Calibration of Detailed Building Energy Simulation Models to Measured Data using Uncertainty Analysis

by Daniel Coakley

Supervisor: Dr. Padraig Molloy and Dr. Marcus Keane

NUI Galway
OÉ Gaillimh

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy,
in the College of Engineering and Informatics

May 2014

Abstract

Over the past few decades, advances in technology, most notably the industrial revolution of the late 18th-century, has brought about dramatic improvements in the socioeconomic circumstances of developed nations. This has also brought with it rapid change in terms of human population, environmental impacts as well as energy consumption. Growth in energy consumption has been largely associated with increased use of finite fossil fuels (oil, coal, gas) in industrialized nations. However, this growth is unsustainable due to the depletion of these natural resources as well as the impact their consumption has on the environment, in terms of carbon dioxide (CO_2) emissions. A shift towards renewable fuels (wind, hydro, solar, geothermal, tidal) is currently underway, but progress remains slow, and the current reliance on fossil fuels for many existing essential technologies (e.g. transport) remains a major barrier to the large-scale transition that is required.

Energy efficiency has the potential to mitigate greenhouse gas emissions (GHG) and provide additional scope for the transition to a sustainable renewables-based energy future. Buildings account for approximately 40% of global energy consumption. Approximately half of this energy requirement stems from space heating and cooling. Studies have shown that savings of up to 40% are possible through the implementation of energy conservation measures (ECM's) and continuous commissioning (CC).

Whole building energy simulation tools have the potential to play a significant role in achieving this goal. However, their widespread adoption in the AEC (Architecture, Engineering and Construction) industry depends on their perceived reliability and the accuracy of their outputs. Currently, simulation tools are used primarily in building design with little integration or comparison to real building operation. It is often found that the actual buildings perform far worse than the design simulation initially predicted. This gap between measured and simulated data needs to be carefully addressed. This thesis proposes a new methodology for calibrating building energy simulation (BES) models to measured data including the incorporation of parameter uncertainty into final model predictions and recommendations.

Table of contents

List of Abbreviations

Abbreviation	Description
AEC	Architecture, Engineering and Construction
AHU	Air Handling Unit
ANN	Artificial Neural Network
ASHRAE	American Society of Heating, Refrigeration and Air-Conditioning Engineers
BAS	Building Automation System
BEPS	Buildings Energy Performance Simulation
BER	Building Energy Rating
BES	Building Energy Simulation
BIM	Building Information Model
BLC	Building Life Cycle
BMS	Building Management System
BREEAM	Building Research Establishment's Environmental Assessment Method
BWM	Box Whisker Mean
CAD	Computer Aided Design
CASBEE	Comprehensive Assessment System for Building Environmental Efficiency
CBE	Center for the Built Environment
CEC	California Energy Commission
CFD	Computational Fluid Dynamics
CHP	Combined Heat and Power
ChW	Chilled Water
CIBSE	Chartered Institute of Building Services Engineers
COP	Coefficient of Performance
CVRMSE	Cumulative Variation of Root Mean Squared Error
DCV	Demand Controlled Ventilation
DDCV	Dual-Duct Constant Volume
DDVAV	Dual-Duct Variable Air Volume
DECC	Department of Energy and Climate Change
DHW	Domestic Hot Water
DOE	Department of Energy
DSA	Differential Sensitivity Analysis
DSP	Dynamic (Thermal) Simulation Programs
ECM	Energy Conservation Measure
ECO	Energy Conservation Opportunity
EDA	End-use Disaggregation
EMCS	Energy Management and Control Systems
EMS	Energy Monitoring System
EPBD	Energy Performance of Buildings Directive
EPI	Energy Performance Indicator
ERI	Environmental Research Institute
ERV	Energy Recovery Ventilation
ESRU	Energy Systems Research Unit

EUI	Energy Use Intensity
FDD	Fault Detection and Diagnosis
FEMP	Federal Energy Management Protocol
HVAC	Heating, Ventilation and Air Conditioning
HW	Hot Water
IAI	International Alliance for Interoperability
IAQ	Indoor Air Quality
IDF	Input Data File (EnergyPlus input syntax for simulation)
IFC	Industry Foundation Class
IMPVP	International Performance Measurement & Verification Protocol
IPCC	Intergovernmental Panel on Climate Change
ISO	International Organization for Standardization
JSBC	Japan Sustainable Building Consortium
LBNL	Lawrence Berkeley National Laboratory
LEED	Leadership in Energy and Environmental Design
LHMC	Latin-Hypercube Monte-Carlo
MBE	Mean Bias Error
MC	Monte-Carlo
MCC	Motor Control Center
MCSA	Monte-Carlo Sensitivity Analysis
MM	Mixed Mode
MPE	Model Parameter Estimation
MSE	Mean Squared Error
NEEAP	National Energy Efficiency Action Plan
NUI	National University of Ireland
NWS	National Weather Service
O&M	Operation & Maintenance
PAYS	Pay-As-You-Save
PIPA	Post Implementation Performance Analysis
PMV	Predicted Mean Vote (Fanger)
PPD	Predicted Percentage Dissatisfied (Fanger)
RESEM	Renewable Energy Savings Estimation Method
RH	Relative Humidity
RMI	Rocky Mountain Institute
RSA	Regional Sensitivity Analysis
RTU	Roof-Top Unit
SA	Sensitivity Analysis
SCM	Source Control Management
SDCV	Single-Duct Constant Volume
SDVAV	Single-Duct Variable Air Volume
SEAP	Simplified Energy Analysis Procedure
SSA	Systematic Sensitivity Analysis
SSE	Sum of Squared Errors
STEM	Short-Term Energy Monitoring

TABS	Thermally Activated Building Systems
TMY	Typical Meteorological Year
UA	Uncertainty Analysis
UC	University of California
UFAD	Underfloor Air Distribution
UQ	Uncertainty Quantification
US DOE	United States Department Of Energy
USGBC	United States Green Building Council
VAV	Variable Air Volume
VDA	Visual Data Analysis
VFD	Variable Frequency Drive
VOC	Volatile Organic Compounds
VSF	Variable Speed Fan
WBCSD	World Business Council for Sustainable Development
Wh	Watt-hour
WHO	World Health Organisation

List of Tables

List of Figures

Declarations

This thesis or any part thereof, has not been, or is not currently being submitted for any degree at any other university.

Daniel Edward Coakley

The work reported herein is as a result of my own investigations, except where acknowledged and referenced.

Daniel Edward Coakley

Acknowledgements

"Sometimes our light goes out but is blown into flame by another human being. Each of us owes deepest thanks to those who have rekindled this light."

– Albert Schweitzer, Philosopher

Over the last few years, I have been extremely fortunate to have developed a loyal network of friends, colleagues, mentors and associates, who have been a constant support throughout my research. To those who have offered their time, assistance, friendship and knowledge, I am eternally grateful.

In particular, I would like to thank my supervisors, Dr. Padraig Molloy, and Dr. Marcus Keane. They have provided encouragement, motivation and direction at each stage of my research, and have always been available to provide assistance with any issues I've faced during my PhD.

Secondly, Dr. Paul Raftery, deserves a special thanks. His knowledge, enthusiasm, expertise and technical guidance were instrumental in supporting me throughout my work. He has always been a mentor, offering guidance when it was needed, as well as being a loyal friend and colleague.

The support of my Graduate Research Committee (GRC) has also played a crucial role in mentoring progress and ensuring timely completion of my research. I would like to thank Prof. Sean Leen for chairing this committee and ensuring critical deadlines were met. I would also like to thank Prof. Peter McHugh and Dr. Mark Bruzzi for their supporting role on this committee.

I am lucky to have also had the support of my colleagues within the Informatics Research Unit for Sustainable Engineering (IRUSE), as well as the wider network of staff and researchers in the College of Engineering and Informatics at NUI Galway. I would particularly like to thank all the IRUSE researchers, past and present, who have supported me for the past few years – Magda, Ignacio, Marco, Andrea, Ray, Jesus, Luis, Ed, Ciara, Martin and Gearóid. I am also grateful to my friends from Civil Engineering (R105) and Mechanical Engineering (CIMRU), for the tea breaks and the social occasions, which always provided a welcome break. To all my friends – in Ireland and abroad – thank you for being there.

I would also like to thank the buildings, technical and administrative staff who provided their time and resources to enable the successful completion of this work. In particular, I would like to thank Mr. Noel O'Connor, Mr. Aodh Dalton and Mr. Gerry Nolan, for their efforts in supporting this work, and their enthusiasm in supporting the student learning experience on campus. I would also like to express my gratitude to the administrators and technical staff in both Civil and Mechanical engineering, who have dealt with countless queries from myself and others over the last number of years, in particular Ms. Bríd O'Flaherty, Ms. Maura McDonagh and Ms. Carmel King (Civil Engineering), as well as Ms. Jane Bowman and Ms. Sharon Gilmartin (Mechanical Engineering).

Thank you also to Prof. Padraic O'Donoghue and Prof. Sean Leen, for access to the facilities and resources in the Departments of Civil and Mechanical Engineering. I am also extremely grateful for the support provided by the College of Engineering and Informatics Fellowship, which funded this research.

Finally, I would like to thank my family, for always being there for me, and for supporting me throughout my academic career to date. To my parents, Noelline and Ted, and brother, Alan – thank you for everything.

Daniel Coakley

May 2014

Chapter 1: Introduction

"Anyone who believes in indefinite growth in anything physical, on a physically finite planet, is either mad or an economist."

– Kenneth E. Boulding, Economist

1.1 Introduction

Globally, the topic of energy is a central concern for many individuals, businesses, corporations and governments. An abundance of energy supply to meet demand is of critical importance in sustaining life and promoting growth. However, there currently exist a number of energy challenges which pose threats to future sustainability of life and economic growth. This chapter will focus on a number of key areas, including:

- Global challenges: The section deals with the current major challenges faced by humanity in the context of energy, including; population growth, urbanisation, fossil fuel depletion, security of energy supply, and the threat of climate change;
- Buildings and Energy: This section examines the building sector in the context of energy consumption, focussing in particular on; energy efficiency in buildings, policy and legislation relating to buildings, and barriers to energy conservation in the built environment;
- Whole Building Energy Modelling: This section details the potential for the use of energy simulation in improving energy efficiency in buildings as well as the current issues relating to energy modelling.

Following this discussion, the topic of this research will be discussed briefly, along with an overview of the proposed work and outline of the thesis.

1.2 Global Challenges

This section will examine energy consumption in a global context in order to identify the key challenges that must be addressed to provide sustainable energy future. These include:

- Population growth;
- Urbanisation;
- Increasing demand for fossil fuels;
- Security of energy supply;
- Environmental Issues.

1.2.1 Population Growth

The world human population has increased dramatically over the past 12,000 years, with growth rates increasing since the early 1800’s (Figure 1-1). The earth has never before sustained such a large human population. This change has naturally created strains on our planet's natural resources, necessary for sustaining life and growth.

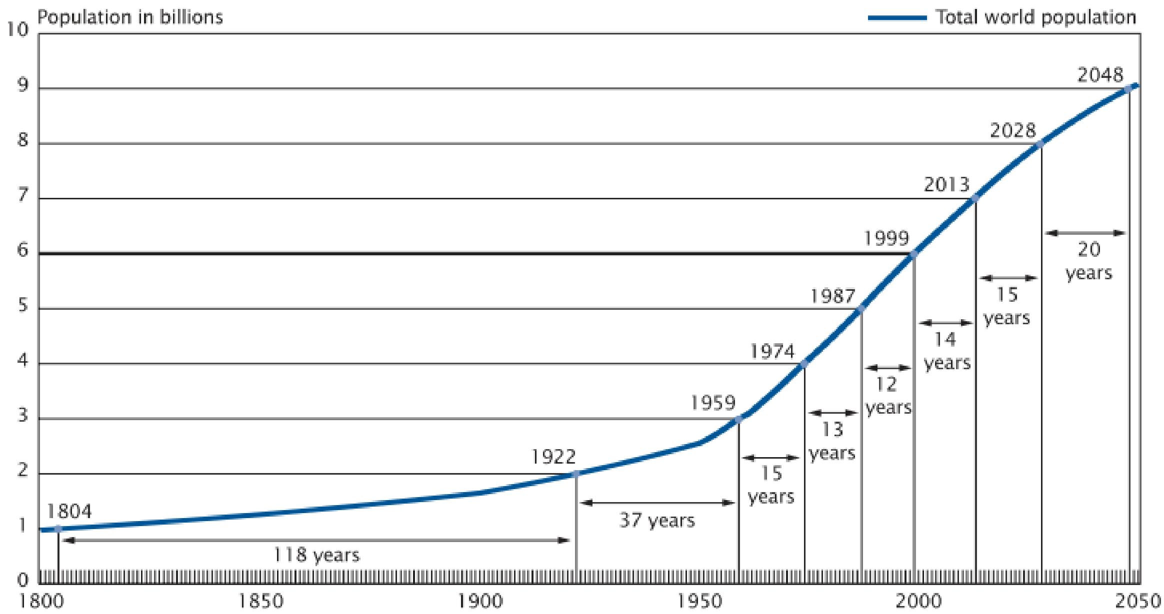


Figure 1-1: World Population Growth, 1800-2050, time to successive billions (Rowe et al. 2004)

Current world population is in excess of 7 billion people. According to the 2012 revision of the official United Nations (U.N.) population estimates and projections (United Nations 2013), this figure is projected to increase by almost one billion people within the next twelve years, reaching 8.1 billion in 2025, and to exceed 9 billion by 2050, with much of this growth coming from non-OECD countries, primarily India and China. While population growth rate is starting to slow, there are new challenges presented by increasing *urbanisation* of the human population and associated increase in *energy use intensity* (EUI).

1.2.2 Urbanisation of human population

As well as population growth, humanity is also experiencing a major shift towards urbanisation; that is an increasing percentage of the population living in concentrated urban areas. This trend is strongly correlated with industrialisation, and therefore can be seen to be highest in modern developed nations. At present, over half of the world's population live in urban areas. By 2050, the population living in urban areas is projected to gain 2.6 billion, increasing from 3.6 billion in 2011 to 6.3 billion 2050. Thus, the urban areas of the world are expected to absorb all the population growth expected over the next four decades while at the same time drawing in some of the rural population. (United Nations 2011)

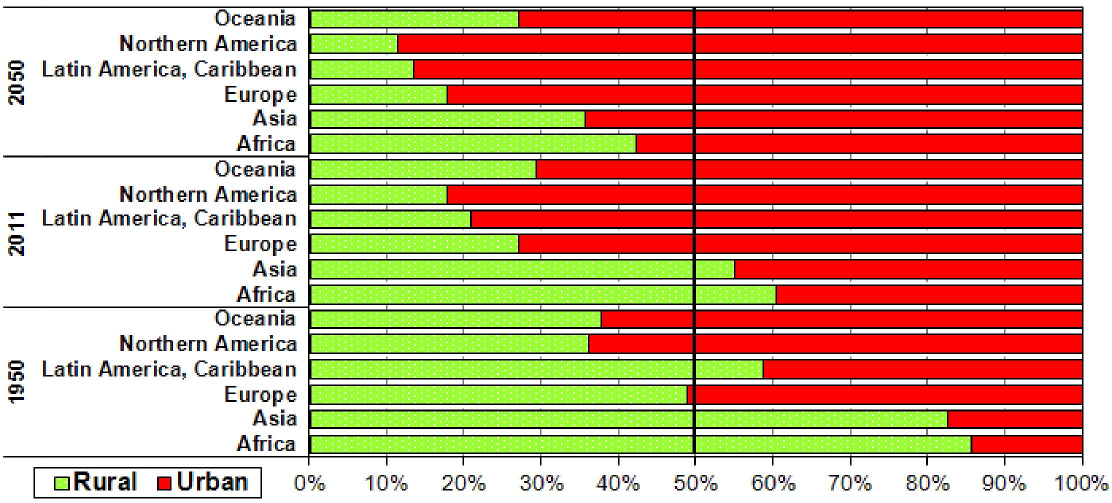


Figure 1-2: Urban and rural population by development regions, 1950, 2011 and 2050 (United Nations 2011)

This trend is primarily driven by concentration of investment and employment opportunities in urban areas. As with population growth trends, India and China are projected to dominate the growth of urban population over the next few decades, with up to 98% of population growth in China attributed to urban population growth from 2000-2050 (Figure 1-3).

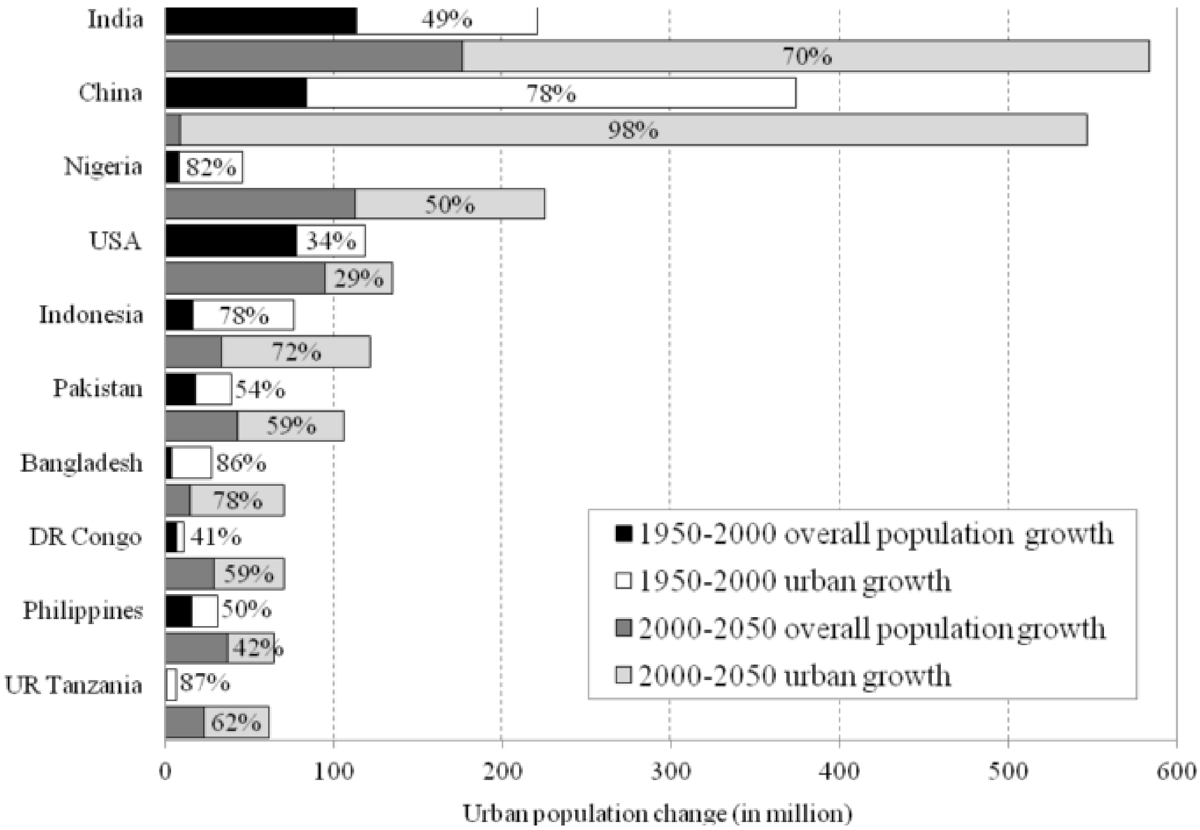


Figure 1-3: Contribution of demographic and urban growth to urbanization among the ten countries with the largest increase in their urban population between 2000 and 2050, 1950-2050 (United Nations 2011)

While urbanisation presents opportunities for inhabitants, it also presents significant challenges:

- **Urban Heat Island (UHI):** This phenomenon of increased temperature rise in urban areas, resulting in a "heat island" effect in built up areas (see Figure 1-4). The annual mean air temperature of a city can be 1–3°C warmer than its surroundings. In the evening, the difference can be as high as 12°C (Akbari et al. 1992). This results in increasing summertime peak energy demand, air conditioning costs (Hassid and Santamouris 2000), air pollution and greenhouse gas emissions, as well as heat-related illness and mortality. The prospect of increased climate change will only serve to exacerbate the current problems associated with UHI.

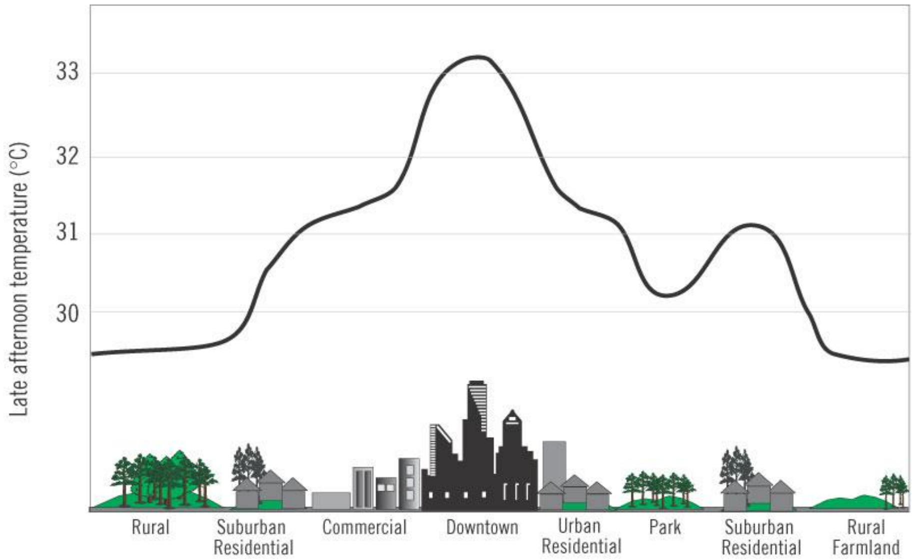


Figure 1-4: Urban Heat Island (UHI) Effect (Akbari et al. 1992)

- **Increased Energy Use Intensity (EUI) per Capita**: Urbanisation has the added effect of increasing energy consumption per capita. This is due to a combination of factors. Personal transportation in rural areas generally entails far less energy use than urban transportation does (i.e. cars, buses and taxis). Higher density living induces substitutions of modern production processes and techniques for more traditional methods (e.g. farming, food production and processing, industrial production). Finally, food must be transported longer distances to urban consumers than to rural agricultural customers (Jones 1989). This phenomenon has been the focus of numerous studies, with many recorded energy use intensities for urban dwellers of the order of 10 times greater than their rural counterparts (Crompton and Wu 2005; Dhakal 2009). This is of particular concern again in countries such as China which currently has the highest rate of urbanisation in the world, at just over 50%, with its urban population expected to reach the one billion mark by 2030.

1.2.3 Increasing Demand for Fossil Fuels

As a direct result of increasing population growth, industrialisation and increasing urbanisa tion, there has been a dramatic increase in energy demand worldwide, primarily driven by Asian countries in recent decades (Figure 1-5). According to the IEA World Energy Outlook, the world's primary energy demand has increased by 58% in 25 years, from about 7.2 billion TOE (tonne of oil equivalent) in 1980 to about 11.4 billion TOE in 2005. Current energy demand is expected to increase by 55% by 2030 at an average annual rate of 1.8%, if the present energy trends continue (IEA 2007). The OECD (Organization for Economic Cooperation and Development) countries used to be the centre of energy demand. However, in IEA energy projections to 2035 (IEA 2011), the focus for this growth is predicted to shift largely to non-OECD countries (e.g. China, India), which are projected to account for 90% of demand growth.

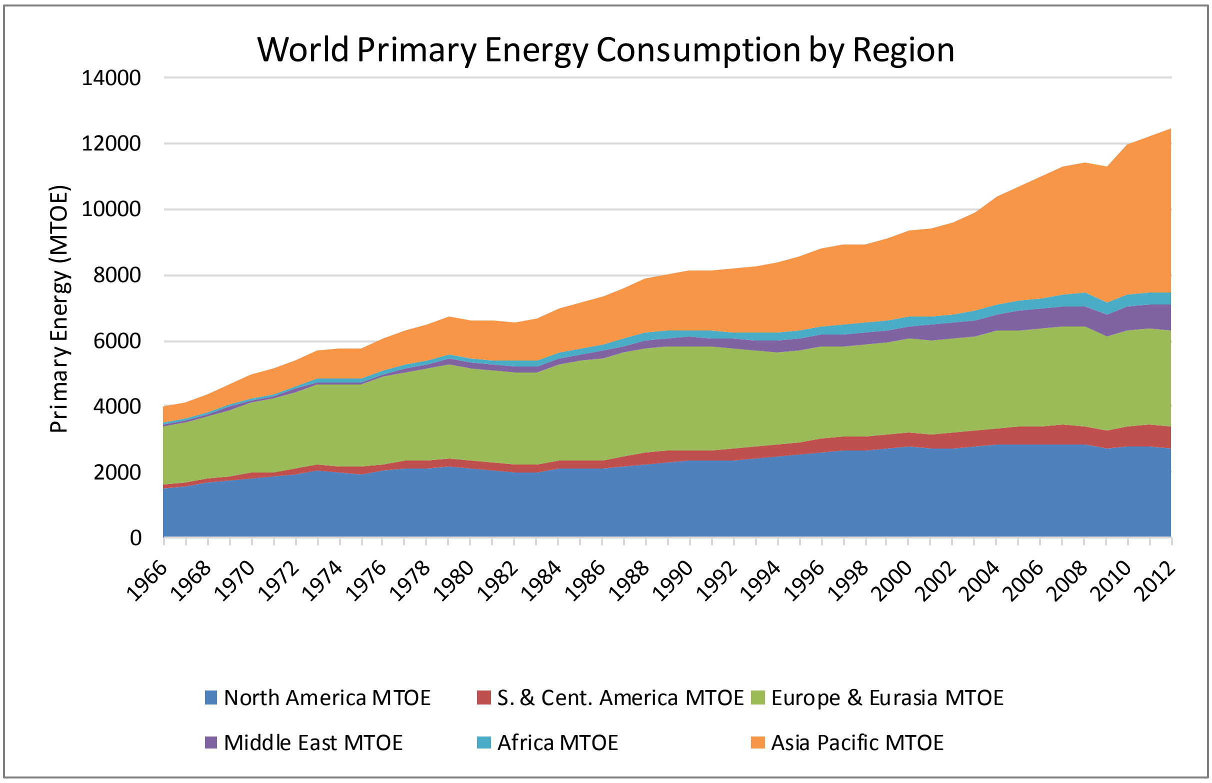


Figure 1-5: World primary energy consumption by region, 1966-2012 (British Petroleum (BP) 2012)

At present, this growth in demand is primarily served by an abundant supply of fossil fuels (coal, oil and natural gas), with over 80% of energy being supplied by these sources (see Figure 1-6). However, should this growth trend continue, supply and cost of these fuels will become major issues within the coming decades, particularly in countries which do not have a domestic fuel economy (e.g. Ireland – see section 1.2.4.1). While renewables are growing progressively, it is expected they will not reach the levels of growth required to meet the fossil fuel 'gap'.

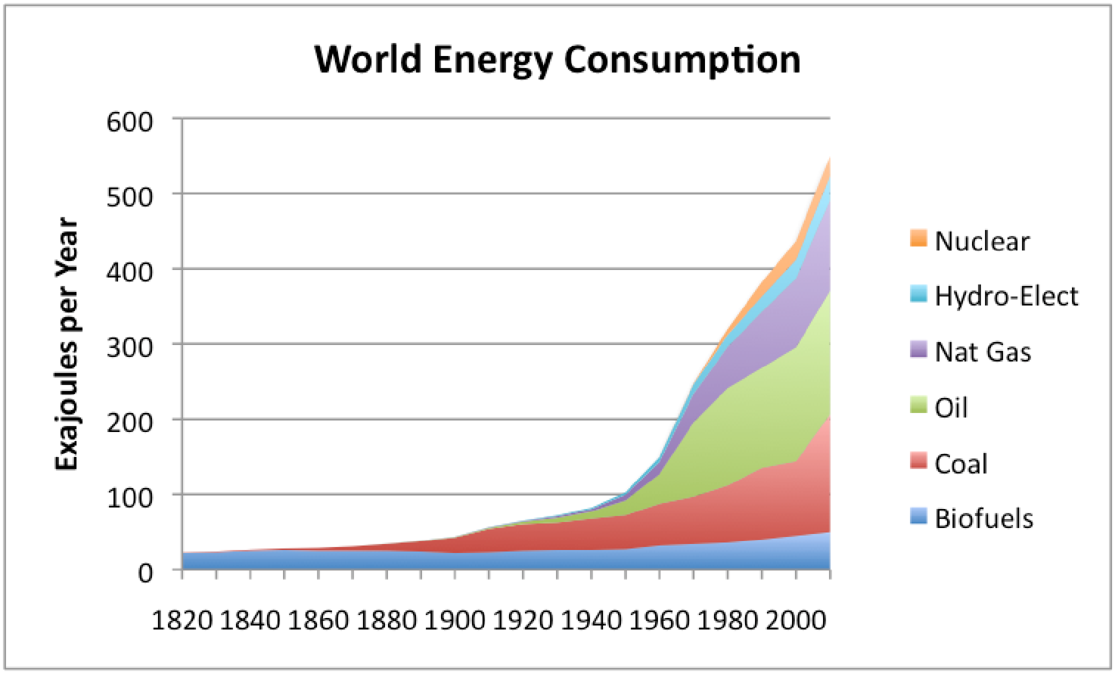


Figure 1-6: World Energy Consumption by Source (British Petroleum (BP) 2012)

Currently, the requirement for growth in fossil fuel production to meet predicted demand is unsustainable economically, socially and environmentally.

1.2.3.1 Peak Oil

There has been much speculation in the past over peak oil, the point in time when the maximum rate of global petroleum extraction is reached, after which the rate of production enters terminal decline. It is suspected that this point in production has already been reach for conventional oil resources. Global oil production has seen weak growth since 2007, with inflation-adjusted price back to levels last seen during the 1978 oil crisis (see Figure 1-7).

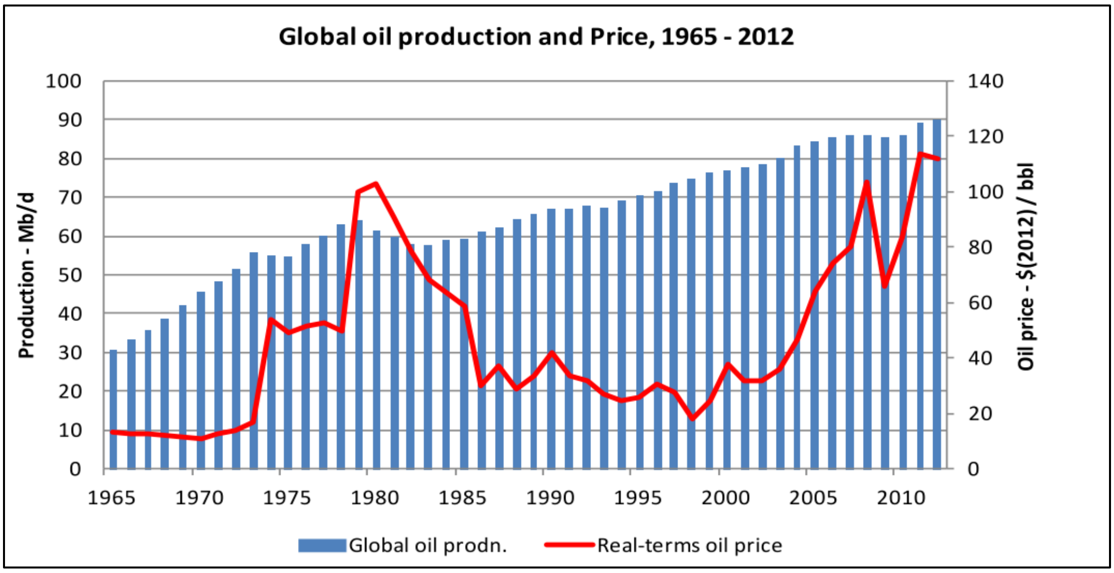


Figure 1-7: Global Oil Production & Price, 1965-2012 (BP)

There have been developments in the area of unconventional resources (i.e. difficult to produce) such as shale gas and oil. However, these are currently the subject of debate in terms of environmental compatibility (e.g. impact of fracking) and social acceptability. Concerns have been raised over air quality, noise pollution, earth tremors, groundwater pollution and land-surface scarring.

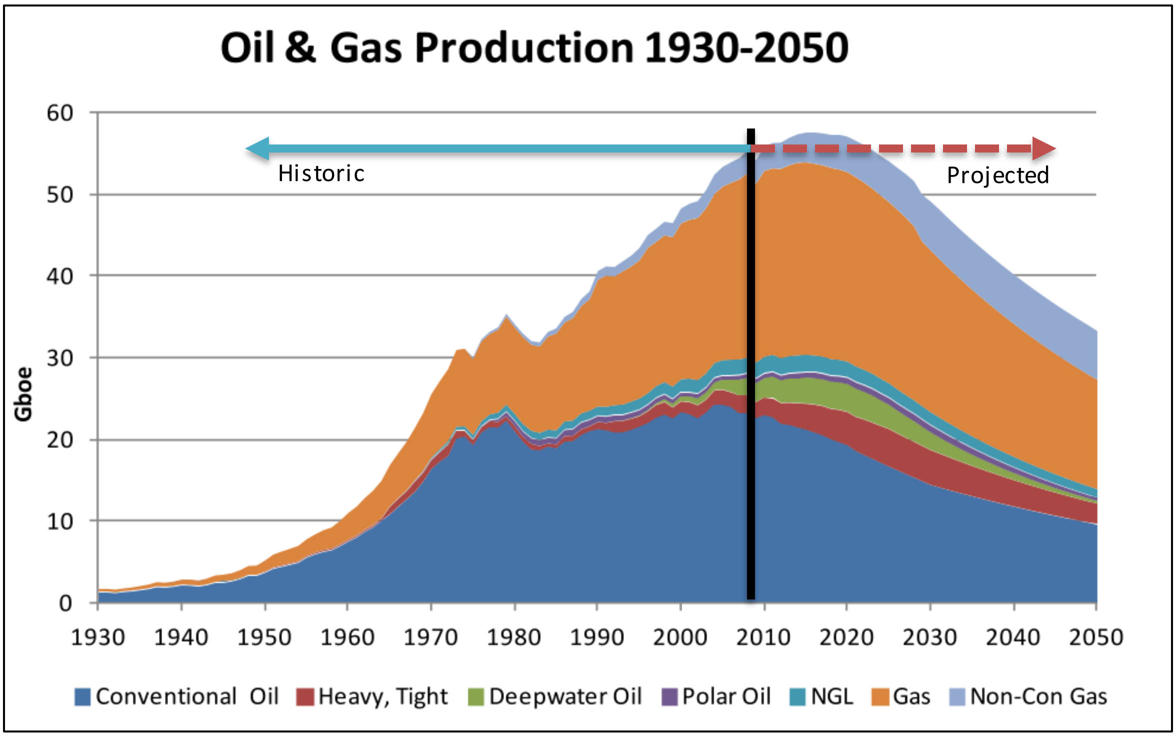
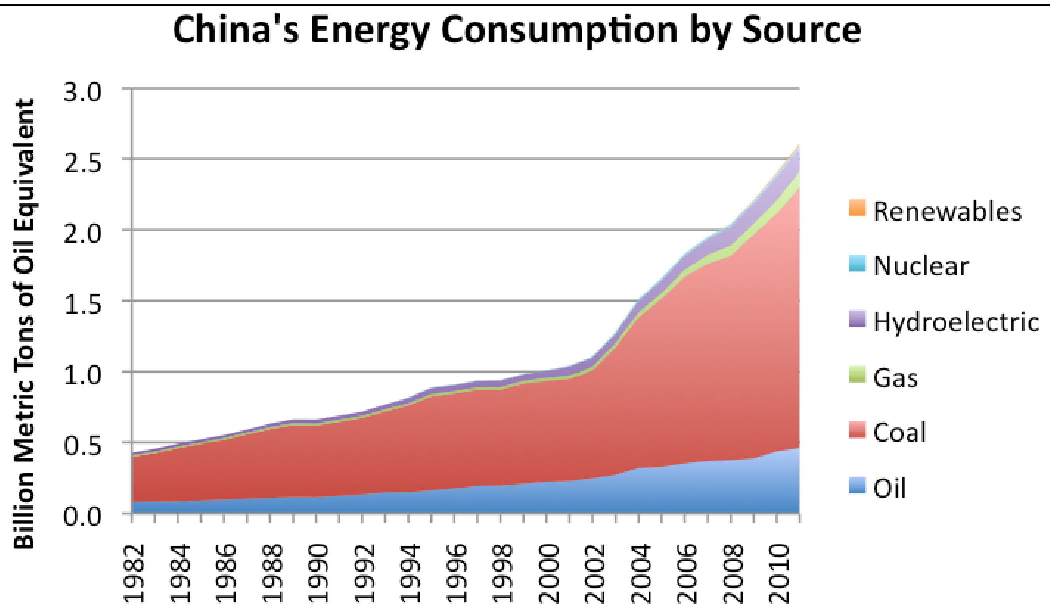
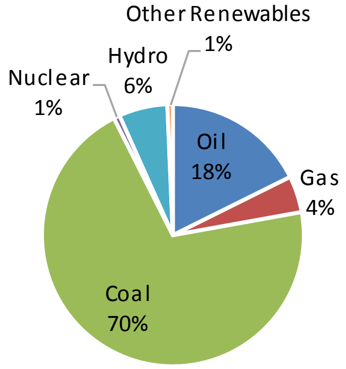


Figure 1-8: Oil & Gas Production, 1930-2050 (Source: Colin Campbell, 2009)

It is predicted that conventional oil production is already on the verge of decline, with some growth in natural gas and unconventional resources (shale oil and gas) helping to meet demand growth. However, these are also expected to hit peak production around 2020 (Figure 1-8). Of particular interest again is China, which has been increasing production of coal (see Figure 1-9) and has embarked on an accelerated programme of building coal-fired power plants, adding net capacity of 50GW in 2012 alone (Meade 2013; Muller 2013), and accounted for half of the world's coal consumption in 2013. With predicted population growth and urbanisation rates in China, this increasing shift towards coal production creates a worrying scenario in terms of growth in greenhouse-gas emissions and associated climate change uncertainty.





Source	MtOE	%
Oil	461.8	17.67
Gas	117.6	4.50
Coal	1839.4	70.39
Nuclear	19.5	0.75
Hydro	157	6.01
Other Renewables	17.7	0.68
Total	2613	

Figure 1-9: China's Energy Consumption by Source (British Petroleum (BP) 2012)

1.2.4 Security of Energy Supply

Security of energy supply is becoming a major issue for many governments, particularly with rising competition from emerging economies like India and China, as well as persistent political instability in many energy producing regions (e.g. Iraq, Iran). This instability has led to spikes in prices for crude oil in the past, severely impacting developed economies which rely on imported oil for trade and production (see Figure 1-10). The U.S. has been tackling its import dependency through the production of domestic resources such as shale gas and oil and recently moved from a net gas importer to a net gas exporter. In the EU, more than half (54.1 %) of the EU-27's gross energy consumption in 2010 came from imported sources. As well as fuel import dependency, the European Union (EU) faces additional challenges in the form of a fragmented energy market and requirement to reduce GHG emissions and shift to cleaner renewable fuels to combat climate change.

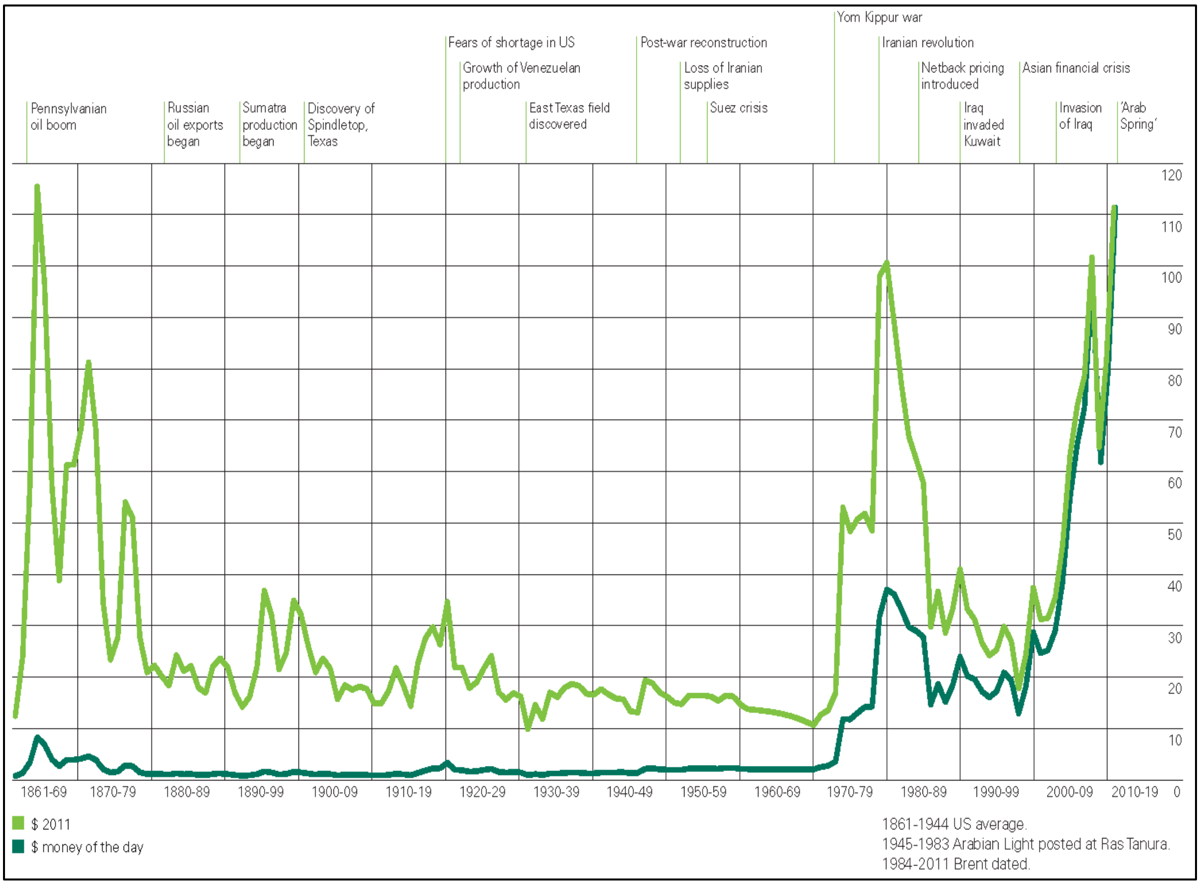


Figure 1-10: Crude Oil Price Fluctuation during World Events (British Petroleum (BP) 2012)

1.2.4.1 *Ireland's Energy Supply*

Currently over 90% of Irish energy requirements are imported. Combined with a peripheral location and small market scale, this current reality leaves Ireland highly vulnerable to supply disruption and imported price volatility. Security of energy supply is a global issue and the

European Union's growing reliance on energy imports increases Ireland's overall energy vulnerability.

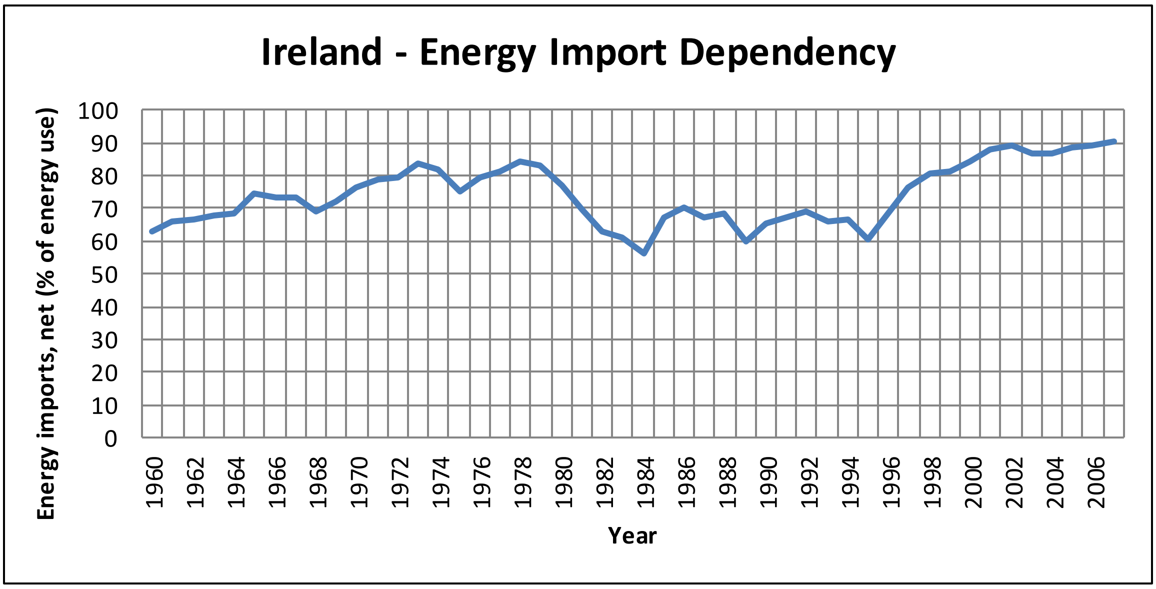


Figure 1-11: Energy Import Dependency in Ireland

1.2.5 Environmental Issues

As well as the lack of resources to meet physical demand, there also exists the issue of global warming brought on by increased concentrations of greenhouse gases in the earth's atmosphere. It is widely accepted that anthropogenic factors are a major contributor to climate change (Stern 2006). The current level of greenhouse gases in the atmosphere is equivalent to around 430 parts per million (ppm) CO_2, compared with only 280ppm before the Industrial Revolution. This has been primarily attributed to the increased use of fossil fuels, particularly since the industrial revolution (see Figure 1-12).

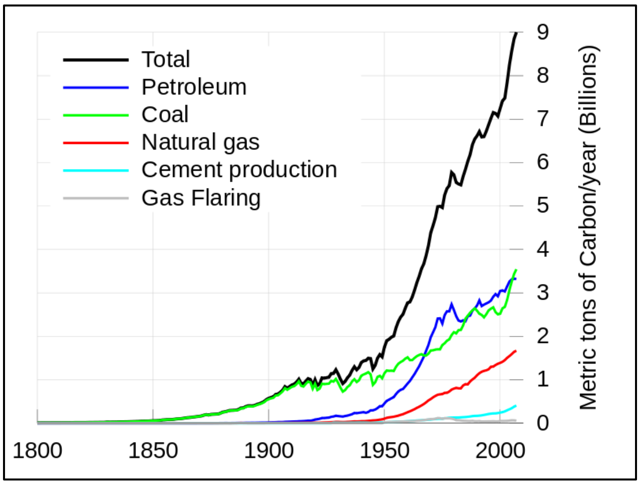


Figure 1-12: Carbon Emissions by Source

In response to growing concerns over this issue, the Intergovernmental Panel on Climate Change (IPCC) was established in 1988 tasked with examining the influence of human activities on climate change. The Fourth Assessment Report of the IPCC estimated that global greenhouse gas (GHG) emissions due to human activities rose by 70% between 1970 and 2004 (Metz et al. 2007; Intergovernmental Panel on Climate Change 2007).

1.2.5.1 Global Temperature Anomalies

There has already been recorded anomalies in global mean air temperature over the past century (Figure 1-13). If global CO_2 emissions continue to grow at the current rate the IPCC predicts a further increase of global mean temperature during the 21st century of about 0.3°C per decade.

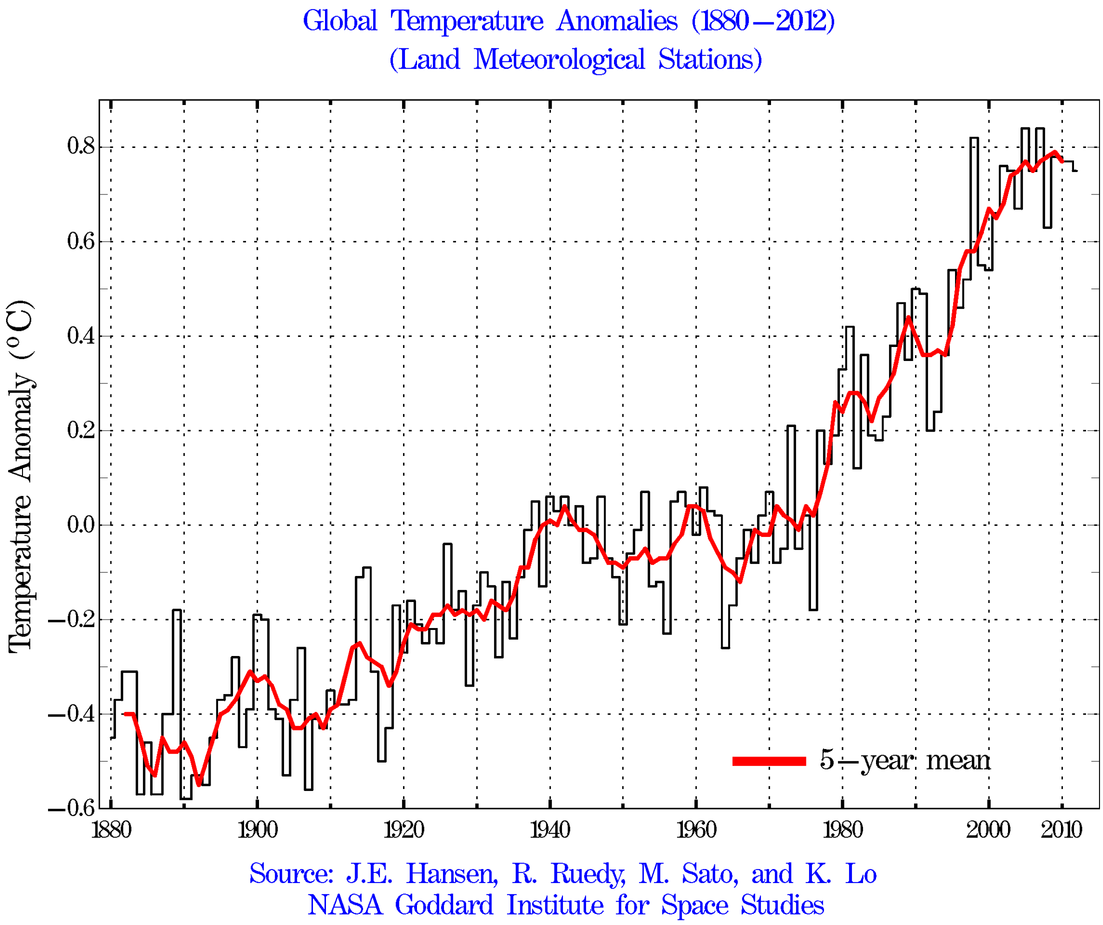


Figure 1-13: Global Mean Temperature Rise, 1999-2008

The impact of anthropogenic climate change has already been seen to have had an impact in other areas also, which are causing increasing concern:

1.2.5.2 Ocean Acidification

The increased concentration of CO_2 in the earth's oceans (Figure 1-14) has become a major focus for earth and ocean scientists. This is of particular importance given the sensitivity of many ocean species (e.g. Krill) to minor changes in ocean chemistry. (Revelle and Suess 1957; Doney et al. 2009). Since Krill are considered an important species in the food chain for a number of dependent predators (e.g. whales, seals and penguins), there is a possibility that this single effect of global warming could have a devastating impact on our planets ecosystem.

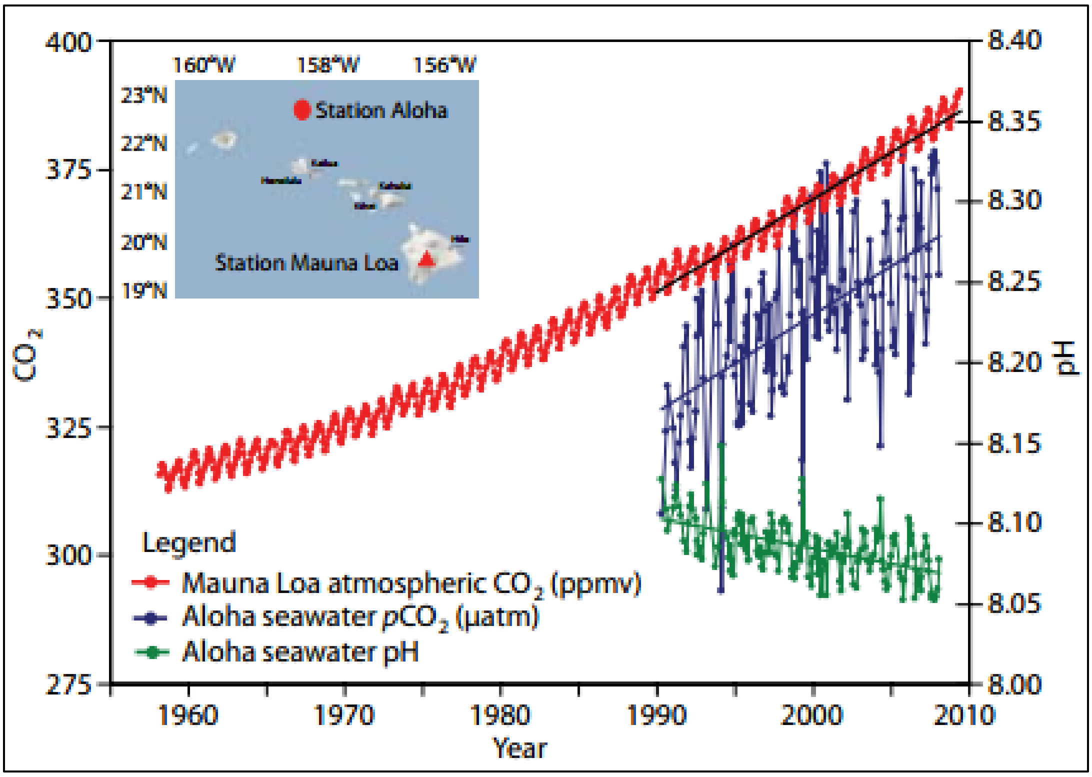


Figure 1-14: Concentration of Carbon Dioxide (CO2) in Ocean, 1960-2010

1.2.5.3 Arctic Sea Ice Retreat

The extent of Summer arctic sea ice declined to unprecedentedly low levels in September 2007 (Stroeve and Serreze 2008), as evidenced in Figure 1-15. This has been attributed to the progressive rise in mean air temperature over the past two centuries. Because of the growing extent of open water in recent summers, ice cover in the following spring is increasingly dominated by thin first-year ice that is more vulnerable to melting out completely during the summer. This thinner ice in spring in turn causes a stronger summer ice albedo feedback through earlier formation of open water areas. This process is repeated and reinforced year after year, accelerating a transition to a seasonally open Artic ocean (Stroeve et al. 2011).

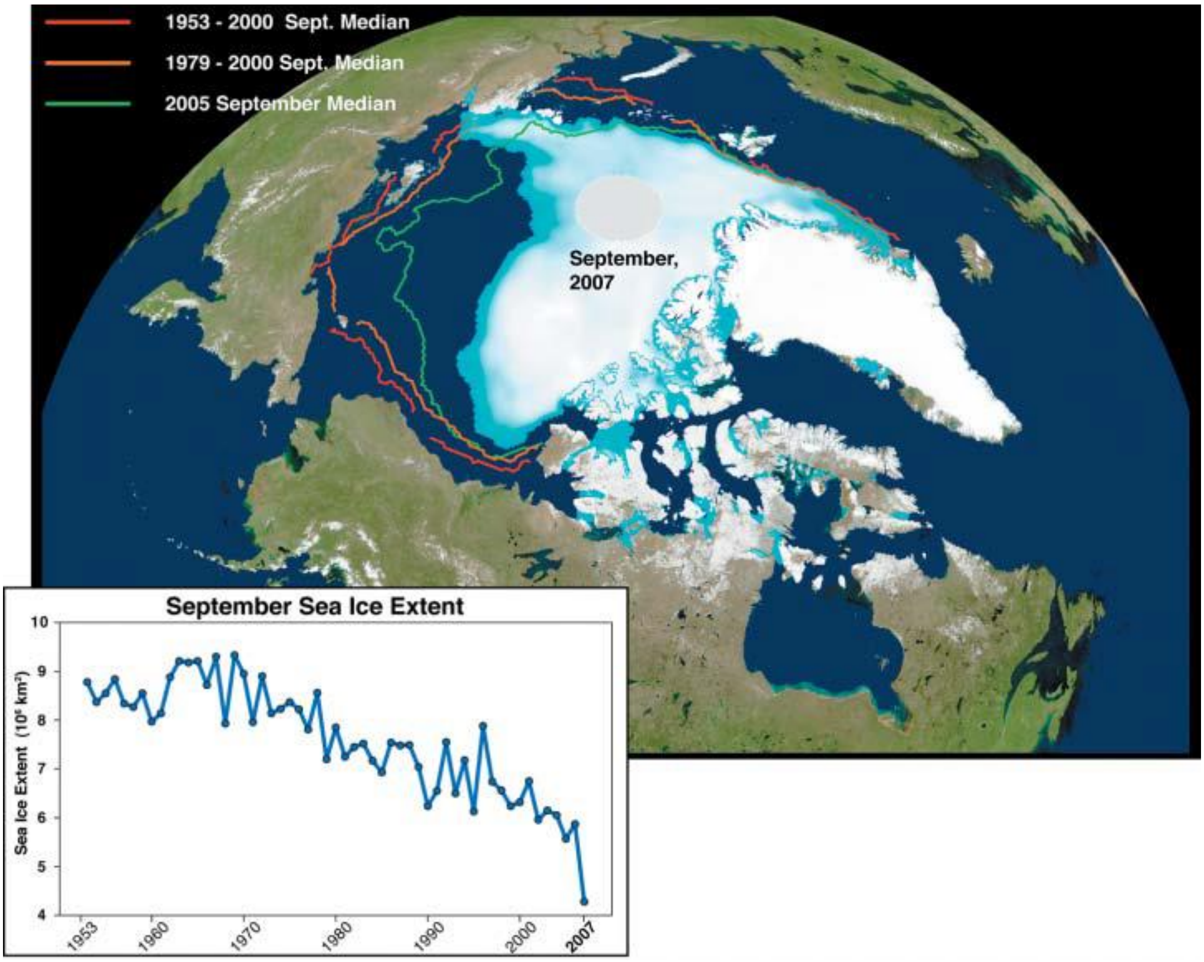


Figure 1-15: Sea ice concentration for September 2007, along with Arctic Ocean median extent from 1953 to 2000 (red curve), from 1979 to 2000 (orange curve), and for September 2005 (green curve). September ice extent time series from 1953 to 2007 is shown to the bottom left (Stroeve and Serreze 2008).

1.2.5.4 Global Response

The United Nations Framework Convention on Climate Change (UNFCCC) was established in 1992 to provide a framework for policy making to mitigate climate change. The Kyoto Protocol (UNFCCC Secretariat 1997) and its successor, the Copenhagen Accord (UNFCCC Secretariat 2009) aimed to establish an international agreement to mitigate GHG emissions, particularly amongst the highest contributors.

1.3 Buildings and Energy

The built environment accounts for approximately 40% of global energy consumption and around 30-40% of greenhouse gas emissions. In the United States, the world's second largest energy consumer after China, the buildings sector accounted for about 41% of primary energy consumption in 2010, compared to 30% by the industrial sector and 29% by the transportation sector (see Figure 1-16). This statistic is reflected similarly across most developed nations. According to the U.S. Department of Energy, total building primary energy consumption in 2009 was about 48% higher than consumption in 1980 with space heating, space cooling, and lighting the dominant end uses in 2010, accounting for close to half of all energy consumed in the buildings sector.

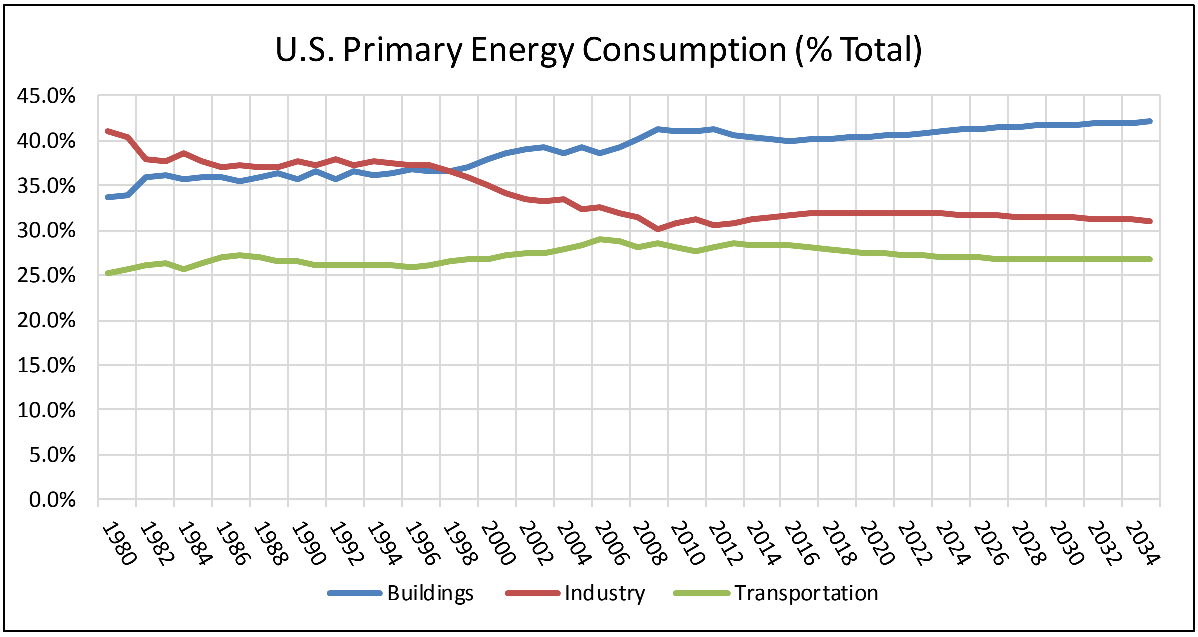


Figure 1-16: Buildings (Residential & Commercial) Energy Consumption (U.S. DOE)

1.3.1 Energy Efficiency

The high level of energy consumption and GHG emissions in the buildings sector in Europe and the US make it an obvious sector to target to improve energy performance. The justification for focussing on efficiency targets can be summarised by a number of arguments from an individual and societal point of view:

- Lower GHG emissions;
- Reduced energy costs for consumers and avoidance of 'fuel poverty';
- Cheaper than investing in increased energy capacity;
- Improved comfort and indoor air quality (IAQ) for building occupants.

Energy efficiency projects also contribute to the objective of sustainable development and rehabilitation of older buildings, as well as providing employment to the building energy services sector. Efficiency improvements also offer significant potential for improvement in existing processes. Currently only 20–35% of the chemical energy of the fuel burned is typically transformed to useful energy (see Figure 1-17). This is primarily due to losses in conversion, distribution and end-use process efficiency.

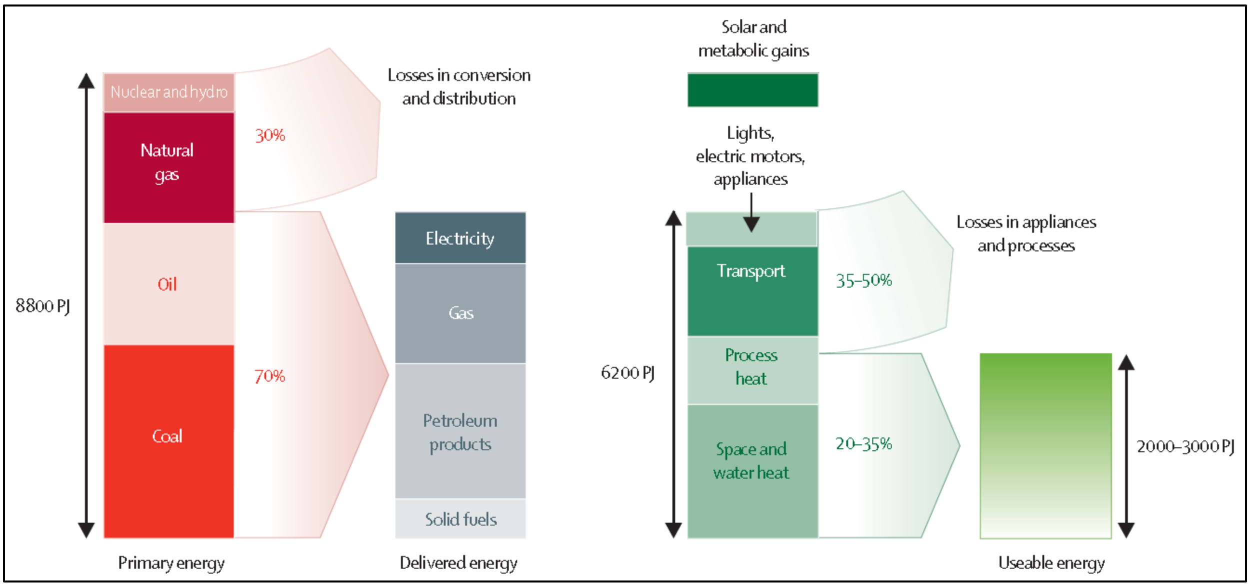


Figure 1-17: Schematic of UK Energy Flow (Wilkinson et al. 2007)

1.3.2 Policy and Legislation

The proliferation of energy consumption and CO_2 emissions in the built environment has made energy efficiency and savings strategies a priority objective for energy policies in most countries (Pérez-Lombard et al. 2008). The US and EU have increased efforts to reduce building energy consumption through prescriptive approaches by introducing stringent standards and codes of practice. The Energy Performance of Buildings Directive (EPBD) in the EU (2002), for example, required the obligatory energy certification of new and existing buildings as well as display of this certification and other relevant information in public buildings. In addition, the Energy End-use Efficiency and Energy Services Directive (2006) requires member states to draw up national action plans to achieve 1% yearly energy savings in the retail, supply and distribution of electricity, natural gas, urban heating, and other energy products including transport fuels.

In June 2000 the EU's European Commission launched the European Climate Change Programme (ECCP). The goal of the ECCP is to identify, develop and implement all the necessary elements of an EU strategy to implement the Kyoto Protocol. It was under this programme that the Emissions Trading Scheme was set up whereby countries could purchase Carbon credits in order to comply with Kyoto targets.

At a national level, these EU directives have translated into a number of country-level national energy action plans. For example, the “Green Deal” which forms part of Britain’s new Energy Bill, aims to revolutionise energy efficiency of British properties. The new framework will enable private firms to offer energy efficiency improvements to their homes and businesses at no upfront cost, and recoup payments through a charge in instalments on their energy bill. In Ireland, the National Energy Efficiency Action Plan (NEEAP) aims to deliver a 20% energy saving target by 2020, including a 33% reduction in public sector energy use. Similar to the ‘Green Deal’, the NEEAP includes a framework for energy performance contracting (EPC)[1], under a Pay-As-You-Save (PAYS) model.

Furthermore, in response to the EU EPBD Directive, Ireland have introduced legislation making Building Energy Rating (BER) certificates a mandatory requirement for all new buildings constructed and all buildings being sold or let after 1st January 2009 (Building Control Authority 2013). The revised building regulations also provide for the introduction of a methodology for building energy performance assessment in the case of new dwellings commencing on or after 1 July 2006 as required by Articles 3, 4 and 5 and Annex of the EPBD. Building certification can help overcome the “first cost” barrier of energy efficiency measures by integrating the operational costs of each building into its market value.

[1] Energy (Savings) Performance Contracts (EPC): a contractual agreement between the beneficiary and the provider (normally an ESCO) of an energy efficiency improvement measure, where investments in that measure are paid for in relation to a contractually agreed level of energy efficiency improvement. An ESCO is a natural or legal person that delivers energy services and/or other energy efficiency improvement measures in a user’s facility or premises

1.3.3 Barriers to Change

A number of barriers to the implementation of energy efficiency measures in buildings have been identified: (Evander et al. 2004; Deringer et al. 2004; Carbon Trust 2005; Yao et al. 2005; Levine et al. 2007; UNEP 2009; DECC 2012).

Table 1-1: Barriers to adoption of Energy Conservation Measures (ECM's)

Category	Barrier
Economic	Upfront cost of measures.
	Length of time required for measures to pay back.
	Poor ratio of investment cost to value of energy savings.
	Costs or risks that are not captured directly in financial flows (e.g. hardware/software incompatibilities, performance risks etc.)
Political and Structural	Market structures and constraints that prevent a consistent trade-off between specific efficiency investment and energy saving benefits
	Structural characteristics of political, economic, energy system which make efficiency investment difficult.
Behavioural and Organisational	Behavioural characteristics of individuals and companies that hinder energy efficiency technologies and practices;
	Difficulty in planning and carrying out work
Information	Lack of knowledge and awareness about the benefits.

While some of these barriers relate to broader market and political factors, there are also barriers relating to lack of information, awareness and ability to effectively assess and plan ECM's. Some of these challenges may be addressed by streamlining the process of identifying and analysing various energy efficiency, through the use of whole building energy simulation.

1.4 Whole Building Energy Modelling

Whole building energy simulation relates to the practice of modelling detailed building processes such as heating, cooling, lighting, ventilation and water use. Energy Simulation tools have been used since the early 1960's to analyse the thermal behaviour and energy consumption in buildings. Initially, they were primarily used in the design stage to optimise the design of the building envelope and HVAC systems. More recently, building energy simulation (BES) models have been employed in the post-construction stage of the building life-cycle for a number of purposes, such as:

- Design alternative evaluation (Trčka and Hensen 2010);
- Design Optimisation (Larsen and Filippín 2008; Attia et al. 2012)
- Benchmarking of Building Energy Consumption (Bertagnolio and Lebrun 2008);

- Continuous Commissioning (Claridge 2004; Liu et al. 2003);
- Operation Optimisation (Sun & A. Reddy 2005);
- Simulation-assisted Building Control (Clarke et al. 1993; Coffey et al. 2010);
- Technical and economical evaluation of Energy Conservation Measures (ECM's) (Waltz 2000; Iqbal & Al-Homoud 2007).

In order for BES models to be used with any degree of confidence, it is necessary that the existing model closely represent the actual behaviour of the building under study. Therefore, the purpose of model calibration is to reduce the discrepancies between building energy simulation (BES) and measured building performance.

1.4.1 Available Tools

There are currently a large number of tools available for assessing energy use in buildings. The applicability of these tools will, of course, depend on the requirement of the final simulation model, as discussed in Section 3.3. A comprehensive list of over 400 software tools for evaluating energy efficiency, renewable energy and sustainability in buildings is provided by US DOE (2013). These include databases, spreadsheets, component and systems analyses, and whole-building energy performance simulation programs. Crawley et al. (2008) presented a comparison of the main features and capabilities of the top 20 tools available in 2008. While many of these features have since been expanded, the broad comparisons remain applicable.

1.4.2 Issues with whole building energy simulation.

According to Tupper et al. (2011), modellers rarely complete accurate, quality calibration of energy models for existing buildings due to:

- The lack of understanding and consistent use of standardized methods;
- Building energy modelling being an over-specified problem;
- The expense and time needed to obtain the required hourly sub-metered data, which is usually not available;
- The lack of integrated tools and automated methods that could assist calibration.

Since the calibration problem is itself over-parameterised and under-determined, it is impossible to find an exact, unique solution. As a result, calibration methodologies and results are often not discussed in detail in many case studies and an approach in which the analyst tunes, or "fudges" (Troncoso 1997), some of the myriad of parameters until the model meets the acceptance criteria is commonly used. 'Furthermore, when a model is established as being calibrated (i.e., the user reports that the accuracy for electricity consumption is ±5% per month), the author does not reveal the techniques used other than stating the final result. Hourly and daily values are seldom reported. In cases where error estimates are presented, the methods and equations used to obtain comparisons are not.' (Haberl & Bou-Saada 1998). These ad-hoc, subjective approaches are not systematic and not explicitly evidence-based.

1.5 Problem Statement

Building simulation models provide a means of understanding building operation as well as optimising performance and making informed decisions on ECMs (Energy Conservation Measures). However, due to the complexity of the built environment and prevalence of large numbers of independent interacting variables, there is a great deal of difficulty in achieving an accurate representation of a real-world building operation to a reasonable degree of accuracy.

One means of increasing the accuracy of simulation models is by means of 'calibration' using actual measured data to verify outputs. However, to date, there is no general consensus as to how computer simulation models should be calibrated or how their accuracy can be verified. The basic issue is the fact that the calibration problem is over-parameterised and underdetermined, i.e. there are vast arrays of input parameters (building system variables, occupancy/load schedule variables, environmental factors etc.) and limited monitored data to determine the impact and interaction of these parameters in a computer model. Furthermore, the calibration problem is under-defined and may be satisfied by countless unique solutions. Currently, a model can be considered to be 'calibrated' if it achieves output accuracy to within 5% of actual measured output data (e.g. monthly energy use, kWh). However, this does not account for inconsistencies in hourly and sub-hourly model behaviour, nor does it address potential inaccuracies present in model inputs.

1.6 Research Question

Can an analytical optimization approach be used to further enhance evidence-based approaches to BES model development and calibration? How can model input uncertainty be propagated through the modelling process in order to quantify risk during performance evaluation of energy conservation measures (ECM's)?

1.7 Overview of the Proposed Approach

The main objective of this thesis is to develop and improve on existing analytical methodologies which may be applied to calibration models to achieve a greater degree of real-world correlation, thereby increasing the reliability of the final calibrated model. This will involve an extensive review of literature relating to energy use in buildings, energy modelling, and building energy model calibration. A calibration methodology which builds on the best available tools and applications is proposed. This proposed approach utilises systematic evidence-based model development, in conjunction with analytical optimisation procedures to produce accurate reliable models capable of simulating real building operational performance. This methodology will be applied to an existing building in order to demonstrate its viability.

1.8 Thesis Outline

The remaining chapters of the thesis are as follows:

- Chapter 2 provides an extensive background and literature review on whole building energy simulation, including: performance criteria in buildings, energy simulation and current approaches to simulation calibration.
- Chapter 3 describes the proposed BES calibration process, including: model preparation, data collection, evidence-based model development and improvement, sensitivity analysis, and parametric simulation.
- Chapter 4 provides an overview of the proposed methodology as applied to a real building, including detailed descriptions of the tools and processes employed at each stage of the process.
- Chapter 5 describes the results at each stage of the calibration process as well as a summary of the novel visualization techniques employed during the process.
- Chapter 6 describes the research conclusions as well as ideas for future development.

Chapter 2:
Literature review

"If I have seen further it is by standing on the shoulders of Giants."

– Isaac Newton, Physicist

2.1 Introduction

This chapter will examine the role of Building Energy Performance Simulation (BEPS) in the context of design and operation of buildings, with specific focus on the following:

- Energy Performance Criteria in Buildings (see Section 2.2);
- Simulation and the Built Environment (See Section 2.3);
    - Types of Models and Inverse Methods;
    - Building Energy Performance Simulation (BEPS) Tools;
    - Benefits of Calibrated Simulation;
    - Problems with BEPS and Model Calibration;
    - Uncertainty in Building Simulation;
- Current approaches to BEPS calibration (see Sections 2.4 - 2.6);
    - Manual Approaches;
    - Automated Approaches;
- Conclusions and Proposed Approach (see Sections 2.7 and 0).

Buildings represent complex systems with high levels of interdependence on many dynamic external sources (weather, occupancy etc.). In addition, the optimisation of building systems requires balancing of sometimes contradictory objectives in terms of energy efficiency and overall performance. These performance criteria will be examined in detail in Section 2.2.

Building Energy Performance Simulation (BEPS) tools provide an efficient means of conducting performance-based analysis and optimisation, taking into account the multitude of complex model interdependencies, internal and external inputs as well as various performance objectives. These tools are compared in more detail in Section 2.3, along with the current problems associated with the use of BEPS, primarily model calibration (i.e. minimising discrepancies between measured and simulated data). In addition, the importance of risk and uncertainty is highlighted, particularly in the context of building simulation and energy conservation measure (ECM) evaluation.

A detailed review of current approaches to model calibration is presented in Section 2.4, focussing on the various analytical and mathematical/statistical tools employed by practitioners to date. This is followed by a discussion on both the problems and the merits of the presented approaches, along with a recommendation for future procedural approaches to BEPS calibration, thus forming a basis for the methodology presented in this thesis.

2.2 Energy Performance Criteria in Buildings

There are three main performance criteria which drive energy use in buildings:

- Thermal Comfort;
- Ventilation and Indoor Air Quality (IAQ);
- Visual and Aural Comfort.

In many respects, the objective of optimising energy performance in buildings is at odds with the goal of optimising these performance criteria. For example, increasing thermal comfort (through increased HVAC design capacity, operation schedules, set points etc.) may have the effect of increasing energy consumption and cost. Conversely, the goal of energy performance optimisation is the reduction in energy consumption through the implementation of energy conservation measures (ECM's) which may include measures which have a negative impact on thermal comfort. Therefore, it is important to understand the principles which govern these metrics in order to strike a balance of satisfactory performance in terms of both comfort and energy performance.

2.2.1 Thermal Comfort

One of the primary functions of buildings is to create and maintain a comfortable environment for its occupants. Thermal comfort is defined as "*the condition of mind which expresses satisfaction with the thermal environment*" (ASHRAE 2004a). It is beyond the scope of this thesis to discuss in depth the factors affecting thermal comfort in the built environment. However, there is an extensive body of scientific literature which deals specifically with this topic (Fanger 1970; McIntyre 1978; McIntyre 1980; Fisk 1981; Nicol and Humphreys 2002; de Dear 2004). For the purpose of this review though, it is useful to understand the main factors which affect thermal comfort and how they relate to controllable parameters in the built environment. There are six primary factors which affect conditions for thermal comfort. These may be split into (1) physical or environmental factors and (2) physiological or personal factors (see Table 2-1):

Table 2-1: Factors affecting thermal comfort

Physical / Environmental	Physiological / Personal
• Air temperature (t_a)	• Clothing level (*clo*)
• Relative humidity (h)	• Activity level / metabolic rate (*met*)
• Mean radiant temperature (t_r)	
• Air Velocity (v_a)	

Initial studies on thermal comfort and formulation of thermal indices began in the 1920's. The Bedford Scale (Bedford 1936) and subsequent thermal sensation scale (ASHRAE 2004a) form the basis for today's subjective/field study comfort studies (Table 2-2).

Table 2-2: Thermal Sensation Scale

Thermal Sensation Scale	Bedford Scale
+3 Hot	Much too Warm
+2 Warm	Too warm
+1 Slightly Warm	Comfortably warm
0 Neutral	Comfortable
-1 Slightly cool	Comfortably cool
-2 Cool	Too cool
-3 Cold	Much too cold

While the scale in Table 2-2 represents the subjective opinions of the occupants surveyed, a deterministic comfort model has also been devised to approximate these responses based on empirical studies carried out in controlled climate chambers. Fanger's Predicted Mean Vote (PMV) approach (Fanger 1970) is the most widely adopted analytical approach today, and forms the basis for the international standard, ISO 7730, on the ergonomics of the thermal environment (ISO 2005) as well as ASHRAE Standard 55 on thermal environmental conditions for human occupancy (ASHRAE 2004a). The PMV is essentially an index that predicts the mean value of the votes of a large group of persons on the above 7-point thermal sensation scale (see Table 2-2). Since PMV predicts the mean of the comfort votes from a large group, a more practical measure for building performance assessment is the Predicted Percentage Dissatisfied (PPD) which indicates the fraction of a sample population that will be dissatisfied with the thermal environment (i.e. those that would vote $\geq$2 or $\leq$-2 on the thermal sensation scale).

PMV can be directly linked with specific design and control parameters (CIBSE 2006a):

- **Temperature**: Thermal comfort is influenced by a combination of mean room air temperature and radiant temperature. These temperatures are the most influential variables affecting comfort.
- **Air Movement and Draught**: Air speed, temperature, direction and fluctuation influence the perception of draught. This is a particularly important consideration when designing naturally ventilated buildings.

- **Humidity**: This is less important in moderate thermal environments where humidity in the range of 40-70% is generally acceptable. However, it is an important consideration in terms of microbial growth or static electricity.

2.2.2 Ventilation and Indoor Air Quality (IAQ)

Indoor air quality (IAQ) is influenced by air infiltration, ventilation efficiency as well as indoor and outdoor pollution (dust, contaminants etc.). In normal working environments, fresh air is an important requirement to provide air for respiration as well as diluting and removing any contaminants which may be present. These contaminants may be naturally occurring (e.g. CO_2 from human respiration) or artificial (e.g. Smoke, volatile organic compounds (VOC's) released by paints and plastics) (WHO 1983; WHO 1989).

Natural or mechanical ventilation is an easy way to control IAQ in most buildings, provided the outdoor air used for ventilation is acceptable. There are standard recommendations available for ventilation rates for different space types where the main contaminants are occupation odours and carbon dioxide by-products of respiration (ASHRAE 2004b; CIBSE 2006a; ISO 2007). In the absence of accurate systems for measuring occupancy, HVAC systems will often over-compensate for ventilation requirements, ignoring over-ventilation losses at partial occupancy (Doty 2009). This may be controlled more effectively by including procedures for automated compensation for reduced occupancy (e.g. CO_2 sensors or proportional damper control).

Ventilation effectiveness is also an important consideration relating to the quality of the air supplied to a space, and more specifically, to the occupant. This measure is based on whether the air is heated or cooled upon discharge and whether it is discharged at ceiling or floor level.

2.2.3 Visual and Acoustic Comfort

Visual comfort relates to the effective use of day-lighting in buildings, characterised primarily by the day-lighting factor, defined as "*the ratio of the daylight illumination at a given point on a given plane due to the light received directly or indirectly from a sky of assumed or known illuminance distribution, to the illumination on a horizontal plane due to an unobstructed hemisphere of the sky*" (Hopkinson et al. 1966). In simple terms, working spaces require sufficient light (defined in terms of illuminance, or lux), with adequate brightness. It is also important to consider 'glare' (reflection due to direct light sources) in the work-space, particularly when designing office spaces which require the use of personal computer screens which may be significantly impacted by glare, causing visual discomfort and reduced productivity (Abdou 1997; Hua et al. 2011). There are also significant gains to be made, both in terms of visual comfort and energy consumption, through the use of daylighting in place of artificial lighting (Bodart and Herde 2002).

The main requirement for acoustic comfort is the absence of distracting noise which detrimentally affects the ability to carry out work. Noise is defined by frequency (hertz, Hz) and sound pressure (decibels, dB). Design curves (Noise Rating, NR curves) are available with recommended acoustic criteria for various building types (CIBSE 2006b).

2.2.4 Conclusions

Buildings are designed to provide an appropriate and comfortable environment for its occupants, in terms of thermal, visual and aural comfort. It is also expected to provide safe working conditions, free of natural and artificial contaminants, through the use of sufficient ventilation. It is important to consider these requirements when designing low-energy buildings or when considering energy conservation measures (ECM's) so that energy efficiency does not come at the expense of occupant comfort or safety. In this regard, simulation tools can play a significant role. With the aid of expert input, they can be used to effectively find the optimum balance between comfort, cost and efficiency. In addition, simulation tools can provide additional insight and information to designers considering novel alternative design approaches, where prior information may not be readily available.

2.3 Simulation and the Built Environment

In order to understand Building Energy Simulation, it is necessary to understand scientific models in general. According to Saltelli et al. (2008), models can be:

- *Diagnostic* or *Prognostic*: Diagnostic models are used to identify the nature or cause of some phenomenon. In other words, it may be used to better understand the laws which govern a given system. Prognostic models, on the other hand, are used to predict the behaviour of a system, given a set of well-defined laws governing that system.
- *Law-Driven or Data-Driven:* Law-Driven (or forward) models apply a given set of laws (e.g. gravity, heat/mass transfer etc.) which govern a system, in order to predict its behaviour given system properties and conditions. Data-Driven (or inverse) models work on the opposite approach, using system behaviour as a predictor for system properties. Therefore, data-driven models can be used to describe a system with a minimal set of adjustable inputs (Young et al. 1996). In contrast, law driven models are often over-parameterised, in that they require more inputs than available data can support. However, the advantage of law-driven models is that they offer the ability to model system behaviour given a set of previously unobserved conditions, while data-driven models would require prior data in order to model behaviour. A simplified comparison of law-driven and data driven models is presented in Figure 2-1 (Florita and Henze 2013).

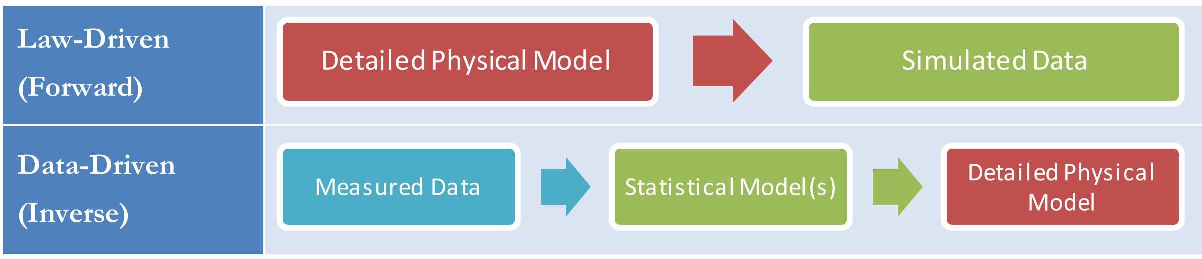


Figure 2-1: Law-Driven (Forward) models vs. Data-Driven (inverse) models

Building Energy Simulation (BES) models, as used in building design, can generally be classified as prognostic law-driven models in that they are used to predict the behaviour of a complex system given a set of well-defined laws (e.g. energy balance, mass balance, conductivity, heat transfer, human physiology etc.).

Conversely, Data-driven (Inverse) approaches, in the context of building energy modelling, refer to methods which use monitored data from the building to produce models which are capable of accurately predicting system behaviour. Inverse methods for energy use estimation in buildings can be broadly classified into three main approaches (Reddy and Andersen 2002):

i. **Black-Box**: This refers to the use of simple mathematical or statistical models (e.g. regression, neural-networks etc.) which relate a set of influential input parameters (e.g. occupancy, weather) to measured outputs. Model input coefficients are determined such that they produce an algorithm with the ability to predict system behaviour. It is important to note that these input coefficients have no relationship with the actual physical environment.

ii. **Grey-Box / Parameter Estimation**: Grey box approaches differ from black-box approaches in that they use certain key or aggregated system parameters identified from a physical system model.

iii. **Detailed Model Calibration**: The final approach uses a fully-descriptive law-driven model of a building system and tunes the various inputs to match the measured data. This approach is the primary focus of this research as it provides a number of benefits over black and grey-box approaches.

2.3.1 Building Energy Performance Simulation (BEPS) Tools

Whole building energy simulation tools allow the detailed calculation of the energy required to maintain specified building performance criteria (see Section 2.2), under the influence of external inputs such as weather, occupancy and infiltration. Detailed heat-balance calculations are carried out at discrete time-steps based on the physical properties of the building and mechanical systems as well as these dynamic external inputs (weather, occupancy, lighting and equipment loads etc.). These calculations are generally performed over the course of a full year. These tools generally fall into the category of prognostic law-driven simulation tools. Some of the main tools which will be discussed during the course of this review are:

- DOE-2 (Winkelmann et al. 1993) is a freeware building energy simulation tool which predicts the hourly energy use and energy cost of a building given hourly weather information, a building geometric and HVAC description, and utility rate structure. It development was funded by the U.S. Department of Energy (DOE), hence the name.

- EnergyPlus (Crawley and Lawrie 2001) is an advanced whole building energy simulation tool, developed on the basis of work carried out on DOE-2. It incorporates the same functionality as DOE-2, producing hourly (or sub-hourly) energy costs of a building given system input information. It also incorporates many advanced features not available in DOE-2, such as multi-zone airflow and extensive HVAC specification capabilities.

- TRNSYS (Klein et al. 1979) is a transient system simulation program with a modular structure which implements a component-based simulation approach. Component may be simple systems like pumps or fans, or complex systems such as multi-zone buildings.
- ESP-r (ESRU 1974) is an integrated modelling tool for the simulation of the thermal, visual and acoustic performance of buildings. Similar to EnergyPlus and DOE-2, ESP-r requires user-specified information regarding building geometry, HVAC systems, components and schedules. It supports explicit energy balance in each zone and at each surface as well as incorporating inherent uncertainty and sensitivity analysis capabilities.

The above four simulation programs represent the most common tools encountered in conducting this review. However, many more tools are available, some of which are tailored specifically to certain tasks (e.g. HVAC simulation, solar gain, daylighting etc.). Crawley et al. (2008) presents a comparison of the main features and capabilities of the top 20 tools available in 2008.

2.3.2 Benefits of BEPS

While the initial focus of BEPS tools was primarily on the design phase, simulation is now becoming increasingly relevant in latter post-construction phases of the building life-cycle (BLC), such as commissioning and operational management and control (Augenbroe 2002). Since BEPS models are based on physical reality, rather than arbitrary mathematical or statistical formulations, they have a number of inherent advantages. One of the primary benefits of detailed simulation models over statistical models is their ability to predict system behaviour given previously unobserved conditions. This allows for analysts to make alterations to the building design or operation, while simultaneously monitoring the effect on system behaviour and performance. Despite the potential benefits and the significant progress which has been made in the development of advanced simulation programmes capable of modelling complex systems and environments, there still remain a number of problems which inhibit their widespread adoption.

2.3.3 Problems with BEPS and Model Calibration

At present, building energy performance simulation models (BEPS) are under-utilised within the AEC industry for a number of reasons, some of which were highlighted in a recent Rocky Mountain Institute (RMI) study (2011). These can be broadly grouped into two main categories, modelling and calibration, as described in Table 2-3.

Table 2-3: Model Development and Calibration Issues

BEPS Modelling Issues	BEPS Calibration Issues
• **Standards**: Lack of understanding and consistent use of standardized methods; • **Expense**: The time, knowledge, expertise and cost required to develop accurate models of building geometry and HVAC systems; • **Integration**: Poor integration between various 3D modelling software packages (such as Autodesk Revit and ArchiCAD) and BEPS simulation packages (such as EnergyPlus, TRNSYS and Modelica).	• **Standards**: Lack of explicit standards for calibration criteria – current guidelines only specify acceptable error ranges for yearly whole-building simulation, but do not account for input uncertainty, sub-metering calibration, or zone-level environmental discrepancies. • **Expense**: The expense and time needed to obtain the required hourly sub-metered data, which is usually not available; • **Simplification**: Calibration is an over-specified and under-determined problem. There are thousands of model inputs but relatively few measurable outputs with which to assess the model accuracy; • **Inputs**: Lack of high-quality input data required for detailed models; • **Uncertainty**: There are currently few studies which account for uncertainty in model inputs and predictions, thus leading to a lack of confidence in BES outputs; • **Identification**: Problems identifying the underlying causes of discrepancies been model predications and measured data. • **Automation**: Lack of integrated tools and automated methods that could assist calibration;

Numerous studies (Karlsson et al. 2007; Turner and Frankel 2008; Scofield 2009) have indicated discrepancies, often significant (up to 100% differences), between BEPS model-predicted and the actual metered building energy use. This undermines confidence in model predictions and curtails adoption of building energy performance tools during design, commissioning and operation. In order for BEPS models to be used with any degree of confidence, it is necessary that the existing model closely represent the actual behaviour of the building under study. This can be achieved through model calibration, the purpose of which is to reduce the discrepancies between BEPS prediction and measured building performance. However, the calibration of forward building energy performance simulation (BEPS) programs, involving thousands of input parameters, to commonly available building energy data is a highly under-determined problem which yields a non-unique solution (Kaplan, Jones, et al. 1990; Carroll and Hitchcock 1993). As a result, calibration methodologies and results are often not discussed in detail in many case studies. An approach in which the analyst tunes, or "fudges" (Troncoso 1997), some of the myriad of input parameters until the model meets the acceptance criteria is commonly used. This is not conducive to the development of reliable building energy simulation models. The current approaches to model calibration, as well as their limitations is discussed in further detail in Section 2.4-2.7.

2.3.4 Validation of BEPS

The validation of building energy simulation models is currently based on compliance with published statistical acceptance criteria, as shown in Table 2-4. These criteria vary depending on whether models are calibrated to monthly or hourly measured data, and are based on standard statistical indices (as discussed in further detail in Section 2.5.4.1):

- Mean-Bias Error, MBE (%) [*Refer to Equation (2.3),* Section 2.5.4.1]: MBE is a good indicator of the overall bias in the model. It captures the mean difference between measured and simulated data points. However, positive bias tends to cancel out negative bias (cancellation effect). Hence, a further measure of model error is also required.

- Coefficient of Variation of Root Mean Square Error, CV RMSE (%) [*Refer to Equation 2.4,* Section 2.5.4.1]: This index allows one to determine how well a model fits the data by capturing offsetting errors between measured and simulated data. It does not suffer from the cancellation effect.

Table 2-4: Acceptance criteria for calibration of BEPS models

Standard/Guideline	Monthly criteria		Hourly criteria	
	MBE	$CVRMSE_{(monthly)}$	MBE	$CVRMSE_{(hourly)}$
ASHRAE Guideline 14 (ASHRAE 2002)	5%	15%	10%	30%
IPMVP (EVO 2007)	20%	-	5%	20%
FEMP (US DOE 2008)	5%	15%	10%	30%

Currently, building energy simulation models are generally considered ‘calibrated’ if they meet the criteria set out by ASHRAE Guideline 14 (ASHRAE 2002). This means that once there is reasonable agreement between measured and simulated data, the model may be deemed ‘calibrated’ according to current international acceptance criteria for BEPS models.

However, in order to holistically address the topic of model calibration it is important to also consider the issue of *model uncertainty*, particularly for indeterminate models of complex systems.

This is an important issue which is often neglected in BEPS calibration studies published to date and is not accounted for by any means in the current BEPS validation criteria. A detailed discussion on this topic follows in Section 2.3.5.

2.3.5 Uncertainty in Building Simulation

Models of complex systems are notoriously difficult to validate and have been the subject of much scientific discussion and debate in terms of quality and uncertainty (Funtowicz and Ravetz 1990). Much of the reason for this debate stems from the fact that models of complex systems represent essential simplifications and simulation constraints. In other words, "the portion of the world captured by the model is an arbitrary enclosure of an otherwise open, interconnected system" (Rosen 1991). This is particularly true when the purpose of the model is to provide some insight into the non-observable parts of the system. Thus mathematical formalisations of partially-observed experiments, even for well-defined or closed systems, can generate non-equivalent descriptions of these system (i.e. models whose outputs are compatible with the same set of observations but whose structures are not reconcilable with one another) (Saltelli et al. 2008). This has also been referred to as equifinality (Beven 1993; Aronica et al. 1998) or model indeterminacy (Oreskes et al. 1994; Saltelli et al. 2008).

The built environment in particular presents a complex challenge in terms of energy modelling and accurate prediction. Any given building is characterised by a multiplicity of parameters including materials properties, occupancy levels, equipment schedules, HVAC and plant operation, climate and weather. These represent diverse sources of model parameter uncertainty. However, this does not illustrate the entire range of potential uncertainty encapsulated by any given building model. Numerous studies have focused on this problem (Macdonald et al. 1999; Macdonald and Strachan 2001; de Wit and Augenbroe 2002a; Macdonald 2002; Moon 2005), although few published case studies incorporate this work into their analyses. De Wit (de Wit and Augenbroe 2002a) classified the various sources of uncertainty in building performance simulation as follows:

- **Specification Uncertainty**: arising from incomplete or inaccurate specification of the building or systems modelled. This may include any exposed model parameters such as; geometry, material properties, HVAC specifications, plant and system schedules etc.
- **Modelling Uncertainty: s**implifications and assumptions of complex physical processes. These assumptions may be explicit to the modeller (zoning, stochastic process scheduling) or hidden by the tool (calculation algorithms).

- **Numerical Uncertainty:** errors introduced in the discretisation and simulation of the model.
- **Scenario Uncertainty:** external conditions imposed on the building, including outdoor climate conditions and occupant behaviour.

It is important that these sources of uncertainty are identified and quantified when assessing model predicted performance. This is particularly important given the 'equifinality' of simulation models (i.e. multiple disparate models may provide the same results). Depending on the application of the BEPS model, it is important to know the degree of uncertainty associated with particular elements of the model or underlying mathematical formulation. This thesis deals primarily with 'specification' and 'modelling' uncertainty and how this can be systematically propagated throughout the simulation model development process.

2.4 Current Approaches to BEPS Calibration:

The main approaches to building energy performance simulation (BEPS) model calibration were first classified by Clarke et al. (1993) and adopted in a later literature review of calibration programs, tools and techniques by Reddy (2006). The four classes proposed are:

i. Calibration based on manual, iterative and pragmatic intervention;

ii. Calibration based on a suite of informative graphical comparative displays;

iii. Calibration based on special tests and analytical procedures;

iv. Analytical/mathematical methods of calibration.

These classifications have been further extended in this review. In general, it was found that approaches to the tuning of simulation models to measured data can be more broadly defined as either *manual* or *automated*.

1. **Manual** – these approaches predominantly rely on iterative pragmatic intervention by the modeller. These include any methods which employs no form of automated calibration through mathematical/statistical methods or otherwise.
2. **Automated** – automated approaches may be described as having some form of automated (i.e. not user driven) process to assist or complete model calibration.

Both manual and automated approaches may employ specific *analytical tools or techniques* to assist in the calibration process (see Section 2.4.1), while automated approaches employ *mathematical and statistical techniques* to reach their goal (see Section 0).

2.4.1 Analytical Tools and Techniques

These can be broadly classified as manual user-driven techniques, but may also be employed as part of an automated calibration process. A list of the main calibration tools and techniques has been compiled following an extensive review of methodologies and applications over the past three decades. For clarity, these are divided into four main categories and presented alongside the relevant key publications in Table 2-5:

- Characterization Techniques: techniques based on the characterisation of the physical and operational characteristics of the building being modelled;
- Advanced Graphical Methods: the use of graphical representations of building data or statistical indices;

- Model Simplification Techniques: techniques which aim to reduce the complexity of simulation models by reducing or aggregating the number of simulation variables;
- Procedural Extensions: the use of standard processes or techniques to improve the simulation and/or model calibration process.

An exhaustive list of the papers mentioned in Table 2-5 is available in Appendix D.1.

Table 2-5: Analytical Tools & Techniques

Acronym	Name	Description	Key Papers
Characterisation Techniques			
AUDIT	Detailed audit	Detailed audits are often conducted prior to building model development in order to gain a better knowledge of the building systems and characteristics (Geometry, HVAC systems, Lighting, Equipment, and Occupancy Schedules).	(Waltz 1992; Shapiro 2009)
EXPERT	Expert Knowledge / Templates / Model Database	Approaches which utilise: • Expert knowledge or judgement as a key element of the process • Prior definition of typical building templates • Database of typical building parameters and components in order to reduce the requirement for user inputs during model development.	(Lebot 1987; Hitchcock et al. 1991; Reddy et al. 2007a; Reddy et al. 2007b)
INT	Intrusive Testing	Intrusive techniques require some intervention in the operation of the actual building, such approach is 'Blink Tests' where-by groups of end-use loads (e.g. plugs loads, lighting etc) are turned on and off in a controlled sequence in order to determine their overall impact on the baseline building load.	(Soebarto 1997)
HIGH	High-Res Data	Data is recorded at hourly (or sub-hourly) levels as opposed to utilising daily load profiles or monthly utility bill data.	(Clarke et al. 1993; Haberl and Bou-Saada 1998; Raftery, Keane and Costa 2011; Coakley et al. 2012)
STEM	Short-Term Energy Monitoring	Metering equipment is used to record on-site measurements for a short period of time (>2 weeks). This may be used in identifying typical energy end-use profiles and/or base-loads.	(Subbarao 1988; Manke et al. 1996; Lunneberg 1999)

Acronym	Name	Description	Key Papers
Advanced Graphical Methods			
3D	3D-Graphical Comparison Techniques	Three-dimensional graphs are used to aid comparison and/or calibration of measured and simulated data. This technique allows users to visualise large quantities of data, compared to traditional 2-D scatter plots etc. which are overwhelmed when analysing large quantities of data points.	(Haberl and Bou-Saada 1998).
SIG	Signature Analysis Methods	Signature analyses techniques are a specific type of graphical analysis technique, typically used by HVAC simulation engineers to identify faulty parameters in Air-Handling Unit (AHU) simulation. They may also be used to develop optimised operation and control schedules. Signature analysis methods are commonly used for the calibration of models based on the simplified energy analysis procedure (SEAP)	(Liu et al. 2003; Liu and Liu 2011)
STAT	Statistical Displays	This refers to the graphical representation of statistical indices and comparisons for easier interpretation. This includes data comparison techniques such as carpet plots, box-whisker mean (BWM) plots and monthly percent difference time-series graphs.	(Haberl and Bou-Saada 1998).
Model Simplification Techniques			
BASE	Base-Case Modelling	The base-case model refers to the use of measured base-loads to calibrate the building model. Base-loads refer to minimum, or weather independent, electrical and gas energy consumption. Calibration is carried out during the base-case when heating and cooling loads are minimal and the building is dominated by internal loads, thus minimising impact of weather dependent variables.	(Yoon and Lee 1999; Yoon et al. 2003)
MPE	Model Parameter Estimation	Deduction of overall aggregate (or lumped) parameters (such as U-values) using non-intrusive measured data.	(Reddy et al. 1999)
PARRED	Parameter Reduction	This involves reducing the requirement for detailed input for variable schedules (e.g. plug loads, lighting, occupancy, equipment etc.). Day-Typing is one such approach which works by analysing long-term data and reducing this to manageable typical day-type schedules (e.g. weekday's vs. weekends, winter vs. summer). Zone-Typing may also be used to reduce large models into similar thermal zones (e.g. Core, Perimeter, Offices, Unoccupied spaces etc.)	(Kaplan, McFerran, et al. 1990; Bronson et al. 1992; Hadley 1993; Raftery, Keane, O'Donnell, et al. 2011)
DISSAG	Data Disaggregation	Data disaggregation refers to the application of non-intrusive techniques to de-couple multiple measured data streams (e.g. energy end-use data from whole-building electrical energy consumption)	(Akbari et al. 1988; Akbari 1995; Akbari and Konopacki 1998)

Acronym	Name	Description	Key Papers
Procedural Extensions			
EVIDENCE	Evidence-Based Model Development	For the purpose of this review, evidence-based approaches may be described as those that implement a procedural approach to model development, making changes according to source evidence rather than ad-hoc intervention. Strictly, this approach should account for adjustments to model parameters in a structured fashion (e.g. using version control software).	(Bou-Saada and Haberl 1995; Raftery, Keane, O'Donnell, et al. 2011; Raftery, Keane and Costa 2011)
SA	Sensitivity Analysis	Sensitivity analysis procedures may be employed in some studies to assess the influence of input parameters on model predictions. This information may be used to identify important parameters for measurement or detailed investigation.	(Lomas and Eppel 1992; Lam and Hui 1996; Saltelli et al. 2004; Westphal and Lamberts 2005; Eisenhower and O'Neill 2012)
UQ	Uncertainty Quantification	This refers to assessment of parameter uncertainty as part of the calibration process. This information may be used to directly assist in model calibration or provide a basis for risk quantification within the results (e.g. uncertainty related risk quantification in ECM analysis).	(Macdonald et al. 1999; Macdonald 2002; de Wit and Augenbroe 2002a; O'Neill et al. 2012; Heo, R Choudhary, et al. 2012)

2.4.2 Mathematical/Statistical Techniques

Modern mathematical and statistical methods are increasingly being employed to assist the calibration process. Applications which employ one or more of these techniques at any stage in the process, have been classified as automated approaches within this framework. Some of the mathematical/statistical approaches employed in calibration studies to date are summarised in Table 2-6, under the following two main categories:

- Optimisation Techniques: This covers the general methods used to optimise prediction performance of any type of model;
- Alternative Modelling Techniques: This section covers alternatives to detailed model calibration, described as black-box or grey-box approaches.

Table 2-6: Mathematical and Statistical Calibration Techniques

Acronym	Name	Description	Key Papers
Optimisation Techniques			
BAYES	Bayesian Calibration	Bayesian calibration is an alternative statistical approach to model calibration. The approach offers the advantage of naturally accounting for uncertainty in model prediction through the use of prior input distributions.	(MacKay 1994; Kennedy and O'Hagan 2001; O'Hagan 2006; Booth et al. 2013)
OBJECT / PENALTY	Objective/Penalty Function	Most mathematical techniques employ some form of optimisation function to reduce the difference between measured and simulated data. An objective function may be used to set a target of minimising, for example, the mean square error between measure and simulated data. Conversely, a penalty function may also be employed to reduce the likelihood to deviating too far from the base-case.	(Carroll and Hitchcock 1993; Sun and Reddy 2006; Reddy et al. 2007a)
Alternative Modelling Techniques			
ANN	Artificial Neural Networks	Neural networks are computational models consisting of an interconnected group of artificial neurons. They are used for modelling complex relationships between inputs and outputs or for finding patterns in data	(Kalogirou and Bojic 2000; Mihalakakou et al. 2002; Karatasou et al. 2006; Neto and Fiorelli 2008)
PSTAR	Primary and Secondary Term Analysis and Re-normalisation	Calibration procedure based on analysis of data from special STEM tests applied to decomposed primary and secondary building energy flows.	(Subbarao 1988; Burch et al. 1990; Balcomb et al. 1993)
META	Meta Modelling	The use of computationally efficient analytical surrogate models which emulate the performance prediction of their complex engineering-based counterparts.	(Eisenhower, O'Neill, Narayanan, et al. 2012; Manfren et al. 2013)
SEAP	Simplified Energy Analysis Procedure	The simplified energy analysis procedure refers to the use of simplified engineering models to represent the building. This may be accomplished by dramatically reducing the number of zones or AHU's in the model by grouping them together.	(Knebel 1983; Katipamula and Claridge 1993; Liu and Claridge 1998)
SYS	Systems Identification	This technique refers to the process of constructing models based only on the observed behaviour of the system (outputs) and a set of external variables (inputs), instead of constructing a detailed model based on 'first principles' of well-known physical variables.	(Goodwin and Payne 1977; Ljung 1987; Liu and Henze 2005)

2.5 Summary of Manual Calibration Developments

Over the past three decades, many procedures have been proposed for the calibration of whole building energy performance simulation models. This section examines manual calibration procedures, chronologically highlighting the new techniques employed by various authors as well as where these techniques have been adapted and advanced.

2.5.1 Characterisation Techniques

Waltz (1992) claims that the single most important factor in developing accurate computer models of existing buildings is developing an intimate knowledge of the physical and operational characteristics of the building being modelled. This section of the review covers techniques which have been used to develop an understanding of these characteristics..

2.5.1.1 Building and Site Audits

An energy audit can be defined as a process to evaluate where a building or plant uses energy, and identifies opportunities to reduce consumption. There is an existing consensus on the definition of three typical levels of building audit (Thumann and Younger 2008):

- **Level 1 – Walkthrough**: This generally implies a tour of the facility and visual inspection of energy using systems. This also includes an evaluation of energy consumption data to analyse energy use quantities and pattern, as well as providing comparisons to industry averages or benchmarks.
- **Level 2 – Standard Audit**: Energy uses and losses are quantified through a more detailed review and analysis of equipment, systems and operational characteristics. On-site measurements may be used to quantify and assess efficiency of energy end-users. This audit also includes an economic analysis of energy conservation measures (ECM's).
- **Level 3 - Investment Grade**: This includes a more detailed review of energy use by function as well as a comprehensive evaluation of energy-use patterns. Energy simulation software is employed to predict year-round energy use, accounting for weather and system variables. The method also accounts for system interactions to prevent over-estimation of savings.

A summary of the main deliverables for the three levels of energy audits is presented in Figure 2-2 (ASHRAE 2011).

Level 1	Level 2	Level 3
• Rough Costs and Savings for ECM's • Identify Capital Projects	• End-Use Breakdown • Detailed Analysis • Cost and Savings for ECM's • O&M Changes	• Refined Analysis • Additional Measurements • Hourly Simulation

Figure 2-2: Deliverables for Energy Audit Levels 1, 2 and 3 (ASHRAE 2011)

Lyberg (1987) provides a comprehensive handbook on energy auditing procedures, defining the auditing process as "*a series of actions, aiming at breaking down into component parts and quantifying the energy used in a building, analysing the applicability, cost and value of measures to reduce energy consumption, and recommending what measures to take*". Lyberg proposes a staged audit process:

1. Building Rating – assessing potential high-potential buildings for audit.
2. Disaggregation of energy consumption (Refer 2.5.3.4).
3. ECO (Energy Conservation Opportunity) identification.
4. ECO (Energy Conservation Opportunity) evaluation.
5. Post Implementation Performance Analysis (PIPA).

An extensive collection of necessary audit templates are also provided in Volume II of the audit handbook (Lyberg 1987), categorised under the following 4 headings: (1) Audit Procedures, (2) Measurement Techniques, (3) Analysis Techniques and (4) Reference Values.

Waltz (1992) suggests two types of survey: (1) observational; and (2) electrical load survey (see Section 2.5.1.2). The observational survey refers to the actual functioning of the buildings control systems as opposed to relying on documentation and as-built drawings. Oftentimes, controls may not be installed as per the design documentation, or operational controls may have been overridden, or have simply failed. The authors also suggest a "late-night" tour of the facility and its HVAC systems to determine 'actual' operating schedules, which often differ from those prescribed in operation & maintenance (O&M) documentation.

CEC (2000) provides a comprehensive guide for reporting investment grade audits of various types of facilities and project types (e.g. Lighting, HVAC, Cogeneration). The guide also includes a copy of sample field data sheets for recording site specific information such as building data, occupancy schedules, lighting and equipment surveys as well as HVAC equipment data.

Ganji and Gilleland (2002) provide assessment of investment grade energy audits and a review of typical cases, identifying several major shortcomings including lack of consistency in auditing, reporting and over-estimation of savings. These shortcomings stem from a number of deficiencies including a lack of expertise and fundamental engineering knowledge on the part of the surveyor. A lack of training in advanced energy simulation software was also identified as an issue, resulting in incorrect outputs in many cases.

Shapiro (2009) also identifies shortcomings in the current approaches to commercial building audits, including a lack of clearly defined boundaries and limitations of simple building audits (Level 1 and Level 2). Shapiro proposes a comprehensive building audit on a room-by-room basis, capturing room-specific opportunities and documenting recommendations in the audit report. Improvements should focus not only on efficiency, but ensuring that the equipment meets the load requirements for the space. An example of a comprehensive lighting audit is given to illustrate how the proposed approach differs from standard walkthrough audits. The author identifies overlit areas and recommends multiple improvements (delamping, occupancy sensors and control changes). In contrast, a typical walthrough audit would record existing equipment but may miss energy reduction recommendations. The proposed comprehensive audit approach is also applied to a case-study office building, identifying 46% potential energy savings, compared with 7% identifiied through a standard walkthrough audit.

To date, a number of standard auditing and energy assessment procedures have been proposed for different industries and applications (CASCADE Consortium 2012):

- AuditAC: Developed as part of a European project "Field Benchmarking and market Development for Audit Methods in Air Conditioning". The project focused on providing tools and information for air-conditioning engineers to identify energy savings in HVAC systems. (Adnot 2007)
- IEA Annex 11: Comprehensive handbook on energy auditing procedures developed in conjunction with the International Energy Agency (IEA) (Lyberg 1987).
- AS/NZS 3598:2000: Standard developed by Australia and New Zealand energy authorities, targeting the commercial and industrial sector. The standard sets out minimum requirements for commissioning and conducting energy audits which identify cost effective opportunities to improve efficiency and effectiveness in the use of energy. (Standards Australia 2000).
- RP-351 Energy Audit Input Procedures and Forms: General ASHRAE procedures for energy auditing including an assessment of existing audit procedures. (ASHRAE 1983).

- ASHRAE Procedures for Commercial Building Energy Audits: Standard for energy companies conducting energy audits of commercial buildings, including definitions of Level 1, 2 and 3 audits (ASHRAE 2011).
- EINSTEIN Audit Methodology: The EINSTEIN (Expert-system for an Intelligent Supply of Thermal Energy in Industry) audit methodology focuses on large scale consumers with high thermal energy (heating and cooling) demand in a low and medium temperature ranges up to 400°C (e.g. manufacturing industry, desalination plants, district heating and cooling networks). (EnergyXperts 2012).

2.5.1.2 Short-Term End-Use Monitoring (STEM)

STEM refers to the application of specialized software and hardware tools to systematically gather and analyse data typically over a short (typically two week) period to evaluate the performance of building energy systems, such as HVAC, controls, and lighting. Diagnostics based on short-term monitoring can clarify how the systems in a building actually perform, as well as highlighting key energy end-users.

A study by the Tishman Research Corporation (TRC 1984) on the calibraton of a DOE-2 office model to measured data was the first identifiable study which incorporated short-term end use monitoring to increase the accuracy of model inputs. Measurement errors for sensors were also accounted for in the study, showing an acceptance of potential uncertainties in the measured end-use values as opposed to solely model inputs.

Waltz (1992) suggests measuring instantaneous power draw for every electrical panel or piece of equipment using a hand-held power factor meter. This is particularly important when high levels of accuracy are required, for example in high-rise multi-zone office builings. Kaplan et al. (1990; 1990) suggest calibrating models to short typical periods as opposed to full year data, for example one month during a heating and cooling season. The authors incorporate short-term energy monitoring during these periods to assist calibration. Statistical analysis is applied to these short-term monitored end-uses to generate manageable DOE-2 schedules for lighting, equipment, occupancy setpoints etc. In this regards, monitored data is used to generate DOE-2 inputs and validate outputs. A similar approach is adopted by Soebarto (1997) for calibrating models to utility bill data using only two to four weeks measured data. The procedure requires the use of STEM in order to develop a set of energy end-use profiles, including; electrical energy, heat energy and, indoor temperatures. The author also proposes the use of intrusive blink-tests (see 2.5.1.4). Short-term monitoring has since been used in a number of studies to assist in identifying input parameters (Lunneberg 1999; Coakley et al. 2012).

2.5.1.3 High-Resolution Data

Clark et al. (1993) investigated the use of calibrated ESP-r simulation to investigate the performance of passive solar components (PASSYS). The study was differentiated by its use of high-quality, high-resolution data and empirical evidence for model calibration and validation. First, a sensitivity analysis (SA) is carried out to quantify uncertainty bands associated with model predictions and associated parameter sensitivities. This information is used to design an experiment to capture a high-quality data set with which to quantify model residuals and identify their cause. The authors also highlight the importance of uncertainties when extrapolating from test-cell scenarios to full-scale application.

A study by Norford et al. (1994) investigated the two-fold differences between a simulation model at design stage and actual operation for a low-energy office building. Focus was placed on high levels of instrumentation (100 sensors polled 200-300 times an hour) to provide hourly averages of ambient and interior conditions as well as energy consumption of HVAC and tenant equipment. The study conluded that differences were mainly due to unanticipated tenant energy consumption (64%), increased HVAC operation beyond design schedule (24%) and specification errors in HVAC equipment, building fabric and infiltration (12%). This highlights the importance of occupant behaviour in determining model performance as well as the need for sufficient instrumentation to monitor this behaviour if it is a significant factor in determining building performance.

2.5.1.4 Intrusive Testing

An approach has been developed for determining characteristic building parameters using controlled heating and cooling tests over short periods of 3-5 days (Subbarao 1988; Manke et al. 1996). This test consists of a period of co-heating to determine an estimate for the building heat-loss co-efficient, and cool-down to provide an estimate for the effective thermal time constant of the building.

Soebarto (1997) presents an approach for calibrating models to utility bill data using only two to four weeks measured data. A series of 'on-off tests' (or Blink Tests) were utilised to determine lighting and plug loads. In these tests, all electrical loads were turned off for a short period, and back on again. This equipment 'on-off' cycling is carried out in a predetermined pattern while recording electrical energy use on a data logger, in order to accurately determine the load profile for various equipment end-users without the need for individual sub-metering. This method resulted in an hourly calibration accuracy of 6.7% CV (RMSE) for whole building electricity and 1% for chilled water energy use.

2.5.2 Advanced Graphical Approaches

In the past, graphical techniques were confined to simple time-series plots (Hunn et al. 1992). With the increasing availability of detailed measured data and requirement to better understand this information, there has been extensive work carried out in the area of graphical data representations.

2.5.2.1 3-D Comparitive Plots

Bronson et al (1992) proposed a means of calibrating hourly building energy models to non-weather dependent (or scheduled) loads using novel comparitive three-dimensional graphics which allowed hourly differences to be viewed for the entire simulation period. Day-typing was also used to assist in the calibration process. The authors reported that the availability of comparative three-dimensional surface plots significantly improved the ability to view small differences between the simulated and measured data, which allowed for the creation of a "super-tuned" DOE-2 simulation that matched the electricity use within 1%. The process of identifying and fixing unknown "misfits" between the simulation and the measured data was significantly enhanced by the use of the plots.

Bou-Saada and Haberl (1995) propose the use of 3D surface plots (see Figure 2-3) and statistical indices (Refer Section 2.5.2.2) to provide a global view of the differences between measured and computed hourly values in order to help identify time-dependent patterns in discrepancies between measured and simulated data.

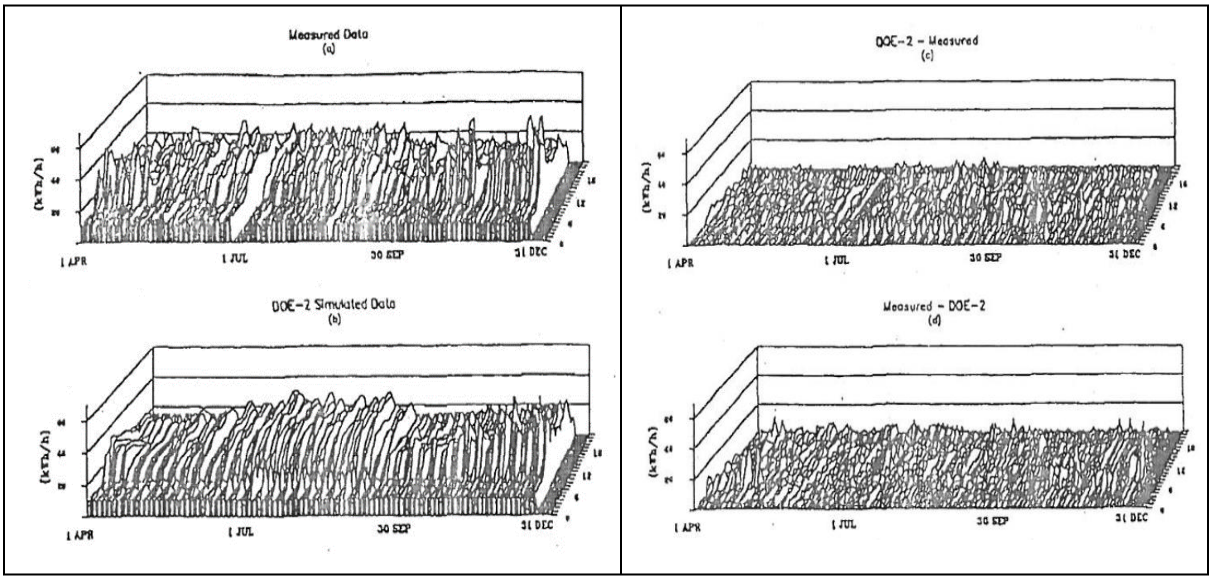


Figure 2-3: Comparative three-dimensional plots (Month, Hour, kWh/h) showing (a) measured data, (b) simulated data (DOE-2), (c) simulated (DOE-2) - measured data, (d) measured – simulated (DOE-2) data (Haberl and Bou-Saada 1998)

McCray et al. (1995) propose another graphical method to calibrate a DOE2.1 model to one year of 15-minute interval data for whole-building energy use. The Visual Data Analysis (VDA) method allows the modeler to quickly review the simulation results and make iterative changes to the models.

A number of later studies focused on further developing this approach by means of visual comparitive displays (Bronson et al. 1992; Haberl et al. 1993; Bou-Saada and Haberl 1995; Haberl et al. 1996; Haberl and Abbas 1998a; Haberl and Bou-Saada 1998; Haberl and Abbas 1998b).

Christensen (Christensen 1984) originally proposed the use of colour contour plots (or Energy Maps, EMAPS) to help display hourly data from a commercial building. Haberl et al. (1996) adopted this technique in developing graphical compartive displays with time-sequenced contour plots. Raftery and Keane (2011) proposed the use of carpet contour plots (see Figure 2-4) as a means of speeding up the identification of major discrepancies between modelled and simulated data as well as a useful tool for fault detection.

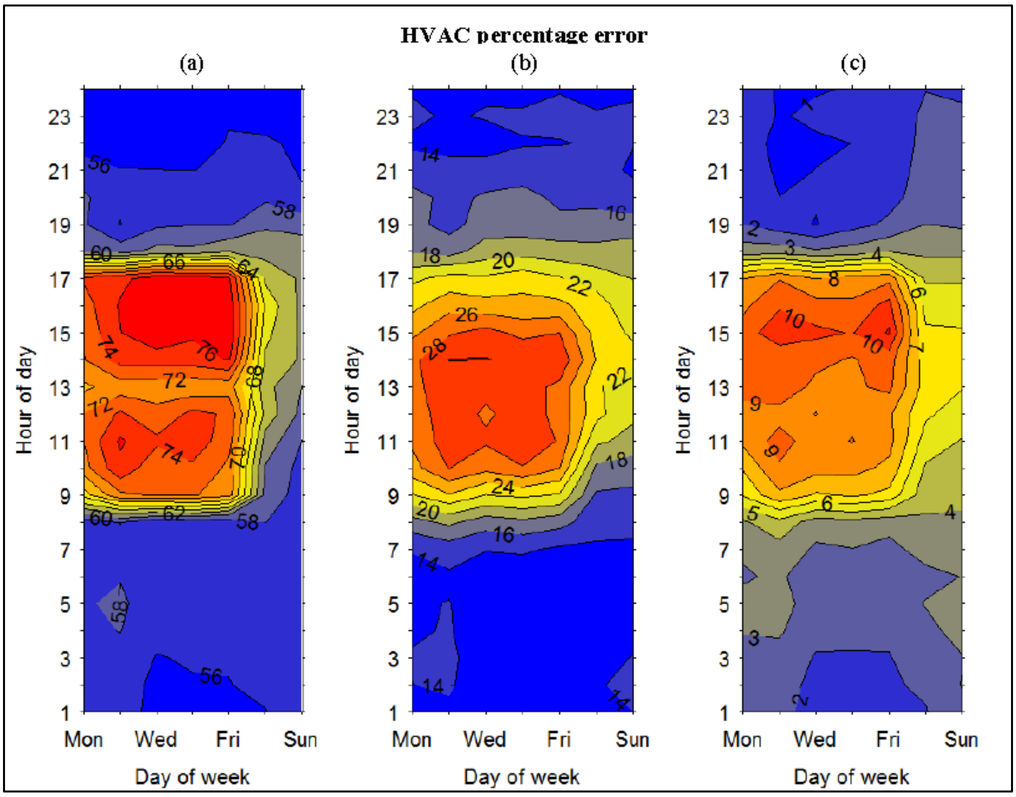


Figure 2-4: Mean percentage error between measured and simulated HVAC electricity consumption against hour of day and day of week for 2007. From left to right: a) model uses lighting and plug load schedules for a typical office, b) model updated to use measured lighting and plug load data in the model, c) the final model. (Raftery and Keane 2011)

2.5.2.2 Graphical Statistical Indices

Graphical Statistical Indices refer to the graphical representation of statistical indices through the use of graphical techniques. One such approach is binned box-whisker mean plots (Abbas 1993;

Bou-Saada and Haberl 1995) which display maximum, minimum, mean, median, 10th, 25th, 75th and 90th percentile points for each data bin given a period of data. These plots eliminate data overlap and allow for more informative statistical characterisation of the dense cloud of data points. The authors also proposed the use of temperature bin analysis, 24-hour weather day-type analysis and 52-week bin analysis. Further examples can be found in a number of more recent case studies illustrating the importance of effectively conveying statistical information behind calibration studies (Yoon et al. 2003; Wilde and Tian 2009; Raftery and Keane 2011; Raftery, Keane and Costa 2011; Wang et al. 2012).

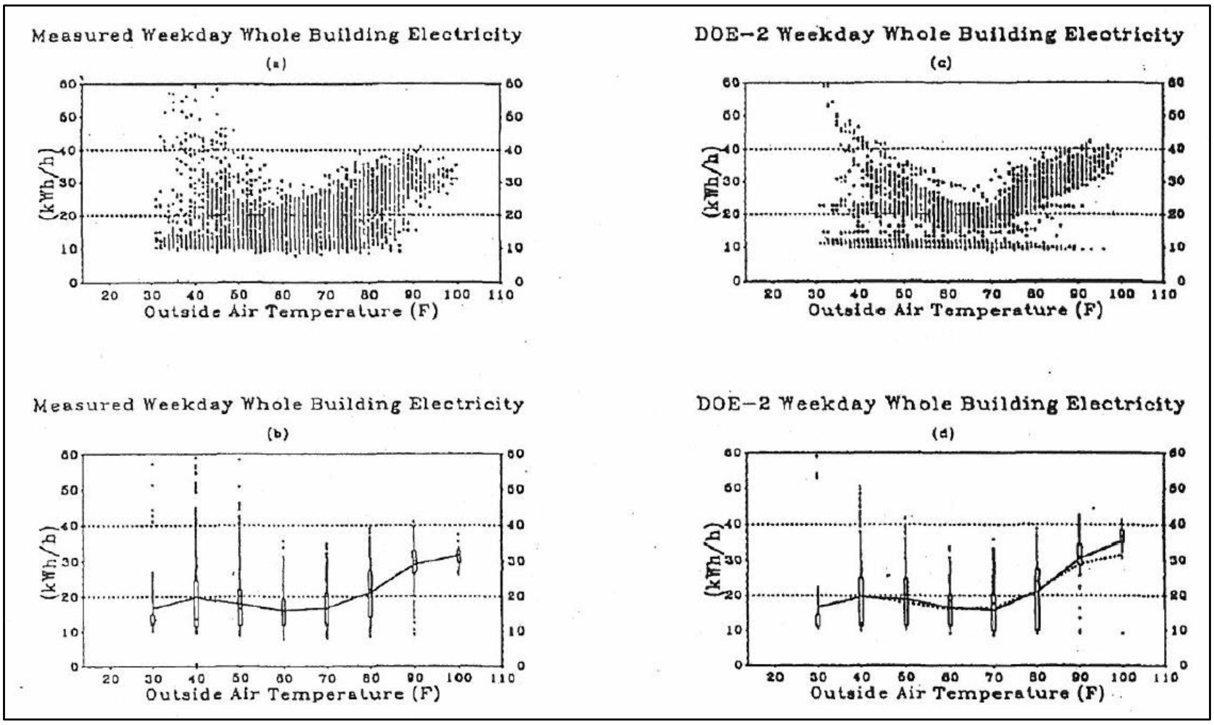


Figure 2-5: Weekday temperature bin calibration plots. The figure shows the measured and simulated hourly weekday data as scatter plots against temperature in the upper plots and as binned box-whisker-mean plots in the lower plots (Haberl and Bou-Saada 1998)

2.5.2.3 Signature Analysis

One of the major issues in tackling building energy calibration is the issue of accurately modelling heating and cooling energy consumption. Katipamula and Claridge (1993) proposed an approach for developing simplified system models for retrofit analysis, based on the work of Knebel (1983) on the Simplified Energy Analysis Procedure (SEAP) (see Section 2.6.2.4). This was later extended to account for calibration and devlopment of optimised control strategies (Liu and Claridge 1998). Based on this work, a process was devloped to include graphical signatures of heating and cooling energy consumption (Wei et al. 1998; Liu et al. 1998; Yu and Chan 2005). These graphic signatures would allow simulation engineers to identify the impacts of different input parameters (weather, occupancy, outside air intake, system type etc.) on an AHU's heating and cooling energy consumption. In addition, the technique may be used by commissioning engineers to identify faulty parameters, and develop optimised operation and control schedules.

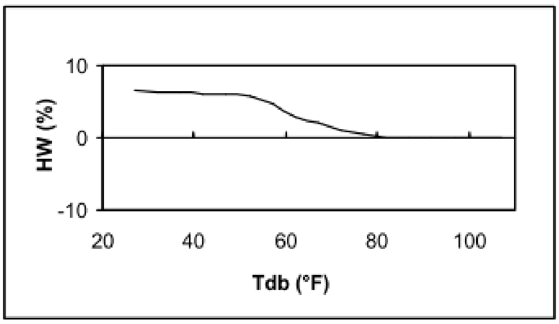


Figure 2-6: Sample Hot-Water (HW) energy use calibration signature, showing HW(%) vs. outside dry-bulb temperature (Tdb) (Liu et al. 2003)

Liu et al. (2003) propose a step-by-step procedure for the manual calibration of simulation models, based on the definition of two characteristic signatures:

- Calibration Signature: normalized plot of the difference between measured energy consumption values and the corresponding simulated values as a function of outdoor air temperature. For a given system type and climate, the graph of this difference has a characteristic shape that depends on the reason for the difference.

$$Calibration\ Signature = \frac{-Residual}{Maximum\ measured\ energy}\ x\ 100\% \tag{2.1}$$

- Characteristic Signature: By simulating the building with one value for an input parameter (the "baseline" run), then changing that input parameter by a given amount and rerunning the simulation, the "residuals" between these two simulations can be calculated, normalized, and plotted versus outdoor air temperature, producing a characteristic signature. By matching the observed signature with the published characteristic signature, the analyst is given clues to the factors that may be contributing to the errors he or she is observing.

$$Characteristic\ Signature = \frac{Change\ in\ energy\ consumption}{Maximum\ energy\ consumption}\ x\ 100\% \tag{2.2}$$

The study provides charateristic signatures for a number of standard systems including single-duct constant-air-volume (SDCV), single-duct variable-air-volume (SDVAV), dual-duct constant-air-volume (DDCV) and dual-duct variable-air-volume (DDVAV) air handling Units (AHUs) and three representative climates in California: Pasadena, Sacramento and Oakland.

G. Liu and Mingsheng Liu (2011) provide a rapid two-stage calibration procedure for simplified energy models (Refer 2.6.2.4), based on the use of calibration signatures. A simplified model of a high-rise office building is developed and calibrated to two weeks worth of measured data. This model is then used to simulate the hourly heating and cooling energy consumption for the building. Calibration signatures are then used to compared measured and simulated data in order to give an indication of which parameters should be changed and the corresonding magnitude of change required. A second stage of calibration requires the fine-tuning of these parameters to obtain a better overall fit of the model to measured data. Comparison of the results of this simulation with the measured data gave monthly CV(RMSE) values of 10.3% and 3.7%, and NMBE values of 2.2% and 1.4%, for cooling and heating respectively. These are within the ASHRAE criteria for model calibration (ASHRAE 2002). However, hourly comparisons gave monthly CV(RMSE) values of 31.0% and 28.0%, and NMBE values of 1% and 0.6%, for cooling and heating respectively. Even with detailed field measurements, these results for hourly calibration are outside the tolerances specified by ASHRAE Guideline 14-2002.

The authors conclude that this is a simplified example, but serves to highlight a number of issues with the calibration process. Firstly, this type of parameter tuning is typical of the general approach to model calibration, and while it may serve to produce a model which demonstrates sufficient overall accuracy when compared to measured data, it is probably not a good representation of the actual building being analysed. It is also highly dependent on analyst knowledge and skill, data availability, and allowed time-frame. The authors also point out that the satisfaction of hourly ASHRAE calibration criteria is quite difficult, even when high levels of measured data are available. It is also questionable as to whether it is even useful (or appropriate) to fine-tune a model to a very high degree of accuracy when employing generalised model assumptions and typical operation profiles.

In summary, advanced graphical techniques are an essential part of the calibration toolkit, but are not sufficiently robust to be used in isolation. Therefore, it is necessary to use a selection of tools instead.

2.5.3 Model Simplification Techniques

The methods described below rely on some form of model simplification to reduce simulation complexity or calibration requirements.

2.5.3.1 Base-Case Modelling

This process relies on the key concept of a detailed base-load energy consumption determination (Yoon and Lee 1999; Yoon et al. 2003) using the swing-season base load analysis recommendation by Lyberg (1987), where the term 'base load' refers to the minimum (or weather-independent) electricity or gas consumption. The swing-season calibration then fine-tunes simulation inputs when heating and cooling loads are minimal and building behaviour is dominated by internal loads. This provides a unique way to calibrate non-weather dependent data. Yoon et al. (2003) illustrate how this step-wise base-case modelling and swing-case calibration has been applied to a 83,212m^2 commercial building in Seoul, Korea. The final simulation gave an annual NMBE of 2.3% and CV (RMSE) of 3.6% for monthly data.

2.5.3.2 Model Parameter Estimation

The process of macro-parameter estimation refers to the process of deducing overall values for aggregated individual building parameters using non-intrusive monitored data, as opposed to intrusive tests described in section 2.5.1.4. Reddy et al. (1999) propose the use of an inverse method for estimating building and ventilation parameters through non-intrusive monitoring of heating and cooling energy use in large commercial buildings. As discussed in a later review paper by Reddy (2006), the procedure involves deducing the loads of an ideal one-zone building from the monitored data and then, in the framework of a mechanistic macro-model, using a multistep linear regression approach to determine the regression coefficients (along with their standard errors), which can finally be translated into estimates of the physical parameters. This procedure is applied to two different building geometries at two different climatic locations, to estimate six physical parameter values, including the overall building heat loss coefficient. The approach has been found to yield very accurate results (regression R^2 coefficients of 0.97-0.99), particularly when combined with daily data over an entire year.

2.5.3.3 Parameter Reduction (Day-Typing and Zone-Typing)

The process of parameter reduction or simplification relies on the statistical characterisation of complex inputs in order to reduce the number of inputs in a model. One approach which has been used extensively is day-typing, in which building energy use is characterised on a daily profile, rather than on an hourly basis. This approach allows for the definition of typical days (e.g. weekdays, weekends, and holidays) which can be used to characterise building energy use, thus condensing a large quantity of complex measured building data into relatively few input points or schedules.

Kaplan et al. (1990; 1990) use day-typing to group days with reasonable uniform non-HVAC load shapes. Zone-Typing (i.e. grouping similar zones) is used to further apply these day-types across multiple zones. Bronson et al. (1992) uses day-typing routines (for occupancy and equipment scheduling) to calibrate a DOE-2 simulation model. Hadley (1993) uses a combination of principal component analysis and cluster analysis to identify distinctive weather day types (which represent repeatable weather conditions that typically occur at each site) from one year of National Weather Service (NWS) station data. HVAC system energy consumption data for each day are then grouped by these weather day types, and daily total and hourly load profiles were developed for each day type.

Raftery et al. (2011; 2011) incorporate zone-typing to separate thermal zones in such a way as to minimise inaccuracies incurred by representing multiple actual thermal zones in a building with a single large zone in the model. This is achieved by assigning thermal zones in the model based on four major criteria (see Figure 2-7): (1) space function, (2) position relative to exterior, (3) available measured data, and (4) space conditioning method.

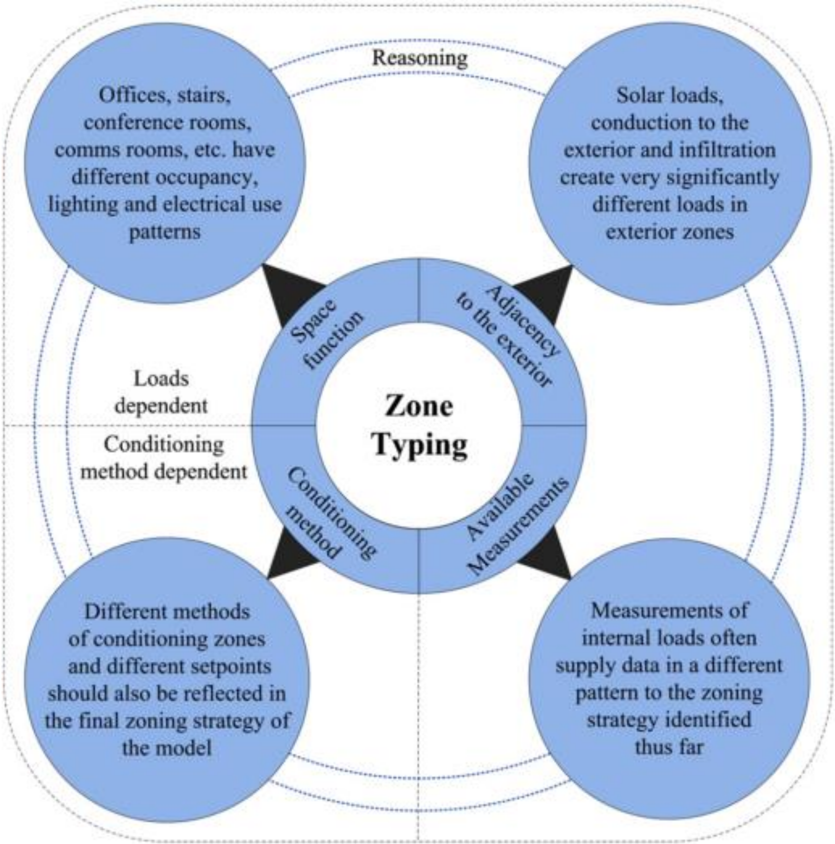


Figure 2-7: Zone Typing (Raftery, Keane, O'Donnell, et al. 2011)

2.5.3.4 Data disaggregation

Disaggregation is the splitting up of the total building energy consumption into its component parts. There are a number of reasons as to why this is done, i.e., to focus on specific energy flows and identify areas for retrofit and conservation. Lyberg (1987) proposes data disaggregation as part of a staged audit process as a means of focussing attention on high-importance areas. This can help limit subsequent auditing to the areas where the most productive retrofits could be carried out. This step will directly assist in the identification of energy-conservation opportunities (ECO's).

Akbari (1988; 1995) developed an algorithm to disaggregate short-interval (hourly) whole building electrical load into major end-uses. The End-Use Disaggregation (EDA) algorithm utilises statistical characteristics of measured hourly, whole-building load and its inferred dependence on temperature to produce hourly load profiles for air-conditioning, lighting, fans, pumps and miscellaneous loads. Regression models are developed for each hour of the day for major day types (see 2.5.3.3) between measured building energy use and outdoor dry-bulb temperature. Since the temperature dependency of the building may change with season, the author suggests using two season specific (summer and winter) sets of temperature regression coefficients. The regression constant for these models are assumed to provide an indication of the weather-independent energy use, while the slope represents weather-dependent behaviour. Since the regression models provide no information about the breakdown of the temperature-independent load, it is simply pro-rated against loads predicted by simulation as well as on-site measurements. The approach is applied to numerous retail and commercial facilities. (Akbari et al. 1988; Akbari 1995; Akbari and Konopacki 1998). The authors conclude that this is a useful approach for buildings in which the whole-building temperature dependent load is primarily due to the HVAC system (i.e. only the HVAC load is sensitive to outdoor temperature). This assumption may be applied to large offices and commercial buildings, but not to buildings characterised by non-HVAC end-uses such as refrigeration (which is weather dependent).

2.5.4 Procedural Extensions

The following section describes procedural tools and techniques used to assist in improving the overall calibration process.

2.5.4.1 Improved Statistical Comparisons

In the early years of building simulation, simple precent difference calculations had been the primary means of comparing measured and simulated data (Diamond and Hunn 1981; Kaplan, McFerran, et al. 1990; Bronson et al. 1992). However, as noted by Diamond & Hunn (1981) this often led to a compensation effect, whereby over-estimations cancelled out under-estimations. Bou-Saada and Haberl (1995) proposed the adoption of standardised statistical indices which better represent the performance of a model (J. Kreider and Haberl 1994; J. F. Kreider and Haberl 1994; Bou-Saada and Haberl 1995):

- Mean Bias Error (MBE) (%): This is a non-dimensional bias measure (i.e. sum of errors), between measured and simulated data for each hour.

$$MBE(\%) = \frac{\sum_{i=1}^{N_p}(m_i - s_i)}{\sum_{i=1}^{N_p}(m_i)} \tag{2.3}$$

Where; m_i and s_i are the respective measured and simulated data points for each model instance '*i*' and N_p is the number of data points at interval '*p*' (i.e. $N_{monthly}$=12, N_{hourly}=8760).

- Root Mean Square Error (RMSE) (%): The root mean square error is a measure of the variability of the data. For every hour, the error, or difference in paired data points is calculated and squared. The sum of squares errors (SSE) are then added for each month and for the total periods and divided by their respective number of points yielding the mean squared error (MSE); whether for each month or the total period. A square root of the result is then reported as the root mean squared error (RMSE).
- Coefficient of Variation of Root Mean Square Error CV(RMSE) (%): This is essentially the root mean squared error divided by the measured mean of the data. CV(RMSE) allows one to determine how well a model fits the data; the lower the CV(RMSE), the better the calibration.

$$CV\ RMSE(\%) = \frac{\sqrt{\frac{\sum_{i=1}^{N_p}(m_i - s_i)^2}{N_p}}}{\bar{m}} \tag{2.4}$$

Where; m_i and s_i are the respective measured and simulated data points for each model instance '*i*'; N_p is the number of data points at interval '*p*' (i.e. $N_{monthly}$=12, N_{hourly}=8760) and $\bar{m}$ is the average of the measured data points.

2.5.4.2 Evidence-based development

Manual approaches to model calibration generally rely on manual pragmatic user intervention to 'fine-tune' individual parameters to achieve a calibrated solution. However, these changes are often not tracked or recorded, and are rarely reported. This results in a situation whereby the calibration process relies heavily on user knowledge, past experience, statistical expertise, engineering judgement, and an abundance of trial and error (Bou-Saada and Haberl 1995). In order to improve the reliability and reproducibility of the calibration process it is necessary to keep a history of the decisions made along with the evidence on which these decisions were based (Bou-Saada and Haberl 1995; Raftery, Keane, O'Donnell, et al. 2011). This allows future users to review the entire calibration process and the evidence on which the model is based. In addition, changes to the input parameters should only be made according to available evidence and clearly defined priorities (Raftery, Keane, O'Donnell, et al. 2011).

A number of studies incorporate systematic evidence-based model development at the core of the calibration process (Bou-Saada and Haberl 1995; Yoon and Lee 1999; Yoon et al. 2003; Monfet et al. 2009; Parker et al. 2012)

2.5.4.3 Sensitivity Analysis (SA)

Sensitivity analysis has been employed in recent calibration efforts to identify parameters of greatest influence on energy end-use in a building. There are a number of available techniques available for conducting sensitivity analyses, depending on the particular requirements and application (e.g. single vs. multiple parameters). For detailed descriptions of tools and techniques, refer in particular to the work of Saltelli et al. (2002; 2005) on this particular subject.

Clarke et al. (1993) used two sensitivity analysis techniques to determine uncertainty bands associated with ESP-r predictions. Differential Sensitivity Analysis (DSA) was used to determine total uncertainty band as the root mean squared summation of individual uncertainties due to each input parameter. Monte-Carlo Sensitivity Analysis (MCSA) was used to determine the total uncertainty band by perturbing all the input parameters simultaneously. These sensitivity methods have been incorporated into ESP-r simulation software (ESRU 1974; Macdonald et al. 1999) for the purpose of uncertainty analysis.

Westphal and Lamberts (Westphal and Lamberts 2005) present a calibration study of a 26,264m^2 public office building, combining a building energy audit, model sensitivity analysis and manual tuning of influential parameters. The study concludes with an electricity consumption prediction within 1% of the measured values within four iterations of the base case model.

2.5.4.4 Uncertainty Quantification

As identified by Carroll and Hitchcock (1993), there may exist multiple solutions which may produce good overall agreement with measured data even though individual parameters are incorrectly defined. Hence, if using these inputs to infer any sort of meaning (e.g. for ECM analysis), it is important to account for uncertainty in these inputs. Reddy (2006), states that uncertainties in building simulation generally arise from four main sources:

i. Improper input parameters;

ii. Improper model assumptions;

iii. Lack of robust and accurate numerical algorithms;

iv. Error in writing simulation code.

While sources (ii)-(iv) deal directly with the simulation program and internal algorithms and assumptions, source (i) depends on the accuracy (and uncertainty) of the available input information. Since the validation of model algorithms is covered extensively in other studies (Neymark et al. 2002; Henninger et al. 2004; Ryan and Sanquist 2012), this review will focus on contributions to the identification of error and uncertainty in model input parameters, and how this has been applied to model calibration.

As discussed in the previous section 2.5.4.3, Clarke et al. (1993) used sensitivity analysis to determine uncertainty bands associated with ESP-r predictions of internal air-temperature in his PASSYS test-cell experiments. In this case, uncertainty bands were quite narrow, reflecting the level of control of the experiments in terms of ESP-r input parameters. It was shown, however, that uncertainty bands were largely temperature-dependent, due primarily to the uncertainty in conservatory air temperature prediction. This was due to instrument accuracy for solar radiation measurement (varying by as much as ±3%).

Lomas and Eppel (Lomas and Eppel 1992) discuss the application of three sensitivity analysis techniques (DSA, MCSA, SSA) to determining the relative sensitivities, in both hourly and daily average model predictions (using ESP-r, HTB2 and SERI-RES), due to the uncertainties in over 70 input parameters. Lomas et al. (1997) conducted an extensive review of dynamic thermal simulation programs (DSPs) comparing measurements with predictions and accounting for experimental uncertainty. The authors state that total model uncertainty has two components: (1) measurement errors, as above – which are easy to identify; and (2) uncertainties in program input data – which is more difficult to calculate. This difficulty is due to the large number of inputs which require quantification of associated uncertainty, as well as the propagation of this uncertainty through the DSP to determine the overall prediction uncertainty.

De Witt and Augenbroe (2002) address uncertainties in building performance evaluations and their potential on design decisions. The authors examine uncertainties in material properties as well as those stemming from model simplifications. They suggest a statistical screening technique (using Monte-Carlo Analysis) to determine which sources have dominant effects on the outcome of the simulation. The procedure is illustrated for a simple building envelope and considered parameters such as wind speed, indoor air distribution, and envelope material and heat transfer coefficients.

Reddy et al. (2007a; 2007b) identified the necessity for uncertainty analysis, which had been over-looked in many calibration studies, particularly in ECM analysis applications. In this work, uncertainty is addressed by assigning ranges of variation to influential input parameters and a

Latin-hypercube Monte-Carlo (LHMC) simulation is carried out to produce multiple possible solutions. The author selects the top 20 solutions, rather than selecting a single solution, to produce a range of values for the predicted performance of ECM's (rather than a single value). Overall, the authors found the relative uncertainty (or fractional difference) between actual and predicted values to be in the range of 25-50%. However, in most cases, the actual savings are usually contained in the range predicted. In conclusion, the authors suggest that one should not rely on calibrated simulations which predict savings of less than 10% (as associated uncertainty could account for up to 50% of this value).

2.6 Summary of Automated Calibration Developments

The following summarises the major developments in automated calibration of building energy performance simulation models over the last three decades.

2.6.1 Optimisation Techniques

2.6.1.1 Objective Function

The first automated calibration technique called RESEM (Renewable Energy Savings Estimation Method) was used for evaluation of ECM's using pre-retrofit and post-retrofit data. (Carroll et al. 1989; Hitchcock et al. 1991; Pal et al. 2002). The tool is based on a previously developed set of *knowledge-based expert rules* designed to bridge simulation models with measured utility bill data (Lebot 1987). RESEM (Retrofit Energy Savings Estimation Model) uses a self-contained energy simulation program similar to DOE-2, called RESegy. The goal of the project was to provide a simple cost-effective solution for ECM analysis by staff with little or no energy simulation expertise. As such, it relied on a database of expert knowledge for the development of building prototypes and parameter defaults based on minimal information from the user. The tool was benchmarked against DOE-2 using a simple base-case building. Comparisons of monthly heating and cooling loads (including peak loads) as well as electrical and gas energy consumption, as computed by DOE-2.1E and RESEM, were performed.

Lavigne (2009) implemented a similar DOE-2 based assisted calibration process using built-in engineering rules as well as optimization algorithms based on a Maquardt-Levenberg non-linear least squares method. Two real case studies are presented and calibrated to monthly utility bill data by tuning a set of user-defined parameters until acceptable limits are reached. In the presented case studies, this was acheieved in 2-3 iterations, achieving a monthly and annual difference in measured and simulated energy consumption of 10.9% and -1.1% respectively.

2.6.1.2 Penalty Function

Based on their original experience with RESEM, Carroll and Hitchcock (1993) introduced a more generic approach to systematically adjusting ("tuning") the parameters of a simulatable building description in order to match simulated performance to metered utility data. The underlying method is based on the minimisation of differential terms between measured and simulated data. In addition, the approach incorporates a weighting function to describe the relative importance of any single term within the minimisation function, thus maintaining reasonable parameter values during the calibration process. The paper also addresses two other important issues:

- Existence – it may not be possible to find an exact match between measured and simulated performance (i.e. the simulation model does not represent exactly what happens in the real building). Therefore, rather than identify an exact solution, the authors suggest finding a minimum quantity based on the normalized difference between predicted and actual consumption.
- Uniqueness – there may be many solutions which match the defined minimization criteria. This can be addressed by providing additional matching constraints in the minimization function, thus reducing the number of possible solutions. The authors suggest the use of a penalty function term which increases quadratically with the difference between each input parameter and its corresponding preferred value.

The approach utilises a prototype building generator to assist in the creation of the initial building model. The tuning process relies on some knowledge of the building to decide on parameters for adjustment, based on associated uncertainties classified during a building audit.

More recently, a methodology has been developed for the systematic calibration of energy models that includes both parameter estimation and determination of uncertainty in the calibration simulation (Sun and Reddy 2006; Reddy et al. 2007a). Based on the building type, the user must heuristically define a set of influential parameters and schedules which correspond to defined input parameters in the building model. These parameters are then assigned 'best-guess estimates' and 'ranges of variation' in order to generate an uncertainty-based search space. A coarse search of this space is carried out using a Monte-Carlo (MC) simulation approach to identify strong and weak parameters. This is achieved by coupling a blind Latin-Hypercube Monte-Carlo (LHMC) search with a Regional Sensitivity Analysis (RSA). This allows the analyst to fix weak parameters and specify narrower bounds of variability for influential parameters to further refine the search space and the corresponding promising vector solutions. By adopting this multi-solution approach, predictions may be made about the effect of changes to a building (ECM analysis) while providing an associated uncertainty of these predictions. This approach is applied to three case study office buildings using the DOE-2 software and calibrating to monthly utility bills (Reddy et al. 2007b).

2.6.1.3 Bayesian Calibration

It is important to consider prediction and uncertainty analysis for systems which are approximated using complex mathematical models. Bayesian calibration methods (MacKay 1994; Kennedy and O'Hagan 2001) can be used to naturally incorporate these uncertainties in the calibration process, , including the remaining uncertainty over the fitted parameters. These uncertainties may be propagated through the model using probabilistic sensitivity analysis (Lomas and Eppel 1992). Bayesian calibration methods also attempt to correct for any inadequacy of the model which is revealed by a discrepancy between the observed data and the model predictions from even the best-fitting parameter values. In addition Bayesian methods have the ability to combine multiple sources of information at varying scales and reliabilities (Booth et al. 2013).

Kennedy and O'Hagan (2001) present a generic approach for the Bayesian calibration of computer models. The method is illustrated by using data from a nuclear radiation release at Tomsk-7 chemical plant, and from a more complex simulated nuclear accident exercise.

Booth et al. (2013) suggest a hierarchical framework in which a top-down (macro-level) statistical model is used to infer energy consumption for (micro-level) representative individual dwellings from publically available energy consumption statistics. In this approach, a Bayesian regression method is employed for the top-down statistical model in order to account for uncertainties in the macro-level data.

2.6.2 Alternative Modelling Techniques

2.6.2.1 Artificial Neural Networks (ANN)

While not strictly used to calibrate energy models, artificial neural networks (ANN) have been proposed as a prediction method for building energy consumption. Neto & Fiorelli (2008) compared the use of EnergyPlus and artificial neural networks (ANN) in simulating energy consumption for an administration building at the University of Sao Paulo, Brazil. The results showed that EnergyPlus consumption forecasts had an error range of ±13% for 80% of the tested database. The authors concluded that the major source of uncertainties in the detailed model predictions are related to proper evaluation of lighting, equipment and occupancy. An adequate evaluation of the coefficient of performance (COP) for the unitary air conditioners serving the space also plays a very significant role in the prediction of the energy consumption of a building. The ANN models, based on simple (temperature-only input) and complex (temperature/ relative humidity/solar radiation inputs) neural networks showed a fair agreement between measured and predicted energy consumption forecasts and actual values, with an average error of about 10%. While the ANN model required less manual input, it can only predict energy consumption based on past performance and therefore requires a large historic set of training data for adequate performance. Therefore, any operation changes or retrofit measures would require re-training using a new data set. Finally, the ANN model cannot provide the same insights as a detailed energy model as it is not based on physical input parameters. However, the authors conclude that there is merit in further investigating the potential for using ANN to improve methodologies for evaluation of energy consumption in air-conditioned buildings (e.g. as an substitute for complex schedule input in detailed energy models).

2.6.2.2 PSTAR

The PSTAR (Primary and Secondary Term Analysis and Re-normalisation) method, originally proposed by Subbarao et al. (1988) and later refined and extended by Burch et al (1990), and Balcomb et al (1993), utilises data from a short-term energy monitoring (STEM) test (Refer to 2.5.1.2). In this approach, adjustments are made to major energy flows rather than to individual input parameters. This is achieved by identifying all the heat flows relevant for the building using a three stage STEM testing procedure:

1. Steady-state heat loss during constant heat input (Night)
2. Thermal mass using cool-down test (Night)
3. Effective solar gain by analysing change in heating/cooling load (Day)

A re-normalisation procedure (using a linear least squares method) is used to define the primary flows and subsequently compute secondary term flows thus enabling the definition of a dynamic energy balance representation of the building system. The process of data analysis and calibration from a set of defined STEM data can be automated, and is reported to yield reasonable results. However, it is dependent upon measurement accuracy. Infiltration heat loss is the major source of uncertainty and may require continuous tracer gas measurements.

2.6.2.3 Meta-Modelling

Currently, building energy simulation models are primarily used at the building design stage, usually for the purpose of energy code compliance certification (e.g. LEED, BREEAM). As building energy models become more accurate and numerically efficient, model-based optimization of building design and operation is becoming more practical. This model-based optimisation generally requires the combination of a whole-building energy simulation model with an optimisation tool. However, this tends to be time consuming due to the simulation and analysis time required for each model iteration. It also often leads to suboptimal results because of the detail and physical complexity of the energy model.

Eisenhower et al. (2012) present an approach which aims to cut the complexity of the optimisation problem, by reducing the detailed simulation model to a simple mathematical meta-model. The method begins by sampling the parameter space of the building model around the baseline values. This is done by applying a uniform distribution and a corresponding range

(±20%) of the baseline parameter value, and then using quasi-Monte Carlo (deterministic) sampling approach to provide samples within this distribution. Numerous simulations (~3000) are performed using this sample data, and an analytical meta-model is then fit to the output data. Once this process is complete, optimization can be performed using different optimization cost functions or optimization algorithms with very little computational effort. Uncertainty and sensitivity analysis is also performed to identify the most influential parameters for the optimization. A case study is explored using an EnergyPlus model of an existing building which contains over 1000 parameters. When using a cost function that penalizes thermal comfort and energy, 45% annual energy reduction is achieved while simultaneously increasing thermal comfort by a factor of two.

Manfren (2013) proposes an approach for calibration and uncertainty analysis in building simulation models based on the use of 'grey-box' meta-modelling techniques, combining data-driven 'black-box' models with detailed law-driven 'white-box' simulation models. This approach is applied to a real case-study office building for the verification and control of energy saving measures results. In addition, the approach is used to create a validated building simulation model for design and operational optimisation. The proposed methodology employs three models to achieve this goal: (1) simple piece-wise regression model trained on real data, (2) a Gaussian process meta-model trained on computer simulation data and calibrated with respect to piece-wise regression data, and (3) a detailed simulation model directly fitted to real data. The authors propose the development of the 'black-box' Gaussian meta-model which allows performing optimisation, uncertainty and sensitivity analysis in an easier and more computationally efficient manner compared with the original 'white-box' simulation model, while maintaining comparable results. This meta-model is also used for calibrating the detailed model input variables with respect to normalised observed data (outputs). Since this approach uses computationally-efficient black-box models, it can be easily integrated with multivariate real measured data. It may also be extended to incorporate highly multivariate inputs and multiple outputs within a real-time simulation environment. The paper concludes that this approach of combining data-driven and law-driven procedures has the potential to increase the potential usefulness, transparency and applications of models for simulation-based design and optimisation of buildings.

2.6.2.4 Simplified Energy Analysis Procedure (SEAP)

In the early years of commercial building energy performance simulation, many solutions were quite complex and required specialists and main frame computers to run. Detailed physical models (e.g. DOE-2, EnergyPlus) also tend to be over-parameterized and can often require significant effort, experience and time to provide an accurate representation of the building. In response, simplification procedures were proposed to increase computational efficiency.

Turiel et al. (1984) also proposed a simplified method of commercial building energy analysis utilizing a database of previous DOE 2.1A simulations to predict the outcome of other simulations. This approach is applied to an office building with very accurate results for heating, cooling and total energy use.

Knebel (1983) proposed the Simplified Energy Analysis procedure (SEAP) in order to reduce model complexity and calibration effort. This simplification is achieved in several ways:

- The building is assumed to have only two zones (one core and one perimeter);
- Average daily data and steady-state models are used for simulation and analysis;
- One large air-handling unit (AHU) is substituted for numerous smaller ones for each zone. This is only done with similar types of AHUs.

This has been successfully applied to a number of campus and commercial buildings with great success (Katipamula and Claridge 1993; Liu and Claridge 1995; Liu and Claridge 1998; Liu et al. 2004). This approach has since been combined with the use of signature analysis techniques (see 2.5.2.3) to help minimise the expertise needed to calibrate such a model (Liu et al. 2003; Liu and Liu 2011).

2.6.2.5 Systems Identification

This technique refers to the process of constructing models based only on the observed behaviour of the system (outputs) and a set of external variables (inputs), instead of constructing a detailed model based on "first principles" of well-known physical variables. Systems identification is based on work first started by Goodwin and Payne (1977). However, the first systematic procedure using computation tools was developed by Lennart Ljung (1987). Typically, the objective is to build a so called "black box" or "grey box" model in situations where a very detailed model would be costly and overly complex. Systems identification methods are very effective when significant amounts of data are available, as is the case with modern IT systems and advanced HVAC controls. This approach also involves an iterative procedure aimed at finding the best fit solution for model inputs.

Liu and Henze (2005) applied system identification techniques to find best-tuned input settings of detailed building energy performance simulation models. This is based on a two-stage calibration process which aims to minimise the root-mean square error (RMSE) between real and simulated data. However, instead of manually adjusting the identified tuning parameters, optimisation algorithms are applied instead. Nielsen and Madsen (2006) present a grey-box approach for modelling the heat consumption in district heating systems. Their approach utilises theoretical based identification of an overall model structure, followed by data-based modelling which is used to identify details of the model.

2.7 Conclusions

Buildings represent complex systems with high levels of interdependence on many external sources. The design, analysis and optimisation of modern building systems may benefit greatly from the implementation of Building Energy Performance Simulation (BEPS) tools at all stages of the building life-cycle (BLC). However, studies have found discrepancies between modelled and measured energy use in many cases where BEPS has been used to model real buildings. This undermines confidence in building simulation tools and inhibits widespread adoption.

Calibration aims to minimise discrepancies between measured and simulated data. However, due to the sheer number of inputs required for detailed building energy simulation and the limited number of measured outputs, calibration will always remain an indeterminate problem which yields a non-unique solution. Numerous approaches to model calibration have been suggested employing various combinations of analytical and/or mathematical and statistical techniques. However, no consensus has been reached on standard calibration procedures and methods that can be used generically on a wide variety of buildings. In addition, many of the current approaches to model calibration rely heavily on user knowledge, past experience, statistical expertise, engineering judgement, and an abundance of trial and error. Furthermore, when a model is established as being calibrated, the author often does not reveal the techniques used, other than stating the final result.

2.7.1 Deficiencies in current approach to model calibration

In summary, the issues with calibrated simulation can be broken down into seven main areas, as previously mentioned in Table 2-3:

- **Standards**: the lack of a consensus standard on simulation calibration. There are guidelines which specify broad ranges of allowable error for building energy models. However, these are over-simplified, in that they do not account for issues such as input uncertainty / inaccuracy or the model fit to zone-level environmental data. In addition, there are no standard guidelines for model development, which leads to fragmentation of the practice of energy modelling;
- **Expense**: Due to the fragmentation of the energy modelling process, it tends to require significant effort for both model development as well as model calibration. There is no integrated standard tool-chains or file formats at present, and building data required for modelling is often unavailable. Therefore, significant expense can be incurred in building auditing, metering and model development;

- **Simplification**: One of the problems with detailed building energy simulation is the fact that they require thousands of inputs for model definition. In practice, many of these inputs are simply un-attainable or may not be practicably measureable. In addition, the data on which these models are validated is limited, generally confined to single measurements for whole building heat energy and electrical loads. Therefore, it is said that the calibration problem, as it relates to detailed models, is over-specified (i.e. too many inputs) and under-determined (i.e. too few validation points). This is a difficult problem to address, as it requires the simplification of detailed models while maintaining accuracy;

- **Inputs**: In any modelling environment, the quality of outputs are only as good as the inputs available (Garbage-in, Garbage-out). In the case of building energy modelling, the sheer number of inputs required makes it impossible to obtain accurate measurements for all parameters. In such cases, it is necessary to find ways of quantifying these parameters to a reasonable degree of accuracy without compromising model output quality;

- **Uncertainty**: Since, building energy modelling requires a degree of approximation and simplification, it is important to account for this when presenting model outputs. As shown in Section 3, there are many sources of uncertainty in building energy modelling. One of the primary sources of model uncertainty is parameter spcification uncertainty, which relates to the degree of uncertainty around each input parameter. This is often disregarded in BEPS calibration case studies, leading to questions over the accuracy of the model outputs;

- **Identification**: The calibration process, at present, can often be described as an ad-hoc procedure requiring numerous iterations of manual pragmatic user intervention based on knowledge or expert judgement. Generally, this procedure is not well defined, in that the analyst decides on model changes based on personal jusgement as opposed to quantifyable evidence. This is often difficult to define, though some studies in the literature have attempted to provide procedures for identifyaing and correcting calibration issues;

- **Automation**: With the level of manual pragmatic user intervention required during all steps of the calibration process, it is clear that any degree of automation would greatly aid this process. However, since many procedures require human knowledge or input, this can be difficult.

2.7.2 Addressing calibration deficiencies

Table 2-7 below highlights the main calibration techniques described in this review and how each provides a basis for addressing some of these issues (Coakley et al. 2014).

Table 2-7: Addressing model calibration issues

Category	Classification	Approach	Calibration Issues						
			Standards	Expense	Simplification	Inputs	Uncertainty	Identification	Automation
Manual	**Characterisation Techniques**	AUDIT	X			X			
		EXPERT			X	X			
		INT				X			
		HIGH				X			
		STEM				X			
	Advanced Graphical Methods	3D						X	
		SIG			X			X	
		STAT	X						
	Model Simplification Techniques	BASE				X			
		MPE			X	X			
		PARRED			X	X			
		DISSAG			X	X			
	Procedural Extensions	EVIDENCE	X			X		X	
		SA	X			X		X	
		UQ					X		
Automated	**Optimisation Techniques**	BAYES		X			X		X
		OBJECT	X						
	Alternative Modelling Techniques	ANN		X	X				X
		PSTAR			X				
		META		X	X				
		SEAP		X	X				
		SYS			X	X			

Based on the above extensive literature review, it is evident that the current approach to calibrating a model is at best based on an optimisation process used to identify multiple solutions within a parameter space identified from a knowledge-base of templates of influential parameters (Reddy et al. 2007a). At worst, it is based on an ad-hoc approach in which the analyst manually tunes the myriad of parameters until a solution is obtained.

Furthermore, there is a lack of accounting for uncertainty in many of the calibration studies to date. This is a fundamental problem which needs to be addressed in order to provide a robust solution which accounts for the fact that there are uncertainties associated with all model input parameters. This can only be addressed by accepting that the calibration problem may be satisfied by multiple solutions within the parameter uncertainty space. By improving the accuracy of parameter measurement and definition, it is possible to reduce this uncertainty, be never eliminate it completely. Therefore, in order to satisfy full transparency of scientific outputs, this uncertainty should also be conveyed in the end result of the calibration process.

The methodology proposed by Reddy et al. (2007a; 2007b) uses building and HVAC templates, and thus, does not account for the fact that each building is unique. In addition, it does not provide the analyst with a comparison between measured and simulated building performance at a detailed level (e.g., energy consumption at a sub-utilities level). Thus, it is not an acceptable approach if the aim is to obtain detailed results to drive the development of simulation tools and best practice modelling techniques. Despite its limitations, this study provides an excellent basis for further work on analytical optimisation of the calibration process.

2.8 Proposed Approach

Based on the review of existing scientific literature on model calibration as well as numerous case studies, an approach has been developed which aims to promote a formalised approach to BEPS model calibration. The proposed calibration methodology incorporates a number of the more effective techniques highlighted in the literature review (see Figure 2-8).

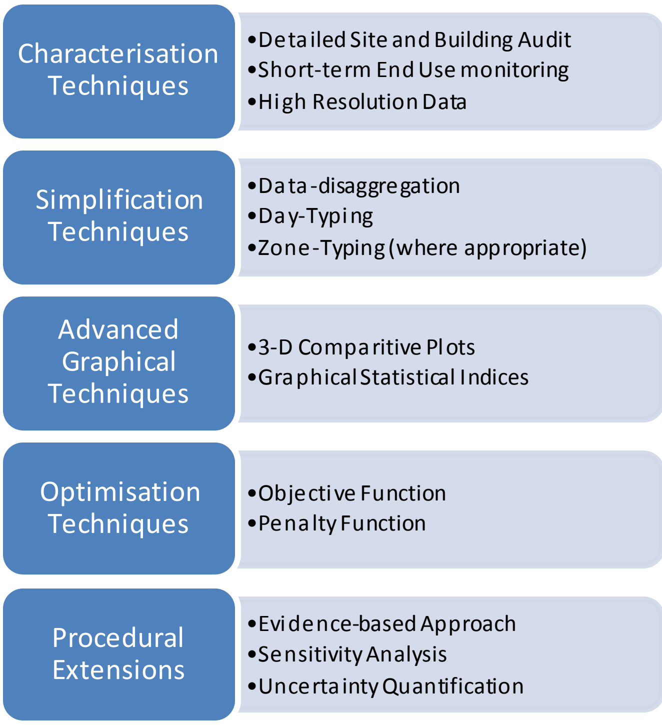


Figure 2-8: Techniques incorporated in proposed approach to model calibration

In terms of building and site characterisation techniques, the adoption of a detailed *investment grade audit* was deemed to be very important. A number of studies have illustrated that this not only provides the analyst with a greater degree of information and familiarity with the building (Waltz 1992), but also may reveal valuable energy conservation opportunities which may not be apparent from the building simulation (e.g. excessive lighting, staff habits and training etc.) (Raftery 2009; Shapiro 2009). If available, short-term end-use monitoring (STEM) can also provide a greater degree of granularity in measured end-use data, particularly in cases where sub-metering is not available or is not an option. The use of STEM allows for the disaggregation of end-use data (e.g. lighting, plug loads, major equipment), providing an invaluable input for detailed energy simulation tools in addition to detailed electrical equipment audits. It also provides information relating to end-use energy consumption profiles which may not be evident at a whole-building metering level. Finally, the use of high-resolution quality data has been identified as a necessity for reliable BEPS model calibration.

Clearly, any high-resolution model requires some simplifications in order to be manageable. Therfore, the use of day-typing and zone-typing present useful means of simplifying dynamic load schedules and complex thermal zone combinations.

Advanced graphical techniques are another useful tool in any approach to model calibration, allowing for quick intuitive visualisation of detailed building data and comparisons between measured and simulated data. Therefore, the use of 3-D comparitive plots (e.g. carpet plots) and graphical statistical indices willl be included in the proposed calibration methodology where deemed useful or necessary.

It must be understood that it will never be possible to accurately assemble all the information required to satisfy a law-driven model of any complex system. Therefore, assumptions, approximations and simplifications are required. However, while these approximations and simplifications are required, it is important that they are evident when presenting any results or recommendations. For this reason, the evidence-based calibration approach (Raftery, Keane, O'Donnell, et al. 2011) is deemed the most appropriate framework for carrying out any BEPS calibration study. Model inputs are explicity linked to source evidence and documentation, thus delivering a more transparent and reliable final product to the client.

It is important that any approach to model calibration also account for uncertainty. Therefore, each model input will be assigned a range of variation based on the uncertainty of the source evidence for that parameter. In the absence of a detailed knowledge base of parameter uncertainty in buildings, this is the best available approach. These uncertainties are propogated through the building energy model by sampling multiple inputs within the bands of uncertainty about the baseline value. The result of this process is the ability to deliver a model that can also represent risk in the context of performance predictions and decision-making.

Finally, it is clear from the literature review that the production of accurate building energy simulation models is a costly and time-consuming process. However, uncertainty weightings and optimisation approaches can be leveraged to reduce some of this time requirement. By quantifying acceptable model performance in terms of statistical goodness-of-fit as well as acceptable uncertainty, it is possible to identify approporiate solutions at a much earlier stage in the calibration process without unneccessarily exceeding the required criteria with negligible gains.

By adopting these techniques, the proposed methodology promises to address a number of the deficiencies outlined in Table 2-3 and Section 2.7.1:

- **Standards**: the proposed approach follows an iterative evidence-based methodology, minimising reliance on analyst knowledge and subjective interpretation. This is achieved by defining steps which need to be followed throughout the methodology, based on mathematical and statistical processes as opposed to subjective interpreattion where possible;
- **Expense**: the proposed methodology aims to reduce the time and expense required to calibrate detailed energy models by prescribing the steps that need to be followed at each stage, and promoting a statistical optimisation approach in combination with manual iterative model improvement;
- **Simplification**: the proposed methodology aims to provide the modeller with an insight into the important model input variables which influence model output. By concentrating on measurement and identification of these variables, less time is wasted on less important model inputs, which may not have a large influence on final model uncertainty. While it is not possible to simplify the physical processes being modelled, this procedure aims to simplify the identification of the important variables which define those physical processes;
- **Inputs**: inputs are assigned best-guess estimates based on information gathered from detailed audits, surveys or building documentation. For each source of model inputs, there is an associated *uncertainty*;
- **Uncertainty**: model input uncertainty is accounted for in the proposed methodology by assigning ranges of variation to each input parameter. This uncertainty is propogated through the calibration process, resulting in multiple solutions which capture the associated input uncertainty;
- **Identification**: the proposed approach uses a number of mathematical and graphical procedures to help identify sources of model error throughout the calibration process;
- **Automation**: it is currently not possible to fully automate the calibration of building energy models due to the unique nature of each building, and subsequent requirement

for manual iterative development at early stages of the model development process. Howeer, by incorporating mathematical and statistical methods, it is possible to improve the efficiency of the calibration procedure. This approach incorporates sensitivity analysis and mathematical optimisation to imporve calibration efficiency;

The proposed approach integrates a number of unique features into a streamlined process:

- Each model input requires the definition of source evidence and associated reliability;
- Parameter uncertainty is incorporated during model development through the definition of baseline values as well as standard deviations;
- Reliabilities are applied to source information in order to quantify this uncertainty;
- Sensitivity analysis is used to identify influential model parameters by leveraging associated uncertainty and Monte Carlo simulations;
- Multiple solutions are generated as opposed to a single solution in order to propagate and capture inherent model parameter uncertainties.

Chapter 3: Methodology

"It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience."

– Albert Einstein, Physicist

3.1 Introduction

This chapter describes the proposed analytical optimisation approach for the calibration of detailed Building Energy Simulation (BES) models. In order to avoid an over-reliance on analyst knowledge and judgement, this methodology follows a clear evidence-based structure and proven statistical methods. This can be broken down into the following steps:

1. Preparation;
2. Data Collection and Classification (incl. Uncertainty Quantification);
3. Evidence-based BES model development;
4. Sensitivity Analysis (Optional);
5. Iterative Model Improvement;
6. Latin Hypercube Monte-Carlo (LHMC) Sampling;

Figure 3-1 provides a generic overview of the entire process. A detailed description of each step is provided in more comprehensive detail in the following sections.

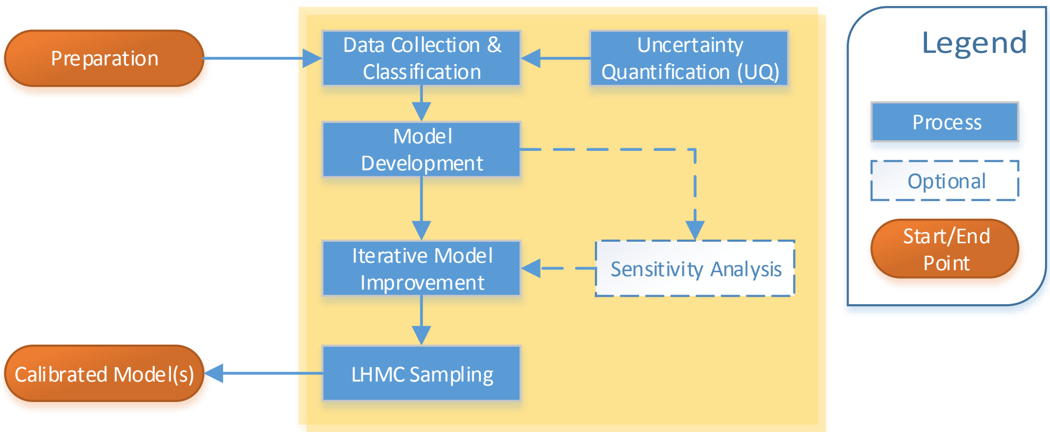


Figure 3-1: General Calibration Methodology Overview

3.2 Methodology Overview

This thesis presents a methodology pertaining to the calibration of detailed building energy simulation models to measured data. Data pertaining to the building construction, systems and operating schedules is collected and prepared for analysis. This data is used to develop an initial Building Energy Simulation (BES) model. This model is iteratively refined and updated using gathered information. Model evolution is tracked using version control software. This process continues until all relevant sources of building information have been fully utilised or incorporated in the model. Uncertainties are then applied to model parameters, based on the reliability of source information, and used to generate a set of random Monte-Carlo (MC) simulation trials. A sensitivity analysis may be carried out to determine the most influential parameters. The results of this analysis provide a basis for further parameter investigation and model refinement. Finally, a batch of MC simulation trials is performed and uncertainty analysis is carried out on the results of these trials.

The following sections provides a detailed overview of each element of the process:

- Section 3.3 (*Preparation*) details how initial preparations are carried out before beginning the calibration process.
- Section 3.4 (*Data Collection*) describes the data required for this methodology as well as how this data is stored, prepared and classified.
- Section 3.5 (*Evidence-Based BES Model Development*) outlines the recommended steps in carrying out an evidence-based BES model development.
- Section 3.6 (*Regional Sensitivity Analysis*) details the criteria for performing a sensitivity analysis on model inputs in order to determine the most influential parameters.
- Section 3.7 (*Iterative Model Improvement*) describes how the BES model is further refined by obtaining additional information for highly influential or uncertain parameters.
- Section 3.8 (*Latin-hypercube Monte-Carlo (LHMC) Search*) outlines the process of generating sets of random simulation trials within the model/parameter uncertainty space. The final models are then validated against the initial *acceptance criteria* and their inherent uncertainty is quantified.

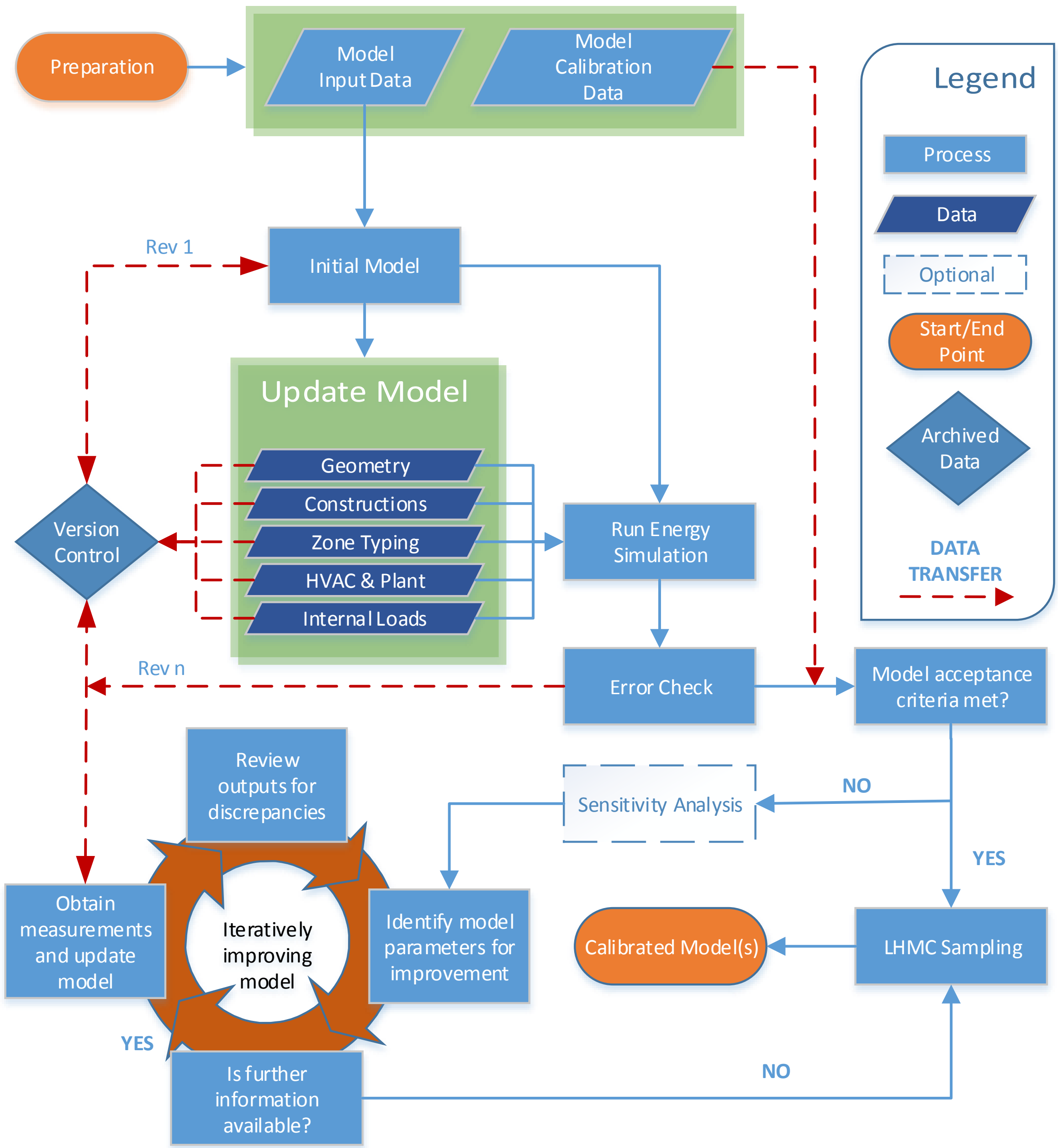


Figure 3-2: Detailed calibration methodology flowchart

3.3 Preparation

In any Building Energy Simulation project, preparation and organisation are key to the project's success. Following these preparation guidelines will enable a more logical workflow through the iterative simulation process and increase the quality and reliability of the results.

3.3.1 Modelling strategy

It is vital to consider the purpose of the model and available resources at the initial preparation stage as this will be a major influence on the chosen modelling strategy and acceptance criteria. For example, if the purpose of the model is to benchmark the building in terms of ideal performance against similar buildings, then a simplified modelling strategy may be adopted. In the case of a more comprehensive retrofit analysis, this may not be appropriate and more detailed survey and assessment will be required. When deciding on a modelling strategy, the following information should be considered:

- Purpose of the Model: What are the requirements for the model (i.e. benchmarking, ECM analysis, fault detection, commissioning etc.)
- Resources Available: What resources are available to commit to the project (time, finance, people, computation)
- Acceptance Criteria: At what stage is a model deemed acceptable. At present, models are considered 'calibrated' when they meet standard statistical acceptance criteria (ASHRAE 2002; EVO 2007). However, the client or analyst may choose to adopt different acceptance criteria depending on the purpose of the model. These criteria should be decided upon at the outset as they will be used to validate the model during development.

3.3.2 Source Control Management (SCM)

It is good practice to use version control software to track and manage the iterative changes throughout the calibration process. Changes are then identified by 'revision numbers' and can easily be identified at any future point in the process. This is useful for a number of reasons:

- Model Reliability – each change can be tracked to a particular point and associated comment or source of evidence
- Backup of Model and Revision history – model revisions can be managed on a remote location (e.g. secure server) to enable backup security.

- Multiple users – multiple people can work on the same model without damaging project integrity.
- Improvement tracking – changes to the model results can be explicitly linked to specific changes to the model inputs.

Virtually any version control software can be used, however, there are several essential capabilities that are required (Raftery 2009): (1) Maintain a change log or revision history of the entire project; (2) Allow for comment entry to the change log with each revision; (3) No limitations on file sizes or file types; (4) No limitation on total project size; (5) Allow nested folders (to facilitate evidence storage); (6) Remote access allowing multiple analysts to work on the same project simultaneously.

There are several tools available that enable source control management (SCM). The suitability of any particular model will depend on the requirements and resources of the project. There are a number of key properties of SCM systems which should be considered when choosing a suitable system for a project:

1. Repository Model: this describes the link between the working copy of a file and the source repository. There are two types of system, each with its' own advantages and disadvantages:
    - Client-server: A central (server) repository hosts the source project. Clients (users) must push/pull files from the server whenever modifications are made in order to propagate the most up-to-date version.
    - Distributed: Every user has a complete copy of the repository stored locally. Changes can be made to the local repository but there is no specified centralised/canonical copy of the repository.
2. Performance: the speed and reliability of the system
3. Access: user requirements in accessing and modify the repository
4. Size: the disk-space requirement for the repository. This may be of particular concern in large projects.

5. Branch Handling: Branching refers to the duplication of the source repository such that modifications can happen in parallel. For example, in the case where one wants to test the viability of the implementation of a particular modification to the building model (i.e. implementation of an ECM) without affecting the ongoing source model development. Branches can later be *merged* back into the 'central' repository.

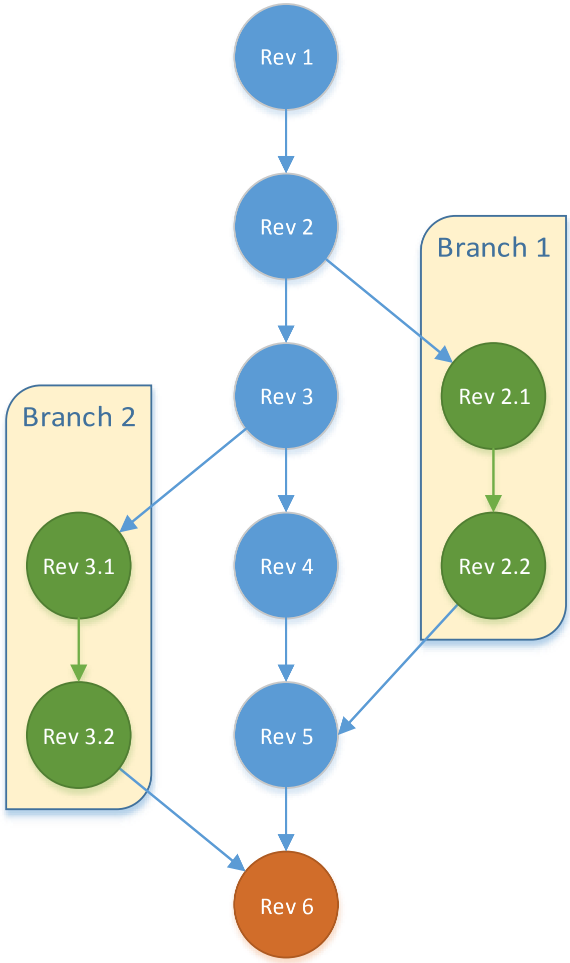


Figure 3-3: Branching in SCM Software

6. Concurrency Handling: This describes how modifications are managed by the version control software to prevent conflicts during simultaneous /concurrent editing:
    - Lock - users must check out or 'lock' a file for editing before they are allowed to make any changes
    - Merge – users may make changes but are informed of possible conflicts when committing their changes to the repository.
7. Platform: This refers to the supported operating systems.

Table 3-1: Comparison of features in SVN / GIT repositories

Software	SVN	GIT
Model	Client-Server Based	Distributed
Performance	Slower since network traffic is required to execute operations (Diff check, Commit, Merge etc.). Reliability may be an issue if server connection is unavailable.	Fast, since all operations (except push and fetch) are executed locally. Reliable since repository is stored locally an individual user machines, thus acting as a natural backup.
Access	Users require full commit access to the repository	Users do not need commit access to view/access a GIT repository. Repository owner controls the source which means users have version control of their own work
Size	Repositories can be quite large since an SVN working directory contains two copies of each file, one working copy and another hidden in .svn to aid SVN operations	Small - GIT working directories only contain one index file and single copies of all files
Branch Handling	Integrated in core functionality of GIT	Can be cumbersome and confusing
Concurrency Model	Merge or Lock	Merge
Platform	Unix, Windows, Mac OS	Unix, Windows, Mac OS

3.3.2.1 Source repository structure

Once a version control repository has been chosen and installed, the next step is to create and structure the source repository. This will vary from project to project. However, a suggested structure is provided in Figure 3-5 which is open to adaptation. A calibration toolkit (Figure 3-4), linked to a version control repository, has been proposed to streamline this process. This is currently under development (Refer to Appendix D.2).

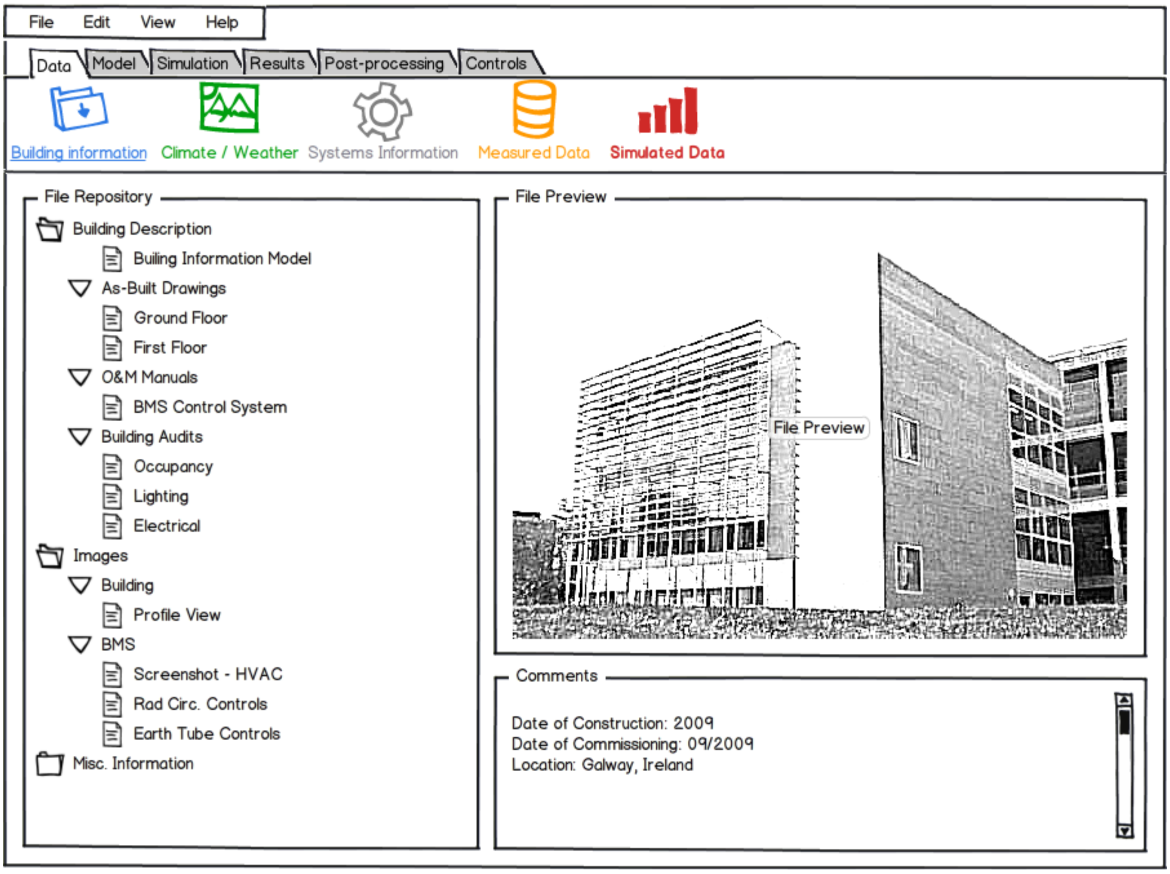


Figure 3-4: Calibration toolkit mock-up

This source repository should be used to store all files associated with the model calibration process, including input files (txt, idf, dwg, pdf, xls etc.) and output files (csv, xls, png etc.). All files are tracked with a version number once they are loaded into the file repository. Any modifications to these files will be tracked, and new version number issued. This applies to both text-based and binary file-types.

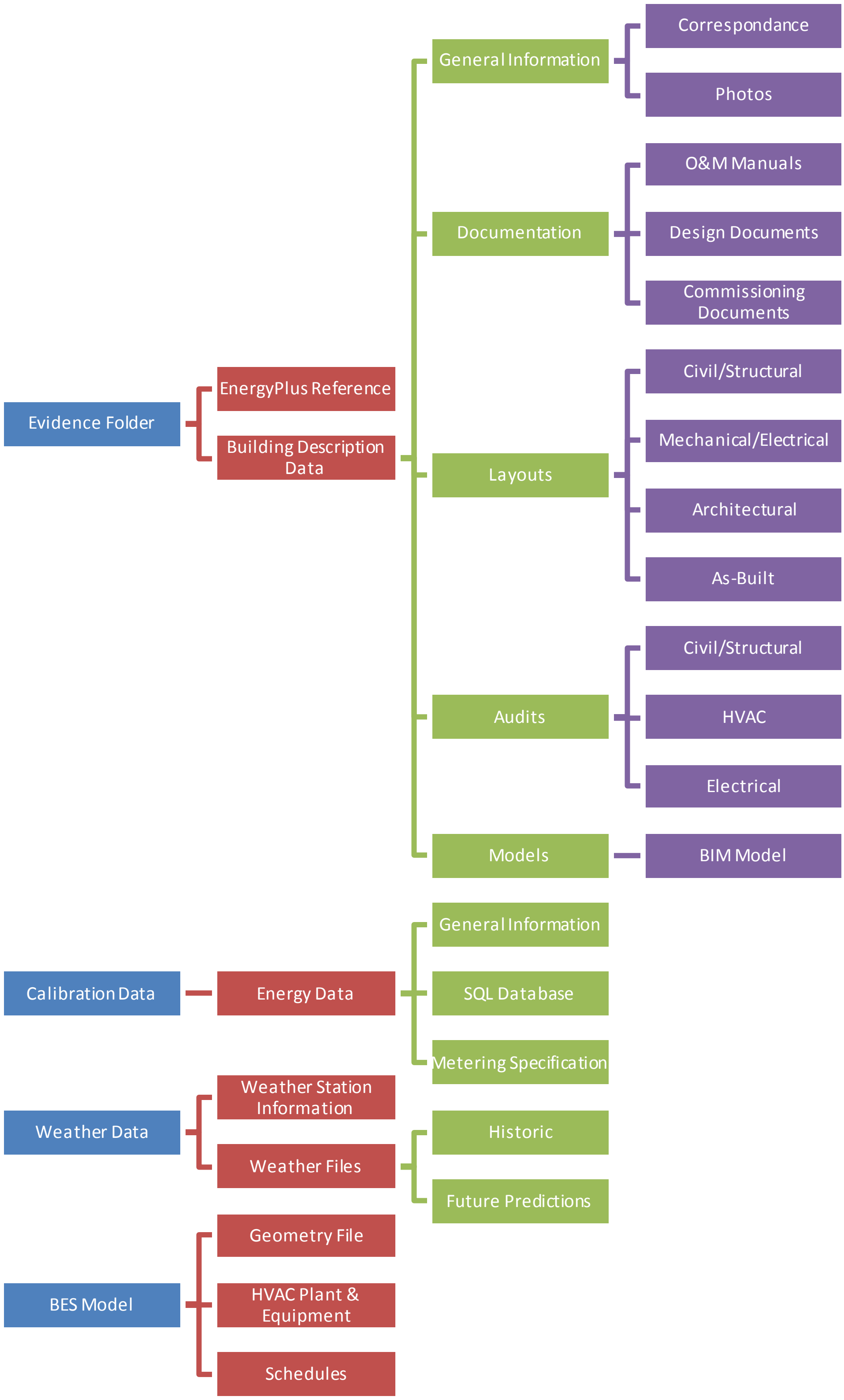


Figure 3-5: Source Repository Structure

3.3.2.2 Building Information Model

A Building Information Models (BIM) is a data-model for a particular building, which contains multi-disciplinary data (e.g. mechanical, electrical, structural data) specific to that building. It includes a specification for the relationships and inheritances for each of the properties of the building components it describes (walls, floors, roof etc.).

A BIM can provide a useful starting point for the creation of a building energy model. However, they can often contain redundant information which may add to the complexity of the final model, thus increasing run time or causing the model to crash if not amended. O'Donnell et al. (2013) provide a detailed discussion on transforming BIM to BEM, focussing on the generation of basic building geometry for the NASA Ames Sustainability Base. This study serves to highlight the complexity of the transformation process, even for a well-defined BIM model. For this reason, the scope of this research does not extend to the BIM to BEM transformation process, as this area of research is still at a relatively early stage, and requires significant effort and collaboration within industry to define the robust standards required to make the link a reality. Recent work on standardisation of BIM models has resulted in the formation of IFC[2] and gbXML[3] industry standards for building data representation.

It should be noted that an accurate BIM model, if available, may provide an invaluable source of detailed model information. However, the complexities involved in translating these models to building energy models means that it is often easier to recreate the BEM from scratch in order to avoid potential errors.

[2] http://www.buildingsmart.org/standards/ifc

[3] http://www.gbxml.org/

3.4 Data Collection

This section describes the process of collecting the input data required to initiate model development and calibration. The data required can be divided into

(1) Model Input data – this is the data required to specify input parameters within the simulation environment

(2) Model Calibration data – this is the data which will be used to calibrate and validate the results of the simulation model. This may also be referred to as the 'training data'.

In order to ensure the validity of this data, we carry out 'Data proofing and classification'. This step is described in further detail in Section 3.4.3

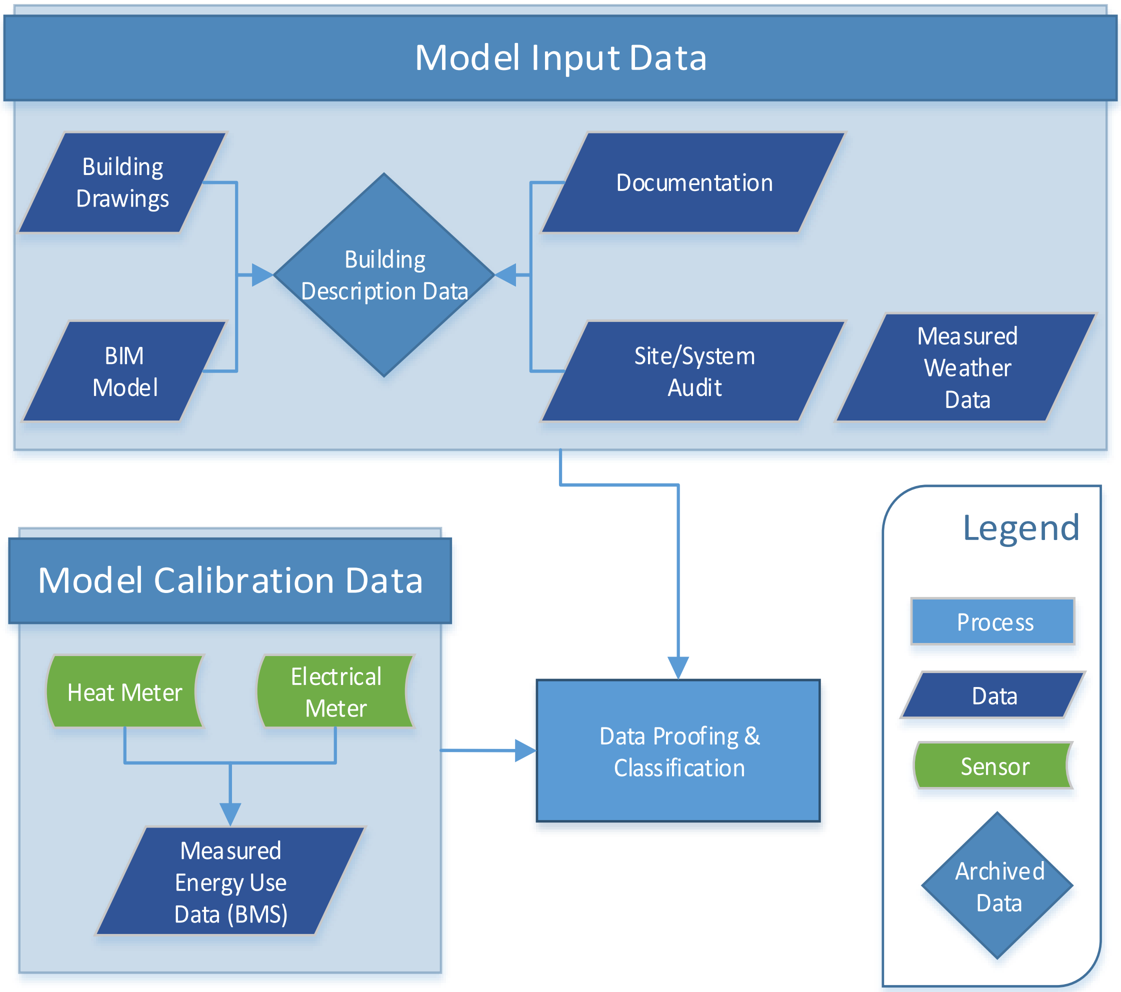


Figure 3-6: Data Collection for Model Calibration

3.4.1 Model Input Data

This section describes the process of collecting the input data required to initiate model development and calibration.

3.4.1.1 Building Description Data

In order to create an accurate representation of the chosen building within the BES software, it is necessary to first gather and record detailed information pertaining to the building, systems, environment and occupants. Sources of this information may include, but are not limited to, the following:

- Building Information Model (BIM);
- As-Built Drawings;
- Commissioning Documents;
- Operation & Maintenance (O&M) Manuals;
- Spot measurements (as recorded by clamp-on meters, hand-held sensors etc.)
- Site Survey;
    - Photographs;
    - Interviews (Building Manager, Staff, Occupants);
    - Nameplates (HVAC Systems, Sensors, Supply Systems);
    - Shading/Exposure;

3.4.1.2 Site Survey

The site survey is an important step as it will allow for first-hand verification of the validity of stock information. In the case of missing documentation, site surveys may also provide the only source of information. Some important steps for the initial site survey are outlined below:

- Check that the as-built drawings conform to the actual building layout and function;
- Check the O&M manuals against equipment on site as these may have been modified, upgraded or decommissioned;
- Check that the required data monitoring and recording systems are in place and are operational;
- Interview operators and occupants to confirm operating schedules for systems and equipment and identify operating problems or special conditions which should be replicated by the calibrated model;

- Conduct spot/short-term measurements (Refer to Table 3-2);
- Perform additional building audits as necessary to account for any changes to building or system specification.

Further regular audits, during different times of the day, week or year may also be required to determine actual occupancy and equipment load schedules as these may be time-dependent. Once data has been collected over a number of days/weeks at selected intervals, typical values may be assigned to the schedules for different time periods. A summary of survey information is presented in Figure 3-7.

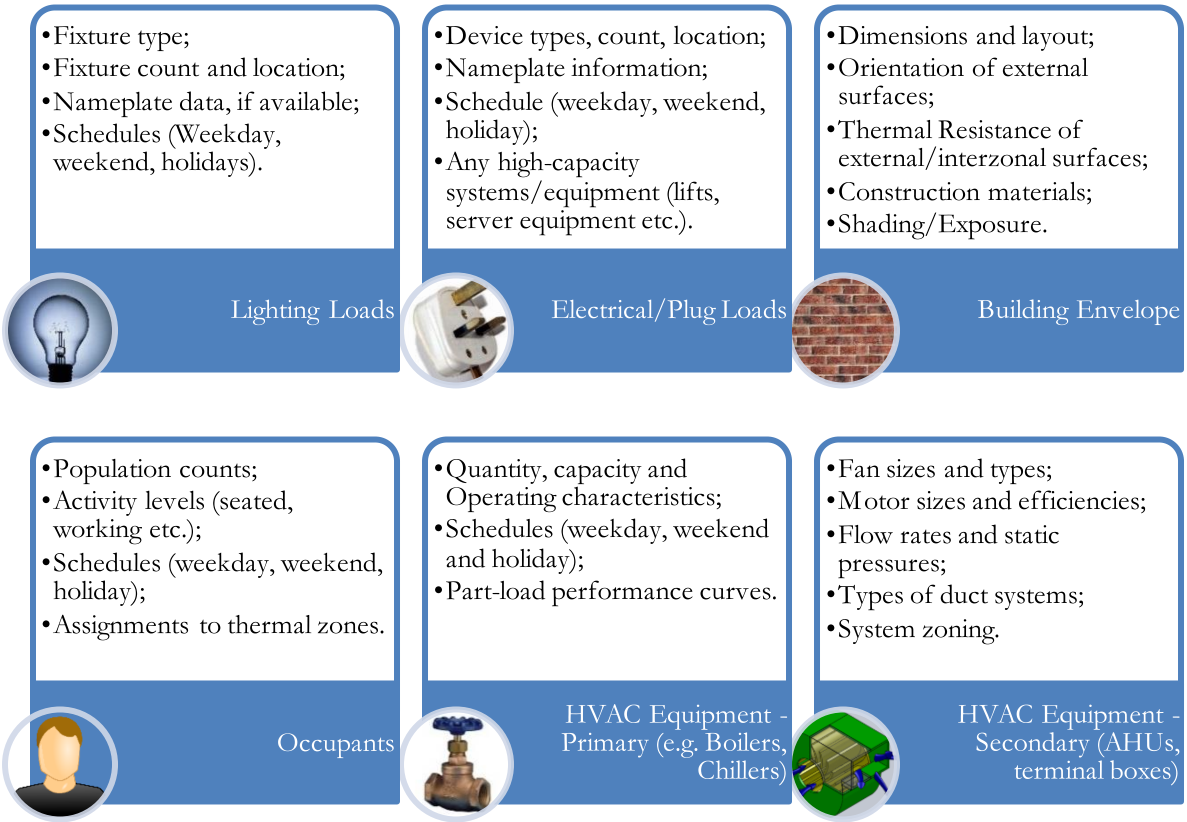


Figure 3-7: Survey information (adapted from ASHRAE Guideline 14, 2002)

Spot measurements may provide a useful means of gathering information about specific systems or building properties without the need for expensive metering or installations. Table 3-2 lists some valuable tests which may be carried out to gather additional input information for the model.

Table 3-2: In-situ / spot measurements (adapted from ASHRAE Guideline 14, 2002)

Category	Test	Equipment
Lighting	• Fixture power; • Operating Schedule.	• Clamp-on electrical meter – take short-term (<1 week) measurements for lighting circuit.
Electrical/Plug Loads	• Fixture power; • Operating Schedule; • Baseline Load.	• Clamp-on electrical meter – take short-term (<1 week) measurements for electrical circuit.
HVAC Systems	• Space temperature and humidity; • Air/Water flows; • Static pressures; • Duct temperatures; • Motor power; • Duct leakage.	• Hand-held Temperature/RH meter; • Flow/Air velocity meter; • Differential Pressure sensor.
Building Ventilation and Infiltration	• Air flow through outside ducts.	• Building Pressure Test; • Tracer gas test.

3.4.1.3 Weather Data

Accurate weather data for the calibration period is required to simulate building response to external conditions. Ideally, this data should be measured at a local weather station. Where onsite measurement is not available, data from local weather stations that are no more than 10km away, without a significant change in elevation, can be used. The weather station should maintain a historical record of environmental variables measured in hourly intervals (or more frequently) for the calibration period. This information is outlined below (with desired accuracy in parentheses):

- Dry-bulb temperature (± 0.1˚C);
- Wet-bulb temperature (± 0.1˚C) or relative humidity (± 2%);
- Wind speed (±0.1ms-1 (0.5 – 10ms-1); ± 1% (10 - 50ms-1); ± 2% (> 50ms-1));
- Wind direction (± 5˚);
- Total global solar radiation (± 2.5%);
- Barometric pressure (± 50Pa).

There are other potentially useful variables to measure, depending on the climate and type of systems installed in the building. For example:

- Rainfall;
- Snow presence and depth;
- Direct solar radiation.

3.4.2 Model Calibration Data

The 'model calibration data' is the data that will be used to verify the result of the simulation model. The accuracy of simulation models is usually assessed on the basis of their correlation with metered whole building energy consumption data (heating, cooling and electrical energy). Depending on the model requirements (ECM assessment, Benchmarking etc.), varying degrees of stringency may be applied to these calibration criteria. For the purpose of this thesis, we assume that a detailed model is required, calibrated to hourly measured data.

As part of the site survey, the analyst should ensure that the required data is collected at the frequency and accuracy required (Refer Section 3.3.1). At a minimum, the following should be recorded for any model calibration:

- Heating energy consumption, kWh, if applicable (Monthly)
- Cooling energy consumption, kWh, if applicable (Monthly)
- Electrical energy consumption, kWh (Monthly)

3.4.3 Data Proofing and Classification:

Using visual or statistical screening methods, check that the measured data coming from the BMS and monitoring devices is of the required standard for final hourly calibration as per ASHRAE Guideline 14-2002. Install any extra equipment/sensors where necessary.

Data should be classified according its source (Figure 3-8). These classifications are used to assign ranges of uncertainty, where explicit accuracies are unavailable, in the later stages of the methodology. It is useful to represent this information in tabular format with the following fields (see Appendix C.1):

- Input Data (e.g. Lighting Schedule);
- Classification (e.g. Spot-measurements);
- Comments (e.g. Date of measurement, times, equipment used and associated accuracy).

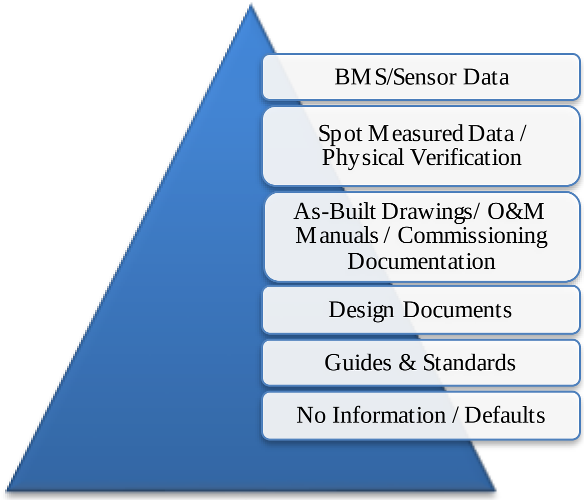


Figure 3-8: Hierarchy of source information

Table 3-3 lists some of the recommended categories for the source hierarchy. Each category also has an associated ranking (class) and range of variation (ROV, %):

$$ROV = \frac{3.\sigma}{\mu} \qquad (3.1)$$

This ROV represents the total heuristically estimated deviation (σ) from the mean value (μ). These are currently preliminary estimations based on prior experience. However, in the absence of any more detailed database of building parameter uncertainties, this provides an adequate means of generating an initial assumption.

Table 3-3: Data Categories

SOURCE	CLASS	ROV (%)
BMS/Sensor Data	1	2
Spot-Measured Data	2	5
Physically Verified Data	2	5
As-Built Drawings	3	10
O&M Manuals	3	10
Commissioning Documents	3	10
Design Documents	4	15
Guides & Standards	5	30
Reference Manual / Default Values	6	40
No Available Information	7	50

There are essentially two types of data source:

- **Qualitative sources** (e.g. drawings, guides, standards etc.): Data captured from such sources are assigned mean values and ranges of variation according to their type (e.g. discrete, continuous or multi-dimensional variables).
- **Quantitative sources** (e.g. sensor, BMS and spot-measured data): This data also has an associated measurement accuracy, which may need to be factored into the uncertainty quantification. In addition, there is an associated time-dependent variation (i.e. measurements will vary depending on the time, frequency and duration of their measurement). Therefore, this time-dependency and instrument accuracy will need to be combined in order to compute the associated uncertainty for quantitative data sources. A detailed discussion on uncertainty combinations may be found in Appendix B-5 of the IPMVP: Concepts and Options for Determining Energy and Water Savings Volume 1 (Efficiency Valuation Organisation (EVO) 2010)

Continuous Variables are assigned a range of variation based on the ranking assigned to the source information for that input variable. For example, Figure 3-9 shows the probability density function (pdf) for a typical building material. In this case, the material has a mean conductivity of 0.04 $W/m.K$ with a 30% range of variation as this information was sourced from standard tables of material properties (Guides & Standards).

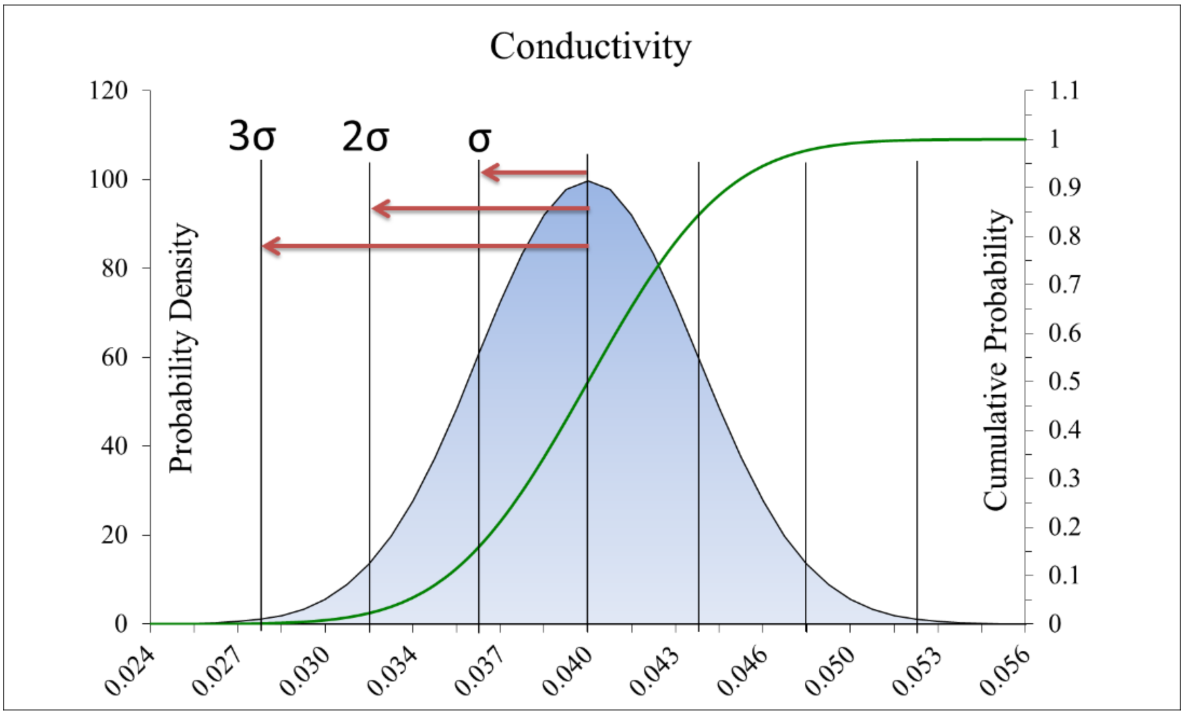


Figure 3-9: Batt insulation conductivity probability density function (pdf)

Discrete Variables may be characterised by assigning low, high and base-case estimates for each discrete input parameter, with a bias towards the base-case value. This bias can be achieved in practice by taking a similar approach to that used for continuous parameters and adopting a discrete value based on the sample point on the pdf curve (see Figure 3-10).

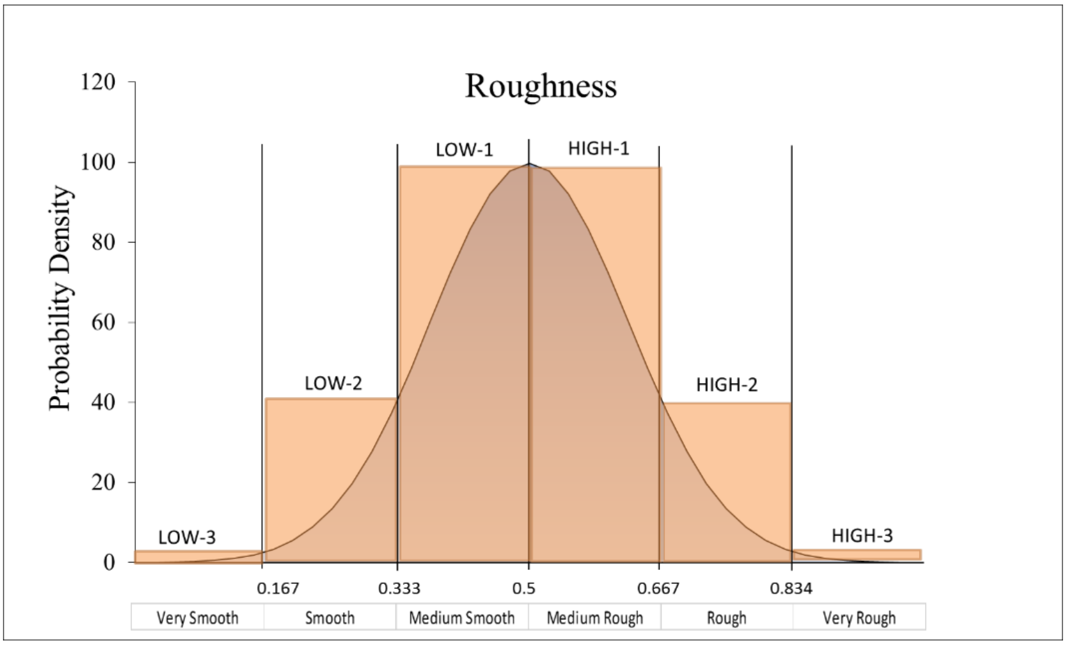


Figure 3-10: Batt insulation roughness pdf and associated discrete steps

Multi-Dimensional Variables such as lighting, infiltration and occupancy schedules may be sampled by splitting the schedule into discrete subsections and then sampling the relevant probability density function at each discrete time segment (see Figure 3-11). Further discussion on the discretisation of continuous and multi-dimensional variables can be found in Burhenne et al. (2010; 2011) and Reddy et al. (2007a)

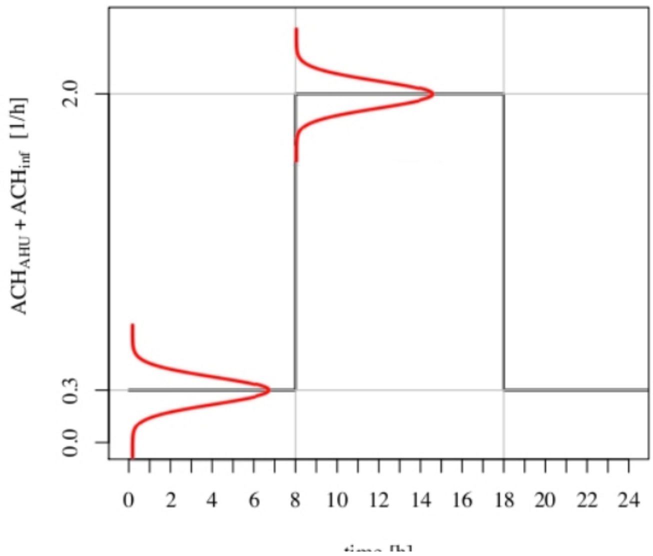


Figure 3-11: Schedule for combined zone infiltration and ventilation
(Burhenne et al. 2010)

3.5 Evidence-Based BES Model Development

This methodology calls for an iterative evidence-based development of the Building Energy Simulation (BES) model. This process is outlined in Figure 3-12 and described in further detail in the following sections.

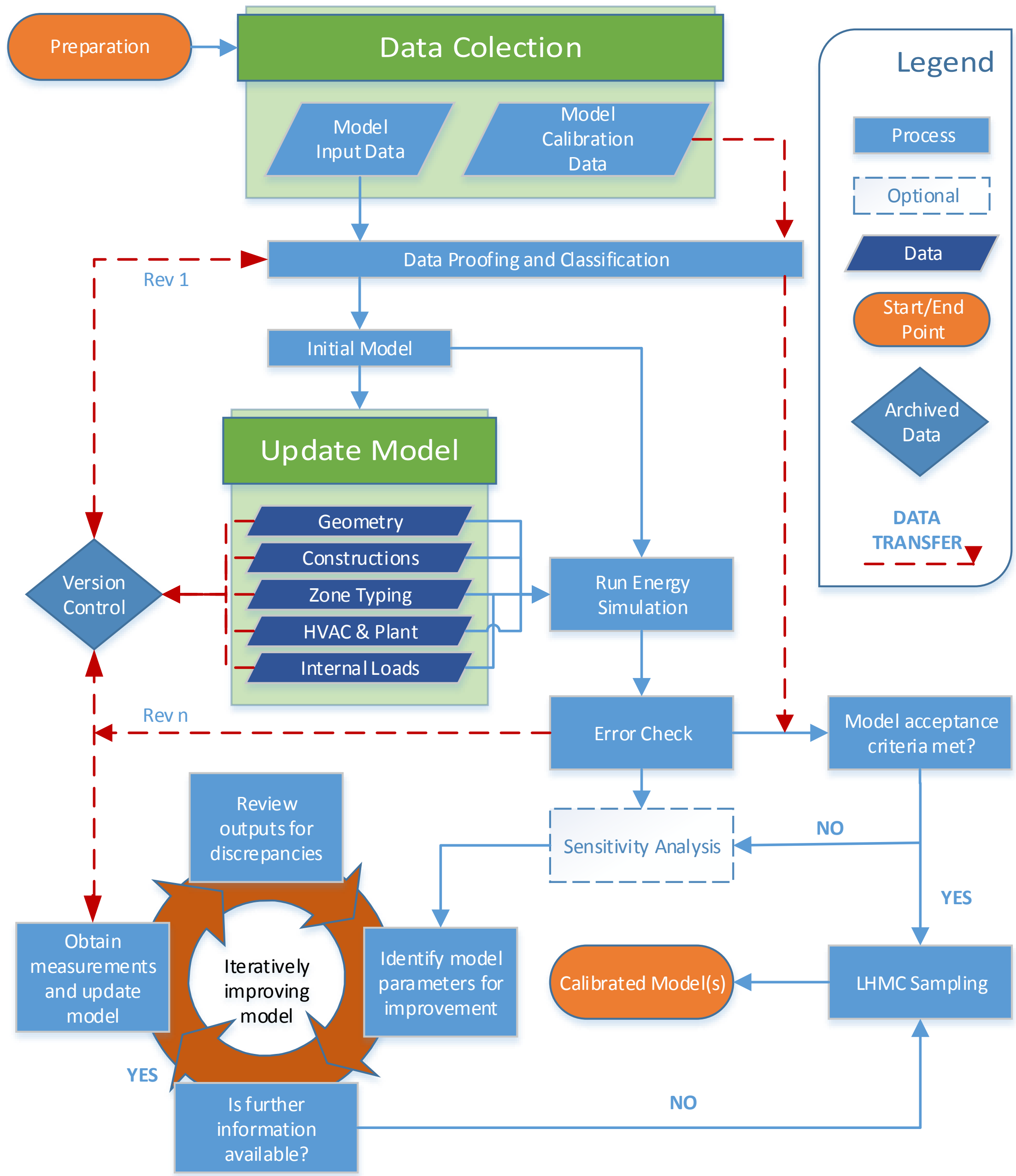


Figure 3-12: Adapted Evidence-based BES model development methodology (Raftery, Keane, O'Donnell, et al. 2011)

3.5.1 Initial Model

The first step in this section of the process is to create the initial Building Energy Simulation model. This can be completed using any of the model development tools described in Section 4.3.1 (e.g. EnergyPlus, TRNSYS, ESP-r). Using the initial information gathered in 3.4.1 (Model Input Data), create the initial model in the building energy simulation environment. Once complete, this model should be committed to the Version Control repository as '*Revision 1*'.

3.5.2 Update Model

The next step is an update of the initial model to include detailed information about the building, its systems and operation. Each iterative improvement should be committed to the version control repository with a comment, identifying the modification to the initial model and the source of evidence associated with this modification. By linking modifications to clearly defined sources within the evidence hierarchy, it will be possible to later quantify the inherent model uncertainty and perform a more robust sensitivity analysis to identify the most influential parameters which merit further investigation.

3.5.2.1 Geometry

If using an existing model (architectural or engineering BIM), it is essential to check that the model conforms to the actual building, as layout or function may have changed since the original model was developed. If a site survey has been carried out as per section 3.4.1.2, check that the building measurements match with those in the as-built layout or design model. Where there are inconsistencies, update the BES model geometry to reflect these discrepancies, referencing the source survey or documentation.

3.5.2.1 Environment

As discussed in section 3.4.1.3, accurate weather data is essential to simulating building performance, particularly for highly weather-dependent building types (e.g. offices, schools etc.). These building are generally more susceptible to outside weather conditions due to the nature of their design and operation, compared to factories or other industrial-type facilities which may tend to be dominated more by internal loads. Ideally, the model should use weather data collected locally, if available. Otherwise, data from the closest available weather station should be used instead.

3.5.2.2 Constructions and Materials

External and internal surface constructions and materials are added according to as-built construction drawings, if available. Again, survey information is useful here as the building may have been retrofitted or upgraded in the past.

Material information is also required for each construction type. This can be obtained from product specification sheets, if supplied. However, this information is not typically available for older buildings. Therefore, guides and standards may be used to acquire this data (CIBSE 2007; ASHRAE 2009). Alternatively, there are a number of useful software packages which may be used to obtain manufacturer-supplied material information for their products. Tools such as BuildDesk U (BuildDesk Ltd 2013) provide information relating to typical building constructions and material properties for the UK and Republic of Ireland.

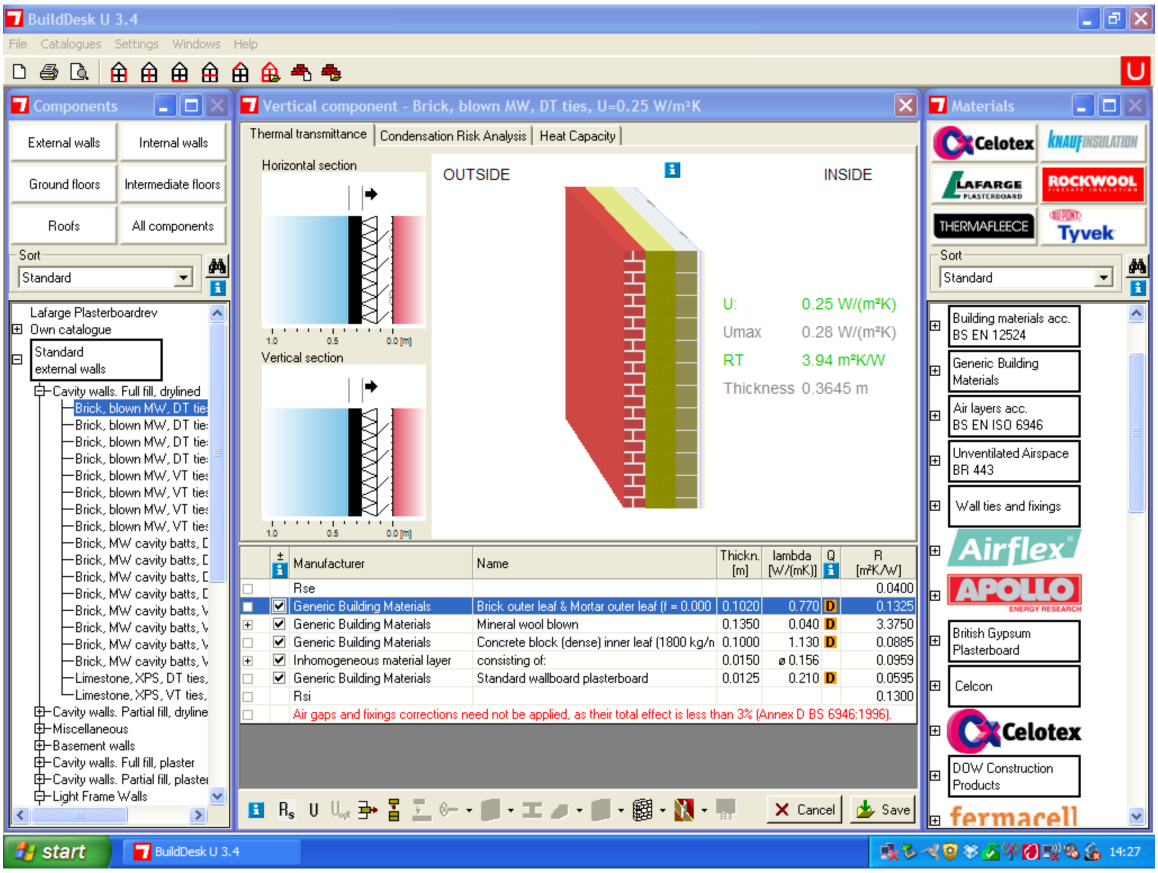


Figure 3-13: Screenshot from BuildDesk U (BuildDesk Ltd 2013)

It should be noted that not all properties are supplied by the manufacturer. Some values are taken from building regulations or standards, or defined manually by BuildDesk staff. However, the tool provides a 'source rating' system similar to the *source hierarchy* defined in the proposed methodology. Table 3-4 highlights the categories of source information quality defined by BuildDesk U.

Table 3-4: Source quality categories defined by BuildDesk U (BuildDesk Ltd 2013)

Class	Description
A	Data is entered and validated by the manufacturer or supplier. Data is continuously tested by 3rd party.
B	Data is entered and validated by the manufacturer or supplier. Data is certified by 3rd party.
C	Data is entered and validated by the manufacturer or supplier.
D	Information is entered by BuildDesk without special agreement with the manufacturer, supplier or others.
E	Information is entered by the user of the BuildDesk software without special agreement with the manufacturer, supplier or others.

3.5.2.3 Zone Typing

In order to maintain reasonable computational time, certain model simplifications are required. Since models depend on the definition of distinct zones in order to perform heat-balance calculations, a simplified representation of the real building is likely to be required here. Depending on the building under study, the analyst will need to decide on a suitable zoning strategy. A zoning strategy helps to simplify the model by aggregating similar thermal zones within the model. Numerous strategies have been proposed to handle this task (Souza and Alsaadani 2012):

- Zoning based on different spatial activities, spatial performance or building usage;
- Zoning based on different HVAC requirements and/or controls;
- Zoning based on different solar gains;
- Zoning based on temperature stratification;
- Zoning based on a combination of the above (Raftery, Keane, O'Donnell, et al. 2011).

Many of these strategies are described in considerable detail in relevant simulation reference documentation (CIBSE 1998; iSBEM 2006; DesignBuilder Software Ltd 2011; US Department of Energy 2011b). The determination of the best practice for defining a zoning strategy for typical buildings is outside the scope of this study. Furthermore, the most suitable zoning strategy is highly dependent on the requirements on (1) the function and size of the building, (2) the requirement of the BES model (benchmarking, ECM analysis, fault detection, etc.) and (3) the computational and time resources available.

3.5.2.4 HVAC and Plant Information

The next step is to update the plant and heating, ventilation and air-conditioning (HVAC) equipment in the model. This element of the process deserves careful consideration as it is likely to have a significant impact on the performance of the model. Information collected from as-built drawings, mechanical layouts and site visits should be collated and assessed for accuracy. Systems will need to be defined for each zone served. Information pertaining to the system components will also be required, including (Raftery 2009):

- Fans: Type, maximum airflow, pressure, operating efficiency, part load curve;
- Coils: Type, heating/cooling capacity, air/water on/off design temperatures, maximum air/water flow-rates, operating set-point(s);
- Motors: Type, maximum power, operating efficiency;
- Pumps: Type, maximum flow rate and head, part load curve, operating set-point(s);
- Boilers: Type, capacity, thermal efficiency, water flow rate and temperature, parasitic electrical consumption, part load curve, operating set-point(s);
- Chillers: Type, capacity, nominal coefficient of performance, design fluid temperature and flow rate conditions, capacity ratio curve, part load curve, operating set-point(s);
- Cooling towers: Type, capacity, fan power, design fluid temperature and flow rate conditions, part load curve (for variable speed cooling towers), operating set-point(s);

3.5.2.5 Internal Loads

The final step in this phase of model preparation is the addition of internal loads to the model. This will include the following:

- Lighting Loads – the lighting intensity level for the building (defined in Watts or Watts/area) and associated operation schedule. If lighting the lighting circuit is not sub-metered, this data can be obtained from design documents or verified by a lighting audit as part of the site visit. Base-loads may also be attained by sampling the electrical energy consumption for the lighting circuit at different times of the day/week/year.
- Electrical/Plug Loads – this incorporates all of the equipment loads (defined in Watts or Watts/area) for the building (computers, displays, printers, etc.) as well their operating schedule. As with the lighting schedule, if the electrical circuit is not sub-metered, an 'equipment audit' will be necessary (see section 4.4.1.2 (b)). Base-loads values may also be

attained by sampling the electrical energy consumption for the plug/equipment circuit at different times of the day/week/year.

- Occupancy – this is the occupancy level (defined as number of people, or people/area) for the building. This is frequently the most difficult parameter to accurately assess. However, valuable information may be obtained from other aggregated data sources (temperature profiles, CO_2 profiles, security/access information, RFID trackers, PIR sensors, PC usage, Wi-Fi network traffic etc.).

3.5.3 Error Check

At each stage in the model iteration, an error check should be carried out to verify the validity of results. Similar to the proofing of measured building data in 3.4.3, visually checking the outputs after each iteration can provide valuable feedback in identifying next steps for model improvement. Some features to look out for include:

- **Anomalous Data**: Abnormal peaks/dips in heat/electrical energy consumption data which may point to problems with the weather file or a possible scheduling defect.
- **Load profiles**: visually check the load profiles for heat/electrical energy consumption as well as zone temperatures to see how they compare to expected benchmarks or actual building data. It is useful to compare these profiles for:

    i. Hour of Day

    ii. Day of Week

    iii. Month of Year

  This is useful in determining where scheduling faults are present, for example during holiday periods or weekends. Factors of difference in these profiles may point to a more fundamental scaling error. Ensure you check the unit measurements in the model input file.

- **Surface/Construction warnings or errors**: These can occur where model geometry has been adapted from a previously created model or software version. It is important that these errors are identified and corrected to ensure model consistency.
- **Weather Profiles**: Check the input weather data for errors, specifically in relation to outside air temperature, humidity and direct/diffuse solar radiation.

3.5.4 Test for acceptance

Once the error check is complete and any modelling faults have been corrected, the analyst should test for correlation to the measured data set. This involves the calculation of the following dimensionless indices for the current model (ASHRAE 14 2002):

$$NMBE(\%) = \frac{\sum_{i=1}^{N_p}(m_i - s_i)}{\sum_{i=1}^{N_p}(m_i)} \tag{3.2}$$

Where; m_i and s_i are the respective measured and simulated data points for each model instance '*i*' and N_p is the number of data points at interval '*p*' (i.e. $N_{monthly}$=12, N_{hourly}=8760).

$$CV\ RMSE(\%) = \frac{\sqrt{\frac{\sum_{i=1}^{N_p}(m_i - s_i)^2}{N_p}}}{\overline{m}} \tag{3.3}$$

Where; m_i and s_i are the respective measured and simulated data points for each model instance '*i*'; N_p is the number of data points at interval '*p*' (i.e. $N_{monthly}$=12, N_{hourly}=8760) and $\overline{m}$ is the average of the measured data points.

This check should be performed at each model iteration (revision) and changes tracked on commitment to the version control repository.

If the model meets the acceptance criteria, then the model is considered calibrated and the analyst may proceed to 3.8.5 (*Uncertainty*). If the model fails to meet the acceptance criteria, then further iterative improvement is required. In order to identify areas for improvement, it may be useful to carry out a sensitivity analysis at this point. This will help to identify the most influential model parameters and those with the highest associated uncertainty bands.

3.6 Regional Sensitivity Analysis (Optional)

Sensitivity analysis evaluates the effect of changes in input values or assumptions on a model's results. *Uncertainty analysis* investigates the effects of lack of knowledge and other potential sources of error in the model. When conducted in combination, sensitivity and uncertainty analysis allow model users to be more informed about the confidence that can be placed in model results. (CREP 2009). It is recommended that a test for statistical significance be carried out to determine the factors of greatest influence on the model output. The analyst may then further investigate these variables within the BES model and revise their respective ranges of variation, if necessary. There are a number of sensitivity analysis methods available to the analyst (Tian 2013):

- Step-wise Linear Regression (Tian and de Wilde 2011; Wright et al. 2012; Hygh et al. 2012; Yildiz et al. 2012)
- Screening-based methods (Morris Method) (de Wit and Augenbroe 2002b; Hyun et al. 2008; Heiselberg et al. 2009; Moon 2010; Heo, R. Choudhary, et al. 2012; Garcia Sanchez et al. 2014)
- Variance-based methods (Mechri et al. 2010; Spitz et al. 2012; Shen and Tzempelikos 2013)
- Meta-model based methods (de Wilde and Tian 2010; Tian and de Wilde 2011; Eisenhower, O'Neill, Fonoberov, et al. 2012; Tian and Choudhary 2012)

In this case, I will focus on variance-based methods, as these methods can account for both input uncertainty and the contribution of the interaction between uncertain factors (Saltelli et al. 2010). The most widely used methods are Sobol, FAST, Random Balance Design, and the Monte Carlo method. Using a variance-based global sensitivity analysis (GSA) (Saltelli et al. 2004; Saltelli et al. 2008) the fractional contribution of each input factor to the variance of the model output (Y) is calculated. Interaction terms are also accounted for. In variance-based methods, the output variance V(Y) can be decomposed to the sum of a 'top-marginal variance' and 'bottom-marginal variance' (Grasman and Straten 1994; Ratto et al. 2001). Specifically,

$$V(Y) = V[E(Y|U)] + E[V(Y|U)] \tag{3.4}$$

Where; U is a group of one or more input factors (X_i), $E[.]$ and $V[.]$ denote expected value and variance, respectively. $E[Y|U]$ and $V[Y|U]$ denote the conditional expected value and conditional variance, respectively, of the output, given the input factor(s) U;

The top marginal variance from U is the expected reduction of the variance of Y in the case where U is known and fixed at nominal values. The bottom marginal variance is the expected value of the variance of Y where all inputs but U are known and U remains variable. The first order sensitivity index S_i, representing the average output variance reduction that can be achieved when X_i becomes known or is fixed, can then be defined:

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} \tag{3.5}$$

Where the subset U is reduced to a single factor X_i, *V[.]* denotes variance and *E[Y|X_i]* denotes the conditional variance of the output, given the input X_i.

Higher order (interaction) effects can also be computed with relatively low computational expense, by employing *total sensitivity indices*. The total sensitivity index, S_{Ti} for each input factor X_i collects in one single term all the interactions involving X_i. It is defined as the average output variance that would remain as long as X_i remains un-fixed.

$$S_{Ti} = \frac{E[V(Y|X_{-i} = x^*_{-i})}{V(Y)} \tag{3.6}$$

Where *E[.]* and *V[.]* denote expected value and variance, respectively. $V(Y|X_{-i})$ denotes the conditional variance of the output, given the input factor X_i, and, X_{-i} indicates all factors but X_i.

By estimating both the sensitivity index (S_i) and total sensitivity index (S_{Ti}), it is possible to estimate the direct impact of each input factor X_i on the output variance *V(Y)* as well as the overall impact through interactions with other input factors.

By narrowing the ranges of variation of influential variables and freezing weak parameters, the analyst may perform a more robust search of the refined parameter space. The test for acceptance (3.5.4) is repeated to determine the goodness-of-fit (GOF) of the refined solutions. However, it is unclear at this time whether a sensitivity analysis will produce significant improvements in model accuracy. In addition, high-dimensional model sensitivity analysis requires the application of advanced techniques to account for their complex interaction structure. Hence, this discussion on the subject is limited to the above provided reference.

3.7 Iterative Model Improvement

At this stage, the analyst must go through a process of iteratively improving the model by adjusting parameter estimates and constraining uncertainty ranges based on higher quality information (i.e. information from more reliable sources). Based on a sensitivity analysis (Section 3.6) or expert judgement, the analyst should identify parameters for improvement. Following, this identification process, sources of further information relating to these parameters must be identified (e.g. spot measurements, additional audits, improved metering etc.). Version control software tracks each model revision. A description of the revision and evidence is provided for each change.

3.8 Latin-Hypercube Monte-Carlo (LHMC) Search

This stage of the statistical calibration methodology involves fixing the ranges of variability of continuous parameters to reduce the dimensionality of the search space. This is followed by a random sampling Latin-Hypercube Monte-Carlo (LHMC) method to identify promising regions of local optima. This requires the following steps:

1) Define best-guess estimates for input parameters
2) Assign ranges-of variation (ROV, %)
3) Generate sample matrix
4) Carry out parametric simulation
5) Uncertainty Analysis

3.8.1 Define best-guess estimates for input parameters

Once, the iterative model improvement process is complete, there now exists a building model which has assigned input values verified by source evidence, where available, and best-guess approximations where no further measurements are available or possible.

3.8.2 Assign ranges of variation

In this step, the analyst should define ranges of variation for uncertain or unknown input variables based on the hierarchy defined in section 3.4.3. For continuous parameters, ranges of variation are characterised by probability density functions (pdf) bounded by upper 95th and lower 5th probability threshold values. Discrete variables are characterized by minimum, maximum and base-case values. The analyst must also account for multi-dimensional variables, such as occupancy schedules which are dependent on time of day, day of week as well as time of year. Such parameters may be discretized for specified periods and assigned ranges of variation based on available information.

3.8.3 Generate sample matrix

In a Monte Carlo analysis, a large number of evaluations of the model are performed with randomly sampled model inputs (Saltelli et al. 2008). Using the Monte-Carlo approach produces an input sample such as;

$$M = \begin{bmatrix} z_1^{(1)} & z_2^{(1)} & \cdots & z_r^{(1)} \\ z_1^{(2)} & z_2^{(2)} & \cdots & z_r^{(2)} \\ \cdots & \cdots & \cdots & \cdots \\ z_1^{(N-1)} & z_2^{(N-1)} & \cdots & z_r^{(N-1)} \\ z_1^{(N)} & z_2^{(N)} & \cdots & z_r^{(N)} \end{bmatrix} \tag{3.7}$$

Where; z represents the input variables; N the sample size and M the corresponding vector solution matrix.

The number of samples generated will depend on a number of factors:

- Model complexity: number of model input parameters and degree of associated uncertainty;
- Resources: time and computation resources available to conduct sampling. A larger samples size, and increased computation time, will most likely result in a better-fitting model. However, this is subject to the law of diminishing returns, in that the final model fit does not follow a direct linear relationship with the number of samples generated.

As an initial approximation of number of samples that can be conducted, it is useful to check the time required to run a single model, and dividing this into the total time available to run the random search. For example,

Time Available: 3 days = 3*24*60*60 = 259200s

Time required (per model) = 80s

Total Samples = Time Available / Time Required = 3240

It should be noted that due to the law of diminishing returns, it may be more economical to run fewer models, with iterative refinement, as opposed to one large sample run. This decision will ultimately be dependent on the model complexity, resources available and level of calibration required (Refer section 3.3.1).

3.8.4 Parametric Simulation

In this step, sample trials are run using a batch simulation tool (Wetter 2001; Zhang and Korolija 2010; Yi Zhang et al. 2012). This generates the desired output vector, Y, for each row of input matrix:

$$Y = \begin{bmatrix} y^1 \\ y^2 \\ \cdots \\ y^{N-1} \\ y^N \end{bmatrix} \tag{3.8}$$

3.8.5 Uncertainty Characterisation of Results

Since calibration is a highly underdetermined problem, more than one solution may satisfy the objective function. In addition, it is erroneous to assume that a solution that satisfies the objective function is therefore implicitly correct. This is due to the many degrees of freedom that may produce good calibration overall even though the individual parameters may be incorrectly identified. As such, it is important to convey these uncertainties, as propagated throughout the above model development process, to the client or decision-maker. A detailed discussion on expression of uncertainty relating to energy savings predictions can be found in the IPMVP Guidelines: Volume 1 (Efficiency Valuation Organisation (EVO) 2010)

Using the set of parametric simulations generated in the previous step (Section 3.8.4), it is possible to illustrate the accuracy and variation of model output predictions. This requires the ranking each solution based on statistical goodness-of-fit (GOF) criteria which are calculated using the statistical indices described in section 3.5.4 (ASHRAE 2002).

$$\mathrm{GOF_A} = \left[\frac{\left(w_{kWh}^2 A_{kWh}^2 + w_{deg}^2 A_{deg}^2 + w_{kWh(H)}^2 A_{kWh(H)}^2\right)}{\left(w_{kWh}^2 + w_{deg}^2 + w_{kWh(H)}^2\right)}\right]^{\frac{1}{2}} \tag{3.9}$$

$$\mathrm{GOF_B} = \left[\frac{\left(w_{kWh}^2 \mathrm{B}_{kWh}^2 + w_{deg}^2 B_{deg}^2 + w_{kWh(H)}^2 \mathrm{B}_{kWh(H)}^2\right)}{\left(w_{kWh}^2 + w_{deg}^2 + w_{kWh(H)}^2\right)}\right]^{\frac{1}{2}} \tag{3.10}$$

Where; 'A' represents Coefficient of Variation of the Root Mean Square Error (CV RMSE); 'B' represents Normalised Mean Bias Error (NMBE); w_{kWh} and $w_{kWh(H)}$ are weighted ratios which represent the ratio of the annual cost of electricity and hot water use, respectively, in *kWh* divided by the total annual utility cost; w_{deg} is a weighted ratio which represents the mean zone air temperature (average room temperature weighted by room volume where multiple room temperature sensors are present) in *°C* and ($w_{kWh}+w_{kWh(H)}+w_{deg}=1$). The weighting applied to w_{deg} is at the discretion of the analyst. A higher weighting will bias the model calibration towards agreement with internal zone temperatures as opposed to energy consumption and vice versa. A default value of 0.2 is recommended in order to focus the model fit towards accuracy of energy consumption data, while filtering out highly unrepresentative models in terms of actual building operation. The inclusion of w_{deg} in the calibration criteria therefore reduces the risk of choosing a less representative BES model.

Since, the 'best solution' will depend on the weighting assigned to the CV and NMBE indices; we therefore introduce a consolidated index as follows:

$$\mathrm{GOF_{TOTAL}} = \left[\frac{(w_{CV}^2 GOF_{CV}^2 + w_{NMBE}^2 GOF_{NMBE}^2)}{(w_{CV}^2 + w_{NMBE}^2)}\right]^{1/2} \tag{3.11}$$

Where ($w_{CV} + w_{NMBE} = 1$).

In practice, it is likely that building energy managers would prefer the calibration to capture the overall mean energy consumption more accurately that the daily or monthly variation (as described by the CV value) (Reddy et al. 2007a). This is reflected in the ASHRAE Guideline 14-2002 (ASHRAE 2002) recommendation for a 1:3 weight for w_{CV}:w_{NMBE}. These weights may be adapted depending on the model purpose (see Section 3.3.1). Using these indices, it is possible to filter out parameter vectors that result in high GOF numbers. Simulation trials with lower GOF numbers represent parameter vectors that provide a closer match to measured utility data.

Depending on the results of this characterisation, the analyst may perform a further refinement of the search space. Section 5.4.2.4 of the ASHRAE-14 document (ASHRAE 14 2002) stipulates that the calibrated computer simulation model should be accurate to within 10% for the NMBE and 30% for CV(RMSE) relative to hourly measured data (see Table 2-4). In case a calibration run yields relatively few solutions (<0.1% of total) meeting these criteria, the analyst should consider re-evaluating the stipulated ranges of variation of some of the influential parameters. This entails performing a Regional Sensitivity Analysis (see Section 3.6) to determine the influential variables that require re-evaluation.

3.9 Conclusions

This chapter presents a novel methodology for the calibration of detailed buildings energy simulation models, based on an analytical optimisation approach. The approach differs from existing approaches to model calibration in a number of ways:

- Structured evidence based approach: the methodology endorses the principle of evidence-based model development and iterative improvement, thus increasing the reproducibility and reliability of the final models;

- Systematic development: Version control techniques are employed to systematically capture changes made during model development. This also improves the reproducibility and reliability of results;
- Sensitivity Analysis: The methodology includes a means of capturing the most influential model input variables in order to guide the iterative model improvement process. Input variables with a combination of high a sensitivity index and high uncertainty are targeted for further investigation;
- Parametric Analysis: The calibration process includes a bounded parametric study based on input parameter uncertainty. This allows for the capture of a range of results for calibrated simulation models as opposed to the single output provided by existing approaches;
- Uncertainty Characterisation: The methodology includes a measure of parameter uncertainty based on source evidence, which is propagated through the model development process to enable risk and uncertainty quantification of final model predictions. Other sources of uncertainty (e.g. measurement accuracy, modelling uncertainty) are currently not captured within this approach, and deserve further consideration for future work (see Section 6.3.1).

The following chapter demonstrates the application of this methodology to a demonstrator building in order to evaluate the viability and potential pitfalls of such an approach.

Chapter 4: Case Study

"It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."

– Arthur Conan Doyle, Physicist and Writer

4.1 Chapter Introduction

This chapter describes the calibration of a detailed whole-building energy simulation model of a mixed-mode building to hourly measured data. This demonstrates the application of an evidence-based analytical optimisation approach described in a previous chapter. The methodology is applied to a mixed-mode library building at the National University of Ireland, Galway.

A detailed simulation model is developed and calibrated to measured building data for heating and electrical energy consumption as well as zone temperatures. This is achieved by assigning probability distribution functions to continuous model parameters, generating simulation trials based on random sampling of these distributions and ranking solutions based on a calculated goodness-of-fit. The chapter concludes with a discussion of the key findings and difficulties encountered during this study.

4.2 Case Study: Nursing Library

This section will cover background information relating to the case study building, in particular:

- Building Description;
- HVAC Systems;
- Stock Information and Measured Data;
- Building Management System (BMS).

4.2.1 Building Description

The Nursing Library is a newly constructed building at the National University of Ireland, Galway located in Galway, Ireland. The 3-storey building contains a library and study areas, as well as a computer room on the ground floor. Completed in 2009, the building has a gross floor area of approximately 700m^2. A set of floor plan layouts are available in Appendix B.3. Also, it should be noted that the location of the building ensures easy access for site surveys and further measurement.

Key Facts	
Year	2009
Area	700m^2
HVAC	Mixed – Mechanical/ Natural Vent
Data	Heat, Electrical, Temperature, CO_2
Documentation	As-Built Drawings, O&Ms

Figure 4-1: NUI, Galway Nursing Library

The building is currently a focus of research in areas such as whole building energy simulation, energy performance of earth tube systems, and calibrated CFD modelling of naturally ventilated spaces.

4.2.2 HVAC systems

This mixed mode building has a dedicated outside air system (DOAS) for ventilation and both automatically and manually operated windows for natural ventilation. The DOAS draws air through an earth tube system (Figure 4-2) to moderate the air temperature in the winter and summer months. Stand-alone direct exchange units (Figure 4-3) cool the computer room on the ground floor. Convective hot water baseboard heaters maintain indoor temperatures outside of the summer months. Campus-wide district hot water supplies all of the heating systems in the building.

Figure 4-3: Direct electrical heat-exchangers

Figure 4-2: Earth Tube / Earth-Air Heat Exchanger

4.2.3 Stock information and measured data

The quality of stock information about the building is very high due to its recent construction, and the attention paid to the building during commissioning by researchers at the University. High quality as-built drawings and detailed information on materials and constructions are available. In addition, O&M information and detailed design criteria is available for all the HVAC equipment.

The existing building management system (BMS) monitors a significant number of points in this building. These include:

- Space temperature (*ºC*);
- Space CO_2 levels (*ppm*);
- Electrical Energy Consumption (*kWh*);
- Heat Energy Consumption (*kWh*).

The electrical panel also explicitly separates electricity consumption by end-use (e.g. HVAC, lighting and plug loads), although these are not logged independently.

4.2.4 Building Management System

The building is monitored and controlled by means of a central building management system (BMS). Access to the Building Management System was obtained through the Building and Estates Office at NUI Galway, via a secured remote desktop connection. This enabled logging and archiving of BMS Data. Initially the data was not archived for a sufficient period to be of use in annual calibration of whole building energy models. The BMS was modified in April 2011 to allow for archiving of long-term data.

This was achieved using an automated historian service included in the BMS software (See Figure 4-4 and Figure 4-5). Scheduled data archiving was carried out on a weekly basis, commencing at 4:00 a.m. At the time of writing, the only available archive file format was individual comma separated value (csv) files. These files were downloaded on a weekly basis and loaded into a MySQL database for ease of access and analysis (see Appendix A.2 and A.3).

Figure 4-4: Screenshot of BMS Interface (Cylon)

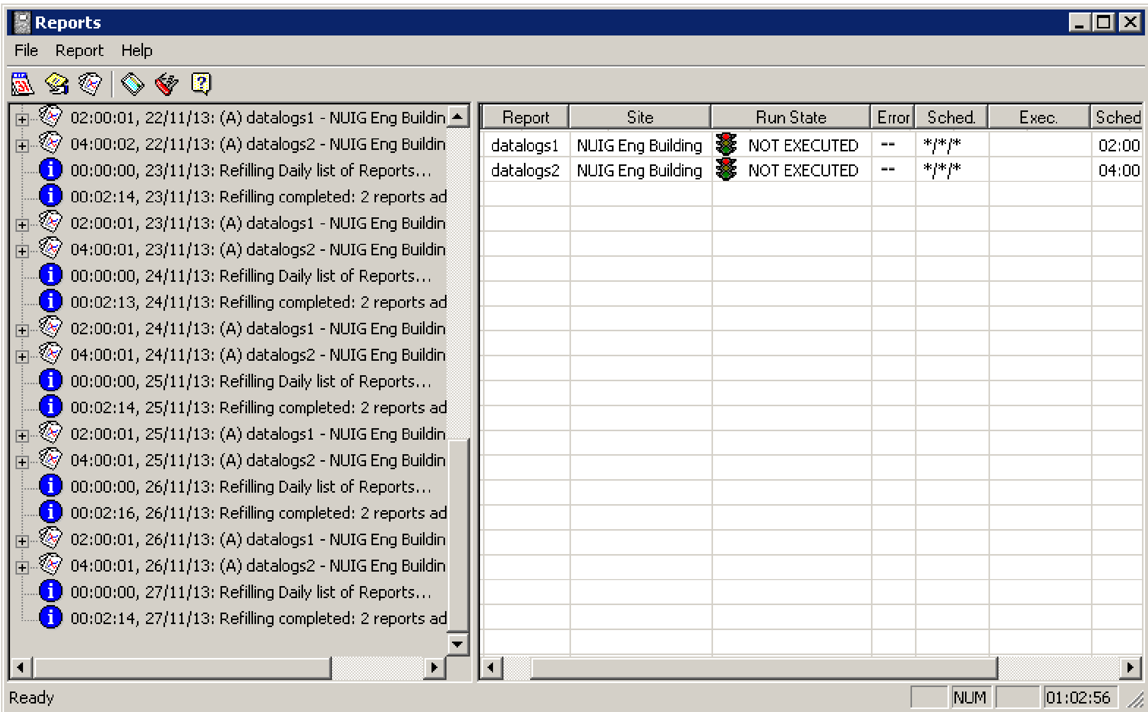


Figure 4-5: Cylon Automated Reporting Software

4.3 Preparation

At the outset of this project, it was decided that the primary function of this study would be to demonstrate the efficacy of the proposed calibration methodology, as well as highlighting the challenges associated with the calibration of detailed BES models in typical office buildings.

4.3.1 Simulation Tool Selection

In order to help identify which piece of software or software combination would be most suitable, a list of requirements was prepared;

- **Function**: capable of modelling the thermal response of a building given a set of physical building parameters and system inputs.
- **Accuracy**: The accuracy of the product is essential as we plan to use some of the response outputs to compile response algorithms through mathematical regression.
- **Flexibility**: We need to be able to model real-world data and systems as accurately as possible. Therefore, data entry flexibility is essential.
- **Clarity**: Results of analysis should be presented clearly and concisely in a usable format. This includes transparency of calculations used during simulation.
- **Usability**: We require software which is user-friendly and will not require in-depth knowledge for effective use.
- **Integration**: Ideally, it would be useful if we could use the software with other mainstream products such as AutoCAD and Microsoft Excel formats.
- **Adaptability**: Where the software does not match our requirements exactly, it will be necessary to adapt the program for our use. For example, the use of real-world data may not be supported on all commercial applications. However, such an allowance may be built-in if the product is open-source.
- **Support**: It is essential that product support is readily available. Detailed documentation of product features and user-guide are essential. Support for software bugs is also a necessity should problems arise with the software in the future.

Based on the above requirements, a number of mainstream simulation tools were shortlisted (see Table 4-1):

Table 4-1: Comparison of software tools based on defined requirements

	Function	Accuracy	Flexibility	Clarity	Usability	Integration	Adaptability	Support	Total
EnergyPlus (Version 7.0)	5	5	4	5	3	4	4	4	34
ESP-r	5	5	4	5	3	4	4	3	33
IES-VE	4	4	3	3	5	4	1	4	28
TRNSYS	3	3	3	5	3	4	4	3	28
eQuest (DOE-2.2)	2	4	2	4	3	3	3	2	23
ECOTECT (Autodesk)	2	3	1	3	4	4	2	3	22

The tool that provided the closest match to our requirements were EnergyPlus (Crawley and Lawrie 2001) and ESP-r (Strachan 2000). Further investigation was required to identify the most suitable tool given the anticipated requirements of our chosen case study (Refer 3.3.1). Maile el al. (2010) provide a comparison of BEPS tools based on the requirements for use during building operation. Based on these requirements, as well as availability of local expertise, EnergyPlus was deemed to be the most appropriate tool for the task, and was selected for use in the presented case study.

4.3.2 Modelling Strategy

Given the available resources, in terms of computing power, measurement data, and available time and requirements, a simplified modelling strategy was adopted. This approach called for a single-zone model of the existing building, for the following reasons:

- Metered energy-use data was only available at the whole-building level;
- The heating system in the building was implemented on a simple North/South circuit, metered at the whole-building level. Given the control strategy, which utilizes aggregated mean space temperatures, it was decided that a simple zoning strategy would be more suitable to replicate this system;
- The methodology requires parametric simulation involving thousands of simulation trials. Therefore, it is necessary to reduce model complexity, where possible. A refined model may include more detail at increased computational cost.

Zone-typing was not used in this case as there was insufficient zone-level data available to justify its inclusion in this study. In addition, much of the building is mixed-function, in that the majority of the building spaces serve as general open-plan study and meeting areas. There are a number of offices and a computer lab on the ground floor. However, as these spaces were all served by the same heating system, and their operating schedules were largely similar to the rest of the spaces in the building, it was decided to maintain a single zone model for the sake of simplicity.

4.3.3 Source Control Management

In order to track model development and allow consistent access to information, a file repository (see section 3.3.2) was established on a local server. This was used to archive evidence data such as building drawings, audits, photographs and manuals. A source control repository was also established to track changes to certain files throughout the project (e.g. model geometry). This was achieved through the use of TortoiseSVN (TortoiseSVN Core Development Team 2011). The subversion (SVN) repository is used to track iterative model revisions, as well as keeping a record of associated source evidence (see Figure 4-6).

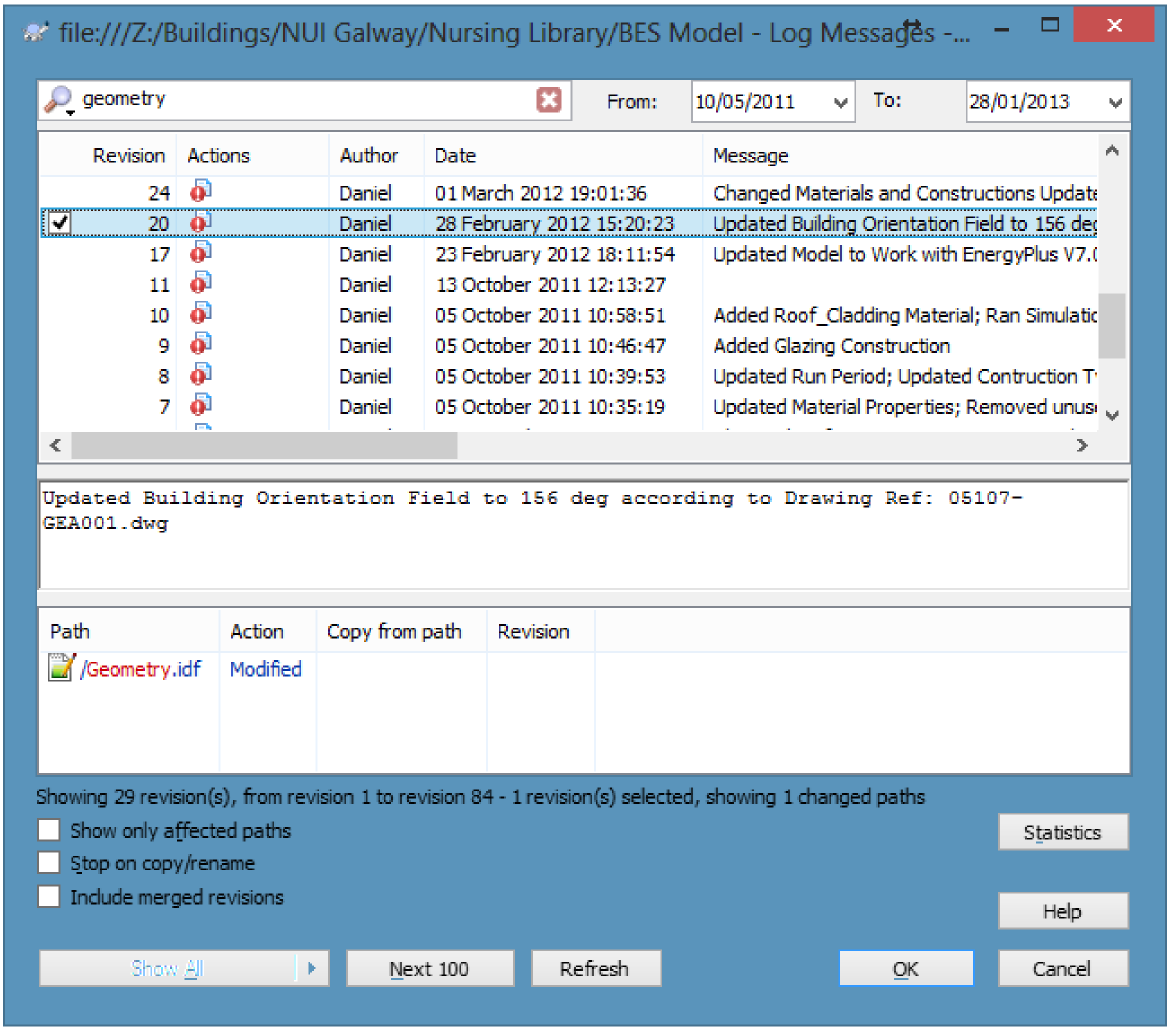


Figure 4-6: SVN source control management

4.4 Data Collection

One of the most important aspects of producing an accurate model of an existing building, is collecting accurate data to inform model inputs and provide sufficient accurate measured data for final model calibration and validation. As per Figure 3-6 in section 3.4.1, this is split into two phases of data collection; Model Input Data, and Model Calibration Data.

4.4.1 Model Input Data

As a Building Information Model (BIM) was not available for this building, it was necessary to collect data from other sources in order to compile sufficient information pertaining to the building, systems, environment and occupants. A set of as-built drawings for structural, mechanical and electrical layouts were obtained from the Buildings and Estates Office at NUI, Galway. In some cases, these drawings were often only made available in pdf format as opposed to their native CAD format (.dwg or .dwf). Mechanical and Electrical Operation and Maintenance (O&M) manuals were also provided for the building. These proved to be a valuable resource in attaining accurate specifications for building and HVAC system components.

4.4.1.1 Occupancy

A site survey is an extremely useful method for gathering detailed information about a building. From an initial visit to the building, it was determined that a number of factors would require further investigation. Firstly, due to the building function as an office/study area, it has a dynamic occupancy profile, with occupancy levels varying widely depending on time of day, day of week and time of year (summer vs. Academic Year). Occupancy is one of the major factors influencing energy use in buildings, particularly in office buildings. In such buildings, occupants are the main driver of heating/cooling energy demand as well as electrical energy consumption, in contrast to large commercial or manufacturing facilities which are dominated by predictable mechanical or equipment loads. Therefore, it is essential that occupancy is accurately determined when modelling office buildings. Unfortunately, this is often quite difficult to monitor in office buildings.

A number of possibilities were investigated but were ruled out due to limitations relating to resources, functionality, or, implementation restrictions:

- Directional People Counter: Since the building only had a single point of entry and exit, a dual-direction infra-red sensor would allow accurate tracking of the number of people present in the building at any given time. However, the cost of this instrumentation proved to be prohibitive for this project.

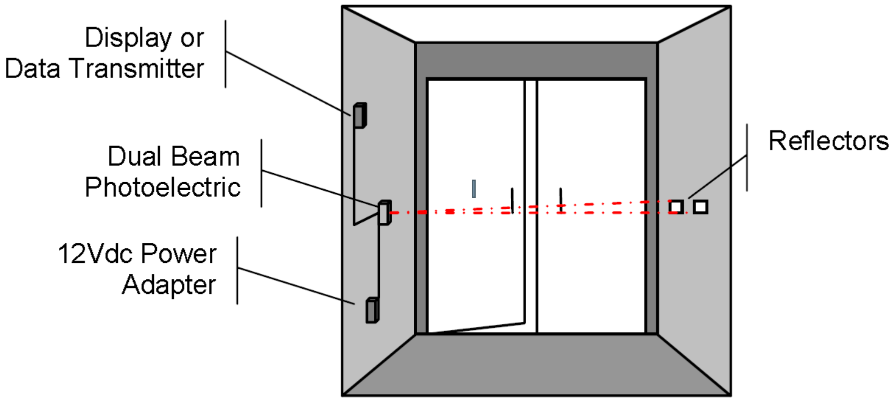


Figure 4-7: Sample configuration of dual-direction people counter

- Single-beam counter: Access to a uni-directional infra-red (IR) sensor was acquired. However, this is only capable of logging intermittent breaks in the IR beam, and does not account for people entering or leaving the building. For this reason, the device is more useful for monitoring foot traffic through an area, rather than actual occupancy;

- Camera monitoring: A camera system was proposed to monitor the main open-plan areas in the building (study areas, computer lab). However, this solution was unavailable due to restrictions relating to individual data privacy;

- Carbon dioxide correlation: It was noted that the level of carbon dioxide in each space was highly correlated with occupancy levels, but is also heavily influenced by levels of outdoor background CO_2, wind speed and infiltration rate. With further investigation this may provide some useful for occupancy inference. However, this was outside of the scope of this project.

As an alternative, audits were conducted in order to determine occupancy levels as well as daily and weekly occupancy profiles. These audits were conducted over the course of a week at functionally different times of the year (i.e. academic and summer periods). As part of this survey, the following information was collected at regular intervals throughout the day, for each day of the week (Refer to Appendix C.2):

- Date and Time;
- Number of people present in each area;
- Number of windows open;
- Number of laptops in use;
- Counter recording (uni-directional people counter).

This data was collected for occupied spaces within the buildings, allowing for the creation of simplified occupancy schedules corresponding to specific times of year (see Table 4-2). There is a clear difference between academic and summer-time occupancy, thus justifying the need for separate schedules for these periods (see Figure 4-8 and Figure 4-9).

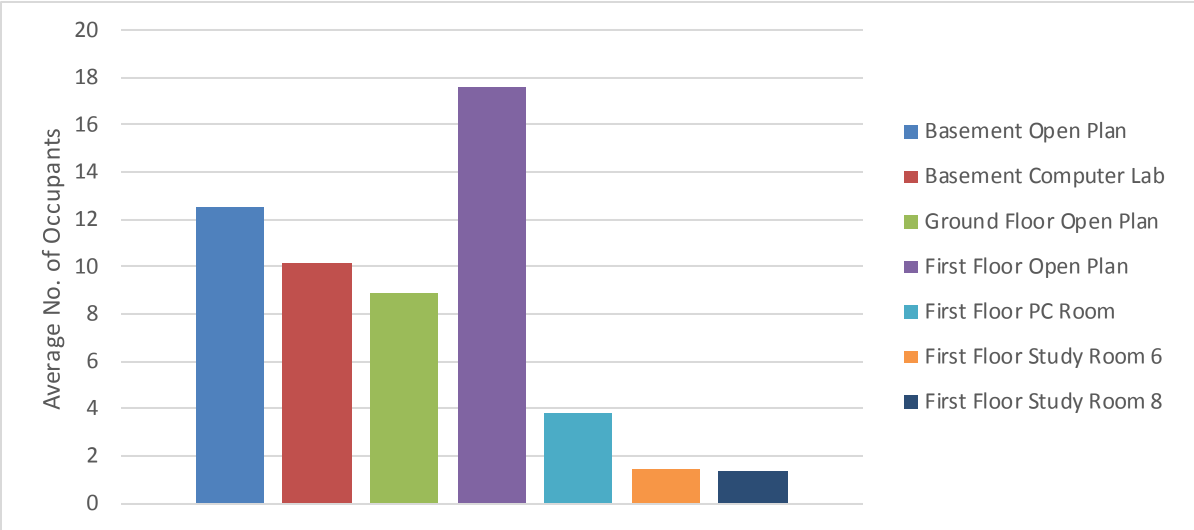


Figure 4-8: Average space occupancy (spring)

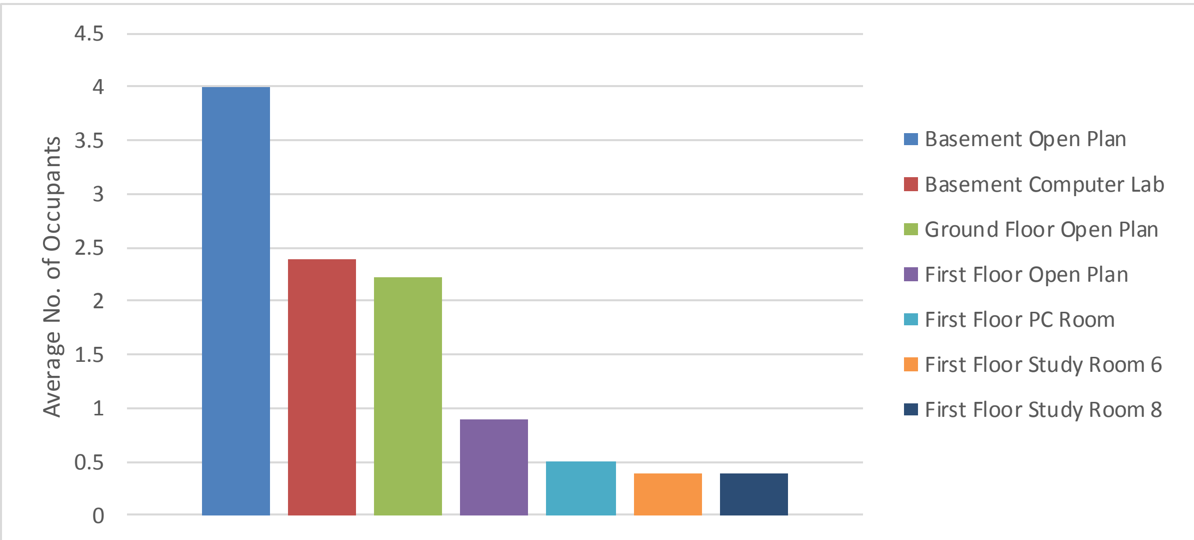


Figure 4-9: Average space occupancy (summer)

Table 4-2: Average space occupancy by room and time of year

Term	Basement Open Plan	Basement Computer Lab	Ground Floor Open Plan	First Floor Open Plan	First Floor PC Room	First Floor Study Room 6	First Floor Study Room 8
Academic	12.55	10.14	8.91	17.59	3.82	1.43	1.33
Summer	4.00	2.39	2.22	0.89	0.50	0.39	0.39

This data was then used to plot the overall daily average occupancy profile based on hour of day, for use in the EnergyPlus occupancy schedules (see Figure 4-10).

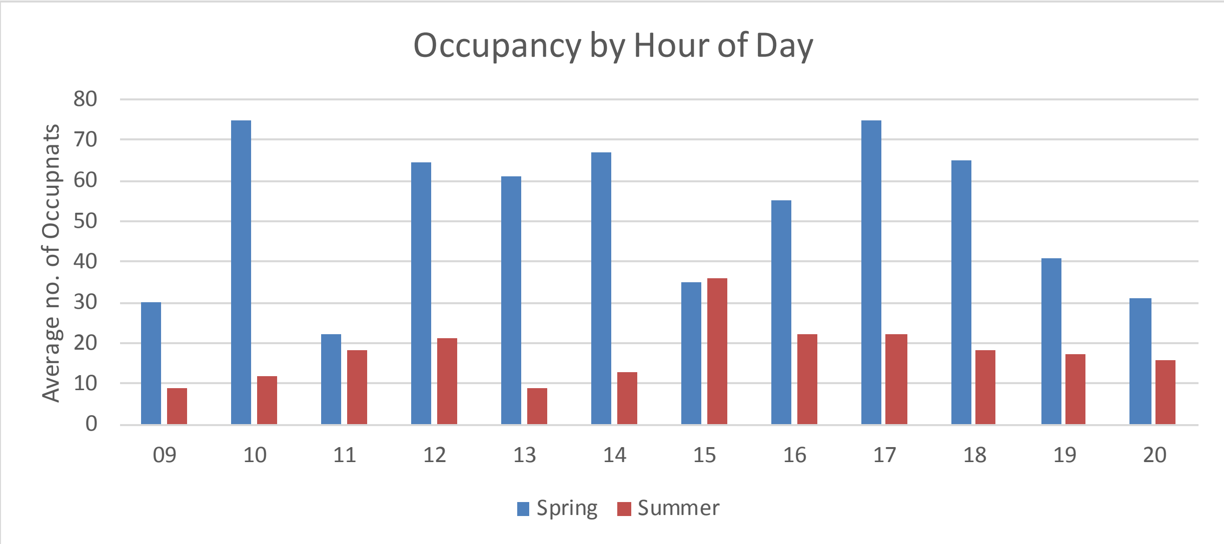


Figure 4-10: Average occupancy by hour of day (spring vs. summer)

As can be seen from the graph, the occupancy level differs greatly between summer and academic year. However, it is difficult to determine a consistent occupancy profile due to the large variations present in the data. This can be explained by:

- Lack of sufficient data: Due to resource and time constraints, it was only possible to conduct a limited number of occupancy audits during the course of this study. In light of the fact that detailed occupancy information may not be available for many buildings, it was decided to accept this limitation rather than assign unfeasible levels of resources to carry out more detailed analysis;

- Functional variation: As mentioned in the introduction, this building functions as an office space, a study, as well as a computer lab. For this reason, occupancy levels can vary widely depending on the time of year due to scheduling of events, classes or exams. At present, these schedules may vary between one semester and the next. In addition, one-off events, such as training courses and exams may skew occupancy levels.

Since occupancy has a large impact on this particular case study, it was decided to explore alternative sources of data, which may help to improve our knowledge of occupancy profiles for the building. Further examination revealed two more potential sources of detailed occupancy data.

Firstly, as mentioned above, scheduling of classes and events plays a large part in determining space occupancy. Therefore, it would be beneficial to tap into the scheduling resource to extract usable information relating to the date, time and number of people scheduled for particular events. It has been shown that such data can play a valuable role in minimising cost and maximising performance, particularly in dynamic office and educational spaces (Curry et al. 2013). Following interviews with staff, this information was found to be logged through a proprietary web-based scheduling service, which is currently not open to data extraction.

A second novel data source related to the usage statistics for the PC's in the computer laboratory. This is currently tracked by Information Solutions and Services (ISS) as part of the IT infrastructure and provision of services, using a management tool called LabStats (Labstats LLC 2013). LabStats tracks user activity for each PC on campus, recording details such as:

- Login Time and Logoff Time;
- Computer Name and IP address (giving accurate positional information);
- Duration.

Table 4-3: Sample of Computer Login History (December 2011)

Login time	Logout time	ID	Computer name	IP address	Duration (hrs.)
13/12/2011 13:24	13/12/2011 13:44	3509	NLDA05	10.210.12.193	0.34
13/12/2011 13:11	13/12/2011 13:14	3509	NLDA05	10.210.12.193	0.06
13/12/2011 13:05	13/12/2011 13:43	3512	NLDA19	10.210.12.176	0.63
13/12/2011 13:00	13/12/2011 13:06	3509	NLDA05	10.210.12.193	0.11
13/12/2011 12:58	13/12/2011 13:21	3531	NLDB05	10.210.12.161	0.38

Snapshot data is also recorded at 10-minute intervals, giving an overview of the total number of PC's in use at each interval.

Table 4-4: Sample snapshot data (December 2011)

Lab Name	Snapshot Time	Total Computers	Computers In Use	Computers Offline
Nursing Library Dept.	13/12/2011 13:40	29	9	0
Nursing Library Dept.	13/12/2011 13:30	29	11	0
Nursing Library Dept.	13/12/2011 13:20	29	12	0
Nursing Library Dept.	13/12/2011 13:10	29	13	0
Nursing Library Dept.	13/12/2011 13:00	29	14	0
Nursing Library Dept.	13/12/2011 12:50	29	12	0
Nursing Library Dept.	13/12/2011 12:40	29	10	0
Nursing Library Dept.	13/12/2011 12:30	29	12	0

This information yields two valuable statistics: (1) Equipment usage profile, and (2) Inferred occupancy profile (see Figure 4-11). While this information only applies to the PC labs in particular, it does present an opportunity to collect detailed measured data where no information was previously readily available. As well as occupancy, the data provides trends which may be used to determine electrical load profiles also, as computers and personal laptops make up a significant proportion of the dynamic electrical load in the building.

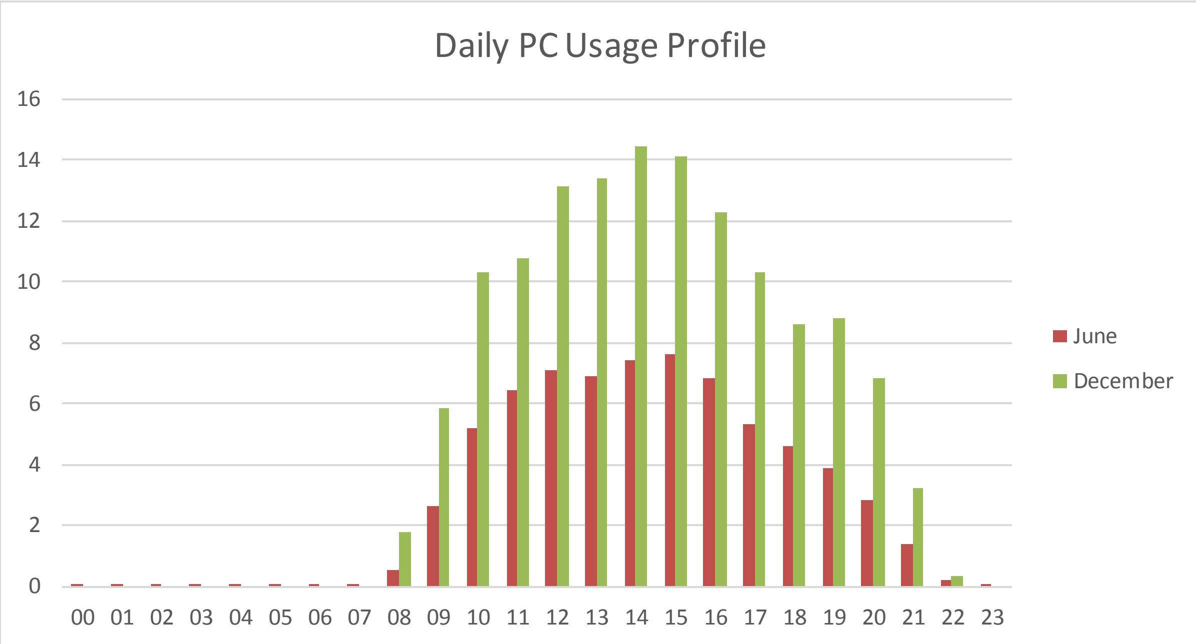


Figure 4-11: Daily PC usage profile (LabStats)

It is also useful to examine how the profile differs between weekdays and weekends (see Figure 4-12 and Figure 4-13). Here it is possible to identify key trends in PC usage, in terms of peak usage times and average number of PC's in use at various time during the day.

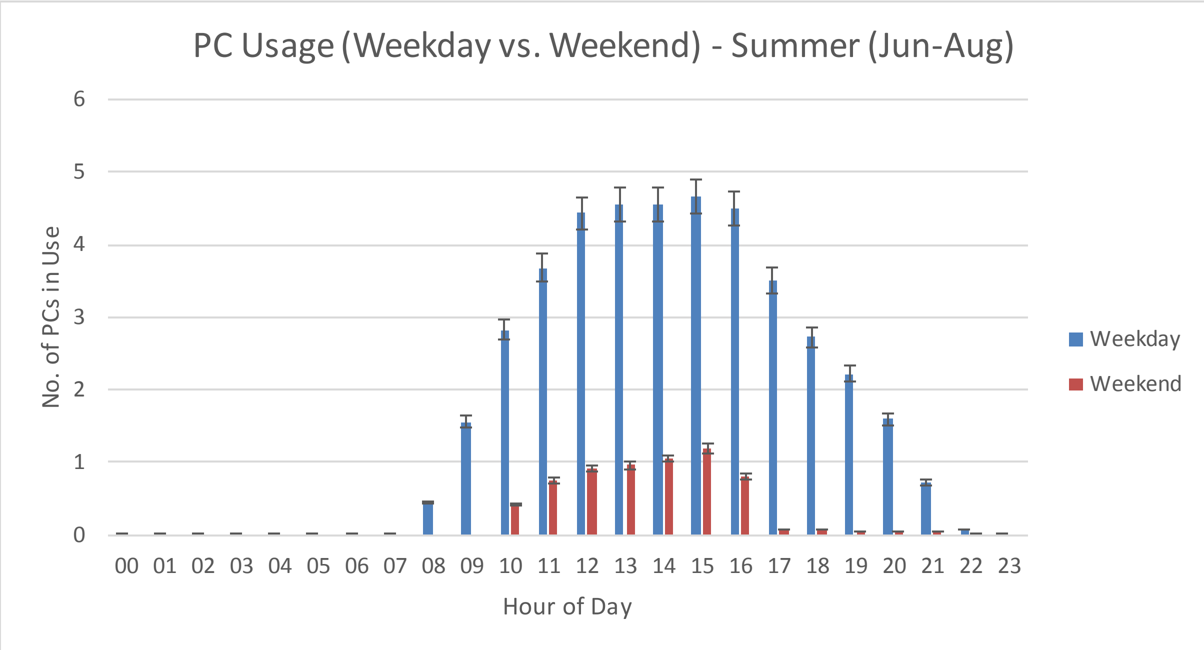


Figure 4-12: PC usage during summer period

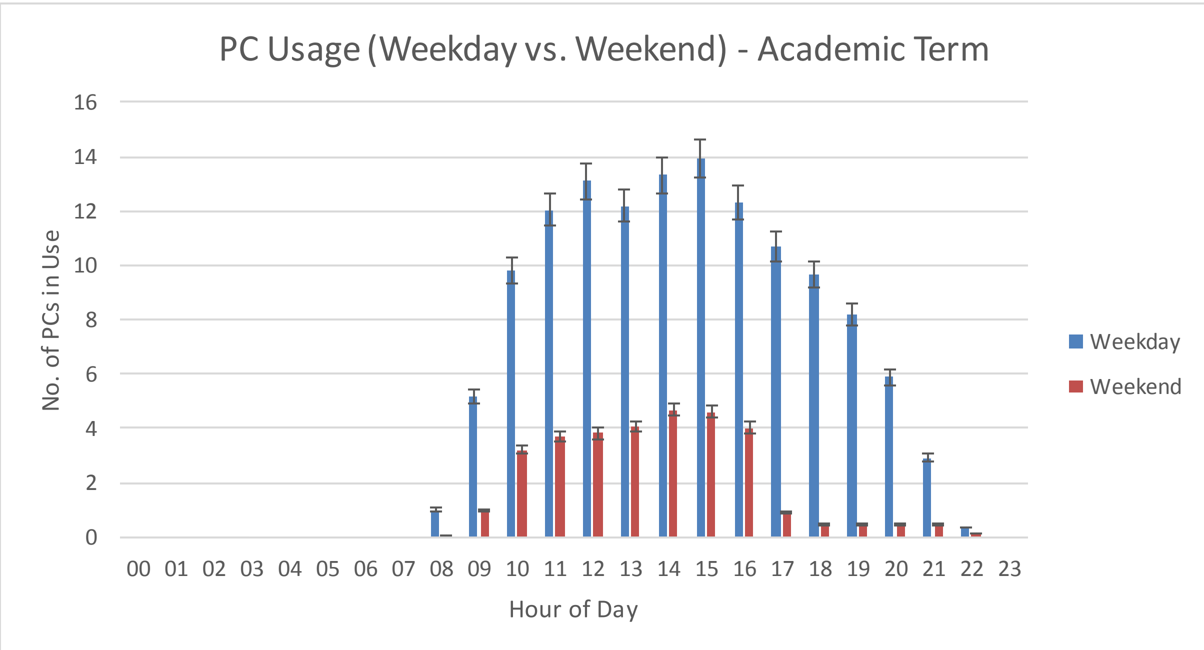


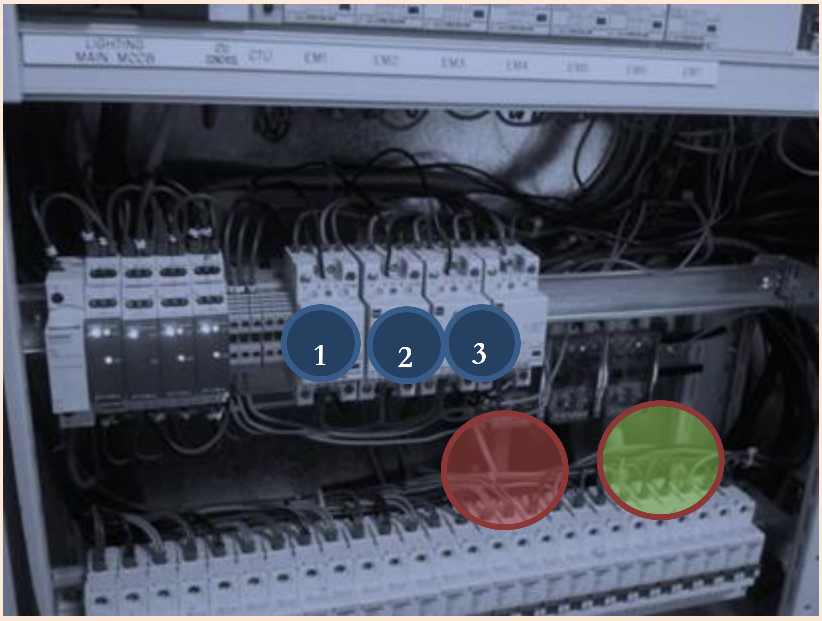
Figure 4-13: PC usage during Academic Term

4.4.1.2 Electrical Loads

At present, electrical energy consumption is only metered at the whole building level. There is no sub-metering of individual end-uses (e.g. lighting, plug loads, HVAC). Therefore, it would be difficult to disaggregate these manually, to provide inputs to a detailed energy simulation model. Furthermore, instrumenting the building to gather this level of data would be costly to implement. At the main electrical incomer (see Figure 4-14), a three-phase (R/S/T) power supply provides power to the building, split into three categories:

- Lighting;
- Power[4] (e.g. AC units, security systems and door locks etc.);
- General Services (all plug sockets);
- Sub-Distribution (e.g. MCC for plant room, Lift, fire alarm).

Services are also further split by floor (Basement, Ground and First).



1 – Basement
2 – Ground Floor
3 – First Floor

- ORANGE – General Services
- GREEN - Lighting

Figure 4-14: Electrical distribution board

[4] Not visible in this photograph as this split is contained in a separate part of the electrical cabinet.

It is therefore possible to use a process of short-term end-use monitoring (see section 2.5.1.2) to gather high-fidelity data for these end-uses. For this purpose, an e-Tracker energy monitor was used to monitor the incoming power supply over the course of two weeks from 25 August 2011 to 8 September 2011. This period was chosen to provide data for both the summer period as well as the start of the 2011-12 academic term.

Figure 4-15: e-Tracker energy monitor

The e-Tracker is capable of monitoring current across three phases, meaning it is only possible to measure one split across all three phases (R, S and T). An initial check was carried out by measuring the **instantaneous** power consumption across each phase for all splits, by connecting the device across each split in turn. This audit was conducted during a period of low-occupancy in order to provide a representation of baseline power consumption.

Table 4-5: Surveyed Power Consumption (all phases)

Panel	Phase			Power (kW)			Total Power (kW)
	R	S	T	R	S	T	
Lighting	7.70	10.70	9.00	1.77	2.47	2.09	**6.33**
Power	0.10	0.40	1.20	0.02	0.09	0.28	**0.39**
General Services	2.10	3.00	2.40	0.48	0.69	0.56	**1.73**
Sub-Distribution							**1.39**
MCC (Plant Room)	1.10	3.10	1.00	0.25	0.72	0.23	1.20
Lift	0.30	0.20	0.20	0.07	0.05	0.05	0.16
Fire Alarm	0.10	0.00	0.00	0.02	0.00	0.00	0.02
TOTAL	**11.40**	**17.40**	**13.80**	**2.62**	**4.02**	**3.20**	**9.84**

This provides an initial breakdown of the main energy end-users in the building (see Figure 4-16). Lighting is the biggest energy consumer, accounting for approximately 64% of the total power consumption at the time of auditing. General services account for around 18%. However, it should be noted that this audit was conducted during a period

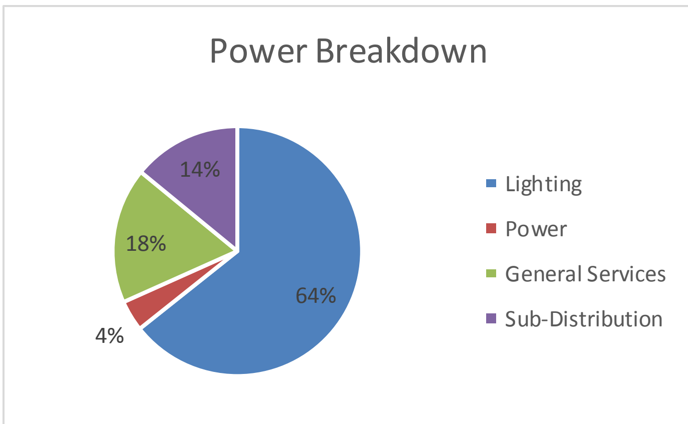


Figure 4-16: Power distribution breakdown

of low occupancy, with relatively few computers or laptops in use. Sub-distribution, which includes always-on systems such as fire alarms, plant room power and emergency lift, accounted for 14%. Finally the power split, which includes air-conditioning units in the PC rooms, security systems and automated door locks, accounted for just 4% of the total power consumption. While lighting accounts for the most dominant distribution load, it is also relatively consistent. Room lighting is always-on once the building is open, while emergency lighting remains on 24/7. Sub-distribution and power loads, by their nature, are always consistent throughout operation. In contrast, general services is highly dependent on occupancy, and thus the most variable. Therefore, it was decided to monitor the general services split for a two-week period, to evaluate variation.

(a) Short-term energy monitoring

The e-Tracker energy monitoring system was installed on 25 August 2011 to monitor individual plug loads (general services circuit) for the Nursing library. Clamp-on meters are connected across each phase (R, S and T) of the plug circuits (as per Figure 4-17) and the unit is set to log the current (I) across each phase at 5-minute intervals. Reference voltages (V) across each phase are also measured in order to convert logged current to power (P) using Equation 4.1.

Figure 4-17: e-Tracker clamp-on meters

$$P = IV \tag{4.1}$$

The results of this audit are summarised in Figure 4-18 below.

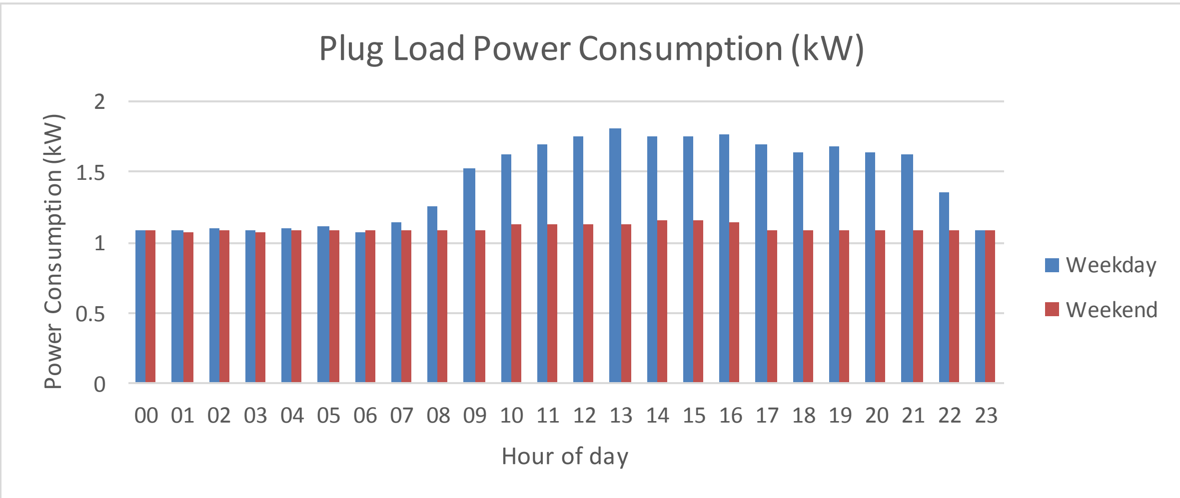


Figure 4-18: Plug load power consumption (weekday vs. weekend)

It is clear that there is very little variation in plug loads over the surveyed weekends, with only a slight increase (<5%) over baseline values. This is due to a significant difference between weekend and weekday occupancy, as evidenced in section 4.4.1.1

(b) Equipment Audit

An equipment audit was also conducted to gather detailed information about the internal loads in the building (see Table 4-6). The following information was logged for each piece of equipment in the building:

- Number of Units;
- Description (incl. model numbers, if applicable);
- Location;
- Rated unit power consumption (W).

Table 4-6: Equipment Audit

No. Units	Description	Location	Unit Power [W]	Total Power [W]	Max Power [W]
22	Dell Optiplex 760	PC Room - Downstairs	255	5610	
22	E190Sb Monitors	PC Room - Downstairs	19	418	
1	HP P4015x	PC Room - Downstairs	20	20	840
5	Dell Optiplex 760	Group PC Room	255	1275	
5	E190Sb Monitors	Group PC Room	19	95	
1	Dell Optiplex 760	Group Study Room 1	255	255	
1	E190Sb Monitors	Group Study Room 1	19	19	
1	Dell Optiplex 760	Group Study Room 2	255	255	
1	E190Sb Monitors	Group Study Room 2	19	19	
1	Elevator	NA	2000	2000	
2	SyncMaster 930XT	First Floor	20	40	
1	Dell Optiplex 760	First Floor	255	255	
1	E190Sb Monitors	First Floor	19	19	
1	Mitsubishi LDT421V	First Floor	230	230	
1	3M Touch Check-In	First Floor	20	20	
1	Dell Optiplex 760	Print Station	255	255	
1	E190Sb Monitors	Print Station	19	19	
1	HP P4015x	Print Station	20	20	840
1	Konica 7222	Print Station	200	200	1500
			TOTAL	**11024**	

4.4.1.3 Lighting

An extensive lighting audit was carried out in order to determine the exact level of lighting fixtures in the building. Most of the study and office spaces are equipped with 36W T8 fluorescent lamps, while hallways and stairwells utilise smaller 18W quad-tube PL lamps.

Figure 4-19: Typical lighting fixture and close-up of T8 fittings (inset)

Table 4-7: Lighting Audit Information

Floor	Location	Type	No. Lighting Fixtures	No. Fittings	Rated Power (W)	Total Power (W)
Basement	PC Room	T5	12	2	36	864
	Open-Plan	T5	16	2	36	1152
Ground Floor	Office	T5	2	2	36	144
	Print Room	T5	2	2	36	144
	Open-Plan	T5	26	2	36	1872
First Floor	Study Room 1	T5	8	2	36	576
	Study Room 2	T5	4	2	36	288
	Study Room 3	T5	4	2	36	288
	Open-Plan	T5	22	2	36	1584
All Floors	Hallways	PL	10	2	18	360
(Emergency)	Stairwells	PL	7	2	18	252
					Sub-Total	**7524**
					Emergency	**612**

The lighting schedule coincides with the building opening hours, activating approximately one hour before opening and switching off one hour after closing. Emergency lighting remains active 24 hours a day. The building and lighting schedule are as follows:

Table 4-8: Lighting Schedule

Day	Building Schedule		Lighting Schedule	
	Open	Close	On	Off
Monday-Friday	8:30:00 AM	10:00:00 PM	7:00:00 AM	11:00:00 PM
Saturday	8:30:00 AM	5:30:00 PM	8:00:00 AM	6:00:00 PM
Sunday	10:00:00 AM	5:30:00 PM	8:00:00 AM	6:00:00 PM

4.4.1.4 HVAC Operation

The operation schedule for the heating, ventilation and air-conditioning (HVAC) system for the building, is controlled by a central building management system (BMS). This provided an accurate source of schedule information pertaining to the operation of these systems (see Figure 4-20)

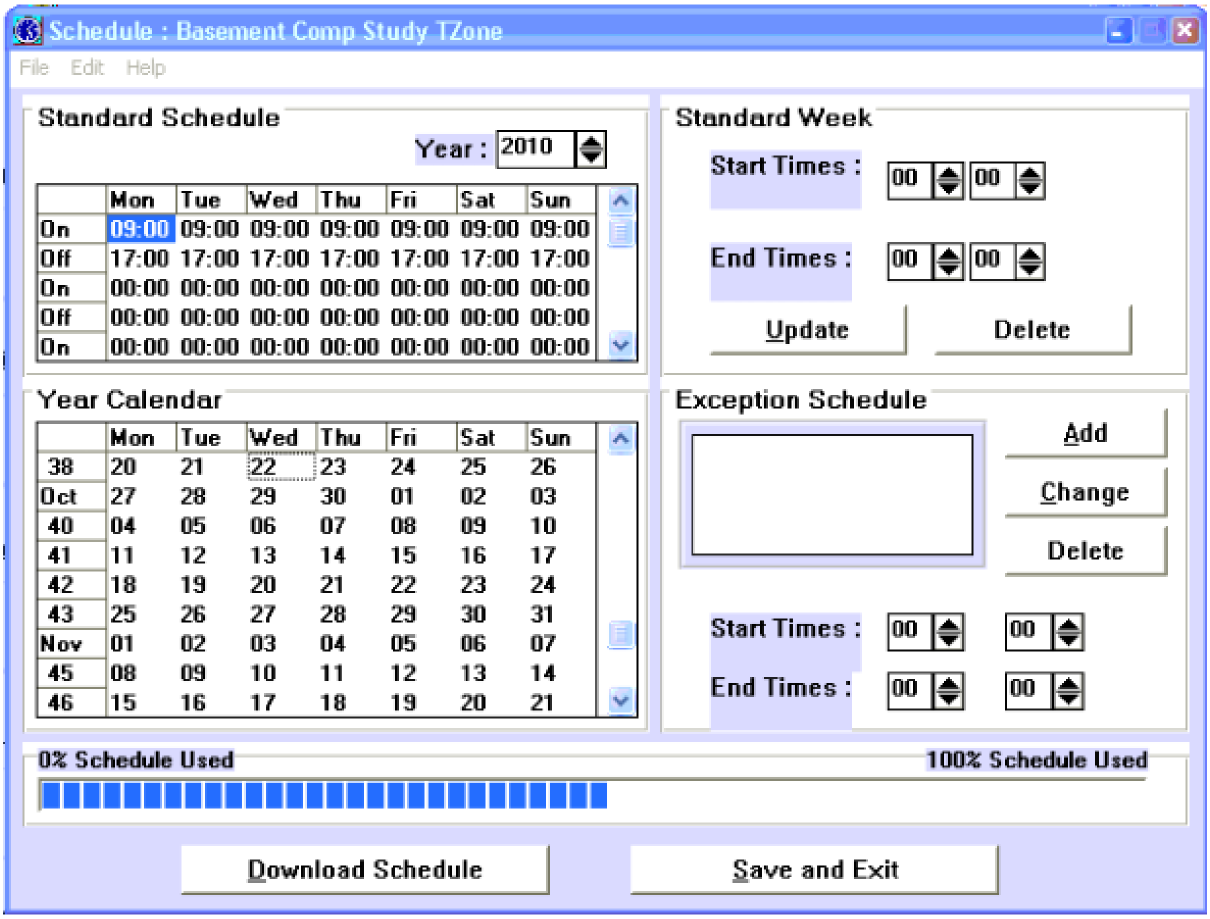


Figure 4-20: HVAC operation schedule for basement computer lab and study space

The HVAC system in the building typically operates on a standard schedule of 09:00-17:00 daily (Monday-Sunday). Exceptions may be added using the 'Exception Schedule' in order to account for holidays or study periods.

4.4.1.5 Natural Ventilation

The building is mixed-mode, meaning that it uses a combination of natural and mechanical ventilation. While the mechanical ventilation is operated automatically by the BMS controllers, the natural ventilation strategy uses a combination on manual and automated windows. Automated windows are controlled by the BMS using window actuators to open/close windows based on ventilation demand. However, this automated procedure was de-activated for the duration of the calibration period due to security issues due to problems controlling operation.

Instead, occupants had control over manually-operable windows in each space. In order to capture the level of natural ventilation, window positions (*W*) were recorded during site audits at various times of the day and at different times during the year (see Appendix C.2). In the absence of any other information pertaining to levels of natural ventilation in the building, this provided some indication as to the level of ventilation at different times of day:

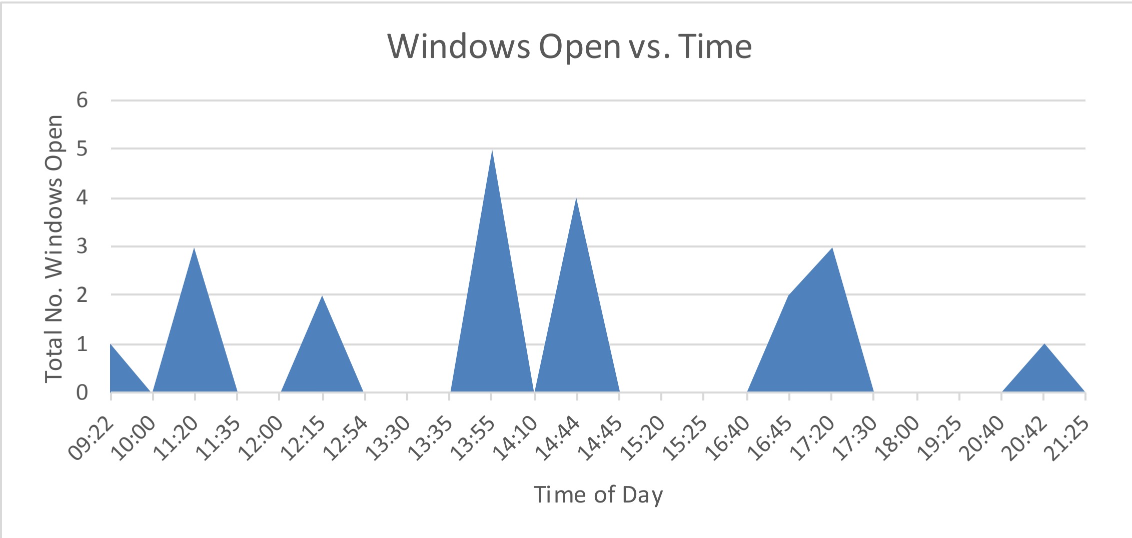


Figure 4-21: Windows open vs. Time of Day (Refer Occupancy Audit - Appendix C.2.)

While the occupancy audit provided some very useful information pertaining to occupancy patterns, it proved difficult to provide any deterministic pattern to the human-controlled natural ventilation strategy in the building. There are a number of reasons for this:

- Human behaviour in response to their thermal environment is a complex issue, dependent on a number of personal and environmental criteria (see section 2.2.1), in addition of personal psychological and behavioural characteristics of the individual;
- The level of natural ventilation, even with accurate values for number of windows open is also notoriously difficult to predict, as it will vary depending on: wind speed, shading and orientation, angle and direction of wind, angle of window opening etc.;

Due to the relative complexity of this issue, it was deemed to be outside of the scope of this study. However, there a number of research papers which attempt to quantify levels of natural ventilation under such uncertainty, which provide an excellent basis for a follow-up study in this area (Brager et al. 2004; Hyun et al. 2008; Haldi and Robinson 2011). Further discussion on this topic can be found in the future work section of this thesis (see Section 6.3.5).

4.4.1.6 Weather Data

In order to accurately simulate the building response to external conditions, it was necessary to gather detailed weather data for the full simulation period. Weather data from the campus weather station[5] was used for this purpose. This measures the following (to the specified accuracy, in parentheses):

- Dry-bulb temperature (± 0.5°C);
- Relative humidity (± 2%);
- Barometric pressure (± 50Pa);
- Wind speed [±0.1ms^{-1} (0.3 – 10ms^{-1}); ± 1% (10 - 55ms^{-1}); ± 2% (> 55ms^{-1})];
- Wind speed -3s gust [(±0.1ms^{-1} (0.3 – 10ms^{-1}); ± 1% (10 - 55ms^{-1}); ± 2% (> 55ms^{-1})];
- Wind direction [± 2% (>5ms^{-1})];
- Global solar irradiance (<5%);
- Diffuse solar radiation (<15%);
- Barometric pressure (± 50Pa).

Figure 4-22: NUIG weather station

All data is logged at one-minute intervals, with the exception of rainfall data, which is logged at hourly intervals. This is logged to a remote server and downloaded via FTP (File Transfer Protocol) on a weekly basis (see Appendix A.1.1). This data is then transferred to a MySQL database, by means of SQL script for analysis (see Appendix A.1.2 and A.1.3). Weather data is then converted to EPW format using the EnergyPlus 'Weather Converter' software (US Department of Energy 2011a). This software uses input data to compute missing values for:

- Dew Point Temperature (°C);
- Direct Normal Radiation (Wh/m2);
- Illuminance (lux);
- Sky Cover.

[5] NUIG weather station, See http://weather.nuigalway.ie.

The entire weather data conversion process is summarized in Figure 4-23.

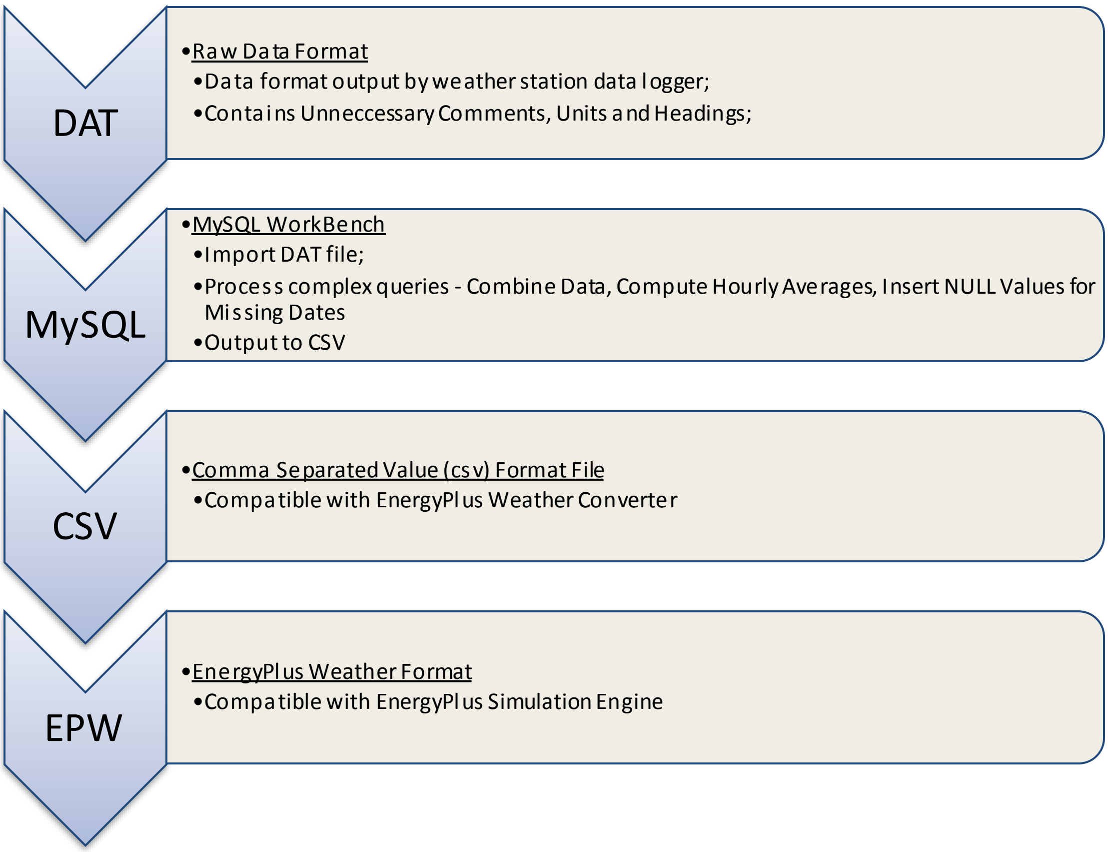


Figure 4-23: Weather Data conversion process

At present, this process requires a significant degree of data transformation from source to simulation, as is the case with much of the model input and calibration data outlined in this study. However, as part of this thesis, this process has been automated as much as possible through the use of batch scripts (executable as scheduled tasks in Windows OS). In addition, a number of checks for data consistency have been included in the MySQL script, prior to conversion to EPW (EnergyPlus Weather) format.

4.4.2 Model Calibration Data

Model calibration data incorporates any data which will be used to calibrate and validate the final Building Energy Simulation (BES) model. In this case, since we are attempting to validate a detailed calibration methodology, the model will be calibrated to high fidelity data for:

- Heat Energy Consumption, *kWh* (Daily);
- Electrical Energy Consumption, *kWh* (Hourly);
- Zone Temperature, °C (Hourly).

This data is attained from the Building Management System (BMS). This data is recorded daily on a remote BMS server archive. Each individual sensor data point is stored in a separate comma-separated value (.csv) file with a unique identifier. Data is recorded using the following format:

T_1, j, $u_{1,1}$, $u_{1,2}$, …, $u_{1,j}$ T_2, j, $u_{2,1}$, $u_{2,2}$, …, $u_{2,j}$ … … … T_i, j, $u_{i,1}$, $u_{i,2}$, …, $u_{i,j}$	**Where**: T_i, Timestamp at row i, interval 1; j, Number of samples = 1024; $u_{i,j}$, Sensor value at row i, interval j;

With over 60 individual sensors for this building recording 1024 values daily to separate csv files, this resulted in approximately 22 million data points over the course of 12 months, with approximately 90% of this information duplicated due to overlapping time periods. A C-script program was written for initial pre-processing and loading this data into a MySQL database (see Figure 4-24). This program reads the data from each csv file and uploaded this data to the MySQL database, removing any duplicate values in the process. Once the data was loaded into a database, it was then possible to carry out further pre-processing and data consistency checks. Further detail on this process can be found in Appendix A.3.

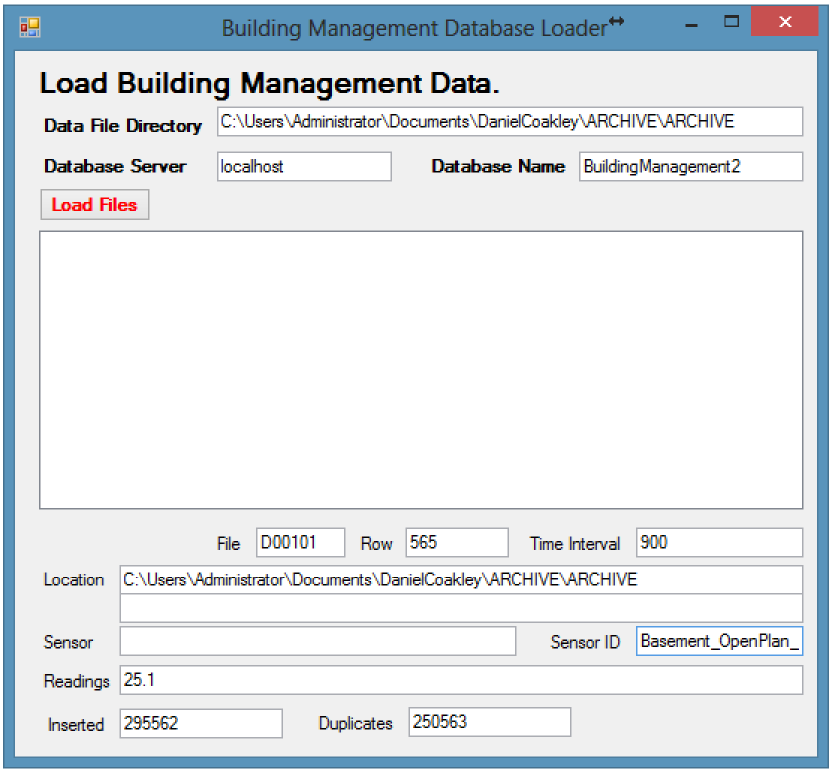


Figure 4-24: Building Management Data Loader

Before comparing measured and simulated data, it is first necessary to proof the data and perform any necessary pre-processing steps. Measured data commonly contains one or more of the following deficiencies; unnecessary fields, inconsistent data, missing time periods, and, missing information. As a general rule, where data is missing or incorrect for short time-periods (<6 hours), simple interpolation was used to generate missing values. Where long-term data is missing, these periods were excluded from final goodness-of-fit (GOF) calculations (i.e. no comparisons are carried out for these periods).

4.4.2.1 Heat Energy Consumption

Measured heat energy consumption from the low-pressure hot-water (LPHW) meter is recorded using cumulative daily, weekly and monthly totals as well as current rates. Unfortunately, heat energy consumption is only recorded on the heat meter in increments of 40kWh. This means that it is impossible to compare hourly averages from EnergyPlus with measured heat energy consumption. Thus, for the purpose of this study, daily totals for heat energy consumption were used in place of hourly values. While this resolution is sufficient for the purpose of the demonstration of this methodology, it may be outside the bounds of acceptable error for the reporting of potential savings in the case of ECM evaluation (see section 6.3.1).

4.4.2.2 *Electrical Energy Consumption*

Measured electrical data from the Building Management System is similarly recorded using cumulative daily (kWh), weekly (kWh) and monthly (kWh) totals as well as current rates (kW). Data from the 'present electrical use' sensor was taken and averaged to compute an hourly power consumption (kWh). Simulated electrical data in EnergyPlus is output in Watt-hours. This was also converted to hourly average power consumption (kWh).

4.4.2.3 *Zone Temperatures*

Since this study uses a single-zone BES model, it was necessary to compute a single averaged zone temperature from the measured room temperatures in the building. It was decided that the best approach here was to use a floor area weighted average. Areas such as storage rooms, hallways and stairwells are excluded from this list as they are not currently monitored or controlled by the building management system (BMS). A summary of these weightings is presented in Table 4-9.

Table 4-9: Room weightings

Room	Area (m^2)	Weighting
Basement Computer Room	56.44	0.1284
Basement Open Plan	78.63	0.1788
Basement Comms Room	5.47	0.0124
GFlr North Library	77.45	0.1761
GFlr South Library	46.49	0.1057
GFlr Copy Room	7.35	0.0167
GFlr Office	18.13	0.0412
FFlr PC Room	15.86	0.0361
FFlr Study Rm 6	10.80	0.0246
FFlr Study Rm 8	16.00	0.0364
FFlr Open Plan	107.07	0.2435
Total Floor Area	439.69	1.0000

4.4.3 Data Proofing and Assessment

The next step involves checking the quality of the measured data which will be used for the purpose of model calibration, as well as classifying the various data sources. It is important that this data is consistent as it will be used to test the validity of the final model. This step is also useful in familiarising oneself with the building and its load profile.

4.4.3.1 *Heat Energy Consumption*

The first check is heat energy (kWh) consumption. As mentioned in 4.4.2.1 , this output is recorded in 40kWh increments on the BMS. In this case, it would be inaccurate to use an hourly rate as the data resolution is insufficient. Therefore, daily heat energy consumption was used instead. An initial audit of the heat meter data from the BMS revealed some discrepancies in the logged flow and return temperature for the low-pressure hot water supply. During the evening, the meter was recording temperature spikes of up to 105°C (see Figure 4-25).

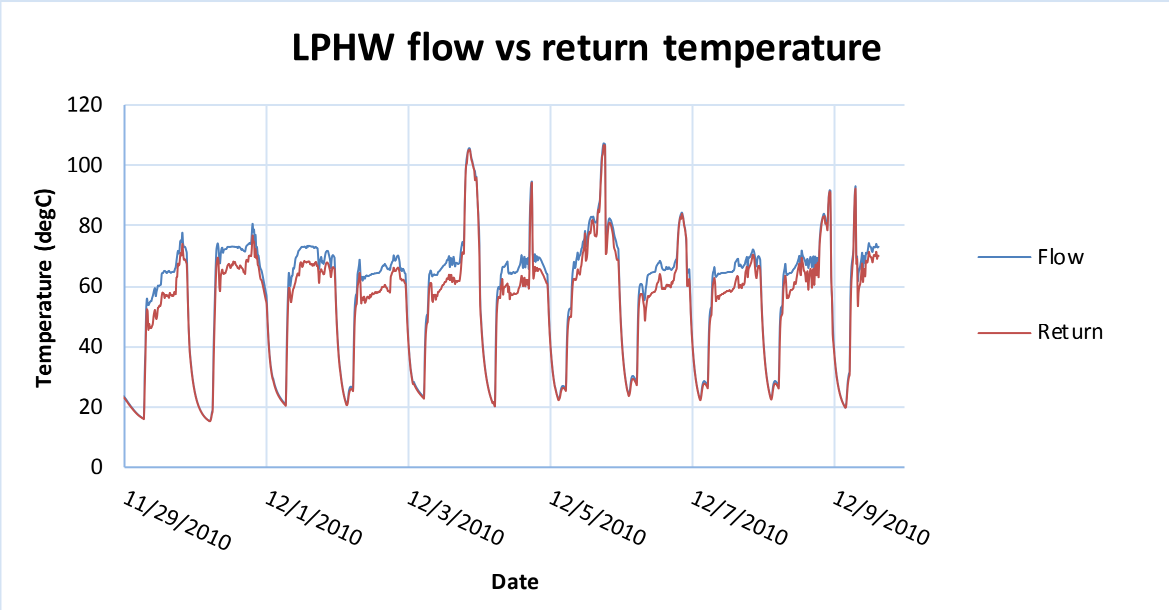


Figure 4-25: Anomaly in flow vs. return temperature for LPHW meter

Further inquiries to the buildings and estates office revealed two issues:

- Heat Meter Calibration: The heat meter had not been calibrated after installation, meaning some of the readings may have been inaccurate. Calibration was completed on 15 December 2011;
- Stuck Flow Valve: The building was serviced by a three-way valve supplying hot water to the adjacent main library building. However, this valve had seized causing excess hot water to be supplied to the nursing library during periods of low demand.

Following calibration, the data from the heat meter was examined again. In addition, wired thermo-couples (HOBO sensors) were used to measure the temperature at the surface of the flow and return headers. This data was used to verify the output from the calibrated heat meter. Figure 4-26 and Figure 4-27 present the total cumulative daily heat energy consumption for the building for the period 01/01/2012 – 31/12/2012. The high level of correlation between heat energy consumption and outside air temperature is clearly evident here.

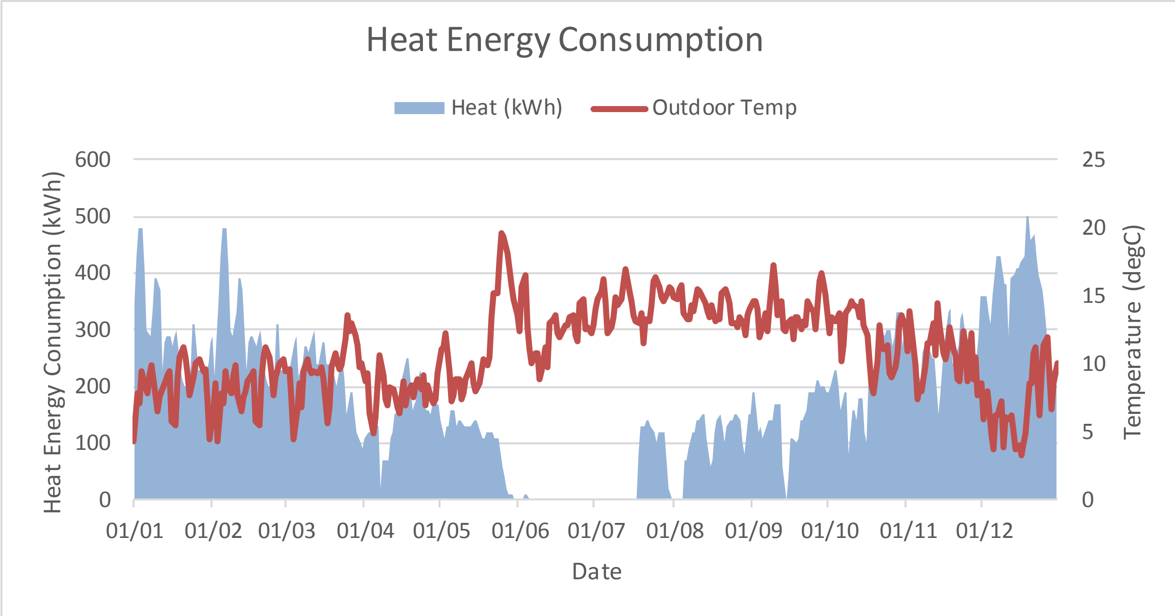


Figure 4-26: Daily heat energy consumption

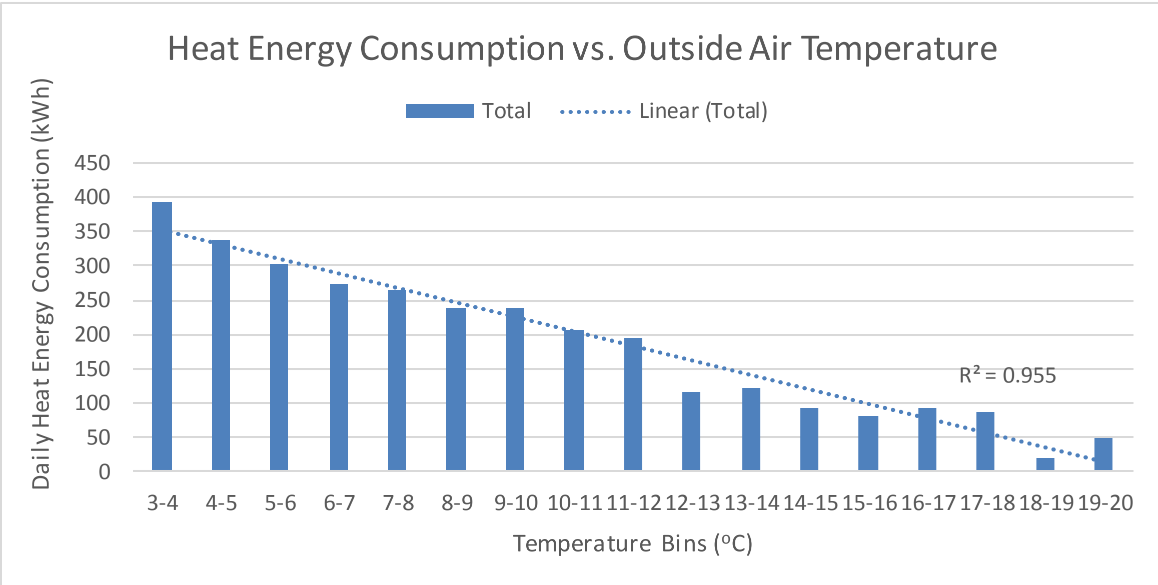


Figure 4-27: Correlation between heat energy consumption and outside air temperature

4.4.3.2 Electrical Energy consumption

This section will examine the overall trends in measured electrical energy consumption (kW). This check is also useful in determining the buildings electrical baseload, the load profile and also electrical equipment schedules. It is useful to examine the recorded daily average electrical energy consumption to identify unexpected dips/peaks. Figure 4-28and Figure 4-29 present a summary of this data.

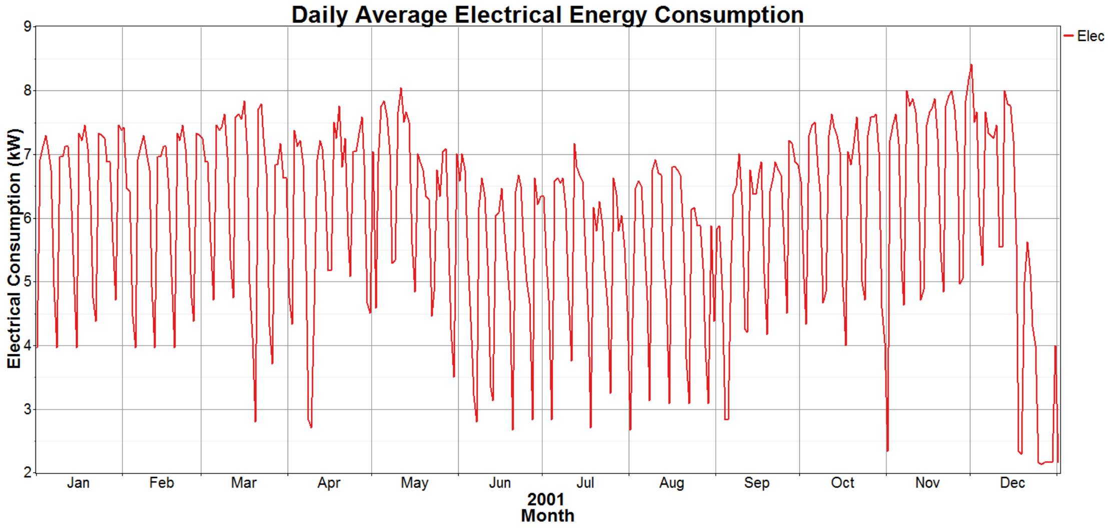


Figure 4-28: Daily average electrical power consumption (kW)

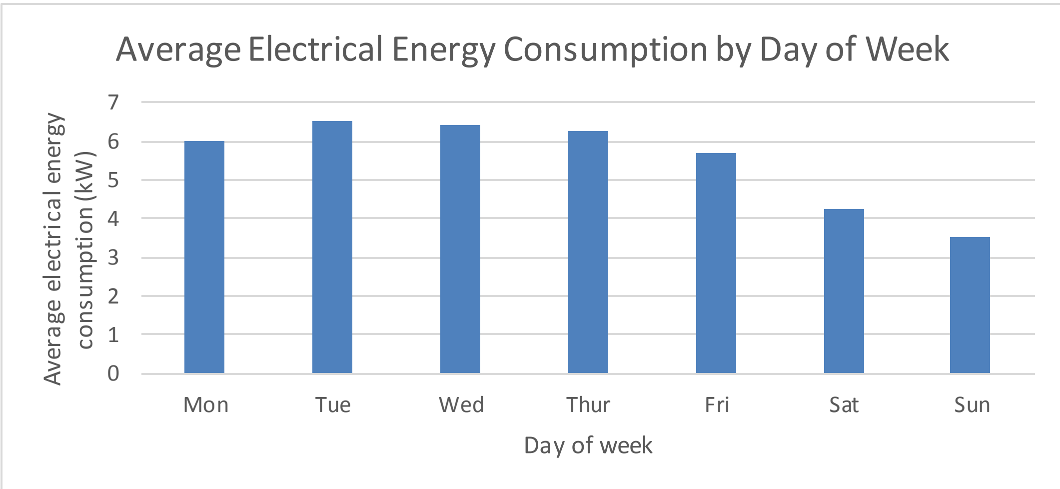


Figure 4-29: Average Electrical Energy Consumption by Day of Week

On inspection, the following details became evident:

- Overall electrical load profile remained largely consistent throughout the year with a slight drop during the summer months (Jun/Jul/Aug) when most students are on holidays;
- Since occupant presence has not had a significant impact on the yearly load profile, it was inferred that other electrical consumers (lighting and equipment) are the most influential factors. An earlier site visit revealed that the main variable electrical load are laptops and mobile phone chargers used by students;
- There was a significant dip in late December. However, this was expected as the building was closed on these dates;

- From examination of individual weeks in more detail, a more variable electrical load profile above the building baseload is evident. Peaks tend to occur on Mondays with lowest demand occurring on Friday. However, the relative difference is minimal. There is also a 30-40% drop in electrical demand at weekends, with lowest dips (2-3kW) occurring on Sunday (see Figure 4-29).

Figure 4-30 and Figure 4-31 compare weekday and weekend electrical energy consumption, for academic (01/01 to 31/05, 01/09 to 31/12) and summer (01/06 to 31/08) terms respectively.

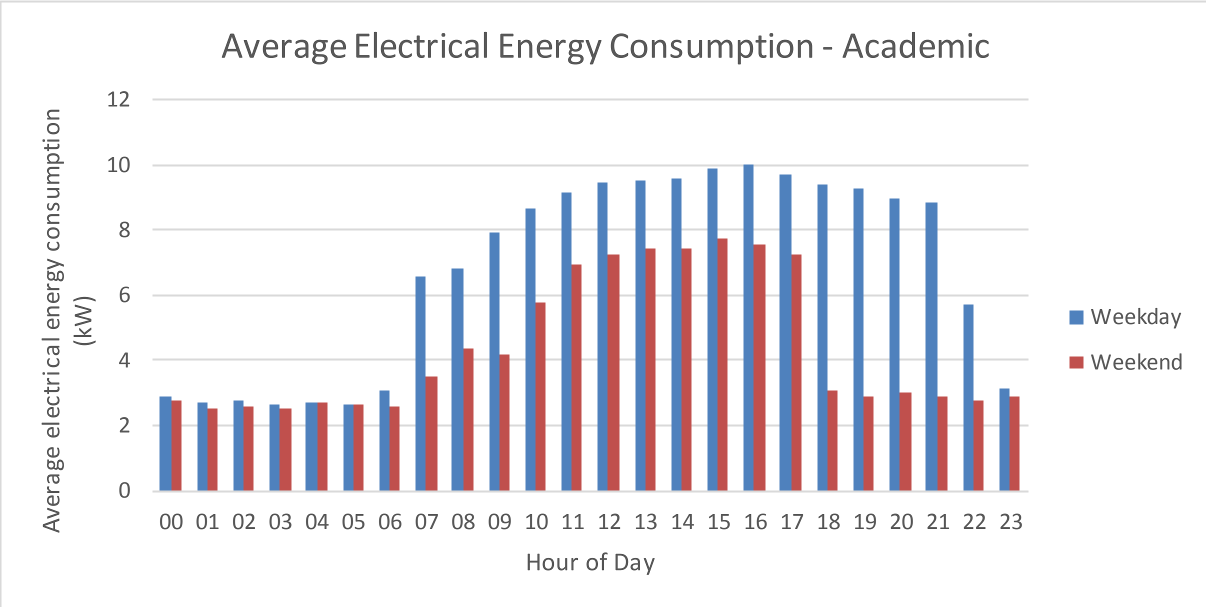


Figure 4-30: Hourly electrical energy consumption (kW) profile – academic term

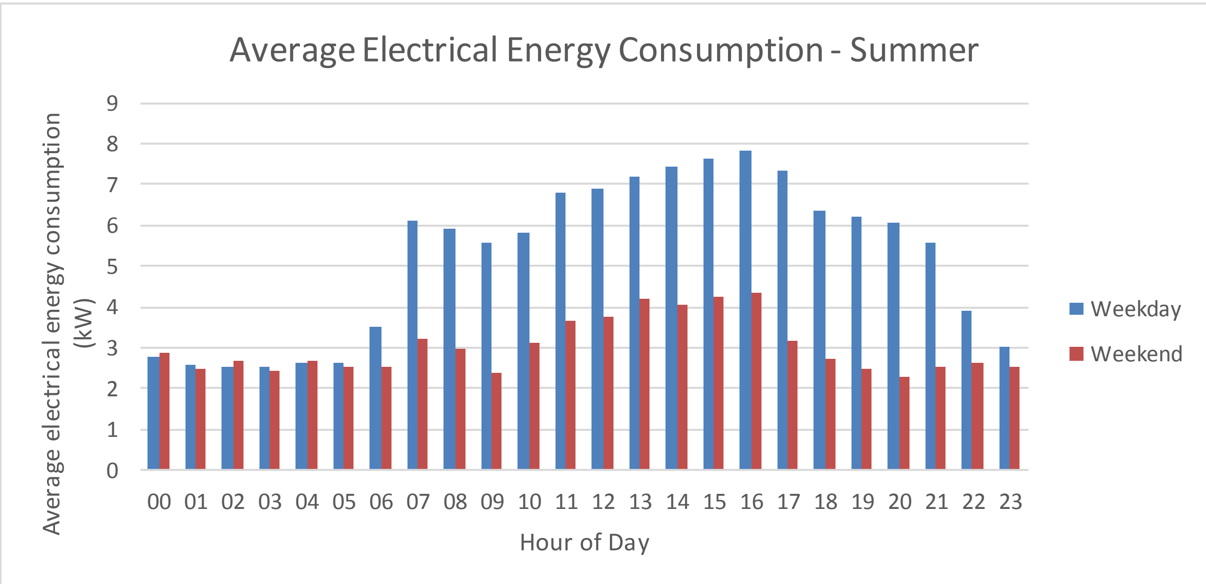


Figure 4-31: Hourly electrical energy consumption (kW) profile - summer

From the above graphs, the difference between academic and summer-time electrical load patterns is clearly evident. Summer-time peak loads are approximately 20% lower that the equivalent academic term peaks. In addition, weekend loads are far lower, peaking at just over 20-30% above baseline energy consumption. The Summer-time load pattern is also less defined, with gradual growth and decline, compared with a distinct rise and drop in consumption during the academic term.

4.4.3.3 Scheduling

A carpet plot is useful to help identify or confirm electrical equipment schedules as well as providing a useful tool for fault detection (i.e. equipment scheduling faults). Figure 4-32 provides an excellent visual summary of this information. It is possible to identify the on/off times for the main building equipment (07:00 – 23:00 on weekdays, 08:00-18:00 on Saturdays, 10:00-18:00 on Sundays). Periods with anomalous data are also clearly visible (as highlighted in red).

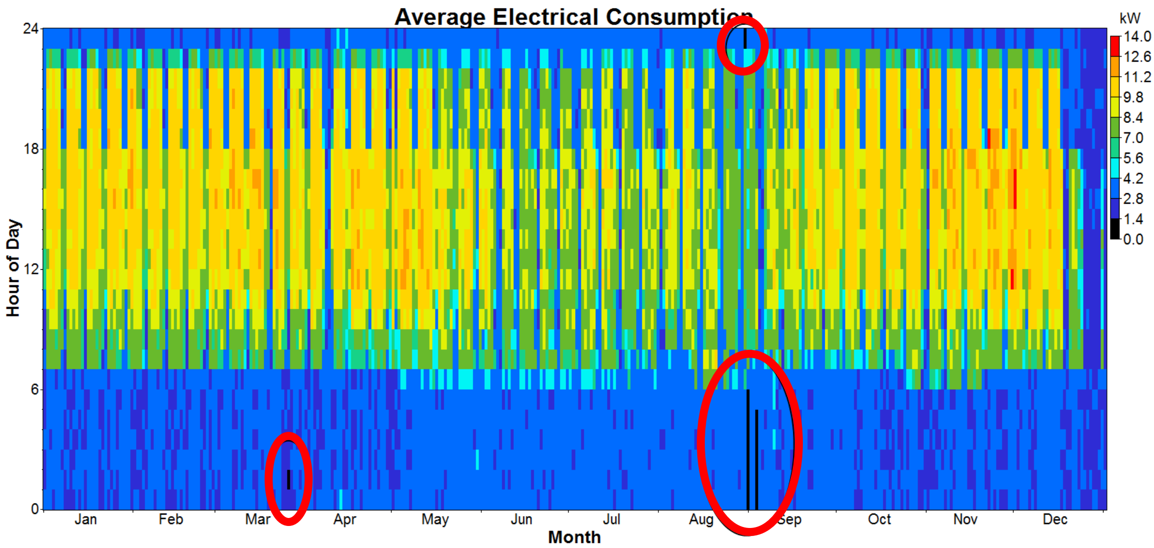


Figure 4-32: Carpet plot of measured electrical power (kW) consumption

4.4.3.4 Zone Temperatures

The final check is yearly temperature data. This is recorded at 15-minute intervals and averaged over each hour. An initial area plot of the daily average room temperature highlights two distinct dips in average room temperature in August/September.

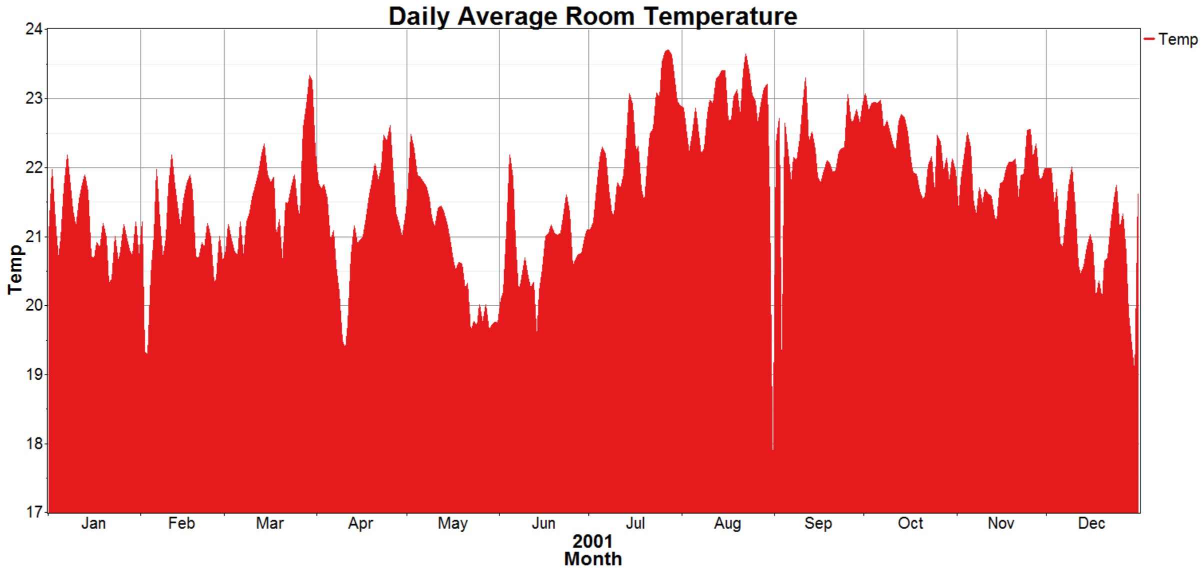


Figure 4-33: Plot of Average Measured Room Temperature (deg C)

Further investigation, by means of a Box-whisker mean plot provides a more revealing statistical summary of the data. It is now clear that there are anomalous outliers in the months of March, August and September. It should be noted that the BWM plot does not provide any insight as to the exact timing of these anomalies.

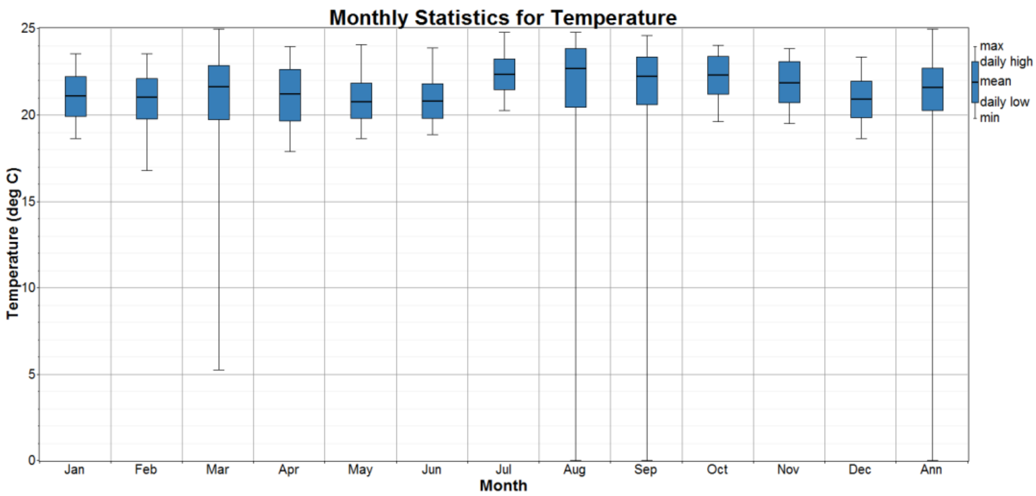


Figure 4-34: Monthly statistics for Measured Room Temperature (BWM Plot)

A final check is carried out by means of a carpet plot (or heat map). Average Room Temperature is plotted against 'Hour of Day' and 'Month of Year'. This provides the most intuitive graphical representation of the measured temperature data. The anomalous data points are immediately identifiable, as highlighted by the red markers in Figure 4-35.

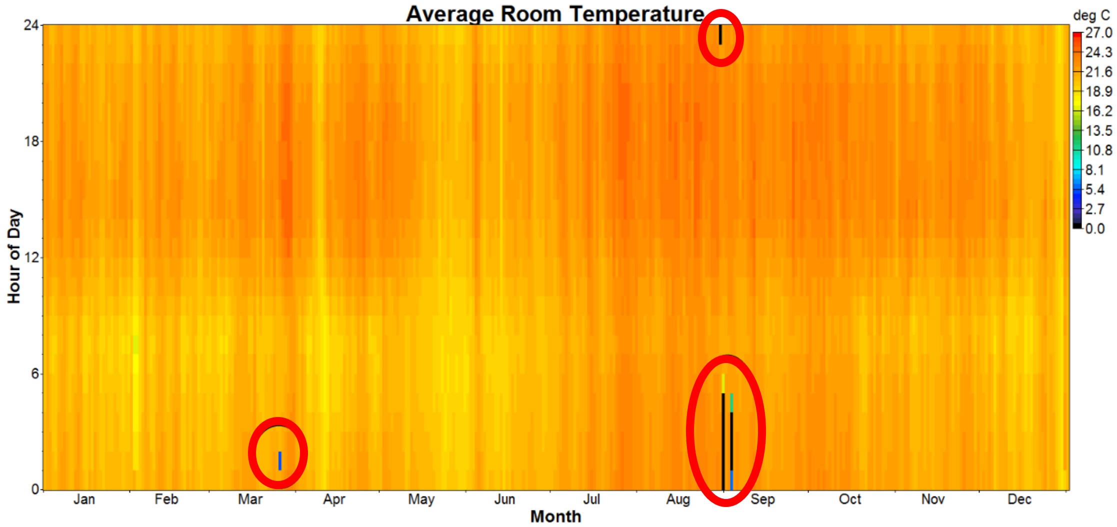


Figure 4-35: Heat Map of Measured Average Room Temperature

4.4.4 Data Collection Discussion

This case study highlighted a number of important points in relation to data collection for detailed BES model development and calibration. Firstly, in relation to **data quality**, it is evident from this that there may often be errors in measured data. These errors may be present due to a combination of the following;

- **Sensor accuracy**: In some cases, sensors may not record data to the level of detail required for model calibration. For example, in this building, the heat meter was only capable of logging heat energy consumption at increments of 40kWh.
- **Sensor Calibration**: Sensor 'drift', or degradation of measurement accuracy, is a common problem in buildings, requiring the regular calibration of sensing equipment in order to maintain optimum performance, in terms of sensor accuracy. However, in many cases, this sensor calibration is not carried out regularly, leading to inaccurate values being logged.
- **Mis-mapping of Data**: Building management systems (BMS) provide an efficient means of automatically maintaining defined conditions within buildings. However, the specification of BMS operation strategies requires a high level of knowledge and expertise, particularly in complex buildings. In some cases, sensors may be mapped incorrectly, or have incorrect measurement units specified during initial set-up. This may be difficult to diagnose, as it may not be initially evident from the BMS data.

Secondly, in relation to data collection and transformation, this study has highlighted a number of issues with this process in terms of gathering the high level of data required for BES calibration:

- **Lack of available information**: Detailed BES tool, such as EnergyPlus, require high levels of information relating to all aspects of building design and operation. However, due to high levels of fragmentation in the process of building design, construction and commissioning, this information is often lost or not sufficiently documented.
- **Data format**: Building data is typically recorded in systems or software using proprietary data formats. Therefore, access to this data relies on users having access to the systems as well as the expertise required for their operation;
- **Data archiving**: Building data is typically only used for operation, and is therefore not archived for long periods, unless specified by the client. Since BES models require a minimum of 12 months data for calibration, it is often not feasible to employ them if an immediate assessment is required;
- **Data Transformation**: Finally, transformation of data to formats suitable for model calibration requires significant levels of effort and knowledge. Figure 4-41 and Figure 4-42 at the end of this chapter specify the main elements of the BES calibration tool-chain, highlighting the level of complexity and sheer number of steps required to complete this process.

4.5 Evidence Based Model Development

This methodology follows an iterative evidence-based model development process. This section covers the major milestones at each stage of development as well as the changes or processes required at each stage.

4.5.1 Initial Model

An initial model is constructed to form the basis for future iterations. This was constructed using basic information regarding the building location, geometry and orientation, as gathered from as-built drawings and on-site measurement surveys (Refer 4.4.1). As discussed in 4.3.2, it was decided to take a simplified modelling approach in order to reduce computational expense during the Monte-Carlo simulation stage. In addition, there was no long-term detailed sub-metering available for the building. The initial single-zone model (see Figure 4-36) was developed using Google SketchUp and OpenStudio Plugin for EnergyPlus.

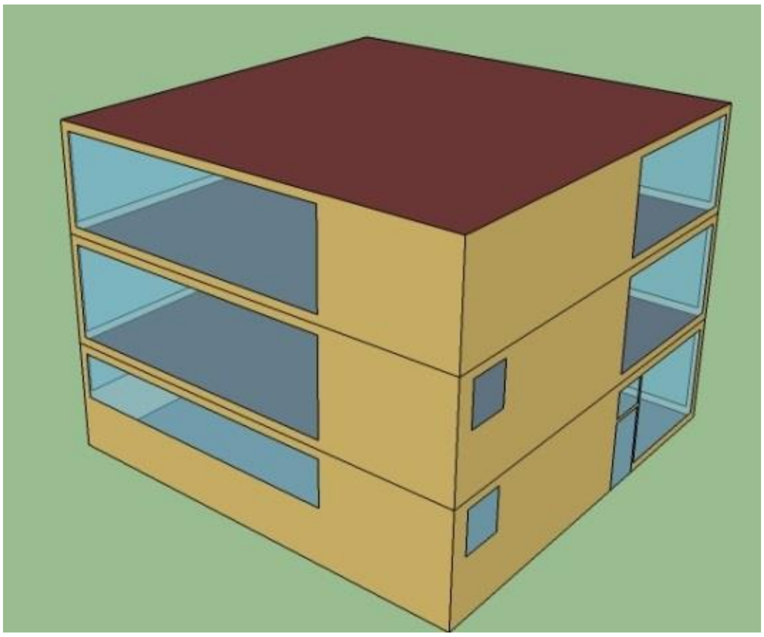


Figure 4-36: Initial SketchUp Model (Note: Three floors are shown, but building is modelled as one zone)

Basic zone information and HVAC details are also added to the model. In this case, HVACGenerator (Raftery et al. 2012) was used to assign HVAC and system variables to the building. This tool outputs macro-format simulation files compatible with EnergyPlus V7.0.

4.5.2 Update Model

The initial model is updated based on currently available information, as per the process outline in section 3.5.2. This update process ensures that all available information is accounted for within the model. Any discrepancies been on-site measurements and as-built drawings was reflected in the updated model, and referenced in the version control repository.

4.5.2.1 Geometry

Model geometry was attained from as-built drawings procured from the Buildings and Estates office. These drawings were checked for accuracy by means of on-site measurements and spot-checks.

4.5.2.2 Construction & Materials

Construction and material information is often difficult to ascertain to a high level of confidence. However, due to the recent construction of the Nursing Library, and relatively high level of available documentation, it was possible to define most construction and material classes with a high level of confidence. Where no data was available for certain materials (e.g. roof cladding, floor covering etc.), properties were derived from standards (British Standards Institution 2000; CIBSE 2007).

4.5.2.3 HVAC and Plant Information

Information pertaining the heating, ventilation and air-conditioning (HVAC) systems was obtained from Operation & Maintenance manuals, commissioning documents and interviews with the Building Manager. This information was added using HVACGenerator (Raftery 2009; Raftery et al. 2012), an excel-based VBA program used to generate all the necessary EnergyPlus objects required to run a simulation (See Figure 4-37).

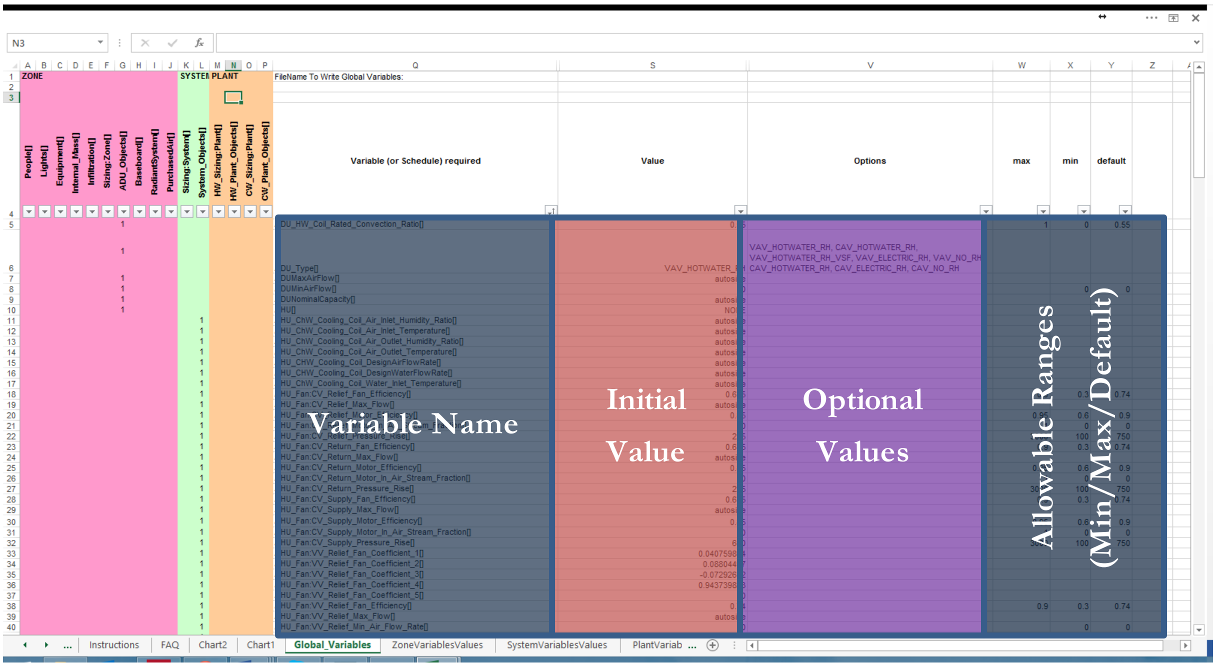


Figure 4-37: HVACGenerator Interface

4.5.2.4 Infiltration and Natural Ventilation

There are two sources of air-changes in buildings:

- Infiltration: unwanted air exchange through leaks and cracks in the building envelope;
- Ventilation: air-changes required or induced by the occupant or building management system (BMS) through vents or windows;

Both of these sources may be modelled in EnergyPlus. In this case study, infiltration was modelled using a simple infiltration schedule specifying the base-line air-changes per hour based on the assumed air-tightness of the building.

Ventilation may be modelled in two ways in EnergyPlus:

- The Zone Airflow group objects provide a simplified method of modeling ventilation based on user-defined assumptions (e.g. flow rate per person (m^3/s/person), flow rate per zone floor area ($m^3/s/m^2$), air-changes per hour (ach)).

- Airflow Network provides more detailed modelling of airflow through windows, cracks, etc. based on temperature differences, wind speeds & pressures, etc.

At the initial preparation stage, it was assumed that the ground heat exchanger would be operational during the course of study, and would therefore need to be modelled. At that time, the ground heat exchange module within EnergyPlus was not compatible with the '*Airflow Network*' module, so the decision was taken to use the *Zone Airflow* method instead. An airflow was specified using the 'Flow/Person' method, specifying approximately 10 l/s/per person, as per the recommended thermal comfort criteria for University/Office spaces (CIBSE 2006b).

4.5.2.5 Ground Heat Exchanger

There is an earth tube (ground) heat exchanger present in this building (see section 4.2.2), which was to be modelled using the GroundHeatExchanger object in EnergyPlus. However, in the winter of 2010, during a period of extreme frost, the fluid in the heat exchanger froze, causing the pipes to rupture. The heat exchanger was sealed off and the building continued to operate using mixed-mode ventilation. Therefore, the ground heat exchanger was not included in the final building model, and is not subject to discussion in the research.

4.5.2.6 Internal Loads

Internal loads were approximated using information gathered through site visits, audits and spot checks (Refer Appendix 0).

- Occupancy was defined for summer and academic terms, utilising information gathered from detailed occupancy audits (Appendix C.2) and PC usage data (Refer section 4.4.1.1).

- Infiltration loads were initially added based on standard guidelines for offices and university buildings (CIBSE 2005) as no additional information was available.

Lighting and Equipment loads were applied to the model based on information gathered during the initial site survey and lighting equipment audit (see sections 4.4.1.2 and 4.4.1.3).

4.5.3 Error Check

The next step is to perform an error check to verify the validity of the current version of the BES model. After the initial model was created, at Revision 1, the IDF file was processed using the EnergyPlus EP-Launch program to check for warnings or severe errors. A number of problems were identified and rectified at this stage:

- **Report Variables**: Added correct report variables for comparison with measured energy use data (Heating Load, Electrical Load) as well as Average Zone Temperature Output;
- **Removed Unnecessary Objects**: Unused material and constructions remaining in model idf file were removed;
- **Fixed incorrect AHU scheduling**: An error in idf input caused the simulation to maintain a fixed zone temperature setpoint all day.

4.5.4 Test for Acceptance

In this step, the results of the simulation model are compared with the measured utility data using the pre-defined acceptance criteria. Models are assessed using two main statistical indices; cumulative variation of root mean squared error (CV RMSE) and the normalized mean bias error (NMBE), as well as a dimensionless goodness-of-fit (GOF) index (see section 3.8.5). These indices are computed for hourly values for the following utility level measurements:

- Whole Building Electrical Energy Consumption (kWh);
- Whole Building Heat Energy Consumption (kWh);
- Average Zone Temperature (°C).

It should be noted that this test for acceptance should also account for measurement uncertainty, as discussed in section 6.3.1. This is not accounted for within the scope of this study. However, in order to provide a more accurate picture of the confidence associated with simulation outputs, measurement uncertainty should also be accounted for in these calculations. A possible approach for dealing with this uncertainty is described in section 6.3.1. Further details can be found in the IPMVP guidelines (Efficiency Valuation Organisation (EVO) 2010) for evaluation of energy and water savings.

4.6 Iterative Model Improvement

If the model fails to meet the acceptance criteria, then further iterative improvements are required until the defined acceptance criteria are met. In this case, acceptance criteria were not met at the first iteration, thus validating the requirement for further model improvement. The BES model was iteratively updated to reflect new information collected during continuous data gathering. Additional information was gathered, where possible, for highly uncertain or influential variables. In this case, these variables were identified manually. However, an automated procedure based on the use of sensitivity indices has also been defined in Section 3.6. A sample of these revisions is catalogued in Table 4-10. Each revision was tracked and linked to source information using version control software (TortoiseSVN Core Development Team 2011) as shown in Figure 4-38.

Table 4-10: Examples of revisions during iterative model improvement process

Rev	Description
2	The simulated heat energy consumption was significantly lower than measured heat energy consumption, particularly in the winter period. Initial simulations had been carried out using the default weather file (tmy file) for Shannon Airport, Ireland. However, given the high level of correlation between outside air temperature and heat energy consumption (Figure 4-27), it was deemed a necessity to use actual local **weather** measurements for the calibration period.
4	Initial model used default **constructions** and materials for Irish office buildings. These were later modified using more reliable sources of evidence (e.g. Manufacturer specification sheets, as-built drawings).
11	During initial scoping and simulation, **occupancy** was identified as a highly uncertain and influential variable. This is due to the nature of the building, which serves primarily as a study and office space. Therefore, occupancy patterns can be highly irregular, depending on time of year, day of week and time of day. Initial assumptions for occupancy patterns were updated in Rev. 11 to account for findings gathered during occupancy audits (see section 4.4.1.1).
24	A discrepancy between measured and simulated heating energy consumption profiles was identified during the iterative model development process. This was attributed to a modelling error in specifying the **HVAC system schedule**. This was corrected according to information pertaining to HVAC schedules attained from interviews with the building manager as well as measured BMS data.

Figure 4-38: Sample revisions tracked using Tortoise SVN version control software

4.7 Latin-Hypercube Monte-Carlo (LHMC) Search

Following the iterative update process, the model is subject to a LHMC (latin hypercube monte-carlo) search. The purpose of this process is to further refine the model parameters by simulating the model within a defined search space. This search space is derived from the prescribed mean values and uncertainty ranges for model input parameters.

4.7.1 Define best-guess estimates for input parameters

- The final output of the iterative model development phase is a building model which has assigned input values corroborated with source evidence, where available, and best-guess approximations where no further measurements are available or possible. In this case, an evidence sheet was prepared to track this information (see Table 4-11). A more comprehensive version of this worksheet is presented in Appendix C.3. This sheet provides a complete list of every object in our EnergyPlus model, grouped according to EnergyPlus Input/Out Reference Class. Each field value is recorded in this sheet and linked to source evidence. A drop-down menu allows the user to select the '*class*' of source evidence based on a hierarchy (see

Table 3-3 lists some of the recommended categories for the source hierarchy. Each category also has an associated ranking (class) and range of variation (ROV, %):

$$ROV = \frac{3.\sigma}{\mu} \qquad (7.1)$$

This ROV represents the total heuristically estimated deviation (σ) from the mean value (μ). These are currently preliminary estimations based on prior experience. However, in the absence of any more detailed database of building parameter uncertainties, this provides an adequate means of generating an initial assumption.

Table 3-3: Data Categories

SOURCE	CLASS	ROV (%)
BMS/Sensor Data	1	2
Spot-Measured Data	2	5
Physically Verified Data	2	5
As-Built Drawings	3	10
O&M Manuals	3	10
Commissioning Documents	3	10
Design Documents	4	15
Guides & Standards	5	30
Reference Manual / Default Values	6	40
No Available Information	7	50

).

Table 4-11: Sample Parameters and Information Classification

OBJECT	FIELD	INITIAL VALUE	FURTHER INFORMATION	CLASS	ROV (%)	STD DEV
Schedule:Compact	Field 3	145	EnergyPlus I/O Reference	6	40	19.33
Material 1 - Insulation	Conductivity {W/m-K}	0.04	Approved Document L1 Conservation of fuel and power in dwellings (2002)	5	30	0.004
Material 1 - Insulation	Specific Heat {J/kg-K}	1450	BS EN 12524	5	30	145
Material 2 - Concrete	Conductivity {W/m-K}	2.5	BS EN 12524	5	30	1
Material 2 - Concrete	Specific Heat {J/kg-K}	1000	BS EN 12524	5	30	100
WindowMaterial:Glazing	Thickness (m)	0.003	Drawings / 0517-GEA-001	3	10	0.0001
WindowMaterial:Glazing	Conductivity {W/m-K}	0.9	Default Value	6	40	0.12
Lights	Lighting Level {W}	13100	Lighting Audit 03/03/2011	2	5	3000
ElectricEquipment	Watts per Zone Floor Area	7	Electrical Audit 04/03/2011	2	5	2
Fan:VariableVolume	Fan Efficiency	0.74	Mechanical O&M Manual	3	10	0.05

4.7.2 Assign ranges of variation

Ranges of variation are automatically assigned to each parameter based on the 'class' of source evidence. This range of variation is linked to the prescribed uncertainty associated with each source of evidence.

- For continuous parameters, ranges of variation (ROV) are characterised by probability density functions bounded by upper 95th and lower 5th probability threshold values (see Figure 4-39). The ROV (%) is used to calculate a standard deviation for *continuous parameters* using the following equation:

$$\sigma = \frac{\mu * ROV}{3} \tag{7.2}$$

- Discrete variables are characterized by minimum, maximum and base-case values;
- Multi-dimensional variables, such as occupancy schedules are discretized for specified periods (academic weekday, academic weekend, summer weekday, and summer weekend) and assigned ranges of variation based on available information.

4.7.3 Generate sample matrix

Using the mean values for each parameter and assigned ranges of variation, a sample input matrix is generated using an automated R-script (Refer to Appendix A.2). This matrix is generated by sampling inputs randomly from the defined ranges, assuming parameters are normally distributed (see Figure 4-39). Alternative distributions (e.g. uniform, triangular etc.) may also be specified. However, these are less common, and were not included within the scope of this thesis.

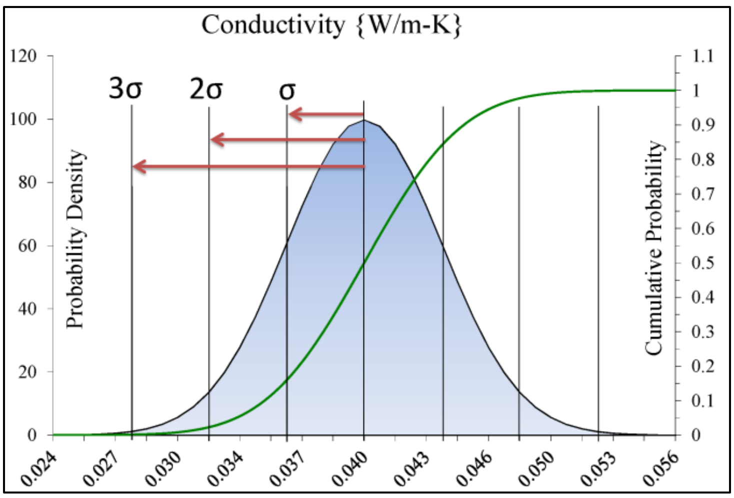


Figure 4-39: Probability density function (PDF) for Insulation conductivity

This step results in the creation of a job matrix in the form:

$$M = \begin{bmatrix} z_1^{(1)} & z_2^{(1)} & \cdots & z_r^{(1)} \\ z_1^{(2)} & z_2^{(2)} & \cdots & z_r^{(2)} \\ \cdots & \cdots & \cdots & \cdots \\ z_1^{(N-1)} & z_2^{(N-1)} & \cdots & z_r^{(N-1)} \\ z_1^{(N)} & z_2^{(N)} & \cdots & z_r^{(N)} \end{bmatrix}$$

Where; z represents the input variables; N the sample size and M the corresponding vector input matrix.

4.7.4 Parametric Simulation

In this step, the sampled job file from 4.7.3 is processed using a batch simulation tool, jEPlus (Yi Zhang et al. 2012). This is used to compute and compile simulation outputs. jEPlus was chosen in this case for the following reasons:

- **Custom job import**: jEPlus allows the user to import custom job files, meaning that sampling can be customized based on user requirements. In this case, it was necessary to randomly sample hundreds of input parameters prior to running each simulation. This strategy differs to typical batch simulations tools which focus on design optimization;
- **Parallel simulations**: jEPlus allows the user to run many simulations in parallel, utilizing all available computing cores on the computers central processing unit (CPU). This was critical as it allowed us to run many more simulations, thus increasing the robustness of the bounded grid search;
- **Flexible support for non-numeric parameters**: Since jEPlus uses string values (in the form @@*parameter*@@) to identify which parameters are variable, it is possible to change any type of parameter within the model input file;
- **File handling**: jEPlus is fully customizable to allow for retention or deletion of files created during the simulation process (e.g. input.idf, outputs.eso etc.). This is useful when file storage limits are an issue, or there is a requirement to keep some element of the simulation process for future reference;
- **Job indexing**: jEPlus automatically indexes all completed jobs, along with a record of warnings, errors and simulation outputs. A problem was encountered when dealing with large batch simulations (>10,000 runs) whereby this index would fail to compile. An indexing script was created in Java to construct this indexing file;

- **Manual post-processing**: It is possible to define customized process functions for handling results of the simulation process. At the time of writing, this function was not implemented, but was due for inclusion in the next release of jEPlus (v.1.3).

Using a template simulation file with strings in place of parameter values, jEPlus enables the automatic processing of large arrays of simulation trials. An initial template file was created by inserting unique strings in place of parameter values in HVACGenerator and setting this as a simulation file within the jEPlus GUI (Zhang and Korolija 2010).

The first step is to create the Parameter Tree. Since optimisation problems often contain multiple dependent parameters, a tree structure is used to represent the hierarchal structure of parameters. However, for the purpose of this case study, we are specifying independent continuous parameters. In order to validate this project in jEPlus, it is still necessary to create a multi-layered parameter tree as shown in Table 4-12:

Table 4-12: Sample parameter tree (jEPlus)

ID	Object	Name	Description	μ	σ	n
P1	Fan:VariableVolume	Fan:VariableVolume	Fan Efficiency	0.74	0.098667	3
P2	Schedule:Compact	Schedule:Activity	Field 3	145	19.3333	3
P3	Material	N_Lib_Batt_Insulation_01	Conductivity {W/m-K}	0.04	0.004	3
P4	Material	N_Lib_Batt_Insulation_01	Specific Heat {J/kg-K}	1450	145	3
P5	Material	N_Lib_Struct_Concrete_01	Conductivity {W/m-K}	2.5	0.25	3
P6	Material	N_Lib_Struct_Concrete_01	Specific Heat {J/kg-K}	1000	100	3
P7	WindowMaterial:Glazing	Clear 3mm	Thickness (m)	0.003	0.0001	3
P8	WindowMaterial:Glazing	Clear 3mm	Conductivity {W/m-K}	0.9	0.12	3
P9	Lights	Lighting	Lighting Level {W}	13100	1746.667	3
P10	ElectricEquipment	Equipment	Watts per Zone Floor Area (W)	7	0.933333	3
				Total Number of Jobs		59049

Where; μ = Mean, σ = Standard Deviation, and, n = Number of Samples

In the above example, it is easy to see how the parameters sample space can quickly become unmanageable when utilising a full factorial (i.e. Design of Experiments) approach. Taking 3 unique samples from 10 continuous parameters will require 59,049 simulations to compute the entire parameter space (Full Factorial Experiment: $3^{10} = 59{,}049$). For this reason, a pseudo-random sampling approach is used to generate a sample matrix of jobs. This sampling process is currently executed externally using R (see section 4.7.3). Once this sample file has been created, jEPlus is used to process the simulations and generate the results (see Figure 4-40).

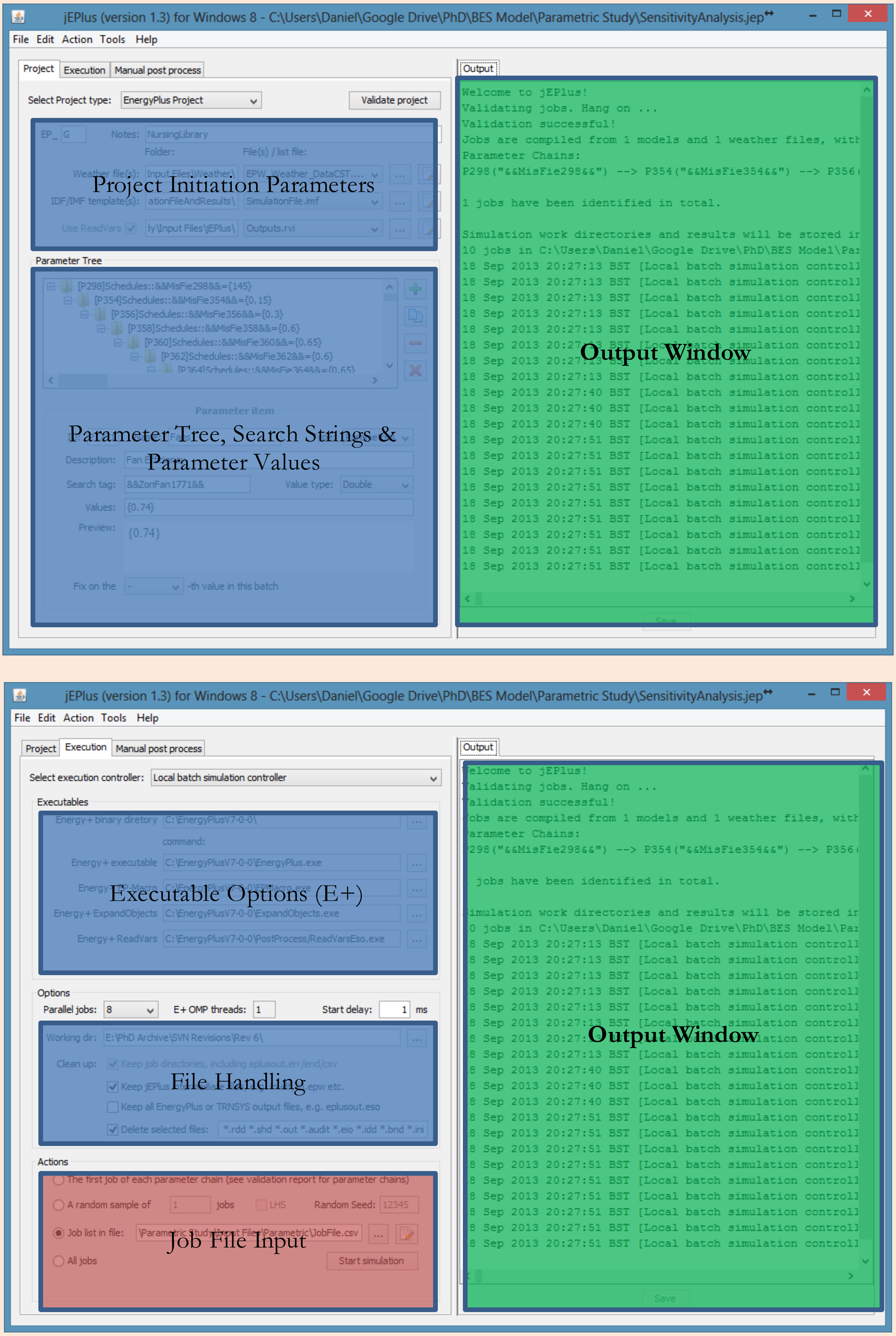


Figure 4-40: jEPlus simulation interface

Once the job file has been processed, this generates the desired output vector, Y, for each row of input matrix:

$$Y = \begin{bmatrix} y^1 \\ y^2 \\ \cdots \\ y^{N-1} \\ y^N \end{bmatrix}$$

These results are collected and transferred to a MySQL database for analysis.

4.8 Post-Processing Data

In order to compare our measured and simulated data directly, it is first necessary to perform some post-processing of the data. This is carried out within MySQL, using automated scripts developed specifically for this case study (see Table 4-13).

Table 4-13: Data post-processing references

Data Type	Measured	Simulated
Zone Temperature (°C)	Refer Appendix A.7	Refer Appendix A.10
Electrical Energy (kWh)	Refer Appendix A.8	Refer Appendix A.10
Heat Energy (kWh)	Refer Appendix A.8	Refer Appendix A.10

4.9 Uncertainty Characterisation of Results

The final step in this process is to recalculate the Goodness-of-Fit (GOF) for each of the simulated models, using the equations outlined in 4.5.4. These outputs are then ranked according to the best-fitting models, thus giving us a set of refined models. Based on these models, it is possible to plot the predicted simulation values (i.e. heat, electrical energy consumption) along their associated prediction uncertainty, based on the uncertainty of the input values which have now been propagated through the model.

4.10 Conclusions

In this section I have discussed how the proposed methodology is applied to a case study building, including the software and analyses used throughout the entire process.

- Section 4.2 give a brief background of the building, its location and the HVAC and plant systems;
- Section 4.3 presents the modelling strategy and source control management (SCM) architecture;

- Section 4.4 discusses the various sources of building information, for both model development and calibration. Data visualisation techniques are used to distil useful information from the vast quantitates of accessible building data. A discussion of the common problems with data collection for detailed model calibration is also presented, with particular emphasis on: data quality, uncertainty and transformation;

- Section 4.5 discusses the implementation of the evidence-based model development process;

- Section 4.6 discusses the iterative model improvement process and the tracking of changes using SCM software. Sensitivity Analysis was not incorporated as part of this demonstration due to time constraints in implementation. However, a combined Sensitivity and Uncertainty analysis will be the subject of further investigation (see Section 6.3.7);

- Section 4.7 illustrates how the parametric study is conducted. Inputs are sampled based on the associated uncertainty of source information, thus allowing for the generation of a LHMC search space, within which all possible (plausible) solutions lie;

- Section 4.8 discusses how this data is post-processed. Given the vast quantities of data generated using this approach, it is necessary to give special attention to how this data is handled and analysed in an effective manner;

- Section 4.9 illustrates how the generated models are assessed for accuracy, using a goodness-of-fit (GOF) approach.

This entire process currently requires a number of software tools and iterative data transformations. To clarify any ambiguous elements of this process, a summary of the tools and techniques used in the model development and parametric study are presented in Figure 4-41 and Figure 4-42 respectively.

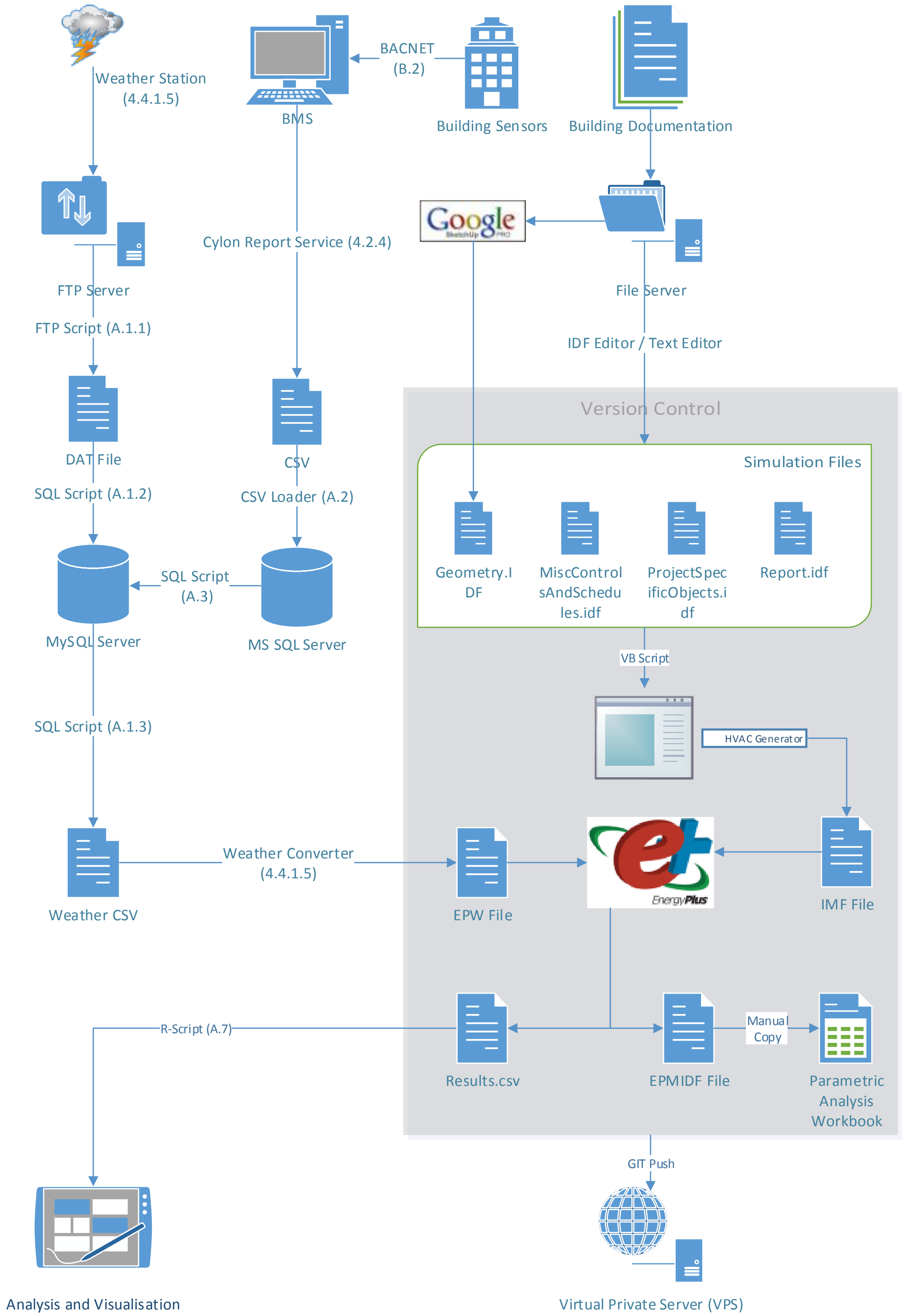


Figure 4-41: Model Development Tool Chain

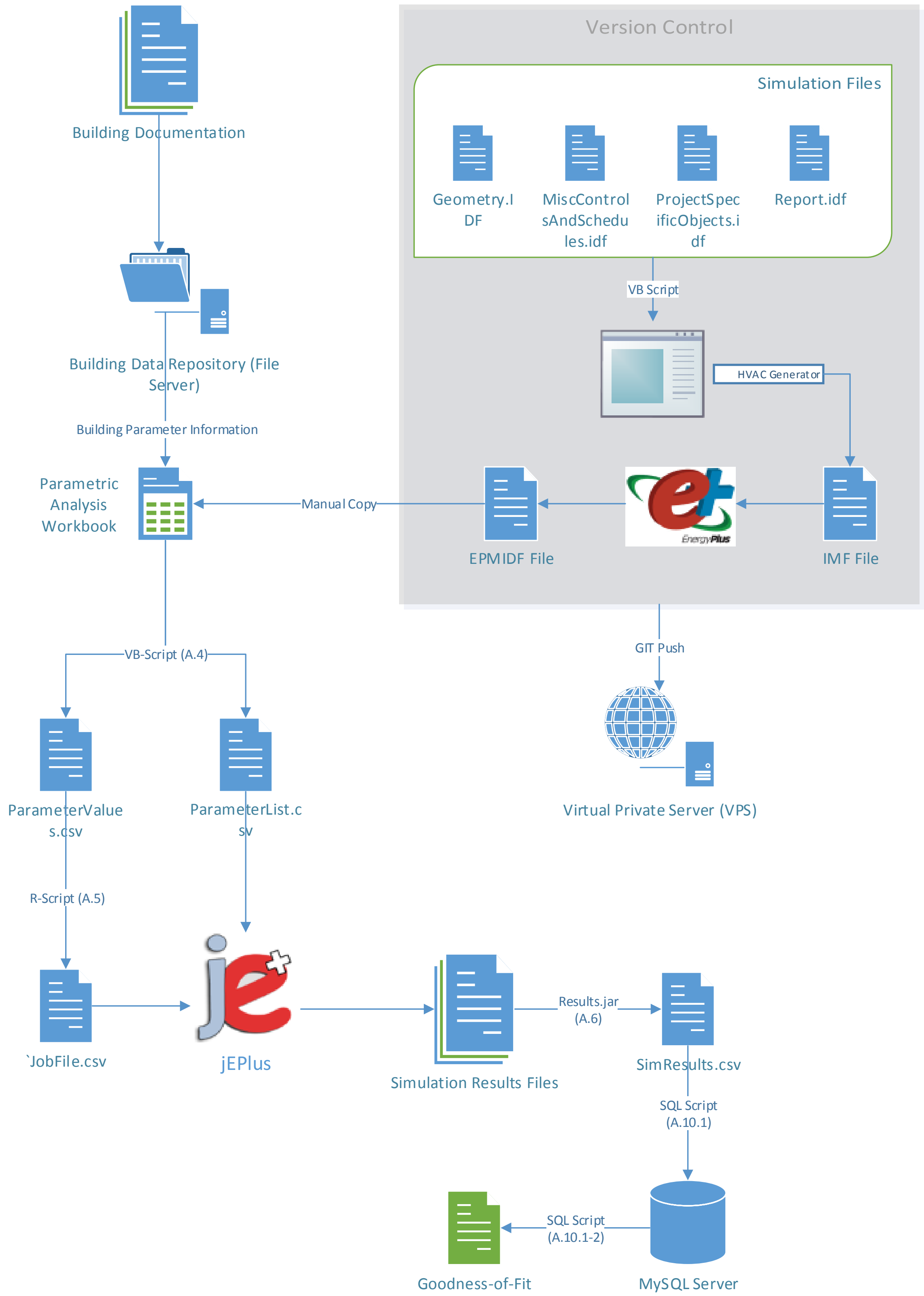


Figure 4-42: Parametric Model Development and Analysis Toolchain

Chapter 5: Results

"Prediction is very difficult, especially about the future."

– Neils Bohr, Physicist

5.1 Chapter introduction

This chapter will focus on the results of the calibration effort, taking into account the modifications at each stage of the calibration process. A summary of the data analysis techniques is presented in section 5.2. This is followed by a discussion of the results of following elements of the proposed methodology, as applied to the demonstrator building presented in Chapter 4:

- Evidence-based model development;
- Iterative model improvement;
- Bounded grid search;

In addition, a comprehensive overview of data visualization techniques is presented in section 5.4, covering intuitive methods for presenting the high volume of information generated during the model simulation and calibration process. The chapter concludes with a discussion of the final results of the process.

5.2 Data analysis techniques

The comparison of simulated data with measured data requires the use of a number of data analysis techniques, which are described in this section. As previously noted in section 4.5.4, the model performance is assessed based on the prediction accuracy for three primary building metrics:

- Whole Building Electrical Energy Consumption (kWh);
- Whole Building Heat Energy Consumption (kWh);
- Average Zone Temperature (°C).

Measured and simulated data is compared on an hourly basis for the entire year in order to compute a single set of statistical indices which describe the model accuracy. These statistical indices include; NMBE (Normalised Mean Bias Error), CVRMSE (Cumulative Variation of Root mean Square Error) and GOF (Goodness-of-Fit), as described in section 3.5.3.

In addition, a number of other data comparison techniques are employed in order to assess the simulation model at each model iteration. Firstly, a Biased Percentage Error (%) is calculated for each hour:

$$\text{Biased Percentage Error } (\%) = \frac{M_i - S_i}{M_i} \tag{6.1}$$

Where M_i and $\mathrm{S_i}$ refer to the measured and simulated value at any instance *i* respectively. This value is used to indicate the relative difference between measured and simulated values, where a negative difference implies over-estimation, and vice-versa. However, due to the presence of negative values, this metric may suffer from compensation effects when average over any time-period (i.e. negative errors cancel out positive errors). Therefore, it is also necessary to calculate an absolute difference in order to give a non-biased estimation of error at each time-step.

$$\text{Absolute Percentage Error} = \sqrt{\left(\frac{\mathrm{M_i} - \mathrm{S_i}}{\mathrm{M_i}}\right)^2} \qquad (6.2)$$

This gives an absolute difference between measured and simulated variables, and thus is a better indicator of actual model performance when averaged over a particular period.

5.3 Evidence-Based Model Development

The evidence-based model development process allows for the systematic improvement of the initial model through a combination of statistical analysis and advanced visualisation techniques. The model results following the 'Model Update' phase of this process are presented in Figure 5-1. The first bar-chart indicates the current performance of the model using the three statistical indices (NMBE, CVRMSE and GOF) calculated for each of the primary building metrics (Heat Energy Consumption, Electrical Energy Consumption, and Average Zone Temperature). The second chart indicates the relative change in % error for each performance index over the previous model revision. This is useful in tracking the relative impact of modifications made during the iterative model development process, providing a clear indicator when a change has had a significant positive/negative impact on the model performance. Finally, the third chart graphs the performance of each model revision. This can be correlated to changes made at each stage in the development process (see Table 5-1).

Table 5-1: Results of Model Update Process

Revision	GOFT	GOFA	GOFB	CVheat	CVelec	CVtemp	MBheat	MBelec	MBtemp	Comment
1	73.9	103.54	55.88	110.36	103.14	23.96	79.56	49.05	8.18	Initial test using Chicago weather data and default constructions. Using district heating plant loop. Default Equipment, Occupancy and Lighting Schedules specified.
2	48.33	70.4	34.23	106.04	59.31	11.36	68.21	18.26	5.54	Changed to local weather file
3	48.25	70.18	34.26	106.22	58.89	11.44	69.04	17.54	5.48	Fixed warnings, Changed site location to Galway, fixed latitude, longitude, Added AHU Schedule Type Limit - Any_Number, Added Ground Temperature Profile,Removed unused constructions
4	64.15	76.22	57.99	109.54	66.47	13.34	72.56	54.58	7.41	Changed materials from default constructions
5	64.33	81.79	54.8	116.09	71.94	11.23	79.99	47.3	4.86	Added Lighting Schedule
6	58.76	78.01	47.77	114.92	66.86	11.69	78.64	37	5.41	Added Electrical Equipment Schedule
7	46.57	66.42	34.23	113.52	49.11	12.51	77.14	2.94	6.28	Revised lighting schedule according to audit data
8	44.28	62.07	33.47	112.29	42.02	12.3	75.65	0.11	6.22	Broke out electrical schedules as per electrical distribution board.
9	43.37	59.75	33.62	112.26	37.51	11.92	75.78	2.83	5.9	Added charts to enable visualisation of Simulation Difference on BarPlots
10	45.36	62.5	35.15	110.65	43.9	11.61	74	14.45	5.53	Revised General Services and Power Schedules based on Available Audit/BMS Data
11	44.03	60.95	33.9	103.15	45.65	11.49	68.5	17.15	5.55	Added Detailed Occupancy Schedule based on Audit Data
12	40.32	56.68	30.34	103.15	37.97	11.49	68.5	1.46	5.55	Changed to Total (Building) Electric Demand - No HVAC Electric Demand
13	39.98	55.75	30.45	103.93	35.6	11.64	68.66	2.31	6.22	Changed Surface Convection Algorithm
14	39.98	55.76	30.44	103.95	35.6	11.85	68.62	2.31	6.35	Set convergence limits and SIte location
15	38.27	54.37	28.32	100.1	35.6	12.93	63.81	2.31	7.5	Switched to Watts per Zone Floor Area for Lighting and Equipment Gains.
16	48.53	63.79	39.93	125.48	35.43	16.06	90.14	1.85	8.54	Moved to HVACgenerator. Adjusted Radiant Fraction from People and Equipment
17	33.59	52.52	20.09	94.05	36.19	10.98	43.39	6.62	7.16	Revised ACH from 0.5 to 2.5 and changed infiltration schedule
18	31.83	49.5	19.36	92.91	31.14	11.04	43.68	0.37	7.31	Corrected Academic Weekend Lighting and Equipment Schedules
19	31.07	47.49	19.8	92.97	26.8	11.1	44.18	3.37	7.47	Corrected SUmmer Weekends to account for Sunday closures
20	30.71	46.41	20.13	92.56	24.62	11.09	44.81	3.82	7.47	Corrected Bank Holidays and Special Day Schedule
21	30.51	45.83	20.28	93.13	22.64	11.1	45.34	3.15	7.48	Corrected Christmas Holidays and September Bank Holidays
22	30.56	45.87	20.33	93.3	22.56	11.05	45.43	3.31	7.43	Corrected Electrical Schedule at Dec 16
23	29.34	44.19	19.38	89.07	22.56	10.08	43.73	0.79	6.5	Added Baseboard Heating Object
24	32.31	43.66	25.71	85.84	24.16	18.8	55.13	8.76	16.88	Corrected baseboard heater node connection. Re-scheduled infiltration. Changed heating and cooling setpoints for baseboard and ADU system
25	29.51	40.38	23.1	77.29	24.1	17.49	49.06	8.57	15.65	Increased max infiltration value to compensate for winter design day.

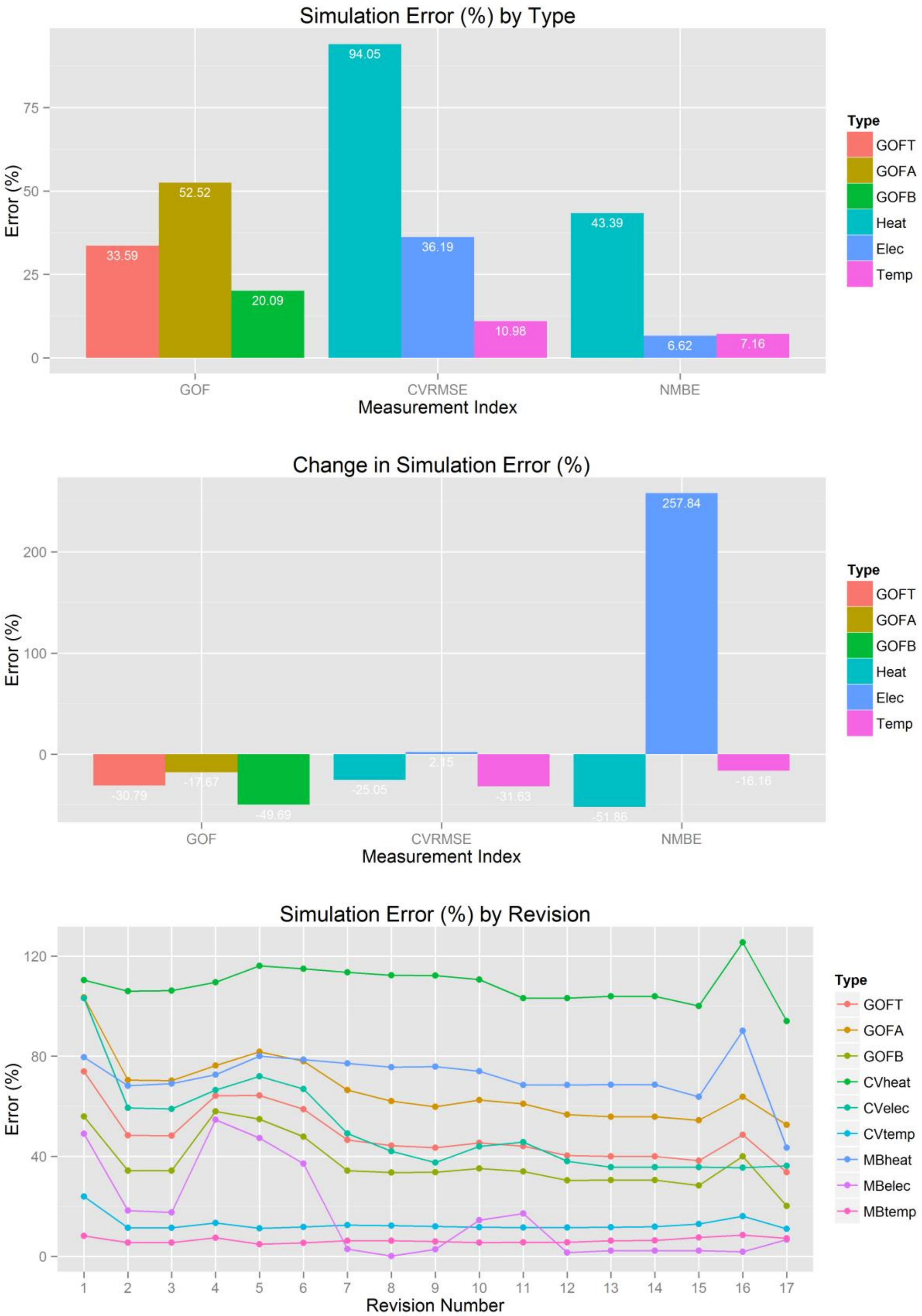


Figure 5-1: Model Results for Revision 17 following completion of Model Update Process

At the end of the model update process, a number of discrepancies were still evident, particularly in the building heat energy consumption (CVRMSE > 90%) and electrical energy consumption (CVRMSE >30%). These issues are addressed in the 'iterative model improvement' phase.

5.4 Iterative Model Improvement

For the purpose of succinctness, this section will focus on the process of iterative improvement process for one model parameter, as this sufficiently demonstrates the step-by-step process involved. As above, a perceivable discrepancy between measured and simulated electrical energy consumption has been identified following the model update process. Having exhausted all currently available measured data and documentation, it is now necessary to perform a more in-depth model analysis. This is achieved by using statistical comparison techniques in combination with automated data visualisation.

5.4.1 Data Visualisation

When examining raw simulation data, and comparing it to measured building data, it is often difficult to understand the reason behind discrepancies between these values. For example, in the case of comparing measured and predicted electrical energy consumption data, it becomes impossible to derive any conclusions as to the root-cause of issues when comparing the entire un-processed data set (see Figure 5-2).

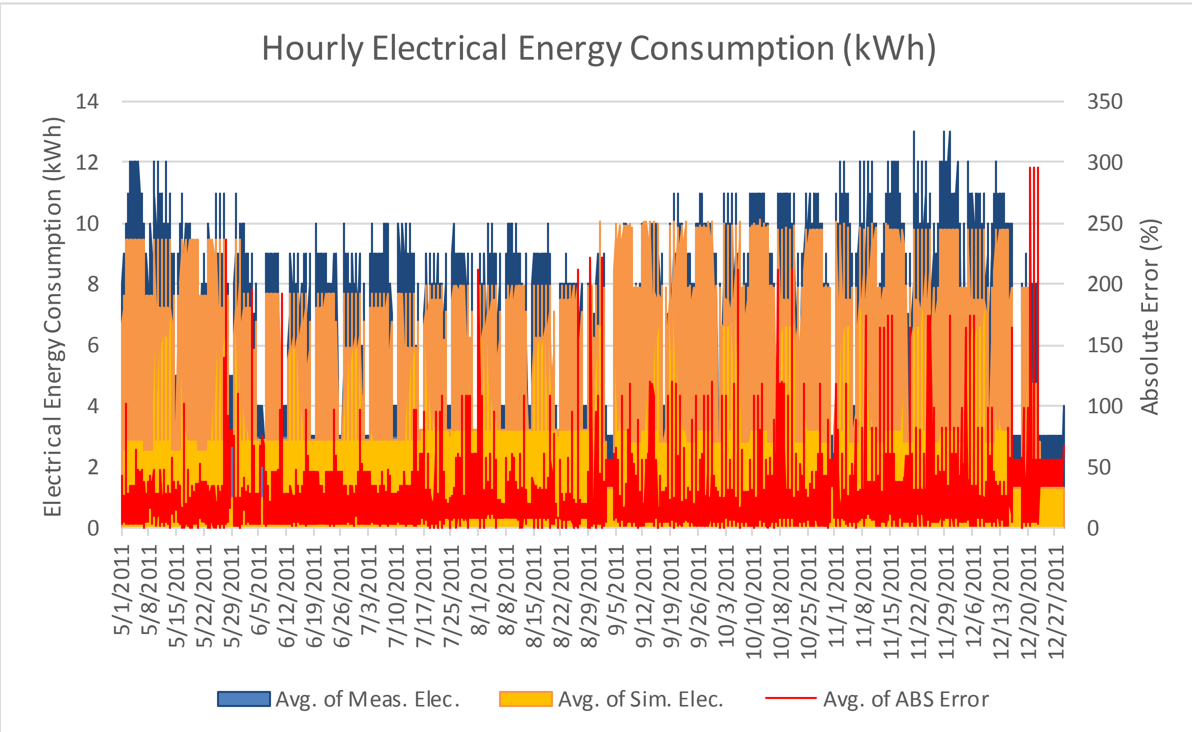


Figure 5-2: Hourly Electrical Energy Consumption (Raw Data)

Conversely, it is also difficult to root-cause issues when data is over-summarised, or presented in a manner which may be confusing or lead to incorrect conclusions.

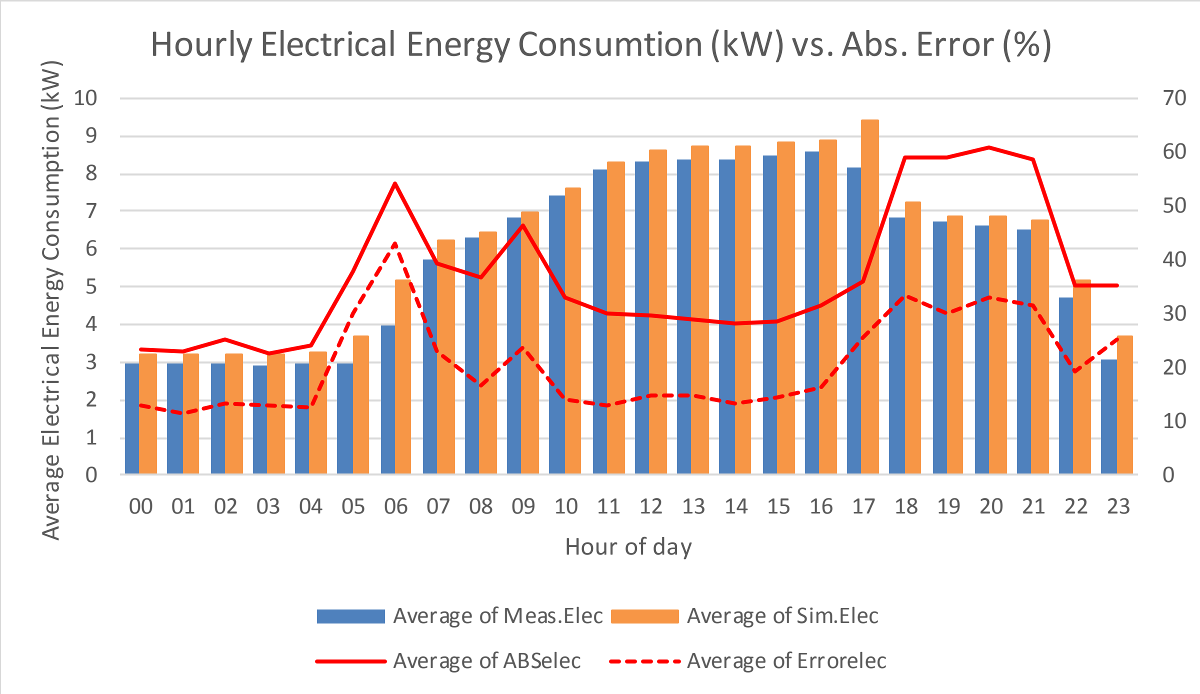


Figure 5-3: Hourly Electrical Energy Consumption, Binned Analysis

In Figure 5-3, the same data is presented using an hourly summary. At first glance, the simulated data appears to correlate very well with the building measurements. However, when the bias error (Errorelec) and average absolute error (ABSelec) are overlaid for each hour, a different picture begins to emerge. Average Absolute error for each hour is in the range of 20-60%, while bias percentage error is in the 15-40% range. This indicates there is a problem with the model, but does not give any clear indication as to the possible root-cause. This example also highlights the importance of using absolute error values to avoid the type of error compensation effects evident in the lower values for mean bias error here.

In order to get a better representation of the root-cause of the problems with the current model, it is necessary to break down the data further into relevant 'bins' or groups of similar data. This can be done by attaching qualitative metadata to the model, for example:

- Time of Day (Morning, Evening, Night-time);
- Day of Week – Weekday vs. Weekend;
- Month of Year – Academic Term (Sep-Dec, Jan-May) vs. summer (Jun-Aug).

Other useful considerations may include holidays, or class timetables to indicate whether the building is highly occupied. When this information is used to break down the model further, a definite pattern begins to form:

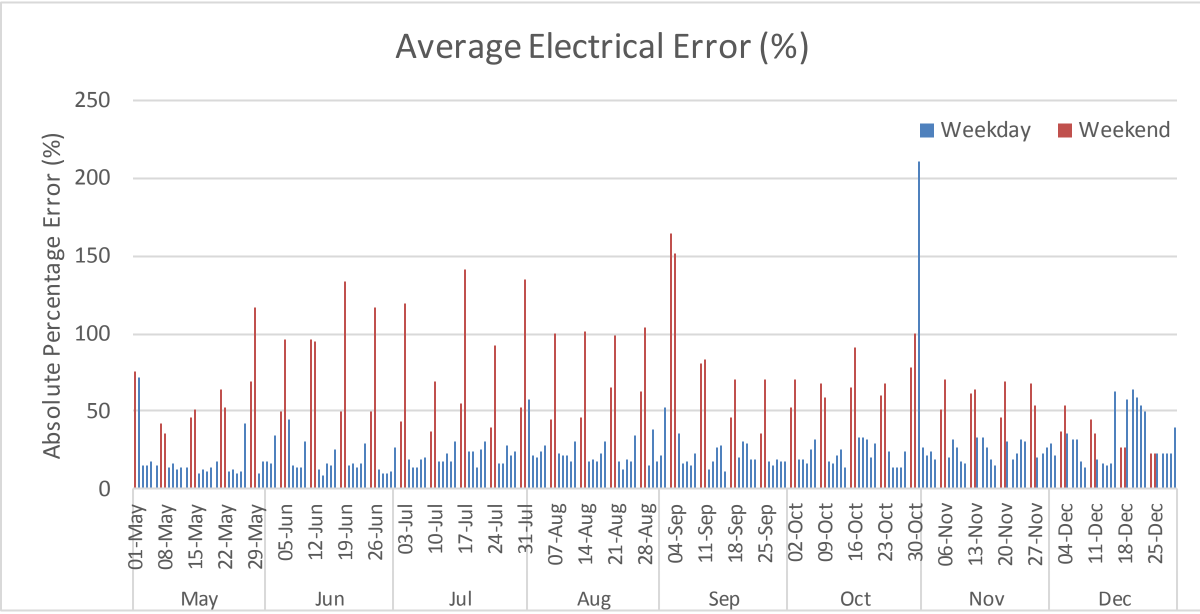


Figure 5-4: Electrical Energy Consumption - Absolute % Error (Weekday vs Weekend)

Looking at Figure 5-4, it is clear that weekends account for a significant proportion of the errors >90%. Furthermore, by attaching bins/groups to this data set, it is possible to begin to filter through the data set in order to gather more insight into the issue.

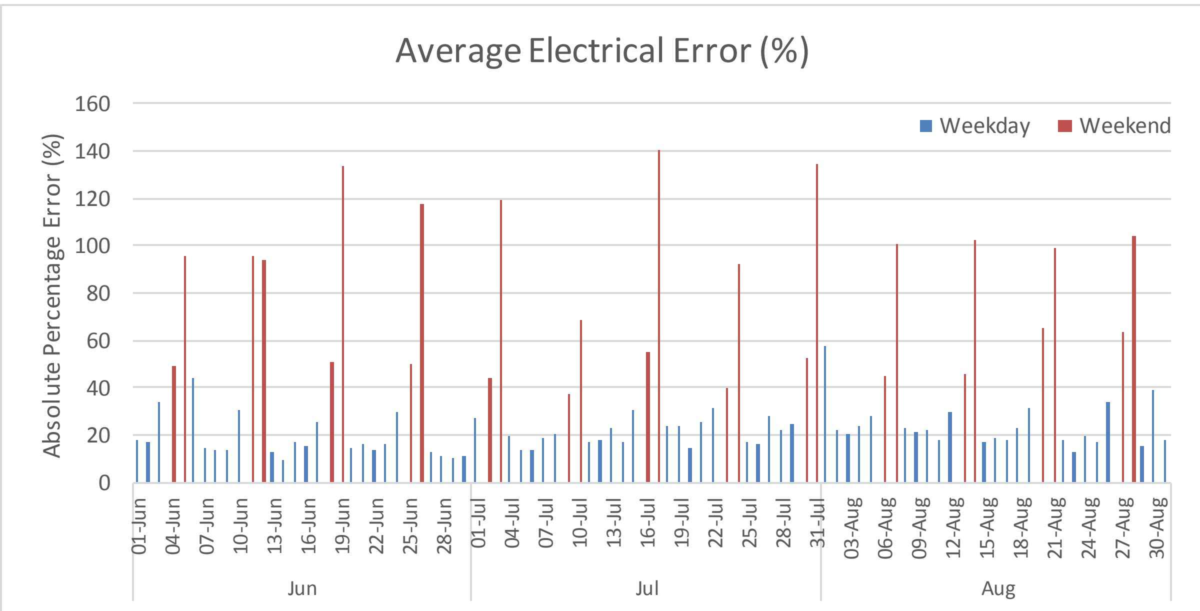


Figure 5-5: Electrical Energy Consumption - Absolute % Error (Summer Weekday vs. Weekend)

It is clear that it is necessary to perform strategic interrogation of model data in order to root-cause any potential issues. Statistical analysis and data visualisation techniques (e.g. MS Excel

Powerview, Pivot charts and Pivot tables) are a useful means of performing this analysis by drilling down into large data sets to infer meaning. However, as has been demonstrated, it is important to remain cognisant of the potential pitfalls of inferring the incorrect conclusions through the use of statistical indices (mean bias error vs. absolute error) or over-summarisation of results.

This still leaves the issue of analysing data in an efficient structured manner. While Pivot Charts and Tables provide an efficient means of pre-assessing large data sets and inferring knowledge from this data, it is also a time-consuming process to filter through numerous graphs and models for each simulation assessment. Fortunately, there exist more advanced data analysis techniques, such as surface plots, which allow the 3-D visualisation of data, thus increasing the amount of information which can be interpreted from a single graphic.

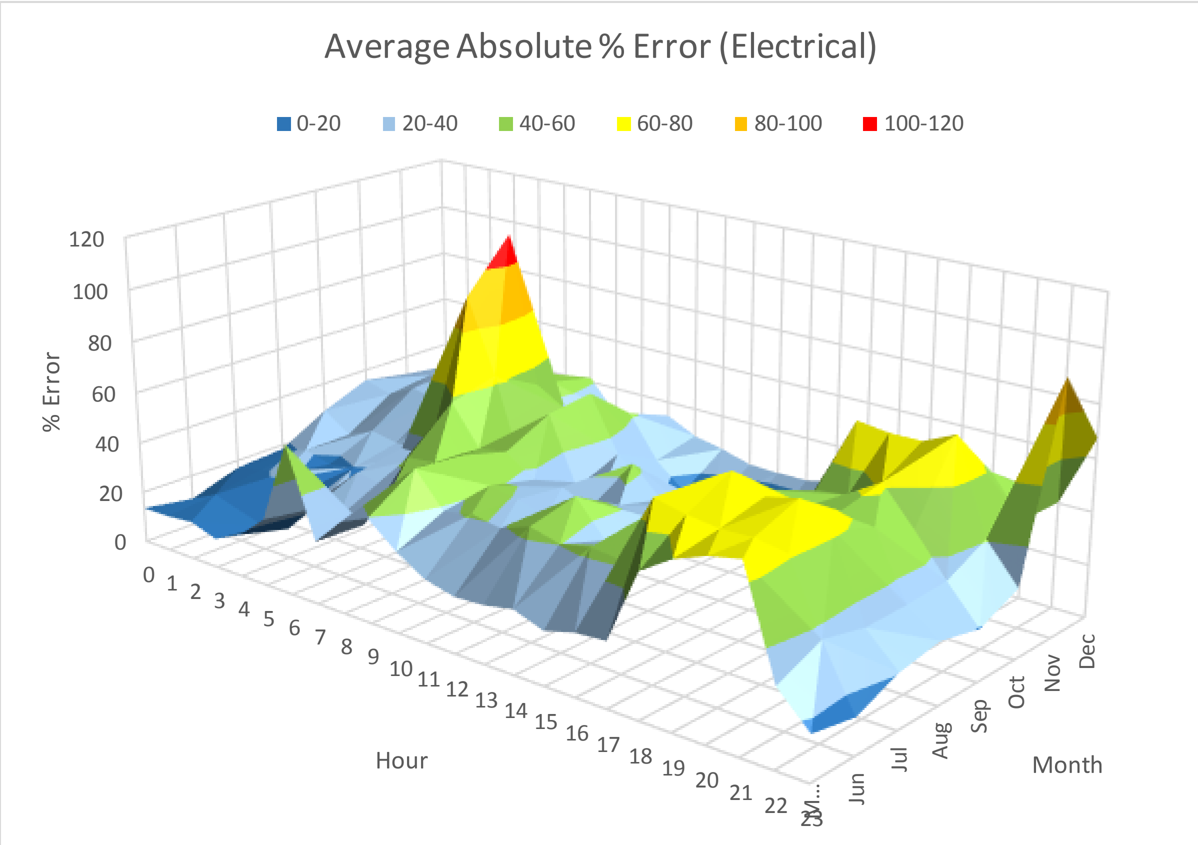


Figure 5-6: Electrical Energy Consumption - Absolute % Error (Surface Plot) – Rev. 17

Figure 5-6 is an example of how the information can be visualised more effectively in this manner. By utilising the z-axis (depth) to display the response variable (i.e. absolute % error), it is possible to further contextualise the data set by adding another variable to the graph. This allows absolute error to be displayed against hour of day and month of year in an easily understandable manner. It is immediately apparent where major discrepancies lie within the model (e.g. Morning, evening, November, December). While visually appealing and easier to interpret, one downside of this form of visualisation is that is can sometimes be difficult to read effectively given the limits of display perspective. It is often necessary to rotate the display in order to examine areas hidden behind peaks in the chart. An alternative is the use of carpet contour plots, which display the same information on a 2-D surface using colours to differentiate the levels in the response variable (z-axis).

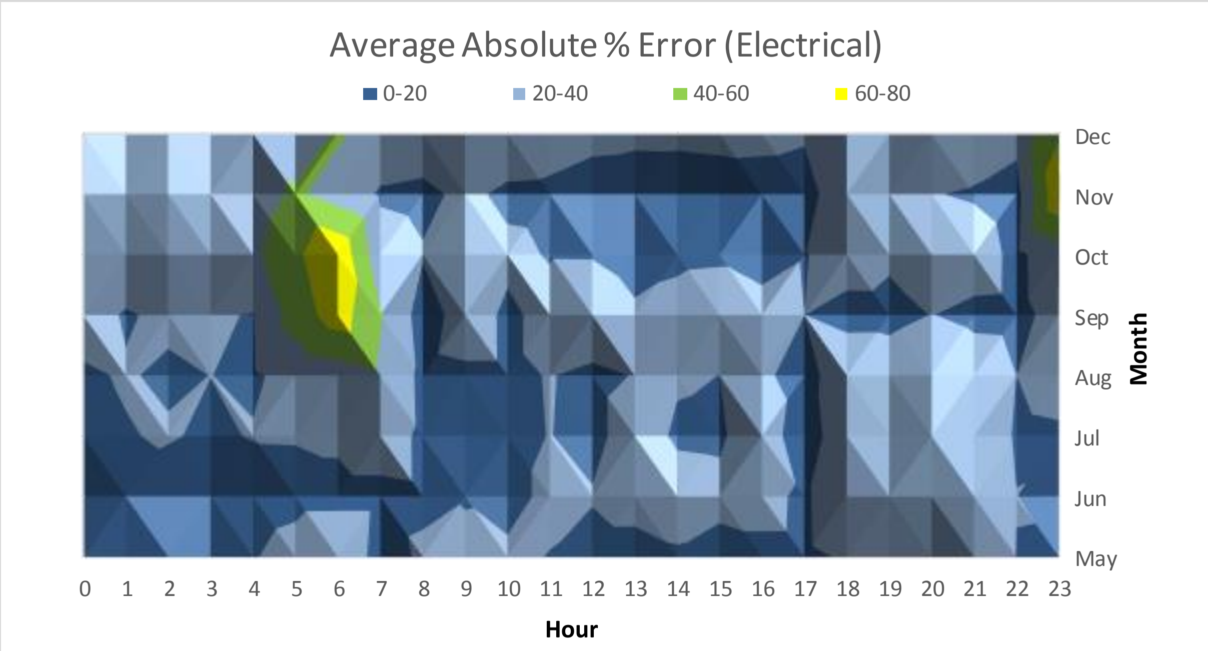


Figure 5-7: Absolute % Error (Electrical Energy Consumption) – Rev. 17

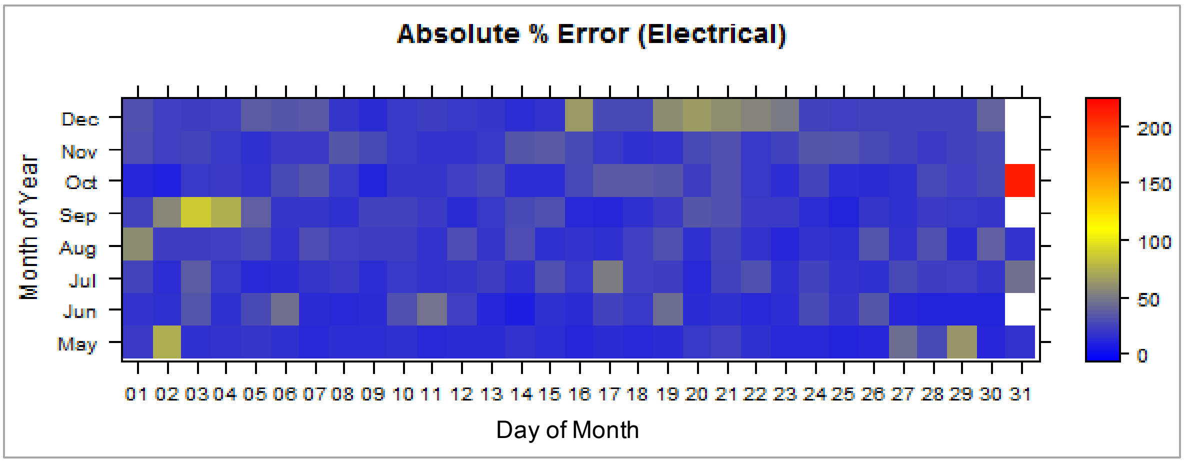


Figure 5-8: Absolute % Error (Electrical Energy Consumption) carpet plot – Rev. 17

The above contour plots (Figure 5-7 and Figure 5-8) illustrate the same information as the 3-D surface plot in Figure 5-6, using colour scales to represent the mean absolute error for different hours of the day and months of the year. It is immediately apparent how this graph may be used as a more effective tool for identifying model discrepancies. From a cursory examination of the plot, the following model characteristics are identifiable:

- Major discrepancies (~100% abs. error) exist in September and October at 06:00;
- Major discrepancies (70-90% abs. error) exist in November and December at 23:00;
- The model performs better during operational hours (10:00-17:00);
- There is a higher percentage error (~40%) for operational hours primarily during the summer months (Jun-Aug/Sep);
- There is a higher model discrepancy in the evenings (18:00-21:00);
- The fit overnight is reasonably good (0-20% abs. error) apart from the period Oct-Dec when errors increase to the 20-40% range.

A final plot which can prove extremely useful is Absolute Error against month and date (See Figure 5-9). This is useful in identifying particular dates which display high levels of error, for example in the case of holidays or scheduled closures. In Figure 5-9, it is possible to clearly identify a number of dates which merit further investigation (May 2, May 29, Aug. 1, Sep. 2-4, Oct. 31, Dec 16, and Dec. 19-23).

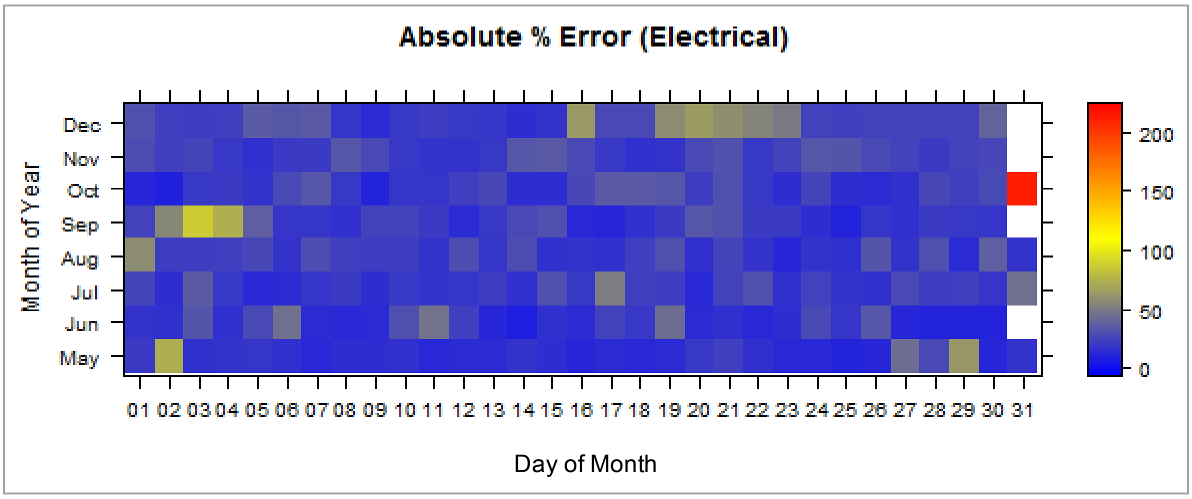


Figure 5-9: Electrical Energy Consumption - Absolute Error (%) by Month and Date – Rev. 17

5.4.2 Visual Summary

In order to allow convenient assessment of model iterations in a uniform manner, statistical summary data is calculated automatically and presented visually, in accordance with the generic post-processing script which has been developed (see Appendix A.9). An example of this visualisation is presented in Figure 5-10 to Figure 5-12. This one-page summary presents a significant amount of data relating to model performance in the context of electrical energy consumption. From the graphs, it is possible to identify:

i. Monthly summary of absolute error (%) by day of week and hour of day;

ii. Overall summary of absolute error (%) by day of week and hour of day;

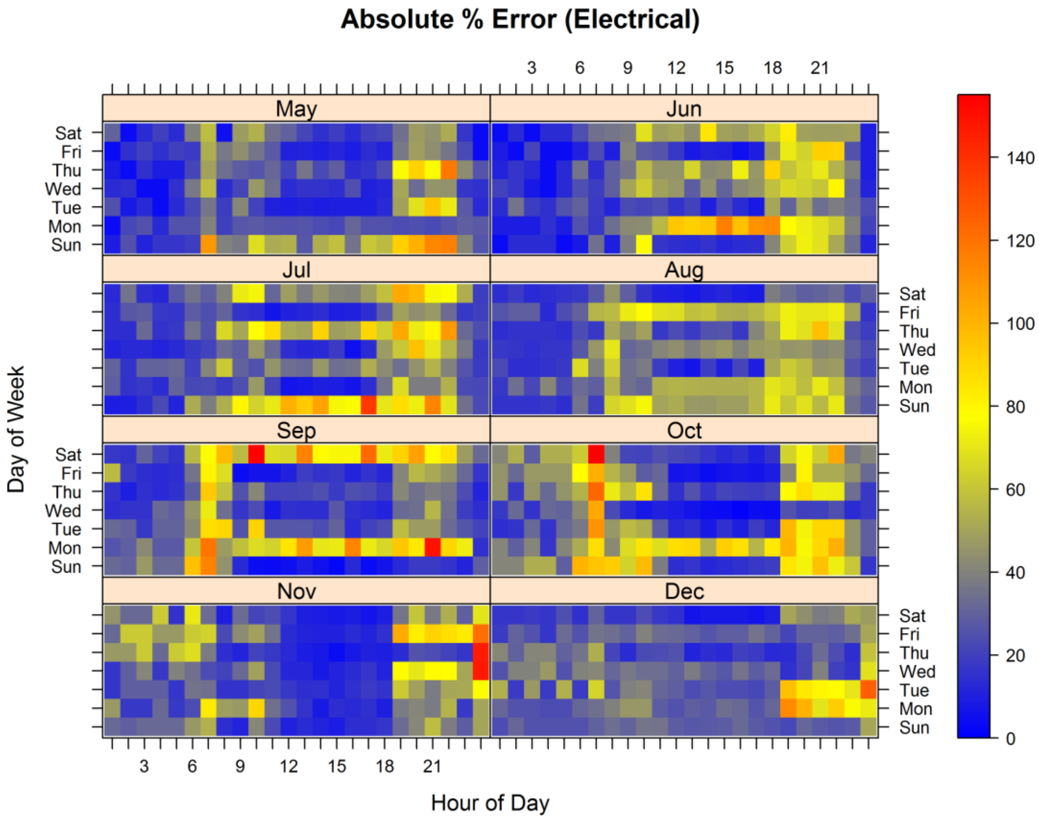
iii. Overall summary of absolute error (%) by month and hour of day.

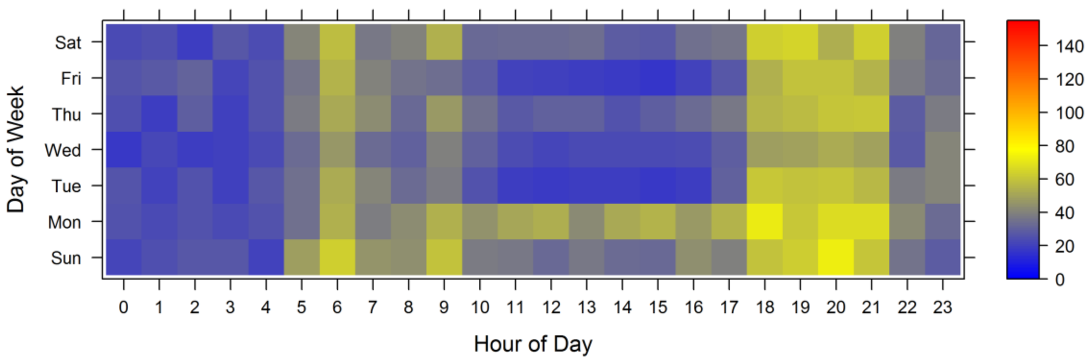
This makes it possible to quickly identify problem areas which need to be addressed.

Table 5-2: Model issues identified using carpet plot data visualisation

Graph	Description
1	• Weekend operational profile (Jul, Aug, Sep) • Evening consumption (18:00-22:00) for all months. • Morning consumption (7:00) in August, September and October • Night-time consumption (23:00) in November • Evening consumption for Monday, Tuesday in December
2	• Evening consumption (18:00-22:00) • Morning consumption (06:00-09:00) • Mondays – check for incorrect holiday, special day scheduling
3	• Evening consumption (18:00-22:00) • Morning consumption (06:00-09:00) – particularly Sept – Oct. • Night-time consumption (23:00) in Nov – Dec.

Therefore, it is evident that the use of carpet plots, as opposed to standard line and bar charts, makes it possible to convey more information in a single graphic, and thus diagnose issues more effectively.





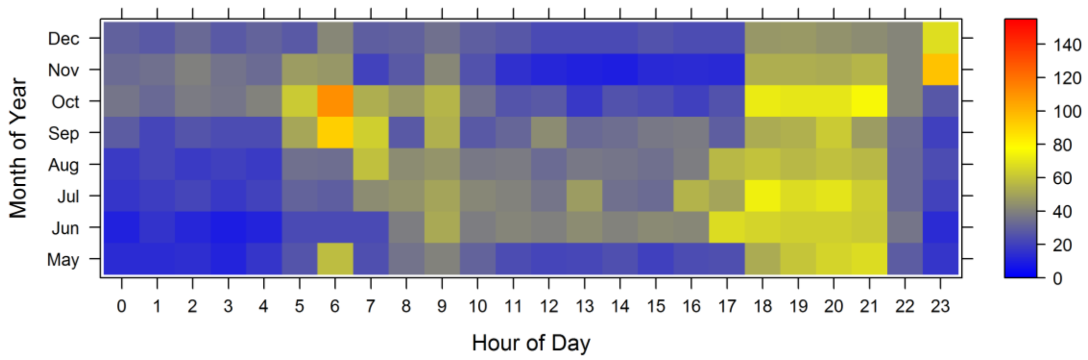
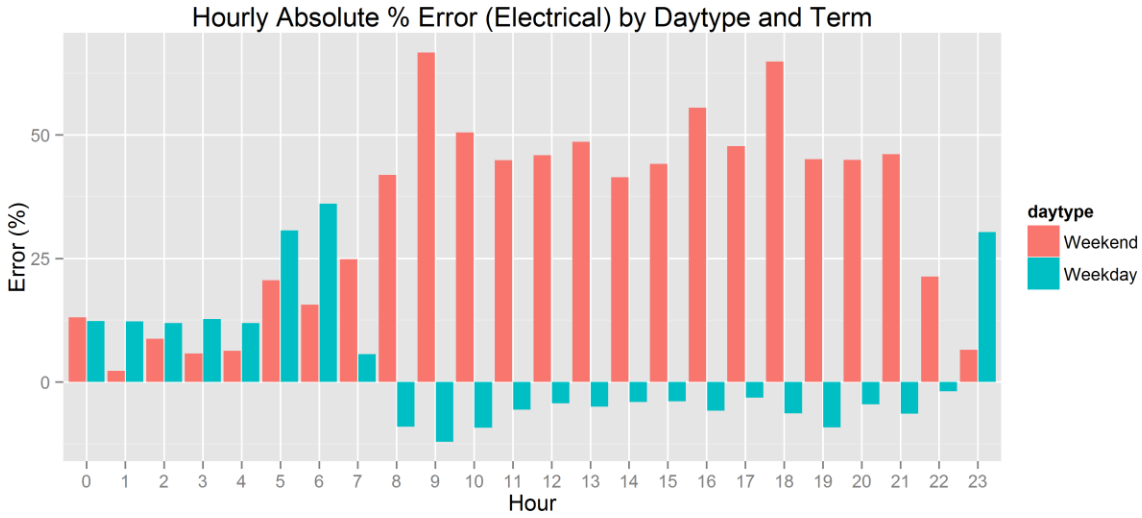


Figure 5-10: Absolute Electrical Error (%) – Rev. 17



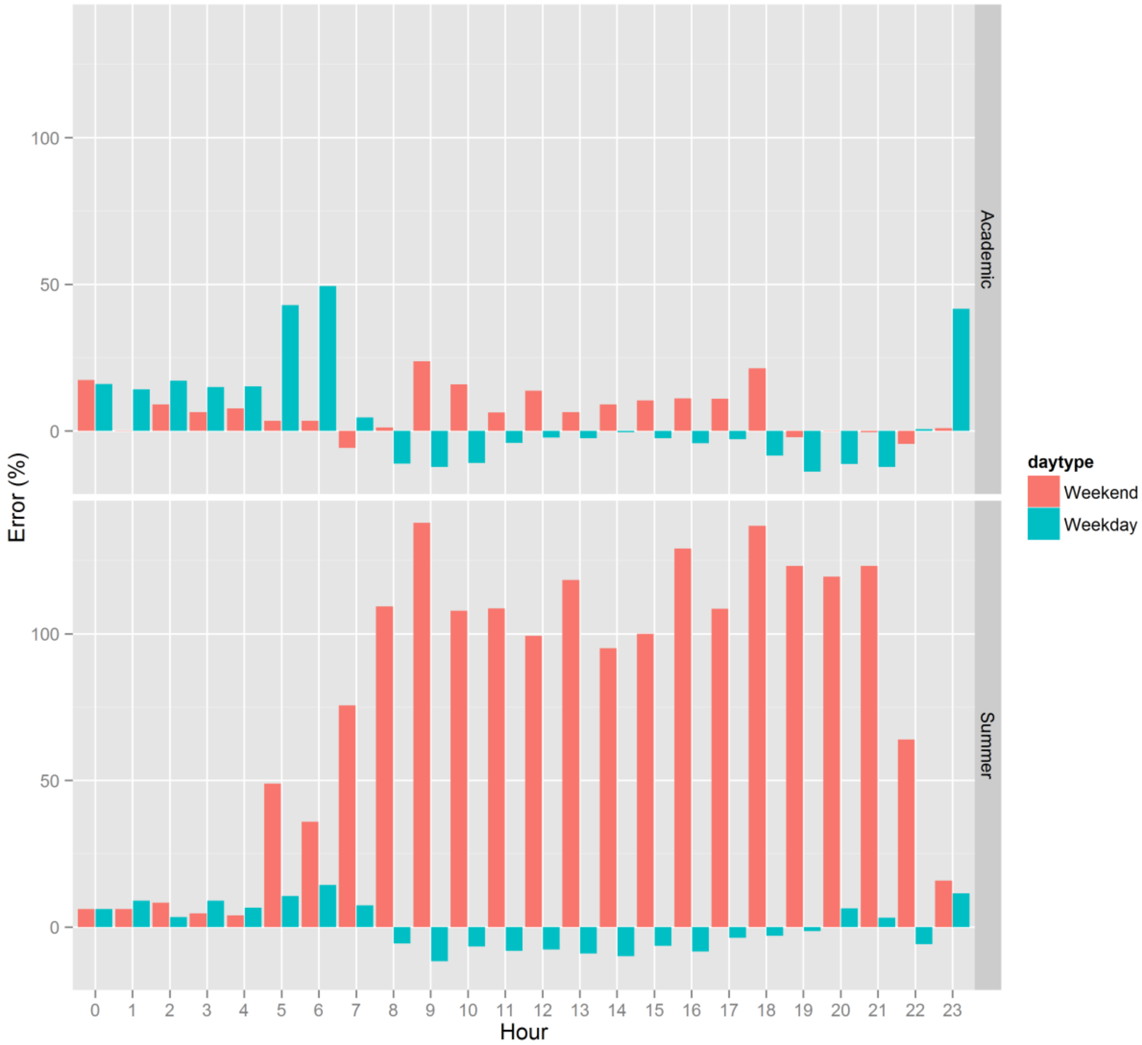


Figure 5-11: Summary of Bias % Error by Day type and Term – Rev. 17

Figure 5-12: Summary of Bias % Error by Term and Month – Rev. 17

By combining the contour plots from Figure 5-10 with the statistical summary charts in Figure 5-11 and Figure 5-12, it is possible to effectively diagnose the main issues affecting model performance. This technique of visual interpretation is extremely efficient is diagnosing issues, in addition to highlighting other potential problems such as poor correlations between measured and simulated trends during particular periods. Such problems may be overlooked if utilising statistical methods as the sole means of diagnosing model performance.

5.4.3 Sensitivity Analysis

While the visual summary approach provides an excellent means of diagnosing issues with the model throughout the iterative improvement process, it also suffers from a heavy reliance on analyst knowledge and an ability to read and interpret such visual summaries. Therefore, mathematical and statistical indicators can provide a useful counter-balance, and add further to the information available to the analyst with which to make a judgement. An (optional) sensitivity analysis procedure has been proposed, but not implemented, in the prescribed methodology. Unfortunately, due to time constraints, it was not possible to include a comprehensive SA within the scope of the case study. However, a framework for its inclusion is outlined in 3.6 and a discussion on how this may be improved in future work is described in section 6.3.7.

5.4.4 Final Results of Model Update Process

Despite an intensive building audit, high levels of measured data and a detailed systematic model improvement process, there are still discrepancies evident in the simulated model (see Figure 5-13 to Figure 5-19). This serves to highlight the level of detail required to model a building to a sufficient degree of accuracy as well as the need for an approach to addressing the uncertainty of individual input factors.

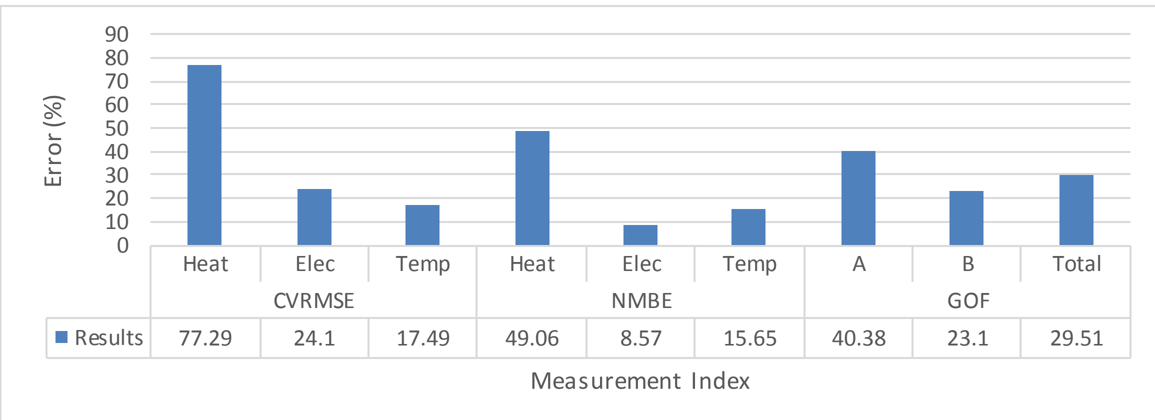


Figure 5-13: Residual errors (%) after model update process

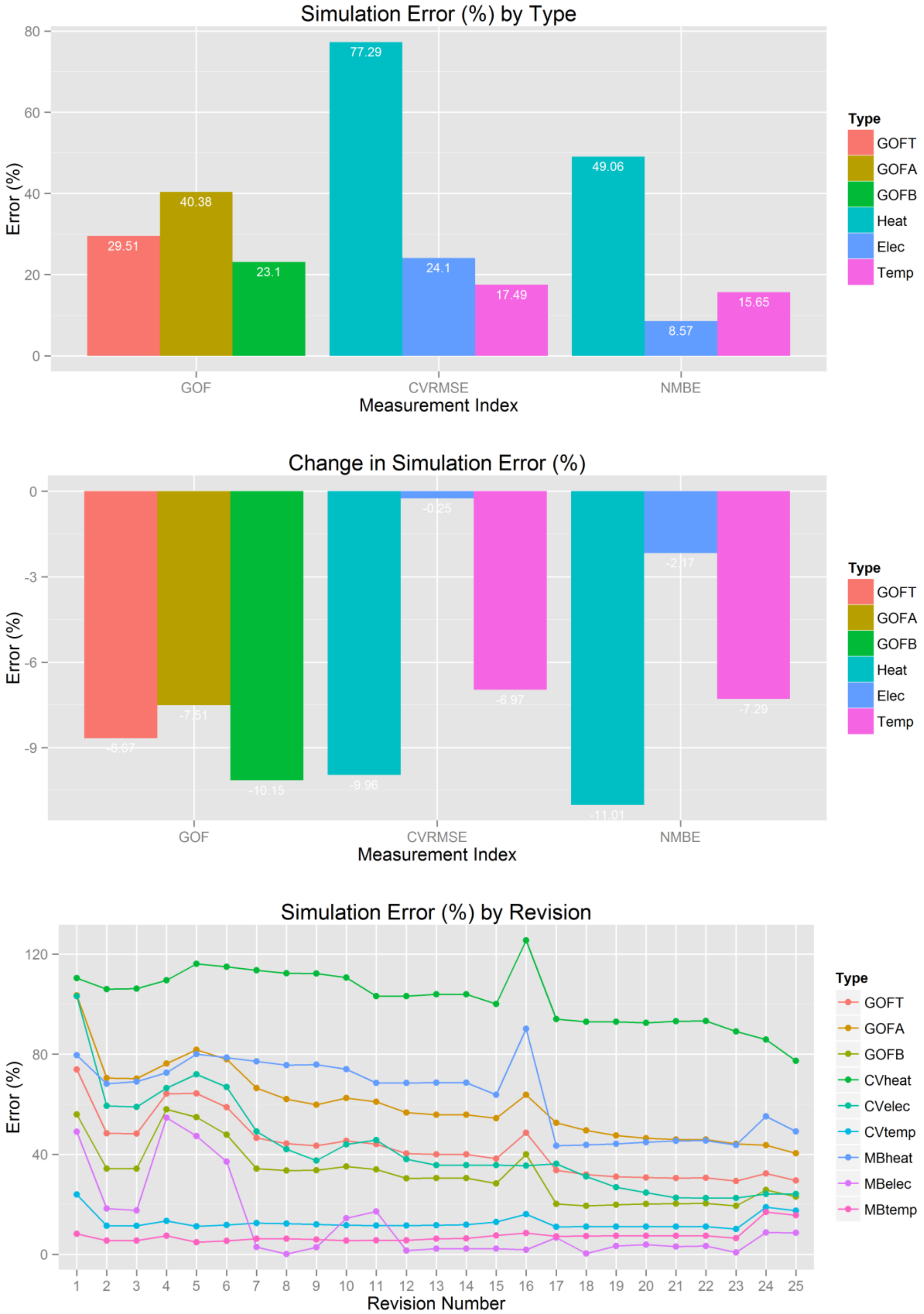
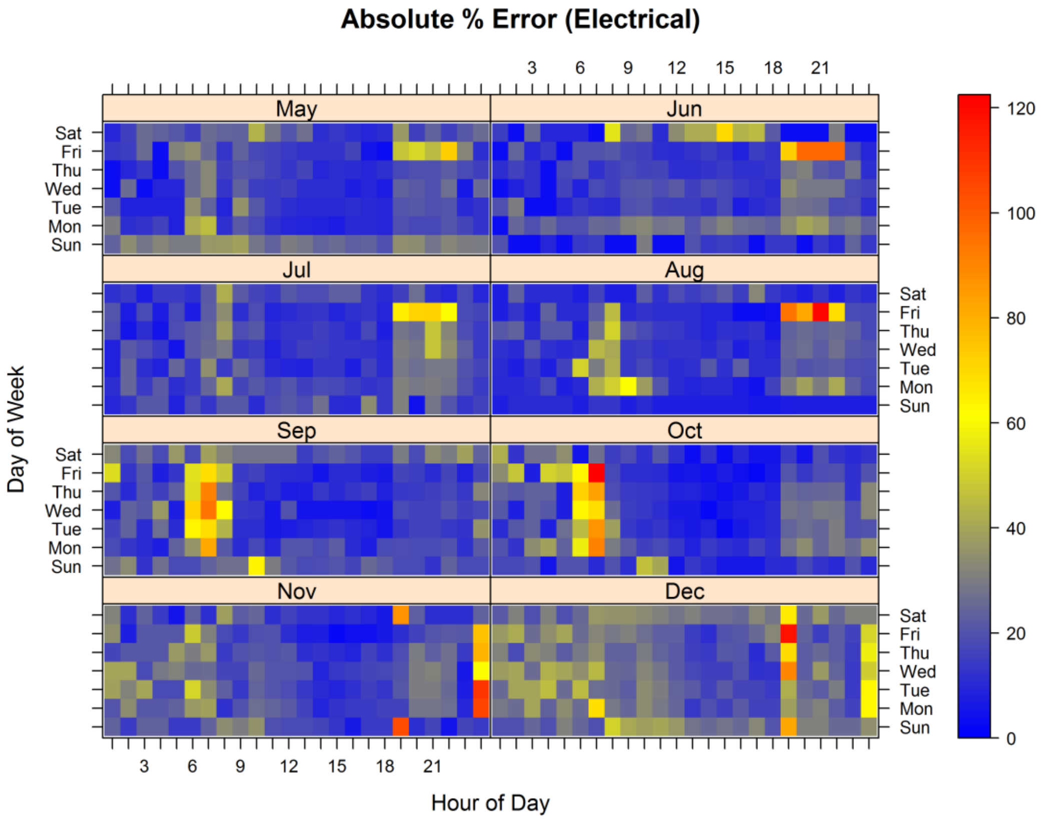
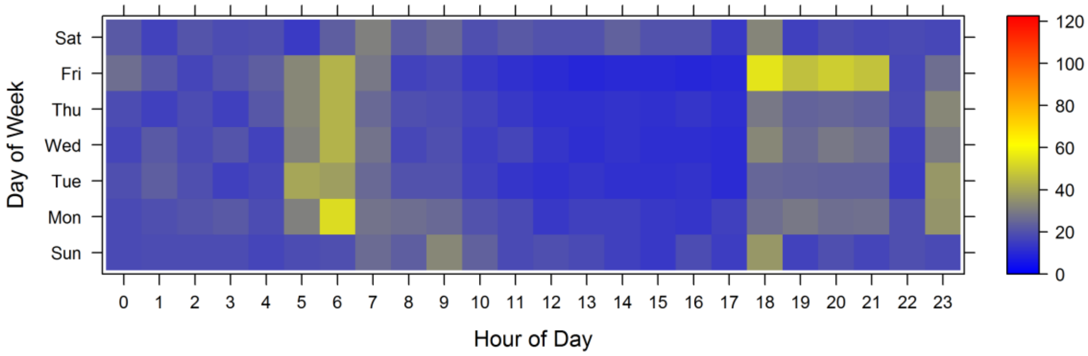


Figure 5-14: Summary of model errors (%) - Rev 25





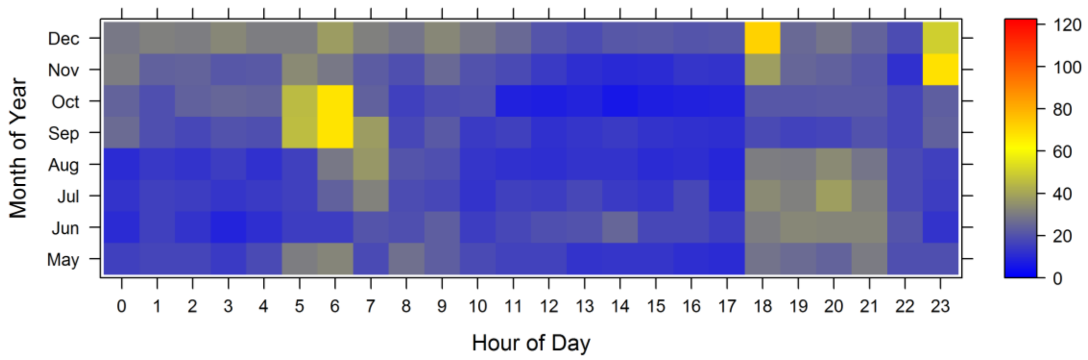


Figure 5-15: Absolute % Error (Electrical) - Rev 25

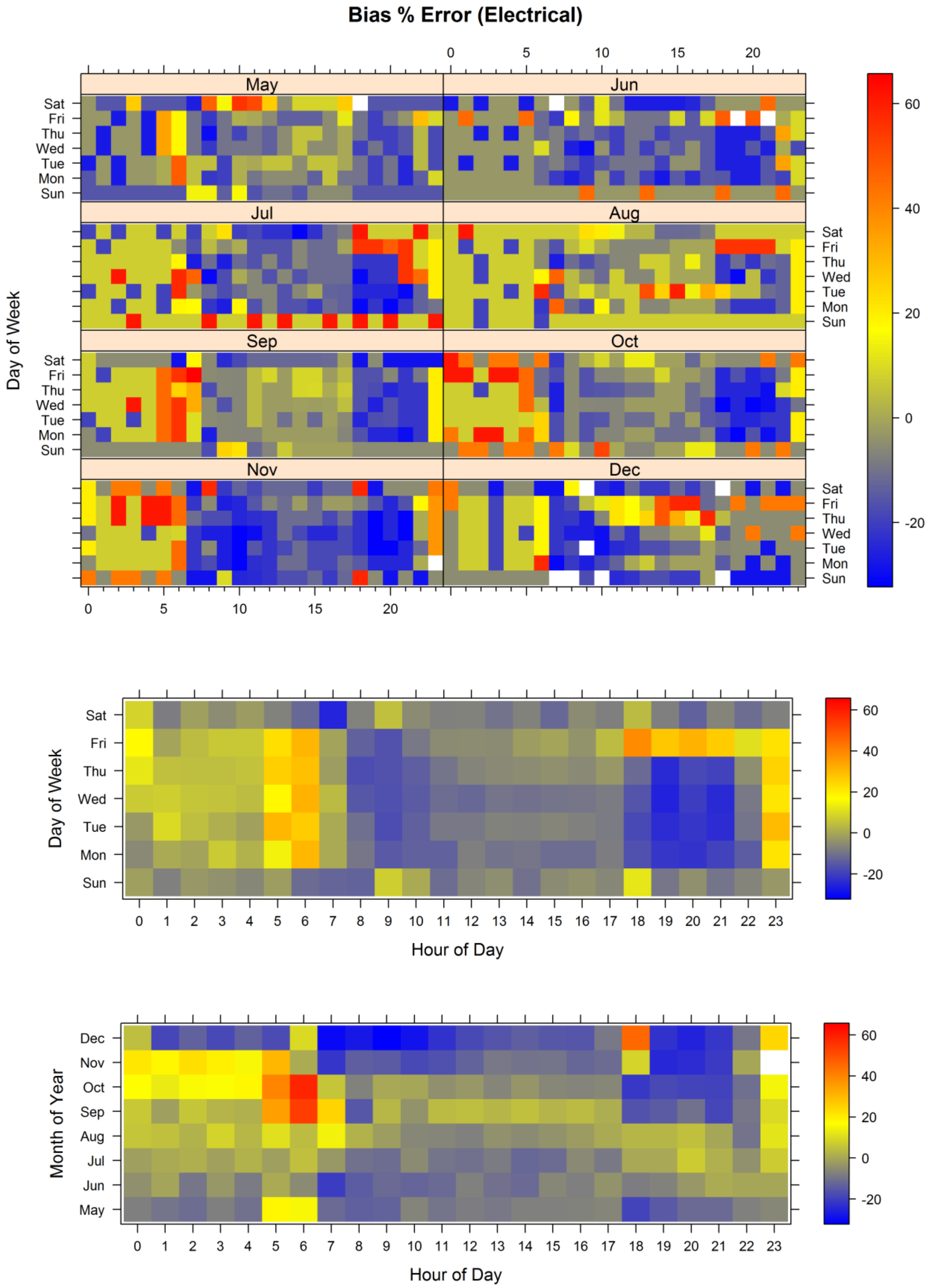


Figure 5-16: Bias % Error (Electrical) - Rev 25

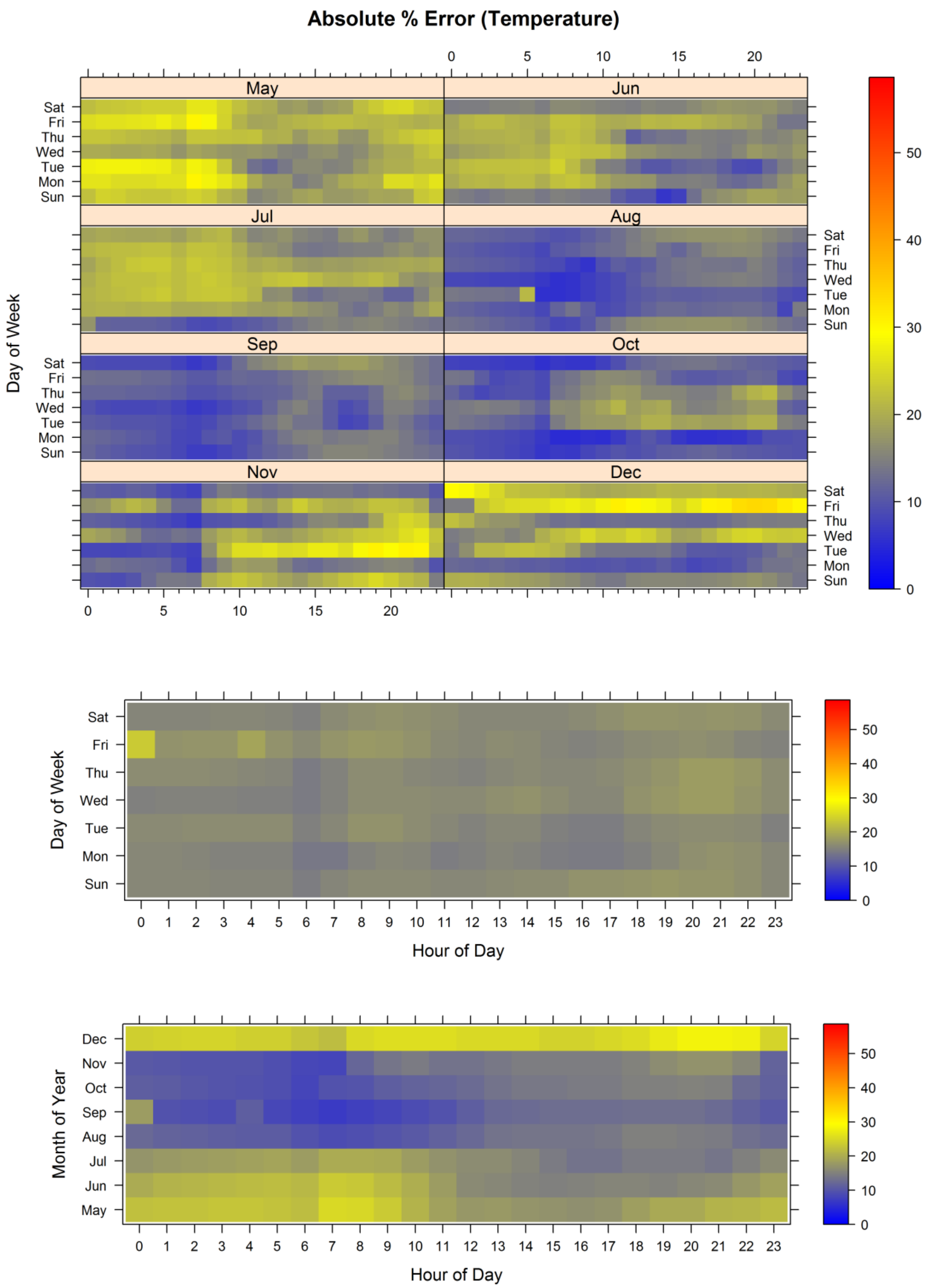
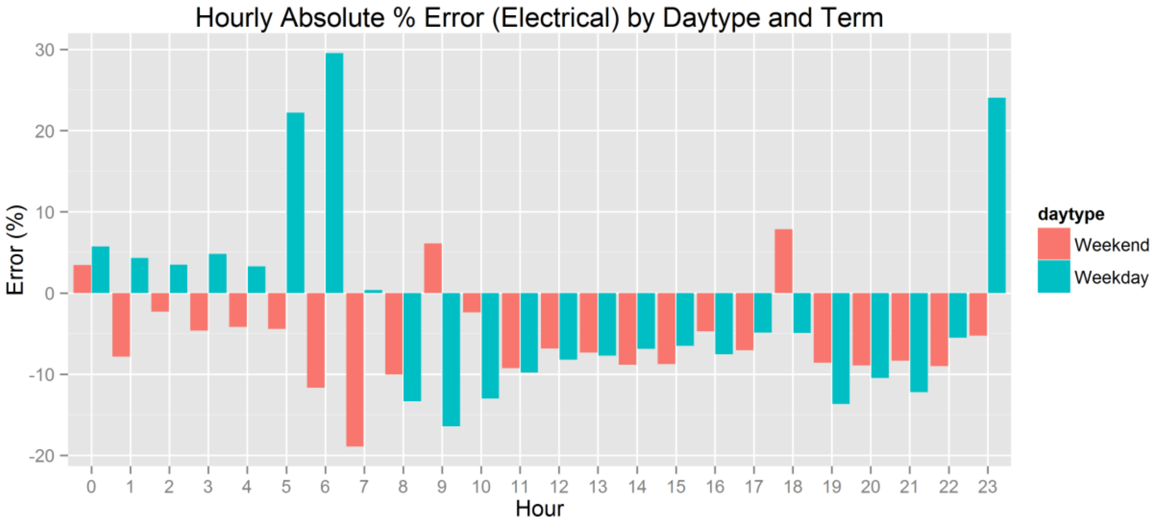


Figure 5-17: Absolute % Error (Temperature) - Rev 25



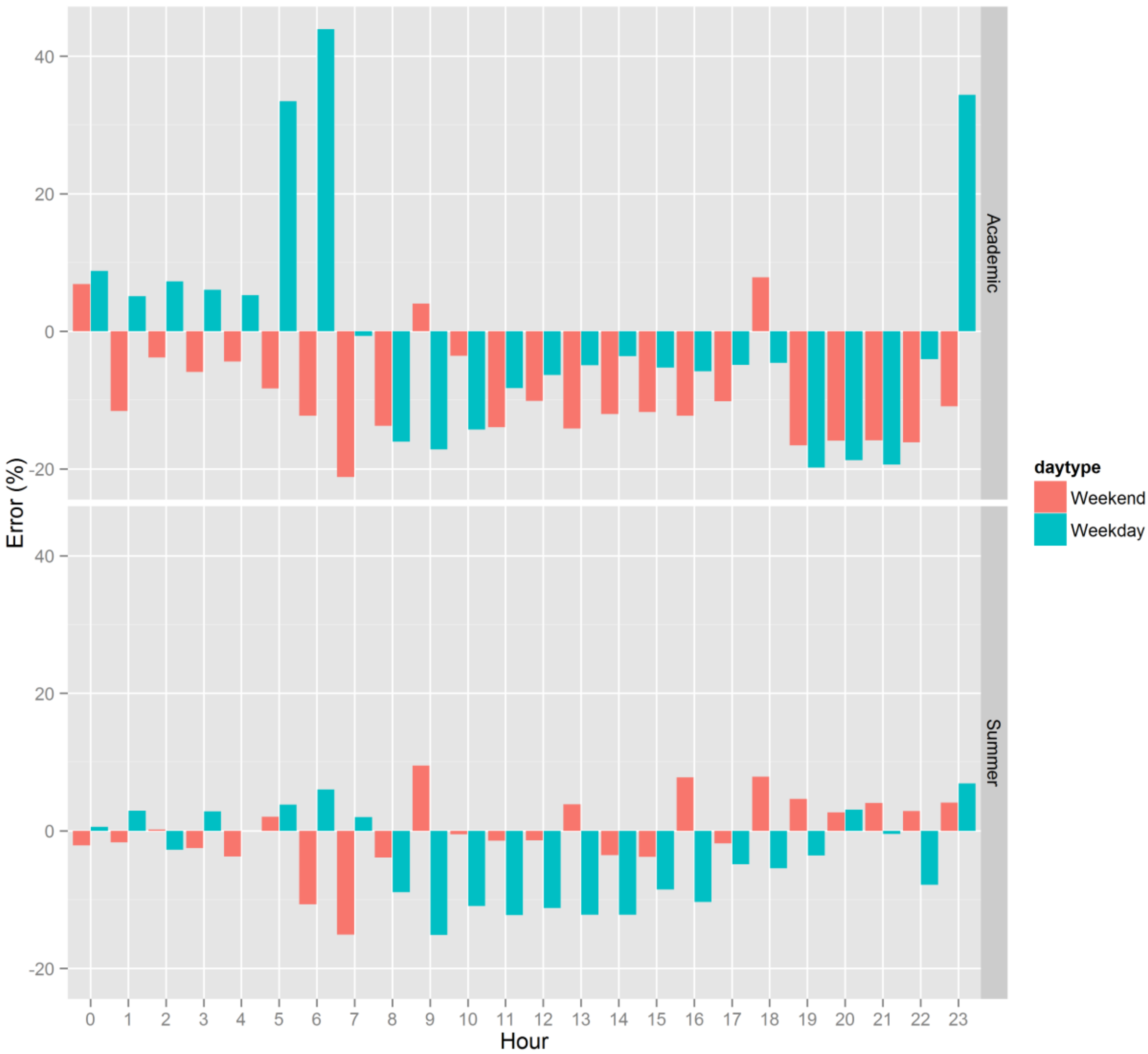
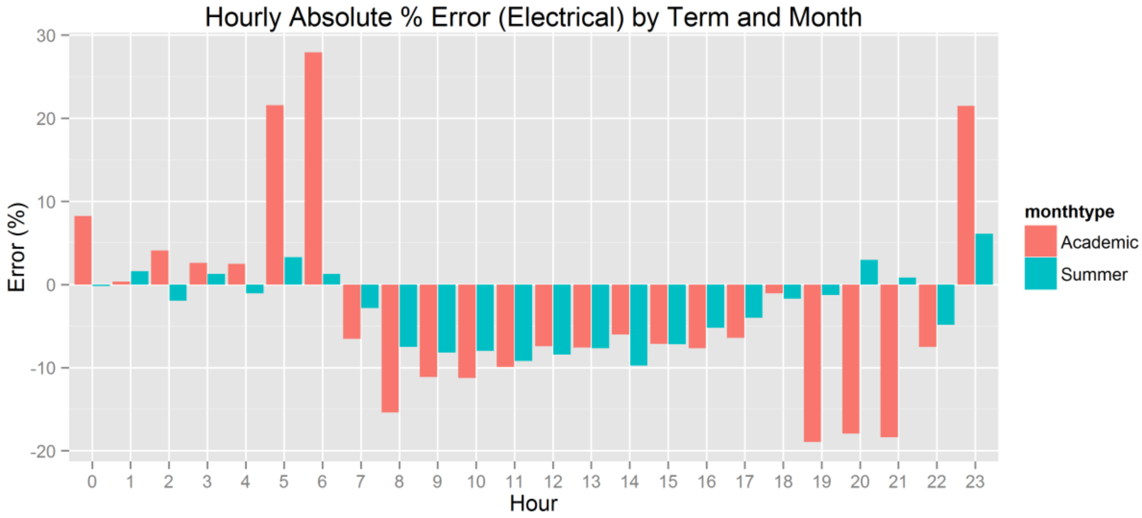


Figure 5-18: Hourly Absolute % error by day type and term (Electrical) - Rev 25



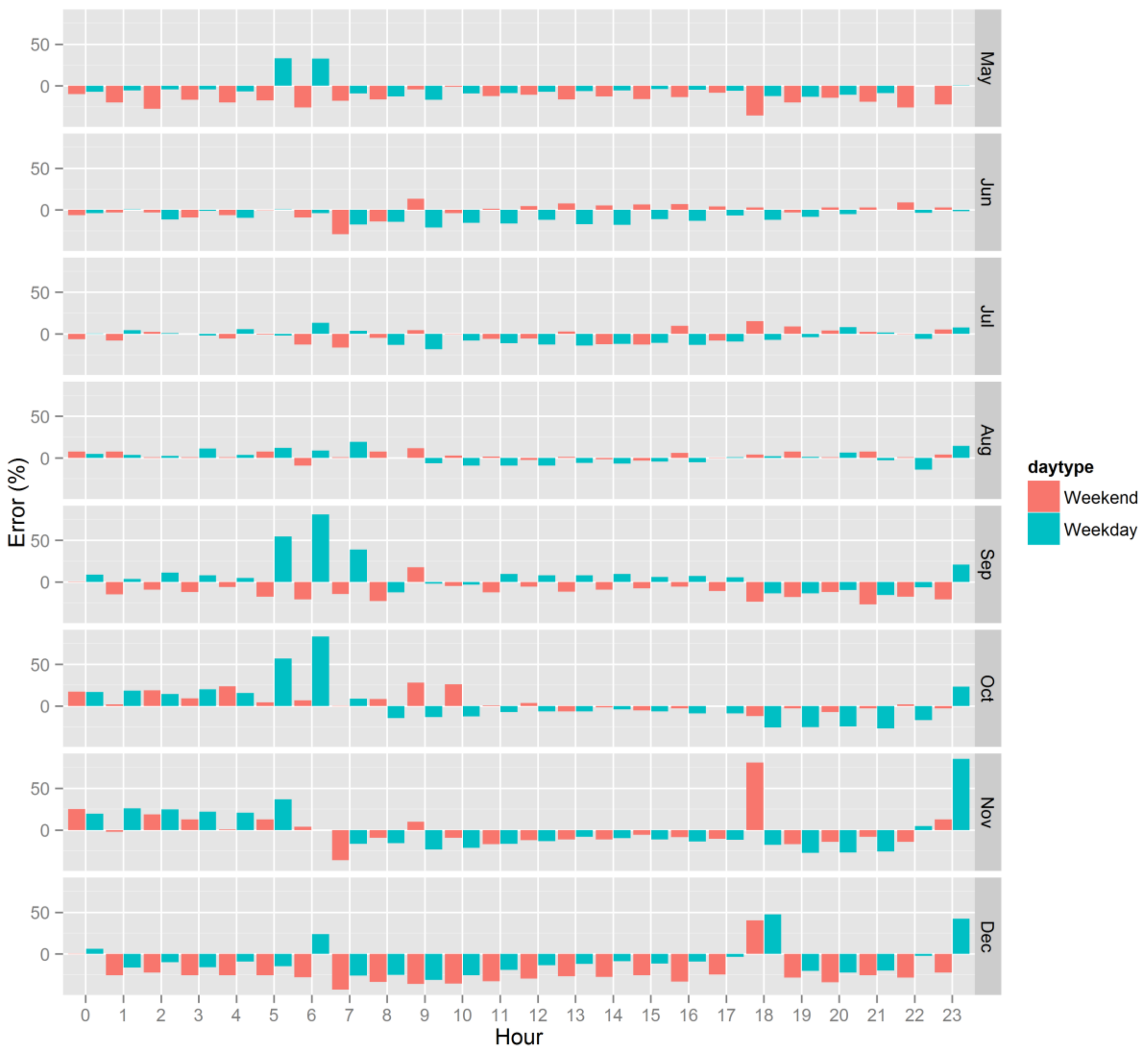


Figure 5-19: Hourly Absolute % error by term and month (Electrical) - Rev 25

5.4.5 Results Discussion

It should be noted that some error is unavoidable without compromising the representative nature of the model. For example, some discrepancies are due to volatile fluctuations in energy consumption throughout particular periods during the year. It is possible to modify the model to capture these fluctuations (e.g. by increasing the resolution of lighting and equipment schedules, and further segregation of time periods), but this would yield a highly over-specified model. In such a case, the model may represent the particular data set to a high degree of accuracy, but would be of little practical use in modelling generic predicted building performance.

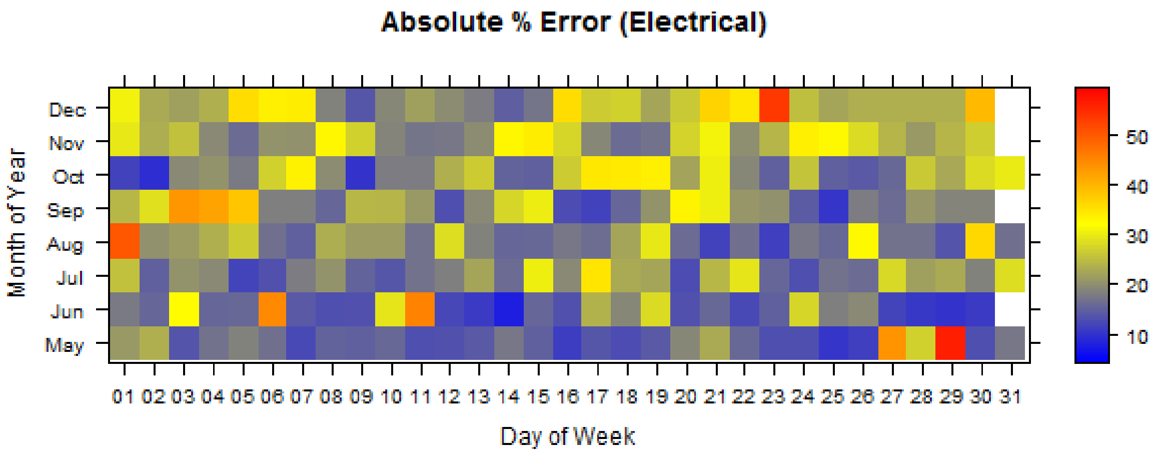


Figure 5-20: Remaining Model discrepancies after Iterative Model Update - Rev. 23

In this case, further refinement of the model would require interpolation from existing measured building data in order to obtain a better fit. It is theoretically possible to generate individual schedules for each day of the year and using the difference between measured and simulated data in order to compute the 'correct' input at each hour. However, as discussed previously, this would yield an over-specified solution. Therefore, the challenge is to develop a model which captures major characteristics in building energy consumption trends, while maintaining a reasonable level of input detail. Another issue which may result in discrepancies is the granularity of available measured and simulated data (see Figure 5-21). In this case study, the building metering was capable of capturing average electrical energy use to the nearest kW for each hour, which simulated data may be generated at any frequency or detail level.

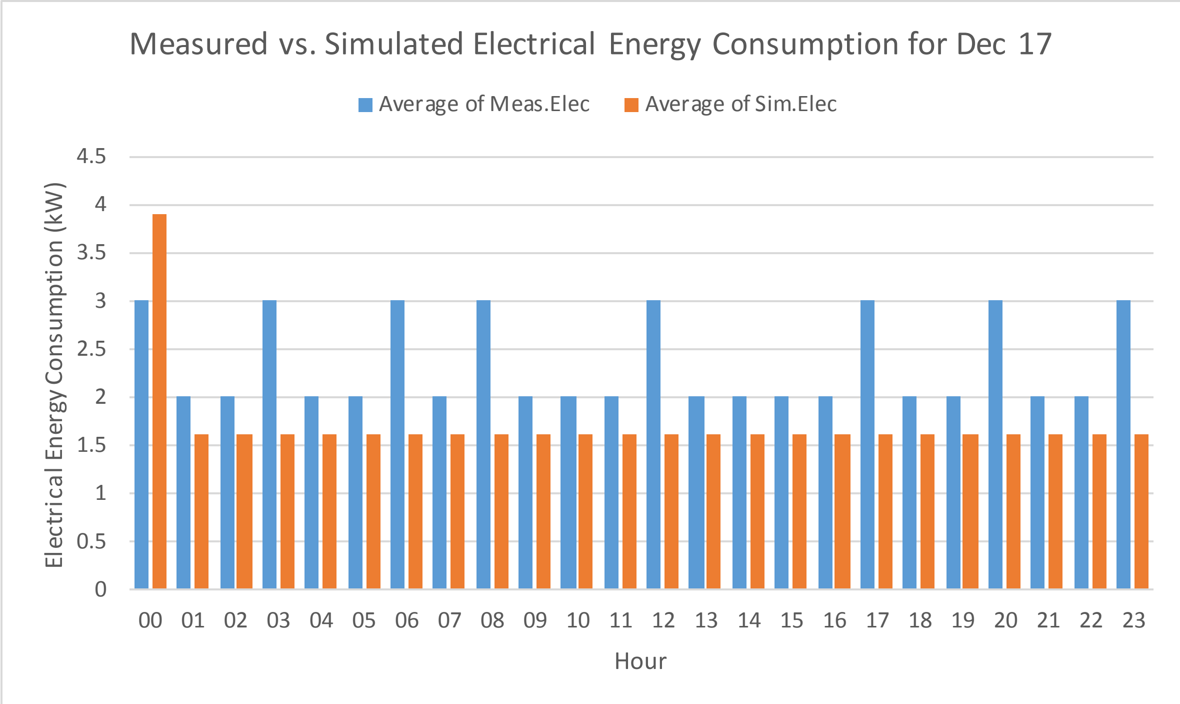


Figure 5-21: Measured vs. Simulated Electrical Energy Consumption (Dec 17 2011)

As evidenced in Figure 5-21, this granularity in data can have a significant impact on relative error values, particularly in a building with relatively low levels of energy consumption. Measured values are only logged to the nearest whole number, meaning that relative error may increase by up to 33% depending on the value recorded on a particular instance.

5.5 Latin Hypercube Monte-Carlo (LHMC) Search

At this stage, the input uncertainty specified during to model development stage is propagated through the model and used to generate the range of potential predictions. The top 10 ranking predictions, in terms of overall goodness-of-fit are automatically generated by comparing each of the randomly-generated simulation outputs with the measured building data. At present, these predictions only account for uncertainty within a limited number of model parameters, including:

- Occupancy schedules;
- Lighting and equipment Schedules;
- Material properties;
- HVAC equipment parameters.

By varying the inputs for each of these parameters randomly within the defined uncertainty ranges, it is possible to propagate this input uncertainty through to the model predictions. This process can be used to find the model (and associated inputs) which best represents the calibration problem, or can simply be used as a means of quantifying the risk and uncertainty associated with individual model predictions.

5.5.1 Calibrating model inputs

This approach can be used as a means of calibrating individual model inputs by fine-tuning their values around the best-guess estimate, within the defined ranges of uncertainty, and bounded by realistic maximum and minimum values. By ranking the best-fitting models according to goodness-of-fit (GOF) criteria (see section 3.8.5), it is possible to filter the models with the most likely values for individual parameters (see Table 5-3).

Table 5-3: Top 10 best-fitting simulation runs

Job	CVRMSE			NMBE			GOF		
	Heat	*Elec*	*Temp*	*Heat*	*Elec*	*Temp*	*A*	*B*	*Total*
915	69.936	8.671	15.114	53.081	7.707	14.683	31.949	24.545	29.867
515	69.938	9.993	15.148	53.045	9.179	14.728	32.251	24.923	30.186
659	71.409	9.374	14.522	54.525	8.493	14.172	32.727	25.347	30.646
706	71.212	9.759	14.578	54.345	8.929	14.266	32.731	25.39	30.66
489	72.903	7.917	14.563	55.63	6.855	14.227	33.068	25.429	30.919
7	72.362	8.911	14.229	55.452	7.984	13.887	33.031	25.606	30.937
369	72.405	9.06	14.49	55.803	8.14	13.944	33.083	25.794	31.023
627	73.107	8.236	14.1	56.263	7.212	13.709	33.213	25.771	31.113
766	72.337	10.07	14.99	55.057	9.258	14.516	33.287	25.782	31.171
64	73.416	8.488	14.277	56.35	7.502	13.831	33.397	25.874	31.276

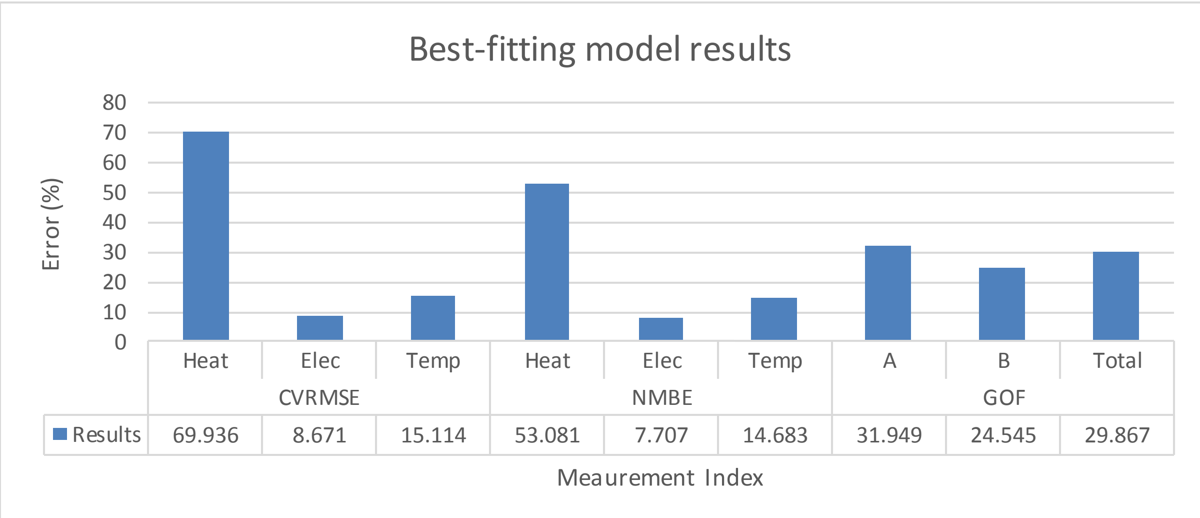


Figure 5-22: Best-fitting model results (Job Ref. 915)

Figure 5-22 presents the results of the best-fitting model, according to overall goodness-of-fit (GOF) criteria. The model represents the optimal values for the model 'tuning' parameters. This does not guarantee that the adjustments are correct. However, since the calibration problem is undetermined, it will never be possible to accurately define every input parameter. Therefore, this method simply provides a means of automatically tuning these values within acceptable ranges, in order to provide a better match to measured data.

Table 5-4: Model input parameter adjustments

ID	Name	Description	Search String	Mean	StdDev	Final	% change
P1098	Surface Construction Elements	Thickness	&&GeoThi1098&&	0.1	3.33E-03	0.1028	2.84
P1099	Surface Construction Elements	Conductivity	&&GeoCon1099&&	0.04	0.004	0.0435	8.74
P1100	Surface Construction Elements	Density	&&GeoDen1100&&	15	1.5	14.8857	-0.76
P1101	Surface Construction Elements	Specific Heat	&&GeoSpe1101&&	1450	145	1503.7886	3.71
P2101	Internal Gains	Lighting Level	&&ZonLig2101&&	7524	125.4	7560.7690	0.49
P2119	Internal Gains	Watts per Zone Floor Area	&&ZonWat2119&&	13	0.21666	13.2702	2.08
P2084	Airflow	Air Changes per Hour	&&ZonAir2084&&	2.5	0.33333	3.3729	34.92
P2103	Airflow	Flow Rate per Person	&&ZonFlo2103&&	0.00236	3.15E-04	0.0024	0.79
P2331	Fans	Fan Efficiency	&&ZonFan2331&&	0.74	0.037	0.6994	-5.49
P2332	Fans	Pressure Rise	&&ZonPre2332&&	750	37.5	791.4427	5.53
P2337	Fans	Motor Efficiency	&&ZonMot2337&&	0.9	0.045	0.8548	-5.02

5.5.2 Quantifying risk and uncertainty

Figure 5-23 illustrates how the methodology can be used to incorporate input uncertainty into predicted results. By allowing the input parameters to vary within the bounds of uncertainty, a range of predictions is generated. The graph shows the average predicted values as well as the maximum and minimum prediction within these uncertainty limits.

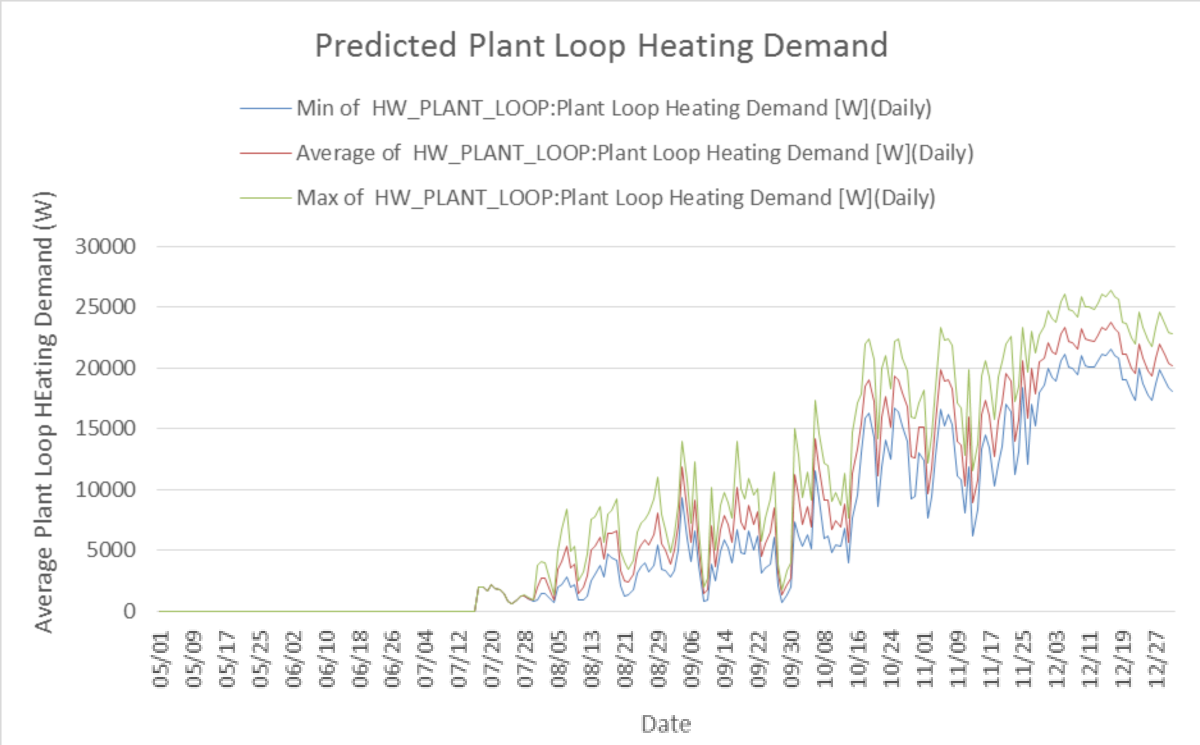


Figure 5-23: Uncertainty-based predicted plant-loop heating demand

The advantage of producing a range of possible results as such is that it clearly indicates the level of inherent uncertainty present in the model predictions, as opposed to a single model prediction. This is beneficial in cases such as retrofit assessment and ECM evaluation where the analyst needs a means of quantifying the risks of under-performance of these measures.

This data can also be represented as a cumulative probability plot, as shown in Figure 5-24. From this graph, we can say that the best estimate of average heat energy consumption for this day is 15.18kW (point estimate) with a 90% probability that the true value falls roughly within ±20% of this value. This is determined from the distributions of the simulated (predicted) heat energy consumption. In actuality, the measured daily heat energy consumption for 1st November was 300kWh, representing an hourly average value of 12.5kW, or around 8.7% cumulative probability (see Figure 5-25). However, it should be noted that due to the resolution of the heat energy measurement available, the true value could also be up to 340kWh (since the measurement resolution is 40kWh). This would represent an hourly average heat energy consumption of 14.167kW or around 30% cumulative probability. Therefore, it is important to also consider the measurement resolution of the metered data.

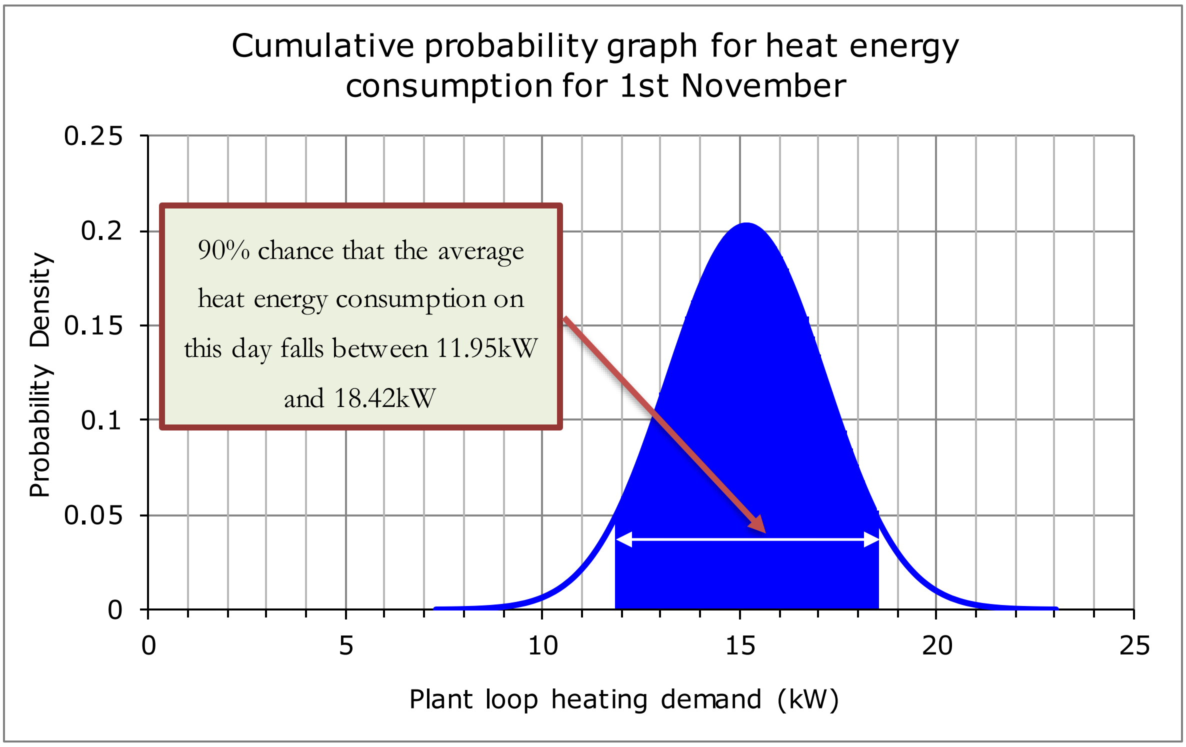


Figure 5-24: Cumulative probability graph for heat energy consumption

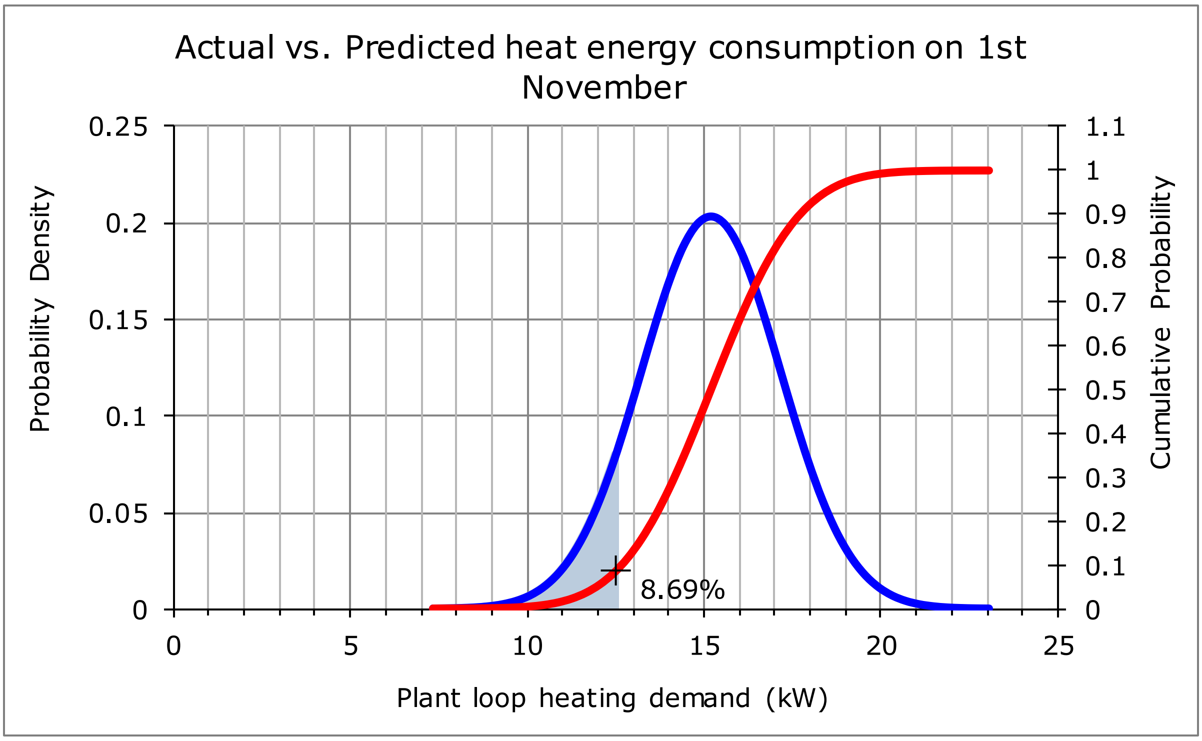


Figure 5-25: actual vs predicted heat energy consumption for 1st November, represented on cumulative probability graph

The above data for heat energy consumption still has a relatively high degree of uncertainty associated with predictions, due largely to the fact that there is still a high degree of uncertainty pertaining to the model input parameters. While this is not ideal, it does serve to demonstrate the level of detail required to reduce model uncertainty, as well as the potential performance range associated with building energy simulation model predictions. This is in line with findings from a US Green Building Council report comparing LEED building performance with baseline energy use predictions (Turner and Frankel 2008). This also further serves to highlight the need for high quality data in order to achieve accurate model predictions, as well as the need to express uncertainty along with model predictions.

5.6 Conclusions

This chapter presents the results of the application of the proposed evidence-based analytical optimisation approach to a mixed-mode University building. The results indicate that this is a viable approach to model development and calibration incorporating parameter uncertainty quantification. There are a number of novel aspects to the presented results:

- Measured and simulated data are automatically compared statistically at each model iteration in order to guide the development process. This process has been streamlined as much as is practically possible in order to aid the model calibration process;
- A novel set of statistical summary graphs have been developed which provide an overview of the current performance of each model simulation, as well as a means of tracking progress;
- Results are automatically archived to a source control repository for later retrieval, if required;
- A parametric study (LHMC search) is carried out based on model input uncertainty, and results are ranked according to goodness-of-fit (GOF) values for heating and electrical energy consumption, as well as zone temperatures;
- The results of the ranking process may be used to fine-tune individual model parameters based on analysis of the best-fitting models. This may be achieved by performing a frequency distribution analysis of input parameters for best fitting models (see further discussion in section 6.3.8);

- The process also enables risk and uncertainty quantification for final calibrated models, which is useful in assessing the performance risks associated with the implementation of energy conservation measures (ECM's);

There are still a number of improvements required in order to streamline this process, which currently requires significant manual effort at various stages of the process. This will be addressed in the future work section (see section 6.3)

Chapter 6:
Conclusions and Future Work

"Research is to see what everybody else has seen, and to think what nobody else has thought."

– Albert Szent-Gyorgyi, Physiologist

6.1 Chapter Overview

This section will present the overall conclusions and recommendations of this work, focussing specifically on the novel aspects of each of the main thesis chapters:

- Literature Review;
- Methodology;
- Case Study;
- Results.

6.1.1 Literature Review

The presented literature review provides a comprehensive overview of existing approaches to model calibration, covering both manual and automated approaches. Numerous approaches to model calibration have been suggested. However, no consensus standards exist. In addition, many of the current approaches to model calibration rely heavily on user knowledge, past experience, statistical expertise, engineering judgement, and an abundance of trial and error. Furthermore, when a model is established as being calibrated, the author often does not reveal the techniques used, other than stating the final result.

A systematic approach to model calibration is proposed, utilising the most promising mathematical, statistical and visualisation techniques identified during the literature review process. The proposed approach relies on a combination of:

1. Detailed site characterisation methods;
2. Model simplification techniques;
3. Advanced graphical techniques;
4. Optimisation techniques;
5. Evidence-based procedure;
6. Sensitivity Analysis;
7. Uncertainty quantification.

These techniques are adapted, extended and incorporated into a comprehensive sytematic evidence-based analytical optimisation approach for model calibration.

6.1.2 Methodology

This thesis proposes a new methodology for the calibration of detailed simulation models to measured data using a systematic, evidence-based approach. The proposed approach differs from existing approaches to model calibration in a number of ways:

- Detailed model development: The proposed approach uses detailed building and HVAC system information, as opposed to building templates, to create an accurate initial representation of the building;
- Structured evidence based approach: the proposed approach follows a systematic model evolution process, whereby each model input is updated according to a source of evidence. Each model update is logged and referenced using version control software;
- Source classification and uncertainty: Sources are assigned to a hierarchy depending on their expected level of accuracy, and used to classify parameter values and associated distributions;
- Sensitivity Analysis (SA): parameters which merit further investigation are automatically identified using an (optional) SA approach, combining source reliability classification and sensitivity indices;
- Parametric Analysis: since the calibration problem is under-determined, it is possible for many unique solutions to exist. Therefore, the proposed approach included a parametric grid search whereby a range of possible solutions are identified, as opposed to a single solution. This ;
- Uncertainty Characterisation: The methodology includes a measure of parameter uncertainty based on source evidence, which is propagated through the model development process to enable risk and uncertainty quantification of final model predictions.

6.1.3 Case Study

The presented case study highlights a number of important points:

- Data visualisation methods: the chapter presents an overview of various means of representing the high volume of data available in the building domain. Automated methods for producing these visual displays are also developed.
- Data transformation and tool-chain: the chapter also highlights the complexity of the data transformation process required for model calibration, primarily due to the use of

proprietary software and systems in the AEC (Architecture , Engineering & Construction) domain, as well as a fragmentation of the building design, commissioning and operation processes.

- Automated processing: the process employed in the presented case study is automated as far as practically possible given the time and resource limitations. Scripts and scheduled tasks are used at various stages in the model development and calibration process. These are presented in Figure 4-41 and Figure 4-42, and included in the thesis appendices.

6.1.4 Results

The results of the calibration process highlight the potential benefits of the proposed methodology in terms of systematically identifying model improvements, tracking changes and visualising the final results. There are a number of important points relating to the presented results:

- A summary of model error is automatically generated after each model iteration in order provide an overview of the current performance, as well as a means of tracking progress;
- Statistical summary graphs are used in conjunction with more modern colour carpet contour plots in order to effectively identify potential sources of model error;
- The results of the ranking process may be used to fine-tune individual model parameters based on analysis of the best-fitting models;
- An illustration of the uncertainty-based model predictions is also presented. This enables risk and uncertainty quantification for final calibrated models, which is useful in assessing the performance risks associated with the implementation of energy conservation measures (ECM's);

6.2 Discussion

This research has illustrated a number of key findings in relation to the realities of calibration of BES models to detailed measured data. The author concludes that the following merit discussion:

i. Data Acquisition and data quality
ii. Uncertainty in Building energy simulation models
iii. Statistical performance metrics

6.2.1 Data Quality and Accuracy

The reliability and accuracy of 'calibrated' BES models depends on the quality of the measured data used to create the model, as well as the accuracy and limitations of the tools used to simulate the building and its' systems.

Throughout the course of this study, it has been found that it is very difficult to obtain the level of data required for detailed calibration, even in modern buildings with relatively large quantities of data readily available. In addition, Building Management Systems are often not configured to collect and archive monitored data points. Since data storage and archiving incurs additional cost, it is typically up to the client to explicitly request this service from the BMS installers.

Limitations of simulation tools also influence the results and thus impact simulated performance data significantly. These limitations are either embedded in the simulation tool or caused by the particular use of a tool and are then included in input data (Maile 2010).

6.2.2 Uncertainty in BES Models

As highlighted by (Kaplan, McFerran, et al. 1990)), it will never be possible to identify the exact solution to the calibration problem. Due to its' highly underdetermined nature, it will always yield a non-unique solution (Carroll and Hitchcock 1993). This case study serves to highlight the level of uncertainty associated with individual model input parameters and, consequently, the final calibrated model.

This uncertainty creates a vast multi-dimensional solution space. The calibration approach outlined in this paper recognises this problem and uses random sampling techniques in an attempt to identify a selection of optimum solutions rather than just one.

Therefore, rather than using only one plausible calibrated solution to make predictions about the effect of intended energy conservation measures (ECMs), we use a small number of the most plausible solutions. Not only does this make it more likely to obtain a more robust prediction of the energy and demand reductions, but this will also allow us to determine the associated prediction uncertainty for each solution (Reddy et al. 2007a).

6.2.3 Hourly vs. Monthly Calibration

Currently, most studies analyse model error using monthly data (Reddy 2006). However, this approach may hide inaccuracies which only appear at hourly or daily resolutions (Raftery, Keane and Costa 2011). Therefore, this methodology attempts to reconcile detailed hourly measurements with simulated data to provide a model which accurately represents actual building design and operation, while also accounting for realistic value ranges and parameter uncertainty.

6.3 Future Work

This work forms the basis for an automated systematic approach to model calibration. However, in order to enable more widespread adoption of building energy models, it is necessary to first tackle some of the barriers to model calibration. This section provides a sample of some of the future work required to further streamline this process and help to increase the adoption and reliability of energy simulation models in the AEC industry.

6.3.1 Measurement Uncertainty

According to the IPMVP guidelines (Efficiency Valuation Organisation (EVO) 2010), uncertainty can arise from three sources: modelling, sampling and measurement:

- Modelling: errors in mathematical modelling of system;
- Sampling: sampling errors arise when a portion of a population of actual values is measured, or a biased sampling approach is used (e.g. occupancy in the presented case study was sampled at random intervals, but it was infeasible to continuously monitor occupancy. Therefore, sampling bias is introduced.);
- Measurement: accuracy of sensors, data tracking errors, sensor drift, imprecise measurements etc. can all lead to increased measurement uncertainty. This is important, as the model is being calibrated to these measurements, so they are assumed to be 100% accurate, which is not correct.

In this thesis, I have discussed how to address modelling uncertainty (see section 2.3.5), particularly specification uncertainty (i.e. the uncertainty associated with model input parameters). However, it is important to also consider the uncertainty on the other side of the equation – measurement uncertainty. This can arise from:

- Poor meter precision, or insufficient resolution of data;
- Poor placement of metering equipment, so it does not capture a representative measurement;
- Data telemetry or communication errors.

Errors associated with measurement uncertainty should be incorporated when using calibrated models as a baseline for predicting energy savings, as the confidence of the model predictions may be affected by this uncertainty. The IPMVP guidelines (Efficiency Valuation Organisation (EVO) 2010) provide a rough guideline as to acceptable uncertainty pertaining to savings predictions for energy conservation measures (ECM's). Specifically, the savings need to be larger than twice the *standard error* of the *baseline* value. If the *variance* of the *baseline* data is excessive, the unexplained random behaviour in energy use of the facility or system is high, and any single savings determination is unreliable. Where you cannot meet this criterion, consider using:

- more precise measurement equipment;
- more independent variables in any mathematical model;
- larger sample sizes, or;
- an IPMVP Option that is less affected by unknown variables.

This guidance provides a reasonable starting point for inclusion of a formal procedure for capturing measurement uncertainty within the proposed calibration methodology.

6.3.2 BIM to BES

Building Information Modelling (BIM) represents a significant step towards the aggregation of the type of detailed information required for the construction of building energy simulation (BES) models. However, BIM models are typically built on modelling software using proprietary data formats (e.g. ArchiCAD, AutoDesk AutoCAD). These formats require transformation in order to become compatible with building energy simulation tools (such as EnergyPlus). A number of open building information transfer protocols (e.g. ifc, gbXML) have been developed specifically to tackle the problem of building data storage and transformation. However, the conversion of proprietary building models to open formats (such as ifc) often results in data corruption or loss of information. Therefore, at present, building energy models are often constructed after BIM models on separate modelling tools. This is both inefficient (due to duplication of work) and detrimental to the quality of the final model (due to potential failure to transfer elements of the BIM model to the BES model). This issue requires coherent industry engagement in order to increase compatibility between BIM and BES tools.

6.3.3 Automated Schedule Generation

One of the most time consuming elements of model development is the creation of detailed schedules for building parameters, such as occupancy, infiltration, lighting and equipment loads. Part of the problem is the effort required to characterise a building such that the model captures past trends, while remaining sufficiently generic so as to be capable of also predicting future performance (i.e. not over-specified). This can be a difficult task to perform manually, requiring creation and adjustment of detailed building schedules for each characteristic period (e.g. weekdays, weekends, holidays, half-days etc.). However, it is possible to use measured building data as a 'training set' which may be used to identify the main characteristic time-periods, and automatically construct appropriate schedules for these time-periods.

6.3.4 Occupancy Profiles

Occupancy tends to be one of the more difficult parameters to capture in buildings without expensive sensing equipment or security systems. It is also one of the most influential variables relating to energy-use and efficient building control. Therefore, it would be useful to devise strategies for using existing building and environmental data in order to infer human presence in spaces. It was noted during this study that there was a strong correlation between Carbon dioxide correlation and human presence. : However, CO_2 concentration is also heavily influenced by levels of outdoor background CO_2, wind speed and infiltration rate. With further investigation though this may provide some useful for occupancy inference, particularly if combined with other corollary sources of information. For example, the PC usage trends presented in 4.4.1.1 are a useful source of data which are not currently captured within the building management system. Other sources of data may include, but are not limited to, temperature data, Wi-Fi usage patterns, security systems (IR sensors), cell-phone radio connections, GPS data, social media (and automated social check-in applications – e.g. Facebook, FourSquare, Google Android Location tracking).

6.3.5 Occupant-driven Natural Ventilation

Occupant-driven natural ventilation, in terms of air-changes per hour (ACH), can be notoriously difficult to predict due to the following (Hyun et al. 2008):

- Stochastic nature of weather: Natural air- flows have two driving forces: buoyancy and wind. Since wind speed and direction change rapidly and the temperature difference between indoor and outdoor also fluctuates diurnally as well as annually, it is difficult to quantify such influences on natural ventilation. Hajdukiewicz et al. (2013) present a methodology for calibrating a CFD model of a naturally-ventilated space to measured data using a range of field measurements (e.g. air temperature, air velocity etc.). While this methodology is focussed on the design of naturally ventilated buildings, it provides an insight into the complex inter-relationships between factors which affect infiltration and ventilation rates in naturally ventilated buildings.

- Occupant's behaviour: An occupant's schedule in opening/closing windows and doors has a significant role in controlling the natural airflow rate. This behaviour is the result of a combination of human thermal comfort characteristics (personal and environmental factors) as well as individual psychological characteristics. These behavioural influences have not been measured extensively in field research. Coakley et al. (2013) presents a study of human response to thermal conditions, based on the work of P.O. Fanger (Fanger 1970; de Dear 2004). However, further study is required in order to develop a reliable method of correlating natural ventilation rates with these environmental and behavioural characteristics, to within a reasonable degree to accuracy in the absence of direct measurement. Simply assuming that natural ventilation occurs only through crack infiltration (when all the doors and windows are closed) is not realistic and measures only the minimum airflow rate. This has been the subject of a number of recent academic publications in the field of building energy research. (Bourgeois et al. 2006; Tanimoto et al. 2008; Hoes et al. 2009; Yun et al. 2009; Antretter et al. 2011; Haldi and Robinson 2011; Robinson and Haldi 2012; Dar et al. 2012; Fabi et al. 2013)
- Building components: Each building in unique in terms of its construction (quality of workmanship, air-tightness) and components (e.g. doors, windows, ventilation ducts, etc.) In addition, each dwelling has a different local environment, such as the natural/urban surroundings, orientation, distance from the ground, shading etc..
- Uncertainties in simulation parameters: Simulation parameters such as discharge coefficient, flow coefficient, etc., used to simulate airflow phenomena inside and around the building cannot be predicted accurately.

It is therefore infeasible to fully address this topic within the scope of this research. However, it will form the focus of future investigative studies, particularly in the area of thermal comfort and occupant behaviour field studies, following on from the work presented in Coakley et al. (2013).

6.3.6 Artificial Neural Networks

Traditional whole building energy simulation requires a large number of detailed inputs and simplifications, particularly in relation to annual scheduling of variable input parameters. In the presented case study, occupancy and infiltration were major influencing factors in determining building performance. However, due to the high level of uncertainty and variability associated with these parameters, it is impossible to reduce them to standard yearly schedules without compromising model quality.

A preferable approach may involve the use of machine learning algorithms, or Artificial Neural Networks (ANN) which can help predict occupancy and infiltration based on a set of training data. This ANN model may then be substituted for the existing deterministic model schedules. Alternatively, the ANN model may be used to generate these deterministic model schedules to a greater degree of accuracy. There is active research in this area (Edwards et al. 2012; Edwards 2013; Sterling et al. 2014), which has the potential to significantly reduce the dimensionality of the model input space, while allowing for better representation of variable model inputs.

6.3.7 Combined Sensitivity/Uncertainty Analysis

Uncertainty analysis (UA) may be used to identify how the uncertainty in an output can be allocated to uncertainty in the input parameters of a process or model, while sensitivity analysis (SA) can be used to identify the input parameters which exhibit a high degree of influence on the model output. These procedures are often performed in conjunction with one another, and their combination is particularly conducive to the LHMC approach presented in this thesis. There is a case for further investigation of how parameter uncertainty and sensitivity indices may be combined in order to further improve the iterative model development process, as well as the final LHMC search. For example, a simple approach may be to combine the input uncertainty (U_i) and the input sensitivity (S_i) to provide an overall parameter index that could be used to guide further investigation of individual parameters which exhibit a large degree of influence over the model outputs, or have a high degree of uncertainty, or a combination of both. There are a number of academic papers which focus on the application of formal UA and SA procedures to building energy modelling (V. Corrado and Mechri 2009; Hopfe 2009; Struck and Kotek 2009; Vincenzo Corrado and Mechri 2009; Zhao et al. 2011; Eisenhower and O'Neill 2012; Spitz et al. 2012).

6.3.8 Input parameter frequency distribution

While a combined SA/UA procedure can provide guidance as to which input parameters require further investigation, it does not provide any guidance as to the optimal value for these parameters. Therefore, it may be useful to carry out a frequency analysis of the distribution of the sampled input parameters from the LHMC search (Section 3.8), and compare the values found in the top-ranking solutions to the specified probability density function. This may help highlight major discrepancies between assumed input distributions and optimal values. Such discrepancies may indicate an error in specification of the mean or distribution type for the parameter. Such an analysis may also be automated, by specifying acceptable ranges for discrepancies between specified and optimal probability density functions (e.g. mean, $\mu \pm 20\%$ and standard deviation, $\sigma \pm 30\%$).

6.3.9 Short-term weather forecast integration

In order to use energy simulation models for the purpose of model-predictive control (MPC), it is essential to have accurate up-to-date weather and climate information for the region of interest. In this respect, predicted weather data, for a period of at least 3 days, will allow accurate simulation of predicted building performance, thus enable improved response.

This type of control has the ability to reverse the current building control dynamic, from reactive, to pro-active. In other words, the building is no longer a passive observer of external conditions, but rather is able to actively respond to changing external influences. This may also include variables such as occupancy and external economic influences. There are a number of websites which provide local weather forecasting services. However, we are particularly interested in those services which also allow interaction with their data through open API's, such as that provided by forecast.io (see Appendix A.12).

Bibliography

Abbas, M. 1993. Development of graphical indices for building energy data. M.S. Thesis. Energy Systems Rpt No. ESL-TH-93/12-02. Texas A&M University, College Station.

Abdou, O. 1997. Effects of luminous environment on worker productivity in building spaces. *Journal of architectural engineering* 3(3), pp. 124–132.

Adnot, J. 2007. Field benchmarking and Market development for Audit methods in Air Conditioning (AUDITAC) final report. *Intelligent Energy Europe*.

Akbari, H. 1995. Validation of an algorithm to disaggregate whole-building hourly electrical load into end uses. *Energy* 20(12), pp. 1291–1301.

Akbari, H., Davis, S., Huang, J., Dorsano, S. and Winnett, S. 1992. Cooling our communities: a guidebook on tree planting and light-colored surfacing.

Akbari, H., Heinemeier, K., Flora, D. and LeConiac, P. 1988. Analysis of commercial whole-building 15-minute-interval electric load data. *ASHRAE Transactions* 94(2).

Akbari, H. and Konopacki, S. 1998. Application of an end-use disaggregation algorithm for obtaining building energy-use data. *Journal of solar energy engineering* 120(3), pp. 205–210.

Antretter, F., Mayer, C. and Wellisch, U. 2011. An approach for a statistical model for the user behaviour regarding window ventilation in residential buildings. In: *Proceedings of the 12th International IBPSA Conference*. Sydney, Australia, pp. 14–16.

Aronica, G., Hankin, B. and Beven, K. 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Advances in Water Resources* 22(4), pp. 349–365.

ASHRAE 2009. ASHRAE Handbook: Fundamentals (SI Edition). American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA 30329.

ASHRAE 2002. Guideline 14-2002: Measurement of Energy and Demand Savings. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA 30329.

ASHRAE 2011. Procedures for Commercial Building Energy Audits. 2nd Editio. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA.

ASHRAE 1983. RP-351-Energy Audit Input Procedures and Forms. Atlanta, GA.

ASHRAE 2004a. Standard 55-2004: Thermal environmental conditions for human occupancy. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA 30329.

ASHRAE 2004b. Standard 62.1-2004: Ventilation for Acceptable Indoor Air Quality. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA 30329.

Attia, S., Gratia, E., De Herde, A. and Hensen, J.L.M.M. 2012. Simulation-based decision support tool for early stages of zero-energy building design. *Energy & Buildings* 49(0), pp. 2–15.

Augenbroe, G. 2002. Trends in building simulation. *Building and Environment* 37(8-9), pp. 891–902.

Balcomb, J., Burch, J. and Subbarao, K. 1993. Short-term energy monitoring of residences. *ASHRAE Transactions* 99(2), pp. 935–944.

Bedford, T. 1936. The warmth factor in comfort at work. Industrial health research board report No. 76.

Bertagnolio, S. and Lebrun, J. 2008. Simulation of a building and its HVAC system with an equation solver: application to benchmarking. *Building Simulation* 1(3), pp. 234–250.

Beven, K. 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources* 16(1), pp. 41–51.

Bodart, M. and Herde, A. De 2002. Global energy savings in of offices buildings by the use of daylighting. *Energy and Buildings* 34(5), pp. 421–429.

Booth, a. T., Choudhary, R. and Spiegelhalter, D.J. 2013. A hierarchical Bayesian framework for calibrating micro-level models with macro-level data. *Journal of Building Performance Simulation* 6(4), pp. 1–26.

Bourgeois, D., Reinhart, C. and Macdonald, I. 2006. Adding advanced behavioural models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control. *Energy and Buildings* 38(7), pp. 814–823.

Bou-Saada, T.E. and Haberl, J.S. 1995. An improved procedure for developing calibrated hourly simulation models. In: *Proceedings of the 5th International IBPSA Conference*. Montréal, Canada.

Brager, G.S., Paliaga, G. and de Dear, R. 2004. Operable windows, personal control and occupant comfort. *ASHRAE Transactions* 110(2), pp. 17–35.

British Petroleum (BP) 2012. BP Statistical Review of World Energy. London, United Kingdom.

British Standards Institution 2000. BS EN 12524:2000 - Building materials and products. Hygrothermal properties. Tabulated design values.

Bronson, D.J., Hinchey, S.B., Haberl, J.S. and O'Neal, D.L. 1992. Procedure for calibrating the DOE-2 simulation program to non-weather-dependent measured loads. In: *ASHRAE Winter Meeting*. Anaheim, CA, USA, pp. 636–652.

BuildDesk Ltd 2013. BuildDesk U.

Building Control Authority 2013. Building Control Regulations (Part L). Ireland.

Burch, J., Subbarao, K., Lekov, A., Warren, M. and Norford, L. 1990. Short-term energy monitoring in a large commercial building. *ASHRAE Transactions* 96(1), pp. 1459–1477.

Burhenne, S., Elci, M. and Jacob, D. 2010. Sensitivity Analysis with Building Simulations to Support the Commissioning Process.

Burhenne, S., Jacob, D. and Henze, G.P. 2011. Sampling based on Sobol' Sequences for Monte-Carlo Techniques Applied to Building Simulations. In: *Proceedings of the 12th International IBPSA Conference*. Sydney, Australia.

Carbon Trust 2005. The UK Climate Change Programme: potential evolution for business and the public sector.

Carroll, W.L., Birdsall, B.E., Hitchcock, R.J. and Kammerud, R.C. 1989. RESEM: An evaluation tool for energy retrofits in institutional buildings. In: *Proceedings of the 1st International IBPSA Conference*. Vancouver, Canada.

Carroll, W.L. and Hitchcock, R.J. 1993. Tuning simulated building descriptions to match actual utility data: methods and implementation. *ASHRAE Transactions* 99(2), pp. 928–934.

CASCADE Consortium 2012. Deliverable D2.1 CASCADE Methodology for Energy Efficient Airports.

CEC 2000. Guide to Preparing Feasibility Studies of Energy Efficiency Projects. Report P400-00-002. California Energy Commission (CEC), California, USA.

Christensen, C. 1984. Digital and color energy maps for graphic display of hourly data. In: *Proc. of 9th Natl. Passive Solar Conference*. Columbus, Ohio.

CIBSE 1998. AM11:2008 Building energy and environmental modelling Applications Manual. The Chartered Institution of Building Services Engineers (CIBSE) ed.

CIBSE 2006a. CIBSE Guide A: Environmental design. 7th Ed. The Chartered Institution of Building Services Engineers (CIBSE) ed. London, United Kingdom.

CIBSE 2005. CIBSE Guide B: Heating, ventilating, air conditioning and refrigeration. The Chartered Institution of Building Services Engineers (CIBSE) ed. London, United Kingdom.

CIBSE 2007. CIBSE Guide C: Reference Data. The Chartered Institution of Building Services Engineers (CIBSE) ed. London, United Kingdom.

CIBSE 2006b. KS6:2006 Comfort. The Chartered Institution of Building Services Engineers (CIBSE) ed. London, United Kingdom.

Clarke, J.A., Strachan, P.A. and Pernot, C. 1993. An approach to the calibration of building energy simulation models. *ASHRAE Transactions* 99, pp. 917–917.

Coakley, D., Corry, E. and Keane, M.M. 2013. Validation of Simulated Thermal Comfort using a Calibrated Building Energy Simulation (BES) model in the context of Building Performance Evaluation & Optimisation. In: *ICEBO - International Conference for Enhanced Building Operations*. Montréal, Canada.

Coakley, D., Raftery, P. and Keane., M.M. 2014. A review of methods to match building energy simulation models to measured data. *Renewable and Sustainable Energy Reviews*.

Coakley, D., Raftery, P. and Molloy, P. 2012. Calibration of Whole Building Energy Simulation Models: Detailed Case Study of a Naturally Ventilated Building Using Hourly Measured Data. In: *Building Simulation and Optimization*. ISBN: 978-1-897911-42-6, Loughborough, UK, pp. 57–64.

Corrado, V. and Mechri, H.E. 2009. Uncertainty and Sensitivity Analysis for Building Energy Rating. *Journal of Building Physics* 33(2), pp. 125–156.

Corrado, V. and Mechri, H.E. 2009. Uncertainty and Sensitivity Analysis for Building Energy Rating. *Journal of Building Physics* 33, pp. 125–156.

Crawley, D. and Lawrie, L. 2001. EnergyPlus: creating a new-generation building energy simulation program. *Energy & Buildings* 33(4), pp. 319–331.

Crawley, D.B.D., Hand, J.W.J., Kummert, M. and Griffith, B.T. 2008. Contrasting the capabilities of building energy performance simulation programs. *Building and Environment* 43(4), pp. 661–673.

CREP 2009. Guidance on the Development, Evaluation, and Application of Environmental Models.

Crompton, P. and Wu, Y. 2005. Energy consumption in China: past trends and future directions. *Energy Economics*.

Curry, E., Donnell, J.O., Corry, E., Hasan, S., Keane, M. and Riain, S.O. 2013. Linking Building Data in the Cloud: Integrating Cross-Domain Building Data using Linked Data. *Advanced Engineering Informatics* 27(2), pp. 206–219.

Dar, U.I., Georges, L., Sartori, I. and Novakovic, V. 2012. Influence of user-behaviour on the performance of the building and the energy supply system: Investigation of heating. In: *First Building Simulation and Optimization Conference*. Loughborough, UK, pp. 364–371.

De Dear, R. 2004. Thermal comfort in practice. *Indoor air* 14(s7), pp. 32–9.

DECC 2012. The Green Deal - A summary of the Government's proposals.

Deringer, J., Iyer, M. and Huang, Y. 2004. Transferred just on paper? Why doesn't the reality of transferring/adapting energy efficiency codes and standards come close to the potential. In: *Proc. 2000 ACEEE Summer Study on Energy Efficiency in Buildings*. Pacific Grove, CA.

DesignBuilder Software Ltd 2011. DesignBuilder Simulation + CFD Training Guide.

Dhakal, S. 2009. Urban energy use and carbon emissions from cities in China and policy implications. *Energy Policy* 37(11), pp. 4208–4219.

Diamond, S. and Hunn, B. 1981. Comparison of DOE-2 computer program simulations to metered data for seven commercial buildings. *ASHRAE Transactions* 87(1), pp. 1222–1231.

Doney, S., Balch, W., Fabry, V. and Feely, R. 2009. Ocean Acidification: A Critical Emerging Problem for the Ocean Sciences. *Oceanography* 22(4), pp. 16–25.

Doty, S. 2009. Building Operations: Balancing Energy Efficiency with Indoor Air Quality. *Energy Engineering* 106(4), pp. 31–52.

Edwards, R.E. 2013. Automating Large-Scale Simulation Calibration to Real-World Sensor Data. University of Tennessee, Knoxville.

Edwards, R.E., New, J. and Parker, L.E. 2012. Predicting future hourly residential electrical consumption: A machine learning case study. *Energy and Buildings* 49, pp. 591–603.

Efficiency Valuation Organisation (EVO) 2010. International Performance Measurement & Verification Protocol: Concepts and Options for Determining Energy and Water Savings Volume 1. 1(January).

Eisenhower, B. and O'Neill, Z. 2012. Uncertainty and sensitivity decomposition of building energy models. *Journal of Building Performance Simulation* 5(3), pp. 37–41.

Eisenhower, B., O'Neill, Z., Fonoberov, V.A. and Mezić, I. 2012. Uncertainty and sensitivity decomposition of building energy models. *Journal of Building Performance Simulation* 5(3), pp. 171–184.

Eisenhower, B., O'Neill, Z., Narayanan, S., Fonoberov, V. a. and Mezić, I. 2012. A methodology for meta-model based optimization in building energy models. *Energy and Buildings* 47(0), pp. 292–301.

EnergyXperts 2012. Experts System for an Intelligent Supply of Thermal Energy in Industry and other Large-Scale Applications [Online]. Available at: http://www.einstein-energy.net/index.php.

ESRU 1974. ESP-r. Energy Systems Research Unit, University of Strathclyde, Glasgow, Scotland G1 1XJ.

European Parliament 2006. 2006/32/EC: Energy end-use efficiency and energy services Directive.

European Parliament 2002. Directive 2002/91/CE of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings.

Evander, A., Sieböck, G. and Neij, L. 2004. Diffusion and development of new energy technologies: lessons learned in view of renewable energy and energy efficiency end-use projects in developing countries.

EVO 2007. International Performance Measurement & Verification Protocol. Efficiency Valuation Organisation.

Fabi, V., Andersen, R.V., Corgnati, S.P. and Olesen, B.W. 2013. A methodology for modelling energy-related human behaviour: Application to window opening behaviour in residential buildings. *Building Simulation* 6(4), pp. 415–427.

Fanger, P.O. 1970. Thermal comfort. Analysis and applications in environmental engineering. Danish Press.

Fisk, D. 1981. Thermal control of buildings. *Applied Science*, pp. 73–74.

Florita, A.R. and Henze, G.P. 2013. Whole Building Fault Diagnostics - A Bayesian Approach. In: *Intelligent Building Operations Workshop*. Boulder, CO.

Funtowicz, S. and Ravetz, J. 1990. Uncertainty and quality in science for policy. Vol. 15. Springer.

Ganji, A.R. and Gilleland, B. 2002. Investment Grade Energy Audit. In: *25th World Energy Engineering Congress*. Atlanta, GA, pp. 1–5.

Garcia Sanchez, D., Lacarrière, B., Musy, M. and Bourges, B. 2014. Application of sensitivity analysis in building energy simulations: Combining first- and second-order elementary effects methods. *Energy and Buildings* 68, pp. 741–750.

Goodwin, G. and Payne, R. 1977. Dynamic system identification: experiment design and data analysis. Academic Press, Inc., San Diego, California.

Grasman, J. and Straten, G. 1994. Predictability and non-linear modelling in natural sciences and economics.

Haberl, J. and Abbas, M. 1998a. Development of graphical indices for viewing building energy data: Part I. *Journal of Solar Energy Engineering* 120(3), pp. 156–161.

Haberl, J. and Abbas, M. 1998b. Development of graphical indices for viewing building energy data: Part II. *Journal of solar energy engineering* 120(3), pp. 162–167.

Haberl, J., Sparks, R. and Culp, C. 1996. Exploring new techniques for displaying complex building energy consumption data. *Energy and Buildings* 24(1), pp. 27–38.

Haberl, J.S., Bronson, D.J., Hinchey, S.B. and O'Neal., D.L. 1993. Graphical tools to help calibrate the DOE-2 simulation program to non-weather dependent measured loads. *ASHRAE Journal* 35(1), pp. 27–32.

Haberl, J.S.J. and Bou-Saada, T.E. TE 1998. Procedures for calibrating hourly simulation models to measured building energy and environmental data. *Journal of solar energy engineering* 120(3), pp. 193–204.

Hadley, D. 1993. Daily variations in HVAC system electrical energy consumption in response to different weather conditions. *Energy and buildings* 19(3), pp. 235–247.

Hajdukiewicz, M., Geron, M. and Keane, M.M. 2013. Calibrated CFD simulation to evaluate thermal comfort in a highly-glazed naturally ventilated room. *Building and Environment* 70, pp. 73–89.

Haldi, F. and Robinson, D. 2011. The impact of occupants' behaviour on building energy demand. *Journal of Building Performance Simulation* (March 2013), pp. 37–41.

Hassid, S. and Santamouris, M. 2000. The effect of the Athens heat island on air conditioning load. *Energy and Buildings* (June 1997).

Heiselberg, P., Brohus, H., Hesselholt, A., Rasmussen, H., Seinre, E. and Thomas, S. 2009. Application of sensitivity analysis in design of sustainable buildings. *Renewable Energy* 34(9), pp. 2030–2036.

Henninger, R.H., Witte, M.J. and Crawley, D.B. 2004. Analytical and comparative testing of EnergyPlus using IEA HVAC BESTEST E100-E200 test suite. *Energy and Buildings* 36(8), pp. 855–863.

Heo, Y., Choudhary, R. and Augenbroe, G. 2012. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings* 47, pp. 550–560.

Heo, Y., Choudhary, R. and Augenbroe, G.A. 2012. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings* 47, pp. 550–560.

Hitchcock, R.J., Carroll, W.L. and Birdsall, B.E. 1991. RESEM: Retrofit Energy Savings Estimation Model Reference Manual. Applied Science Division, Lawrence Berkeley Laboratory (LBL), University of California, Berkeley, CA 94720.

Hoes, P., Hensen, J.L.M., Loomans, M.G.L.C., de Vries, B. and Bourgeois, D. 2009. User behavior in whole building simulation. *Energy and Buildings* 41(3), pp. 295–302.

Hopfe, C.J. 2009. Uncertainty and sensitivity analysis in building performance simulation for decision support and design optimization. Technische Universiteit Eindhoven.

Hopkinson, R., Petherbridge, P. and Longmore, J. 1966. Daylighting. Heinemann.

Hua, Y., Oswald, A. and Yang, X. 2011. Effectiveness of daylighting design and occupant visual satisfaction in a LEED Gold laboratory building. *Building and Environment* 46(1), pp. 54–64.

Hunn, B., Banks, J. and Reddy, S. 1992. Energy analysis of the Texas capitol restoration. In: *Proceedings of the 8th Symposium on Improving Building Systems in Hot and Humid Climates*. Dallas, TX, pp. 165–173.

Hygh, J.S., DeCarolis, J.F., Hill, D.B. and Ranji Ranjithan, S. 2012. Multivariate regression as an energy assessment tool in early building design. *Building and Environment* 57, pp. 165–175.

Hyun, S., Park, C. and Augenbroe, G. 2008. Analysis of uncertainty in natural ventilation predictions of high-rise apartment buildings. *Building Services Engineering Research and Technology* 29(4), pp. 311–326.

IEA 2007. World Energy Outlook 2007.

IEA 2011. World Energy Outlook 2011.

Intergovernmental Panel on Climate Change 2007. Climate Change 2007 Synthesis Report.

iSBEM 2006. User Guide to iSBEM (iSBEM version 1.2.a).

ISO 2007. ISO 15251:2007 - Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics. International Standards Organisation.

ISO 2005. ISO 7730:2005 Ergonomics of the thermal environment — Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria. International Standards Organisation.

Jones, D. 1989. Urbanization and energy use in economic development. *The Energy Journal* 10(4), pp. 29–44.

Kalogirou, S.A. and Bojic, M. 2000. Artificial neural networks for the prediction of the energy consumption of a passive solar building. *Energy* 25(5), pp. 479–491.

Kaplan, M., Jones, B. and Jansen, J. 1990. DOE-2.1 C model calibration with monitored end-use data. *Proceedings from the ACEEE 1990 Summer Study on Energy Efficiency in Buildings* 10, pp. 115–125.

Kaplan, M., McFerran, J., Jansen, J. and Pratt, R. 1990. Reconciliation of a DOE2. 1C model with monitored end-use data for a small office building. *ASHRAE Transactions* 96(1), pp. 982–993.

Karatasou, S., Santamouris, M. and Geros, V. 2006. Modeling and predicting building's energy use with artificial neural networks: Methods and results. *Energy and Buildings* 38(8), pp. 949–958.

Karlsson, F., Rohdin, P. and Persson, M.L. 2007. Measured and predicted energy demand of a low energy building: important aspects when using Building Energy Simulation. *Building Service Engineering Research and Technology* 28(3), pp. 223–235.

Katipamula, S. and Claridge, D. 1993. Use of simplified system models to measure retrofit energy savings. *Journal of Solar Energy Engineering* 115(2), pp. 57–68.

Kennedy, M.C. and O'Hagan, A. 2001. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63(3), pp. 425–464.

Klein, S.A., Beckman, W.A., Mitchell, J.W., Duffie, J.A., Duffie, N.A., Freman, T.L., Mitchell, J.C., Braun, J.E., Evans, B.L., Kummer, J.P. and others 1979. TRNSYS, A transient system simulation program. Solar Energy Laboratory, University of Wisconsin-Madison.

Knebel, D. 1983. Simiplified energy analysis using the modified bin method. American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), Atlanta, GA, USA.

Kreider, J. and Haberl, J. 1994. Predicting hourly building energy usage. *ASHRAE Journal* 36(6), pp. 72–81.

Kreider, J.F. and Haberl, J.S. 1994. Predicting hourly building energy use: the great energy predictor shootout: overview and discussion of results. *ASHRAE Transactions* 100(2), pp. 1104–1118.

Labstats LLC 2013. Labstats [Online]. Available at: http://www.labstats.com/.

Lam, J.C. and Hui, S.C.M. 1996. Sensitivity analysis of energy performance of office buildings. *Building and Environment* 31(1), pp. 27–39.

Larsen, S. and Filippín, C. 2008. An experience on integrating monitoring and simulation tools in the design of energy-saving buildings. *Energy & Buildings* 40(6), pp. 987–997.

Lavigne, K. 2009. Assisted calibration in buildings simulation–Algorithm description and case studies. In: *Proceedings of the 11th International IBPSA Conference*. Glasgow, Scotland.

Lebot, B. 1987. Investigation of an Expert System to Calibrate A Computer Simulation Program.

Levine, M., Urge-Vorsatz, D., Blok, K., Geng, L., Harvey, D., Land, S., Levermore, G., Mongameli Mehlwana, A., Mirasgedis, S., Novikova, A., Rlling, J. and Yoshino, H. 2007. Residential and commercial buildings, Climate Change 2007: Mitigation, Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. New York, NY, U.S.A.

Liu, G. and Liu, M. 2011. A rapid calibration procedure and case study for simplified simulation models of commonly used HVAC systems. *Building and Environment* 46(2), pp. 409–420.

Liu, M. and Claridge, D. 1995. Application of calibrated HVAC system models to identify component malfunctions and to optimize the operation and control schedules. *Solar Engineering*, pp. 209–209.

Liu, M. and Claridge, D.E. 1998. Use of calibrated HVAC system models to optimize system operation. *Journal of solar energy engineering* 120(2), pp. 131–138.

Liu, M., Claridge, D.E., Bensouda, N., Heinemeier, K., Lee, S.U. and Wei, G. 2003. High Performance Commercial Building Systems Manual of Procedures for Calibrating Simulations of Building Systems. California Energy Commission (CEC), California, USA.

Liu, M., Song, L., Wei, G. and Claridge, D.E. 2004. Simplified Building and Air-handling Unit Model Calibration and Applications. *Journal of Solar Energy Engineering* 126(1), p. 601.

Liu, M., Wei, G. and Claridge, D. 1998. Calibrating AHU Models Using Whole Building Cooling and Heating Energy Consumption Data. In: *Proceedings of the ACEEE 1998 Summer Study on Energy Efficiency in Buildings*. pp. 229–241.

Liu, S. and Henze, G.P. 2005. Calibration of building models for supervisory control of commercial buildings. In: *Proceedings of the 9th international building performance simulation association (IBPSA) conference*. Montréal, Canada.

Ljung, L. 1987. System identification: Theory for the user. Prentice-Hall, Englewood Cliffs, New Jersey 07632.

Lomas, K.J. and Eppel, H. 1992. Sensitivity analysis techniques for building thermal simulation programs. *Energy and Buildings* 19(2), pp. 1–44.

Lomas, K.J., Eppel, H., Martin, C.J. and Bloomfield, D.P. 1997. Empirical validation of building energy simulation programs. *Energy and Buildings* 26(3), pp. 253–275.

Lunneberg, T. 1999. Improving simulation accuracy through the use of short-term electrical end-use monitoring. In: *Proceedings of the 6th International IBPSA Conference*. Kyoto, Japan, pp. 1–8.

Lyberg, M. 1987. Source Book for Energy Auditors: Vol's 1 & 2. Swedish Council for Building Research, Stockholm, Sweden.

Macdonald, I., Clarke, J.A. and Strachan, P.A. 1999. Assessing uncertainty in building simulation. In: *Proceedings of the 6th International IBPSA Conference*. International Building Performance Simulation Association, Kyoto, Japan.

Macdonald, I. and Strachan, P. 2001. Practical application of uncertainty analysis. *Energy & Buildings* 33(3), pp. 219–227.

Macdonald, I.A. 2002. Quantifying the Effects of Uncertainty in Building Simulation. Ph.D. Thesis. University of Strathclyde.

MacKay, D. 1994. Bayesian nonlinear modeling for the prediction competition. *ASHRAE Transactions* 100(2), pp. 1053–1062.

Maile, T. 2010. Comparing Measured and Simulated Building Energy Performance Data. Stanford University.

Maile, T. and Fischer, M. 2010. Formalizing approximations, assumptions, and simplifications to document limitations in building energy performance simulation. Comparing Measured and Simulated Building Energy Performance Data.

Manfren, M., Aste, N. and Moshksar, R. 2013. Calibration and uncertainty analysis for computer models – A meta-model based approach for integrated building energy simulation. *Applied Energy* 103, pp. 627–641.

Manke, J., Hittle, D. and Hancock, C. 1996. Calibrating building energy analysis models using short term test data. In: *Proceedings of the 1996 ASME International Solar Engineering Conference*. San Antonio, TX, pp. 369–378.

Mccray, J.A., Bailey, P.L., Parker, J.L. and Gillman, R. 1995. Using data visualization tools for the calibration of hourly DOE-2 . 1 simulations Calibration of simulation models. In: *Proceedings of the 3rd International IBPSA Conference*. Madison, WI, pp. 461–466.

McIntyre, D. 1980. Indoor climate. Applied Science Publishers.

McIntyre, D. 1978. Three approaches to thermal comfort. *ASHRAE Transactions* 84(1), pp. 101–109.

Meade, R. 2013. China, Coal and Climate Change: A Narrative for Inaction [Online].Insititute of International and European Affairs (IIEA) Available at: http://www.iiea.com/blogosphere/china-coal-and-climate-change-a-narrative-for-inaction.

Mechri, H.E., Capozzoli, A. and Corrado, V. 2010. USE of the ANOVA approach for sensitive building energy design. *Applied Energy* 87(10), pp. 3073–3083.

Metz, B., Davidson, O.R., Bosch, P.R., Dave, R. and Meyer, L.A. (eds. 2007. Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Mihalakakou, G., Santamouris, M. and Tsangrassoulis, A. 2002. On the energy consumption in residential buildings. *Energy and Buildings* 34(7), pp. 727–736.

Monfet, D., Charneux, R., Zmeureanu, R. and Lemire, N. 2009. Calibration of a Building Energy Model Using Measured Data. *ASHRAE Transactions* 115(1), pp. 348–359.

Moon, H. 2005. Assessing mold risks in buildings under uncertainty. Ph.D. Thesis. Georgia Institute of Technology, GA, USA.

Moon, H.J. 2010. Uncertainty Analysis in Mould Spore Transportation and its Application in an Existing Building. *Indoor and Built Environment* 19(3), pp. 355–365.

Muller, E. 2013. China Must Exploit Its Shale Gas. *NY Times*.

Neto, A.H. and Fiorelli, F.A.S. 2008. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and Buildings* 40(12), pp. 2169–2176.

Neymark, J., Judkoff, R., Knabe, G., Le, H.-T., Dürig, M., Glass, a and Zweifel, G. 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.

Nicol, J.F. and Humphreys, M. a. 2002. Adaptive thermal comfort and sustainable thermal standards for buildings. *Energy and Buildings* 34(6), pp. 563–572.

Nielsen, H.A. and Madsen, H. 2006. Modelling the heat consumption in district heating systems using a grey-box approach. *Energy and Buildings* 38(1), pp. 63–71.

Norford, L.K., Socolow, R.H., Hsieh, E.S. and 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'Donnell, J., Maile, T. and Rose, N.C. 2013. Transforming BIM to BEM: Generation of Building Geometry for the NASA Ames Sustainability Base BIM. (January).

O'Hagan, a. 2006. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety* 91(10-11), pp. 1290–1300.

O'Neill, Z., Eisenhower, B., Fonoberov, V. and Bailey, T. 2012. Calibration of a Building Energy Model Considering Parametric Uncertainty. *ASHRAE Transactions* 118(SA-12-C024).

Oreskes, N., Shrader-Frechette, K. and Belitz, K. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*.

Pal, V., Carroll, W. and Bourassa, N. 2002. RESEM-CA: Validation and testing. HPBCS E2P2. High Performance Commercial Building Systems (HPCBS). California Energy Commission (CEC).

Parker, J., Cropper, P. and Shao, L. 2012. A Calibrated Whole Building Simulation Approach to Assessing Retrofit Options for Birmingham Airport. In: *Building Simulation and Optimization Conference*. Loughborough, UK, pp. 49–56.

Pérez-Lombard, L., Ortiz, J. and Pout, C. 2008. A review on buildings energy consumption information. *Energy and Buildings* 40(3), pp. 394–398.

Raftery, P. 2009. Calibrating whole building energy models: An evidence-based methodology and detailed case study. National University of Ireland, Galway.

Raftery, P., Costa, A., Coakley, D. and White, G. 2012. HVACGenerator.

Raftery, P. and Keane, M. 2011. Visualising Patterns in Building Performance Data. In: *Proceedings of the 12th International IBPSA Conference*. Sydney, Australia.

Raftery, P., Keane, M. and Costa, A. 2011. Calibrating whole building energy models: Detailed case study using hourly measured data. *Energy & Buildings* 43(12), pp. 3666–3679.

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

Ratto, M., Tarantola, S. and Saltelli, A. 2001. Sensitivity analysis in model calibration: GSA-GLUE approach. *Computer Physics Communications*.

Reddy, T., Deng, S. and Claridge, D. 1999. Development of an inverse method to estimate overall building and ventilation parameters of large commercial buildings. *Journal of solar energy engineering* 121, pp. 40–46.

Reddy, T.A. 2006. Literature Review on Calibration of Building Energy Simulation Programs: Uses, Problems, Procedures, Uncertainty, and Tools. *ASHRAE Transactions* 112(1), pp. 226–240.

Reddy, T.A. and Andersen, K.K. 2002. An evaluation of classical steady-state off-line linear parameter estimation methods applied to chiller performance data. *HVAC&R Research* 8(1), pp. 101–124.

Reddy, T.A., Maor, I. and Panjapornpon, C. 2007a. Calibrating Detailed Building Energy Simulation Programs with Measured Data-- Part I: General Methodology (RP-1051). *HVAC&R Research* 13(2), pp. 221–241.

Reddy, T.A., Maor, I. and Panjapornpon, C. 2007b. Calibrating Detailed Building Energy Simulation Programs with Measured Data--Part II: Application to Three Case Study Office Buildings (RP-1051). *HVAC&R Research* 13(2), pp. 243–265.

Revelle, R. and Suess, H.E. 1957. Carbon Dioxide Exchange between Atmosphere and Ocean and the Question of an Increase of Atmospheric CO2 During the Past Decades. *Tellus* 9, pp. 18–27.

Robinson, D. and Haldi, F. 2012. Modelling Occupants' Presence and Behaviour–Part II. *Journal of Building Performance Simulation* (March 2013).

Rosen, R. 1991. Life itself: a comprehensive inquiry into the nature, origin, and fabrication of life. Columbia University Press.

Rowe, S.P.M., Dunlop, J., Fowler, T., Greenwell, F., Heaton, L., Hooper, L., Irazi, C., Kingkade, W., Litz, L., Mcdevitt, T., Miles, A., Murray, N., Rowe, P., Sawyer, C., Wannall, W., Yoder, K., Zaslow, D., Johnson, P., Assistant, S., Reil, J., Caban, K., Poe, I., Dickerson, P., Hanson, M., Carroll, G., Nelsa, D., Adams, B., Sweeney, J. and Dove, E. 2004. Global Population Profile : 2002.

Ryan, E.M. and Sanquist, T.F. 2012. Validation of building energy modeling tools under idealized and realistic conditions. *Energy and Buildings* 47, pp. 375–382.

Saltelli, A. 2005. Global Sensitivity Analysis : An Introduction. European Commission, Joint Research Centre, Ispra, Italy.

Saltelli, A. 2002. Sensitivity analysis for importance assessment. *Risk Analysis* 22(3), pp. 579–590.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S. 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications* 181(2), pp. 259–270.

Saltelli, A., Ratto, M., Andres, T. and Campolongo, F. 2008. Global Sensitivity Analysis: The Primer. John Wiley & Sons, Ltd.

Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M. 2004. Sensitivity analysis in practice: a guide to assessing scientific models. Wiley, Hoboken NJ.

Scofield, J.H. 2009. Do LEED-certified buildings save energy? Not really…. *Energy and Buildings* 41(12), pp. 1386–1390.

Shapiro, I. 2009. Energy Audits in Large Commercial Office Buildings. *ASHRAE Journal Journal* 51, pp. 18–27.

Shen, H. and Tzempelikos, A. 2013. Sensitivity analysis on daylighting and energy performance of perimeter offices with automated shading. *Building and Environment* 59, pp. 303–314.

Soebarto, V. 1997. Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings. In: *Proceedings of the 4th International IBPSA Conference*. Prague, Czech Republic.

Souza, C.B. De and Alsaadani, S. 2012. Thermal Zoning in Speculative Office Buildings: Discussing the Connections between Space Layout and Inside Temperature Control. In: *First Building Simulation and Optimization Conference*. Loughborough, UK, pp. 417–424.

Spitz, C., Mora, L., Wurtz, E. and Jay, A. 2012. Practical application of uncertainty analysis and sensitivity analysis on an experimental house. *Energy and Buildings* 55, pp. 459–470.

Standards Australia 2000. AS/NZ 3598:2000 Australian/New Zealand Standard: Energy Audits. Australia/New Zealand.

Sterling, R., Coakley, D. and Keane., M.M. 2014. Improving whole building energy simulation with artificial neural networks and real performance data: A natural ventilation building case study (DRAFT). In: *Building Simulation and Optimization*. London, United Kingdom.

Stern, N. 2006. The economics of climate change: the Stern Review. World Economics. Cambridge University Press, Cambridge, UK.

Strachan, P. 2000. ESP-r: Summary of validation studies. Energy Systems Research Unit, University of Strathclyde. Glasgow, Scotland.

Stroeve, J. and Serreze, M. 2008. Arctic sea ice extent plummets in 2007. *EOS, Transactions, American Geophysica Union* 89(2), pp. 2007–2008.

Stroeve, J.C., Serreze, M.C., Holland, M.M., Kay, J.E., Malanik, J. and Barrett, A.P. 2011. The Arctic's rapidly shrinking sea ice cover: a research synthesis. *Climatic Change* 110(3-4), pp. 1005–1027.

Struck, C. and Kotek, P. 2009. On the Application of Uncertainty and Sensitivity Analysis with Abstract Building Performance Simulation Tools. *Journal of Building Physics* 33(1), pp. 5–27.

Subbarao, K. 1988. Primary and secondary terms analysis and renormalization: A unified approach to building energy simulations and short-term testing. Golden, CO.

Sun, J. and 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.

Tanimoto, J., Hagishima, A. and Sagara, H. 2008. Validation of methodology for utility demand prediction considering actual variations in inhabitant behaviour schedules. *Journal of Building Performance Simulation* 1(1), p. 31.

Thumann, A. and Younger, W. 2008. Handbook of energy audits. 7th Ed. Fairmont Press, Lilburn, GA.

Tian, W. 2013. A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews* 20, pp. 411–419.

Tian, W. and Choudhary, R. 2012. A probabilistic energy model for non-domestic building sectors applied to analysis of school buildings in greater London. *Energy and Buildings* 54, pp. 1–11.

Tian, W. and de Wilde, P. 2011. Uncertainty and sensitivity analysis of building performance using probabilistic climate projections: A UK case study. *Automation in Construction* 20(8), pp. 1096–1109.

TortoiseSVN Core Development Team 2011. TortoiseSVN.

TRC 1984. DOE-2: Comparison with measured data: design and operational energy studies in a new high-rise office building. New York, USA.

Trčka, M. and Hensen, J. 2010. Overview of HVAC system simulation. *Automation in Construction* 19(2), pp. 93–99.

Troncoso, R. 1997. A hybrid monitoring-modeling procedure for analyzing the performance of large central chilling plants. In: *Proceedings of the 6th International IBPSA Conference*. International Building Performance Simulation Association, Prague, Czech Republic, pp. 421–428.

Tupper, K., Franconi, E., Chan, C., Fluhrer, C., Jenkins, M. and Hodgin, S. 2011. Pre-Read for BEM Innovation Summit. In: *Building Energy Modeling (BEM) Innovation Summit.* Boulder, CO.

Turiel, I., Boschen, R., Seedall, M. and Levine, M. 1984. Simplified energy analysis methodology for commercial buildings. *Energy and Buildings* 6(1), pp. 67–83.

Turner, C. and Frankel, M. 2008. Energy Performance of LEED for New Construction Buildings. New Buildings Institute, Washington, DC.

UNEP 2009. Buildings and Climate Change: Summary for Decision-Makers.

UNFCCC Secretariat 2009. Copenhage Accord (COP-15).

UNFCCC Secretariat 1997. Kyoto Protocol to the United Nations Framework Convention on Climate Change.

United Nations 2013. World Population Prospects: The 2012 Revision. New York, USA.

United Nations 2011. World Urbanization Prospects: The 2011 Revision. New York, USA.

US Department of Energy 2011a. Auxiliary EnergyPlus Programs Extra programs for EnergyPlus. California, USA.

US Department of Energy 2011b. Getting Started with EnergyPlus Basic Concepts Manual - Essential Information You Need about Running EnergyPlus. California, USA.

US Department of Energy (DOE) 2013. Building Energy Software Tools Directory: DOE Sponsored Tools [Online]. Available at: http://apps1.eere.energy.gov/buildings/tools_directory/doe_sponsored.cfm [Accessed: 1 July 2013].

US DOE 2008. M&V Guidelines: Measurement and Verification for Federal Energy Projects Version 3.0. US Department Of Energy.

Waltz, J. 1992. Practical experience in achieving high levels of accuracy in energy simulations of existing buildings. In: *ASHRAE Transactions*. pp. 606–617.

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

Wei, G., Liu, M. and Claridge, D. 1998. Signatures of heating and cooling energy consumption for typical AHUs. In: *Eleventh Symposium on Improving Building Systems in Hot and Humid Climates*. Forth Worth, TX, pp. 387–402.

Westphal, F. and Lamberts, R. 2005. Building simulation calibration using sensitivity analysis. In: *Proceedings of the 9th International IBPSA Conference*. Montréal, Canada, pp. 1331–1338.

Wetter, M. 2001. GenOpt®," Generic Optimization Program. In: *Proceedings of the 7th International IBPSA Conference*. pp. 601–608.

WHO 1983. Indoor air pollutants: exposure and health effects. World Health Organisation, Copenhagen, Denmark.

WHO 1989. Indoor Air Quality: Organic Pollutants. World Health Organisation, Copenhagen, Denmark.

De Wilde, P. and Tian, W. 2010. Predicting the performance of an office under climate change: A study of metrics, sensitivity and zonal resolution. *Energy and Buildings* 42(10), pp. 1674–1684.

Wilde, P. and Tian, W. 2009. Identification of key factors for uncertainty in the prediction of the thermal performance of an office building under climate change. *Building Simulation* 2(3), pp. 157–174.

Wilkinson, P., Smith, K. and Beevers, S. 2007. Energy, energy efficiency, and the built environment. *The Lancet* 370(9593), pp. 1175–1187.

Winkelmann, F., Birdsall, B. and Buhl, W. 1993. DOE-2 Supplement: Version 2.1 e. *LBL-34947. Berkeley, Calif.: Lawrence Berkeley Laboratory*.

De Wit, S. and Augenbroe, G. 2002a. Analysis of uncertainty in building design evaluations and its implications. *Energy and Buildings* 34(9), pp. 951–958.

De Wit, S. and Augenbroe, G. 2002b. Analysis of uncertainty in building design evaluations and its implications. *Energy and Buildings* 34(9), pp. 951–958.

Wright, J.A., Wang, M., Brownlee, A.E.I. and Buswell, R.A. 2012. Variable convergence in evolutionary optimization and its relationship to sensitivity analysis. In: *Building Simulation and Optimization*. Loughborough University© IBPSA-England, London, United Kingdom.

Yao, R., Li, B. and Steemers, K. 2005. Energy policy and standard for built environment in China. *Renewable Energy* 30, pp. 1973–1988.

Yi Zhang, Ivan Korolija and Graeme Stuart 2012. jEPlus.

Yildiz, Y., Korkmaz, K., Göksal Özbalta, T. and Durmus Arsan, Z. 2012. An approach for developing sensitive design parameter guidelines to reduce the energy requirements of low-rise apartment buildings. *Applied Energy* 93, pp. 337–347.

Yoon, J. and Lee, E. 1999. Calibration procedure of energy performance simulation model for a commercial building. In: *Proceedings of Building Simulation 1999*. International Building Performance Simulation Association, Kyoto, Japan.

Yoon, J., Lee, E.J. and Claridge, D.E. 2003. Calibration Procedure for Energy Performance Simulation of a Commercial Building. *Journal of Solar Energy Engineering* 125(3), p. 251.

Young, P., Parkinson, S. and Lees, M. 1996. Simplicity out of complexity in environmental modelling: Occam's razor revisited. *Journal of Applied Statistics* 23(2-3), pp. 165–210.

Yu, F.W. and Chan, K.T. 2005. Energy signatures for assessing the energy performance of chillers. *Energy and Buildings* 37(7), pp. 739–746.

Yun, G.Y., Tuohy, P. and Steemers, K. 2009. Thermal performance of a naturally ventilated building using a combined algorithm of probabilistic occupant behaviour and deterministic heat and mass balance models. *Energy and Buildings* 41(5), pp. 489–499.

Zhang, Y. and Korolija, I. 2010. Performing complex parametric simulations with jEPlus. *SET2010-9th International Conference on Sustainable …*.

Zhao, Z. Bin, Xu, W.S. and Cheng, D.Z. 2011. Uncertainty and Sensitivity Analysis of Building Energy Demand Model Based on Occupant Behaviors. *Applied Mechanics and Materials* 71-78, pp. 411–415.

Appendix A: Programs and Scripts

A.1 Weather Data

A.1.1 Download Weather Data (Batch FTP Script)

```
open weather.nuigalway.ie
nuigweather
readonly
!:--- FTP commands below here ---
binary
get "CR1000_Minute_Table.dat"
get "CR1000_Hourly_Table.dat"
pause
disconnect
pause
bye
```

A.1.2 Import and Aggregate in MySQL Database (SQL)

```
#---------------------------------------------------------------------#
# DESCRIPTION
/* This script was created to automate the process of collecting weather
data from the NUIG weather station. Data is collected in two (.dat)
files, containing both hourly and minute interval data. This data is
transferred to temporary tables and merged to a single weather data
table. Missing dates, where present, are inserted with NULL values */
# AUTHOR - Daniel Coakley (daniel.coakley@nuigalway.ie)
# DATE - January 2013
#---------------------------------------------------------------------#

#Insert all Data from Archive/Server to temporary Tables in SQL Database

#HOURLY DATA
DROP TABLE IF EXISTS temp.Weather1;
CREATE TABLE temp.Weather1 (
    Date DateTime Primary Key,
    Rainfall Double (5 , 3 )
);
LOAD DATA LOCAL INFILE 'C:\\Users\\Daniel\\Google Drive\\PhD\\Weather
Data\\Raw Data\\CR1000_Hourly_Table.dat'
INTO TABLE temp.Weather1
FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 4 LINES
(Date, @dummy, Rainfall, @dummy);

#MINUTE DATA
DROP TABLE IF EXISTS temp.Weather2;
CREATE TABLE temp.Weather2 (
    Date DateTime Primary Key,
    AirTemp Double (5 , 3 ),
    RH Double,
    WindSpeed Double (10 , 5 ),
    WindDir Double (10 , 5 ),
    WindGust Double (10 , 5 ),
    BP INT(11),
    SlrTot Double (10 , 5 ),
```

```
    SlrDiff Double (10 , 5 )
);
LOAD DATA LOCAL INFILE 'C:\\Users\\Daniel\\Google Drive\\PhD\\Weather
Data\\Raw Data\\CR1000_Minute_Table.dat'
INTO TABLE temp.Weather2
FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 4 LINES
(Date,@dummy,@dummy,@dummy,AirTemp,RH,@dummy,@dummy,WindSpeed,WindDir,Win
dGust,BP,SlrTot,SlrDiff);

#MERGING TEMPORARY TABLES
DROP TABLE IF EXISTS temp.WeatherTot;
CREATE TABLE temp.WeatherTot (
    Date DateTime Primary Key,
    AirTemp Double,
    RH Double,
    WindSpeed Double (10 , 5 ),
    WindDir Double (10 , 5 ),
    WindGust Double (10 , 5 ),
    BP INT(11),
    SlrTot Double (10 , 3 ),
    SlrDiff Double (10 , 3 ),
    Rainfall Double (4 , 2 )
);
INSERT INTO temp.WeatherTot
SELECT
DATE,avg(AirTemp),avg(RH),avg(WindSpeed),avg(WindDir),avg(WindGust),avg(B
P),avg(SlrTot),avg(SlrDiff),avg(Rainfall)
FROM temp.Weather2
INNER JOIN temp.Weather1
USING(Date)
GROUP BY hour(Date),day(Date),month(Date),year(Date)
ORDER BY Date;

#INSERT MISSING DATES
DROP TABLE IF EXISTS temp.WeatherTotDate;
CREATE TABLE temp.WeatherTotDate (
    Date DateTime Primary Key,
    AirTemp Double,
    RH Double,
    WindSpeed Double (10 , 5 ),
    WindDir Double (10 , 5 ),
    WindGust Double (10 , 5 ),
    BP INT(11),
    SlrTot Double (10 , 3 ),
    SlrDiff Double (10 , 3 ),
    Rainfall Double (4 , 2 )
);
INSERT INTO temp.WeatherTotDate
(Date,AirTemp,RH,WindSpeed,WindDir,WindGust,BP,SlrTot,SlrDiff,Rainfall)
SELECT *
FROM data.datelist
LEFT JOIN temp.WeatherTot
USING(Date)
GROUP BY hour(Date),day(Date),month(Date),year(Date)
ORDER BY Date;

DROP TABLE temp.Weather1;
DROP TABLE temp.Weather2;
CREATE VIEW temp.weatherdata2012 AS
```

```
    SELECT
        DATE,
        avg(AirTemp),
        avg(RH),
        avg(WindSpeed),
        avg(WindDir),
        avg(BP),
        avg(SlrTot),
        avg(SlrDiff),
        avg(Rainfall)
    FROM
        temp.WeatherTotDate
    WHERE
        DATE >= '2012-01-01 00:00'
            AND DATE < '2013-01-01 00:00'
    GROUP BY hour(Date) , day(Date) , month(Date) , year(Date)
    ORDER BY Date;
```

A.1.3 Export Weather Data for Conversion (SQL)

```
SELECT
      Date_Format(DATE, '%Y-%m-%d'),Date_Format(Date,'%H:%i'), AirTemp,
      RH, WindSpeed, WindDir, BP, SlrTot, SlrDiff, Rainfall
FROM temp.WeatherTotDate
INTO OUTFILE 'C:\\Users\\Daniel\\Google Drive\\PhD\\Weather Data\\EPW
Conversion\\EPW_Weather_Data1.csv'
FIELDS TERMINATED BY ','
ENCLOSED BY '';
```

A.2 Loading Archived BMS Data (MS SQL)

Pre-Requisites:

- BuildingManagment Program
- Set up IIS Server on local machine (Refer: http://www.iis.net/)
- Set up Microsoft SQL Server 2008 R2 (http://www.microsoft.com/download/en/details.aspx?id=23650)

Server Details

Currently set up as 'localhost' server on local machine only

Mixed mode SQL server Authentication

- Username: *sa*
- **Password:** ....

Restoring the BMS SQL Database

1. Launch SQL Server Management Studio
2. Right-click 'Databases' in Object Explorer Window
3. Click 'Restore Database…'
   - **To Database:** SQL Database Name (buildingmanagement2)
   - **From Device: File:** buildingmanagement2.bak

Loading Data to BMS SQL Database

1. Download current ARCHIVE folder from BMS server
2. Copy ARCHIVE Folder to Building Management/ARCHIVE
3. Create 'processed' folder at Building Management/ARCHIVE/processed
   - If you are running the program for the second or subsequent time, ensure the processed folder is empty as this may lead to DataHandlingException errors during loading
4. Open BuildingManagement.exe
5. Input Required Fields

- **Database Directory:** ARCHIVE folder location (e.g. C:\Building Management\ARCHIVE\)
- **Database Server:** SQL server name (e.g. localhost)
- **Database Name:** SQL database name (BuildingManagement2)

6. Click 'Load Files'

   WARNING: This procedure may take a few hours to complete due to the number of cells to be checked and duplicate data contained in each csv file. Please allow up to 6 hours to transfer one year's data.

A.3 Extracting BMS Data from MS SQL Database

A.3.1 Export MS SQL Express Server Data (SQL)

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT TOP 10000000000 [SensorId]
      ,[TimeOfEvent]
      ,[Value]
  FROM [buildingmanagement2].[dbo].[SensorData]
  WHERE TimeOfEvent >= '2011-05-01 00:00'
```

A.3.2 Import MS SQL Data to MySQL (SQL)

```
DROP TABLE IF EXISTS buildingmanagement.sensordata;
CREATE TABLE buildingmanagement.sensordata (
    SensorID Varchar(8) Not Null,
    TimeofEvent DateTime Not Null,
    Value Float Null,
    idkey INT(11) AUTO_INCREMENT PRIMARY KEY
);
LOAD DATA LOCAL INFILE 'C:\\Users\\Daniel\\Documents\\PhD\\BES Model\\BMS
Data\\SQL Import Data\\SensorData.csv'
INTO TABLE buildingmanagement.sensordata
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\r\n'
IGNORE 1 LINES;
```

A.4 Exporting Parametric Job Files (VB-script)

```
 Sub ParameterList()
' This subroutine writes a list of the variable parameters to the
ParameterList.csv file.
    ' This file will only contain those parameters which have not been
specified as 'fixed' on the worksheet

    ' Refresh this cell so that it automatically recalculates the current
filepath & name of the sheet
    Range("FullName").FormulaR1C1 = "=MyFullName()"

    ' open writepath for the files
    writepath = Range("FilePath").Value & "Input Files\Parametric\" &
Range("FileName1").Value
    Open writepath For Output As #1

    'Print Headings
    Print #1, "#" & "This is a list of variable model parameter fo input into
jEPlus"
    Print #1, "#" & "ID" & "," & "Name" & "," & "ParameterType" & "," &
"Description" & "," & "SearchString" & "," & "ValueType" & "," & "ValueString"
& "," & "ValueIndex"


    'Count all global variables
    NoVariables =
Worksheets("Variables").Range("E7:E9999").SpecialCells(xlCellTypeConstants,
xlTextValues).Cells.Count

     'prints all global variables
    Dim currentVarID As String
    Dim currentVarName As String
    Dim currentVarParType As String
    Dim currentVarDescription As String
    Dim currentVarSearchString As String
    Dim currentVarValueType As String
    Dim currentVarValueString As String
    Dim currentVarValueIndex As String
    Dim currentrow As Integer
    Dim c As String
  '  c = Range("L:L").Value

    'For Each c In Range("L:L")
    For currentrow = 1 To NoVariables - 1
        currentVarID = Range("ID").Offset(currentrow, 0).Value
        currentVarName = Range("ID").Offset(currentrow, 1).Value
        currentVarParType = 0
        currentVarDescription = Range("ID").Offset(currentrow, 4).Value
        currentVarSearchString = Range("String").Offset(currentrow, 0).Value
        currentVarValueType = Range("VariableType").Offset(currentrow,
0).Value
        currentVarValueString = Range("ID").Offset(currentrow, 5).Value
        currentVarValueIndex = 0
      '  c = Range("L:L").Offset(currentrow, 0).Value

        If Range("ID").Offset(currentrow, 9).Value = "N" And
Range("ID").Offset(currentrow, 10).Value = "Y" Then
```

```
            Print #1, "P" & currentVarID & "," & currentVarName & "," &
currentVarParType & "," & currentVarDescription & "," & currentVarSearchString
& "," & currentVarValueType & "," & "{" & currentVarValueString & "}" & "," &
currentVarValueIndex
        End If
    Next
    Close #1

    ' open writepath for the Parameter Values
    writepath = Range("FilePath").Value & "Input Files\Parametric\" &
Range("FileName2").Value
    Open writepath For Output As #2

    'Print Headings
    Print #2, "ID" & "," & "Name" & "," & "ParameterType" & "," &
"Description" & "," & "SearchString" & "," & "ValueType" & "," & "ValueString"
& "," & "ValueStdDev" & "," & "ValueIndex"


    'Count all global variables
    NoVariables =
Worksheets("Variables").Range("E7:E9999").SpecialCells(xlCellTypeConstants,
xlTextValues).Cells.Count

     'prints all global variables

    Dim currentVarValueStdDev As String

    'For Each c In Range("L:L")
    For currentrow = 1 To NoVariables - 1
        currentVarID = Range("ID").Offset(currentrow, 0).Value
        currentVarName = Range("ID").Offset(currentrow, 1).Value
        currentVarParType = 0
        currentVarDescription = Range("ID").Offset(currentrow, 4).Value
        currentVarSearchString = Range("String").Offset(currentrow, 0).Value
        currentVarValueType = Range("VariableType").Offset(currentrow,
0).Value
        currentVarValueString = Range("ID").Offset(currentrow, 5).Value
        currentVarValueStdDev = Range("StdDev").Offset(currentrow, 0).Value
        currentVarValueIndex = 0
      '  c = Range("L:L").Offset(currentrow, 0).Value

        If Range("ID").Offset(currentrow, 9).Value = "N" And
Range("ID").Offset(currentrow, 10).Value = "Y" Then
            Print #2, "P" & currentVarID & "," & currentVarName & "," &
currentVarParType & "," & currentVarDescription & "," & currentVarSearchString
& "," & currentVarValueType & "," & currentVarValueString & "," &
currentVarValueStdDev & "," & currentVarValueIndex
        End If
    Next
    Close #2
    Dim ret As String
    ret = "E:\Dropbox\PhD\Model\Analysis\Model\Parametric\Input
Files\Parametric\CreateParametricBatch.bat"
    Shell ret
    End

End Sub
```

A.5 Generating sample matrix (R-script)

```
#=============================================================================
#
# CREATE A JOB IMPORT FILE FOR JEPLUS
#
#=============================================================================
#
# Read in the Parameter Data file
#
#setwd("C:/Users/Daniel/Google Drive/PhD/BES Model/Parametric Study/Input
Files/Parametric")
params <- read.csv(file="ParameterValues.csv",header=TRUE,sep=",")

# for each row i in the Data file, we calculate a random parameter value
# based on the specified mean and sd values.

# Step 1: Specify Number of Simulations, s
# This section creates a Simulation Index matrix

s=10 # Number of simulations
Job=1:s # Job Reference
JobIndex=0 # Index Value for jEPlus, leave at 0
WthrIndex=0 # Weather Index Value for jEPlus, leave at 0
m<-matrix(Job)
m<-cbind(m,JobIndex) #cbind is used to add additional columns to a matrix
m<-cbind(m,WthrIndex)

# c combines data into a vector matrix form, e.g, data<-c(1:10)
# t may be used to transpose a vector, if required, e.g. t(data)
# A for loop is initiated to add columns containing parameter values
# (rnorm is used to generate a random normal distribution)

for (i in 1:NROW(params)){
u<-c(rnorm(s,params$ValueString[i],params$ValueStdDev[i]))
m<-cbind(m,u)}

# The job file is output in csv format to the working directory.

write.table(m,"JobFile.csv", sep =",", row.names = FALSE, col.names =
FALSE)

#=============================================================================
#
# ISSUES
# 1.  Script generates Out of Range values which cause EnergyPlus to
#crash occasionally. Needs to be addessed in Parametric Study sheet
#(Define Ranges for all Params)
#=============================================================================
====================

#=============================================================================
#
# CHANGELOG
#
#=============================================================================
```

A.6 Aggregating Results (Java)

In order to post-process the simulation results, it is necessary to compile the generated outputs and export them to the MySQL database. The current version of jEPlus (v.1.3) does not natively support this process. As this process requires the generation of results for up to 500,000 simulations, I ran into difficulty when outputting a summary of simulation results. jEPlus uses a 'for loop' to read simulation results from the results directory, adds these to a large Data array and finally outputs a summary table to the Parent directory. However, due to the sheer size of the array, there is a problem with allocation of Virtual memory within the Java application. Therefore, it was necessary to write an external program in java to complete this step.

The requirements for this package are as follows:

- Loop through **eplusout.csv** files in simulation sub-folders in Results folder
- Append 'Job ID' to first column of eplusout.csv file
- Add entire file to SimResults.csv summary file

This process loops through each Simulation result file, collecting the EnergyPlus outputs and exporting them to a single csv file or to a MySQL database. The following Java program was developed for this purpose:

```
package com.resultscollector;

import java.io.BufferedReader;
import java.io.FileReader;
import java.io.IOException;
import java.util.ArrayList;
import java.util.List;

public class FileArray {

    public String[] readLines(String filename) throws IOException {
        FileReader fileReader = new FileReader(filename);
        BufferedReader bufferedReader = new BufferedReader(fileReader);
        List<String> lines = new ArrayList<String>();
        String line = null;
        while ((line = bufferedReader.readLine()) != null) {
            lines.add(line);
        }
        bufferedReader.close();
        return lines.toArray(new String[lines.size()]);
    }
}


package com.resultscollector;

import java.io.*;
import com.resultscollector.FileArray;
```

```
public class ListFiles {

    public static void main(String[] args) {
        // eg java ListFiles c:\
            try {
                listFiles(new File("Output Files/jEPlus"));
            } catch (Exception e) {
                e.printStackTrace();
            }
    }

    //TODO Need to change the code below to make it work using relative paths
    // i.e. place jar in directory and execute.

    //TODO Add other functions to program -
    //Call jePlus and Start parsing files once jEPlus simulation process is
complete
    //Call MySQLImport to export results file to Database (see JDBC)

    public static void listFiles(File dir) throws Exception {
      boolean append = true;
      File f1 = new File("Output Files/jEPlus/SimResults.csv");
      f1.delete();
        File[] files = dir.listFiles();
        FileArray SimResults = new FileArray();
        FileWriter list = new FileWriter("Output Files/jEPlus/SimResults.csv",
append);
        for (int i = 0; i < files.length; i++) {

            String fileName = files[i].getName();

            // put in your filter here
            if (fileName.endsWith("eplusout.csv")) {

                if (files[i].isFile()) {
                  String Dir = files[i].getParentFile().getName();
                  String Path = files[i].getPath();
                  String[] lines = SimResults.readLines(Path);

                  for (String line : lines) {
                        list.append(Dir + "," + line + "\n");
                  }

                }
            }
            if (files[i].isDirectory()) {
                listFiles(files[i]);
            }
        }
       list.close();
    }

}

//TODO Create .jar file to be executed as part of Simulation process (require
relative path)
```

A.7 Pre-Processing: Temperature Data (SQL)

```
#----------------------------------------------------------------------#
# DESCRIPTION - Aveage Zone Temperature Data (degC)
# AUTHOR - Daniel Coakley (daniel.coakley@nuigalway.ie)
# DATE - January 2013
#----------------------------------------------------------------------#

DROP TABLE IF EXISTS temp_bms.RawRoomtemperature;
CREATE TABLE temp_bms.RawRoomtemperature (
    Date DateTime Primary Key NULL,
    D0010201 Double (5 , 3 ),
    D0010202 Double (5 , 3 ),
    D0010203 Double (5 , 3 ),
    D0010204 Double (5 , 3 ),
    D0010206 Double (5 , 3 ),
    D0010207 Double (5 , 3 ),
    D0010301 Double (5 , 3 ),
    D0010302 Double (5 , 3 ),
    D0010303 Double (5 , 3 ),
    D0010304 Double (5 , 3 ),
    D0010401 Double (5 , 3 )
);
INSERT INTO temp_bms.RawRoomtemperature (Date,
        D0010201, D0010202, D0010203, D0010204, D0010206, D0010207,
        D0010301, D0010302, D0010303, D0010304, D0010401)
SELECT TimeofEvent
     , MAX(CASE WHEN SensorID = 'D0010201' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010202' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010203' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010204' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010206' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010207' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010301' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010302' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010303' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010304' THEN value END)
     , MAX(CASE WHEN SensorID = 'D0010401' THEN value END)
  FROM buildingmanagement.sensordata
GROUP BY TimeofEvent
ORDER BY TimeofEvent;

DROP TABLE IF EXISTS temp_bms.AverageRoomtemperature;
CREATE TABLE temp_bms.AverageRoomtemperature (
    Date DateTime Primary Key NULL,
    D0010201 Double (5 , 3 ),
    D0010202 Double (5 , 3 ),
    D0010203 Double (5 , 3 ),
    D0010204 Double (5 , 3 ),
    D0010206 Double (5 , 3 ),
    D0010207 Double (5 , 3 ),
    D0010301 Double (5 , 3 ),
    D0010302 Double (5 , 3 ),
    D0010303 Double (5 , 3 ),
    D0010304 Double (5 , 3 ),
    D0010401 Double (5 , 3 )
);
INSERT INTO temp_bms.AverageRoomtemperature (Date,
```

```
        D0010201, D0010202, D0010203, D0010204, D0010206,
        D0010207, D0010301, D0010302, D0010303, D0010304,
        D0010401)
SELECT Date,
        avg(D0010201), avg(D0010202), avg(D0010203), avg(D0010204),
avg(D0010206),
        avg(D0010207), avg(D0010301), avg(D0010302), avg(D0010303),
avg(D0010304),
        avg(D0010401)
FROM temp_bms.RawRoomtemperature
GROUP BY hour(Date),day(Date),month(Date),year(Date)

ORDER BY Date;
DROP TABLE IF EXISTS buildingmanagement.zonetempaverage;
CREATE TABLE buildingmanagement.zonetempaverage (
    Date DateTime Primary Key NULL,
    ZoneTemp Double (5 , 3 ),
    D0010201 Double (5 , 3 ),
    D0010202 Double (5 , 3 ),
    D0010203 Double (5 , 3 ),
    D0010204 Double (5 , 3 ),
    D0010206 Double (5 , 3 ),
    D0010207 Double (5 , 3 ),
    D0010301 Double (5 , 3 ),
    D0010302 Double (5 , 3 ),
    D0010303 Double (5 , 3 ),
    D0010304 Double (5 , 3 ),
    D0010401 Double (5 , 3 )
);
INSERT INTO buildingmanagement.zonetempaverage (Date, ZoneTemp, D0010201,
D0010202, D0010203, D0010204, D0010206, D0010207, D0010301, D0010302,
D0010303, D0010304, D0010401)

SELECT Date,
((D0010201*0.1284)+(D0010202*0.1788)+(D0010203*0.0124)+(D0010204*0.1761)+
(D0010206*0.1057)+(D0010207*0.0167)+(D0010301*0.0412)+(D0010302*0.0361)+
(D0010303*0.0246)+(D0010304*0.0364)+(D0010401*0.2435)), D0010201,
D0010202, D0010203, D0010204, D0010206, D0010207, D0010301, D0010302,
D0010303, D0010304, D0010401

FROM temp_bms.AverageRoomtemperature
GROUP BY hour(Date),day(Date),month(Date),year(Date)
ORDER BY Date;

DROP TABLE IF EXISTS temp_bms.AverageRoomtemperature;
CREATE TABLE temp_bms.AverageRoomtemperature (
    Date DateTime Primary Key NULL,
    D0010201 Double (5 , 3 ),
    D0010202 Double (5 , 3 ),
    D0010203 Double (5 , 3 ),
    D0010204 Double (5 , 3 ),
    D0010206 Double (5 , 3 ),
    D0010207 Double (5 , 3 ),
    D0010301 Double (5 , 3 ),
    D0010302 Double (5 , 3 ),
    D0010303 Double (5 , 3 ),
    D0010304 Double (5 , 3 ),
    D0010401 Double (5 , 3 )
);
INSERT INTO temp_bms.AverageRoomtemperature (Date,
        D0010201, D0010202, D0010203, D0010204, D0010206,
```

```
        D0010207, D0010301, D0010302, D0010303, D0010304,
        D0010401)
SELECT Date,
        avg(D0010201), avg(D0010202), avg(D0010203), avg(D0010204),
avg(D0010206),
        avg(D0010207), avg(D0010301), avg(D0010302), avg(D0010303),
avg(D0010304),
        avg(D0010401)
FROM temp_bms.RawRoomtemperature
GROUP BY hour(Date),day(Date),month(Date),year(Date)

ORDER BY Date;
DROP TABLE IF EXISTS buildingmanagement.zonetempaverage;
CREATE TABLE buildingmanagement.zonetempaverage (
    Date DateTime Primary Key NULL,
    ZoneTemp Double (5 , 3 ),
    D0010201 Double (5 , 3 ),
    D0010202 Double (5 , 3 ),
    D0010203 Double (5 , 3 ),
    D0010204 Double (5 , 3 ),
    D0010206 Double (5 , 3 ),
    D0010207 Double (5 , 3 ),
    D0010301 Double (5 , 3 ),
    D0010302 Double (5 , 3 ),
    D0010303 Double (5 , 3 ),
    D0010304 Double (5 , 3 ),
    D0010401 Double (5 , 3 )
);
INSERT INTO buildingmanagement.zonetempaverage (Date, ZoneTemp, D0010201,
D0010202, D0010203, D0010204, D0010206, D0010207, D0010301, D0010302,
D0010303, D0010304, D0010401)

SELECT Date,
((D0010201*0.1284)+(D0010202*0.1788)+(D0010203*0.0124)+(D0010204*0.1761)+
(D0010206*0.1057)+(D0010207*0.0167)+(D0010301*0.0412)+(D0010302*0.0361)+
(D0010303*0.0246)+(D0010304*0.0364)+(D0010401*0.2435)), D0010201,
D0010202, D0010203, D0010204, D0010206, D0010207, D0010301, D0010302,
D0010303, D0010304, D0010401

FROM temp_bms.AverageRoomtemperature
GROUP BY hour(Date),day(Date),month(Date),year(Date)
ORDER BY Date;
```

A.8 Pre-Processing: Electrical & Heat Energy Consumption Data (SQL)

```
#----------------------------------------------------------------------#
# DESCRIPTION - Aveage Hourly Electrical Energy Consumption (kWh)
# AUTHOR - Daniel Coakley (daniel.coakley@nuigalway.ie)
# DATE - January 2013
#----------------------------------------------------------------------#

DROP TABLE IF EXISTS buildingmanagement.AverageElectrical;

CREATE TABLE buildingmanagement.AverageElectrical (
    Date DateTime Primary Key NULL,
    D0010225 Double (5 , 3 )
);

INSERT INTO buildingmanagement.AverageElectrical (Date, D0010225)
SELECT TimeofEvent, avg(Value)
FROM buildingmanagement.sensordata
WHERE SensorID='D0010225'
GROUP BY
hour(TimeofEvent),day(TimeofEvent),month(TimeofEvent),year(TimeofEvent)
ORDER BY TimeofEvent;

#----------------------------------------------------------------------#
# DESCRIPTION - Aveage Daily Heat Energy Consumption (kWh)
# AUTHOR - Daniel Coakley (daniel.coakley@nuigalway.ie)
# DATE - January 2013
#----------------------------------------------------------------------#

DROP TABLE IF EXISTS buildingmanagement.DailyHeat;

CREATE TABLE buildingmanagement.DailyHeat (
    Date DateTime Primary Key NULL,
    D0010117 Double (5 , 3 )
);

INSERT INTO buildingmanagement.DailyHeat (Date, D0010117)
SELECT TimeofEvent, sum(Value)
FROM buildingmanagement.sensordata
WHERE SensorID='D0010117'
GROUP BY day(TimeofEvent),month(TimeofEvent),year(TimeofEvent)
ORDER BY TimeofEvent;
```

A.9 Post-Processing: Simulation Data (R)

```
#-----------------------------------------------------------------------#
# DESCRIPTION -  Simulation Results Post-Processing Script
# AUTHOR - Daniel Coakley (daniel.coakley@nuigalway.ie)
# DATE - January 2013
#-----------------------------------------------------------------------#


Simulation Results (Sample)
========================================================
The following page presents a summary of current model simulation
results, illustrating simulation error by type, where:
* GOFT = Goodness-of-Fit (Total)
  * GOFA = Goodness-of-Fit (CVRMSE)
  * GOFB = Goodness-of-Fit (NMBE)
* Heat = Heating Energy Consumption (kWh)
* Elec = Electrical Energy Consumption (kWh)
* Temp = Average Zone Temperature (degC)

Further detailed visualisations are available for each simulation run,
showing absolute error for various facets (hour, weekday, month).

## Initial set-up

Set working directory and load libraries.
```{r init.parameters, echo=TRUE,
warning=FALSE,error=FALSE,message=FALSE}
setwd("E:/Dropbox/PhD/Model/Analysis")
library(ggplot2)
library(lattice)
library(latticeExtra)
library(reshape)
require(grid)
require(gridExtra)
```

Read in data: weather data, measured building data and simulation data.
```{r Reading Data - Measured,
echo=TRUE,cache=TRUE,error=FALSE,message=FALSE,warning=FALSE}
weather <- read.csv("Model\\Calibration Data\\WeatherData2011.csv")
hourly  <- read.csv("Model\\Calibration Data\\HourlyData2011.csv")
hourly  <- na.omit(hourly)
heat    <- read.csv("Model\\Calibration Data\\DailyHeat2011.csv")
```

Read in data: weather data, measured building data and simulation data.
```{r Reading Data - Simulated,
echo=TRUE,cache=FALSE,error=FALSE,message=FALSE,warning=FALSE}
#sim     <- read.csv("Model\\Current Model\\BaseModel.csv")
sim     <- read.csv("Model\\Current
Model\\SimulationFileAndResults\\SimulationFile.csv")

```

Read in results of previous runs, and determine whether this is a first
run.
```

````
```{r Reading Last Run,
echo=TRUE,error=FALSE,warning=FALSE,message=FALSE,results='asis',eval=TRU
E}
models  <- read.csv("Results.csv")
try(i<- max(models$Revision))
if (i==-Inf || i=="") {
  n=1
} else {
  n=i+1
}
```

Set weightings for various statistical indices and measurements in order
to determine goodness-of-fit (GOF)
```{r GOF, echo=TRUE,cache=TRUE,error=FALSE,message=FALSE,warning=FALSE}
WHeat   <-  0.3   #Heating energy consumption (kWh)
WElec   <-  0.6   #Electrical energy consumption (kWh)
WTemp   <-  0.1   #Average Temperature (degC)
WCV     <-  0.4   #CVRMSE weighting
WNMBE   <-  0.6   #NMBE weighting
```

Multiplot function (for use with ggplot)
```{r Multiplot Function,eval=TRUE,cache=TRUE,echo=TRUE}
# Multiple plot function
#
# ggplot objects can be passed in ..., or to plotlist (as a list of
ggplot objects)
# - cols:   Number of columns in layout
# - layout: A matrix specifying the layout. If present, 'cols' is
ignored.
#
# If the layout is something like matrix(c(1,2,3,3), nrow=2, byrow=TRUE),
# then plot 1 will go in the upper left, 2 will go in the upper right,
and
# 3 will go all the way across the bottom.
#
multiplot <- function(..., plotlist=NULL, file, cols=1, layout=NULL) {
  require(grid)

  # Make a list from the ... arguments and plotlist
  plots <- c(list(...), plotlist)

  numPlots = length(plots)

  # If layout is NULL, then use 'cols' to determine layout
  if (is.null(layout)) {
    # Make the panel
    # ncol: Number of columns of plots
    # nrow: Number of rows needed, calculated from # of cols
    layout <- matrix(seq(1, cols * ceiling(numPlots/cols)),
                    ncol = cols, nrow = ceiling(numPlots/cols))
  }

 if (numPlots==1) {
    print(plots[[1]])

  } else {
    # Set up the page
    grid.newpage()
````
````

```
    pushViewport(viewport(layout = grid.layout(nrow(layout),
ncol(layout))))

    # Make each plot, in the correct location
    for (i in 1:numPlots) {
      # Get the i,j matrix positions of the regions that contain this
subplot
      matchidx <- as.data.frame(which(layout == i, arr.ind = TRUE))

      print(plots[[i]], vp = viewport(layout.pos.row = matchidx$row,
                                      layout.pos.col = matchidx$col))
    }
  }
}
```


Pre-Processing

Tidy up simulation file: Change column headers, fix timestamp format
```{r cleaning, echo=TRUE, warning=FALSE,error=FALSE,message=FALSE}
names(sim)[1] <- "Date"
heat$Date<-strftime(heat$Date,"%d/%m/%Y %H:%M")
sim$Date<-strftime(strptime(sim$Date," %m/%d  %H:%M:%Y"),"%d/%m/2011
%H:%M")
```

Next, merge the data in one single data frame
```{r Merging, echo=TRUE, warning=FALSE,error=FALSE,message=FALSE}
HourlyTotal <- merge(hourly,sim,by=c("Date"),all.x=TRUE)
DailyTotal  <- merge(heat,HourlyTotal,by=c("Date"),all.x=TRUE)
DailyTotal  <- DailyTotal[order(as.Date(DailyTotal$Date,
format="%d/%m/%Y")),]
```

Assign variable names
```{r Assign Variable Names, echo=TRUE,
warning=FALSE,error=FALSE,message=FALSE}
try(names(HourlyTotal)[names(HourlyTotal) ==
"HW_DISTRICT_HEATING.District.Heating.Rate..W..Hourly."] <- "Sim.Heat")
try(names(DailyTotal)[names(DailyTotal) ==
"HW_DISTRICT_HEATING.District.Heating.Rate..W..Hourly."] <- "Sim.Heat")
try(names(DailyTotal)[names(DailyTotal) ==
"HW_PLANT_LOOP.Plant.Loop.Heating.Demand..W..Daily."] <- "Sim.Heat")
try(names(HourlyTotal)[names(HourlyTotal) ==
"CHW_DISTRICT_COOLING.District.Cooling.Rate..W..Hourly."] <- "Sim.Cool")
try(names(HourlyTotal)[names(HourlyTotal) ==
"Whole.Building.Total.HVAC.Electric.Demand..W..Hourly."] <- "Sim.HVAC")
try(names(HourlyTotal)[names(HourlyTotal) ==
"Whole.Building.Total.Electric.Demand..W..Hourly."] <- "Sim.Elec")
#try(names(HourlyTotal)[names(HourlyTotal) ==
"Whole.Building.Total.Building.Electric.Demand..W..Hourly."] <-
"Sim.Elec")
try(names(HourlyTotal)[names(HourlyTotal) ==
"NLIB_ZONE_01.Zone.Mean.Air.Temperature..C..Hourly."] <- "Sim.Temp")

try(names(DailyTotal)[names(DailyTotal) == "D0010117"] <- "Meas.Heat")
try(names(HourlyTotal)[names(HourlyTotal) == ""] <- "Meas.Cool")
try(names(HourlyTotal)[names(HourlyTotal) == ""] <- "Meas.HVAC")
try(names(HourlyTotal)[names(HourlyTotal) == "ZoneElectric"] <-
"Meas.Elec")
```

```
try(names(HourlyTotal)[names(HourlyTotal) == "ZoneTemp"] <- "Meas.Temp")

try(names(HourlyTotal)[names(HourlyTotal) ==
"Environment.Outdoor.Dry.Bulb..C..Hourly."] <- "Meas.DryBlb")
try(names(HourlyTotal)[names(HourlyTotal) ==
"Environment.Outdoor.Wet.Bulb..C..Hourly."] <- "Meas.WetBlb")
try(names(HourlyTotal)[names(HourlyTotal) ==
"Environment.Direct.Solar..W.m2..Hourly"] <- "Meas.DirSol")
try(names(HourlyTotal)[names(HourlyTotal) ==
"Environment.DayType.Index....Hourly."] <- "Meas.DayTyp")

try(HourlyTotal$Sim.Elec<-HourlyTotal$Sim.Elec/1000) # Convert W to kWhr
(as measured)
try(DailyTotal$Sim.Heat<-DailyTotal$Sim.Heat*(24/1000)) # Convert W to
kWhr (as measured)
```
```

Define factors and levels
```{r Define factors, echo=TRUE, warning=FALSE,error=FALSE,message=FALSE}
daylevels=c("Sun","Mon","Tue","Wed","Thu","Fri","Sat")
daytypelevels=c("Weekend","Weekday","Weekday","Weekday","Weekday","Weekda
y","Weekend")
monthlevels=c("May","Jun","Jul","Aug","Sep","Oct","Nov","Dec")
#monthlevels=c("Jan","Feb","Mar","Apr","May","Jun","Jul","Aug","Sep","Oct
","Nov","Dec")
monthtypelevels=c("Academic","Academic","Academic","Academic","Academic",
"Summer","Summer","Summer","Academic","Academic","Academic","Academic")

HourlyTotal$weekdaynum<-as.POSIXlt(HourlyTotal$Date)$wday+1
HourlyTotal$hour<-as.numeric(format(strptime(HourlyTotal$Date,
format='%d/%m/%Y %H:%M'), format="%H"))
HourlyTotal$monthnum<-format(strptime(HourlyTotal$Date, format='%d/%m/%Y
%H:%M'), format="%b")

DailyTotal$weekdaynum<-weekdays(as.Date(DailyTotal$Date,'%d/%m/%Y
%H:%M'))
DailyTotal$hour<-as.numeric(format(strptime(DailyTotal$Date,
format='%d/%m/%Y %H:%M'), format="%H"))
DailyTotal$monthnum<-format(strptime(DailyTotal$Date, format='%d/%m/%Y
%H:%M'), format="%b")

HourlyTotal <- HourlyTotal[order(as.Date(HourlyTotal$Date,
format="%d/%m/%Y %H:%M")),]
DailyTotal <- DailyTotal[order(as.Date(DailyTotal$Date, format="%d/%m/%Y
%H:%M")),]
#DailyTotal <- DailyTotal[complete.cases(DailyTotal),]

DailyTotal$month <- as.factor(DailyTotal$monthnum)
HourlyTotal$month <- as.factor(HourlyTotal$monthnum)
DailyTotal$weekday <- as.factor(DailyTotal$weekdaynum)
HourlyTotal$weekday <- as.factor(HourlyTotal$weekdaynum)

HourlyTotal$month <- ordered(HourlyTotal$month, levels = monthlevels)
HourlyTotal$weekday <- ordered(HourlyTotal$weekday, labels = daylevels)
HourlyTotal$daytype<-factor(HourlyTotal$weekdaynum)
levels(HourlyTotal$daytype)=daytypelevels
HourlyTotal$monthtype<-factor(HourlyTotal$month, levels = monthlevels)
levels(HourlyTotal$monthtype)=monthtypelevels

DailyTotal$month <- ordered(DailyTotal$month, levels = monthlevels)
DailyTotal$weekday <- ordered(DailyTotal$weekday, labels = daylevels)
```
```

`````
```

## Post-processing

Define error calculation methods
```{r Errors, echo=TRUE, warning=FALSE,error=FALSE,message=FALSE}

## Relative absolute error (RAE) ------------------

ABSerror <- function(x,y,na.rm=TRUE) {
  ifelse (x == 0 | y == 0,
          ABS <- "NA",
          ABS <- abs((((((x-y)/y)*100)^2)^0.5))
          )
}

BiasError <- function(x,y,na.rm=TRUE) {
  ifelse (x == 0 | y == 0,
          ABS <- "NA",
          ABS <- ((x-y)/y)*100)
}

## Absolute error (MAE) ------------------
MAerror <- function(x,y,na.rm=TRUE) {
  MAE <- abs(x-y)
}
```

Calculate error values
```{r Results, echo=TRUE, warning=FALSE,error=FALSE,message=FALSE}
try(HourlyTotal$ABStemp <-
as.numeric(ABSerror(HourlyTotal$Sim.Temp,HourlyTotal$Meas.Temp)))
try(HourlyTotal$ABSelec <-
as.numeric(ABSerror(HourlyTotal$Sim.Elec,HourlyTotal$Meas.Elec)))
try(HourlyTotal$Errorelec <-
as.numeric(BiasError(HourlyTotal$Sim.Elec,HourlyTotal$Meas.Elec)))
try(DailyTotal$ABSheat <-
as.numeric(ABSerror(DailyTotal$Sim.Heat,DailyTotal$Meas.Heat)))
try(DailyTotal$MAEheat <-
as.numeric(MAerror(DailyTotal$Sim.Heat,DailyTotal$Meas.Heat)))


vars1<-c("month","Sim.Elec","Meas.Elec")
MonthlyTotalElec<-HourlyTotal[vars1]
colnames(MonthlyTotalElec)<-vars1
MonthlyTotalElec<-aggregate(. ~ month, data = MonthlyTotalElec, sum)

vars2<-c("month","Sim.Temp","Meas.Temp")
MonthlyTotalTemp<-HourlyTotal[vars2]
colnames(MonthlyTotalTemp)<-vars2
MonthlyTotalTemp<-aggregate(. ~ month, data = MonthlyTotalTemp, mean)

vars3<-c("month","Sim.Heat","Meas.Heat")
MonthlyTotalHeat<-DailyTotal[vars3]
colnames(MonthlyTotalHeat)<-vars3
MonthlyTotalHeat<-aggregate(. ~ month, data = MonthlyTotalHeat, sum)

MonthlyTotal<-cbind(MonthlyTotalHeat,MonthlyTotalElec,MonthlyTotalTemp)
`````

````
```


## Result Visualisation
```{r Plotting Loop,eval=FALSE,echo=FALSE}
# summarydata <-
as.data.frame(cbind(HourlyTotal$ABStemp,HourlyTotal$weekday,HourlyTotal$m
onth,HourlyTotal$hour,HourlyTotal$DryBlb))
# write.csv(summarydata,"Model\\SummaryData.csv")
# summarydata <- read.csv("Model\\SummaryData.csv")
# DailyTotal$weekday <- as.factor(DailyTotal$weekday)
# summarydata$weekday <- as.factor(HourlyTotal$weekday)
# summarydata$month <- ordered(HourlyTotal$month, levels = monthlevels)
# summarydata$weekday <- ordered(HourlyTotal$weekday, labels = daylevels)
```

### Temperature Data
```{r Plotting -
Hourly,echo=TRUE,warning=FALSE,error=FALSE,message=FALSE}
dftemp<-with(HourlyTotal, tapply(ABStemp, list(hour, weekday, month),
mean,na.rm=T))
dftemp2<-with(HourlyTotal, tapply(ABStemp, list(hour, weekday),
mean,na.rm=T))
dftemp3<-with(HourlyTotal, tapply(ABStemp, list(hour, month),
mean,na.rm=T))

max <- max(dftemp,na.rm=T)
min <- 0
seq <- seq(min,max,max/200)

plot1<-contourplot(ABStemp ~ hour * weekday | month,
                   data = HourlyTotal,
                   cuts = 200,
                   labels=TRUE,
                   contour=FALSE,
                   drop.unused.levels =
lattice.getOption("drop.unused.levels"),
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq,
                   xlab = "",
                   ylab = "Day of Week",

col.regions=colorRampPalette(c("blue","yellow","red")),
                   main = "Absolute % Error (Temperature)",
                   layout=c(2,4),
                   as.table= TRUE
)

plot2<-contourplot(dftemp2,
                   aspect=0.3,
                   cuts = 200,
                   contour=FALSE,
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq,
                   xlab = "Hour of Day",
                   ylab = "Day of Week",
````

```
col.regions=colorRampPalette(c("blue","yellow","red")),
)

plot3<-contourplot(dftemp3,
                   aspect=0.3,
                   cuts = 200,
                   #labels=TRUE,
                   contour=FALSE,
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq,
                   xlab = "Hour of Day",
                   ylab = "Month of Year",
                   col.regions=colorRampPalette(c("blue","yellow","red"))
)

png('Results\\Current\\Temperature.png', width = 3600, height = 5000,
units = "px", res = 400)
print(plot1, position = c(0,.5,1,1), more=TRUE)
print(plot2, position = c(0,.25,1,.5), more=TRUE)
print(plot3, position = c(0,0,1,.25), more=FALSE)
dev.off()
```
```

### Electrical Data
```{r Plotting -
Electrical,echo=TRUE,warning=FALSE,error=FALSE,message=FALSE}
dfelec<-with(HourlyTotal, tapply(ABSelec, list(hour, month), mean,
na.rm=T))
dfelec2<-with(HourlyTotal, tapply(ABSelec, list(hour, weekday), mean,
na.rm=T))

max <- max(dfelec,na.rm=T)
seq <- seq(min,max,max/200)

plot4<-contourplot(ABSelec ~ hour * weekday | month,
                   data = HourlyTotal,
                   cuts = 200,
                   labels=TRUE,
                   contour=FALSE,
                   drop.unused.levels =
lattice.getOption("drop.unused.levels"),
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq,
                   xlab = "",
                   ylab = "Day of Week",

col.regions=colorRampPalette(c("blue","yellow","red")),
                   main = "Absolute % Error (Electrical)",
                   layout=c(2,4),
                   as.table= TRUE)

# plot5<-contourplot(ABSelec ~ hour * weekday,
#                    aspect=0.3,
#                    data = HourlyTotal,
#                    cuts = 200,
#                    #labels=TRUE,
```
```

```
#                     contour=FALSE,
#                     region = TRUE,
#                     pretty=FALSE,
#                     xscale.components = xscale.components.subticks,
#                     at=seq,
#                     xlab = "Hour of Day",
#                     ylab = "Day of Week",
#
col.regions=colorRampPalette(c("blue","yellow","red"))
# )


plot5<-contourplot(dfelec2,
                   aspect=0.3,
                   cuts = 200,
                   #labels=TRUE,
                   contour=FALSE,
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq,
                   xlab = "Hour of Day",
                   ylab = "Day of Week",
                   col.regions=colorRampPalette(c("blue","yellow","red"))
)


# plot6<-contourplot(ABSelec ~ hour * month,
#                     aspect=0.3,
#                     data = HourlyTotal,
#                     cuts = 200,
#                     #labels=TRUE,
#                     contour=FALSE,
#                     region = TRUE,
#                     pretty=FALSE,
#                     xscale.components = xscale.components.subticks,
#                     at=seq,
#                     xlab = "Hour of Day",
#                     ylab = "Month of Year",
#
col.regions=colorRampPalette(c("blue","yellow","red"))
# )

plot6<-contourplot(dfelec,
                   aspect=0.3,
                   cuts = 200,
                   #labels=TRUE,
                   contour=FALSE,
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq,
                   xlab = "Hour of Day",
                   ylab = "Month of Year",
                   col.regions=colorRampPalette(c("blue","yellow","red"))
)

png('Results\\Current\\ABSElectrical.png', width = 3600, height = 5000,
units = "px", res = 400)
print(plot4, position = c(0,.5,1,1), more=TRUE)
print(plot5, position = c(0,.25,1,.5), more=TRUE)
```

```
print(plot6, position = c(0,0,1,.25), more=FALSE)
dev.off()
````


Electrical Data - Error (%)
```{r Plotting -
Electrical2,echo=TRUE,warning=FALSE,error=FALSE,message=FALSE}
dfelec2<-with(HourlyTotal, tapply(Errorelec, list(hour, month), mean,
na.rm=T))
dfelec<-with(HourlyTotal, tapply(Errorelec, list(hour, weekday), mean,
na.rm=T))

max2 <- max(dfelec2,na.rm=T)
min2 <- min(dfelec2,na.rm=T)
seq2 <- seq(min2,max2,max2/200)

plot10<-contourplot(Errorelec ~ hour * weekday | month,
                   data = HourlyTotal,
                   cuts = 200,
                   labels=TRUE,
                   contour=FALSE,
                   drop.unused.levels =
lattice.getOption("drop.unused.levels"),
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq2,
                   xlab = "",
                   ylab = "Day of Week",

col.regions=colorRampPalette(c("blue","yellow","red")),
                   main = "Bias % Error (Electrical)",
                   layout=c(2,4),
                   as.table= TRUE)

plot11<-contourplot(dfelec,
                   aspect=0.3,
                   cuts = 200,
                   #labels=TRUE,
                   contour=FALSE,
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq2,
                   xlab = "Hour of Day",
                   ylab = "Day of Week",
                   col.regions=colorRampPalette(c("blue","yellow","red"))
)


plot12<-contourplot(dfelec2,
                   aspect=0.3,
                   cuts = 200,
                   #labels=TRUE,
                   contour=FALSE,
                   region = TRUE,
                   pretty=FALSE,
                   xscale.components = xscale.components.subticks,
                   at=seq2,
                   xlab = "Hour of Day",
```

```
                    ylab = "Month of Year",
                    col.regions=colorRampPalette(c("blue","yellow","red"))
)


png('Results\\Current\\Electrical.png', width = 3600, height = 5000,
units = "px", res = 400)
print(plot10, position = c(0,.5,1,1), more=TRUE)
print(plot11, position = c(0,.25,1,.5), more=TRUE)
print(plot12, position = c(0,0,1,.25), more=FALSE)
dev.off()
````

Electrical Data - Statistical Difference
```{r Plotting -
Electrical3,echo=TRUE,warning=FALSE,error=FALSE,message=FALSE}
dfelec2<-with(HourlyTotal, tapply(Errorelec, list(hour, month), mean))
max2 <- max(dfelec2,na.rm=T)
min2 <- min(dfelec2,na.rm=T)
seq2 <- seq(min2,max2,max2/200)

plot13<-ggplot(HourlyTotal, aes(x=factor(hour),
y=Errorelec,fill=daytype,xlab="Hour"),na.rm=T) +
              stat_summary(fun.y = "mean", geom =
"bar",position=position_dodge(0.95),na.rm=T) +
              xlab("Hour") + ylab("Error (%)") + ggtitle("Hourly Absolute
% Error (Electrical) by Daytype and Term")

plot14<- plot13 + facet_grid(monthtype ~ .) + ggtitle("")


plot15<-ggplot(HourlyTotal, aes(x=factor(hour),
y=Errorelec,fill=monthtype,xlab="Hour"),na.rm=T) +
              stat_summary(fun.y = "mean", geom =
"bar",position=position_dodge(0.95),na.rm=T) +
              xlab("Hour") + ylab("Error (%)")+ ggtitle("Hourly Absolute
% Error (Electrical) by Term and Month")


plot16<- plot13 + facet_grid(month ~ .) + ggtitle("")



png('Results\\Current\\Electrical4.png', width = 3600, height = 5000,
units = "px", res = 400)
grid.arrange(plot15,plot16,heights=1:2,ncol=1)
dev.off()
png('Results\\Current\\Electrical3.png', width = 3600, height = 5000,
units = "px", res = 400)
grid.arrange(plot13,plot14,heights=1:2,ncol=1)
dev.off()
````


Heating Data
```{r Plotting -
Heating,echo=TRUE,warning=FALSE,error=FALSE,message=FALSE}
dfheat<-with(DailyTotal, tapply(MAEheat, list(weekday, month), mean))
max <- max(dfheat,na.rm=T)
seq <- seq(min,max,max/200)
```

````
plot7<-contourplot(MAEheat ~ weekday * month,
                   data = DailyTotal,
                   cuts = 200,
                   labels=TRUE,
                   contour=FALSE,
                   drop.unused.levels =
lattice.getOption("drop.unused.levels"),
                   region = TRUE,
                   pretty=FALSE,
                   at=seq,
                   xlab = "",
                   ylab = "Day of Week",

col.regions=colorRampPalette(c("blue","yellow","red")),
                   main = "Absolute Error (Heat)",,
                   as.table= TRUE)

plot8 <- bwplot(MAEheat~weekday,data=DailyTotal,as.table= TRUE)
plot9 <- bwplot(MAEheat~month,data=DailyTotal,as.table= TRUE)

png('Results\\Current\\Heat.png', width = 3600, height = 5000, units =
"px", res = 400)
print(plot7, split=c(1,1,1,3), more=TRUE)
print(plot8, split=c(1,2,1,3), more=TRUE)
print(plot9, split=c(1,3,1,3), more=FALSE)
#print(plot10, split=c(1,4,1,4), more=FALSE)
dev.off()
```

Plot to screen
```{r Plotting - Hourly Temperature,
echo=TRUE,warning=FALSE,error=FALSE,message=FALSE,eval=FALSE}
## Temperature Error Plots
print(plot1)
print(plot2, width = 3600, height = 1500, units = "px", res = 400)
print(plot3, width = 3600, height = 1500, units = "px", res = 400)
```

```{r Plotting - Hourly Electrical,
echo=TRUE,warning=FALSE,error=FALSE,message=FALSE,eval=FALSE}
## Electrical Error Plots
print(plot4)
print(plot5)
print(plot6)
```

```{r Plotting - Daily Heat,
echo=TRUE,warning=FALSE,error=FALSE,message=FALSE,eval=FALSE}
## Heating Error Plots
print(plot7)
print(plot8)
print(plot9)
```

```{r
Stats,eval=FALSE,echo=FALSE,warning=FALSE,error=FALSE,message=FALSE}
#qplot(ABStemp,hour,data=HourlyTotal,facets=month~.,binwidth=2)
#qplot(ABStemp,data=HourlyTotal,facets=month~.,binwidth=2)
#qplot(ABStemp,data=HourlyTotal,facets=weekday~.,binwidth=2)
```
````
````

```
## Statistical summary

Calculate the overally Mean Bias Error (MBE) and CV root mean square
error (CVRMSE) for model
```{r Goodness-of-Fit Calcs,
echo=TRUE,warning=FALSE,error=FALSE,message=FALSE}

MBError <-function (Qobs, Qsim)
{
  Qsim <- Qsim[!is.na(Qobs)]
  Qobs <- Qobs[!is.na(Qobs)]
  Qobs <- Qobs[!is.na(Qsim)]
  Qsim <- Qsim[!is.na(Qsim)]
  if (length(Qobs) == 0 || length(Qsim) == 0)
    return(NA)
  MB <- abs(((sum(Qsim) - sum(Qobs))/sum(Qobs)))
  return(MB)
}

CVRMSError <-function (Qobs, Qsim)
{
  Qsim <- Qsim[!is.na(Qobs)]
  Qobs <- Qobs[!is.na(Qobs)]
  Qobs <- Qobs[!is.na(Qsim)]
  Qsim <- Qsim[!is.na(Qsim)]
  if (length(Qobs) == 0 || length(Qsim) == 0)
    return(NA)
  CVRMSE <- (sqrt((sum((Qsim-Qobs)^2)/length(Qobs)))/mean(Qobs))
  return(CVRMSE)
}

try(MBelec<-MBError(HourlyTotal$Meas.Elec,HourlyTotal$Sim.Elec))
try(MBheat<-MBError(DailyTotal$Meas.Heat,DailyTotal$Sim.Heat))
try(MBtemp<-MBError(HourlyTotal$Meas.Temp,HourlyTotal$Sim.Temp))

try(CVelec<-CVRMSError(HourlyTotal$Meas.Elec,HourlyTotal$Sim.Elec))
try(CVheat<-CVRMSError(DailyTotal$Meas.Heat,DailyTotal$Sim.Heat))
try(CVtemp<-CVRMSError(HourlyTotal$Meas.Temp,HourlyTotal$Sim.Temp))

try(Monthly_MBelec<-
MBError(MonthlyTotal$Meas.Elec,MonthlyTotal$Sim.Elec))
try(Monthly_MBheat<-
MBError(MonthlyTotal$Meas.Heat,MonthlyTotal$Sim.Heat))
try(Monthly_MBtemp<-
MBError(MonthlyTotal$Meas.Temp,MonthlyTotal$Sim.Temp))

try(Monthly_CVelec<-
CVRMSError(MonthlyTotal$Meas.Elec,MonthlyTotal$Sim.Elec))
try(Monthly_CVheat<-
CVRMSError(MonthlyTotal$Meas.Heat,MonthlyTotal$Sim.Heat))
try(Monthly_CVtemp<-
CVRMSError(MonthlyTotal$Meas.Temp,MonthlyTotal$Sim.Temp))

try(GOFA<-
((((WHeat^2)*(CVheat^2))+((WElec^2)*(CVelec^2))+((WTemp^2)*(CVtemp^2)))/(
(WHeat^2)+(WElec^2)+(WTemp^2)))^(0.5))
try(GOFB<-
((((WHeat^2)*(MBheat^2))+((WElec^2)*(MBelec^2))+((WTemp^2)*(MBtemp^2)))/(
(WHeat^2)+(WElec^2)+(WTemp^2)))^(0.5))
try(GOFT<-
(((((WCV^2)*(GOFA^2))+((WNMBE^2)*(GOFB^2)))/((WCV^2)+(WNMBE^2)))^(0.5)))
```
```

```
```

Summary of model performance, and change from previous models.
```{r Summary Plots,echo=TRUE,warning=FALSE,error=FALSE,message=FALSE}
Measlevels=c("GOF","CVRMSE","NMBE")
Typelevels=c("GOFT","GOFA","GOFB","Heat","Elec","Temp")

StatsPerc<-data.frame(Measure =
factor(c("GOF","GOF","GOF","CVRMSE","CVRMSE","CVRMSE","NMBE","NMBE","NMBE
"),levels=Measlevels,ordered=TRUE),Type =
factor(c("GOFT","GOFA","GOFB","Heat","Elec","Temp","Heat","Elec","Temp"),
levels=Typelevels,ordered=TRUE),Error =
round(c(GOFT,GOFA,GOFB,CVheat,CVelec,CVtemp,MBheat,MBelec,MBtemp)*100,2))

current =
data.frame(round(cbind(GOFT,GOFA,GOFB,CVheat,CVelec,CVtemp,MBheat,MBelec,
MBtemp)*100,2))
current2 =
data.frame(round(cbind(GOFT,GOFA,GOFB,CVheat,CVelec,CVtemp,MBheat,MBelec,
MBtemp,Monthly_CVheat,Monthly_CVelec,Monthly_CVtemp,Monthly_MBheat,Monthl
y_MBelec,Monthly_MBtemp)*100,2))

if (n>1) {
# Change in Error
mdata   <- melt(models[,2:11], id=c("Revision"))
colnames(mdata)<-c("Revision","Type","Error")
last=models[i,3:11]
change<-((current-last)/last)*100

StatsChange<-data.frame(Measure =
factor(c("GOF","GOF","GOF","CVRMSE","CVRMSE","CVRMSE","NMBE","NMBE","NMBE
"),levels=Measlevels,ordered=TRUE),Type =
factor(c("GOFT","GOFA","GOFB","Heat","Elec","Temp","Heat","Elec","Temp"),
levels=Typelevels,ordered=TRUE),Error =
round(c(change$GOFT,change$GOFA,change$GOFB,change$CVheat,change$CVelec,c
hange$CVtemp,change$MBheat,change$MBelec,change$MBtemp),2))

# Summary Plots

summary1<-ggplot(StatsPerc, aes(x=Measure, y=Error, fill=Type)) +
  geom_bar(stat="identity", position="dodge") +
  labs(title="Simulation Error (%) by Type",x = "Measurement Index", y =
"Error (%)") +
  geom_text(aes(label=Error), vjust=1.5,
colour="white",position=position_dodge(.9), size=3)

summary2<-ggplot(StatsChange, aes(x=Measure, y=Error, fill=Type)) +
  geom_bar(stat="identity", position="dodge") +
  labs(title="Change in Simulation Error (%)",x = "Measurement Index", y
= "Error (%)") +
  geom_text(aes(label=Error), vjust=1.5,
colour="white",position=position_dodge(.9), size=3)

# Adding new data to data frame for Revision Plot

cdata<-data.frame(Revision = n,Type =
factor(c("GOFT","GOFA","GOFB","CVheat","CVelec","CVtemp","MBheat","MBelec
","MBtemp")), Error = StatsPerc$Error)
newmodel<-rbind(mdata,cdata)
```

````
summary3<-ggplot(newmodel, aes(x=factor(Revision),y=Error, colour=Type,
group=Type)) + geom_line() + geom_point() +
 labs(title="Simulation Error (%) by Revision",x = "Revision Number", y =
"Error (%)")

print(StatsPerc)
png('Results\\Current\\Summary.png', width = 3600, height = 5000, units =
"px", res = 400)
multiplot(summary1, summary2,summary3,cols=1)
dev.off()

multiplot(summary1, summary2,summary3,cols=1)
} else {

# Summary Plot for first model

summary1<-ggplot(StatsPerc, aes(x=Measure, y=Error, fill=Type)) +
  geom_bar(stat="identity", position="dodge") +
  labs(title="Simulation Error (%) by Type",x = "Measurement Index", y =
"Error (%)") +
  geom_text(aes(label=Error), vjust=1.5,
colour="white",position=position_dodge(.9), size=3)

png('Results\\Current\\Summary.png', width = 3600, height = 2000, units =
"px", res = 400)
plot(summary1)
dev.off()

print(StatsPerc)
plot(summary1)
}


```

Archiving of results
```{r Archiving,
echo=TRUE,results='hide',warning=FALSE,error=FALSE,message=FALSE,prompt=T
RUE}

write.csv(HourlyTotal, "Results\\Current\\HourlyData.csv")
write.csv(DailyTotal, "Results\\Current\\DailyData.csv")
continue<-readline("Would you like to commit this revision to Archive?
(y/n)")

if(continue=="y"){
  #Append to file
  comment<-readline("Provide some information about this simulation? ")
  summary<-c(n,current2,comment)
  FF <- as.matrix(t(summary))
  write.table(FF, file = "Results.csv", sep = ",",
             col.names = FALSE, append=TRUE)
  print(paste("Complete - results added to summary sheet as Revision",n))

  #Make a copy of the model
  flist <- list.files("Model\\Current Model", full.names = TRUE)
  dir.create(paste("Model\\Simulation Data\\Rev",n))
  file.copy(flist, paste("Model\\Simulation Data\\Rev",n))

  flist <- list.files("Model\\Current Model\\SimulationFileAndResults",
full.names = TRUE)
````

```
  dir.create(paste("Model\\Simulation Data\\Rev
",n,"\\SimulationFileAndResults", sep=""))
  file.copy(flist, paste("Model\\Simulation Data\\Rev
",n,"\\SimulationFileAndResults", sep=""))

  #Make a copy of the results
  flist <- list.files("Results\\Current", full.names = TRUE)
  dir.create(paste("Results\\Rev",n))
  file.copy(flist, paste("Results\\Rev",n))

} else {
  print("Complete - no results have been added to analysis summary")
}
```
```` ``` ````
````

A.10 Post-Processing: Simulation Data (SQL)

A.10.1 Import Simulation Data

```
# Turn off Strict Trans Table Mode to Allow for Non-Default Table values

SET SESSION sql_mode='';

/* First, we import the raw EnergyPlus Output data using the
auto_inrement Integer
as the Primary Key.
Import Values for Date, Hourly Hot Water Plant (kWh), Hourly Electric
Consumption (kWh) and Mean Air Temperature (degC) */

DROP TABLE IF EXISTS temp_bes.eplusout;
CREATE TABLE temp_bes.eplusout (
    idkey INT(11) PRIMARY KEY,
    Job INT(11) Not Null,
    Date Varchar(20) Not Null,
    HWPlant Double Not Null,
    Electric Double,
    MeanAir Double
);

LOAD DATA LOCAL INFILE 'C:\\Users\\Daniel\\Google Drive\\PhD\\BES
Model\\Parametric Study\\jEPlus\\output\\SimResults.csv'
INTO TABLE temp_bes.eplusout
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
IGNORE 1 LINES;
```

A.10.2 Run Hourly Comparisons

```
# Turn off Strict Trans Table Mode to Allow for Non-Default Table values

SET SESSION sql_mode='';

/* First, we import the raw EnergyPlus Output data using the
auto_inrement Integer
as the Primary Key.
Import Values for Date, Hourly Hot Water Plant (kWh), Hourly Electric
Consumption (kWh) and Mean Air Temperature (degC) */

DROP TABLE IF EXISTS temp_bes.eplusout;
CREATE TABLE temp_bes.eplusout (
    idkey INT(11) PRIMARY KEY auto_increment,
    Job INT(11) Not Null,
    Date Varchar(20) Not Null,
    HWPlant Double Not Null,
    Electric Double,
    MeanAir Double
);

LOAD DATA LOCAL INFILE 'E:\\PhD Archive\\Parametric\\Rev 4\\Output
Files\\jEPlus\\SimResults.csv'
INTO TABLE temp_bes.eplusout
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
IGNORE 1 LINES
```

```
(Job,Date,HWPlant,Electric,MeanAir);

/* Step 2 - Fix Time Values. EnergyPlus outputs using 24hr format and
represents
the last hour of the simulated day as 24:00:00 which is an invalid
timestamp as this
is actually 00:00 on the following day. Therefore, we replace 24:00 with
00:00.*/

DROP TABLE IF EXISTS temp_bes.eplusouttime;
CREATE TABLE temp_bes.eplusouttime (
    idkey INT(11) Not Null auto_increment PRIMARY KEY,
    Job INT(11),
    Date Varchar(20) Not Null,
    HWPlant Double,
    Electric Double,
    MeanAir Double
);

INSERT INTO temp_bes.eplusouttime (Job,Date,HWPlant,Electric,MeanAir)
SELECT Job,REPLACE(Date, '24:00:00', '00:00:00'),HWPlant,Electric,MeanAir
FROM temp_bes.eplusout;

/* Step 3 - Change the Date from String to Date Format*/


DROP TABLE IF EXISTS temp_bes.eplusoutdateformat;
CREATE TABLE temp_bes.eplusoutdateformat (
    idkey INT(11) Not Null auto_increment PRIMARY KEY,
    Job INT(11),
    Date Varchar(20) Not Null,
    HWPlant Double,
    Electric Double,
    MeanAir Double
);

INSERT INTO temp_bes.eplusoutdateformat
(Job,Date,HWPlant,Electric,MeanAir)
SELECT Job,STR_TO_DATE(Date,'%m/%d %H:%i:%s'),HWPlant,Electric,MeanAir
FROM temp_bes.eplusouttime;

/* Step 4 - Add the year*/

DROP TABLE IF EXISTS temp_bes.eplusoutyear;
CREATE TABLE temp_bes.eplusoutyear (
    idkey INT(11) Not Null auto_increment PRIMARY KEY,
    Job INT(11),
    Date DateTime Not Null,
    HWPlant Double,
    Electric Double,
    MeanAir Double
);

INSERT INTO temp_bes.eplusoutyear (Job,Date,HWPlant,Electric,MeanAir)
SELECT Job,REPLACE(Date, '0000-', '2011-'),HWPlant,Electric,MeanAir
FROM temp_bes.eplusoutdateformat;

DROP TABLE IF EXISTS bes.eplusoutdateformat;
DROP TABLE IF EXISTS bes.eplusouttime;
```

```
/* Step 5 - Hours 01 to 23 from our Simulated Data now represent Hours 00
to 22
from our Measured Data. Hour 00 represents Hour 23. Therefore we will
Subtract 1 hour from each time period
between 1 and 23 and add 23 to time period 00*/

DROP TABLE IF EXISTS temp_bes.eplusouttime1;
CREATE TABLE temp_bes.eplusouttime1 (
    idkey INT(11) Not Null auto_increment PRIMARY KEY,
    Job INT(11),
    Date Varchar(20) Not Null,
    HWPlant Double,
    Electric Double,
    MeanAir Double
);

INSERT INTO temp_bes.eplusouttime1 (Job,Date,HWPlant,Electric,MeanAir)
SELECT Job,ADDTIME(Date, '23:00:00'), HWPlant, Electric, MeanAir
FROM temp_bes.eplusoutyear
WHERE Hour(Date)='0';

DROP TABLE IF EXISTS temp_bes.eplusouttime2;
CREATE TABLE temp_bes.eplusouttime2 (
    idkey INT(11) Not Null auto_increment PRIMARY KEY,
    Job INT(11),
    Date Varchar(20) Not Null,
    HWPlant Double,
    Electric Double,
    MeanAir Double
);

INSERT INTO temp_bes.eplusouttime2 (Job,Date,HWPlant,Electric,MeanAir)
SELECT Job,SUBTIME(Date, '1:00:00'), HWPlant, Electric, MeanAir
FROM temp_bes.eplusoutyear
WHERE Hour(Date) between '1' and '23';

# Step 6 - Join both tables

DROP TABLE IF EXISTS temp_bes.eplusouttime3;
CREATE TABLE temp_bes.eplusouttime3 (
    idkey INT(11) Not Null auto_increment PRIMARY KEY,
    Job INT(11),
    Date Varchar(20) Not Null,
    HWPlant Double,
    Electric Double,
    MeanAir Double
);

INSERT INTO temp_bes.eplusouttime3 (Job,Date,HWPlant,Electric,MeanAir)
SELECT Job,Date,HWPlant,Electric,MeanAir FROM temp_bes.eplusouttime1
UNION
SELECT Job,Date,HWPlant,Electric,MeanAir FROM temp_bes.eplusouttime2
ORDER BY Job, Year(Date), Month(Date), Day(Date), hour(Date);


# Step 7 - Move to Complete Table

DROP TABLE IF EXISTS temp_bes.eplusout;
CREATE TABLE temp_bes.eplusout (
    idkey INT(11) Not Null auto_increment PRIMARY KEY,
    Job INT(11),
```

```
    Date Varchar(20) Not Null,
    HWPlant Double,
    Electric Double,
    MeanAir Double
);

INSERT INTO temp_bes.eplusout (Job,Date,HWPlant,Electric,MeanAir)
SELECT Job,Date,HWPlant,Electric,MeanAir
FROM temp_bes.eplusouttime3;

# Step 8 - Drop unused temporary transitional tables

DROP TABLE IF EXISTS temp_bes.eplusoutdateformat;
DROP TABLE IF EXISTS temp_bes.eplusouttime;
DROP TABLE IF EXISTS temp_bes.eplusouttime1;
DROP TABLE IF EXISTS temp_bes.eplusouttime2;
DROP TABLE IF EXISTS temp_bes.eplusouttime3;
DROP TABLE IF EXISTS temp_bes.eplusoutyear;



/* In this Query, we will create a table for comparing Hourly Measured
data to Hourly
Simulated Data and Computing a GOF for each hour */


# Combine Hourly Measured Data with Hourly Simulated data

DROP TABLE IF EXISTS comparisons.hourly;
CREATE TABLE comparisons.hourly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    Date DateTime Not Null,
    Sim_ZoneTemp Double (5 , 3 ) Not Null,
    ZoneTemp Double (5 , 3 ) Not Null,
    Sim_ZoneElectric Double (5 , 3 ) Not Null,
    ZoneElectric Double (5 , 3 ) Not Null,
    ABSe_Temp Double (10 , 5 ),
    ABSe_Elec Double (10 , 5 )
);

# Create an Index of Measured/Simulated Data tables to speed up query
performance

DROP INDEX index_1 ON temp_bes.eplusout;
DROP INDEX index_2 ON buildingmanagement.hourlydata;

create index index_1 on temp_bes.eplusout (Date);
create index index_2 on buildingmanagement.hourlydata (Date);

INSERT INTO comparisons.hourly
(Job,Date,Sim_ZoneTemp,ZoneTemp,Sim_ZoneElectric,
ZoneElectric,ABSe_Temp, ABSe_Elec)
SELECT
Job, DATE, MeanAir, ZoneTemp, Electric/1000, ZoneElectric,
ABS(((ZoneTemp-MeanAir)/(MeanAir))*(100)),
ABS(((ZoneElectric-(Electric/1000))/(ZoneElectric))*(100))
FROM temp_bes.eplusout
#WHERE Date > '2011-05-01'
INNER JOIN buildingmanagement.hourlydata
#WHERE Date > '2011-05-01'
```

```
USING(Date)
WHERE Date > '2011-05-01';


DROP TABLE IF EXISTS comparisons.daily;
CREATE TABLE comparisons.daily (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    Date DateTime Not Null,
    Sim_Heat Double (6 , 3 ) Not Null,
    Meas_Heat Double (6 , 3 ) Not Null,
    ABSe_Heat Double (10 , 5 )
);

INSERT INTO comparisons.daily (Job,Date,Sim_Heat,Meas_Heat,ABSe_Heat)
SELECT
Job, DATE,(HWPlant*(24/1000)),D0010117,ABS((((HWPlant*(24/1000))-
D0010117)/(D0010117))*(100))
FROM temp_bes.eplusout
INNER JOIN buildingmanagement.dailyheat
USING(Date)
WHERE Date > '2011-05-01';


# Calculate NMBE and CVRMSE

# Calculate Measured - Simulated

DROP TABLE IF EXISTS comparisons.GOFcalcs;
CREATE TABLE comparisons.GOFcalcs (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    ABSe_Temp Double (10 , 5 ) Not Null,
    NMBE_Temp Double (10 , 5 ) Not Null,
    CVRMSE_Temp Double (10 , 5 ) Not Null,
    ABSe_Elec Double (10 , 5 ) Not Null,
    NMBE_Elec Double (10 , 5 ) Not Null,
    CVRMSE_Elec Double (10 , 5 ) Not Null
);


INSERT INTO comparisons.GOFcalcs (Job,ABSe_Temp, NMBE_Temp, CVRMSE_Temp,
ABSe_Elec, NMBE_Elec, CVRMSE_Elec)

SELECT
Job,
ABSe_Temp,
ABS((SUM(ZoneTemp-Sim_ZoneTemp)/SUM(ZoneTemp))*100),
(POW((SUM(POW((Sim_ZoneTemp-ZoneTemp),2)/5834)),0.5)/AVG(ZoneTemp))*100,
ABSe_Elec,
ABS((SUM(ZoneElectric-Sim_ZoneElectric)/SUM(ZoneElectric))*100),
(POW((SUM(POW((Sim_ZoneElectric-
ZoneElectric),2)/5834)),0.5)/AVG(ZoneElectric) )*100
FROM comparisons.hourly
# WHERE ABSe_Temp is not null and ABSe_Elec is not null
GROUP BY JOB;


DROP TABLE IF EXISTS comparisons.GOFdaily;
CREATE TABLE comparisons.GOFdaily (
    idkey INT(11) auto_increment PRIMARY KEY,
```

```
    Job INT(11) Not Null,
    ABSe_Heat Double (10 , 5 ) Not Null,
    NMBE_Heat Double (10 , 5 ) Not Null,
    CVRMSE_Heat Double (10 , 5 ) Not Null
);


INSERT INTO comparisons.GOFdaily (Job,ABSe_Heat, NMBE_Heat, CVRMSE_Heat)

SELECT
Job,
ABSe_Heat,
ABS((SUM(Sim_Heat-Meas_Heat)/SUM(Meas_Heat))*100),
(POW((SUM(POW((Sim_Heat-Meas_Heat),2)/244)),0.5)/AVG(Meas_Heat))*100
FROM comparisons.daily
# WHERE ABSe_Temp is not null and ABSe_Elec is not null
GROUP BY JOB;

# Total Goodness of Fit - Hourly

DROP TABLE IF EXISTS comparisons.GOF;
CREATE TABLE comparisons.GOF (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
      CVRMSE_Heat Double (10 , 3 ) Not Null,
      CVRMSE_Elec Double (10 , 3 ) Not Null,
      CVRMSE_Temp Double (10 , 3 ) Not Null,
      NMBE_Heat Double (10 , 3 ) Not Null,
      NMBE_Elec Double (10 , 3 ) Not Null,
      NMBE_Temp Double (10 , 3 ) Not Null,
      GOF_Heat Double (10 , 3 ) Not Null,
      GOF_Elec Double (10 , 3 ) Not Null,
      GOF_Temp Double (10 , 3 ) Not Null,
      GOF_A Double (10 , 3 ) Not Null,
      GOF_B Double (10 , 3 ) Not Null
);

INSERT INTO comparisons.GOF
(Job,CVRMSE_Heat, CVRMSE_Elec, CVRMSE_Temp, NMBE_Heat, NMBE_Elec,
NMBE_Temp, GOF_Heat, GOF_Elec,
GOF_Temp, GOF_A, GOF_B)

SELECT
Job, CVRMSE_Heat, CVRMSE_Elec, CVRMSE_Temp, NMBE_Heat, NMBE_Elec,
NMBE_Temp,
POW(((POW(0.4,2)*POW(NMBE_Heat,2))
+(POW(0.6,2)*POW(CVRMSE_Heat,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.4,2)*POW(NMBE_Elec,2))
+(POW(0.6,2)*POW(CVRMSE_Elec,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.4,2)*POW(NMBE_Temp,2))
+(POW(0.6,2)*POW(CVRMSE_Temp,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.3,2)*POW(CVRMSE_Heat,2))+(POW(0.6,2)*POW(CVRMSE_Elec,2))+(POW
(0.1,2)*POW(CVRMSE_Temp,2)))/
(POW(0.3,2)+POW(0.6,2)+POW(0.1,2)),0.5),
POW(((POW(0.3,2)*POW(NMBE_Heat,2))+(POW(0.6,2)*POW(NMBE_Elec,2))+(POW(0.1
,2)*POW(NMBE_Temp,2)))/
(POW(0.3,2)+POW(0.6,2)+POW(0.1,2)),0.5)
FROM comparisons.GOFcalcs
```

```
INNER JOIN comparisons.GOFdaily
USING (Job);

# Total Goodness of Fit - Final

DROP TABLE IF EXISTS comparisons.GOFTotal;
CREATE TABLE comparisons.GOFTotal (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    NMBE_Temp Double (10 , 3 ) Not Null,
    CVRMSE_Temp Double (10 , 3 ) Not Null,
    GOF_Temp Double (10 , 3 ) Not Null,
    NMBE_Elec Double (10 , 3 ) Not Null,
    CVRMSE_Elec Double (10 , 3 ) Not Null,
    GOF_Elec Double (10 , 3 ) Not Null,
    GOF_A Double (10 , 3 ) Not Null,
    GOF_B Double (10 , 3 ) Not Null,
    GOF_Total Double (10 , 3 ) Not Null
);

INSERT INTO comparisons.GOFTotal
(Job,CVRMSE_Heat, CVRMSE_Elec, CVRMSE_Temp, NMBE_Heat, NMBE_Elec,
NMBE_Temp, GOF_Heat, GOF_Elec,
GOF_Temp, GOF_A, GOF_B, GOF_Total)

SELECT
Job, CVRMSE_Heat, CVRMSE_Elec, CVRMSE_Temp, NMBE_Heat, NMBE_Elec,
NMBE_Temp, GOF_Heat, GOF_Elec,
GOF_Temp, GOF_A, GOF_B,
POW(((POW(0.6,2)*POW(GOF_A,2))
+(POW(0.4,2)*POW(GOF_B,2)))
/(POW(0.6,2)+POW(0.4,2))
,0.5)
FROM comparisons.GOF;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Hourly\\GOF_Heat.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.gof
ORDER BY GOF_Heat
LIMIT 10;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Hourly\\GOF_Elec.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.gof
ORDER BY GOF_Heat
LIMIT 10;


SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Hourly\\GOF_Temp.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.gof
ORDER BY GOF_Temp
```

```
LIMIT 10;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Hourly\\GOF_Total.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.gof
ORDER BY GOF_Total
LIMIT 10;
```

A.10.3 Run Monthly Comparisons

```
/* In this Query, we will create a table for comparing Monthly Measured
data to Monthly
Simulated Data and Computing a GOF for each hour */

#Step 1: Create a Table of Monthly Measured Data

DROP TABLE IF EXISTS buildingmanagement.monthlydata;
CREATE TABLE buildingmanagement.monthlydata (
    idkey INT(11) auto_increment PRIMARY KEY,
    Date DateTime Not Null,
    ZoneTemp Double (5 , 3 ) Not Null,
    ZoneElectric Double (5 , 3 ) Not Null
);

INSERT INTO buildingmanagement.monthlydata (Date,ZoneTemp,ZoneElectric)
SELECT
DATE, AVG(ZoneTemp), SUM(D0010225)
FROM buildingmanagement.zonetempaverage
INNER JOIN buildingmanagement.averageelectrical
USING(Date)
WHERE DATE BETWEEN '2011-05-01 00:00:00' AND '2012-01-01 00:00:00'
GROUP BY Month(Date)
ORDER BY year(Date), month(Date), Day(Date), Hour(Date)
limit 10000;

DROP TABLE IF EXISTS buildingmanagement.monthlyheat;
CREATE TABLE buildingmanagement.monthlyheat (
    idkey INT(11) auto_increment PRIMARY KEY,
    Date DateTime Not Null,
    Monthly_Heat Double (5 , 3 ) Not Null
);

INSERT INTO buildingmanagement.monthlyheat (Date,Monthly_Heat)
SELECT
DATE,SUM(D0010117)
FROM buildingmanagement.dailyheat
WHERE DATE BETWEEN '2011-05-01 00:00:00' AND '2012-01-01 00:00:00'
GROUP BY Month(Date)
ORDER BY year(Date), month(Date), Day(Date), Hour(Date)
limit 10000;

DROP TABLE IF EXISTS buildingmanagement.totalmonthlydata;
CREATE TABLE buildingmanagement.totalmonthlydata (
    idkey INT(11) auto_increment PRIMARY KEY,
    Date DateTime Not Null,
    ZoneTemp Double (5 , 3 ) Not Null,
```

```
    ZoneElec Double (5 , 3 ) Not Null,
    ZoneHeat Double (5 , 3 ) Not Null
);
INSERT INTO buildingmanagement.totalmonthlydata
(Date,ZoneHeat,ZoneElec,ZoneTemp)
SELECT
DATE,Monthly_Heat,ZoneElectric,ZoneTemp
FROM buildingmanagement.monthlydata
INNER JOIN buildingmanagement.monthlyheat
USING(Date)
GROUP BY Month(Date)
ORDER BY year(Date), month(Date), Day(Date), Hour(Date)
limit 10000;

#Step 2: Create a Table of Monthly Simulated Data

DROP TABLE IF EXISTS temp_bes.monthly;
CREATE TABLE temp_bes.monthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    Date DateTime Not Null,
    Sim_ZoneHeat Double (5 , 3 ) Not Null,
    Sim_ZoneElec Double (5 , 3 ) Not Null,
    Sim_ZoneTemp Double (5 , 3 ) Not Null
);

INSERT INTO temp_bes.monthly
(Job,Date,Sim_ZoneHeat,Sim_ZoneElec,Sim_ZoneTemp)
SELECT
Job, DATE, SUM(HWPlant/1000),SUM(Electric/1000), AVG(MeanAir)
FROM temp_bes.eplusout
GROUP BY Job, Month(Date);


#Step 3: Combine Monthly Measured Data with Monthly Simulated data

DROP TABLE IF EXISTS comparisons.monthly;
CREATE TABLE comparisons.monthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    Date DateTime Not Null,
    Sim_ZoneHeat Double (5 , 3 ) Not Null,
    ZoneHeat Double (5 , 3 ) Not Null,
    Sim_ZoneElec Double (5 , 3 ) Not Null,
    ZoneElec Double (5 , 3 ) Not Null,
    Sim_ZoneTemp Double (5 , 3 ) Not Null,
    ZoneTemp Double (5 , 3 ) Not Null,
    ABSe_Heat Double (5 , 3 ) Not Null,
    ABSe_Elec Double (5 , 3 ) Not Null,
    ABSe_Temp Double (5 , 3 ) Not Null
);

INSERT INTO comparisons.monthly
(Job,Date,Sim_ZoneHeat,ZoneHeat,Sim_ZoneElec,
ZoneElec,Sim_ZoneTemp,ZoneTemp,ABSe_Heat,ABSe_Elec,ABSe_Temp)
SELECT
Job, DATE, Sim_ZoneHeat,ZoneHeat,Sim_ZoneElec,
ZoneElec,Sim_ZoneTemp,ZoneTemp,
ABS(((ZoneHeat-(Sim_ZoneHeat))/(ZoneHeat))*(100)),
ABS(((ZoneElec-(Sim_ZoneElec))/(ZoneElec))*(100)),
ABS(((ZoneTemp-Sim_ZoneTemp)/(ZoneTemp))*(100))
```

```
FROM temp_bes.monthly
INNER JOIN buildingmanagement.totalmonthlydata
USING(Date);

# Step 4: Calculate NMBE and CVRMSE


DROP TABLE IF EXISTS comparisons.GOFcalcs_monthly;
CREATE TABLE comparisons.GOFcalcs_monthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    ABSe_Heat Double (10 , 5 ) Not Null,
    NMBE_Heat Double (10 , 5 ) Not Null,
    CVRMSE_Heat Double (10 , 5 ) Not Null,
    ABSe_Elec Double (10 , 5 ) Not Null,
    NMBE_Elec Double (10 , 5 ) Not Null,
    CVRMSE_Elec Double (10 , 5 ) Not Null,
    ABSe_Temp Double (10 , 5 ) Not Null,
    NMBE_Temp Double (10 , 5 ) Not Null,
    CVRMSE_Temp Double (10 , 5 ) Not Null
);

INSERT INTO comparisons.GOFcalcs_monthly (Job,ABSe_Heat, NMBE_Heat,
CVRMSE_Heat,ABSe_Elec, NMBE_Elec, CVRMSE_Elec,
ABSe_Temp, NMBE_Temp, CVRMSE_Temp)

SELECT
Job,
ABSe_Heat,
ABS((SUM(Sim_ZoneHeat-ZoneHeat)/SUM(ZoneHeat))*100),
(POW((SUM(POW((ZoneHeat-Sim_ZoneHeat),2)/8)),0.5)/AVG(ZoneHeat))*100,
ABSe_Elec,
ABS((SUM(Sim_ZoneElec-ZoneElec)/SUM(ZoneElec))*100),
(POW((SUM(POW((Sim_ZoneElec-ZoneElec),2)/8)),0.5)/AVG(ZoneElec))*100,
ABSe_Temp,
ABS((SUM(Sim_ZoneTemp-ZoneTemp)/SUM(ZoneTemp))*100),
(POW((SUM(POW((Sim_ZoneTemp-ZoneTemp),2)/8)),0.5)/AVG(ZoneTemp))*100
FROM comparisons.monthly
GROUP BY JOB;


DROP TABLE IF EXISTS comparisons.GOFmonthly;
CREATE TABLE comparisons.GOFmonthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    CVRMSE_Heat Double (10 , 3 ) Not Null,
    CVRMSE_Elec Double (10 , 3 ) Not Null,
    CVRMSE_Temp Double (10 , 3 ) Not Null,
    NMBE_Heat Double (10 , 3 ) Not Null,
    NMBE_Elec Double (10 , 3 ) Not Null,
    NMBE_Temp Double (10 , 3 ) Not Null,
    GOF_Heat Double (10 , 3 ) Not Null,
    GOF_Elec Double (10 , 3 ) Not Null,
    GOF_Temp Double (10 , 3 ) Not Null,
    GOF_A Double (10 , 3 ) Not Null,
    GOF_B Double (10 , 3 ) Not Null
);

INSERT INTO comparisons.GOFmonthly
(Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,
NMBE_Temp,GOF_Heat,GOF_Elec,GOF_Temp,GOF_A,GOF_B)
```

```
SELECT
Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,NMBE_Temp,
POW(((POW(0.4,2)*POW(NMBE_Heat,2))
+(POW(0.6,2)*POW(CVRMSE_Heat,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.4,2)*POW(NMBE_Elec,2))
+(POW(0.6,2)*POW(CVRMSE_Elec,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.4,2)*POW(NMBE_Temp,2))
+(POW(0.6,2)*POW(CVRMSE_Temp,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.3,2)*POW(CVRMSE_Heat,2))+(POW(0.6,2)*POW(CVRMSE_Elec,2))+(POW
(0.1,2)*POW(CVRMSE_Temp,2)))/
(POW(0.3,2)+POW(0.6,2)+POW(0.1,2)),0.5),
POW(((POW(0.3,2)*POW(NMBE_Heat,2))+(POW(0.6,2)*POW(NMBE_Elec,2))+(POW(0.1
,2)*POW(NMBE_Temp,2)))/
(POW(0.3,2)+POW(0.6,2)+POW(0.1,2)),0.5)
FROM comparisons.GOFcalcs_monthly;

# Total Goodness of Fit - Final

DROP TABLE IF EXISTS comparisons.GOFmonthly_Total;
CREATE TABLE comparisons.GOFmonthly_Total (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    CVRMSE_Heat Double (10 , 3 ) Not Null,
    CVRMSE_Elec Double (10 , 3 ) Not Null,
    CVRMSE_Temp Double (10 , 3 ) Not Null,
    NMBE_Heat Double (10 , 3 ) Not Null,
    NMBE_Elec Double (10 , 3 ) Not Null,
    NMBE_Temp Double (10 , 3 ) Not Null,
    GOF_Heat Double (10 , 3 ) Not Null,
    GOF_Elec Double (10 , 3 ) Not Null,
    GOF_Temp Double (10 , 3 ) Not Null,
    GOF_A Double (10 , 3 ) Not Null,
    GOF_B Double (10 , 3 ) Not Null,
    GOF_Total Double (10 , 3 ) Not Null
);

INSERT INTO comparisons.GOFmonthly_Total
(Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,
NMBE_Temp,GOF_Heat,GOF_Elec,GOF_Temp,GOF_A,GOF_B,GOF_Total)

SELECT
Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,NMBE_Temp,GOF
_Heat,GOF_Elec,GOF_Temp,
GOF_A,GOF_B,
POW(((POW(0.6,2)*POW(GOF_A,2))
+(POW(0.4,2)*POW(GOF_B,2)))
/(POW(0.6,2)+POW(0.4,2))
,0.5)
FROM comparisons.GOFmonthly;


SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Heat.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
```

```
ORDER BY GOF_Heat
LIMIT 10;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Elec.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
ORDER BY GOF_Elec
LIMIT 10;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Temp.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
ORDER BY GOF_Temp
LIMIT 10;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Total.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
ORDER BY GOF_Total
LIMIT 10;
```

A.11 Post-Processing: Monthly Simulation Data (SQL)

```
/* In this Query, we will create a table for comparing Monthly Measured
data to Monthly
Simulated Data and Computing a GOF for each hour */

#Step 1: Create a Table of Monthly Measured Data

DROP TABLE IF EXISTS buildingmanagement.monthlydata;
CREATE TABLE buildingmanagement.monthlydata (
    idkey INT(11) auto_increment PRIMARY KEY,
    Date DateTime Not Null,
    ZoneTemp Double (5 , 3 ) Not Null,
    ZoneElectric Double (5 , 3 ) Not Null
);

INSERT INTO buildingmanagement.monthlydata (Date,ZoneTemp,ZoneElectric)
SELECT
DATE, AVG(ZoneTemp), SUM(D0010225)
FROM buildingmanagement.zonetempaverage
INNER JOIN buildingmanagement.averageelectrical
USING(Date)
WHERE DATE BETWEEN '2011-05-01 00:00:00' AND '2012-01-01 00:00:00'
GROUP BY Month(Date)
ORDER BY year(Date), month(Date), Day(Date), Hour(Date)
limit 10000;

DROP TABLE IF EXISTS buildingmanagement.monthlyheat;
CREATE TABLE buildingmanagement.monthlyheat (
    idkey INT(11) auto_increment PRIMARY KEY,
    Date DateTime Not Null,
    Monthly_Heat Double (5 , 3 ) Not Null
);

INSERT INTO buildingmanagement.monthlyheat (Date,Monthly_Heat)
SELECT
DATE,SUM(D0010117)
FROM buildingmanagement.dailyheat
WHERE DATE BETWEEN '2011-05-01 00:00:00' AND '2012-01-01 00:00:00'
GROUP BY Month(Date)
ORDER BY year(Date), month(Date), Day(Date), Hour(Date)
limit 10000;

DROP TABLE IF EXISTS buildingmanagement.totalmonthlydata;
CREATE TABLE buildingmanagement.totalmonthlydata (
    idkey INT(11) auto_increment PRIMARY KEY,
    Date DateTime Not Null,
    ZoneTemp Double (5 , 3 ) Not Null,
    ZoneElec Double (5 , 3 ) Not Null,
    ZoneHeat Double (5 , 3 ) Not Null
);
INSERT INTO buildingmanagement.totalmonthlydata
(Date,ZoneHeat,ZoneElec,ZoneTemp)
SELECT
DATE,Monthly_Heat,ZoneElectric,ZoneTemp
FROM buildingmanagement.monthlydata
INNER JOIN buildingmanagement.monthlyheat
USING(Date)
GROUP BY Month(Date)
```

```
ORDER BY year(Date), month(Date), Day(Date), Hour(Date)
limit 10000;

#Step 2: Create a Table of Monthly Simulated Data

DROP TABLE IF EXISTS temp_bes.monthly;
CREATE TABLE temp_bes.monthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    Date DateTime Not Null,
    Sim_ZoneHeat Double (5 , 3 ) Not Null,
    Sim_ZoneElec Double (5 , 3 ) Not Null,
    Sim_ZoneTemp Double (5 , 3 ) Not Null
);

INSERT INTO temp_bes.monthly
(Job,Date,Sim_ZoneHeat,Sim_ZoneElec,Sim_ZoneTemp)
SELECT
Job, DATE, SUM(HWPlant/1000),SUM(Electric/1000), AVG(MeanAir)
FROM temp_bes.eplusout
GROUP BY Job, Month(Date);


#Step 3: Combine Monthly Measured Data with Monthly Simulated data

DROP TABLE IF EXISTS comparisons.monthly;
CREATE TABLE comparisons.monthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    Date DateTime Not Null,
    Sim_ZoneHeat Double (5 , 3 ) Not Null,
    ZoneHeat Double (5 , 3 ) Not Null,
    Sim_ZoneElec Double (5 , 3 ) Not Null,
    ZoneElec Double (5 , 3 ) Not Null,
    Sim_ZoneTemp Double (5 , 3 ) Not Null,
    ZoneTemp Double (5 , 3 ) Not Null,
    ABSe_Heat Double (5 , 3 ) Not Null,
    ABSe_Elec Double (5 , 3 ) Not Null,
    ABSe_Temp Double (5 , 3 ) Not Null
);

INSERT INTO comparisons.monthly
(Job,Date,Sim_ZoneHeat,ZoneHeat,Sim_ZoneElec,
ZoneElec,Sim_ZoneTemp,ZoneTemp,ABSe_Heat,ABSe_Elec,ABSe_Temp)
SELECT
Job, DATE, Sim_ZoneHeat,ZoneHeat,Sim_ZoneElec,
ZoneElec,Sim_ZoneTemp,ZoneTemp,
ABS(((ZoneHeat-(Sim_ZoneHeat))/(ZoneHeat))*(100)),
ABS(((ZoneElec-(Sim_ZoneElec))/(ZoneElec))*(100)),
ABS(((ZoneTemp-Sim_ZoneTemp)/(ZoneTemp))*(100))
FROM temp_bes.monthly
INNER JOIN buildingmanagement.totalmonthlydata
USING(Date);

# Step 4: Calculate NMBE and CVRMSE


DROP TABLE IF EXISTS comparisons.GOFcalcs_monthly;
CREATE TABLE comparisons.GOFcalcs_monthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
```

```
    ABSe_Heat Double (10 , 5 ) Not Null,
    NMBE_Heat Double (10 , 5 ) Not Null,
    CVRMSE_Heat Double (10 , 5 ) Not Null,
    ABSe_Elec Double (10 , 5 ) Not Null,
    NMBE_Elec Double (10 , 5 ) Not Null,
    CVRMSE_Elec Double (10 , 5 ) Not Null,
    ABSe_Temp Double (10 , 5 ) Not Null,
    NMBE_Temp Double (10 , 5 ) Not Null,
    CVRMSE_Temp Double (10 , 5 ) Not Null
);

INSERT INTO comparisons.GOFcalcs_monthly (Job,ABSe_Heat, NMBE_Heat,
CVRMSE_Heat,ABSe_Elec, NMBE_Elec, CVRMSE_Elec,
ABSe_Temp, NMBE_Temp, CVRMSE_Temp)

SELECT
Job,
ABSe_Heat,
ABS((SUM(Sim_ZoneHeat-ZoneHeat)/SUM(ZoneHeat))*100),
(POW((SUM(POW((ZoneHeat-Sim_ZoneHeat),2)/8)),0.5)/AVG(ZoneHeat))*100,
ABSe_Elec,
ABS((SUM(Sim_ZoneElec-ZoneElec)/SUM(ZoneElec))*100),
(POW((SUM(POW((Sim_ZoneElec-ZoneElec),2)/8)),0.5)/AVG(ZoneElec))*100,
ABSe_Temp,
ABS((SUM(Sim_ZoneTemp-ZoneTemp)/SUM(ZoneTemp))*100),
(POW((SUM(POW((Sim_ZoneTemp-ZoneTemp),2)/8)),0.5)/AVG(ZoneTemp))*100
FROM comparisons.monthly
GROUP BY JOB;


DROP TABLE IF EXISTS comparisons.GOFmonthly;
CREATE TABLE comparisons.GOFmonthly (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    CVRMSE_Heat Double (10 , 3 ) Not Null,
    CVRMSE_Elec Double (10 , 3 ) Not Null,
    CVRMSE_Temp Double (10 , 3 ) Not Null,
    NMBE_Heat Double (10 , 3 ) Not Null,
    NMBE_Elec Double (10 , 3 ) Not Null,
    NMBE_Temp Double (10 , 3 ) Not Null,
    GOF_Heat Double (10 , 3 ) Not Null,
    GOF_Elec Double (10 , 3 ) Not Null,
    GOF_Temp Double (10 , 3 ) Not Null,
    GOF_A Double (10 , 3 ) Not Null,
    GOF_B Double (10 , 3 ) Not Null
);

INSERT INTO comparisons.GOFmonthly
(Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,
NMBE_Temp,GOF_Heat,GOF_Elec,GOF_Temp,GOF_A,GOF_B)

SELECT
Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,NMBE_Temp,
POW(((POW(0.4,2)*POW(NMBE_Heat,2))
+(POW(0.6,2)*POW(CVRMSE_Heat,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.4,2)*POW(NMBE_Elec,2))
+(POW(0.6,2)*POW(CVRMSE_Elec,2)))
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.4,2)*POW(NMBE_Temp,2))
+(POW(0.6,2)*POW(CVRMSE_Temp,2)))
```

```
/(POW(0.4,2)+POW(0.6,2)),0.5),
POW(((POW(0.3,2)*POW(CVRMSE_Heat,2))+(POW(0.6,2)*POW(CVRMSE_Elec,2))+(POW
(0.1,2)*POW(CVRMSE_Temp,2)))/
(POW(0.3,2)+POW(0.6,2)+POW(0.1,2)),0.5),
POW(((POW(0.3,2)*POW(NMBE_Heat,2))+(POW(0.6,2)*POW(NMBE_Elec,2))+(POW(0.1
,2)*POW(NMBE_Temp,2)))/
(POW(0.3,2)+POW(0.6,2)+POW(0.1,2)),0.5)
FROM comparisons.GOFcalcs_monthly;

# Total Goodness of Fit - Final

DROP TABLE IF EXISTS comparisons.GOFmonthly_Total;
CREATE TABLE comparisons.GOFmonthly_Total (
    idkey INT(11) auto_increment PRIMARY KEY,
    Job INT(11) Not Null,
    CVRMSE_Heat Double (10 , 3 ) Not Null,
    CVRMSE_Elec Double (10 , 3 ) Not Null,
    CVRMSE_Temp Double (10 , 3 ) Not Null,
    NMBE_Heat Double (10 , 3 ) Not Null,
    NMBE_Elec Double (10 , 3 ) Not Null,
    NMBE_Temp Double (10 , 3 ) Not Null,
    GOF_Heat Double (10 , 3 ) Not Null,
    GOF_Elec Double (10 , 3 ) Not Null,
    GOF_Temp Double (10 , 3 ) Not Null,
    GOF_A Double (10 , 3 ) Not Null,
    GOF_B Double (10 , 3 ) Not Null,
    GOF_Total Double (10 , 3 ) Not Null
);

INSERT INTO comparisons.GOFmonthly_Total
(Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,
NMBE_Temp,GOF_Heat,GOF_Elec,GOF_Temp,GOF_A,GOF_B,GOF_Total)

SELECT
Job,CVRMSE_Heat,CVRMSE_Elec,CVRMSE_Temp,NMBE_Heat,NMBE_Elec,NMBE_Temp,GOF
_Heat,GOF_Elec,GOF_Temp,
GOF_A,GOF_B,
POW(((POW(0.6,2)*POW(GOF_A,2))
+(POW(0.4,2)*POW(GOF_B,2)))
/(POW(0.6,2)+POW(0.4,2))
,0.5)
FROM comparisons.GOFmonthly;


SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Heat.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
ORDER BY GOF_Heat
LIMIT 10;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Elec.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
ORDER BY GOF_Elec
LIMIT 10;
```

```
SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Temp.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
ORDER BY GOF_Temp
LIMIT 10;

SELECT
    *
INTO OUTFILE 'E:\\PhD Archive\\Parametric\\GOF\\Monthly\\GOF_Total.csv'
FIELDS TERMINATED BY ',' LINES TERMINATED BY '
' FROM
    comparisons.GOFmonthly_Total
ORDER BY GOF_Total
LIMIT 10;
```

A.12 Weather Forecast Data – Forecast.io (R)

```
#---------------------------------------------------------------------------#
# DESCRIPTION -  Weather Forecasting Access Script
# AUTHOR - Daniel Coakley (daniel.coakley@nuigalway.ie)
# DATE - October 2013
#---------------------------------------------------------------------------#

Weather Forecasting
========================================================

In order to use energy simulation models for the purpose of model-
predictive control (MPC), it is essential to have accurate, up-to-date,
weather and climate information for the region of interest. In this
respect, predicted weather data, for a period of at least 3 days, will
allow accurate simulation of predicted building performance, thus enable
improved response.

This type of control has the ability to reverse the current building
control dynamic, from reactive, to pro-active. In other words, the
building is no longer a passive observer of external conditions, but
rather is able to actively respond to changing external influences. This
may also include variables such as occupancy and external economic
influences.

## Available forecast data

There are a number of websites which provide local weather forecasting
services. However, we are particularly interested in those services which
also allow interaction with their data through open API's, such as that
provided by [forecast.io](http://forecast.io).

![Forecast.io provides an open API for developers](Images/forecastio.png)

## Downloading the data

Using forecast.io's open API, it is possible to develop applications to
read and use forecast data, by means of an **API Key**. You can sign up
for a devloper account [here](https://developer.forecast.io/), which
allows you to make up to 1000 calls to this API per day.

Obviously, it is also necessary to programatically access this data in
order to make use of it for your application. Recently,
[rud.is](http://rud.is/) developed
[RForecastIO](http://rud.is/b/2013/09/08/rforecastio-simple-r-package-to-
access-forecast-io-weather-data/), a simple R package which enables users
to access Forecast.IO data through R. A simple implementation is shown
below, thanks again to the guide on [rud.is
blog](http://rud.is/b/2013/09/08/rforecastio-simple-r-package-to-access-
forecast-io-weather-data/).

## Implementation

This is currently a simple implementation, following on from the above
blog post. However, I will be updating this in the coming months,
specifically with regard to integrating this data with building energy
performance simulation data.

First, install and include all necessary R-packages;
```

````
```{r
Packages,echo=TRUE,results='hide',warning=FALSE,error=FALSE,message=FALSE
}
library("devtools")
library("RJSONIO")

install_github("Rforecastio", "hrbrmstr")

library(Rforecastio)
library(ggplot2)
```
```{r Directory,echo=FALSE, results='hide',warning=FALSE}
setwd("E:/Dropbox/PhD/Model/Analysis/Weather")
```

Read in necessary parameters: API key, latitude and longitude. I am using
Galway, Ireland in this example.
```{r Input Params, echo=TRUE,
results='hide',warning=FALSE,error=FALSE,message=FALSE}

# NEVER put credentials or api keys in script bodies or github repos!!
# the "config" file has one thing in it, the api key string on one line
# this is all it takes to read it in

fio.api.key = readLines("forecast.io")
my.latitude = "53.2737969"
my.longitude = "-9.05177989"

fio.list <- fio.forecast(fio.api.key, my.latitude, my.longitude)
```

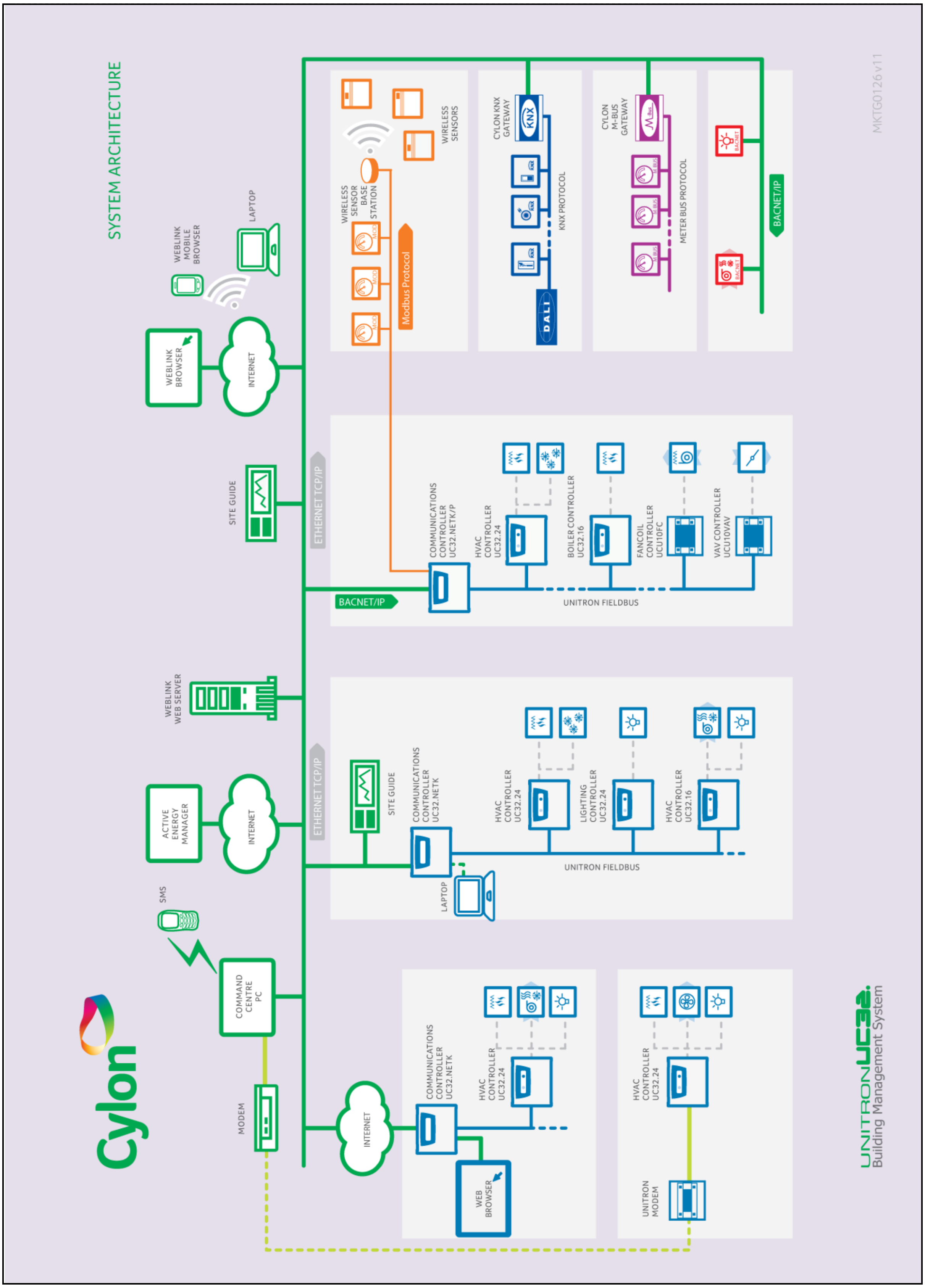
Set up the plot area
```{r Plot Params, echo=TRUE, results='hide',warning=FALSE}
forecast.x.min <- ISOdatetime(1970,1,1,0,0,0) + unclass(Sys.time())
forecast.x.max <- max(fio.list$hourly.df$time)
if (forecast.x.min > forecast.x.max) forecast.x.min <- forecast.x.max
fio.forecast.range.df <- data.frame(xmin=forecast.x.min,
xmax=forecast.x.max,
                                    ymin=-Inf, ymax=+Inf)
```

## Results
Finally, we can plot the results. Here, I am just plotting humidity
(green), temperature (red) and dewpoint (blue) for the next two days.

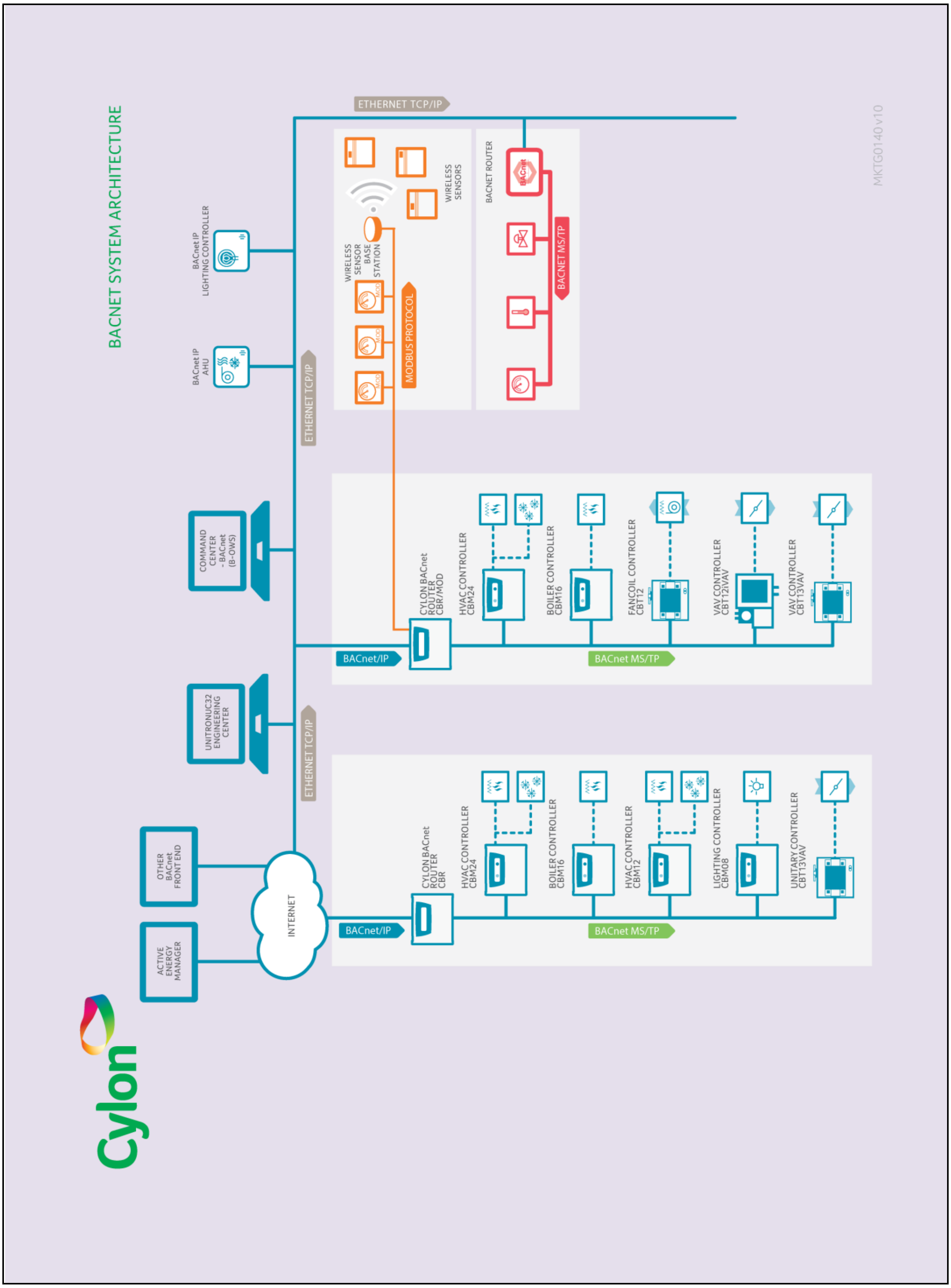
```{r Plot, echo=TRUE, results='hide',warning=FALSE}
fio.gg <- ggplot(data=fio.list$hourly.df,aes(x=time, y=temperature))
fio.gg <- fio.gg + labs(y="Readings", x="Time")
fio.gg <- fio.gg + geom_rect(data=fio.forecast.range.df,
                             aes(xmin=xmin, xmax=xmax,
                                 ymin=ymin, ymax=ymax),
                             fill="yellow", alpha=(0.15),
                             inherit.aes = FALSE)
fio.gg <- fio.gg + geom_line(aes(y=humidity*100), color="green")
fio.gg <- fio.gg + geom_line(aes(y=temperature), color="red")
fio.gg <- fio.gg + geom_line(aes(y=dewPoint), color="blue")
fio.gg <- fio.gg + theme_bw()
fio.gg
```
````

Appendix B: Drawings & Schematics

B.1 Cylon BMS - System Architecture



B.2 BACnet - System Architecture



B.3 Nursing Library - Layouts

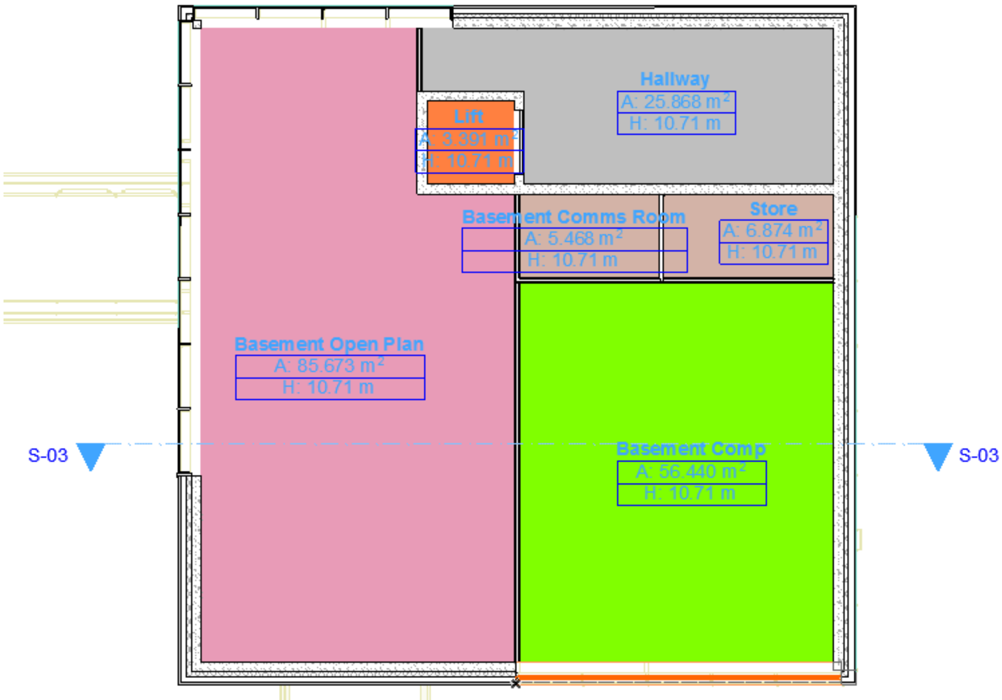


Figure 6-1: Basement layout

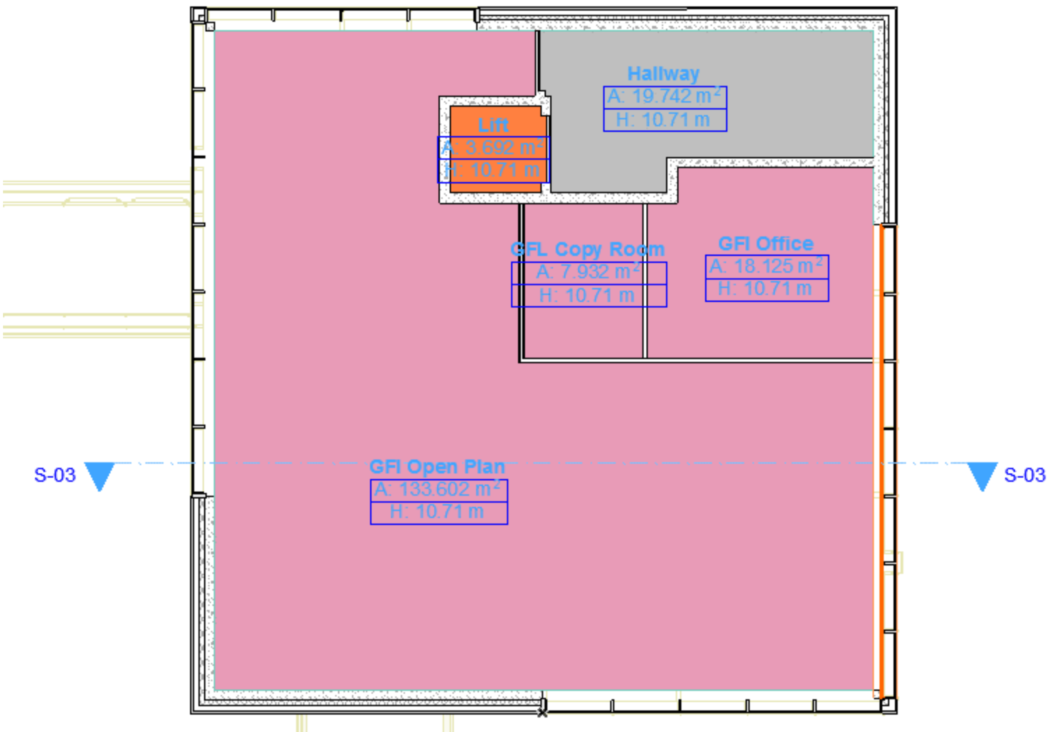


Figure 6-2: Ground Floor layout

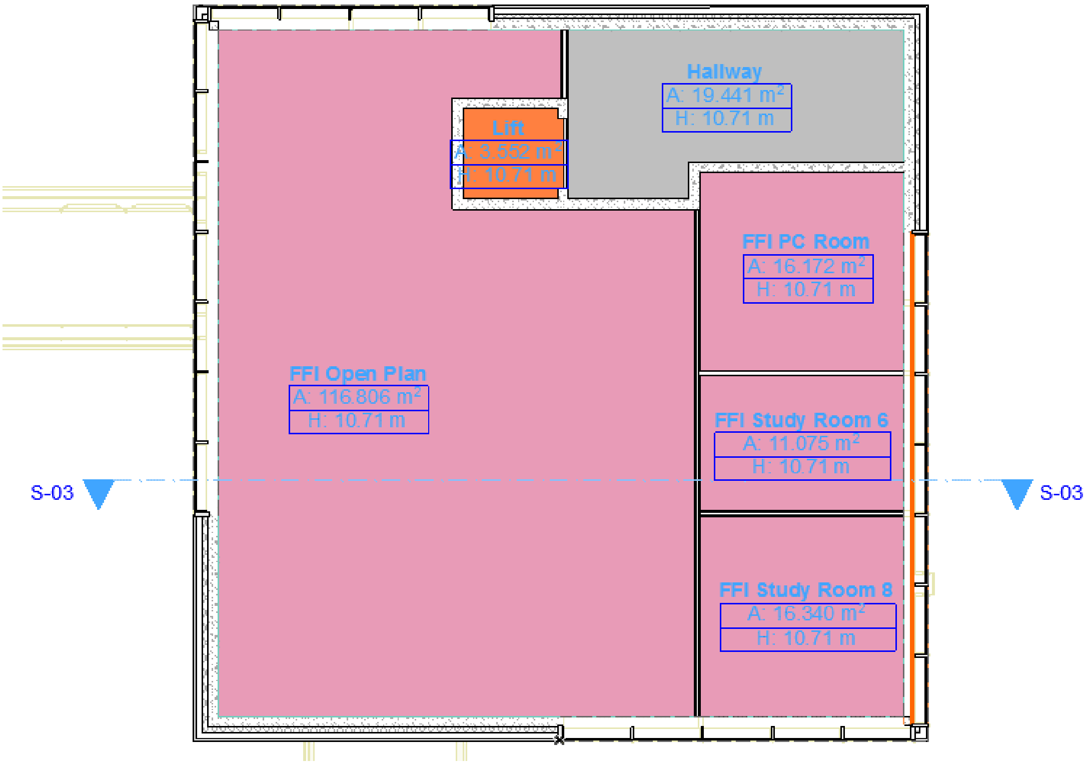


Figure 6-3: First Floor Layout

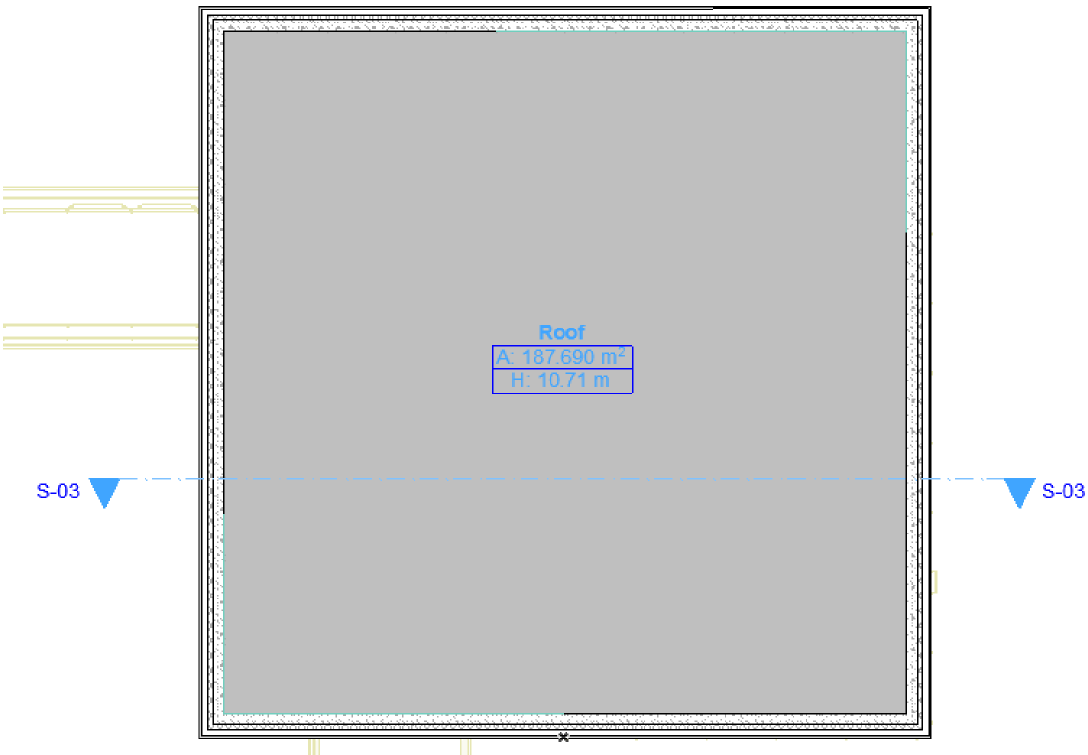


Figure 6-4: Roof Layout

Appendix C:
Model Development

C.1 Source Hierarchy

Category	Source type	Ref	Description	Associated file(s) or folder(s)	Revision
1	**BMS / Logged Measured Data**	a	BMS Electrical Energy Consumption Data	[Database]\D0010208	14
		b	BMS Heat Energy Consumption Data	[Database]\D0010117	2
		c	BMS Room Temperature Data	[Database]\'Varies'	4
		d	Labstats PC Usage	\Input Data\Audits\Electrical Data\Labstats\SnapshotData.xlsx	11
		e	Weather Data	\Calibration Data\Weather\WeatherData.csv	3
2	**Spot-measured data**	f	Spot-measured electrical data (Aug - Sep 2011)	\Input Data\Audits\Electrical Data\Electrical Audit - Sept 2011.xlsx	14
		g	Spot-measured flow/return HDR temperature (Dec 2011)	\BMS Data\LPHW\HOBO Comparisons\HOBO-BMS Comparison.xlsx	N/A
		h	Occupancy counter - spot-checks during surveys	\Input Data\Audits\Occupancy\Occupancy Survey Sheet.xlsx	11
3	**Surveys and physical verification**	i	Geometry verified by physical inspection	N/A	1
		j	Office occupancy counts (8 surveys - Academic / Summer)	\Input Data\Audits\Occupancy\Occupancy Survey Sheet.xlsx	11
		k	Photographic evidence	\Input Data\Photos	2
		l	Equipment Survey	\Input Data\Audits\Equipment\Equip Audit.xlsx	2
4	**Interviews**	m	Building Manager Interview (N. O'Connor)	Meetings\Correspondance	2,11,14
5	**As-built documentation**	n	Architectural as-built drawings	\Input Data\Layouts\Achitect	1
		o	Engineering As-Built Drawings	\Input Data\Layouts\Engineeing	1
		p	Services Layouts	\Input Data\Layouts\Services	1
		q	Structural As-Built Drawings	\Input Data\Layouts\Structural	1
6	**Operation & Maintenance manuals**	r	BMS Controls Layout & Operation	\Input Data\O&M Manuals\BMS\Boilerhouse controls.pdf	2
		s	BMS O&M Manual	\Input Data\O&M Manuals\Mechanical O&M Manual\Section 2\3 BMS System\	2
7	**Guides and Standards**	t	EnergyPlus Input/Output Reference v7.0	InputOutputReference.pdf	1
		u	ASHRAE 2005 Handbook of Fundamentals	N/A	
		v	BS EN 12524:2000	N/A	7,9,10
		w	BRE 443 (2006) - Conventions for U-Value Calculations	N/A	
		x	CIBSE Guide B - Heating, ventilating, air-conditioning and refrigeration	N/A	
		y	CIBSE Guide C - Reference Data	N/A	7,9,10
8	**Default Values / No Information**	z	EnergyPlus Input/Output Reference v7.0		

C.2 Occupancy Audit (Excerpt)

		Basement Floor					Ground Floor			1st Floor									Misc	
Date	Time	People Study Area	W	People Comp Lab	W	Laptops	People Study Area	W	Laptops	People Study Area	W	PC Meeting Room 1 (5)	W	Group Study Room 2 (6)	W	Group Study Room 2 (8)	W	Laptops	Counter	Out
07/08/2011	14:45	22		20		12	7		?	28		4		1		3		10	41897	41997
	17:30	33		7			9			25		3		1		0			42558	
	21:25	6		4			6			6		1		0		0			43051	
08/03/2011	12:00	20		7		14	10		7	21		7						18		
	13:30	20		18		14	8		7	21		4		3		3		18	44062	
	14:10	21		4			8			29		4		3		4			44214	44248
	18:00	16		6		5	14		2	17		12		0		5		8	45013	
	19:25	10		2			7		?	19		3		0		0			45176	
09/03/2011	10:00	13		28		3	11		3	17		6		1		3		4	45644	45664
	12:54	15		11		17	9		9	26		1		0		2		12	46208	
	13:55	17		20		15	8		7	21		2	1	3	1	0	1	11	46408	46447
	14:44	18		17			8	2		22		3	2	3	0	0	0		46578	
	17:20	20	1	15	0		12	0	4	26	1	2	1	4	1	3	1	12	47122	47129
	20:42	11	0	10	1	6	8	0	12	12	1	3	1	0	0	0	0	3	47566	
10/03/2011	09:22	3	0	1	0	0	8	0	3	13	0	5	1	2	1	2	0	3	47806	47817
	12:15	25	0	11			10	1	7	25	0	5	1	0	1	0	0	13	48230	48249
	13:55	22	2	12	4	5	14	0	2	17	7	1	1	0	1	0	0	15	48628	48662
11/03/2011	12:15	6	0	20	0	4	10	0	7	14	0	5	0	5	1	6	0	16	50220	50236
	13:55	4		9		5	10		9	13		3		2		0			50424	50426
	15:20	7	0	9	4	2	8	0	5	17	0	5	0	2	0	0	0	5	50554	50557
	16:40	10		9		9	8		7	26		2		2		0		24	50663	50668
	20:40	2		1		0	2		2	8		5		0		0		9	50913	50919
13/03/2011	15:25	7	0	7	0	1	10		1	9		3		0		0		1	51538	51561

C.3 Parametric Analysis worksheet

ID	Class List	Location	Object	Field	Initial Val	Units	Discrete	Op	Fix	jE	Source	Class	ROV (%)	Std Dev	String
1195			Material	1	N_Lib_Screed_01										
1196	Surface Co	Geometry	Material	Name	N_Lib_Screed_01		Discrete	N	Y	Y	Not Appli	NA	NA	NULL	&&GeoNam1196&&
1197	Surface Co	Geometry	Material	Roughnes	MediumRough		Discrete	N	Y	Y	Default Va	6	40	NULL	&&GeoRou1197&&
1198	Surface Co	Geometry	Material	Thickness	0.075		Continuou	N	N	Y	As-Built D	3	10	0.0025	&&GeoThi1198&&
1199	Surface Co	Geometry	Material	Conductiv	1.15		Continuou	N	N	Y	Guides &	5	30	0.115	&&GeoCon1199&&
1200	Surface Co	Geometry	Material	Density	1200		Continuou	N	N	Y	Guides &	5	30	120	&&GeoDen1200&&
1201	Surface Co	Geometry	Material	Specific H	1000		Continuou	N	N	Y	Guides &	5	30	100	&&GeoSpe1201&&
1202	Surface Co	Geometry	Material	Thermal A	0.9		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoThe1202&&
1203	Surface Co	Geometry	Material	Solar Abs	0.73		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoSol1203&&
1204	Surface Co	Geometry	Material	Visible Ab	0.73		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoVis1204&&
1205			Material	1	N_Lib_Screed_02										
1206	Surface Co	Geometry	Material	Name	N_Lib_Screed_02		Discrete	N	Y	Y	Not Appli	NA	NA	NULL	&&GeoNam1206&&
1207	Surface Co	Geometry	Material	Roughnes	MediumRough		Discrete	N	Y	Y	Default Va	6	40	NULL	&&GeoRou1207&&
1208	Surface Co	Geometry	Material	Thickness	0.05		Continuou	N	N	Y	As-Built D	3	10	0.001667	&&GeoThi1208&&
1209	Surface Co	Geometry	Material	Conductiv	1.15		Continuou	N	N	Y	Guides &	5	30	0.115	&&GeoCon1209&&
1210	Surface Co	Geometry	Material	Density	1200		Continuou	N	N	Y	Guides &	5	30	120	&&GeoDen1210&&
1211	Surface Co	Geometry	Material	Specific H	1000		Continuou	N	N	Y	Guides &	5	30	100	&&GeoSpe1211&&
1212	Surface Co	Geometry	Material	Thermal A	0.9		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoThe1212&&
1213	Surface Co	Geometry	Material	Solar Abs	0.73		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoSol1213&&
1214	Surface Co	Geometry	Material	Visible Ab	0.73		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoVis1214&&
1215			Material	1	N_Lib_Struct_Concrete_01										
1216	Surface Co	Geometry	Material	Name	N_Lib_Struct_Concre		Discrete	N	Y	Y	Not Appli	NA	NA	NULL	&&GeoNam1216&&
1217	Surface Co	Geometry	Material	Roughnes	Rough		Discrete	N	Y	Y	Default Va	6	40	NULL	&&GeoRou1217&&
1218	Surface Co	Geometry	Material	Thickness	0.25		Continuou	N	N	Y	As-Built D	3	10	0.008333	&&GeoThi1218&&
1219	Surface Co	Geometry	Material	Conductiv	2.5		Continuou	N	N	Y	Guides &	5	30	0.25	&&GeoCon1219&&
1220	Surface Co	Geometry	Material	Density	2300		Continuou	N	N	Y	Guides &	5	30	230	&&GeoDen1220&&
1221	Surface Co	Geometry	Material	Specific H	1000		Continuou	N	N	Y	Guides &	5	30	100	&&GeoSpe1221&&
1222	Surface Co	Geometry	Material	Thermal A	0.9		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoThe1222&&
1223	Surface Co	Geometry	Material	Solar Abs	0.6		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoSol1223&&
1224	Surface Co	Geometry	Material	Visible Ab	0.6		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoVis1224&&
1225			Material	1	N_Lib_Struct_Concrete_02										
1226	Surface Co	Geometry	Material	Name	N_Lib_Struct_Concre		Discrete	N	Y	Y	Not Appli	NA	NA	NULL	&&GeoNam1226&&
1227	Surface Co	Geometry	Material	Roughnes	Rough		Discrete	N	Y	Y	Default Va	6	40	NULL	&&GeoRou1227&&
1228	Surface Co	Geometry	Material	Thickness	0.275		Continuou	N	N	Y	As-Built D	3	10	0.009167	&&GeoThi1228&&
1229	Surface Co	Geometry	Material	Conductiv	2.5		Continuou	N	N	Y	Guides &	5	30	0.25	&&GeoCon1229&&
1230	Surface Co	Geometry	Material	Density	2300		Continuou	N	N	Y	Guides &	5	30	230	&&GeoDen1230&&
1231	Surface Co	Geometry	Material	Specific H	1000		Continuou	N	N	Y	Guides &	5	30	100	&&GeoSpe1231&&
1232	Surface Co	Geometry	Material	Thermal A	0.9		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoThe1232&&
1233	Surface Co	Geometry	Material	Solar Abs	0.6		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoSol1233&&
1234	Surface Co	Geometry	Material	Visible Ab	0.6		Continuou	N	Y	Y	Guides &	5	30	0	&&GeoVis1234&&
	Material:AirGap Count			3											
1236			Material:A	1	N_Lib_Air_Space										
1237	Surface Co	Geometry	Material:A	Name	N_Lib_Air_Space		Discrete	N	Y	N	As-Built D	3	10	NULL	&&GeoNam1237&&
1238	Surface Co	Geometry	Material:A	Thermal R	0.15		Continuou	N	N	N	Guides &	5	30	0.015	&&GeoThe1238&&
	Material:InfraredTranspa			2											
1240			Material:I	1	IDFGENERATOR IRTMaterial										
1241	Surface Co	Geometry	Material:I	Name	IDFGENERATOR IRTM		Discrete	N	Y	N	Default Va	6	40	NULL	&&GeoNam1241&&
	WindowMaterial:Glazing			12											
1243			WindowM	1	Clear 3mm										
1244	Surface Co	Geometry	WindowM	Name	Clear 3mm		Discrete	N	Y	N	As-Built D	3	10	NULL	&&GeoNam1244&&
1245	Surface Co	Geometry	WindowM	Optical Da	SpectralAverage		Discrete	N	Y	N	Default Va	6	40	NULL	&&GeoOpt1245&&
1246	Surface Co	Geometry	WindowM	Window Glass Spectral Data Se			Continuou	Y	Y	N	Default Va	6	40	NULL	&&GeoWin1246&&
1247	Surface Co	Geometry	WindowM	Thickness	0.003		Continuou	N	N	N	As-Built D	3	10	0.0001	&&GeoThi1247&&
1248	Surface Co	Geometry	WindowM	Solar Tran	0.837		Continuou	N	N	N	Default Va	6	40	0.1116	&&GeoSol1248&&
1249	Surface Co	Geometry	WindowM	Front Side	0.075		Continuou	N	N	N	Default Va	6	40	0.01	&&GeoFro1249&&
1250	Surface Co	Geometry	WindowM	Back Side	0.075		Continuou	N	N	N	Default Va	6	40	0.01	&&GeoBac1250&&
1251	Surface Co	Geometry	WindowM	Visible Tra	0.898		Continuou	N	N	N	Default Va	6	40	0.119733	&&GeoVis1251&&
1252	Surface Co	Geometry	WindowM	Front Side	0.081		Continuou	N	N	N	Default Va	6	40	0.0108	&&GeoFro1252&&
1253	Surface Co	Geometry	WindowM	Back Side	0.081		Continuou	N	N	N	Default Va	6	40	0.0108	&&GeoBac1253&&
1254	Surface Co	Geometry	WindowM	Infrared T	0		Continuou	N	N	N	Default Va	6	40	0	&&GeoInf1254&&

Appendix D: Other

D.1 Calibration Papers - Techniques

Author	Title	Journal / Conference	Year	Tools	Calibration Type	Analytical Techniques	Math. Methods	Building Type	System Type
Subbarao, K.	Primary and secondary terms analysis and renormalization: A unified approach to building energy simulations and short-term testing	Solar Energy Research Inst.	1988	PSTAR	Manual	PSTAR, STEM	N/A	Residential	N/A
Kaplan, MB; McFerran, J; Jansen, J & Pratt, R	Reconciliation of a DOE2. 1C model with monitored end-use data for a small office building	ASHRAE Transactions	1990	DOE-2	Manual	STEM, PARRED	N/A	Office	Mechanical HVAC
Bronson, D.J.; Hinchey, S.B.; Haberl, J.S. & O'Neal., D.L.	A procedure for calibrating the DOE-2 simulation program to non-weather-dependent measured loads	Proc. Of ASHRAE Winter Meeting	1992	DOE-2	Manual	3D	N/A	Government / Large Office	N/A
Waltz, JP	Practical experience in achieving high levels of accuracy in energy simulations of existing buildings	ASHRAE Transactions	1992	N/A	Manual	AUDIT, EXPERT, STAT	N/A	N/A	N/A
Katipamula, S & Claridge, DE	Use of simplified system models to measure retrofit energy savings	Journal of Solar Energy Engineering	1993	SEAP	Manual	SEAP	N/A	Engineering Center	DDCV with VAV
Carroll, W.L. & Hitchcock, R. J.	Tuning simulated building descriptions to match actual utility data: methods and implementation	ASHRAE Transactions	1993	DOE-2	Automated	EXPERT	PENALTY	School	Mixed
Clarke, J. A.; Strachan, P. A. & Pernot, C.	An approach to the calibration of building energy simulation models	ASHRAE Transactions	1993	ESP-r	Manual	HIGH, SA	N/A	Test Cell	Passive Solar (PASSYS)

Author	Title	Journal / Conference	Year	Tools	Calibration Type	Analytical Techniques	Math. Methods	Building Type	System Type
Bou-Saada, T. E. & Haberl, J. S.	An improved procedure for developing calibrated hourly simulation models	Proceedings of the 5th International IBPSA Conference	1995	DOE-2.1D	Automated	EVIDENCE, 3D	N/A	Office (US DOE Forrestal Complex)	Mechanical HVAC
Manke, JM; Hittle, DC & Hancock, CE	Calibrating building energy analysis models using short term test data	Journal of solar energy engineering	1996	BLAST	Manual	STEM	N/A	School	Furnace and air-conditioning units
Soebarto, VI	Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings	Proceedings of the 4th International IBPSA Conference	1997	DOE-2	Manual	INT	N/A	University and Municipal Building	Dual-Duct Variable Air Volume (DDVAV)
Haberl, J. S JS & Bou-Saada, T. E. TE	Procedures for calibrating hourly simulation models to measured building energy and environmental data	Journal of solar energy engineering	1998	DOE-2.1D	Automated	HIGH, 3D	N/A	Office (US DOE Forrestal Complex)	Mechanical HVAC
Reddy, TA; Deng, S & Claridge, DE	Development of an inverse method to estimate overall building and ventilation parameters of large commercial buildings	Journal of solar energy engineering	1999	DOE-2	Manual	MPE	N/A	Commercial Office	Mechanical HVAC
Lunneberg, TA	Improving simulation accuracy through the use of short-term electrical end-use monitoring	Proceedings of the 6th International IBPSA Conference	1999	DOE-2	Manual	STEM	N/A	Office	N/A

Author	Title	Journal / Conference	Year	Tools	Calibration Type	Analytical Techniques	Math. Methods	Building Type	System Type
Yoon, JH & Lee, EJ	Calibration procedure of energy performance simulation model for a commercial building	Proceedings from the Building Simulation Conference	1999	DOE-2.1E	Manual	BASE, EVIDENCE	N/A	Large Commercial Building	Package AHU
Liu, M; Claridge, D E; Bensouda, N; Heinemeier, K; Lee, Seung Uk & Wei, G	High Performance Commercial Building Systems Manual of Procedures for Calibrating Simulations of Building Systems	Report. California Energy Commission	2003	DOE-2	Manual	SIG, STAT	N/A	Office Building / Campus Building /	SDCV / SDVAV / DDCV / DDVAV
Yoon, Jongho; Lee, E. J. & Claridge, D. E.	Calibration Procedure for Energy Performance Simulation of a Commercial Building	Journal of Solar Energy Engineering	2003	DOE-2.1E	Manual	BASE, EVIDENCE, STAT	N/A	Large Commercial Building	Package AHU
Liu, S. & Henze, G. P	Calibration of building models for supervisory control of commercial buildings	Proceedings of the 9th International IBPSA Conference	2005	EnergyPlus / GenOpt /	Auto	N/A	SYS	N/A	AHU / VAV / Icemaking Chiller
Westphal, FS & Lamberts, R.	Building simulation calibration using sensitivity analysis	Proceedings of the 9th International IBPSA Conference	2005	EnergyPlus (v1.2)	Manual	SA	N/A	Office	Mixed Mode
Sun, Jian & Reddy, T. Agami	Calibration of Building Energy Simulation Programs Using the Analytic Optimization Approach (RP-1051).	HVAC&R Research	2006	DOE-2	Auto	UQ	MC	Pseudo-Building	N/A

Author	Title	Journal / Conference	Year	Tools	Calibration Type	Analytical Techniques	Math. Methods	Building Type	System Type
Reddy, T. Agami TA; Maor, Itzhak & Panjapornpon, Chanin	Calibrating Detailed Building Energy Simulation Programs with Measured Data-- Part I: General Methodology (RP-1051).	HVAC&R Research	2007	DOE-2	Auto	EXPERT, UQ	MC	N/A	N/A
Reddy, T. Agami; Maor, Itzhak & Panjapornpon, Chanin	Calibrating Detailed Building Energy Simulation Programs with Measured Data--Part II: Application to Three Case Study Office Buildings (RP-1051).	HVAC&R Research	2007	DOE-2	Auto	EXPERT, UQ	MC	Office	VAV System
Neto, Alberto Hernandez & Fiorelli, Flavio Augusto Sanzovo	Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption	Energy and Buildings	2008	EnergyPlus	Auto	N/A	ANN	University Building	N/A
Raftery, Paul; Keane, Marcus & Costa, Andrea	Calibrating whole building energy models: Detailed case study using hourly measured data	Energy & Buildings	2011		Manual	HIGH, EVIDENCE	N/A	Office	VAV Systems
Liu, Guopeng & Liu, Mingsheng	A rapid calibration procedure and case study for simplified simulation models of commonly used HVAC systems	Building and Environment	2011	SEAP	Manual	SEAP, SIG	N/A	Office	VAV systems, CV system
Raftery, Paul; Keane, Marcus; & O'Donnell, James	Calibrating whole building energy models: An evidence-based methodology	Energy and Buildings	2011	EnergyPlus	Manual	EVIDENCE, PARRED	N/A	N/A	N/A

Author	Title	Journal / Conference	Year	Tools	Calibration Type	Analytical Techniques	Math. Methods	Building Type	System Type
Heo, Y.; Choudhary, R & Augenbroe, GA	Calibration of building energy models for retrofit analysis under uncertainty	Energy and Buildings	2012	EnergyPlus	Auto	UQ	BAYES	Office	Gas Condensing Boiler/ /Natural Ventilation
Coakley, Daniel; Raftery, Paul & Molloy, Padraig	Calibration of Whole Building Energy Simulation Models: Detailed Case Study of a Naturally Ventilated Building Using Hourly Measured Data	Proc. of Building Simulation & Optimisation Conference	2012	EnergyPlus	Auto	HIGH, EVIDENCE, STEM	PEN	University Building	Natural Ventilation
Manfren, Massimiliano; Aste, Niccola & Moshksar, Reza	Calibration and uncertainty analysis for computer models: "A meta-model based approach for integrated building energy simulation"	Applied Energy	2013	N/A	Auto	UQ	BAYES, META	Office Building / Campus Building /	N/A
Booth, A.T.; Choudhary, R. & Spiegelhalter, D.J.	A hierarchical Bayesian framework for calibrating micro-level models with macro-level data	Journal of Building Performance Simulation	2013	N/A	Auto	UQ	BAYES	Residential	Varies
O'Neill, Z & Eisenhower, B	Leveraging the analysis of parametric uncertainty for building energy model calibration	Building Simulation	2013	EnergyPlus 6.0	Auto	SA, UQ	META	Military Barracks	VAV System

Notes

[1] N/A: Not applicable to this case.

[2] X: Information / paper not available at time of writing.

[3] A more comprehensive version of this table can be found on the attached CD (\Appendices\Appendix D\CalibrationPapers.xlsx)

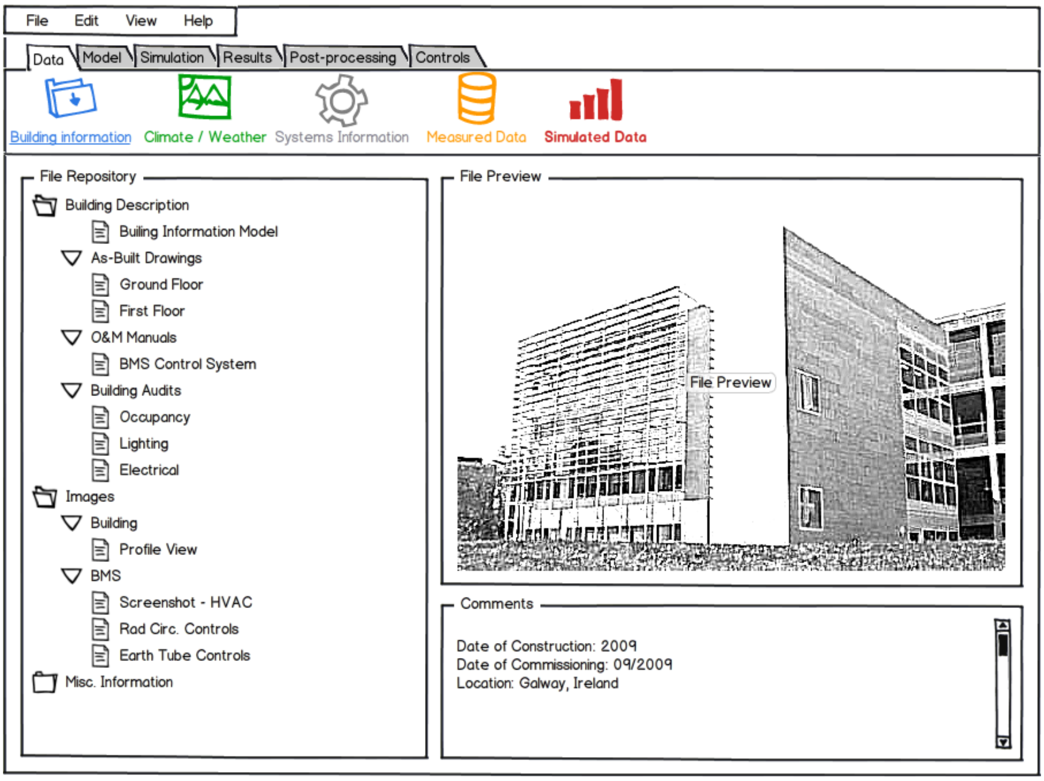
D.2 Calibration Toolkit

This project proposes the creation of an integrated environment to assist in the calibration of Building Energy Simulation (BES) models. The tool will help users to follow the steps proposed in the calibration methodology outlined in this thesis. The following steps will be integrated into the BES calibration toolkit:

- Data Gathering;
- BES Model Development;
- Uncertainty Analysis;
- Parametric Analysis;
- Sensitivity Analysis;
- Visualisation;

Each module is described in further detail below:

Data gathering: The first module serves as a repository for Building Information necessary for the creation and calibration of Building Energy Simulation models. Files pertaining to the building, HVAC systems and environment may be added to the repository and categorised accordingly.



Useful features of this module could include:

- Comment Box – enables users to add and store comments regarding documentation as it is added to the repository.
- Datestamp – add date and timestamp for repository files
- Version Control – Integrated VC tool to track and record changes made to repository files and restore un-wanted changes as necessary.
- Source classification – enables the user to specify the reliability of source documentation. May be a useful feature when conducting parameter uncertainty analysis.
- Data Capture from BMS System – Auto-import data to SQL database.

The main purpose of the Data module is to enable the user to quickly organise, browse and classify source evidence. This will be a useful reference when building and updating the BES model.

BES Model Development: This module would enable the user to access the current BES model and edit various elements of the model. This may be integrated into the toolkit using existing open-source modules or may act as a UI to call various tools as required.

Currently, there is no comprehensive GUI for EnergyPlus, although there is a tool under development (Simergy). The BES model used for our detailed calibration case-study was developed using a combination of the following tools and interfaces:

- Google SketchUp and OpenStudio – Sketchup allows for the creation of building geometry while the OpenStudio plug-in allows users to add simulation information (such as zones, materials and constructions) to the model;
- Note: Since this project began, Google SketchUp has been sold to 'Trimble' and OpenStudio has been developed as a stand-alone package. However, OpenStudio remains available as an open-source tool and integration into third-party tools is actively encouraged.
- HVAC Generator – an excel-based user-interfaced which enables the specification of zone and system specific variable values as well as the amalgamation of EnergyPlus macro input files;
- TortoiseSVN – version control software used to track and record changes to the BES model.

Ideally, these tools would be integrated into a single interface. Alternatively, the BES model module may call the various tools required for model development.

Useful features of this module could include:

- **Version Control** – display current model version along with comments. Automatically prompt user to update the version control repository when saving changes to any of the BES model files.
- **Model Import** – Import Geometry files from various tools (ArchiCAD, SketchUp, AutoCAD etc.) using the standardised .ifc file format and clean/prepare file for simulation.
- **Visual representation** – ability to view a virtual representation of the BES model. A previous MSc SD&D project dealt with the web-based visualisation of building models and ability to view time-series plots of data points associated with various Building Zones. This could serve as a useful extension of this previous work. (Ref. to Peter O'Neill thesis). This visualisation tool could also be a useful method of presenting error values (Measured vs. Simulated data) at various hours of the day (or averaged over any time period) or for visualising measured data in various building zones. Refer to data capture – integration with Measured BMS data.

Parameter Uncertainty Analysis: This module refers to the prediction of uncertainty associated with individual model parameters. Currently, this process is managed using the *Parametric Analysis* excel worksheet. However, this process could be vastly improved by automating the following processes:

- Generate parameter list – at present, I have manually created a list of all objects and parameters contained in our BES model. I have also manually classified these variables as Discrete, Continuous or Multi-Dimensional. Improvements could be made by:
- Automatically generate the list of model parameters from the EnergyPlus epmidf – this could be achieved using the same IDF Editor module contained in EnergyPlus.
- Assign Ranges of Variation – this could be automatically achieved if the objects in the BES module were linked to the Sources contained in the Data Gathering module. For example, if we auto-classify all As-Built Drawings to be 95% accurate, we can auto-assign a ROV to all variables with this linked source evidence.

- Generate parametric study job list – This could be easily achieved by auto-generating a unique string for each model parameter and generating samples based on the assigned ROV using a Java math function.

Parametric Study: This module refers to the process of running a parametric study based on the defined job list from the previous module. This may simply be achieved by calling jEPlus (can be called as a command line function) and generating the desired results.

There are a number of features which would make this module more flexible for the user depending on the general project requirements. For example, we may want to change the EnergyPlus output parameters depending on the type of study we are conducting:

- Calibration: We may just want to calibrate Heat or Electrical Energy use or a combination of different outputs. We may also want to change the resolution of our calibration study – hourly/weekly/monthly calibration.
- Design: In a design case, we may just want to look at some form of cost function analysis when using different insulating materials or glazing surfaces. Here, we may define Energy use outputs along with a custom post-processing function to calculate the Value of construction/retrofit measures.

The above tasks can be handled using jEPlus by implementing a number of tweaks. Therefore, this module could be completed separately by tweaking the jEPlus source code (as discussed in the other MSc proposal) and integrated into the final calibration tool.

Sensitivity Analysis (Optional): This module could be implemented as a means of calculating parameter sensitivity for a given (set of) output(s). At present, we have not implemented a sensitivity analysis procedure as part of our methodology. However, it would certainly be a useful feature in any Building Energy Performance analysis toolkit.