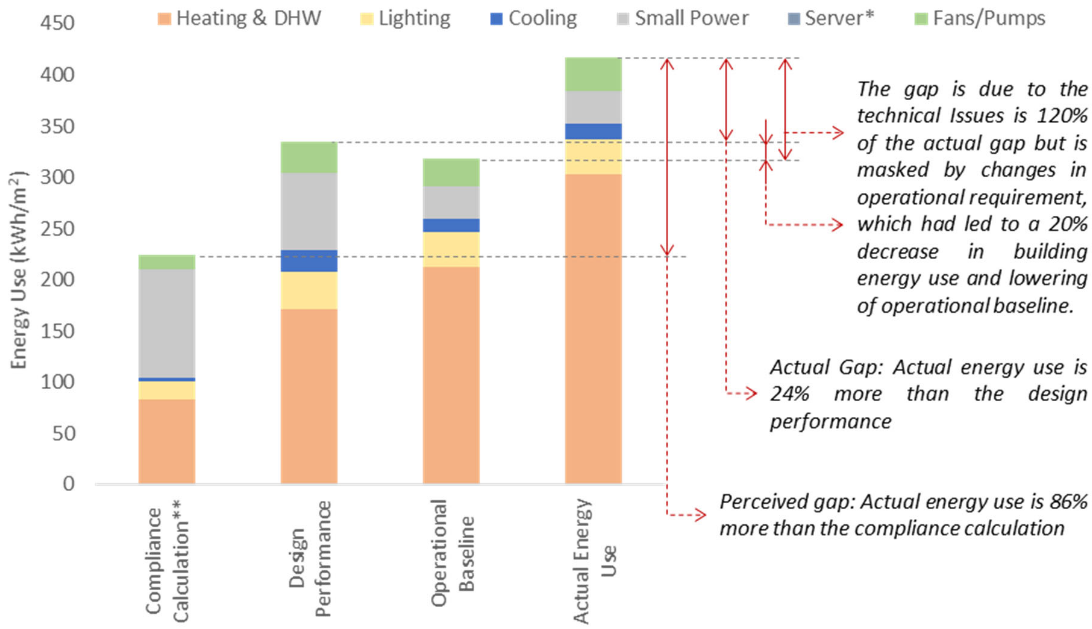


* Small Power includes Server energy use in Compliance Calculation, and Small Power includes Cooling energy use in Actual Energy Use; ** In compliance calculation in the UK, small power use is calculations but is not reported in totals, however they have been reported here for a meaningful comparison against other results.

Figure 7.2: Performance calculations and associated gaps for the school building

7.2.3 Hospital

Figure 7.3 shows the comparison of the operational-baseline against the design and actual performance of the hospital building, and the associated gaps between the projections. In the hospital, compliance calculation results, created from modelling done as part of Building Regulations compliance, contributed to an 86% perceived gap. CIBSE TM54 design performance calculations were done for the calibration-baseline (as-design) models in Section 5.3.1 and were compared against actual energy use to ascertain the actual gap of 24%. When the operational-baseline model was created from the calibrated model, the hospital operational-baseline was lower than the design performance. This was because the design-stage estimates for equipment loads and their use, assumed as per regulatory standards and typical assumptions, were much higher than those seen in the building during actual operations. As a result, 20% of the actual gap due to technical issues was masked by this overestimation. This highlights the specialist nature of hospitals, which have high energy consumption and irregular and often atypical energy use patterns, and are highly dependent on varied/specialised functional requirements. The root causes of the technical performance gap in this building are identified in Table 7.1; actual performance could be much closer to design performance if the technical issues are resolved.

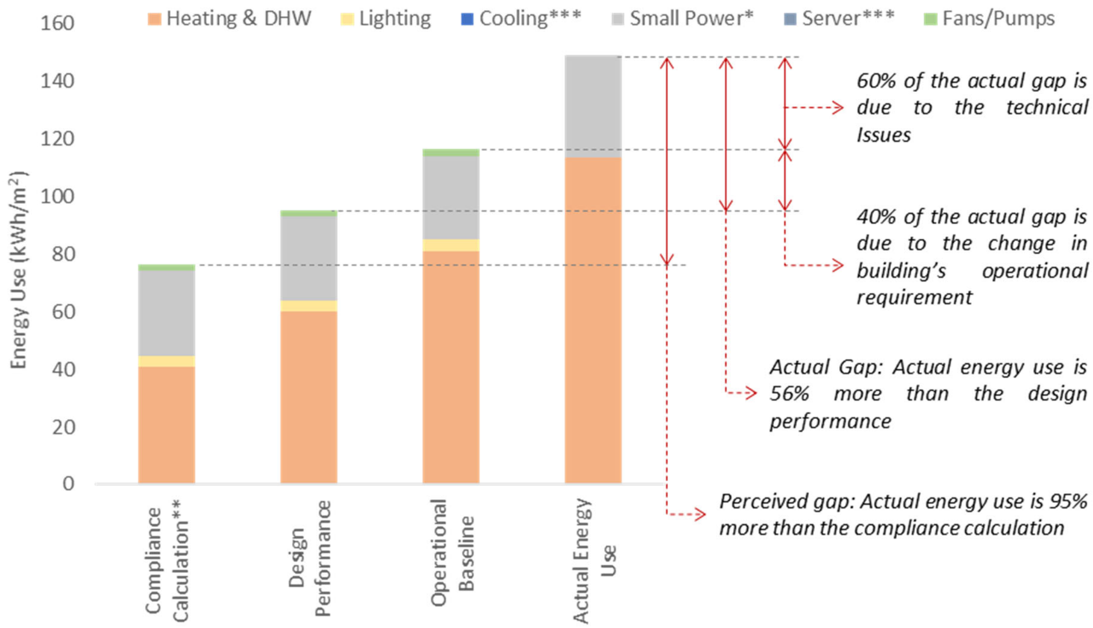


* Server is not present in the building; ** In compliance calculation in the UK, small power use is calculations but is not reported in totals, however they have been reported here for a meaningful comparison against other results.

Figure 7.3: Performance calculations and associated gaps for the hospital building

7.2.4 Apartment

Figure 7.4 shows the comparison of the operational-baseline against the design and the actual performance of the flat in the apartment block building, and associated gaps between the projections. In the apartment, compliance calculation results, created from modelling done as part of Building Regulations compliance, were misinterpreted as a prediction of actual performance, contributing to the 95% perceived gap. CIBSE TM54 design performance calculations were done for the calibration-baseline (as-design) model in Section 5.4.1 and compared against actual energy use to ascertain the actual gap of 56%. Compared to other buildings, the apartment has a noticeable (40%) gap due to operational changes due to the high level of control that occupants have on the operations in the apartment as well as their variable usage patterns. The remaining actual gap (60%), the difference between the operational-baseline and the actual energy use, is the performance gap due to technical issues, such as system issues or poor maintenance. The root causes of the technical performance gap in this building are identified in Table 7.1; however, high operational uncertainties in the flat mean that actual energy use would still be higher even if the technical issues were resolved.



* Small Power includes Fans/Pumps & Lighting in small power use; ** In compliance calculation in the UK, small power use is calculations but is not reported in totals, however they have been reported here for a meaningful comparison against other results; *** Cooling & Server are not present in the building

Figure 7.4: Performance calculations and associated gaps for the typical flat in the apartment block

7.3 Cross-sectoral energy performance issues

Analysing the various issues identified and validated across the case studies by model calibration, that were contributing to the performance gap revealed many common themes and lessons. While drawn from individual cases, these findings are supportive of other similar studies (Burman, 2016; Palmer, et al., 2016b; Palmer, et al., 2016a). Therefore, the issues identified are considered to be endemic in the industry and are likely to have applicability to other buildings and across the construction sector.

7.3.1 Design-stage energy projections

Building Regulation compliance models use simplified calculations intended to ensure that minimum regulatory requirements are met and to benchmark energy use for the entirety of the building stock. The use of regulatory compliance calculations as the design-stage predicted performance may significantly underestimate actual energy use and lead to a perceived gap. This was seen in the school (Figure 7.2), hospital (Figure 7.3) and apartment block (Figure 7.4), where compliance calculation results, typically reported without small power energy use, were orders of magnitude lower than the actual performance.

To avoid this perceived gap, realistic estimation of energy use at the design stage should be mandatory, calculated as per CIBSE TM54 guidelines or other equivalent protocols, accounting for all end-uses in the building alongside realistic operating patterns and occupant behaviour. This was carried out in the public office building, where, in addition to the regulatory compliance requirements, a contractually mandated performance target to beat the operational benchmarks was the driver for the design team to undertake realistic calculations.

7.3.2 Performance targets and contractual accountability

It is common that, after handover, engagement of the design and construction team with the building is minimal. This means that the focus is on delivery and system functionality rather than performance. Shortcomings in commissioning and qualitative design issues subsequently identified largely remain unaddressed. Enhanced engagement for all stakeholders, along with accountability of the people delivering the building, is necessary to ensure that the building performs as expected on both energy and IEQ fronts. Soft Landings Framework and performance contracting approaches can be used as frameworks to achieve this.

Due to a strict operational performance target, the overall energy use in the public office is not just lower than other building types (Figure 7.1) but also much lower than the benchmarks within its building category (Table 4.3). Performance contracting, integrated within a Soft Landings Framework, guaranteed the involvement of the contractor and the designers after the handover to fine-tune the building operations and fix issues related to design, construction and commissioning. This increased accountability also led to robust construction quality checks, thereby mitigating issues such as the significantly high U-values seen in the school (Figure 5.9, Figure 5.10).

One of the challenges in performance contracting is the contractual period, which might determine the sustainability measures used. Also, if key sustainability measures that are beneficial in the longer term are not safeguarded from the start, some may be value-engineered out during the construction process. Another challenge is to objectively define the targets and the metric to use, i.e. ensuring that the metrics are in alignment with the actual intent. For example, targeting net carbon emissions may result in a different approach than when the focus is on energy use only. Also, in certain building types, such as hospitals, current benchmarks have very broad categories. As seen in the case study, within each category the energy demands can vary significantly as per the specialised services provided. In these cases, energy analysis and benchmarking need to look beyond the entire building and consider departments or other sub-spaces as the unit of analysis.

7.3.3 Uncertainties in building use

Operational uncertainties also affect design predictions. Occupant behaviour is uncertain and has a significant effect in cases where occupants have a high degree of control, such as in apartments. Small events can affect a flat's energy use significantly, such as a sharp drop in energy use during holiday time, as in the case study flat in the apartment block. Uncertainty in building use is also a reason for the wide variation in the energy use of the 98 flats (Figure 4.21) within the apartment block. In most large non-domestic buildings, the services are controlled centrally by building managers, limiting the occupant effect to manual overriding of zone thermostats and operation of windows. However, sometimes the building function is a cause of uncertainty, for example, the irregular nature of processes in hospitals making it difficult to identify typical events and average durations of use. The transient nature of occupants and their needs further added to this uncertainty. To address them, sensitivity and scenario analysis can be used to implement informed design decisions and operational safeguards regarding the most important factors.

7.3.4 Transitionally/seasonally occupied buildings and changing user patterns

Buildings continually evolve in the way they are used, and this leads to a difference in how we perceive a building and how it truly operates. Contrary to design assumptions, nowadays schools have higher space-time utilisation. Beyond regular school hours, the case study school had partial occupancy for evening/holiday classes and out-of-hours extracurricular activities. However, the building services were not tuned for low occupancy after-hours use. A centralised system design (one control and sensor for multiple zones) led to multiple spaces (including unoccupied ones) being conditioned during transitional occupancy times. The ability to hydraulically isolate unoccupied areas and use decentralised controls is beneficial in these scenarios.

Flexible space use, including hot-desking in offices, is another practice used to manage transitional occupancy. In the office building, hot-desking was implemented in conjunction with hydraulic isolation. It was assumed that during low occupancy times all users could work in one part of the building, keeping the rest of the building shut. However, due to the departmental nature of the occupant organisation, this was not achieved in practice. Regular users occupied their usual desks within the part of building where their department was located.

The changes in building occupancy patterns and the way people use buildings are determined by changes in operational requirements and are therefore inevitable. However, closer discussions between designers and clients on longer-term use and adaptability of the building can be incorporated during the the design stages.

7.3.5 Use of new and innovative low-carbon technologies

The technologies used across the case studies were partly effective. In the office, the innovative strategy of using free heat from server rooms was not successful due to technical issues with the heat pump. Consequently, the heating load of the building was shifted to the gas boilers, and there was no active comfort cooling. Server cooling was also provided by two backup unitary DX chillers.

In the school, driven by the local council's intention to use and promote low carbon technologies, a biomass boiler was installed. However, this system was not operational post-handover, due to the logistical limitations of managing the fuel for running it and disagreements between the school management and the council. This meant that the expected CO_2 emissions of this building were significantly higher than what was assumed upon completion of the building.

In the hospital too, the new CHP system was planned to be installed at the campus level following a major renovation to maximise the efficiency savings across the facility rather than as a separate system for the new building only. It has not been installed due to budgetary constraints, and the present thermal performance of the building is much worse than expected from a new building.

Therefore, while there was an intent, lack of follow-up and checks meant that the intended benefits were not achieved. At the policy level, steps are required for not only enabling smooth integration of new technologies with conventional practices but also safeguarding them against value engineering.

7.3.6 Protocols for managing building operations

Some of the performance issues in the case studies were due to sub-optimal operations and irregular maintenance. In the office, delays in resolving issues with sensors and automatic roof vents meant that some of the vents were either left open or remained closed.

In the school, sub-optimal operations were partly due to coarse and centralised system design (one control and sensor for many zones) and a lack of user-friendly BMS controls to manage it. A more streamlined building operation and management strategy envisaged in the design stage and incorporated at handover would enable a building to operate reasonably closely to what was assumed at the design stage.

In the apartments, management of the Mechanical Ventilation with Heat Recovery (MVHR) system and replacement of filters was the responsibility of individual occupants. However, the occupants were either not aware of it or did not understand the need/requirement to regularly change them. An effective building operation and management strategy envisaged in design and incorporated at handover would help manage these issues.

7.4 Going beyond energy - Unintended IEQ underperformance

In all case study buildings, energy performance was the only specific quantified performance objective. There was no effective metering, monitoring and reporting strategy for any of the IEQ parameters. However, lower energy use and better IEQ can be divergent objectives. The temperature, CO_2, NO_2 and $PM_{2.5}$ measurements provided insights into the IEQ and suggested their potential energy-related implications. Some of the unintended IEQ performance issues, across the case studies, are discussed in this section.

7.4.1 Thermal comfort

Out of all the case study buildings, the hospital had the most appropriate thermal conditions because the building operated with close control over its indoor environment. However, other buildings, including the public office, school and apartment block, had space heating provision but no mechanical cooling (in school cooling was provided only in selected spaces). This made spaces in those buildings susceptible to overheating risks in hot summers.

In the public office building, there were a few instances where indoor temperatures were recorded in excess of 28°C (Figure 4.4, Figure 4.5). As per the design strategy, night cooling could have kept the building comfortable in summers by maintaining temperatures of 2-3°C less than peak outdoor temperatures. However, due to the limitations of present controls the night cooling strategy became ineffective. To avoid overcooling in the night (and the need for heating the next morning), the vents were only open when the indoor temperature was above 19°C. While this is an energy-saving measure, this restricted night ventilation led to thermal discomfort issues during the hot summer days.

In the school, during the non-heating season, most of the spaces did not suffer from overheating. However, rooms on the south façade, lacking solar controls (blinds/shades) had high heat gains. They were susceptible to overheating risks in hot summers. These issues were worsened by the airtight envelope and inadequate operable windows. The temperature measurements in those spaces sometimes exceeded the maximum temperature threshold of 26°C (CIBSE, 2013a). Overall in the school, while the current overheating risk was not high for long periods of time, in the context of a changing climate and increasing temperatures, the building will need to adapt by using strategies such as night purge ventilation.

7.4.2 Indoor air quality

CO_2 concentrations

CO_2 concentration is considered as a proxy for fresh air intake. Besides the apartment block, which had an MVHR system for minimum fresh air along with operable windows, all other buildings had CO_2-

based controls for ventilation. A mechanical ventilation system provided all the fresh air in the school and hospital, and CO_2-based automatically controlled vents provided natural ventilation in the public office building.

In the controlled hospital environment, CO_2 levels in all monitored wards remained below 950 ppm due to effective mechanical ventilation and high air change rates. To meet these requirements, a mechanical ventilation system provided 10 ach to most medical spaces and 6 ach to examination and measurement rooms. It should be noted that maintaining high IEQ through high air changes and filtration comes at an energy expense. Consequently, fans and pumps used to provide this close control use around 30% of total electricity

However, the indoor CO_2 concentrations on the third floor of the office building reached above 2500 ppm, exceeding the IDA Class II levels for office spaces (CIBSE, 2015). Apart from the effect of stack ventilation that may cause concentration levels to be slightly higher than lower floors, on-site observations found a malfunctioning CO_2 sensor that led to the underestimation of concentration levels on the third floor, meaning the motorised vents were not open when required. There was also the closing of roof vents in winter to avoid cold drafts due to the stack effect and to maintain users' thermal comfort.

$PM_{2.5}$ concentrations

Mechanical ventilation and good airtightness resulted in good control over the filtered fresh air intake. $PM_{2.5}$ and PM_{10} concentrations in all buildings with mechanical ventilation were generally adequate.

However, some air quality issues were seen in the mechanically ventilated apartment block. $PM_{2.5}$ levels in the flat were close to external levels with some spikes due to internal sources. However, they were frequently higher than the WHO annual mean limit of 10 $\mu g/m^3$ (WHO, 2018). The flats used G3 filters; however, as ambient $PM_{2.5}$ levels are typical for London, use of high-grade filtration (F type filters), while slightly increasing electricity use, will help reduce $PM_{2.5}$ levels.

NO_2 concentrations

NO_2 concentrations are primarily driven by external pollution and are an issue in congested urban areas. This was seen across the case study buildings where there was a lack of measures against ambient NO_2. The public office building was naturally ventilated and controlled by CO_2 concentration, and the concentration levels of NO_2 were significantly higher during summer months, as the vents and windows were open. Similarly, in the school, NO_2 concentration levels in the library, which is located next to a main road, were higher than the other monitored zones. In both cases, while the current concentration levels are lower than the WHO limit, the numbers show the potential risk of control strategies that are solely based on CO_2 concentration for IAQ control.

In the hospital, indoor NO_2 levels recorded in a ward very closely followed the external levels, and the recordings were often above the WHO annual mean threshold of 40 µg/m³ (WHO, 2018). This suggests a potential risk of exposure to more than WHO-recommended levels if the external air remains polluted for prolonged periods. Advanced chemical filtration (such as activated carbon) and controls that consider the balance between the requirement for fresh air and protection from outdoor sources of pollution could provide a healthier environment at the same time as saving energy.

7.5 SUMMARY

The level 1A-SIM calibrated simulation model developed in Chapter 5 was useful in identifying and validating the performance issues in the case study buildings. To assess the issues in detail, two things were done:

1. Ascertaining magnitude of the gap: The calibration-baseline model developed for each case study building was modified and simulation was undertaken as per CIBSE TM54 guidelines, to create the 'design performance'. This CIBSE TM54 model (which follows the performance modelling approach – see section 2.1.3) helps to identify the actual gap and distinguishes it from the perceived gap.
   Calculations done for building regulations compliance often underestimate energy use and do not report all energy end-uses in building totals. Comparing results from compliance calculation based on compliance modelling against actual performance provides an inflated perception of the gap, called the perceived gap. Design performance is correct baseline[*] to use for performance gap analysis because it is the realistic estimate of the building's energy use, and therefore the difference between the design performance and the actual performance is the actual gap.
2. Analysing the root causes and their impact: To assess the issues in detail, based on engineering judgement, the issues identified during calibration, listed in Table 5.1, Table 5.2, Table 5.3, Table 5.4 and Table 5.5, are categorised as per their root causes into:
   a. Issues related to changes in operational requirements, or,
   b. Issues arising from technical shortcomings.

   Reverting all the technical issues that were identified in the case study buildings to their design intents, according to information on the technical aspects gathered during data collection, led to the creation of a new operational-baseline. This model simulates the building as it should be running, with all operational changes to meet its functional requirements, but without any technical issues or shortcomings. Common operational issues across the case study buildings included changes/diversity in occupancy and operational patterns, modifications to environmental control set-points and adjustments to operating regimes and maintained schedules due to functional needs.
   Technical issues were specific to each building, however, there were some common factors that contributed to these issues. These issues involved system inefficiencies due to design faults or construction/commissioning oversights or improper operation and maintenance. A

[*] reference point for calculating the performance gap

key aspect between these was that a lot of the issues were related to new and innovative low-carbon technologies.

Analysing all the performance issues, many common themes and lessons were identified which can be considered to be endemic in the industry and are likely to have applicability to other buildings and across the construction sector. These include the following:

- Design-stage projections of energy use: the mistaken interpretation of building regulations compliance calculation as a projection of the energy performance of the building.
- Performance targets and contractual accountability: use of binding performance targets forces streamlining of efforts from the design and construction team well into the operation stages and minimises the risk of value engineering.
- Uncertainty in building use: occupant behaviour is a key determinant in energy performance; facilities, where there is a high degree of occupant control, have been observed not to perform as well as intended.
- Transitionally/seasonally occupied buildings and changing user patterns: longitudinal changes to the typical operations of the buildings, due to changes in functional needs, can test the flexibility of existing systems.
- Use of new and innovative low-carbon technologies: these technologies take time to integrate and to become conventional. Safeguarding and supporting these is necessary until that time.
- Protocols for managing building operations: training of building occupants and managers and subsequent operations and maintenance of buildings by them, if not undertaken regularly and lead to problems that may take long times to detect and fix.

Due to the industry emphasis on improving the energy performance of buildings, energy performance is often the only specific quantified performance objective for which detailed calculations and estimations are done. Effective metering, monitoring, and reporting strategies for any of the IEQ parameters are often never undertaken and evaluated in detail. However, lower energy use and better IEQ can be divergent objectives. Findings from the case studies show potential conflicts within IEQ objectives and between energy and IEQ, especially in new low-energy buildings constructed in polluted urban areas. This is primarily because, during operations, unlike energy, IEQ performance is not an explicit target and therefore is not a focus for building monitoring and fine-tuning.

Common energy-related IEQ issues include

- Summer overheating in airtight spaces without adequate cooling provisions, either mechanically or naturally. These are directly linked to space conditioning energy use.

- Lack of adequate fresh air, especially in heating seasons, where there is a tendency of occupants to keep windows closed. This reduces the burden on the heating system and supply fans (in mechanically ventilated buildings) but affects the health of occupants.
- The use of proper filtration of external air is necessary for pollutants such as $PM_{2.5}$ (driven by external and internal sources) and NO_2 (driven primarily by external traffic), especially in dense urban environments and naturally ventilated buildings. These affect fan efficiencies.

Some of the issues identified here will lead to an increase in the energy demand of the buildings to meet the IEQ requirements. In those instances, where maintaining better IEQ might result in slightly increased energy use, the energy use increase needs to be offset elsewhere.

8 SYNTHESIS OF KNOWLEDGE

Buildings underperform post-completion when compared against the performance predicted during the design stage. The difference between the actual energy use and the design intent is called the *performance gap* (Carbon Trust, 2012; de Wilde, 2018).

The performance gap is a commonly used term in the context of building energy use, but its definition is quite vague. Depending on the baseline[*] chosen or the calculation protocols used, the magnitude and the cause of the gap can vary (Burman, 2016). Due to the assumptions used in defining the input parameters in design models, some variability in simulation outputs is to be expected. However, the scale of the discrepancy reported is very wide and reduces the confidence levels in the results of simulation tools (van Dronkelaar, et al., 2016). A calibrated computer model, which is an 'operationally accurate virtual representation' of the actual building, can be used to investigate the issues and identify measures to fix them.

With this background, this chapter describes the key findings from this thesis, linking them to the context, aim and key objectives of the research described earlier. The central aim of this research was to **develop a simulation-based framework for a procedural examination of the root causes of the performance gap in buildings**, focusing on energy use and energy-related unintended IEQ underperformance. A model calibration assisted methodology was developed and used to achieve that aim. In this chapter, the basis of the creation of the methodology is explained first, followed by reflections on each of the specific objectives (in Section 1.2), i.e.:

1. Analyse the operational performance of buildings for their energy use and the performance of energy-related IEQ parameters.
2. Develop a systematic and replicable framework for M&V, using calibrated building performance simulation models to identify performance issues.
3. Explore model calibration and improved validation approaches for implementation within M&V, when used in the context of performance issues identification.
4. Examine the findings from different types of buildings and determine the cross-sectoral performance issues and their underlying root causes.
5. Assess the unintended IEQ underperformance that might occur when the focus of the design is primarily to meet increasingly stringent objectives of energy efficiency.

[*] reference point for calculating the performance gap

8.1 REFLECTIONS ON METHODOLOGY: PRINCIPLES AND DEVELOPMENT

Most design-stage energy calculations are conducted in the context of compliance modelling, which is primarily the modelling undertaken for regulatory compliance. Performance modelling, on the other hand, using the CIBSE TM54 methodology, makes simulation results (at the design stage) more representative of the actual performance. However, if applied to operational buildings during M&V, performance modelling done as per CIBSE TM54 does not provide a method to separate discrepancies in the operational settings and functional use changes that have evolved over time from the technical issues in the building and its systems. Therefore, it may not give a full picture of any performance gap and its underlying causes (issues related to construction, commissioning and operations). A new measurement and verification framework aiming to close the performance gap should be able to address the following issues:

1. Use an appropriate baseline* to analyse the performance gap: Most design energy calculations are carried out in the context of compliance modelling, and often their results are used as a baseline* to assess the performance gap, which often gives an inflated estimate of the gap. Performance modelling using the CIBSE TM54 methodology makes simulation results more representative of actual performance at the design stage, and therefore can be compared with actual energy use to ascertain the actual gap.
2. Need for structural and procedural verification of the most prominent issues: Conventional measurement and verification protocols generally focus on statistical requirements and are not tied to a framework for a procedural verification of all issues. Therefore, it is not certain that the technical issues uncovered in a building through, e.g., on-site investigations, reflect all or most of the key causes of the performance gap. It is likely that some key issues are identified during the investigations whilst other potential issues are not uncovered. Therefore, there is a requirement for a systematic and robust way to identify most of the issues that cause underperformance.
3. Identify deviations of operating conditions from the design assumptions: Buildings evolve over time, and operating conditions get changed from those assumed at the design stage. Therefore, the framework should identify the deviations that are primarily driven by the building's function and its actual occupancy (such as changes in occupancy hours, occupants maintaining of higher heating set-points and changes in the functional use of the spaces).
4. Identify and quantify the effect of technical issues: Technical problems that cause the performance gap between the design intent and the actual operation (such as thermal bypass

* reference point for calculating the performance gap

in the envelope due to missing insulation, incorrect boiler operations, and faulty sensors preventing dimming of lights) need to be identified, quantified and separated.

These requirements were achieved by making dynamic thermal simulation and model calibration integral to measurement and verification. Use of correct design stage performance modelling approach addressed the first issue of a proper baseline*, and the use of model calibration helped to address the other issues. The systematic calibration-assisted method used was designed to find and validate the building-level performance issues, identifying deviations in operating conditions as well as technical issues separately.

The framework starts with collecting data and comparing actual data against the design information to quantify the magnitude of the gap and to identify discrepancies between the two datasets. Next, a calibration-baseline (as-design) model is created using the design data. In case the design stage performance calculated by the design team is not calculated correctly (e.g. it is not as per CIBSE TM54), then the actual gap is calculated by comparing the performance from this new calibration-baseline (as-design) model against the actual energy use.

In the next stage, the calibration-baseline (as-design) model is calibrated using evidence-based fine-tuning as per the operation-stage information and metered data. The calibrated model's energy results are validated through standard statistical checks (NMBE and C_V(RMSE)) and further cross-validated by checking IEQ parameters (such as zone temperatures) predicted by the calibrated model against the monitored data. The identified changes in the building, its systems and overall operations, that are verified during site investigations, have led to the creation of the calibrated model. Those changes, which helped to create the calibration models, are assumed to be the key factors causing the gap.

In the final stage, an operational-baseline is created and the gap due to technical issues is calculated by reverting the technical changes in the calibrated model to their design intents. For example, if the technical issue was with the building envelope and a higher U-value was used in the calibrated model (based on measurements during building evaluation), then in the operational baseline model, the U-value should be reverted to the lower value used as noted in the design specification or used in the calibration-baseline (as-design) model. Reproduced from Figure 3.8, Figure 8.1 shows the various performance calculations and their associated gaps.

* reference point for calculating the performance gap

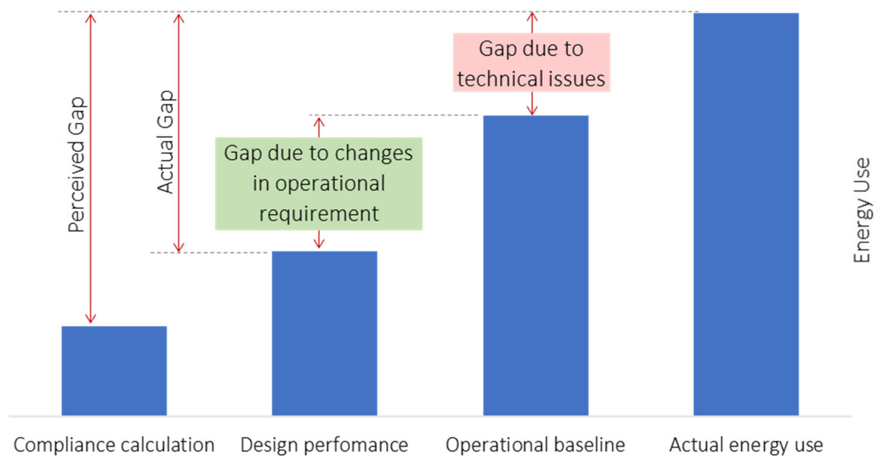


Figure 8.1: Performance calculations and associated gaps

This new methodology, developed to be structured, procedural and replicable, was applied to various typologically different case studies. Their energy and IEQ performance are analysed by focusing on three key aspects:

1. Improvements in model calibration processes for better measurement and verification, and identification of key factors responsible for the performance gap.
2. Analysis of performance gap findings to determine the common issues across the case study buildings which could be endemic across the construction industry.
3. The unintended IEQ underperformance which might occur when the focus of the design is primarily to meet increasingly stringent energy efficiency objectives in new buildings.

8.2 ENERGY AND IEQ PERFORMANCE OF CASE STUDIES

Objective 1: Analyse the operational performance of buildings for their energy use and the performance of energy-related IEQ parameters.

The performance analysis methodology that was developed in this research was applied to five case study buildings across four building sectors – offices, schools, hospitals and apartment blocks. These buildings represent a wide cross-section of the UK building stock and therefore can provide useful insights into the endemic issues in the construction sector that drive the performance gap.

The actual performance of all the buildings reported for one calendar year shows deviations from the design projections. The regulatory compliance calculation results underestimated the actual energy use significantly. This was seen in the school, the hospital, and the apartment block, where the actual energy use was more than 2-4 times what was predicted using the compliance calculations. In the public office, building-specific design calculations were done, and while there were some under-performance issues, it was a very high-performing building among the naturally ventilated open-plan office buildings in the UK. Part of this is due to the performance contracting procurement process followed in the project. This process guaranteed the involvement of the contractor and designers after the handover to fine-tune the building and ensure low-carbon objectives were met. Therefore, the case study public office provides good evidence that if performance contracting was also integrated within a Soft Landings Framework, which promotes extended involvement of the design and contracting team, then it can help to ensure that the design-stage performance objectives are met.

Unlike energy, IEQ performance was not a metered, monitored or reported performance parameter for any of the case study buildings. The temperature, CO_2, NO_2, and $PM_{2.5}$ measurements, done for holistic performance assessments in the case study buildings, provided insights into the IEQ in the buildings. In the public office and the school, there was no provision of comfort cooling (except in areas with servers and high ICT equipment), and instances of summertime overheating were observed.

IAQ in the buildings was dependent on the type of ventilation system and also had seasonal variations. The public office and university office were naturally ventilated, whereas the school, hospital and apartment block provided fresh air using a mechanical ventilation system.

Ventilation controls based on zone CO_2 concentrations (except in the university office building) kept indoor CO_2 levels within acceptable levels, with higher average concentrations observed in winters than summers There was some underperformance seen in the public office in winters due to a combination of closed vents and malfunctioning sensors, which made CO_2 concentrations in some zones exceed the IDA Class II levels.

Externally driven pollutants such as PM levels and NO_2 are difficult to control in naturally ventilated buildings. However, $PM_{2.5}$ in the apartment block, with both mechanical ventilation as well as operable windows, was observed to exceed recommended levels, occasionally driven by both internal and external sources. Other buildings in urban settings had high levels of external PM concentrations, but the filtration in the mechanical ventilation system was effective in keeping the levels low in those buildings. This was, however, not the case for NO_2, which is primarily driven by external traffic-related pollution. In the hospital and to some extent in the apartment block, which are located in dense urban areas, indoor NO_2 levels closely followed external levels and sometimes showed risks of exceeding the WHO (WHO, 2005) thresholds.

8.3 CAUSES OF THE ENERGY PERFORMANCE GAP

Objective 2: Develop a systematic and replicable framework for M&V, using calibrated building performance simulation models to identify performance issues.

In this study, identification of the cause of performance issues in the case study buildings was done under a systematic M&V framework, in three stages.

- First, after analysing the metered and monitored actual data and comparing it against design intent, along with on-site observation, initial areas of deviations were identified.
- Then fine-tuning these deviations in a model calibration process, the deviations were verified as the key causes of the performance issues, Additionally, some new issues were also identified.
- Finally, based on engineering judgement, those causes were classified into operational changes and technical issues and, based on further discussions with stakeholders, they were also attributed to various construction stages.

This distinction of identifying and separating the operational changes from the underperformance caused due to manageable and avoidable technical issues is necessary for effective identification of the root causes of building's technical underperformance. The process followed to meet this objective was iteratively created with the learnings from the objective 3 findings, described in the next section, Section 8.4.

8.3.1 Summary of operational changes in the case study buildings

Buildings and their functions evolve over time; operational changes can be due to occupant behaviour issues or can be needed to adapt the buildings and to enable them to perform their functions. While most of the case study buildings analysed were newly built (the university office was refurbished), the need for operational modifications, intended at the design stage, had already arisen.

In the **public office**, operational changes included the addition of more regular occupants and a diversified occupancy pattern. Additionally, occupant needs determined higher heating set-point temperatures, and the departmental nature of the occupant organisation rendered the use of hot-desking, and subsequent energy efficient operation protocols during out-of-hours use, irrelevant.

In the **school**, besides elevated set-point temperatures, the key operational deviation was the significant transient occupancy during out-of-school hours for extracurricular activities, evening classes and summer schools.

Hospitals are by nature dynamic places that continuously evolve to meet functional needs; different clinical processes have unique needs. In the case study hospital, it was observed that typical operational trends of various spaces, their equipment loads and set-point temperatures changed over time during the year to meet functional needs. These changes included the number of beds in the wards. The high changeover rate of patients added to these changes.

Individual flats in **apartment blocks** have occupant-controlled functions, and their individual needs and preferences affect energy use significantly. The effect of operational modifications in apartments was seen in system controls and when there was a change in occupancy due to guests or occupants going on vacation.

In the **university office building**, the occupancy profile the occupant varied seasonally and was higher than a typical office during out-of-hours use and on weekends. Additionally, because of occupant controls for the space conditioning system, the set point maintained in the spaces was also high, leading to high energy use.

These operational changes could have been minimised with more detailed interactions with the occupants and the designers during the design stage. However, this is not always practically possible. Some lessons that could be drawn from this to feed into larger industry applications, such as the need for incorporating flexibility and adaptability into building and systems design.

8.3.2 Summary of technical issues in the case study buildings

While operational changes can be occupant behaviour related issues or be necessary deviations, technical issues are the major unintended shortcomings that cause underperformance. These generally relate to building design, construction, building systems, and their operations and maintenance.

In the **public office**, a major technical issue was related to the heating system: the primary heating system was malfunctioning. Caused by design stage errors and subsequent modifications to operations, this issue resulted in the inefficient overall performance of the HVAC system. Heat loss due to leakage through poorly sealed vents and at construction junctions was also observed. Additionally, operation and maintenance issues such as faulty sensors and high parasitic loads were also seen.

The centralised system design of the **school** led to inefficient operation during transient occupancy times, both the heating and the ventilation systems were working, leading to extremely high energy use. Additionally, AHU fans were also running at higher than optimal speeds during unoccupied times. On the primary energy front, a biomass boiler, installed for decarbonising energy use, was never used, due to practical and logistical issues of using biomass as fuel. On the construction front, significant thermal bypasses around insulation were seen, and the tested U-values were much higher than the design.

In the **hospital building**, poor demand-controlled ventilation led to excessive auxiliary energy use to meet fresh air needs. However, like the school, a major issue was the primary energy use. A new CHP plant was to be installed for the entire facility rather than just for the new building, to replace the low efficiency, old, steam-based central heating network. However, this has not happened yet.

In the **apartment block**, technical issues were related to maintenance of mechanical ventilation filters, cold bridging in balconies and air leakage under the doors. Besides this, on the primary energy use front, efficiency of the district heating network, estimated at 91% in design calculations, was around 50% in practice.

In the **university office building**, there were significant parasitic loads, around 50%. Another significant issue was the operation of the heating and cooling systems, which were running and conditioning the spaces continuously.

Issues found via the case studies were due to shortcomings at various construction stages (design, construction, commissioning, handover and operation), with some issues compounded by lasting over multiple stages. Operational changes were primarily linked to the design-stage (incorrect assumptions or protocols) or the operation-stage (changes to functional needs). Technical issues, on the other hand, were spread across all the stages, including design-stage calculation errors, construction defects, poor commissioning of systems and controls, lack of proper protocols set at handover, and poor maintenance and operations.

8.4 Reflections on modelling and calibration

Objective 3: Explore model calibration and improved validation approaches for implementation within M&V, when used in the context of performance issues identification.

The methodology used in the performance analysis of these case study buildings, being systematic, procedural and replicable, has building performance modelling and model calibration at the centre of it. The typical process used in calibration and validation of energy models is based on guidance given in M&V protocols, such ASHRAE Guideline 14 and IPMVP. These M&V guidance documents were primarily aimed at providing a procedure for better evaluation of savings due to ECMs. In this application of identifying causes of performance gap, the guidance provided in these M&V protocols regarding calibration can be insufficient. This study analyses the applicability of existing modelling and calibration protocols and suggests ways to improve them. This is achieved in two ways:

1. Undertaking traditional calibration for monthly energy use along with the statistical error indices, validating them with cross-checking of other dependent parameters.
2. For one of the case study buildings, exploring a multi-level framework for checking model calibration that included various levels of detailed calibration and advanced cross-validation.

In the multi-level calibration checking framework, advanced calibration and validation included the use of disaggregated data, integration of temperature cross-validation and using uncertainty-based hybrid validation methods.

Undertaking calibration at multiple temporal resolutions suggests that data granularity can affect the accuracy of the model and the more detailed (hourly) calibration requires significantly more effort and data (when compared to monthly calibration). However, depending on the application most optimal calibration requirement can be selected. For example for the application of identifying performance issues in buildings, the monthly calibrated model created in Chapter 5 can be considered as reasonably accurate. However, if the model is calibrated at a lower resolution (monthly), it may mask some of the underlying issues relevant for more detailed system assessments of occupant behaviour analysis, which can only be uncovered at finer timescales (Chapter 6). The use of disaggregated end-use data minimises cross-compensation of performance issues. Also, analysing disaggregated end-use data provides building use information. Lighting and small power loads can help to identify building use patterns and operational behaviours. Space conditioning loads can help to determine system use patterns. This detailed calibration was done for the university office building, where floor-wise sub-metred data was also available. Some of the diversity arising from space type and the nature of occupants was clarified when a more granular analysis was performed for each floor. These findings at the granular level might be unique to that year, and more data is required to ensure that the observed deviations are a regular occurrence rather than a passing event. Also, detailed granular information and on-site observations during audits help in verifying many of the design aspects, but it can sometimes be challenging due to factors such as access restrictions in sensitive hospital areas. Discussions with building users and facility managers not only help in understanding building usage but also can identify unique events, point towards problem areas and validate any inferences drawn from the data trends. Inputs from stakeholders are required in the procedural evidence-based calibration method, which needs to be followed when a calibrated model is used for performance gap identification. In other words, in evidence-based calibration, the fine-tuning of the inputs in the base model should be based on observed trends only.

Level 2D-ADV calibration, the highest level of calibration, requires cross-validation of dependent parameters (space temperatures) in typical zones and is used to check the robustness of the model. This examination ensures that the calibration model matches not just the actual energy performance but other physical parameters as well. However, there were some limitations. As the heating and cooling were provided by a VRF system, it was not possible to isolate the space heating from cooling energy. This is important in London's climate, where summers are mild and cooling load can be driven by internal gains, especially in an office building. This can often lead to instances of simultaneous heating and cooling provided in different zones. Also, building envelope construction was not documented well during the refurbishment; therefore, the issues and uncertainties regarding material properties could not be assessed in this case.

In case the calibration criteria are not met after all the evidence-based changes, further fine-tuning of inputs must be based on engineering judgement but within reasonable ranges, as numerous solutions could generate a calibrated model. Uncertainty-based calibration can provide a way to deal with issues of data availability and granularity. It provides an alternate validation method of a partially calibrated model. Also, as shown in this case study, it can be used to reinforce confidence in a calibrated model, ensuring that even with input uncertainty, the measured data is within a reasonable output range. In the case study, using the multi-level calibration framework, the residual uncertainty seen in the overall results was not high.

The more granular the data, the better it is; however, calibration at finer timescales requires a large amount of well sub-metered data and significant effort from the modeller to get useful results. Also, the probabilistic validation approach requires a lot of data and computational time to model all scenarios compared to statistical validation. Therefore, depending on the time available, a suitable balance is needed between temporal resolution, data availability, accuracy, effort and the intended use of the calibrated model.

To summarise, integrating calibrated simulation as an investigative tool and validation measure for post-occupancy evaluation processes in an M&V protocol in this study provided three clear benefits:

1. **Identification of major issues**: Conventional post-occupancy evaluations for identifying causes of underperformance which do not undertake calibrated simulation could miss some of the causes. Using a calibration-assisted procedure incorporating all input uncertainty can validate whether, through the initial assessment, all the major issues have been identified or if more detailed site investigation is necessary.

2. **Operational-baseline creation**: A new operational-baseline is necessary to separate the technical causes of underperformance from changes that are necessary for the building to perform its function in practice. This process is only possible when a validated calibration model is created, as it is possible in the calibrated model to revert all the technical issues identified back to their design intent.

3. **Enhanced assessment and analysis**: Due to the use of simulation models, uncertainty and sensitivity analysis can also help in speeding up the identification of causes of the performance gap by providing a hierarchy of the most influential parameters. This is especially useful in cases where initial assessments could not account for all the deviation. Using simulation, quantification of the impact of individual issues and correction measures could be done, exploring the various ‘what if’ scenarios.

8.5 CROSS-SECTORAL ISSUES IN ENERGY PERFORMANCE

Objective 4: Examine the findings from different types of buildings and determine the cross-sectoral performance issues and their underlying root causes.

Analysing cases from multiple buildings across various sectors (office, school, hospital and apartment block) provides an opportunity for a robust assessment of performance issues and corroboration of their causes, with many common themes and lessons applicable across the UK building sector. While drawn from individual cases, some of the issues identified in these cases concur with findings from case study investigation in other sector-specific studies (Burman, 2016; Palmer, et al., 2016b; Palmer, et al., 2016a). Therefore, in this study, which coves various building sectors, the issues identified are considered to be endemic in the industry and are likely to have applicability to other buildings and across the construction sector.

The key cross-sectoral lessons drawn from these cases studies are

- **Design stage energy projections:** The use of regulatory compliance calculations as the design stage predicted performance may significantly underestimate actual energy use and lead to a perceived gap. To avoid this, a realistic estimation of energy use at the design stage should be calculated as per CIBSE TM54 guidelines or other equivalent protocols.
- **Performance targets and contractual accountability:** Performance contracting, integrated within a Soft Landings Framework, guarantees the involvement of the contractor and the designers after the handover to fine-tune building operations and fix issues related to design, construction and commissioning. If these are balanced against the challenges of a finite contractual period and the targets and the metric to use are objectively defined, then it can be a very effective approach.
- **Uncertainties in building use:** Occupant behaviour is uncertain and has a significant effect in cases where occupants have a high degree of control, such as in apartments. In most large non-domestic buildings, services are controlled centrally, limiting the impact of the occupant. To address this issue, sensitivity and scenario analysis can be used to implement informed design decisions and operational safeguards regarding the most important factors.
- **Transitionally/seasonally occupied buildings and changing user patterns:** Contrary to design assumptions, nowadays schools have higher space-time utilisation. The ability to hydraulically isolate unoccupied areas and use decentralised controls is beneficial in these scenarios. Flexible space use, including hot-desking in offices, is another practice used to manage transitional occupancy. This strategy works effectively when the building occupants and their work habits are also structured in that way. In the public office hot-desking was planned but

was not effective, as regular users occupied their usual desks within the part of the building where their department was located.

- **Use of new and innovative low-carbon technologies:** The technologies used across the case studies showed that while there was an intent, lack of follow-up and checks meant that the intended benefits were not achieved. In the office, the innovative strategy of using free heat from server rooms was not successful. In the school, the biomass boiler installed was not operational, due to the enhanced logistics needed to manage the fuel. In the hospital, the new CHP system was not installed due to budgetary constraints.
- **Protocols for managing building operations:** Some of the performance issues in the case studies were due to sub-optimal operations and irregular maintenance. These included delays in resolving issues with sensors in the public office or replacement of ventilation filters in the apartment block. An effective building operation and management strategy envisaged in design and incorporated at handover would help manage these issues.

8.6 UNINTENDED IEQ UNDERPERFORMANCE

Objective 5: Assess the unintended IEQ underperformance that might occur when the focus of the design is primarily to meet increasingly stringent objectives of energy efficiency.

It is a challenge for designers to balance energy efficiency and IEQ performance, due to potential conflicts between these performance objectives. In all case study buildings, energy performance was the only specific quantified performance objective. There was no effective metering, monitoring and reporting strategy for any of the IEQ parameters. However, lower energy use and better IEQ can be divergent objectives. Some of the unintended IEQ performance issues across the case studies included

- **Summertime overheating:** Overheating was found in some of the case study buildings that were highly insulated and airtight, a common feature in new buildings constructed to high energy standards. The temperature measurements in the school exceeded the maximum temperature threshold of 26°C (CIBSE, 2013b), because of the airtight envelope and inadequate operable windows in a south-facing zone (high solar exposure), as well as a lack of solar controls (blinds/shades). Instances of similar summer overheating were also seen in the public office.
- **Conflicts with fresh air supply**: Conditioning of fresh air has significant energy implications and balancing of fresh air requirements against energy use is needed to make a building perform optimally. Fresh air supply in the case study buildings was controlled by CO_2-based sensor controls. The indoor CO_2 concentrations in the public office building reached above 2500 ppm due to sensor malfunctioning as well as the closing of roof vents in winter to avoid cold drafts due to the stack effect and to maintain users' thermal comfort. This exceeds the IDA Class II levels for office spaces (CIBSE, 2015).
- **Ventilation strategies in urban areas, mitigating outdoor air pollution**: A natural ventilation strategy may not be suitable for dense urban environments where external air can be more polluted than indoor air. MV systems provide the necessary controls and create more airtight envelopes. The industry's main metric for assessment of IAQ is currently CO_2 concentrations. Most existing control strategies for ventilation systems also use this metric. In mechanically ventilated buildings, filtration is used to provide a level of protection against outdoor sources of pollution, such as micro-particles. However, some traffic-related pollutants, such as NO_2, are not mechanically filtered. The hospital building was located in a congested urban area but there was a lack of measures against ambient NO_2. Indoor NO_2 levels recorded in a ward very closely followed external levels. These recordings were often above the WHO annual mean threshold of 40 μg/m³ (WHO, 2018). This suggests a potential risk of exposure to more than

WHO-recommended levels if the external air remains polluted for prolonged periods. Advanced chemical filtration (such as activated carbon) and controls that consider the balance between fresh air requirements and protection from outdoor sources of pollution could provide a healthier environment and at the same time save energy in both mechanically and naturally ventilated buildings that rely on automated ventilation.

- **Better filtration for pollutants**: Similar air quality issues were seen in the mechanically ventilated apartment block: $PM_{2.5}$ levels in the flat are close to external levels, with some spikes due to internal sources. However, they were frequently higher than the WHO annual mean limit of 10 µg/m^3 (WHO, 2018). The flats used G3 filters; however, as ambient $PM^{2.5}$ levels are typical for London, the use of high-grade filtration (F type filters), while slightly increasing electricity use, will help reduce $PM^{2.5}$ levels.

These points show potential conflicts within IEQ objectives and between energy and IEQ, especially in new low-energy buildings constructed in polluted urban areas. This is primarily because during operations, unlike energy, IEQ performance is not an explicit target and therefore not a focus for building monitoring and fine-tuning. Issues observed across the case studies have shown the need to address IEQ simultaneously with energy through better design, advanced operational controls and, most importantly, incorporating regular IEQ measurements. In some instances, where maintaining better IEQ might result in slightly increased energy use, the energy use increase needs to be offset elsewhere. The study also shows the importance of contractual accountability to minimise performance issues, building a case for having IEQ in energy performance contracts to mitigate the trade-offs of IEQ against energy performance that lead to unintended health consequences for occupants.

9 CONCLUSION

This detailed analysis of the performance gap for multiple buildings provides insights into the performance issues applicable across the UK building sector. Identifying and verifying performance issues using evidence-based model calibration is a novel approach to increase the likelihood that all key issues are found. A calibrated model, when validated statistically for energy use and also for dependent parameters such as space temperatures, ensures that it could be reasonably used for assessing deviations from the design intents. Using this simulation-based approach enables a systematic identification and classification of the root causes of the performance gap, which would not have been possible otherwise.

This study identified several lessons that can potentially be used to inform and improve current building design practices. The findings regarding performance issues might be somewhat specific to the case studies, especially the technical issues regarding building systems. However, the significance of optimal operations and maintenance of building systems for better energy and IEQ performance has applicability for the building sector in general.

9.1 Key Lessons

9.1.1 Causes of the performance gap

Root causes identified across the case-studies were due to three factors:

1. A perceived gap due to the use of inappropriate design-stage calculation methods;
2. Technical issues with the building, its systems, and their operations; and
3. Operational changes that the building has gone through in order to meet its functional requirements.

The first can be dealt with by improving industry practices for design-stage calculation that account for realistically expected operating conditions. The third can brought by functional necessity (which might be inevitable) or due to occupant behaviour reasons (which could be addressed and resolved with better procedures and education). The second gap (due to technical issues) requires technical attention and correction to ensure that the building functions optimally. The technical issues identified across the case studies were either due to design errors, improper construction and installation, poor commissioning or issues with building systems and low-carbon technologies. It was observed that long-term involvement and responsibility for performance on the part of the design and construction teams is effective in lowering the performance gap. Including this approach through performance contracting ensures that, after handover, technical issues are addressed, and the building is well managed. Additionally, policy measures and safeguards are essential to ensure energy efficiency measures and low-carbon systems, specified at the design stage, will be effectively used in practice.

9.1.2 Using calibrated simulation in performance assessment

Integrating calibrated simulation as an investigative tool and validation measure for post-occupancy evaluation processes in an M&V protocol enables confidence in identification of all major issues and creation of an up-to-date operational-baseline for enhanced assessment and analysis.

Limitations of validation practices can be improved by detailing and enhancing the scope of current standards and guidelines. Current M&V practices need to account for enhanced sub-metering and incorporation of IEQ cross validation checks for model calibration. This can be structured in the form of pre-requisites for various levels of calibration for each temporal resolution. The new framework proposed in this thesis keeps the essence of existing protocols as the base level of calibration. Depending on the intended use of the model, higher calibration levels can be achieved by adding more checks tailored to disaggregated end-uses whilst meeting the minimum validation requirement for IEQ parameters (e.g. number of data streams, duration and frequency of monitoring and number of spaces to be covered). A probabilistic approach can be used as an alternative validation process to overcome some of the issues with data availability and granularity and improve the calibrated model's usability.

9.1.3 Need for safeguarding of low-carbon technologies

Issues with heat pumps in the public office, the non-operational biomass boiler in the school and use of a non-optimal steam-based heating network instead of the planned CHP plant in the hospital meant that the technologies planned to offset carbon emissions were not fulfilling the role envisaged during the design stage due to technical, logistical and budgetary challenges.

To minimise the long-term impact of inefficient systems, steps are required for not only enabling the smooth integration of new technologies with conventional practices but also safeguarding them against value engineering. At the policy level, robust regulatory safeguards, such as measurement and verification of building and system performance in the first few years, are needed to ensure that proposed low or zero-carbon strategies and technologies will be used in practice, within acceptable timelines.

9.1.4 Enhanced stakeholder engagements

Most of the technical issues causing energy performance gaps were due to sub-optimal operation and maintenance issues related to building systems. A more streamlined building operation and management strategy envisaged in design and incorporated at handover would enable a building to operate reasonably close to what is assumed at the design stage.

However, it is common that, after handover, the engagement of the design and construction team with the building is minimal. This means that most of the focus is on delivery and system functionality rather than performance. Shortcomings in commissioning and qualitative design issues subsequently identified remain largely unaddressed. Enhanced engagement with all stakeholders, along with the accountability of people delivering the building, is necessary to ensure that the building performs as expected, both on energy and IEQ fronts. This can be supplemented by a Soft Landings Framework (Way, et al., 2009) or performance contracting approach, in which the designers, contractors and building managers are accountable and are stakeholders in ensuring the operational performance of the building.

This was seen in the public office building, which was the best-performing building and was subject to energy performance contracting. It is suggested that this type of contract can help narrow the performance gap and would be financially practical in projects where the project client has a long-term vested interest in the building.

9.1.5 Addressing energy and IEQ performance holistically

The focus on energy efficiency alone as the performance objective in buildings built with low-energy-use intentions may lead to unintended IEQ concerns (such as poor IAQ and overheating). Issues observed across the case studies have shown the need to address IEQ simultaneously with energy through better design, advanced operational controls and, most importantly, by incorporating regular IEQ measurements. Including IEQ within the performance contracting framework for energy will help address the trade-offs that happen during the operational stage and the unintended health and wellbeing consequences to occupants.

9.2 MAIN CONCLUSION

To ensure buildings perform holistically upon completion, they should be procured under a performance contract that accounts for specific requirements for both energy and IEQ: *EEPC (Environment and Energy Performance Contract)*. EEPC provides a framework for enabling high performance through accountability – various lessons from these case studies, listed below, provide guidance on what operational implementation can be done to check and deliver on performance.

- At the design stage, measurable and verifiable operation performance targets for energy and IEQ performance should be set. These should be supplemented with ongoing risk assessment, from design to completion, of various energy efficiency measurements to safeguard them from value engineering.
- During handover, clear operations and maintenance protocols for efficient and effective running of the building and its systems should be developed.
- In the initial few years of operation, until the building reaches steady modes of operation, enhanced engagement of the design and construction teams should be ensured to analyse, optimise and fine-tune building operations so that it is running at its optimum potential. During this process, and afterward during ongoing operations, a systematic calibration-based M&V protocol should be used to analyse and correct performance issues.
- During the assessment of performance issues and for performance targets, the baseline[*] used should be clearly defined. Ideally, the operational-baseline should be used, i.e. the baseline representing the ideal performance of the building after incorporating any changes done in the building to meet its functional requirements.

The EEPC approach will reduce unintended IEQ underperformance, and ongoing systematic assessment of the performance issues based on their causes will allow a clearer identification of a

[*] reference point for calculating the performance gap

building's potential to perform and measures that are needed to improve the issues. The overall framework proposed in this study, aimed at early stages of post-occupancy, can be easily integrated within existing project procurement routes that focus on operational performance outcomes, such as the Soft Landings Framework. For example, the public office case study building was procured within a Soft Landings Framework.

Most of the processes during design and construction in the Soft Landings Framework run alongside normal procurement practices, without much duplication or extra work. The only elements that require additional effort are aftercare activities needed during the initial years of operation, where the approach presented in this paper could be integrated. Depending on the complexity of the project, the cost of these activities can be estimated at around 0.1% -0.25% of the total contract value (BSIRA, 2012; Morris, 2017). The extra effort and resources required for integrating the Soft Landings Framework or other similar frameworks is minimal and is usually paid back swiftly because of the reduced building operating costs.

9.3 FUTURE WORK

The findings of this research can be enhanced and supplemented by further investigations in the following research areas:

- Holistic analysis of the energy and IEQ interdependence. The current study focuses on energy-related impacts on thermal comfort and some parameters of IAQ. While these two parameters cover most of the important energy-related interdependence, further can include other IAQ parameters as well as lighting and acoustics. This will provide a complete IEQ perspective, which can go beyond energy-related issues, driven by control protocols for building operations that can balance the objectives of achieving high energy efficiency with better IEQ.
- Development of enhanced and standardised modelling and calibration protocols. The methods proposed here require a certain level of modelling expertise and engineering judgement to generate the calibrated model. Additionally, model validation, through multi-parameter checking, requires more standardisation to ensure that the dynamic thermal performance can be accurately used for performance prediction over a wider range of energy and environmental outputs.
- Research that focuses on improvements to model validation criteria. Taking the idea of enhanced cross-validation forward, follow-up research can look at the importance of different

data types for cross-validation could give a better understanding of which parameters to focus on during model calibration.

- Probabilistic uncertainty and sensitivity analysis and exploration of automated techniques for calibration. To facilitate uptake of these projects in the industry, innovative data capture methods from existing buildings and identification of building issues in a semi-automated way will be essential to achieve reasonable accuracy in computer model calibration at a low cost.
- Linking calibrated models to operational buildings for real-time forecasting of performance, and analysis of control strategies.
- Current building design and operation strategy catering to today's climatic conditions show overheating risks in hot summer spells in certain zones. In the context of future climate, this risk can be significantly higher. The future performance can be tested using future climate data in building performance simulations and explore design reliance to mitigate future climate risks.
- Incorporation of detailed HVAC modelling in performance estimations and evaluations, both in design calculations and during operational stage assessments, have the potential to generate more accurate models. However, this additional detail in the modelling needs to be balanced against the uncertainty of those variables.
- Building systems are being operated to integrate high air-change requirements to minimise airborne transmission of infections. This change in approach requires analysis of its impact on energy performance and the performance gap in existing buildings. Additionally, for new buildings, there could be a need to redefine carbon budges and targeted energy benchmarks.

10 References

Abadie, M. & Wargocki, P., 2016. *Indoor Air Quality Design and Control in Low-energy Residential Buildings, Subtask 1, Defining the metrics,* s.l.: IEA EBC Annex 68.

Addy, M. N., Adinyira, E. & Koranteng, C., 2014. *A review of research methodologies in building energy efficiency assessment.* Johannesburg, South Africa, 8th Construction Industry Development Board (CIDB) Postgraduate Conference.

Ahmad, M. & Culp, C. H., 2006. Uncalibrated Building Energy Simulation Modeling Results. *HVAC&R Research,* 12(4), pp. 1141-1155.

Al Horr, Y. et al., 2016. Impact of indoor environmental quality on occupant well-being and comfort: A review of the literature. *International Journal of Sustainable Built Environment,* Volume 5, pp. 1-11.

Andre, P. et al., 2008. From model validation to production of reference simulations: how to increase reliability and applicability of building and HVAC simulation models. *Build. Serv. Eng. Res. Technol.,* 29(1), pp. 61-72.

AQEG, 2004. *Nitrogen Dioxide in the United Kingdom: Summary,* London: Defra.

AQEG, 2005. *Particulate Matter in the UK: Summary,* London: Defra.

ASHRAE, 2011. *Procedures For Commercial Building Energy Audits.* 2nd ed. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

ASHRAE, 2014. *Measurement of energy, demand and water savings - ASHARE Guidelines 14-2014,* Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE).

ASHRAE, 2017. *ANSI/ASHRAE Standard 55, Thermal Environmental Conditions for Human Occupancy,* Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Azar, E. & Menassa, C., 2012. *Sensitivity of Energy Simulation Models to Occupancy Related Parameters in Commercial Buildings.* s.l., Construction Research Congress 2012.

Bannister, P., 2009. *Why good buildings go bad while some are just born that way.* s.l., Ecolibrium.

Batterman, S., 2017. Review and Extension of CO2-Based Methods to Determine Ventilation Rates with Application to School Classrooms. *International Journal of Environmental Research and Public Health,* 14(2), p. 145.

BEIS, 2016. *Building Energy Efficiency Survey 2014-15 (BEES) Overarching Report,* London: Department for Business, Energy & Industrial Strategy, HM Government.

Beizaee, A., 2016. *Measuring and Modelling the Energy Demand Reduction Potential of Using Zonal Space Heating Control in a UK Home. A Doctoral Thesis,* Loughborough, UK: Loughborough University.

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

Booten, C. & Tabares-Velasco, P. C., 2012. *Using EnergyPlus to Simulate the Dynamic Response of a Residential Building to Advanced Cooling Strategies.* Boulder, Colorado, 2nd International Conference on Building Energy and Environment.

Bordass, B., Cohen, R. & Field, J., 2004. *Energy Performance of Non-Domestic Buildings - Closing the Credibility Gap.* Frankfurt, IEECB’04 Building Performance Congress.

Bordass, B., Cohen, R., Standeven, M. & Leaman, A., 2001a. Assessing building performance in use 3: energy performance of the PROBE buildings. *Building Research & Information,* 29(2), pp. 114-128.

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

BRE, 2012. *NCM modelling guide Green Deal Development Edition.* [Online]
Available at: https://www.ncm-sbem.org.uk/
[Accessed March 2018].

BRE, 2016. *Overheating in dwellings,* Watford: Building Research Establishment (BRE).

BRECSU, 1996. *Energy consumption guide 72: Energy consumption in hospitals,* Watford, UK: Building Research Energy Conservation Support Unit (BRECSU).

Brown, M. G., 2006. *Keeping Score Using the Right Metrics to Drive World Class Performance.* 1 ed. New York: Productivity Press.

BSI, 2007. *BS EN 15251: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics,* London: British Standards Institution.

BSI, 2008. *BS EN 15603 Energy performance of buildings. Overall energy use and definition of energy ratings,* London: British Standards Institution.

BSI, 2012. *BS EN 12599:2012 Ventilation for buildings. Test procedures and measurement methods to hand over air conditioning and ventilation systems,* London: British Standards Institution.

BSI, 2017. *BS ISO 17772-1:2017 - Energy performance of buildings. Indoor environmental quality. Indoor environmental input parameters for the design and assessment of energy performance of buildings,* London: British Standards Institution.

BSI, 2019. *BS EN 16798-1:2019 Energy performance of buildings: Ventilation for buildings.,* London: British Standards Institution.

BSIRA, 2012. *Soft Landings procurement and budgets.* [Online]
Available at: https://www.bsria.com/uk/news/article/soft-landings-procurement-and-budgets/
[Accessed 13 04 2020].

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, Proceedings of the CIBSE Technical Symposium 2015, Chartered Institution of Building Services Engineers (CIBSE).

Burman, E., 2016. *Assessing the Operational Performance of Educational Buildings against Design Expectations – A Case Study Approach. Doctoral Thesis,* London: UCL (University College London).

Burman, E., Kimpian, J. & D., M., 2018. Building Schools for the Future: Lessons Learned From Performance Evaluations of Five Secondary Schools and Academies in England. *Frontiers in Built Environment,* Volume 4.

Burman, E., Kimpian, J. & Mumovic, D., 2014. *Analysis of the applicability of the UK National Calculation Methodology to energy efficiency finance of non-domestic buildings: A case study approach.* London, Proceedings of Building Simulation and Optimization 2014, IBPSA England.

Campolongo, F., Saltelli, A. & Cariboni, J., 2011. From screening to quantitative sensitivity analysis. A unified approach. *Computer Physics Communications,* 182(4), pp. 978-988.

Carbon Trust, 2012. *Closing the gap: Lessons learned on realising the potential of low carbon building design,* London: The Carbon Trust.

CEC, 2018. *2019 Building Energy Efficiency Standards for Residential and Nonresidential Buildings,* Sacramento, CA: California Energy Commission.

Chai, T. & Draxler, R. R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development,* Volume 7, pp. 1247-1250.

Chakraborty, D. & Elzarka, H., 2018. Performance testing of energy models: are we using the right statistical metrics?. *Journal of Building Performance Simulation,* 11(4), pp. 433-448.

Chatzidiakou, L., Mumovic, D. & Dockrell, J., 2014. *The Effects of Thermal Conditions and Indoor Air Quality on Health, Comfort and Cognitive Performance of Students,* London: The Bartlett, UCL Faculty of the Built Environment.

Chen, C., Twycross, J. & Garibaldi, J. M., 2017. A new accuracy measure based on bounded relative error for time series forecasting. *PLoS ONE,* 12(3).

Chong, A. & Lam, K. P., 2015. *Uncertainty analysis and parameter estimation of HVAC systems in building energy models.* Hyderabad, 14th Conference of International Building Performance Simulation Association.

Chong, A., Xu, W. & Lam, K. P., 2015. *Uncertainty Analysis in Building Energy Simulation. A Practical Approach.* Hyderabad, India, Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association.

CIBSE, 2008. *CIBSE TM46: Energy benchmarks,* London: The Chartered Institution of Building Services Engineers (CIBSE).

CIBSE, 2012. *CIBSE Guide F: Energy efficiency in buildings,* London: The Chartered Institution of Building Services Engineers.

CIBSE, 2013a. *TM54: Evaluating Operational Energy Performance of Buildings at the Design Stage,* London: The Chartered Institution of Building Services Engineers (CIBSE).

CIBSE, 2013b. *TM52: The Limits of Thermal Comfort: Avoiding Overheating in European Buildings,* London: The Chartered Institution of Building Services Engineers (CIBSE).

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

CIBSE, 2017. *TM59 Design methodology for the assessment of overheating risk in homes,* London, UK: The Chartered Institution of Building Services Engineers (CIBSE).

CIBSE, 2020. *The Energy Benchmarking Tool (Beta Version).* [Online]
Available at: https://www.cibse.org/knowledge/digital-tools/the-energy-benchmarking-tool-(beta-version)
[Accessed 02 12 2020].

CIBSE, 2020. *TM40: Health and wellbeing in building services,* London: The Chartered Institution of Building Services Engineers (CIBSE).

Clarke, J., Strachan, P. & Pernot, C., 1993. Approach to the calibration of building energy simulation models. *ASHRAE Transactions,* 99(2), pp. 917-927.

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,* 37(9), pp. 123-141.

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, First Building Simulation and Optimization Conference (BSO12).

Cohen, R., Ratcliffe, S. & Bannister, P., 2018. *Design for Performance Pilot Programme: Technical Report,* London: BBP (Better Buildings Partnership).

Council Directive, 2018. Directive (EU) 2018/844 of the European Parliament and of the Council of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and Directive 2012/27/EU on energy efficiency (Text with EEA relevance). *Official Journal of the European Union,* 156(19.6.2018), pp. 75-91.

DCLG, 2010. *Approved Document F: ventilation ,* London: Department for Communities and Local Government, UK.

DCLG, 2013. *Approved Document L2A: conservation of fuel and power in new buildings other than dwellings,* London: Department for Communities and Local Government, UK.

DCLG, 2015. *Improving the energy efficiency of our buildings: A guide to display energy certificates and advisory reports for public buildings,* London: UK Department for Communities and Local Government.

DCLG, 2017. *A guide to Energy Performance Certificates for the construction, sale and let of non-dwellings,* London: UK Department for Communities and Local Government.

de Dear, R. & Brager, G. S., 1998. Developing an adaptive model of thermal comfort and preference. *ASHRAE Transaction,* 104(1a), pp. 145-167.

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.

de Wilde, P., 2018. *Building Performance Analysis.* 1st ed. s.l.:Wiley-Blackwell.

de Wilde, P. & Jones, R., 2014. *The building energy performance gap: up close and personal.* Dublin, CIBSE ASHRAE Technical Symposium 2014.

de Wit, S., 2004. Uncertainty in building simulation. In: A. M. Malkawi & G. Augenbroe, eds. *Advanced Building Simulation.* New York: Spon Press, pp. 25-59.

DECC, 2018 . *Energy Consumption in the UK (2018) Chapter 5: Service sector energy consumption in the UK between 1970 and 2017,* London, UK: Department of Energy & Climate Change, UK.

Department for Education, 2003. *Acoustic design of schools: performance standards - Building Bulletin 93,* London: Department for Education.

DesignBuilder Software Ltd., 2019a. *DesignBuilder V6,* Stroud: DesignBuillder Software Ltd..

DesignBuilder Software Ltd., 2019b. *Climate Analytics.* [Online]
Available at: https://analytics.designbuilder.co.uk/
[Accessed 12 11 2019].

DfE, 2006. *Building Bulletin 101: Ventilation of School Buildings,* London, UK: Department for Education.

DOE, 2019. *EnergyPlus.* [Online]
Available at: https://energyplus.net/
[Accessed 02 February 2020].

Edwards, R. E., New, J. & Parker, L. E., 2012. Predicting future hourly residential electrical consumption: A machine learning case study. *Energy and Buildings,* Volume 49, p. 591–603.

Eisenhower, B. et al., 2012. A methodology for meta-model based optimization in building energy models. *Energy and Buildings,* Volume 47, pp. 292-301.

EVO, 2016. *Core Concepts: International Performance Measurement and Verification Protocol,* Washington DC: Efficiency Valuation Organization.

Fabbri, K. & Tronchin, L., 2015. Indoor Environmental Quality in Low Energy Buildings. *Energy Procedia,* Volume 78, pp. 2778-2783.

Fabrizio, E. & Monetti, V., 2015. Methodologies and Advancements in the Calibration of Building Energy Models. *Energies,* 8(4), pp. 2548-2574.

Fanger, P., 1970. *Thermal Comfort.* Copenhagen: Danish Technical Press.

Garrett, A. & New, J., 2016. Suitability of ASHRAE guideline 14 metrics for calibration. *ASHRAE Transactions,* 122(1), p. 469+.

George, A. L. & Bennett, A., 2005. *Case Studies and Theory Development in the Social Sciences.* Cambridge, MA: MIT Press.

Gupta, R. & Kapsali, M., 2015. Empirical assessment of indoor air quality and overheating in low-carbon social housing dwellings in England. *Advances in Building Energy Research,* 10(1), pp. 1-23.

Guyon, G., 1997. *Role of the model user in results obtained from simulation software program.* Prague, 6th International Building Performance Simulation Association Conference.

Haldi, F. & Robinson, D., 2008. On the behaviour and adaptation of office occupants. *Building and Environment,* 43(12), p. 2163–2177.

Heiselberg, P. et al., 2009. Application of sensitivity analysis in design of sustainable buildings. *Renewable Energy,* 34(9), pp. 2030-2036.

Heo, Y., Choudhary, R. & Augenbroe, G., 2012. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings,* Volume 47, pp. 550-560.

HM Government, 2019. *The Climate Change Act 2008 (2050 Target Amendment) Order 2019,* London: HM Government.

Hong, S.-M.et al., 2013. A comparative study of benchmarking approaches for non-domestic buildings: Part 1 – Top-down approach. *International Journal of Sustainable Built Environment,* 2(2), pp. 119-130.

Hopfe, C. & Hensen, J., 2011. Uncertainty analysis in building performance simulation for design support. *Energy and Buildings,* 43(10), pp. 2798-2805.

IEA & UNEP, 2019. *2019 Global Status Report for Buildings and Construction: Towards a zero-emission, efficient and resilient buildings and construction sector,* Geneva, Switzerland: International Energy Agency (IEA) and United Nations Environment Programme (UNEP).

Jain, N. et al., 2017. *Improving the Energy Performance Contracting Process using Building Performance Simulation: Lessons Learnt from a Post Occupancy Investigation of a Case Study in the UK.* San Francisco, USA, Proceedings of the 15th IBPSA (International Building Performance Simulation Association) Conference, Aug. 7-9, 2017.

Ji, Y. & Xu, P., 2015. A bottom-up and procedural calibration method for building energy simulation models based on hourly electricity submetering data. *Energy,* Volume 93, pp. 2337-2350.

Kapalo, P., 2013. *Analysis of ventilation rate and concentration of carbon dioxide in the office.* s.l.:Technical University of Kosice, Slovakia.

Kimpian, J. & Chisholm, S., 2011. *Tracking Design and Actual Energy Use: CarbonBuzz, an RIBA CIBSE platform.* Louvain-la-Neuve, International conference on Passive and Low Energy Architecture (PLEA).

Larsen, T. S., Daniels, O. & Jensen, R. L., 2012. *Evaluation of the indoor environment in 8 Danish passive houses.* Trondheim, Passivhus Nordern.

Logue, J. M., McKone, T. E., Sherman, M. H. & Singer, B. C., 2011. Hazard assessment of chemical air contaminants measured in residences. *Indoor Air,* 21(2), pp. 92-109.

Lunneberg, T. A., 1999. *Improving simulation accuracy through the use of short-term electrical end-use monitoring.* Kyoto, Proceedings of the 6th international IBPSA conference.

MacDonald, I., 2002. *Quantifying the effects of uncertainty in building simulation - PHD Thesis,* Scotland: University of Strathclyde.

Maivel, M., Kurnitski, J. & Kalamees, T., 2015. Field survey of overheating problems in Estonian apartment buildings. *Architectural Science Review,* 58(1).

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.

Martani, C. et al., 2012. ENERNET: Studying the dynamic relationship between building occupancy and energy consumption. *Energy and Buildings,* Volume 47, pp. 584-591.

Mavrogianni, A. & Mumovic, D., 2010. On the Use of Windcatchers in Schools: Climate Change, Occupancy Patterns and Adaptation Strategies. *Indoor and Built Environment,* 19(3), pp. 340-354.

Menezes, A. C., 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,* 97(9), pp. 355-364.

Monfet, D., Charneux, R., Zmeureanu, R. & Lemire, N., 2009. Calibration of a building energy model using measured data. *ASHRAE Transactions,* 115(1), pp. 348-359.

Moore, D. S., Notz, W. I. & & Flinger, M. A., 2015. *The basic practice of statistics.* 7th ed. New York: WH Freeman.

Morgensterna, P. et al., 2016. Benchmarking acute hospitals: Composite electricity targets based ondepartmental consumption intensities?. *Energy and Buildings,* Volume 118, pp. 277-290.

Morris, M. D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics,* 33(2), pp. 161-174.

Morris, N., 2017. *RIBA calls on government to commit to POE.* [Online]
Available at: https://www.architecture.com/knowledge-and-resources/knowledge-landing-page/riba-calls-on-government-to-commit-to-poe
[Accessed 01 05 2020].

Mumovic, D. et al., 2009. A methodology for postoccupancy evaluation of ventilation rates in schools. *Build. Serv. Eng. Res. Technol.,* 30(2), pp. 143-152.

Mustafaraj, G., Marini, D., Costa, A. & Keane, M., 2014. Model calibration for building energy efficiency simulation. *Applied Energy,* Volume 130, pp. 72-85.

Neto, A. H. & Fiorelli, F. A. S., 2008. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and Buildings,* 40(12), p. 2169–2176.

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,* 34(9), pp. 917-931.

Norford, L. K., Socolow, R. H., Hsieh, E. S. & Spadaro, G. V., 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,* 21(2), pp. 121-131.

O'Neill, Z. & Eisenhower, B., 2013. Leveraging the analysis of parametric uncertainty for building energy model calibration. *Building Simulation,* 6(4), pp. 365-377.

OEH, 2019. *Handbook for estimating NABERS ratings,* Sydney: Office of Environment and Heritage (OEH) .

ONS, 2011. *Nomis official labour market statistics.* [Online]
Available at: https://www.nomisweb.co.uk/
[Accessed 08 01 2020].

Palmer, J., Godoy-Shimizu, D., Tillson, A. & Mawditt, I., 2016b. *Building Performance Evaluation Programme: Findings from domestic projects; Making reality match design,* Swindon: Innovate UK.

Palmer, J., Terry, N. & Armitage, P., 2016a. *Building Performance Evaluation Programme, Findings from Non-Domestic Projects,* London: Innovate UK.

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. 100-108.

Parsons, P., 2014. Determining Infiltration Rates and Predicting Building Occupancy Using CO2 Concentration Curves. *Journal of Energy,* Volume 2014, p. 6 pages.

Penna, P. et al., 2015. Multi-stage calibration of the simulation model of a school building through short-term monitoring. *Journal of Information Technology in Construction,* Volume 20, pp. 132-145.

Petersen, S. & Hviid, C. A., 2012. *The European Energy Performance of Buildings Directive: Comparison of calculated and actual energy use in a Danish office building.* Loughborough, Building Simulation and Optimization Conference.

Phillips, T. J. & Levin, H., 2015. Indoor environmental quality research needs for low-energy homes. *Science and Technology for the Built Environment ,* 21(1), pp. 80-90.

Philp, D., Churcher, D. & Davidson, S., 2019. *Government Soft Landings,* Cambridge. UK: Centre for Digital Built Britain.

Piette, M. A. et al., 1994. *Over the energy edge: results from a seven year new commercial buildings research and demonstration project,* Berkeley: LBNL.

Poppendieck, D., Gong, M. & Emmerich, S., 2016. *Characterization of Emissions from Spray Polyurethane Foam. Final Report to U.S. Consumer Product Safety Commission NIST Technical Note 1921,* Gaithersburg, MD: US Department of Commerce, National Institute of Standards and Technology.

Raftery, P., Keane, M. & O'Donnell, J., 2011. Calibrating whole building energy models: An evidence-based methodology. *Energy and Buildings,* Volume 43, p. 2356–2364.

Reddy, T. A., 2006. Literature review on calibration of building energy simulation programs: uses, problems, procedure, uncertainty, and tools. *ASHRAE Transactions,* Volume 112 (Part1), pp. 226-240.

Reddy, T. A. & Maor, I., 2006. *Procedures for Reconciling Computer-Calculated Results With Measured Energy Data ASHRAE Research Project 1051- RP,* Philadelphia: Drexel University.

Reddy, T. A., Maor, I. & Panjapornpon, C., 2006. Calibrating Detailed Building Energy Simulation Programs with Measured Data—Part I: General Methodology. *HVAC&R Research,* pp. 221-241.

RIBA and Hay, R. et al., 2016. *Building Knowledge: Pathways to Post Occupancy Evaluation,* Reading: Value of Architects, University of Reading, RIBA.

RIBA, 2019. *RIBA Sustainable Outcomes Guide Foreword,* London, United Kingdom: Royal Institute of British Architects.

RMI, 2011. *Colloborate and Capitalize: Post-Report from the BEM Innovation Summit ,* Boulder: Rocky Mountain Institute.

Roberti, F., Oberegger, U. F. & Gasparella, A., 2015. Calibrating historic building energy models to hourly indoor air and surface temperatures: Methodology and case study. *Energy and Buildings,* 108(1), pp. 236-243.

Robertson, C. & Mumovic, D., 2014. Building performance in the context of industry pressures. *International Journal of Energy Sector Management,* 8(4), pp. 527-543.

Robertson, J., Polly, B. & Collis, J., 2013. *Evaluation of Automated Model Calibration Techniques for Residential Building Energy Simulation,* Golden, CO: National Renewable Energy Laboratory (NREL).

Royapoor, M. & Roskilly, T., 2015. Building model calibration using energy and environmental data. *Energy and Buildings,* 94(1), pp. 109-120.

Ruiz, G. R. & Bandera, C. F., 2017. Validation of Calibrated Energy Models: Common Errors. *Energies,* 10(10), p. 1587.

Ruiz, G. R., Bandera, C. F., Temes, T. G.-A. & Gutierrez, A. S.-O., 2016. Genetic algorithm for building envelope calibration. *Applied Energy,* Volume 168, pp. 691-705.

Ruiz, R. & Lemort, V., 2014. *Calibration of Building Simulation Models: Assessment of Current Acceptance Criteria.* Frankfurt, 8th International Conference Improving Energy Efficiency in Commercial Buildings (IEECB'14).

Saltelli, A. et al., 2007. *Global Sensitivity Analysis. The Primer.* Chichester, UK: John Wiley & Sons, Ltd.

Saltelli, A., Ratto, M., Tarantola, S. & Campolongo, F., 2004. *Sensitivity analysis practice: A guide to assessing scientific models.* Chichester, England: John Wiley & Sons Ltd..

Samuelson, H. W., Ghorayshi, A. & Reinhart, C. F., 2014. *Post-Occupancy Evaluation and Partial-Calibration of 18 Design-Phase Energy Models.* Atlanta, ASHRAE/IBPSA-USA Building Simulation Conference.

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.

Sheppy, M., Pless, S. & Kung, F., 2014. *Healthcare Energy End-Use Monitoring,* Golden, Colorado: National Renewable Energy Laboratory (NREL).

Shrubsole, C. et al., 2018. Bridging the gap: The need for a systems thinking approach in understanding and addressing energy and environmental performance in buildings. *Indoor and Built Environment,* 28(1), pp. 100-117.

Soebarto, V. I., 1997. *Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings..* Prague, Proceedings of the 4th international IBPSA conference.

Subbarao, K., 1988. *PSTAR: Primary and secondary terms analysis and renormalization: A unified approach to building energy simulations and short-term monitoring,* Golden, CO: Solar Energy Research Institute.

Sun, J. & Reddy, T. A., 2006. Calibration of Building Energy Simulation Programs Using the Analytic Optimization Approach (RP-1051). *HVAC&R Research,* 12(1), pp. 177-196.

Taheri, M., Tahmasebi, F. & Mahdavi, A., 2013. *A case study of optimization-aided thermal building performance simulation calibration Calibration.* Chambéry, 13th Conference of International Building Performance Simulation Association,.

Tanner, R., Henze, G. & Pless, S., 2013. Optimized Control of Automatic Windows for Energy Savings and Occupant Comfort. *ASHRAE Transactions,* Volume 119, pp. 274-287.

Teegavarapu, S. & Summers, J. D., 2008. *Case Study Method for Design Research: A Justification.* New York, Proceedings of IDETC/DTM 2008 ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.

Tian, W. & de Wilde, P., 2011. Uncertainty, sensitivity analysis of building performance using probabilistic climate projections: a UK case study. *Automation in Construction,* 20(8), pp. 1096-1109.

Tuohy, P. G. & Murphy, G. B., 2015. Are current design processes and policies delivering comfortable low carbon buildings?. *Architectural Science Review,* 58(1), pp. 39-46.

Tupper, K. et al., 2011. *Pre-Read for BEM Innovation Summit,* Boulder: Rocky Mountain Institute.

Turner, C. & Frankel, M., 2008. *Energy Performance of LEED for New Construction Buildings,* Vancouver: New Buildings Institute (NBI).

van Dronkelaar, C. et al., 2016. A Review of the Energy Performance Gap and Its Underlying Causes in Non-Domestic Buildings. *Frontiers in Mechanical Engineering,* Volume 1, p. 17.

Vesma, V., 2016. *The Meaning of R-squared.* [Online]
Available at: http://www.enmanreg.org/r2/
[Accessed 23 02 2020].

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.

Wang, S., Yan, C. & Xiao, F., 2012a. Quantitative energy performance assessment methods for existing buildings. *Energy and Buildings,* Volume 55, pp. 873-888.

Wassmer, C. J., 2013. *Energy Model Creation and Calibration Procedures for a Residential Home using Short Term Monitored Data,* Boulder: University of Colorado.

Way, M., Bordass, B., Leaman, A. & Bunn, R., 2009. *The Soft Landings Framework,* London: Building Services Research & Information Association and Usable Build Trust.

Webster, L. et al., 2015. *M&V Guidelines: Measurement and Verification for Performance-Based Contracts Version 4.0,* Washington, DC, USA: U.S. Department of Energy Federal Energy Management Program.

WHO, 1990. *Indoor environment : health aspects of air quality, thermal environment, light and noise,* Geneva: World Health Organization.

WHO, 2005. *WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide,* Geneva: World Health Organization.

WHO, 2018. *Ambient (outdoor) air pollution.* [Online]
Available at: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health
[Accessed 01 05 2020].

Wingfield, J. et al., 2013. *Centenary Quay Fabric and District Heating performance Study,* London: Innovate UK.

Won, D. et al., 2005. *A material emission database for 90 target VOC's.* Beijing, China, 10th International Conference on Indoor Air Quality and Climate.

Wyon, D. P. & Wargocki, P., 2013. How Indoor Environment Affects Performance. *ASHRAE Journal,* 55(3), pp. 46-52.

Yin, R. K., 2014. *Case Study Research: Design and Methods.* 5th ed. Thousand Oaks, CA: Sage.

Yoon, J., Lee, E. J. & Claridge, D. E., 2003. Calibration Procedure for Energy Performance Simulation of a Commercial Building. *Journal of Solar Energy Engineering,* 125(3), pp. 251-257.

Yu, C. & Crump, D., 2010. *'Investigation of an Odour and IAQ Complaint in Newly Built Sustainable Energy Efficient Housing.* Syracuse, NY, 9th Int. Healthy Buildings Conference.

ZCH, 2014. *Closing the gap between design and as-built performance,* London: Zero Carbon Hub.