ENERGY
www.csiro.au

Residential air-conditioner use behaviour & building performance analysis

Mark Goldsworthy, Peter Toscas, Ying Guo and Omid Motlagh

June, 2017

Citation

Goldsworthy, M., Toscas, P., Guo, Y. and Motlagh, O., 2017. Residential air-conditioner use behaviour & building performance analysis. CSIRO Energy.

Copyright and disclaimer

© 2017 CSIRO To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

Important disclaimer

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

Abbreviations and acronyms

ABS	Australian Bureau of Statistics
AC	air conditioning
CDD	cooling degree days
DTS	deemed to satisfy
EUDM	Energy Use Data Model
HDD	heating degree days
NEM	National Energy Market
PCA	principal component analysis
PMV	predicted mean vote
PPD	percentage persons dissatisfied
RBEE	Residential Building Energy Efficiency
RBEES	Residential Building Energy Efficiency Study
RSD	relative standard deviation
SVM	support vector machine

Contents

Executive summary 1

1 Introduction 4

2 Overall thermal performance of the building shell: modelling analysis 6

2.1 Introduction 6

2.2 Comfort versus energy use trade-off 8

2.3 Sensitivity of buildings to occupant behaviours 17

2.4 Discussion & limitations 21

2.5 Summary & next steps 23

2.6 References 24

3 Influence of building parameters on total energy consumption- statistical analysis 26

3.1 Introduction 26

3.2 South-East Queensland monthly energy use data set 26

3.3 Annual energy use across Australia 39

3.4 Discussion & limitations 53

3.5 Summary & next steps 60

3.6 References 60

4 Understanding household space heating behaviour – RBEE data set analysis and support vector machine modelling 62

4.1 Introduction 62

4.2 RBEE data set analysis of space heating behaviour 62

4.3 References 68

5 Linking building indoor and outdoor temperatures – generic predictive model 69

5.1 Introduction 69

5.2 Analysis method 69

5.3 Summary of outputs, limitations and next steps 75

6 Conclusion 77

7 APPENDIX A 78

8 APPENDIX B 80

List of Figures

Figure 2-1 Temperature levels for air-conditioning control. Note apparent temperature used for trigger temperatures 9

Figure 2-2 Contours indicating net heating (positive) and net cooling (negative) climates based on annual subtraction of apparent temperature cooling degree days from heating degree days (HDD-CDD); capital city locations marked 11

Figure 2-3 Annual mean scale parameter for discomfort and energy for selected building parameters plotted for all climate zones across Australia. Colour indicates relative heating (red) or cooling dominated climates (dark blue). Selected capital city climate zones are indicated as labelled. Note: different axis scale for ceiling insulation plot 13

Figure 2-4 Annual mean scale parameters for discomfort and energy for different orientation component pairs plotted for all climate zones across Australia. The left plot shows a component comparison for orientations with N/S components. The right plot shows a component comparison for orientations with E/W components. Colour indicates relative heating (red) or cooling dominated climates (dark blue). 15

Figure 2-5 Mean discomfort (vertical axis) versus mean energy use (horizontal axis) for each simulated building for four different climates. Clockwise from top left: Darwin, Brisbane, Hobart, Melbourne. The colour of the dots indicates the relative magnitude of the sensitivity of discomfort and energy use to different occupant behaviours. 20

Figure 2-6 Annual average percentage persons dissatisfied (%) versus energy use (kWh) as a function of AC usage behaviour (i.e. varying PPD_t) for two selected buildings in the Darwin climate. Both buildings have approximately the same mean energy use averaged across all values of PPD_t, although building A has a significantly lower mean discomfort. 20

Figure 2-7 Sensitivity of discomfort to variations in occupant AC usage behaviour by building thermal parameters for Darwin (tropical climate) in buildings with the lowest 30% of energy consumption. Lower numbers indicate less sensitivity 21

Figure 3-1 Monthly average energy consumption for the different postcodes 35

Figure 3-2 Biplot of first two components for dwelling type 35

Figure 3-3 Biplot of first two components for number of bedrooms. (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms) 36

Figure 3-4 Biplot of first two components for income classes. (I_0_31 = income <$31,000; I_31_65 = $31,000 ≤ Income < $65,000; I_65_104 = $65,000 ≤ Income < $104,000; I_104_156 = $104,000 ≤ Income < $156,000; and I156p = Income ≥ $156,000) 36

Figure 3-5 Biplot of first and third components for building age classes. (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011. Note that the data from Geoscience Australia had no buildings listed as having been built after 2011 for South-East Queensland) 37

Figure 3-6 Biplot of first two components for wall type materials. (W_C_S_M = cavity and solid masonry; W_F_C = fibre cement; W_M_S = metal sheeting; W_T = timber; and W_V_M = veneer masonry) 37

Figure 3-7 Biplot of first two components for floor area by dwelling type. (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two storey apartments; THREE_S_A = area for three storey apartments; and FOUR_P_S_A = area for four or more storey apartments)..... 38

Figure 3-8 Biplot of second and third components for floor area by dwelling type. (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two storey apartments; THREE_S_A = area for three storey apartments; and FOUR_P_S_A = area for four or more storey apartments) 38

Figure 3-9 Biplot of the first and third principal components for equivalised income..... 48

Figure 3-10 Biplot of first and third components for building age classes. (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011) 48

Figure 3-11 Biplot of the first two components for roof type material (IT = imitation tiles; MS = metal sheeting; FC = fibre cement; and T = tiles) 49

Figure 3-12 Biplot of the first and third components for roof type material (IT = imitation tiles; MS = metal sheeting; FC = fibre cement; and T = tiles)..... 49

Figure 3-13 Biplot of the first two components for wall type materials (RCM = reinforced concrete masonry; MS = metal sheeting; FC = fibre cement; T = timber; and CSM = cavity and solid masonry)..... 50

Figure 3-14 Biplot of the third and fourth components for wall type materials (RCM = reinforced concrete masonry; MS = metal sheeting; FC = fibre cement; T = timber; and CSM = cavity and solid masonry)..... 50

Figure 3-15 Biplot of the fourth and fifth components for wall type materials (RCM = reinforced concrete masonry; MS = metal sheeting; FC = fibre cement; T = timber; and CSM = cavity and solid masonry)..... 51

Figure 3-16 Biplot of the first two components for floor area for different dwelling types (SH = separate house; SDD = semi-detached dwelling; A2S = two storey apartments; A3S = three storey apartments; and A4S = four or more storey apartments) 51

Figure 3-17 Biplot of the third and fourth components for floor area for different dwelling types (SH = separate house; SDD = semi-detached dwelling; A2S = two storey apartments; A3S = three storey apartments; and A4S = four or more storey apartments) 52

Figure 3-18 Biplot of the second and third components for tenure type (OWNT = own tenure; RPRT = private rent tenure; RPUT = public rent tenure; and OTHT = other tenure) 52

Figure 3-19 Biplot of first two components for number of bedrooms (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms) and income

classes (I_0_31 = income <$31,000; I_31_65 = $31,000 ≤ Income < $65,000; I_65_104 = $65,000 ≤ Income < $104,000; I_104_156 = $104,000 ≤ Income < $156,000; and I156p = Income ≥ $156,000) 59

Figure 3-20 Biplot of the first and third components for number of bedrooms (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms), floor area for different dwelling types (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two-storey apartments; THREE_S_A = area for three-storey apartments; and FOUR_P_S_A = area for four- or more storey apartments), and building age classes (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011) 59

Figure 3-21 Biplot of the fourth and fifth components for number of bedrooms (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms), floor area for different dwelling types (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two-storey apartments; THREE_S_A = area for three-storey apartments; and FOUR_P_S_A = area for four- or more storey apartments), and building age classes (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011) 60

Figure 4-1 Histogram of the numbers of AC switching on and switching off actions versus ambient temperature (top) and indoor temperature (bottom) 63

Figure 4-2 Histogram of the numbers of AC switching on and switching off actions versus time of the day..... 63

Figure 4-3 Probability of air-conditioner (A/C) 'switch on' action and 'switch off' actions as a function of apparent ambient temperature (left) and apparent indoor temperature (right) for space heating..... 64

Figure 4-4 Probability of air-considitoner 'switch on' action (left) and 'switch off' action (right) as a function of time of day separated by apparent ambient temperature (T_a)..... 66

Figure 4-5 Probability function for AC switch off behaviour as a function of duration switched on..... 66

Figure 4-6 Probability of air-conditioner (A/C) 'switch on' action and 'switch off' actions as a function of apparent ambient temperature (left) and apparent indoor temperature (right) f or space *cooling*. 67

Figure 5-1 Modelling TB for house #10 (-0.9 to -0.7 °C). 70

Figure 5-2 (a) Scree plot showing information captured on the first seven components, (b) maximised averaged silhouette width determining two clusters..... 71

Figure 5-3 Sensitivity of the representative thermal models of the clusters to the input variables 71

Figure 5-4 Buildings scores as split into two clusters..... 72

Figure 5-5 (a) 8 building clusters of homes' scores based on compression of 12 building characteristics, (b) silhouette widths for the 8 building clusters, (c) silhouette widths for h=8 thermal model clusters..... 73

Figure 5-6 Building representative types matched to general thermal models. Rows in (a) correspond to rows in (b). See Appendix B for a description of each building parameter..... 74

Figure 7-1 Comparison of relative scaled discomfort sensitivity versus normalised percentage discomfort for four climate zones. Clockwise from top left: Darwin, Brisbane, Hobart and Melbourne..... 78

Figure 7-2 Comparison of relative scaled energy sensitivity versus normalised energy usage for four climate zones. Clockwise from top left: Darwin, Brisbane, Hobart and Melbourne. 78

Figure 7-3 Sensitivity of discomfort to variations in occupant AC usage behaviour by building thermal parameters for Brisbane (subtropical climate) for buildings with the lowest 30% of energy consumption. Lower numbers indicate less sensitivity. 79

List of Tables

Table 2-1 Summary of reverse-cycle air-conditioner rated performance parameters used..... 8

Table 2-2 Summary of building thermal parameter values considered 11

Table 2-3 Summary of mean parameter effects..... 16

Table 3-1 Variables found to significantly explain variation in average energy consumption using the linear mixed effects analysis for the monthly energy consumption data in South-East Queensland..... 29

Table 3-2 Summary of principal component analyses of the monthly South-East Queensland data 33

Table 3-3 Variables found to significantly explain variation in average energy consumption using the linear mixed effects analysis for the Australia wide annual energy consumption data 40

Table 3-4 Summary of principal component analyses of the annual Australia wide data..... 44

Table 3-5 Summary of effect of different variables for the Australia-wide and South-East Queensland data sets..... 57

Table 4-1 Coefficients for polynomial fits of air-conditioner usage probability according to Eq. (1)..... 65

Table 5-1 Jaccard similarity for 2 clusters (left), 3 clusters (middle), and 4 clusters (right). Columns are built-form clusters. Rows are thermal response clusters. Highlighting shows the strongest correspondences between clusters..... 72

Table 5-2 Jaccard similarity showing one-to-one matching of 8 clusters. Columns are built-form clusters. Rows are thermal response clusters. Highlighting shows the strongest correspondences between clusters..... 73

Table 5-3 Matching predictors with characteristics. See Appendix B for a description of each building parameter..... 75

Executive summary

The aim of Work Stream 4.4 is to gain a better understanding of the factors that influence energy consumption and the use of residential air-conditioning in Australia. In Year 1 of the project, we conducted a literature review, performed an in-depth modelling and statistical analysis of space cooling in particular, and conducted and analysed the results from a survey of air-conditioner usage. In the second year of work, which is described in this report, we continued our analysis of residential air-conditioner usage, specifically by investigating;

i) space heating behaviour;

ii) the trade-off between designing residential buildings to minimise annual energy use for air conditioning (cooling and heating) versus maximising comfort when the buildings are unconditioned;

iii) the comfort and energy use sensitivity of different buildings to variations in occupant behaviour;

iv) the statistical effect of building age (among other parameters) on energy use at monthly resolution across South-East Queensland and annual resolution across the NEM;

v) the extent to which a general model can be developed to estimate indoor conditions given limited building information and the outdoor conditions.

This work was performed through modelling and statistical analysis of existing data sets. New data sets that will be available during FY17/18 will provide new opportunities for analysis of air-conditioner usage behaviour, in particular the linked occupant survey and smart-meter data that is currently being collected.

Key findings are summarised below.

- We studied the relative comfort levels of buildings with minimal or no AC use as compared with the relative energy efficiency of the same buildings with comparatively frequent AC use. A key finding was that for heating dominated climates, there is a close correlation between buildings with high energy efficiency and those that are more comfortable with limited AC. However, for cooling dominated climates, buildings with the same energy efficiency rating can have a wide variation in indoor comfort levels when limited AC is used.

  This suggests that building energy efficiency rating measures that factor in the wide range of occupant behaviours, for example by including both measures of energy usage when the building under assessment is conditioned, as well as occupant comfort when the building is unconditioned, would result in a better overall representation of the thermal performance. This is particularly relevant for understanding the impacts of building codes and standards on occupants across the socio-economic spectrum.

- Our statistical analysis of aggregate South-East Queensland and Australia wide energy data revealed that higher energy users tended to be those in the higher income brackets, those

owning their own property and those in large separate houses. Interestingly, for both data sets the buildings built prior to 1952 appear to have used the least amount of energy, perhaps reflecting the fact that these tend to be smaller than more modern buildings and that they may not have central heating and/or cooling. For the SEQ data, houses built between 2007 and 2011 were associated with less energy use than those built between 2002 and 2006. For the national data set, houses built between 2007 and 2011 were associated with less energy than those built between 1987 and 2001. For SEQ data, houses with fibre cement walls appear to have used less energy than those with timber, cavity or solid masonry walls, while for the national data, the opposite was observed. For the Australia-wide data, buildings with timber and veneer masonry walls appeared to use more energy than those with cavity and solid masonry walls.

One important consideration for the above analysis is that the data analysed is highly aggregated both in space and time. Hence there is the potential for many confounding variables. It would be preferable to have individual household energy data along with information such as building date and the code it was built under, materials, climate data and socio-economic variables for the household. This would increase the likelihood of identifying any differences in energy usage due to different building codes, while controlling for other variables such as local temperature and socio-economics.

- We used energy data from 140 households to assess how the indoor and outdoor conditions and the time of day influenced householders' switching on and off space heating. (In Year 1 of the project a similar analysis was performed for space cooling). In contrast to cooling behaviour, where there was a strong relationship between the outside air temperature and the actions of switching on and off AC, for space heating the effect of temperature was markedly less and time of day was comparatively more important.

  For space cooling, the probability of switching on AC was relatively constant across the day when the effect of ambient temperature was discounted. For the heating behaviour, there were two peaks: one in the early morning, and one in the evening, even before factoring in the cooler ambient temperatures during these periods. For switch-off behaviour, space heating also showed clear peaks in probability at around 9am and between 10pm and 2am, while for space cooling two peaks were also present, but were much more spread out.

  Analysis of the duration of AC running time for heating revealed that households tended to use heating for much shorter durations on average than for cooling. For example, after 5 hours, 20% of air-conditioners were still on when cooling, but only approximately 7% of air-conditioners remained on when heating.

  Overall, these results show that the ways in which occupants operate cooling and heating are different, and hence require different modelling approaches. These results can be used to build behaviour models for use in simulation tools. However, an important limitation of this analysis is that we have used the aggregate behaviour of all households in the study to calculate the event probabilities and run-time values; to identify and group households into

distinct behaviour classifications would require a larger and more diverse sample of households and climates.

- Finally, we also developed a set of generic predictive models to estimate building indoor temperatures given ambient temperature, AC on/off status and solar radiation. These models are expected to be useful in the common situation where the latter are known but the indoor temperature is required as an input for an occupant behaviour model. The models are based on a classification of the building to a high-level grouping using parameters that are likely to be known or able to be estimated, such as the conditioned floor area and star rating. However, an important limitation of this approach was the difficulty of uniquely matching the thermal response to the building characteristics. Given the relatively small number of buildings used for this assessment, further data is needed to be more confident in the applicability of the groupings.

Selected data sets from this work are available for download from the EUDM web platform.

1 Introduction

As outlined in the Year 1 Milestone Report for EUDM Work Stream 4.4, the aim of this work stream is:

> *'…to gain a better understanding of the factors that influence the energy consumption and use of residential AC for cooling. In particular, we aim to consider the extent to which different weather, built environment and demographic variables can be used to predict an occupant's or household's decision to use AC, and the amount of energy consumed when doing so. '*

In Year 2 of the project, this work has been expanded and extended. In particular, we investigated;

i) space heating behaviour

ii) the trade-off between designing residential buildings to minimise annual energy use for air conditioning (cooling and heating) versus maximising comfort when the buildings are not air-conditioned

iii) the comfort and energy use sensitivity of different buildings to variations in occupant behaviour

iv) the statistical effect of building age (among other parameters) on energy use at monthly resolution across South-East Queensland and annual resolution across the NEM

v) the extent to which a general approximate model can be developed to relate outdoor and indoor conditions given limited building information.

This analysis was performed using existing data sets and modelling. We anticipate that new data sets will be available during the final year of the project, particularly through the national surveying program.

Underlying motivations for this work were:

i) space conditioning is a substantial component of total consumption (accounting for 40% of household energy usage (Department of Environment, Water, Heritage and the Arts, 2008)), and space cooling is the key driver of peak energy consumption for most of Australia. Therefore, understanding and predicting current and future air-conditioning (AC) energy consumption is fundamentally important for designing and managing a stable electricity network.

ii) real energy costs are rising, as are climate extremes and climate variability. There is a risk that these factors could lead to significant additional thermal and/or financial stresses, particularly on some consumers. Understanding the interaction between space conditioning, building performance and occupant behaviour is important to help minimise these effects.

iii) the energy market is undergoing a period of significant change. Assessing the potential impact of changes such as new tariff structures, demand–response regulations, incentive schemes or appliance energy-efficiency requirements requires careful consideration of their impact across different households. Models of behaviour, in particular air-conditioning behaviour, are an important tool to help assess this impact.

iv) the National Construction Code specifies design requirements for new buildings. The energy-efficiency specifications of the code make assumptions about how households use space heating and cooling in order to estimate energy consumption of proposed designs. The assumed air-conditioner usage model thus has a direct influence on the building construction industry, on the design of new products, and also on how comfortable future occupants of the building will be. Ensuring that the assumptions used in the code lead to optimal outcomes in as many instances as possible is thus fundamentally important.

For the EUDM program in particular, this work also provides part of the background to inform end-use energy monitoring and data-gathering efforts. It does this by helping to improve our understanding of which factors are most important to record when assessing individual household (air-conditioner) energy consumption. This is especially important given the typically high cost of gathering fine-grained end-user data.

The remainder of the report is divided into six parts. Modelling analysis of the air-conditioning performance of the building shell for a range of Australian climates is described in Section 2. The statistical analysis of monthly whole-of-house energy use data at the postcode aggregate level for South-East Queensland and annual analyses for the wider NEM is described in Section 3. Characterisation of space heating behaviour using data from the Residential Building Energy Efficiency Study (RBEES), and development of a predictive heating behaviour model is described in Section 4. This follows on to Section 5, where a general model for estimating indoor temperature from outdoor temperature using only high-level building information is proposed and tested. Lastly, the key findings from the Year 2 work are consolidated and the proposed next steps for Year 3 are outlined in Sections 6 and 7.

2 Overall thermal performance of the building shell: modelling analysis

2.1 Introduction

In Australia, the National Construction Code (Australian Building Codes Board, 2016) specifies minimum requirements for the design and construction of new buildings and significant renovations. Energy-efficiency requirements form part of this specification. For detached residential buildings (known as Class 1 buildings) there are two basic methods of meeting the code; i) the Deemed to Satisfy (DTS) method, in which the building, in addition to the inclusion of various energy efficient features, must obtain a minimum 5 or 6 star energy rating (depending on the jurisdiction) as assessed using a building simulation software protocol known as NatHERS, or ii) complying with the specific building construction performance specifications (for example insulation R-values and window areas) outlined in the code. It is estimated that a majority of new residential buildings use the NatHERS building simulation star rating approach (Australian Government, n.d.), although there are some doubts as to the exact proportion.

In the DTS approach, the energy star rating of a building is calculated based on the theoretical annual cooling and heating requirements as determined by a computer building simulation. This simulation uses representative climate data, assumed occupancy patterns, separate space conditioning (heating and cooling) and ventilation control strategies for living and bedroom zones. When natural ventilation is insufficient, this control strategy applies continuous conditioning of at least one occupied zone in the building whenever indoor temperatures are above 22.5°C to 28°C (depending on the climate zone) or below 20°C (living spaces), 18°C (bedroom spaces in the evening) or 15°C (bedroom spaces overnight).

As outlined in the NatHERS Protocol (NatHERS National Administrator, 2012), this particular control strategy is not designed to reflect actual operation in any one household. Indeed, no single control strategy could achieve this. Instead, it is designed to give a uniform conservative estimate of the upper limit of energy use for the given building. However, as identified in the Year 1 Milestone Report for this project (Goldworthy & Seongwon, 2016), a number of studies in Australia (Soebarto & Bennetts, 2014) (Strategic Insights, 2014) (Goldsworthy, Towards a residential air-conditioner usage behaviour model for Australia) (Daniel, Soebarto, & Williamson, 2015) (Law, Osman, Ambrose, James, & White, 2014) have identified that many households do not use space conditioning with a regularity anything approaching this upper limit. For example, a recent survey (Strategic Insights, 2014) of households in Queensland found that 34% use air-conditioning (cooling) 'rarely or only on extremely hot (i.e. >34C°) days', while an online national poll gave this percentage as 18% (Goldsworthy, Towards a residential air-conditioner usage behaviour model for Australia). The reasons why usage for many households is less than continuous are many, and include, among others, the financial cost of operation in particular, but also factors such as a preference for fresh outside air.

The fact that actual air-conditioner usage may be significantly different from the assumed usage used for building energy-efficiency assessment has two key potential implications for the regulatory approach;

i) the regulation may be overly restrictive for certain building-occupant combinations, and/or

ii) the regulation may lead to perverse energy, or comfort outcomes if it drives building construction methods towards approaches that are either not energy efficient, or that lead to more uncomfortable buildings, factoring in how households actually use (or do not use AC).

This issue has been previously highlighted by both Williamson et al. (Williamson, Soebarto, & Radford, Comfort and energy use in five Australian award-winning houses: regulated, measured and perceived, 2010) and Kordjamshidi and King (Kordjamshidi & King, 2009). The latter compared house thermal designs in both free-running (unconditioned) and air-conditioned operation in the temperate climates of Sydney and Canberra. They showed that the best thermal design to optimise comfort in a free-running household does not necessarily coincide with the best design to minimise energy use. Based on these findings, they proposed a new rating method to combine both comfort and energy use in the assessment of the building star rating. To date, however, this method has not been adopted.

A review of the literature reveals that although a number of studies (for example (Williamson, Designing houses for the Australian climate: the early research, 2013) (Smith, Aguilar, Gentle, & Chen, 2012) (Ambrose, James, Law, & Osman, 2013) (Karimpour, Belusko, Xing, Boland, & Bruno, 2015) (McLeod & Fay, 2010) (Whitsed & Davidson, 2013) and the NatHERS approach) have investigated the energy use implications from different building construction components, and some (see for example (Miller, Buys, & Bell, 2012)) have considered the comfort implications in unconditioned (passive) buildings. Very few, with the exception of the work noted above, have compared the trade-off between these two factors, and none has considered the effect across all climates of Australia.

Thus, Section 2.2 described results from a comparison of the particular building construction details that lead to occupant comfort (and discomfort) in a building that may be unconditioned for a large portion, or even the majority of the occupied hours, with the particular building construction details that lead to minimal energy use under almost continuous air-conditioning usage in climates across Australia. In addition, we also estimate the relative magnitude of the influence of the different building construction parameters to gauge which parameters are more important.

Following on from this, in Section 2.3 an alternative analysis is performed in which we consider how sensitive different buildings are to variations in occupant air-conditioner usage behaviour. That is, given that there is a wide variation in usage behaviours, we analyse which buildings lead to greater variations in energy use and average comfort levels. The importance of performing this style of analysis, given the large impact of behaviour variability, has been identified recently by Hoes et al. (Hoes, Trcka, & Hensen, 2011).

Both of these analyses are made using results from computer simulations based on a model similar to that used in the NatHERS estimation. This model includes a simple building thermal model, air-conditioner performance models for cooling and heating, as well as corresponding control strategies, and employs the same climate data used in the NatHERS calculations. However, we do not model occupant controlled natural ventilation, but rather only natural infiltration.

The simulation model and analysis approach are described briefly below in Section 2.2.1 and Section 2.3.1, noting that there are some differences in the approaches used for the two analyses. Discussion of the results and limitations of both approaches is given in Section 2.4 and suggested future work is discussed in Section 2.5.

2.2 Comfort versus energy use trade-off

2.2.1 Analysis method

For this study, the simulation model as described in (Goldsworthy, Building thermal design for solar photovoltaic air-conditioning in Australian climates, 2017) was used with selected modifications. This model consists of i) a building model based on a uniform temperature and humidity air node coupled to transfer function models of external walls, ceiling and roof cavity; ii) a vapour-compression air conditioner cooling model, and iii) weather and radiation data processing models. The modifications included i) implementation of a reverse-cycle (both heating and cooling) vapour-compression air-conditioner model, and ii) alternate control strategies for cooling and heating.

Reverse-cycle air conditioning was modelled using the standardised set of performance curves for residential single speed reverse-cycle air-conditioners as documented by Cutler et. al. (Cutler, Winkler, Kruis, Chistensen, & Brandemuehl, 2013). These performance curves relate the total gross cooling capacity of the cooling coil Q_{gross} and the electrical power consumption of the outdoor unit P_{gross} at off-rated conditions to the values of these parameters at the rated condition (Q^*_{gross}, P^*_{gross}). These curves are bi-quadratic functions of the indoor wet bulb and ambient dry bulb temperatures (in cooling) and indoor and ambient dry bulb temperatures (in heating). The net (*i.e.* overall) cooling capacity Q_{net} and net electrical power consumption P_{net} are obtained from the gross values by accounting for the power consumption of the air-supply fan.

In cooling mode, the sensible heat ratio at off-rated conditions was calculated using the bypass factor/apparent dew point approach as used in the EnergyPlus building simulation tool (U.S Department of Energy, 2016). Part load performance was not included, since this would make it necessary to size the air-conditioner individually for every building and climate combination.

The rated values of the net cooling and heat capacity, net coefficient of performance, net sensible heat ratio, evaporator air mass flow rate and fan electrical power consumption used here are given in Table 2-1.

Table 2-1 Summary of reverse-cycle air-conditioner rated performance parameters used

Quantity	Value at rated condition
Total net cooling capacity	10kW
Sensible heat ratio (net)	0.72
Net COP (cooling)	3.3
Total net heating capacity	10.5kW
Net COP (heating)	3.5
Rated evaporator flow rate	200 kg/hr/kW total rated net capacity
Fan power consumption	40 W/kW total gross capacity

For this analysis, we sought to use a conditioning control strategy that was slightly more reflective of how air conditioning is used in residential homes than pure on/off set-point controls, which is more realistic for commercial buildings. However, the chosen control strategy should still result in a high frequency of air-conditioner usage compared with most households. The control thus implemented, summarised in Figure 2-1, consisted of three temperatures levels for each of four different conditions; day cooling, night cooling, day heating and night heating. Here we described the control with reference to the day-cooling operation case; operation for the other cases follows in an analogous fashion from this.

For day cooling, the highest temperature level (28°C in the figure), referred to here as the 'trigger temperature' T_t, was compared to the indoor apparent temperature $\tilde{T}_b$. Apparent temperature combines both temperature and humidity and gives a better indicator of thermal comfort. Here the method of Steadman (Steadman, 1984) was used to calculate apparent temperature. If $\tilde{T}_b > T_t$, then cooling was activated and the reverse-cycle air-conditioner cycled on/off to maintain the building temperature between the upper and lower (dry bulb) temperature set points T_u and T_l (25°C and 22°C in the figure). The signal that indicates that cooling has been 'tripped' was reset at 6am each day. This approach models, in effect, a 'resetting' of the occupant memory of air-conditioner use. (Note, however, that the AC could still turn off prior to this if the indoor temperature was within the set points.) As noted above, a similar scheme was applied to heating, except that the trigger temperature is the lowest temperature for each case (*i.e.* 15°C for day heating).

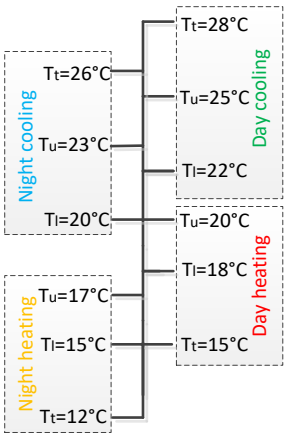

Figure 2-1 Temperature levels for air-conditioning control. Note apparent temperature used for trigger temperatures

To measure occupant discomfort in the *unconditioned building*, we assumed that indoor apparent temperatures above the corresponding upper temperature threshold T_u for the given time of day for cooling, or below the corresponding lower temperature threshold T_l for the given time of day for heating, were 'uncomfortable'. Cooling and heating degree of discomfort factors dd_c and dd_h were thus calculated according to;

$$dd_c = \sum max(\tilde{T}_b - T_u, 0)\Delta t/24 \quad \begin{matrix} T_u = 25°C\ 6am\ to\ 10pm \\ T_u = 23°C\ 10pm\ to\ 6am \end{matrix} \qquad (1)$$

$$dd_h = \sum max(T_l - \tilde{T}_b, 0)\Delta t/24 \quad \begin{matrix} T_u = 18°C\ 6am\ to\ 10pm \\ T_u = 15°C\ 10pm\ to\ 6am \end{matrix}$$

$$dd = dd_c + dd_h$$

with the overall discomfort factor represented by *dd*. In Eq. (1), Δt is the time interval in hours.

To measure air-conditioner energy (electricity) use (en) in the *conditioned building*, we simply summed the air-conditioner power consumption P_{ac} (kW) over each time step over the year for both heating and cooling.

$$en = \sum P_{ac}\Delta t \tag{2}$$

In several results, heating degree days (HDD) and cooling degree days (CDD) are used as indicators of the amount of heating and cooling required respectively in a particular climate. Here we define HDD and CDD according to Eq. (1), where the indoor apparent temperature $\tilde{T}_b$ is replaced with the ambient apparent temperature $\tilde{T}_a$. That is, HDD is the number of degree days where the apparent ambient temperature is below 15°C or 18°C (depending on the time of day). CDD is defined analogously. The difference between HDD and CDD (i.e. HDD – CDD) gives a measure of whether the climate is heating dominated or cooling dominated. Contours of this difference across Australia, given in Figure 2-2, show that by this measure, all of the capitals cities with the exception of Darwin and Brisbane are in heating dominated climates.

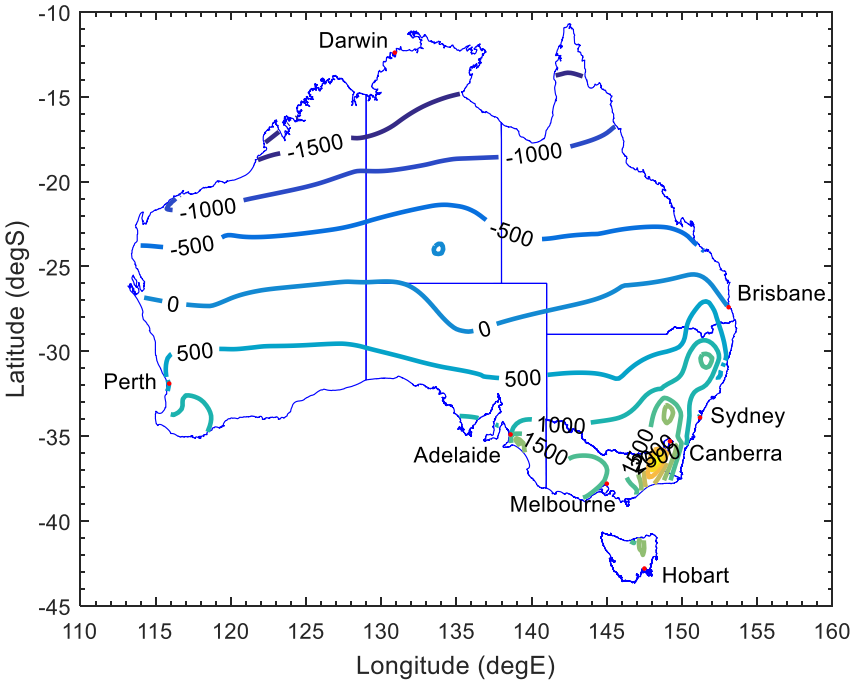

Figure 2-2 Contours indicating net heating (positive) and net cooling (negative) climates based on annual subtraction of apparent temperature cooling degree days from heating degree days (HDD-CDD); capital city locations marked

2.2.2 Parameter selection

Since the aim is to compare the *relative* influence of different building thermal parameter variations on building energy use and comfort across climate zones, it is necessary to frame these parameter variations in the context of 'realistic' parameter ranges. Here we use the same parameter ranges as described in (Goldsworthy, Building thermal design for solar photovoltaic air-conditioning in Australian climates, 2017) and shown in Table 2-2.

With the exception of orientation and wall construction type, each parameter has two values; a notionally 'above average' value and a notionally 'below average' value. These values are not intended to represent the extrema, but rather plausible ranges that cover the majority of Australian detached residential buildings. Ideally these parameter values would be chosen with reference to detailed statistics on actual residential building constructions (either existing or new constructions, depending on the aim). As outlined in (Goldsworthy, Building thermal design for solar photovoltaic air-conditioning in Australian climates, 2017), these parameter values were selected with reference to the published literature, where possible.

When comparing the influence of different parameters, we use Cohen's mean scaled difference (Cohen, 1988) as an indicator of the relative magnitude of the influence of each parameter value (ignoring second and higher order correlations between parameters) on energy use for the conditioned building and discomfort for the unconditioned building.

Table 2-2 Summary of building thermal parameter values considered

Variable	Parameter values

Internal zone thermal capacitance (ratio of air thermal capacitance)	5x, 15x
Infiltration rate	0.5, 2.0 a/hr
Internal load profile (morning and evening peaks)	Average daily load: Low: 4.7/1.95/0.34 W/m² (sensible/latent/radiative) High: 9.4/3.9/0.67 W/m² (sensible/latent/radiative)
Building aspect ratio	1:1, 2:1
Window orientation (degrees from Nth)	0, 45, 90, 135, 180, 225, 270, 315°
Ceiling insulation	None, R3
Roof solar absorptivity	0.3, 0.9
Wall insulation	None, R2
Wall construction (thermal mass)	Light (weatherboard), Medium (brick veneer), Heavy (double brick)
Wall solar absorptivity	0.3, 0.75
Window to floor area ratio	10%, 20%
Window type	Single glazed, double glazed with shading from external roof eave.

Approximately 2000 different building models were composed from a random selection of the parameter values in Table 2-2. Two annual simulations were then run for each of these buildings; one for a conditioned (both cooling and heating) building and one for the same building without any AC. These simulations were run for each of the 69 NatHERS climate zones, resulting in a total of approximately 276,000 simulations. Cohen's mean scaled difference parameter was then calculated for energy use for the 2000 air-conditioned buildings and for discomfort for the same 2000 building without AC.

2.2.3 Results

The scaled mean influence of each building parameter value (excluding area ratio and window type, which had only small *mean* influence, and window orientation which is discussed separately) is summarised across all climate zones in Figure 2-3. In this figure, the horizontal axis is the mean scale parameter for energy and the vertical axis is the mean scale parameter for discomfort.

Since it is preferable to achieve both lower energy use and lower discomfort, a positive mean scale value for a particular parameter indicates that the parameter (the value to the right in Table 2-2) led to better performance (i.e. lower energy use or lower discomfort). Hence, values in the top right quadrant and bottom left quadrants correspond to the ability for that parameter value to be optimised to achieve *both* low mean energy consumption and low mean discomfort. On the other hand, values in the top left or bottom right quadrants correspond to the case where the parameter value that resulted in lower mean energy use gave *higher* mean discomfort (or vice versa).

In this figure, individual points show results for each climate zone and are also colour-coded according to the value of HDD - CDD for the particular climate zone. Red colouring corresponds to heating dominated climates and blue colouring corresponds to cooling dominated climates.

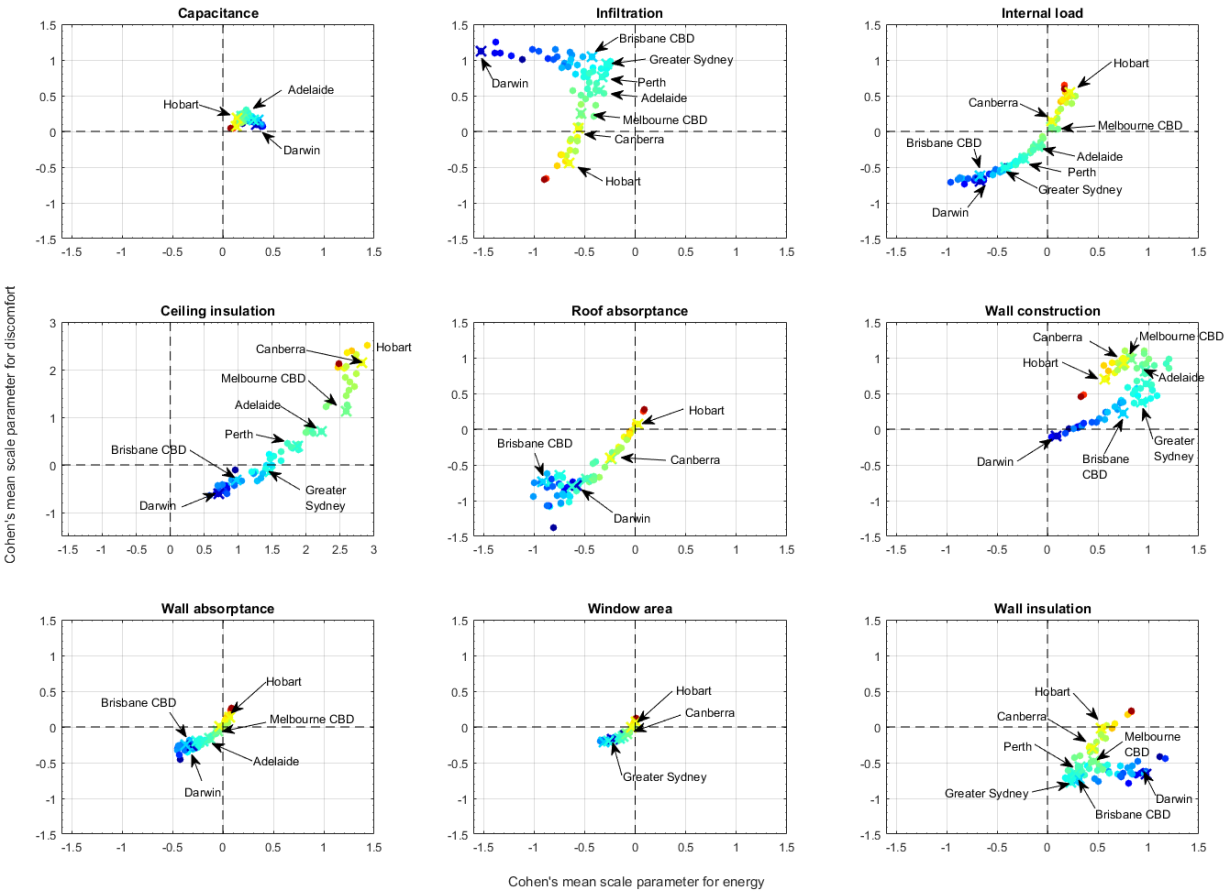

Figure 2-3 Annual mean scale parameter for discomfort and energy for selected building parameters plotted for all climate zones across Australia. Colour indicates relative heating (red) or cooling dominated climates (dark blue). Selected capital city climate zones are indicated as labelled. Note: different axis scale for ceiling insulation plot.

The following trends are apparent;

- higher *natural infiltration* leads to higher mean annual energy use for all climates, but especially for those with greater temperature extremes (i.e. either heating or cooling dominated climates). The climates with lower average energy penalties over the year generally correspond to those that have significant coastal breezes that moderate temperature extremes (for example, Coffs Harbour MO, Sydney RO, Mascot AMO, Mandurah, Swanbourne, Esperance). In the case of comfort, higher infiltration is more beneficial, on average, for climates that are more cooling dominated; this corresponds to the majority of climate zones across Australia.
- for all climates zones, higher *internal capacitance* in the zone results, on average, in both less energy consumption and more comfort, though the magnitude of the effect is relatively small overall.
- for all climates zones, *wall and ceiling insulation* results in lower energy consumption. However, except for five heating dominated climates, wall insulation results in lower levels of comfort in the unconditioned building. For ceiling insulation this pattern is similar, although the number of climates where insulation improves comfort in the unconditioned

building is much greater (approximately 37 zones), with mean annual comfort higher with the insulated ceiling buildings for the more heating dominated climates.

- for *internal loads,* the impact on energy use and comfort is similar; for approximately 17 heating dominated climates, higher internal loads lower energy use and improve comfort, while the reverse occurs for cooling dominated climates.
- for *wall and roof solar absorptivity*, the effects on energy consumption and comfort are also similar as for the case of the internal loads. However, far fewer heating dominated climates (only three in fact) show a net benefit for higher *absorptivity* values. Roof absorptivity has a greater effect than wall absorptivity.
- for *window area,* the effects on energy consumption and comfort are again similar to those for absorptivity and internal loads. However, for this parameter almost all zones have lower mean energy consumption and higher comfort for smaller window areas. The mean effect of window area is also smaller.
- for *wall construction* (in particular wall thermal mass), higher thermal mass leads to lower energy consumption and higher comfort for most climates. These effects are most significant for more moderate climates. For tropical climates (e.g. Darwin) higher thermal mass buildings had a slightly lower mean comfort. Further inspection of the three building types individually revealed minimal difference between the weatherboard and brick veneer constructions, with the majority of the effect resulting from the difference between double brick construction and the other two types.

For window orientation, the mean scale parameter has been calculated for pairs of orientations with opposing North/South or East/West components. For example, North and South orientations are compared with each other, as are NE/SE and NW/SW. The magnitude of the mean scale parameter is directly comparable between all cases. The chosen sign convention was for favourable Southerly and Easterly directions to correspond with positive mean scaled parameter values. That is, if SE is preferable to NE then the pair comparison will have a positive mean scale parameter. Likewise if NE is preferable to NW then the mean scale parameter will be positive for that pair.

Figure 2-4 left shows that for almost all climates the Easterly aspect is preferable to Westerly for both comfort and energy use (i.e. values are in the top right quadrant). However, there are a number of cooling dominated climates where SW is slightly better than SE for energy use, and two alpine zones where NW is slightly better than NE for both energy use and comfort. However, these effects are relatively small.

The right plot in Figure 2-4 tells a slightly more complex story. Firstly, the magnitude of the scale parameters is generally less than for the E-W comparisons, indicating that the latter has more mean effect. For cooling dominated zones SW is preferable to NW for both comfort and energy use, while for heating dominated zones the situation is reversed. For almost all zones NE is preferable to SE for both comfort and energy use, particularly for cooling dominated zones.

For most zones, Southerly is preferable to Northerly for energy use and, for the cooling dominated zones, for comfort as well. However, the mean difference between North and South orientations is generally relatively small; most likely assisted by the presence of eaves over the window which provide summer shading, particularly for the North window orientation, while still allowing winter heat gain.

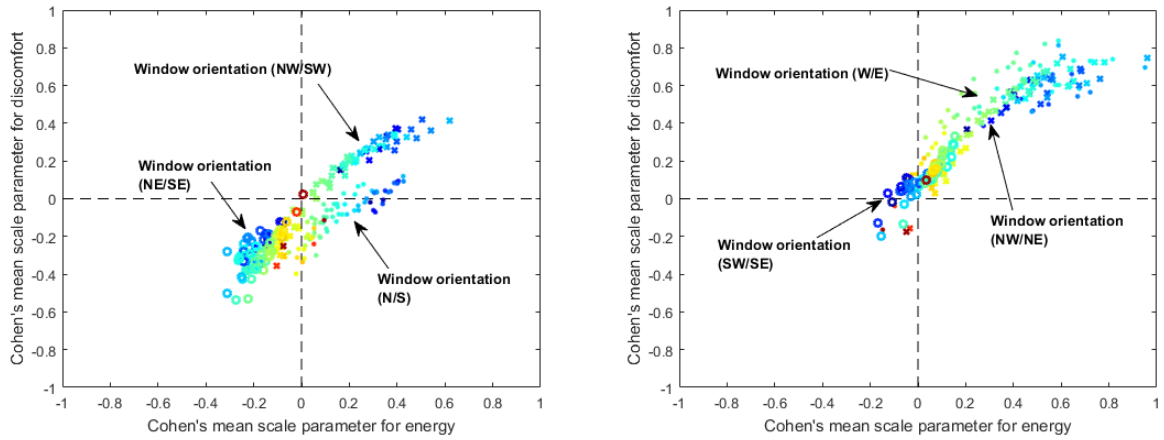

Figure 2-4 Annual mean scale parameters for discomfort and energy for different orientation component pairs plotted for all climate zones across Australia. The left plot shows a component comparison for orientations with N/S components. The right plot shows a component comparison for orientations with E/W components. Colour indicates relative heating (red) or cooling dominated climates (dark blue).

The mean effects of each parameter are summarised in Table 2-3. Once again, it is important to keep in mind that the mean effect of a particular parameter value shows its average effect across all other parameter values. This does not account for parameter effects that might be particularly large only when combined with one particular value of another parameter. For example, window type has a greater effect for certain window orientations and for larger window sizes. However, the mean effect size only accounts for the average effect of window type across all orientations and window sizes.

Table 2-3 Summary of mean parameter effects

Parameter	Qualitative mean effect size across all climates	Summary of annual mean effect
Ceiling insulation	Large	Ceiling insulation lowers energy use for all climates but improves comfort only for heating dominated climates. The climate of Sydney represents an approximate division between these two groupings.
Natural infiltration	Large	Higher infiltration increases energy use on average for all climates but improves comfort in almost all climates (exceptions being Canberra and Hobart). Comfort improvement is greatest for cooling dominated climates while the energy penalty is least for climates with moderating coastal breezes.
Window orientation	Moderate-Large	In general, East and NE orientations are better than West and NW orientations, particularly for cooling dominated climates, for both comfort and energy use. In general S and SW are preferable to N and NW, again, particularly for cooling dominated climates. However, NE is preferable to SE.
Wall construction	Moderate-Large	Double brick wall construction lowers energy use and improves comfort for non-tropical climates.
Roof absorptivity	Moderate-large	Darker coloured roofs lower energy use and improve comfort for heating dominated climates and have the opposite effect for cooling dominated climates. The climate of Hobart represents the approximate division between these two groupings.
Wall insulation	Moderate-large	Wall insulation lowers energy use, though the effect is small for relatively mild climates (e.g. Sydney). Wall insulation reduces comfort for all climates excluding alpine and sub-alpine.
Internal load	Moderate	Higher internal load lowers energy use and improves comfort for heating dominated climates and has the opposite effect for cooling dominated climates. The climate of Melbourne represents the approximate division between these two groupings.
Wall absorptivity	Small	Darker coloured external walls lower energy use and improve comfort for heating dominated climates and have the opposite effect for cooling dominated climates. The climate of Canberra represents the approximate division between these two groupings.
Window area	Small	Larger window area increases energy use and reduces comfort for all climates with the exception of alpine locations.
Internal capacitance	Small	Higher capacitance lowers energy use and improves comfort. The effect is relatively small.
Window type	Small	Double glazed (low solar gain) windows lower energy use for all climates. However, comfort is reduced for heating dominated climates. The climate of Canberra represents the approximate division between these two groupings.
Aspect ratio	Negligible	No statistically significant effect.

2.3 Sensitivity of buildings to occupant behaviours

2.3.1 Analysis method

To analyse the energy use and comfort sensitivity of different buildings to variations in occupant air-conditioning usage behaviour, the same base building and air-conditioner simulation tool as described above was used, but with a different air-conditioner control approach and method of comparing performance. These changes are described below.

2.3.1.1 Occupant air-conditioner control method

We required a method to simulate different air-conditioning usage behaviours. Ideally the method would provide a continuous variation of behaviours, ranging from frequent to infrequent usage. These should be computationally efficient, given the large number of building and climate zone combinations to be considered for each range of behaviours. This rules out the stochastic based models (such as is discussed in Section 3), since these require multiple simulations per case to determine output statistics.

Thus we employed a method based on the ASHRAE percentage persons dissatisfied (PPD) comfort metric (ASHRAE, 2004). This metric, which is a function of six variables (the ambient room temperature, radiant temperature, room relative humidity, room air velocity and the activity and clothing of the occupant), gives an estimate of the percentage of persons who would be uncomfortable (either too hot or too cold). Typically commercial building air-conditioning systems are designed such that fewer than 10% of persons are dissatisfied at any one time (i.e. PPD<10%).

We used this metric to control the air-conditioner usage under the assumption that if PPD is equal to a certain threshold PPD_t, then the percentage of people who are satisfied with the conditions is 100 – PPD_t, and this corresponds to the percentage of people who *would not turn on air conditioning* at those conditions. With this assumption, the value of PPD_t can then be varied between 0 and 100% to obtain a continuous variation in air-conditioner usage behaviour ranging from near continuous air-conditioner usage to no usage.

To implement this control strategy, we use a PPD dead-band ΔPPD = 1% and varied the threshold PPD_t from 10% to 90% in increments of 10%. In this model, each simulation then represents 10% of the population of behaviours.

Fixed occupant activity and clothing levels corresponding to light activity in typical indoor clothing (MET=1.1 and CLO=0.75) and a fixed indoor air velocity corresponding to minimal forced air-movement (1.1m/s) were used. Interior surface radiant temperatures output by the building simulation sub-model were used to calculate the mean radiant temperature using an area-weighted approach for input into the PPD calculation.

Since the PPD does not indicate whether occupants are dissatisfied because of over or under temperature, the Predicted Mean Vote (PMV) (also output via the ASHRAE method) was used to determine whether cooling or heating was required.

2.3.1.2 Performance metrics

In the previous analysis discussed in Section 2.2, two separate simulations were performed for each building and climate combination. One simulation was of a completely unconditioned building used

for the comfort calculation, while the other simulation was of the conditioned building used for the energy use calculation.

Here we only perform *one simulation for each combination of building, climate and PPD_t.* Occupant degree of discomfort and air-conditioner energy usage were both calculated for this one simulation. Energy use was calculated using the same method (i.e. Eq. 2) as used previously. However, for discomfort we use the annual average percentage of persons dissatisfied ($\overline{ppd}$) in the building as the metric. That is, although we are performing simulations with air conditioning that is controlled according to the expected comfort level of the occupants (via different PPD thresholds), the discomfort performance metric is calculated in the same way for all cases. The reasoning here is that we are defining the discomfort metric not as the discomfort of the person who controls the air-conditioner for each simulation case (which is by definition zero since the air conditioning is controlled to always turn on to avoid discomfort *for that person*), but as the discomfort level for all possible (current or future) occupants, factoring in that any one set of occupants i) is comprised of multiple individuals with varying comfort tolerances, and ii) may not always turn on air conditioning despite being 'uncomfortable'.

To assess the comfort and energy use sensitivity of each building and climate combination to variations in occupant behaviour, we used the relative standard deviation (RSD) approach of Hoes et al. (Hoes, Trcka, & Hensen, 2011). The RSDs for energy use and discomfort were calculated according to;

$$RSD_{discomfort} = \frac{\sigma(\overline{pdd_i})}{\mu(\overline{pdd_i})} \tag{3}$$

$$RSD_{energy} = \frac{\sigma(en_i)}{\mu(en_i)}$$

$$\|RSD\| = \sqrt{RSD_{discomfort}^2 + RSE_{energy}^2}$$

where the subscript 'i' corresponds to each of the nine simulations with varying PPD_t. The overall sensitivity of the building-climate combination to occupant behaviour is given by $\|RSD\|$, with higher numbers indicating a greater sensitivity.

2.3.2 Results

Key results are described below with reference to Figures 2–5 to 2–7 and (in the Appendix) Figures 8–1 to 8–3.

Figure 2-5 compares the normalised annual mean discomfort on the vertical axis with the normalised annual mean energy use on the horizontal axis. For a given building-climate combination, the mean is taken across the nine simulations with varying air-conditioner control strategy. Thus each data point represents the results for single building and climate. The dots are coloured by the relative sensitivity parameter $\|RSD\|$. The four plots show results for different climate zones. Clockwise from the top left these are; Darwin (tropical), Brisbane (subtropical), Hobart (cool temperate) and Melbourne (mild temperate). Figures 8-1 and 8-2 show the energy and discomfort relative sensitivity parameters RSD_{energy} and $RSD_{discomfort}$ plotted separately against

the normalised annual mean discomfort and normalised annual mean energy use. Figure 2-6 compares the variation of annual mean PPD and annual energy usage for the different AC control cases for two selected buildings using the Darwin climate. These two buildings, labelled 'A' and 'B' have similar mean annual energy use as indicated by the circled dots in Figure 2-5. However, building A has a much lower PPD when averaged across the nine different AC control cases.

From these figures we can observe that;

- building energy use relative sensitivity increases as the building becomes more energy efficient. Building percentage occupant discomfort relative sensitivity decreases as the building becomes more comfortable. That is, buildings that are most comfortable, tend to have relatively less comfort variation as occupant AC usage behaviour changes. However, buildings that are most energy efficient tend to have relatively more energy use variation, as occupant AC usage behaviour changes. In relative terms, building energy use is more sensitive than building percentage occupant discomfort in predicting occupant behaviour variations.
- For strongly heating dominated climates (e.g. Hobart) there is a correlation between energy efficient buildings and buildings with high occupant comfort. That is, buildings with the lowest mean energy consumption also tend to have the highest mean occupant comfort. This phenomenon decreases as the climate becomes progressively more cooling dominated to the extent that, for tropical climates (e.g. Darwin), there can be a large variation in occupant mean comfort for buildings with the same mean energy use.

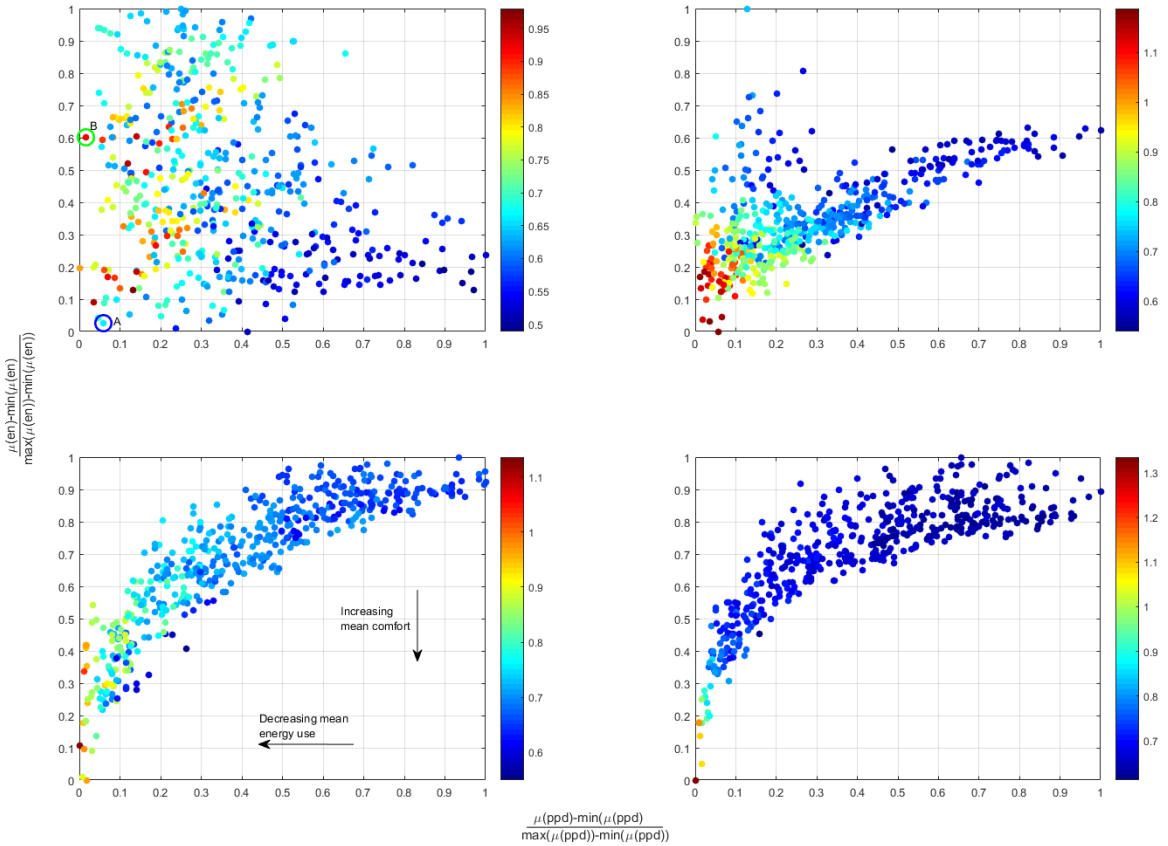

Figure 2-5 Mean discomfort (vertical axis) versus mean energy use (horizontal axis) for each simulated building for four different climates. Clockwise from top left: Darwin, Brisbane, Hobart, Melbourne. The colour of the dots indicates the relative magnitude of the sensitivity of discomfort and energy use to different occupant behaviours.

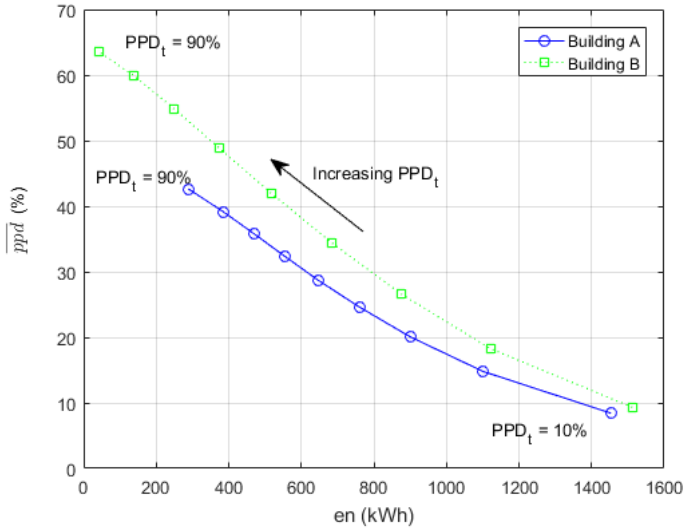

Figure 2-6 Annual average percentage persons dissatisfied (%) versus energy use (kWh) as a function of AC usage behaviour (i.e. varying PPD_t) for two selected buildings in the Darwin climate. Both buildings have approximately the same mean energy use averaged across all values of PPD_t, although building A has a significantly lower mean discomfort.

Figure 2-7 and Figure 7-3 show the comfort sensitivity to variations in occupant AC usage behaviour by building thermal parameters for the buildings with the lowest 30% of energy consumption for

Darwin and Brisbane, respectively. The lowest 30% was chosen since these buildings are of most interest in terms of developing designs that are both energy efficient and comfortable. These two particular locations were chosen, since they have a larger sensitivity to variations in occupant behaviour than the cooler climates.

From these figures we can observe that:

- for buildings in the bottom 30% of energy consumption in Darwin, those with higher natural infiltration, uninsulated walls and ceilings and a light colour roof showed less occupant comfort sensitivity in response to variations in occupant AC usage behaviour. Buildings with less comfort sensitivity generally also had higher mean comfort levels.
- for buildings in the bottom 30% of energy consumption in Brisbane, those with higher natural infiltration, higher thermal mass walls, uninsulated walls and ceilings, and light colour walls and roof showed less occupant comfort sensitivity in response to variations in occupant AC usage behaviour. Buildings with less comfort sensitivity generally also had higher mean comfort levels.

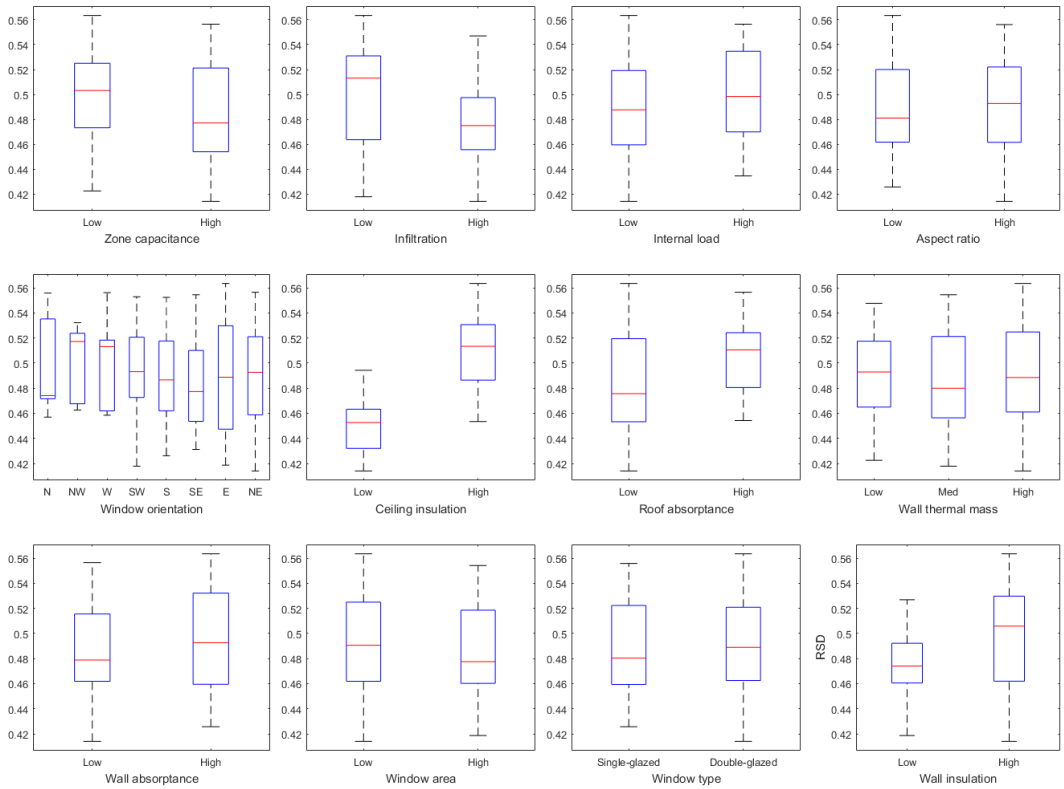

Figure 2-7 Sensitivity of discomfort to variations in occupant AC usage behaviour by building thermal parameters for Darwin (tropical climate) in buildings with the lowest 30% of energy consumption. Lower numbers indicate less sensitivity.

2.4 Discussion & limitations

From the results in Section 2.2 it is clear that there is a trade-off between designing buildings in which occupants feel comfortable without mechanical air conditioning, and buildings that use minimal energy when conditioned. For example, wall insulation only saves energy if occupants use air conditioning regularly and even then, the effect is small in some climates. In contrast, wall

insulation generally *lowers* comfort if the occupants do not use conditioning and there is limited ventilation. Likewise, a well-sealed house is beneficial when air conditioning is used, but is not otherwise for many climates. In the case of ceiling insulation, energy use is generally lowered across all climates, but comfort is only improved for non-cooling dominant climates. On the other hand, for many other parameters the preferred design for a particular climate is clear. For example, light coloured (low solar absorptance) roofs and, to a lesser extent, walls are important for most climates (but not for example Hobart which is heating dominated) for both conditioned and unconditioned buildings.

Although natural infiltration was modelled, the simulations here did not include occupant control of window coverings or natural ventilation (for example opening of windows), as considered in the NatHERS approach. In practice, some occupants are likely to take advantage of these features, *generally* to improve comfort and reduce energy use, while for others concerns over privacy, security or noise are often cited (Soebarto & Bennetts, 2014) as reasons that limit this behaviour. Ideally, simulations with a range of occupant behaviours (including window/blind operation) derived from direct measurements and observations (along the lines of those discussed elsewhere in this report) would be used to assess the sensitivity of different building designs. Unfortunately at present, the data used to develop these models is not representative or extensive enough to develop a full set of different behaviour models. Hence, the simplistic PMV-PPD approach was adopted to model different occupant behaviours. There is also a question of whether such analysis should be conducted using actual (i.e. existing) occupant behaviours, or some hypothetical occupant behaviour that occupants might revert to in the absence of constraints such as energy cost concerns; that is, the *desired* behaviour rather than the actual. There is no easy answer to this question. The fact that cost concerns are likely to remain real for the foreseeable future, and that wholesale behaviour changes are unlikely to occur rapidly, suggests modelling efforts should focus more on actual (current) behaviours.

For a code or standardised approach, the wide range of occupant behaviours and preferences makes the selection of a single 'optimal' solution more difficult. Most buildings will be home to multiple occupants during their lifetime. Regulations that promote building designs that are tolerant to a wider range of behaviours is one approach to solving this problem. Rating measures that allow for, and capture performance across more than one dimension, for example in separate 'free-running' and 'conditioned' operating modes are also important. Technologies, especially those that adapt automatically to different occupant behaviours and conditions, for example windows that adjust solar transmittance, ventilation systems that provide fresh air cooling as required, or conditioning systems that offer a more natural variation in indoor conditions, will also form part of the solution.

The analysis in Section 2.3 shows that buildings in cooling dominated climates are, in general, much more sensitive (in terms of energy use and average comfort) to variations in occupant behaviour. That is, in these climates there is a greater risk that buildings that are deemed energy efficient will lead to poor comfort outcomes for some occupants. The likelihood of the analogous effect occurring for energy (i.e. that buildings that are deemed energy efficient will in fact use lots of energy for some occupants) depends on what usage behaviour was assumed when performing the 'deeming'. If the assumed air-conditioner usage behaviour corresponds to relatively frequent usage, as in the analysis described in Section 2.3, then there is less risk of this occurring.

In this analysis we have considered only a single building zone as opposed to an entire house design, and the relative magnitudes of the parameter effects depend of course on the particular parameter ranges chosen. Use of a single zone was considered necessary to simplify the analysis given the large number of parameters considered, but it also reflects the fact that the majority of detached residential buildings do not have conditioning in all rooms. As a consequence, a whole of house

analysis would be further complicated by the difficultly of defining comfort across both the conditioned and unconditioned zones that might be occupied. This further highlights the fact that building thermal design that is based solely on the energy performance of conditioned zones neglects comfort in the remainder of the house, which, although perhaps is less likely to be occupied during extreme weather conditions, presumably is occupied occasionally.

A final important point is that we have considered cooling and heating via a reverse-cycle air-conditioning system such as a wall mounted, split, or ducted system. In general, the effect of most building thermal parameters is expected to be similar for other forms of cooling and heating, provided that these systems are sized appropriately. However, some variations might result from either different system performance characteristics under varying operating conditions, or the different mechanisms of operation. For example, heating systems that combine both convective and radiative heating (*i.e.* wood combustion and some electric resistance heaters) may yield a slightly different effect on the building parameters due to the radiative heating component, while evaporative cooling systems may yield a different effect due to the higher air-flow and humidity change.

2.5 Summary & next steps

The simulation results described above show that at the most basic level, good building thermal design is essentially the result of the trade-off between two factors; i) whether the climate is heating or cooling dominated, and ii) whether the building operation will be free-running or conditioning dominant.

For heating dominated climates, buildings with higher energy efficiency generally also lead to better comfort even when unconditioned, as compared with buildings with lower energy efficiency. However, for cooling dominated climates this is not necessarily the case. Hence, it is important to consider how occupants will use the building when assessing different designs.

The analysis here used a single zone building model and some basic air-conditioner control strategies to reveal these trade-offs. Noting the particular importance of the occupant behaviour, further work could include:

- implementing a more realistic, detailed set of occupant AC behaviour models (for example drawing on the results discussed in Section 4) and re-running the simulations. This behaviour model should also include models for the use of natural ventilation (i.e. window opening), operation of blinds and occupancy patterns.
- using the new CSIRO survey deployment platform that is planned to be developed in order to run a new set of surveys and thus gain further insights into behavioural aspects such as blind and window use, drivers and motivations and comfort parameters.
- building on the AC usage model data through the analysis of additional house circuit level metering that is planned to be deployed. In particular, we could investigate clustering of different AC behaviours.

2.6 References

Ambrose, M., James, M., Law, A., & Osman, P. a. (2013). *The evaluation of the 5-star energy efficiency standard for residential buildings.* Australia: CSIRO.

ASHRAE. (2004). *Standard 55 Thermal environmental conditions for human occupancy.* Atlanta: American Society of Heating, Refrigeration and Air-conditioning Engineers.

Australian Building Codes Board. (2016). *National Construction Code, Volume 2.* Canberra: ABCB.

Australian Government. (n.d.). *The National Construction Code and State and Territory Regulations*. Retrieved 12 22, 2016, from http://nathers.gov.au/governance/national-construction-code-and-state-and-territory-regulations

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences.* New Jersey: Lawrence Erlbaum Associates.

Cutler, D., Winkler, J., Kruis, N., Chistensen, C., & Brandemuehl, M. (2013). *Improved modeling of residential air conditioners and heat pumps for energy calculations.* Colorado: NREL.

Daniel, L., Soebarto, V., & Williamson, T. (2015). House energy rating schemes and low energy dwellings: the impact of occupant behaviours in Australia. *Energy and Buildings, 88*, 34-44.

Department of Environment, Water, Heritage and the Arts. (2008). *Energy use in the Australian residential sector 1986-2020.* Canberra: Australian Government.

Goldsworthy, M. (2017). Building thermal design for solar photovoltaic air-conditioning in Australian climates. *Energy and Buildings, 135*, 176-186.

Goldsworthy, M. (n.d.). Towards a residential air-conditioner usage behaviour model for Australia. *Energy and Buildings, Submitted manuscript*.

Goldworthy, M., & Seongwon, S. a. (2016). *Predictors of residential air-conditioner cooling behaviour.* Newcastle: CSIRO.

Hoes, P., Trcka, M., & Hensen, J. a. (2011). Optimizing building designs using a robustness indicator with respect to user behaviour. *Proceedings of the 12th Int. IBSPSA conference* (pp. 1710-1717). Sydney: Int. Building Performance Simulation Association.

Karimpour, M., Belusko, M., Xing, K., Boland, J., & Bruno, F. (2015). Impact of climate change on the design of energy efficient residential building envelopes. *Energy and Buildings, 87*, 142-154.

Kordjamshidi, M., & King, S. (2009). Overcoming problems in house energy ratings in temperature climates; a proposed new rating framework. *Energy and Buildings, 41*, 125-132.

Law, A., Osman, P., Ambrose, M., James, M., & White, S. (2014, November). Energy consumption of 100 Australia residential air conditioners. *Ecolibrium*.

McLeod, P., & Fay, R. (2010). Costs of improving the thermal performance of houses in a cool-temperate climate. *Architectural Science Review, 53*(3), 307-314.

Miller, W., Buys, L., & Bell, J. (2012). Performance evaluation of eight contemporary passive solar homes in subtropical Australia. *Building and Environment, 56*, 57-68.

NatHERS National Administrator. (2012). *Nationwide house energy rating scheme - software accreditation protocol.* Canberra: DCCEE.

Smith, G., Aguilar, J., Gentle, A., & Chen, D. (2012). Multi-parameter sensitivity analysis; a design methodology applied to energy efficiency in temperature climate houses. *Energy and Buildings, 55*, 668-673.

Soebarto, V., & Bennetts, H. (2014). Thermal comfort and occupant responses during summer in a low to middle income housing development in South Australia. *Building and Environment, 75*, 19-29.

Steadman, R. (1984). A universal scale of apparent temperature. *J. App. Met., 23*(12), 1674.

Strategic Insights. (2014). *Queensland Household Energy Survey 2013.* Brisbane: Colmar Brunton.

U.S Department of Energy. (2016). *EnergyPlus Version 8.5 Documentation: Engineering Reference.* U.S DOE.

Whitsed, R., & Davidson, P. (2013). Remote assessment of factors influencing housing thermal efficiency in NSW using GIS and aerial photography. *Urban Policy and Research, 31*(4), 473-492.

Williamson, T. (2013). Designing houses for the Australian climate: the early research. *Architectural Science Review, 56*(3), 197-207.

Williamson, T., Soebarto, V., & Radford, A. (2010). Comfort and energy use in five Australian award-winning houses: regulated, measured and perceived. *Building research and infomation, 38*(5), 509-529.

3 Influence of building parameters on total energy consumption- statistical analysis

3.1 Introduction

In recent years, building codes and standards in Australia have changed on a number of occasions. For example, in 2003, 2006 and 2010 the National Construction Code [1] made changes to the energy-efficiency requirements. There is an interest in identifying whether these changes in energy-efficiency requirements have resulted in detectable shifts in energy usage when compared with buildings built under different construction codes and energy-efficiency standards. As the Productivity Commission [2] noted

'There is considerable uncertainty about the extent to which building standards have reduced energy consumption and emissions.'

In this section, we analyse monthly substation energy consumption data from South-East Queensland that was provided by Energex, with a view to identifying if energy consumption changes according to the year that the buildings were constructed, as well as other factors, such as building characteristics and socio-economic factors. In addition, we analyse Australia-wide annual energy usage data collected by the Australian Bureau of Statistics (ABS). In the next subsection (3.2) we outline the data from South-East Queensland, the analysis undertaken and the results. In the following subsection (3.3) we outline the Australia-wide data, the analysis undertaken and the results from the analysis. We end this section with a discussion and identify next steps.

3.2 South-East Queensland monthly energy use data set

3.2.1 Data

For this analysis, we used monthly energy usage data from over 160 postcodes in South-East Queensland. The analysis involved linking four data sets. The postcode data set included monthly consumption data that was available for the period 1 January 2011 to 31 December 2015, the number of domestic customers, the total energy consumed in the postcode and the average energy consumed per customer for each month.

The second data set contained information from over 200 substations in South-East Queensland and included the variables:

- tariff type (e.g. whether customers were on controlled loads and the type of controlled load tariff type)
- energy consumed
- property age based on NMI
- number of NMIs with pools
- domestic and business energy consumption quintiles
- estimated number of households in different income classes

- number of dwellings of different types (e.g. separate house, semi-detached dwelling and flat)
- estimated number of residences with different numbers of bedrooms.

The third data set was the Building Exposure Statistical Area Level 1 data set from the National Exposure Information System of Geoscience Australia [3]. This data set included the following information:

- number of buildings build at different times
- the footprint size of buildings
- floor area of different building types (e.g. separate house, semi-detached house, two storey apartments, three storey apartments and 4+ storey apartments)
- the type of building material used for walls and roofs of buildings
- socio-economic variables such as income and tenure type.

The final data set used in the analysis was the set of meteorological readings from local weather stations in South-East Queensland. This data set was used to derive a number of meteorological variables:

- apparent temperature (incorporates temperature and humidity information) [4]
- 'heating' and 'cooling' degree days (higher values mean more 'heating' or 'cooling' days than those regions with lower values). The 'cooling' degree days were calculated from the half-hourly data as

$$\frac{\sum_{i=1}^{365}\sum_{j=1}^{48}\left(T_{apparent_ij}-T_{Creference}\right)}{48}$$

and the 'heating' degree days were calculated from the half-hourly data as

$$\frac{\sum_{i=1}^{365}\sum_{j=1}^{48}\left(T_{Hreference}-T_{apparent_ij}\right)}{48},$$

where $T_{Hreference}=15^{o}\,\mathrm{C}$ and $T_{Creference}=24^{o}\,\mathrm{C}$.

3.2.2 Analysis method

Many of the variables in the Energex and in the Building Exposure data sets represent subsets of a super variable. For example, the total number of customers on domestic tariffs is broken down into those without controlled load tariffs and those on either tariff 9000 or 9100, or on both tariffs 9000 and 9100. This means that the variables pertaining to tariffs are compositional data [5], as they can be made to represent the proportion of customers with the different tariff types. Importantly, the sum of the subsets of the various domestic tariff customer types should add up to 1 for each of the substation areas. Consequently, the analyses for this report were undertaken taking into account the fact that many of the variables are compositional in nature. The analyses were undertaken using the **R** statistical package [6] and the **compositions** library for that statistical package.

A linear mixed effects [7] modelling approach was used to analyse the data. This is similar to a linear regression analysis in that it relates a response variable, in this case average energy consumption,

to explanatory variables, which are then tested to see if they help explain a 'significant amount' of variation in the response variable. The linear mixed effects model, however, allows some variables to have random coefficients. Here random coefficients were introduced for the intercept, seasonal effects and trend. That is, each postcode was allowed to have a different intercept term in the model, different trend, and different seasonal effects. As noted above, the analysis respected the fact that some of the variables were compositional in nature. Finally, a backward stepwise approach was used to eliminate variables not found to contribute statistically significantly to the explanation of the variation in the response variable.

3.2.3 Results

Table 3-1 lists the variables (and their *p*-values) that were found to explain a sufficient amount of the variation of the monthly energy usage data in South-East Queensland to be statistically significant. (Another three variables were initially included but found not to contribute significantly: residential building footprint, standard deviation of the daily minimum of apparent temperature and HDD.) An examination of Figure 3-1 reveals that for most postcodes, the average energy consumption declined over time, hence the significant trend term in Table 3-1. In addition, in Figure 3-1 there appear to be approximately two peaks in a 12 month period for many postcodes, and this is identified in the significant seasonal effects in the model in Table 3-1. Average energy consumption was higher, the more swimming pools in a postcode. Also, as the standard deviation of the daily maximum apparent temperature increased, this resulted in higher average energy usage.

Table 3-1 Variables found to significantly explain variation in average energy consumption using the linear mixed effects analysis for the monthly energy consumption data in South-East Queensland

Variable	p-value
Dwelling type	0.0045
Number of bedrooms	0.0021
Income	0.0019
Number of pools	< 0.0001
Building age	0.0139
Wall types	< 0.0001
Roof types	0.0394
Building type floor area	0.0010
Standard deviation of daily maximum of Apparent temperature	< 0.0001
Linear trend	< 0.0001
Yearly seasonal effect	< 0.0001
Half-yearly seasonal effect	<0.0001

To get a better understanding of how other variables in Table 3-1 impact average energy usage, we generated biplots [8]. These plot the principal component loadings for each variable and the principal component scores for each variable on the same plot. This allows one to see which observations are more heavily loaded on which variables.

The biplot for the first two components for dwelling type are plotted in Figure 3-2. In this figure, we have also classified each postcode by calculating the average energy usage over 60 months and determining into which quintile they fall across all postcodes. On the biplot charts, the label "1" represents postcodes in the lowest quintile of average energy usage over the 60 months. In contrast, "5" represents those postcodes in the highest quintile of average energy usage over the 60 months. An examination of Figure 3-2 reveals that the first principal component contrasts separate houses with flats and the second principal component contrasts flats and separate houses with semi-detached dwellings. Treating the principal component scores as response variables and the quintiles as explanatory variables, a one-way analysis of variance revealed that:

- for the first principal component scores, the postcodes in the second to fifth quintiles of average energy usage had significantly greater first principal component scores than the postcodes in the first quintile. This suggests that postcodes in the first quintile of average energy usage over

the 60 months tended to have proportionately more flats than the postcodes in the other four quintiles.

- In terms of the second principal components scores, the third to fifth quintiles had significantly lower second principal component scores than the first quintile. This suggests that the postcodes in the third to fifth quintiles of average energy usage over the 60 months tended to have proportionately moresemi-detached dwellings than the postcodes in the first quintile.

For the number of bedrooms, the biplot of the first two components in Figure 3-3 reveals that the first principal component contrasts the four or more bedroom dwellings with the dwellings with no more than two bedrooms. The second principal component contrasts the five or more bedroom dwellings (and to a lesser extent one bedroom dwellings) with the three bedroom dwellings (and to a lesser extent two bedroom dwellings). A one-way analysis of variance for the first principal component scores for the number of bedrooms variables revealed that:

- the postcodes in the second to fifth quintiles of average energy usage had significantly greater first principal component scores than the postcodes in the first quintile. This suggests that postcodes in the first quintile of average energy usage over the 60 months tended to have proportionately more one and two bedroom dwellings than the postcodes in the other four quintiles.

The biplot for the income variables is displayed in Figure 3-4. It can be seen that the first principal component contrasts the highest income bracket with the two lowest income brackets, while the second principal component contrasts the middle-income bracket (and to a lesser extent the fourth highest income bracket) with the highest and the lowest income brackets. The one-way analysis of variance revealed that:

- the postcodes in the highest quintile of average energy usage over the 60 months had higher first principal component scores than the postcodes in the lowest quintile. This suggests that the postcodes in the highest quintile tended to have proportionately more households on the highest income bracket than the postcodes in the lowest quintile, which tended to have proportionately more on the two lowest income brackets.
- For the second principal component scores, the postcodes in the second to fifth quintiles of average energy usage over the 60 months had significantly lower second principal component scores than the postcodes in the first quintile. This suggests that the second to fifth quintiles had proportionately more households on third (and fourth) highest income bracket than the postcodes in the first quintile.

In terms of year of construction, the biplot for the first and third principal components can be found in Figure 3-5. (The second principal components are not presented as there was no meaningful pattern evident.) The first principal component contrasts houses built between 1788 and 1951 with houses built after 1986, particularly after 2001. The third principal component contrasts houses built between 2002 and 2006 with those houses built between 2007 and 2011. The one-way analysis of variance on the first and third principal components revealed that:

- the postcodes in second to fifth quintiles of average energy usage over 60 months had relatively lower first principal component scores than the postcodes in the first quintile. This suggests that the households in the first quintile tended to have proportionately more houses built in the first

age period while those in the top four quintiles tended to have proportionately more from the two most recent age periods.

- for the third principal component scores, the postcodes in the highest quintile had relatively lower third principal component scores than the postcodes in the first quintile. This suggests that the postcodes in the highest quintile tended to have proportionately more houses built between 2002 and 2006, while the postcodes in the lowest quintile tended to have proportionately more houses built between 2007 to 2011.

Figure 3.6 shows the biplot for the first two principal components for wall type materials. The second principal component contrasts fibre cement walls with timber walls and to a lesser extent cavity and solid masonry walls. The one-way analysis of the quintiles on the second principal component scores revealed that:

- the postcodes in the top three quintiles of average energy consumption over the 60 months had significantly lower second principal component scores than the postcodes in the first quintile. This suggests that the postcodes in the lowest quintile had proportionately more fibre cement walls than the top three quintiles.

For roof type material, the one-way analysis of variance of the principal component scores against the quintiles of average energy usage across 60 months did not reveal any statistically significant outcomes. This is perhaps not surprising, given that in Table 3-1 roof type had the largest *p*-value, which was only just significant at a 5% significance level.

The biplot in Figure 3-7 plots the first two principal components for the floor area by dwelling type, while the biplot for the second and third can be found in Figure 3-8. From Figure 3-7, we see that the first principal component contrasts the separate houses and semi-detached dwellings with the two or more storey apartments, while the second principal component contrasts the four or more storey apartments with the two or three storey apartments. For the third principal component, it can be seen from Figure 3.8 that this differentiates between two-storey apartments (and to a lesser extent semi-detached dwellings) and three-storey apartments (and to a lesser extent separate houses). One-way analysis of variance for the first three principal component scores revealed that:

- the postcodes in the top four quintiles of average energy consumption over 60 months had significantly lower values on the first principal component scores. This suggests that these postcodes had proportionately more separate houses and semi-detached dwellings than the postcodes in the lowest quintile.
- for the second principal component scores, the postcodes in the top two quintiles of average energy consumption over 60 months had higher second principal component scores than the postcodes in the lowest quintile. This suggests that the postcodes in the highest two quintiles had proportionately more four or more storey apartment blocks than the postcodes in the lowest quintile.
- for the third principal component scores, it was found that the postcodes in the third and fourth quintiles of average energy consumption over 60 months had significantly higher third principal component scores than the postcodes in the lowest quintile. This suggests that the postcodes in third and fourth quintiles had proportionately more three storey apartment blocks (and to a lesser extent separate houses) than the postcodes in the first quintile, which had

proportionately more two storey apartment blocks (and to a lesser extend semi-detached dwellings).

- the postcodes in the highest quintile had significantly larger third principal component scores than the two lowest quintiles. This suggests that the postcodes in the highest quintile had proportionately more lthree storey apartments (and to a lesser extent separate houses) than the two lowest quintiles.

The principal components analysis results in this section are summarised in Table 3-2. The information in this table should be read by rows. Further discussion of the results in this section is deferred until Section 3.4, so as to make comparisons with the results in the next section.

Table 3-2 Summary of principal component analyses of the monthly South-East Queensland data

	Energy usage quintiles				
Variable	**1st quintile**	**2nd quintile**	**3rd quintile**	**4th quintile**	**5th quintile**
Dwelling type	Proportionately more flats than other quintiles	Proportionately less flats than 1st quintile	Proportionately less flats than 1st quintile	Proportionately less flats than 1st quintile	Proportionately less flats than 1st quintile
	Proportionately less semi-detached dwellings than 3rd to 5th quintiles		Proportionately more semi-detached dwellings than 1st quintile	Proportionately more semi-detached dwellings than 1st quintile	Proportionately more semi-detached dwellings than 1st quintile
Number of bedrooms	Proportionately more 1 and 2 bedroom dwellings than other quintiles	Proportionately less 1 and 2 bedroom dwellings than 1st quintile	Proportionately less 1 and 2 bedroom dwellings than 1st quintile	Proportionately less 1 and 2 bedroom dwellings than 1st quintile	Proportionately less 1 and 2 bedroom dwellings than 1st quintile
Income	Proportionately more on the 2 lowest income brackets				Proportionately more on the highest income bracket
	Proportionately less on the 3rd and 4th highest income brackets than other quintiles	Proportionately more on the 3rd and 4th highest income brackets than the 1st quintile	Proportionately more on the 3rd and 4th highest income brackets than the 1st quintile	Proportionately more on the 3rd and 4th highest income brackets than the 1st quintile	Proportionately more on the 3rd and 4th highest income brackets than the 1st quintile
Building age	Proportionately more properties built prior to 1952 than the other quintiles	Proportionately more properties built after 2011 than the 1st quintile	Proportionately more properties built after 2011 than the 1st quintile	Proportionately more properties built after 2011 than the 1st quintile	Proportionately more properties built after 2011 than the 1st quintile
	Proportionately more properties built between 2007 and 2011 than the 5th quintile				Proportionately more properties built between 2002 and 2006 than the 1st quintile

	Energy usage quintiles				
Variable	**1st quintile**	**2nd quintile**	**3rd quintile**	**4th quintile**	**5th quintile**
Wall type material	Proportionately more fibre cement walls than the 3rd to 5th quintiles		Proportionately less fibre cement walls than the 1st quintile	Proportionately less to fibre cement walls than the 1st quintile	Proportionately less to fibre cement walls than the 1st quintile
Floor area by dwelling type	Proportionately less separate houses and semi-detached dwellings than the other quintiles	Proportionately more separate houses and semi-detached dwellings than the 1st quintile	Proportionately more separate houses and semi-detached dwellings than the 1st quintile	Proportionately more separate houses and semi-detached dwellings than the 1st quintile	Proportionately more separate houses and semi-detached dwellings than the 1st quintile
	Proportionately less 4+ storey apartment blocks than the 4th and 5th quintiles			Proportionately more 4+ storey apartment blocks than the 1st quintile	Proportionately more 4+ storey apartment blocks than the 1st quintile
	Proportionately more 2 storey apartment blocks than the 3rd and 4th quintiles		Proportionately more 3 storey apartment blocks than the 1st quintile	Proportionately more 3 storey apartment blocks than the 1st quintile	
	Proportionately less 3 storey apartment blocks than the 5th quintile	Proportionately less 3 storey apartment blocks than the 5th quintile			Proportionately more 3 storey apartment blocks than 1st and 2nd quintiles

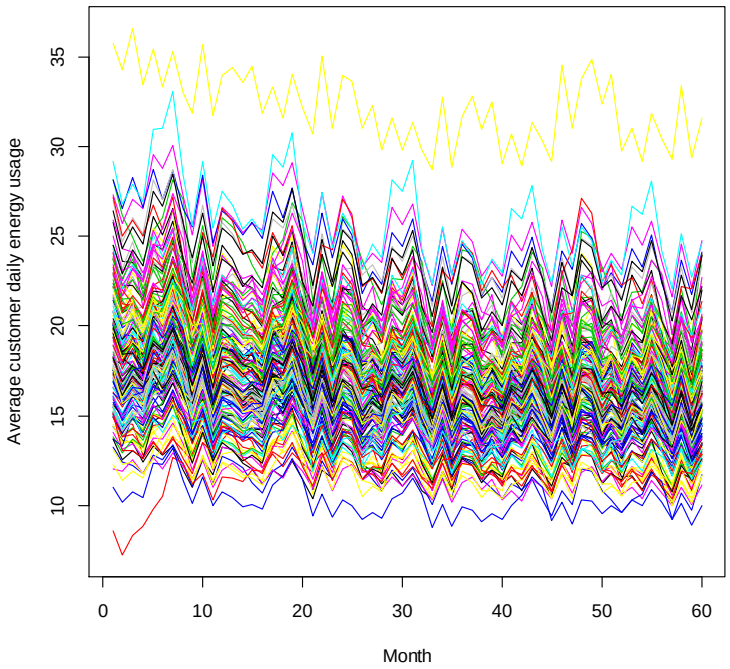

Figure 3-1 Monthly average energy consumption for the different postcodes

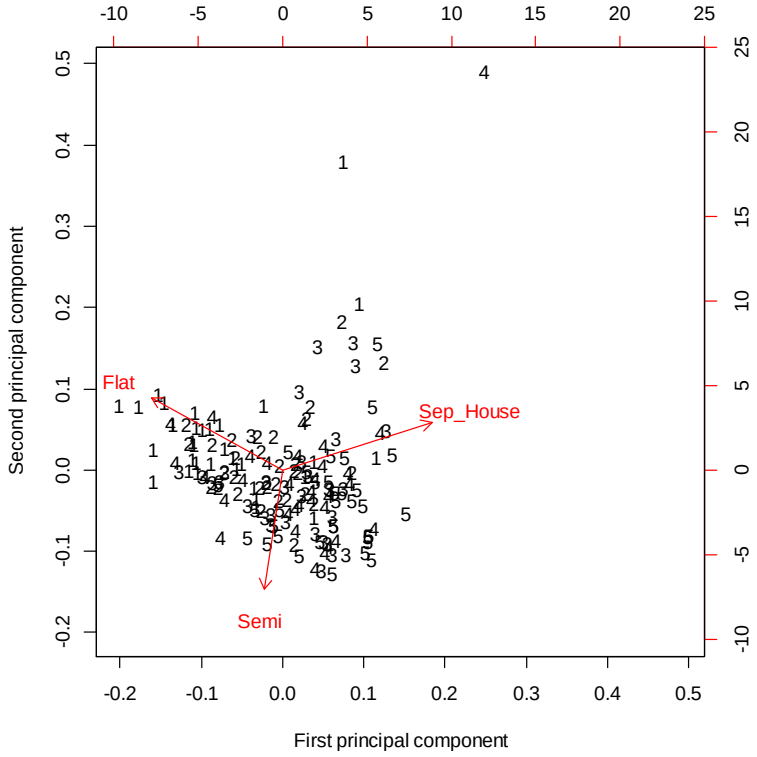

Figure 3-2 Biplot of first two components for dwelling type

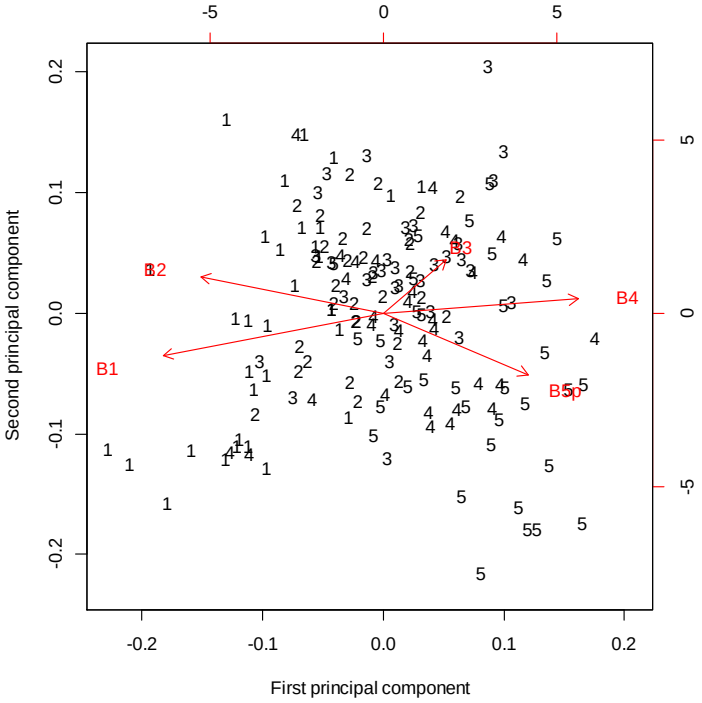

Figure 3-3 Biplot of first two components for number of bedrooms. (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms)

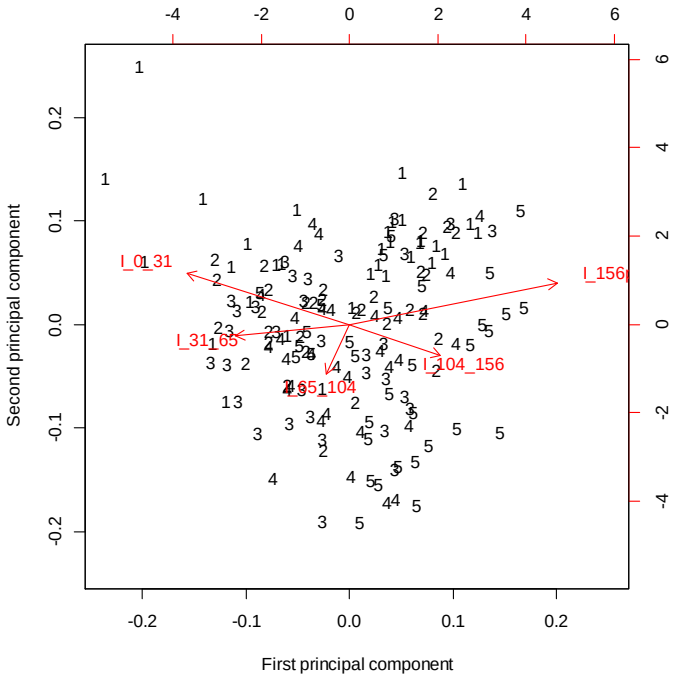

Figure 3-4 Biplot of first two components for income classes. (I_0_31 = income <$31,000; I_31_65 = $31,000 ≤ Income < $65,000; I_65_104 = $65,000 ≤ Income < $104,000; I_104_156 = $104,000 ≤ Income < $156,000; and I156p = Income ≥ $156,000)

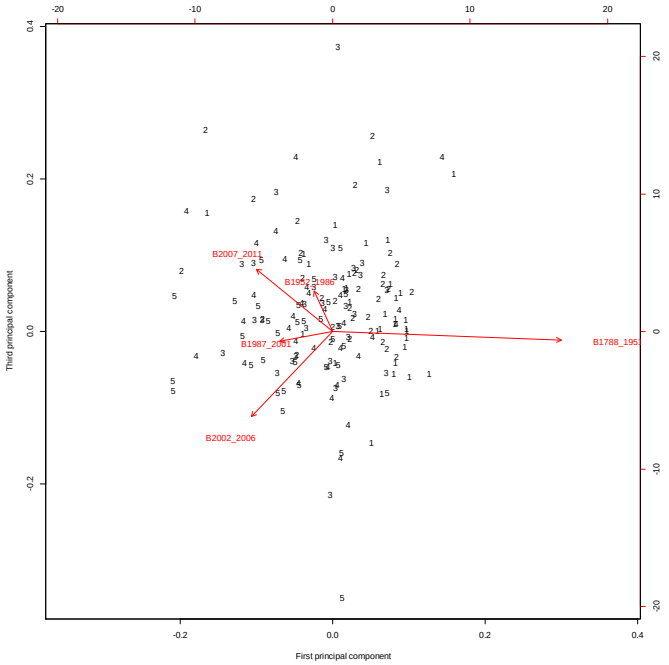

Figure 3-5 Biplot of first and third components for building age classes. (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011. Note that the data from Geoscience Australia had no buildings listed as having been built after 2011 for South-East Queensland)

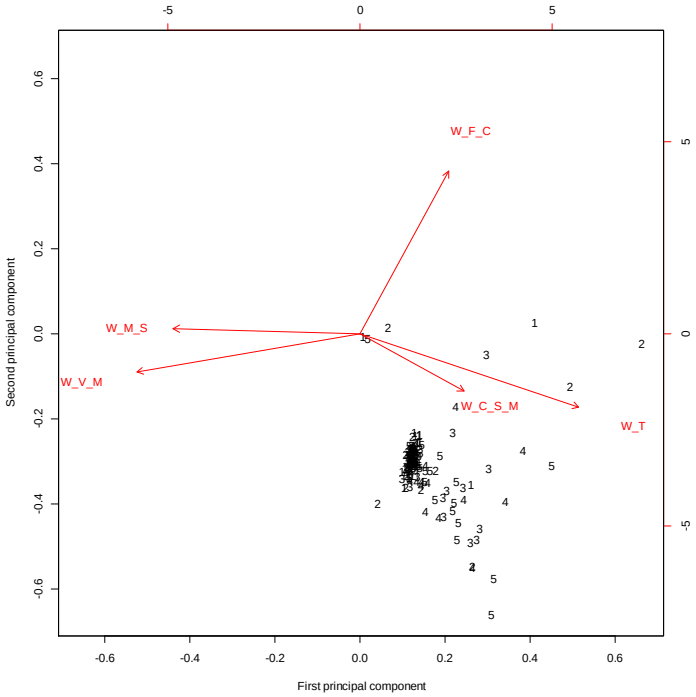

Figure 3-6 Biplot of first two components for wall type materials. (W_C_S_M = cavity and solid masonry; W_F_C = fibre cement; W_M_S = metal sheeting; W_T = timber; and W_V_M = veneer masonry)

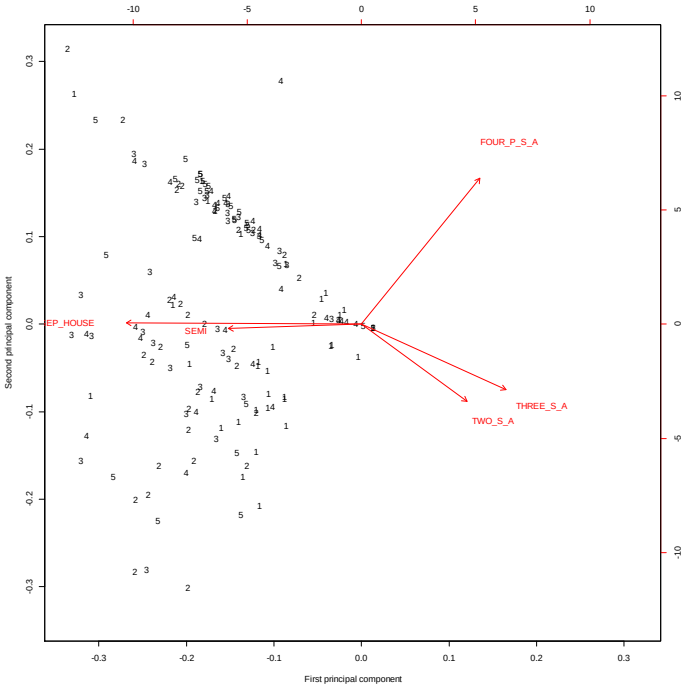

Figure 3-7 Biplot of first two components for floor area by dwelling type. (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two storey apartments; THREE_S_A = area for three storey apartments; and FOUR_P_S_A = area for four or more storey apartments)

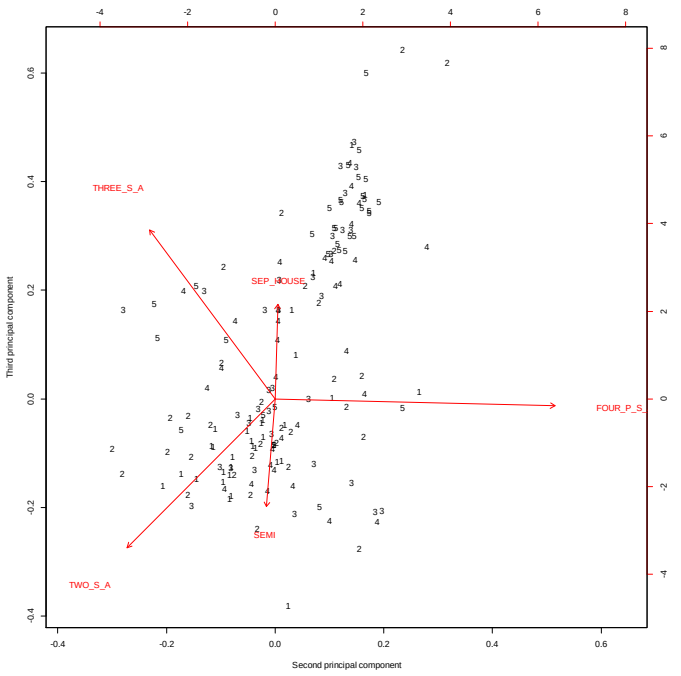

Figure 3-8 Biplot of second and third components for floor area by dwelling type. (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two storey apartments; THREE_S_A = area for three storey apartments; and FOUR_P_S_A = area for four or more storey apartments)

3.3 Annual energy use across Australia[1]

3.3.1 Data

For the Australia-wide annual energy use analysis, three data sets were linked together. Two of these were used in the analysis of the monthly energy usage by postcode in South-East Queensland in the previous subsection. These were the Building Exposure Statistical Area Level 1 data set from the National Exposure Information System of Geoscience Australia, and the meteorological readings from local weather stations. The spatial extent of the data drawn from the Building Exposure data and meteorological data differs between this subsection and the previous one. In this case, we used data from the six states of Australia and the Australian Capital Territory[2].

The third data set was the ABS 4670.0 Household Energy Consumption Survey for non-generating residential meters across Australia (except Northern Territory). The data set contains annual average energy consumption at the statistical area (SA) 2 level for 2010, 2011 and 2012. (Note that data for residential net meters was not included in the study. Many of the SA2s were missing values as a result of only a small number of dwellings in those areas having appropriate meters). This data set also assigned each SA2 to a climate zone. The definition of the climate zones in the ABS data set is:

- Zone 1 – High humid summer, warm winter
- Zone 2 – Warm humid summer, mild winter
- Zone 3 – Hot dry summer, warm winter
- Zone 4 – Hot dry summer, cool winter
- Zone 5 – Warm temperature
- Zone 6 – Mild temperature
- Zone 7 – Cool temperature.

3.3.2 Analysis method

As in Section 3.2, the analysis in this subsection involved the use of linear mixed effects modelling. The average annual energy consumption for 1,913 SA2s were treated as responses, and the explanatory variables included climatic zone, derived meteorological variables from the meteorological data set, building age, building material characteristics and area, and socio-economic variables from the Building Exposure data set. As in the previous subsection (3.2), the compositional nature of variables was respected. As there were only three data points (years) for each SA2, no trend term was included. The annual nature of the data also precluded using seasonal terms.

[1] The data examined here does not include energy data from the Northern Territory.

[2] With the exception that meteorological data from the Northern Territory was included if it represented the nearest recorded weather data for statistical areas from other states.

3.3.3 Results

The statistical output from the linear mixed effects modelling is summarised in Table 3-3. (Other weather-related variables were initially included in the modelling, but due to the high correlation between some of the variables they were removed to avoid computational issues). Note that in Table 3-3 there is a significant interaction between climate zone and cooling degree days (CDD). This means that the main effects for CDD and climate zone are not meaningful, and the interaction effects need to be examined to determine how energy consumption changes with climate zone and CDD. Such an examination revealed that a higher number of CDD in climate zone 5 (warm temperate) and 7 (cool temperate) resulted in significantly less average annual energy consumption compared with the climate zones 1 (high humid summer, warm winter), 2 (warm humid summer, mild winter), 4 (hot dry summer, cool winter) and 6 (mild temperate). In addition, a unit increase in CDD resulted in a greater reduction in energy usage for zone 5 than for zone 3. It should be noted that the average CDD for SA2 regions in zone 5 was a lot smaller than that for the SA2s in region 3, perhaps suggesting that a unit increase in CDD in zone 5 is more likely to move the temperature into a "comfortable" zone, requiring less heating or no heating.

Table 3-3 Variables found to significantly explain variation in average energy consumption using the linear mixed effects analysis for the Australia wide annual energy consumption data

Variable	p-value
Building age	0.0012
Wall types	< 0.0001
Roof type	< 0.0001
Building type floor area	< 0.0001
Equivalised income	< 0.0001
Tenure type	< 0.0001
Cooling degree days	< 0.0001
Climate zone	< 0.0001
Interaction between cooling degree days and climate zone	<0.0001

To get a better understanding of how equivalised income was related to average annual energy consumption for the Australia wide data, the biplot for the first and third principal components can be found in Figure 3-9. From this plot, it can be seen that the first principal component contrasts the highest equivalised income class with the two lowest equivalised income classes. The third principal component contrasts the middle equivalised income class with the lowest equivalised

income class. One-way analysis of variance of the first and third principal component scores (no significant difference was identified for the 2nd principal component scores) showed that:

- the SA2s in the highest quintile of average annual energy usage over the three years had significantly lower first principal component scores than the SA2s in the lowest and third quintile. This suggests that the SA2s in the highest quintile had proportionately more households on the highest equivalised income class while the SA2s in the lowest and third quintile had proportionately more households on the two lowest equivalised income classes.
- for the third principal component scores, the SA2s in the second highest quintile of average energy usage over the three years had significantly higher third principal component scores than the SA2s in the first quintile. This suggests that the SA2s in the second highest quintile had proportionately more households on the middle equivalsed income class than the SA2s in the lowest quintile.
- the SA2s in the highest quintile of average energy usage over the three years had significantly higher third principal component scores than the SA2s in the three bottom quintiles. This suggests that the SA2s in the highest quintile had proportionately more households on the middle equivalised income than the three lowest quintiles, which had proportionately more households on the low equivalised income class.

For building age, the biplot for the first and third principal components can be found in Figure 3-10. The first principal component contrasts buildings built prior to 1952 with buildings built after 1986. The third principal component contrasts buildings built between 1987 and 2001 with buildings built between 2007 and 2011. The one-way analysis of variance of the first and third principal component scores (no significant difference was identified for the 2nd principal component scores) indicated that:

- SA2s in the second lowest and the highest quintile of average energy usage over the three years had lower first principal component scores than the SA2s in the lowest quintile. This suggests that the SA2s in the lowest quintile had proportionately more buildings built prior to 1952, and the second lowest quintile and the highest quintile had proportionately more buildings built after 1986.
- for the third principal component scores, the SA2s in the four highest quintiles had significantly higher third principal component scores than the SA2s in the first quintile. This implies that the SA2s in the four highest quintiles had proportionately more buildings built between 1987 and 2001, while the SA2s in the lowest quintile had proportionately more buildings built between 2007 and 2011.

Figures 3-11 and 3-12 are biplots for the first three principal components for roof type material. The first principal component contrasts tiles and fibre cement with imitation tiles, the second contrasts metal sheeting and tiles with fibre cement, and the third contrasts imitation tiles and tiles with metal sheeting. The one-way analysis of the variance for the first three principal component scores revealed that:

- the SA2s in the two highest quintiles of average energy consumption over the three years had significantly lower first principal component scores than the SA2s in the two lowest quintiles. This suggests that the SA2s in the two highest quintiles had proportionately more imitation roof

tiles and the SA2s in the two lowest quintiles had proportionately more fibre cement and tiled roofs.

- for the second principal component scores, the SA2s in the top four quintiles of average energy usage over the three years had significantly higher second principal component scores than the SA2s in the lowest quintile. This means that the SA2s in the four highest quintiles had proportionately more fibre cement roofs than the SA2s in the lowest quintile, which had proportionately more metals sheeting or tiled roofs.
- for the third principal component scores, the SA2s in the second to fourth highest quintiles had relatively higher third principal component scores than the SA2s in the lowest quintile. This suggests that the SA2s in the lowest quintile had proportionately more metal sheeting roofs, while the SA2s in the second to fourth quintiles had proportionately more imitation tiled and tiled roofs.

The biplots for the first five principal components for wall type materials are found in Figures 3-13 to 3-15. The first principal component is a comparison between reinforced concrete masonry and veneer masonry; the second principal component is a contrast between concrete solid masonry and veneer masonry and metal sheeting; the third principal component differentiates between veneer masonry and metal sheeting, and timber; the fourth principal component distinguishes between timber and concrete solid masonry, and fibre cement; and the fifth principal component contrasts timber and veneer masonry (and, to a lesser extent, reinforced concrete masonry) with concrete solid masonry. The one-way analysis of variance for the first five principal component scores revealed that:

- the SA2s in the second to fourth quintiles of average energy consumption over three years had significantly lower first principal component scores than the SA2s in the first quintile. The SA2s in the third quintile also had significantly lower first principal component scores than the SA2s in the highest quintile. This means the SA2s in the second to fourth quintiles had proportionately more reinforced concrete masonry walls than the SA2s in the first quintile, who had proportionately more veneer masonry walls.
- for the second principal component scores, the SA2s in the two highest quintiles of average energy usage over the three years had significantly lower second principal component scores than the SA2s in the two lowest quintiles. This implies that the SA2s in the two highest quintiles had proportionately more metal sheeting walls than the SA2s in the two lowest quintiles, which had proportionately more cavity and solid masonry and veneer masonry walls.
- for the third principal component scores, the SA2s in the second lowest quintile had significantly higher principal component scores than the SA2s in the lowest quintile. This suggests that the SA2s in the lowest quintile had proportionately more timber walls, while the SA2s in the second lowest quintile had proportionately more veneer masonry and metal sheeting walls.
- for the fourth principal components, the SA2s in the four highest quintiles had significantly higher fourth principal component scores than the SA2 in the lowest quintile. This suggests that the SA2s in the four highest quintiles had proportionately more fibre cement walls and the SA2s in the lowest quintile had proportionately more timber and cavity and solid masonry walls.
- for the fifth principal component scores, the SA2s in the two highest quintiles had significantly lower fifth principal component scores than the SA2s in the lowest quintile. The SA2s in the

highest quintile also had lower values than the SA2s in the second lowest quintile. As a result, the SA2s in the two highest quintiles had proportionately more timber and veneer masonry walls, while the SA2s in the lowest quintile (and to a lesser extent second lowest quintile) had proportionately more cavity and solid masonry walls.

Figures 3-16 and 3-17 are biplots for the first four principal components for floor area for different dwelling types. The first principal component contrasts the separate houses and semi-detached dwellings against the apartments with three or more storeys; the second principal component contrasts apartments with four or more storeys and separate houses with apartments with two storeys; the third principal component differentiates the three-storey apartments (and to a lesser extent separate houses) from the semi-detached dwellings, two-storey apartments and the four- or more storey apartments; and the fourth component contrasts the semi-detached houses with the separate houses (and to a lesser extent the two-storey apartments). The one-way analysis of variance for the first four principal component scores revealed that:

- the SA2s in the highest quintile had significantly higher first principal component scores than the SA2s in the lowest quintile. This suggests that the SA2s in the highest quintile had proportionately more three or more storey apartments and the SA2s in the lowest quantile had proportionately more separate houses and semi-detached dwellings.
- for the second principal component scores, the SA2s in the two highest quintiles had significantly higher second principal component scores than the SA2s in the two lowest quintiles. This suggests that the SA2s in the two highest quintiles had proportionately more four or more storey apartments and separate houses compared with the SA2s in the two lowest quintiles, which had proportionately more two-storey apartments.
- for the third principal component scores, the SA2s in the four highest quintiles had significantly lower third principal component scores than the SA2s in the lowest quintile. The SA2s in the highest two quintiles also had significantly lower scores than the SA2 in the second lowest quintile, and the SA2s in the highest quintile had significantly lower scores than the SA2s in the middle quintile. The general trend is that the SA2s in higher quintiles had proportionately more three storey apartments than the SA2s in the lower quintiles.
- for the fourth principal component scores, the SA2s in the four highest quintiles had significant higher fourth principal component scores than the SA2s in the lowest quintile. This implies that the four highest quintiles had proportionately more separate houses (and to a lesser extent two-storey apartments), while the SA2s in the lowest quintile had proportionately more semi-detached dwellings.

The biplot for the second and third principal components for the tenure types is presented in Figure 3-18. From this figure, it can be seen that the second principal component contrasts private rent tenure with other tenure and the third principal component contrasts own tenure with other tenure. The one-way analysis of variance of the second and third principal component scores revealed that:

- the SA2s in the four highest quintiles had significantly larger scores than the SA2s in the lowest quintile, and the two highest quintiles also had significantly higher scores than the SA2s in the second lowest quintile. This suggests that generally the higher quintiles had proportionately

more “other” tenure types than the lower quintiles, which had proportionately more private rent tenures.

- for the third principal component scores, the SA2s in the two highest quintiles had higher third principal component scores than the SA2s in the lowest quintile. The SA2s in the highest quintiles also had significantly higher scores than the SA2s in the second and third lowest quintiles. This suggests that the SA2s in the two highest quintiles had proportionatelymore “own” tenures than the SA2s in the lower quintiles, which had proportionately more “other” tenures.

The principal components analyses results in this section are summarised in Table 3-4. The information in this table should be read by rows. In the next section we compare the results from this and the previous section, and highlight some of the limitations of the analyses conducted.

Table 3-4 Summary of principal component analyses of the annual Australia wide data

	Energy usage quintiles				
Variable	**1st quintile**	**2nd quintile**	**3rd quintile**	**4th quintile**	**5th quintile**
Equivalised income	Proportionately more on the 2 lowest income classes than the top quintile	Proportionately more on the 2 lowest income classes than the top quintile			Proportionately more on the highest income class than the bottom 2 quintiles
	Proportionately less on the middle income class than the 4th quintile			Proportionately more on the middle income class than the 1st quintile	
	Proportionately more on the lowest income class than the highest quintile	Proportionately more on the lowest income class than the highest quintile	Proportionately more on the lowest income class than the highest quintile		Proportionately more on the middle income class than the bottom 3 quintiles
Building age	Proportionately more buildings built prior to 1952 than the 2nd and 5th quintiles	Proportionately more buildings built after 1986 than the 1st quintile			Proportionately more buildings built after 1986 than the 1st quintile
	Proportionately more buildings	Proportionately more buildings	Proportionately more buildings	Proportionately more buildings	Proportionately more buildings

Variable	Energy usage quintiles				
	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile
	built between 2007 and 2011 than the other quintiles	built between 1987 and 2001 than the 1st quintile	built between 1987 and 2001 than the 1st quintile	built between 1987 and 2001 than the 1st quintile	built between 1987 and 2001 than the 1st quintile
Roof type material	Proportionately more fibre cement and tiled roof than the 4th and 5th quintiles	Proportionately more fibre cement and tiled roof than the 4th and 5th quintiles		Proportionately more imitation roof tiles than the 1st and 2nd quintiles	Proportionately more imitation roof tiles than the 1st and 2nd quintiles
	Proportionately more metals sheeting and tiled roof than the other four quintiles	Proportionately more fibre cement roof than the 1st quintile	Proportionately more fibre cement roof than the 1st quintile	Proportionately more fibre cement roof than the 1st quintile	Proportionately more fibre cement roof than the 1st quintile
	Proportionately more metal sheeting than the 2nd to 4th quintiles	Proportionately more imitation tiles and tiled roof than the 1st quintile	Proportionately more imitation tiles and tiled roof than the 1st quintile	Proportionately more imitation tiles and tiled roof than the 1st quintile	Proportionately more imitation tiles and tiled roof than the 1st quintile
Wall type material	Proportionately more veneer masonry walls than the 2nd to 4th quintiles	Proportionately more reinforced concrete masonry walls than the 1st quintile	Proportionately more reinforced concrete masonry walls than the 1st quintile	Proportionately more reinforced concrete masonry walls than the 1st quintile	
	Proportionately more cavity and solid masonry and metal sheeting walls than the 4th and 5th quintiles	Proportionately more cavity and solid masonry and metal sheeting walls than the 4th and 5th quintiles		Proportionately more metal sheeting walls than the 1st and 2nd quintiles	Proportionately more metal sheeting walls than the 1st and 2nd quintiles
	Proportionately more timber	Proportionately more veneer masonry and			

Variable	Energy usage quintiles				
	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile
	walls than the 2nd quintile	metal sheeting walls than the 1st quintile			
	Proportionately more timber and cavity and solid masonry walls than the other quintiles	Proportionately more fibre cement walls than the 1st quintile	Proportionately more fibre cement walls than the 1st quintile	Proportionately more fibre cement walls than the 1st quintile	Proportionately more fibre cement walls than the 1st quintile
	Proportionately more cavity and solid masonry walls than the 4th and 5th quintiles			Proportionately more timber and veneer masonry walls than the 1st quintile	Proportionately more timber and veneer masonry walls than the 1st quintile
Floor area of dwelling types	Proportionately more separate houses and semi-detached dwellings than the 5th quintile				Proportionately more 3 or more storey apartment blocks than the 1st quintile
	Proportionately more 2 storey apartment blocks than the 4th and 5th quintiles	Proportionately more 2 storey apartment blocks than the 4th and 5th quintiles		Proportionately more 4 or more storey apartment blocks and separate houses than the 1st and 2nd quintiles	Proportionately more 4 or more storey apartment blocks and separate houses than the 1st and 2nd quintiles
	Proportionately less 3 or more storey apartment blocks than the other quintiles	Proportionately more 3 or more storey apartment blocks than the 1st quintile but less so than the 4th and 5th quintiles	Proportionately more 3 or more storey apartment blocks than the 1st quintile but less so than the 5th quintile	Proportionately more 3 or more storey apartment blocks than the 1st and 2nd quintiles	Proportionately more 3 or more storey apartment blocks than the 1st to 3rd quintiles

	Energy usage quintiles				
Variable	**1st quintile**	**2nd quintile**	**3rd quintile**	**4th quintile**	**5th quintile**
	Proportionately more semi-detached dwellings than the other quintiles	Proportionately more separate houses than the 1st quintile	Proportionately more separate houses than the 1st quintile	Proportionately more separate houses than the 1st quintile	Proportionately more separate houses than the 1st quintile
Tenure type	Proportionately more private rent than the 2nd to 5th quintiles	Proportionately more "other" tenure type than the 1st quintile but proportionately more private rent than the 4th and 5th quintiles	Proportionately more "other" tenure than the 1st quintile	Proportionately more "other" tenure than the 1st and 2nd quintiles	Proportionately more "other" tenure than the 1st and 2nd quintiles
	Proportionately more "other" tenure than the 4th and 5th quintiles	Proportionately more "other" tenure than the 5th quintile	Proportionately more "other" tenure than the 5th quintile	Proportionately more "own" tenure than the 1st quintile	Proportionately more "own" tenure than the 3 lowest quintiles

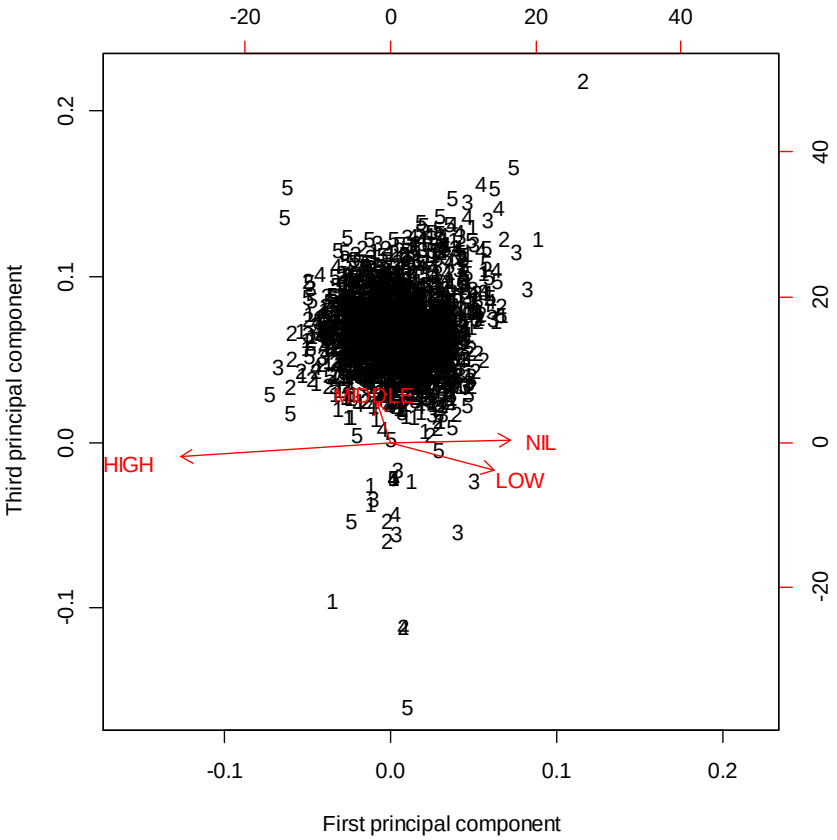

Figure 3-9 Biplot of the first and third principal components for equivalised income

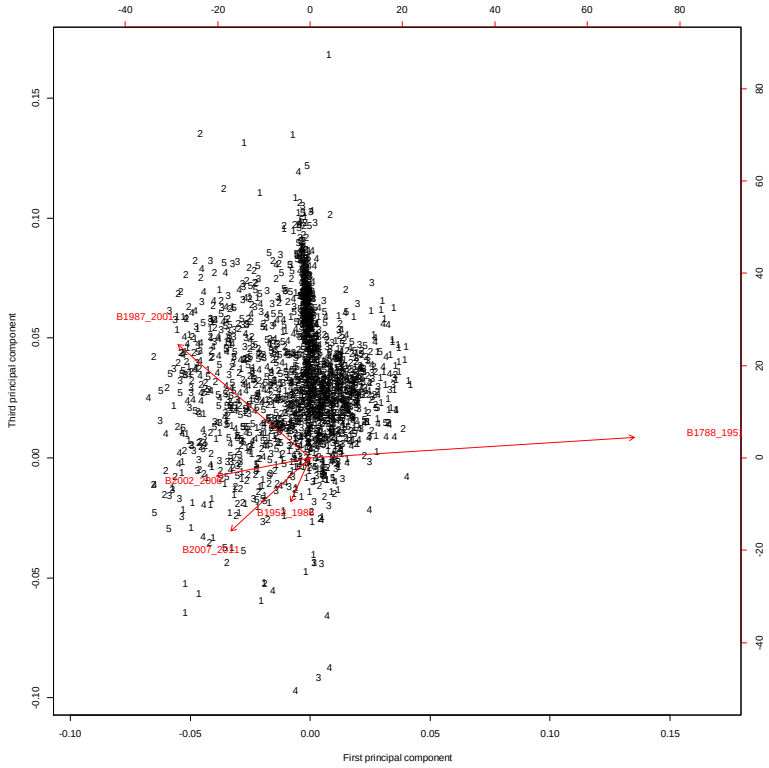

Figure 3-10 Biplot of first and third components for building age classes. (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011)

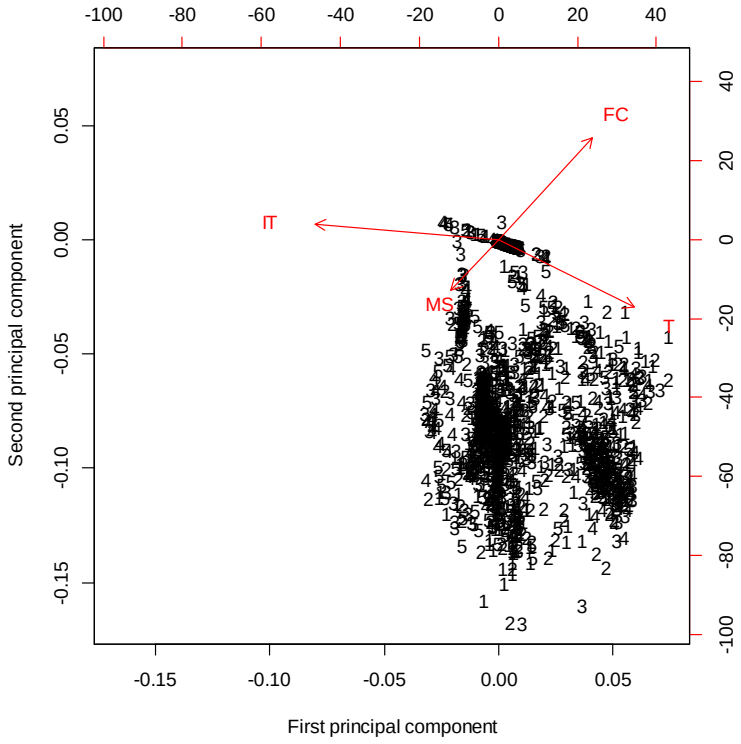

Figure 3-11 Biplot of the first two components for roof type material (IT = imitation tiles; MS = metal sheeting; FC = fibre cement; and T = tiles)

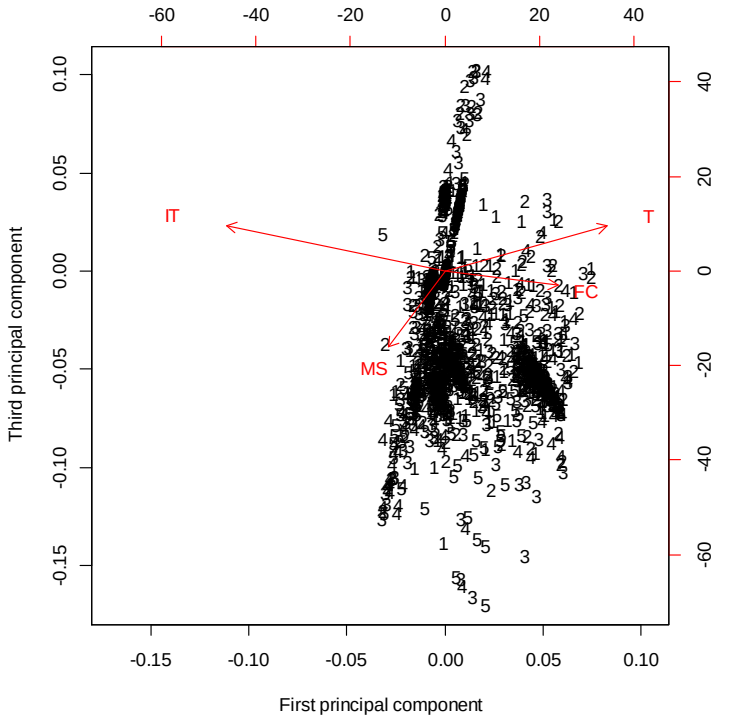

Figure 3-12 Biplot of the first and third components for roof type material (IT = imitation tiles; MS = metal sheeting; FC = fibre cement; and T = tiles)

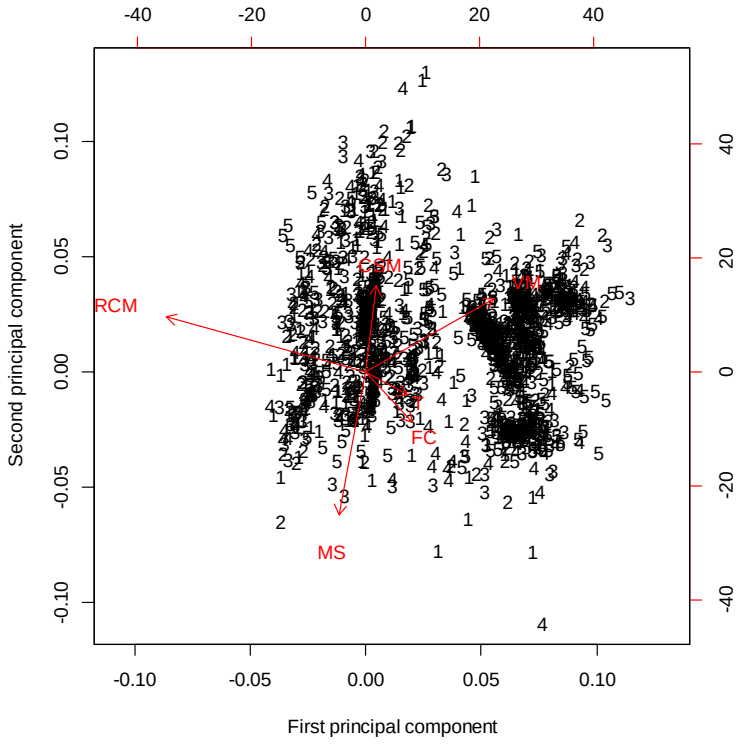

Figure 3-13 Biplot of the first two components for wall type materials (RCM = reinforced concrete masonry; MS = metal sheeting; FC = fibre cement; T = timber; and CSM = cavity and solid masonry)

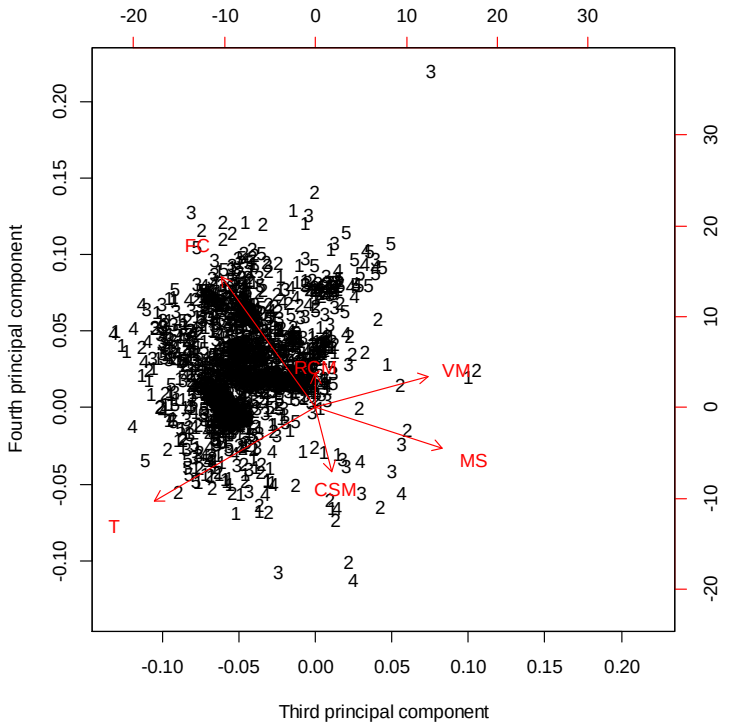

Figure 3-14 Biplot of the third and fourth components for wall type materials (RCM = reinforced concrete masonry; MS = metal sheeting; FC = fibre cement; T = timber; and CSM = cavity and solid masonry)

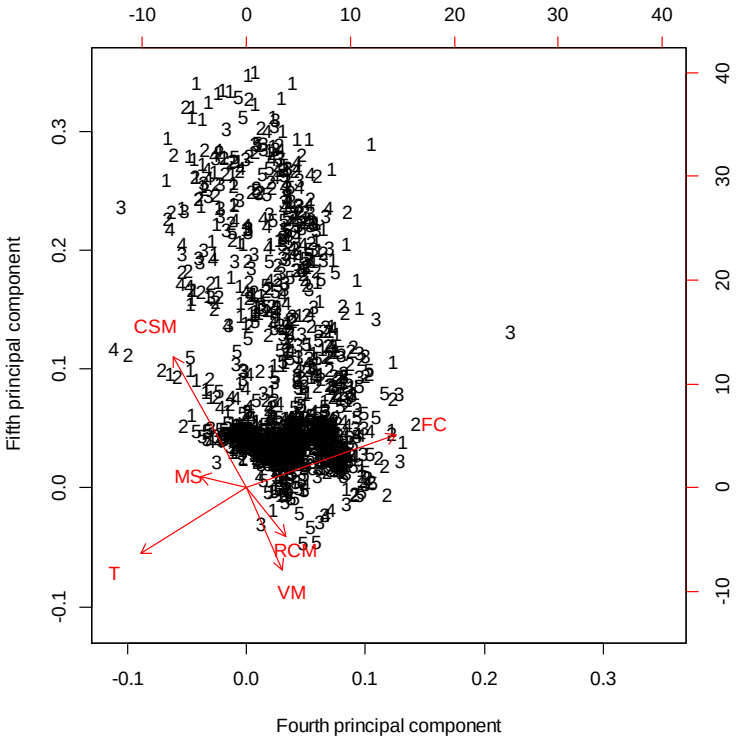

Figure 3-15 Biplot of the fourth and fifth components for wall type materials (RCM = reinforced concrete masonry; MS = metal sheeting; FC = fibre cement; T = timber; and CSM = cavity and solid masonry)

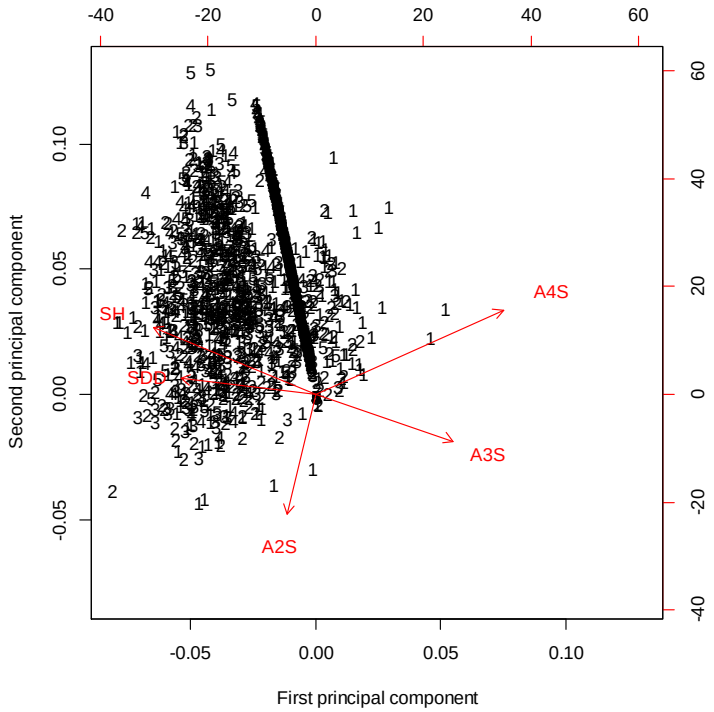

Figure 3-16 Biplot of the first two components for floor area for different dwelling types (SH = separate house; SDD = semi-detached dwelling; A2S = two storey apartments; A3S = three storey apartments; and A4S = four or more storey apartments)

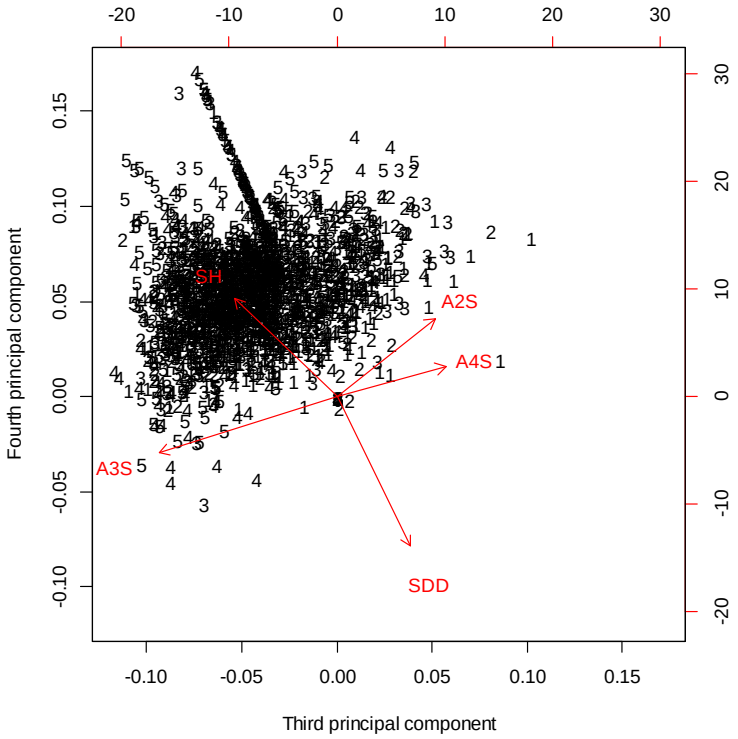

Figure 3-17 Biplot of the third and fourth components for floor area for different dwelling types (SH = separate house; SDD = semi-detached dwelling; A2S = two storey apartments; A3S = three storey apartments; and A4S = four or more storey apartments)

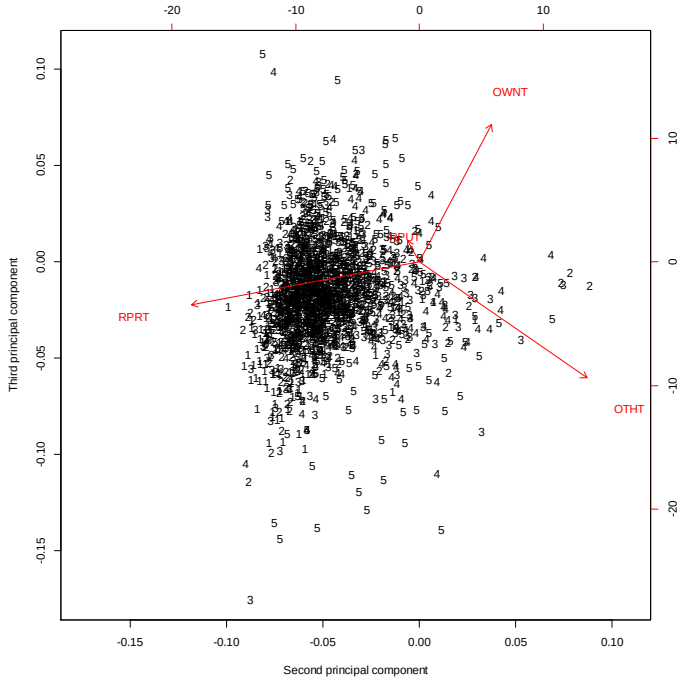

Figure 3-18 Biplot of the second and third components for tenure type (OWNT = own tenure; RPRT = private rent tenure; RPUT = public rent tenure; and OTHT = other tenure)

3.4 Discussion & limitations

To better see how the South-East Queensland (SEQ) and Australia-wide data sets compare in terms of important effects, Table 3-5 summarises the outcomes from the analyses for the various variables in the two data sets. Making comparisons between the output of the South-East Queensland data set and that of the Australia-wide data set is not always possible. This is because not all the variables are directly comparable. For example, income is a variable for which direct comparison between the two data sets is difficult. For the South-East Queensland data set, there are five income classes based on specified dollar figures, whereas in the Australia-wide data set there are four income classes based on unspecified dollar figures. This makes it difficult to undertake direct comparisons between the two data sets. Having said this, the higher energy users tended to be in the higher income brackets than the lower energy users, who tended to be in the lower income brackets.

In terms of the number of bedrooms in a dwelling, in the South-East Queensland data one or two bedrooms tended to equate to lower energy usage. There was no comparable variable for the NEM data.

From the results of the tenure type for the Australia-wide data, it would appear that those with private rent tenures may generally use less energy than those with "other" tenure, and those with "other" tenure generally used less energy than those with "own" tenure.

In terms of building type for South-East Queensland, semi-detached dwellings appeared generally to have had higher energy usage than flats and separate houses, while flats may have used less energy than separate houses.

In terms of floor area of different building types, the two data sets gave broadly similar results. The analyses for both data sets suggested that, in general, separate houses may have used more energy than semi-detached dwellings. For the apartments, the story is not as clear, but two-storey apartment buildings tended to use less energy than separate houses in both data sets.

The impact of building code changes can only be indirectly inferred at best, as the Building Exposure data set provides information on the number of buildings built as five-year blocks in recent years. This means that a major building change could have been implemented in the middle of one of these block periods. Hence, for that block period, houses would have been built based on two or more building codes. Despite this, the work here suggests that the time blocks do provide some guide for understanding whether construction codes have had a significant impact on energy usage. However, it was not possible to include all variables that may impact energy usage. Interestingly, for both data sets, the buildings built prior to 1952 appear to have used the least amount of energy. There could be a number of factors at play here, such as the size of these houses, which tend to be smaller than the more modern buildings; they may not have central heating and/or cooling; the type of material used may have an impact on energy use; and complex interactions between these and other variables that were not captured in the modelling, as the introduction of interaction terms could result in instabilities in the analysis algorithms. For the South-East Queensland data, the houses built between 2007 and 2011 may have used less energy than the houses built between 2002 and 2006. For the Australia-wide data set, the suggestion is that the houses built between 2007 and 2011 used less energy than those built between 1987 and 2001, where there were no national construction code standards.

For the South-East Queensland data, buildings with fibre cement walls appeared to use less energy than those built with timber or cavity and solid masonry. For the Australia-wide data, however, the analysis suggests that houses with fibre cement walls appeared to use more energy than those built with timber or cavity and solid masonry. It could be that these materials have different thermal properties under different climatic and environmental conditions, or that having finer grained data for South-East Queensland made it possible to identify differences that may exist as the seasons change. Further, for the Australia-wide data, houses with timber and veneer masonry walls in general appeared to have higher energy usage than the houses whose walls were built with cavity and solid masonry.

For the Australia-wide analysis, metal sheeting roofs appears to have resulted in the least energy usage, while imitation tiled roofs may have led to the highest energy usage.

For the SEQ data, the results above suggest that energy usage is related to both income and number of bedrooms. To try to get a better handle on how the two variables relate to energy usage together, a biplot for the first two components of the two sets of variables is presented in Figure 3-19. An examination of this figure revealed that the first principal component predominantly splits residences into those with two bedrooms and those with four bedrooms. The second principal component predominantly contrasts the two lowest income classes and the three bedroom residences with the two highest income classes (and to a lesser extent residences with two bedrooms). The one-way analysis of variance of the first principal component scores revealed that the postcodes in the lowest quintile of energy usage had significantly lower first principal component scores than the postcodes in the other four energy usage quintiles. Furthermore, the postcodes in the second and third lowest energy usage quintiles also had significantly lower first principal components scores than the postcodes in the highest energy usage quintile. This suggests that the postcodes in the lowest energy usage quintile were relatively more aligned to residences with two bedrooms than the postcodes in the other four energy usage quintiles, and that the postcodes in the second and third lowest energy usage quintiles were also relatively more aligned to residences with two bedrooms than the postcodes in the highest energy usage quintile. An analysis of variance of the second principal component scores revealed that the postcodes in the middle energy usage quintile had significantly lower scores than the postcodes in the lowest energy usage quintile. This implies that the postcodes in the lowest energy usage quintile had proportionately more households on the two highest income classes, while the postcodes in the middle energy usage quintile had proportionately more households on the two lowest income classes and residences with three bedrooms.

These results for the first two principal component scores and a closer examination of Figure 3-19 revealed that there is a complex interaction between income and number of bedrooms. Many of the postcodes in the lowest energy usage quintile appear in the upper left quadrant of the graph, bracketed by both the highest income class (and to a lesser extent second highest income class) and by residences with two bedrooms (and to lesser extend one bedroom). This suggests that these postcodes had proportionately more households on high income classes and residences with a small number of bedrooms. At the same time, there are many postcodes in the lowest energy usage quintile that appear in the lower left quadrant of the graph that is bracketed by the two lowest income brackets (and to a lesser extent residences with three bedrooms) and by the residences with two bedrooms (and to a lesser extent residences with one bedroom). This suggests that some of the postcodes in the lowest energy usage quintile had proportionately more households on the lower

income brackets and the residences with small number of bedrooms. Hence the postcodes in the lowest quintile of energy usage were predominantly composed of residences with a small number of bedrooms, but that could span a wide variety of income brackets. Further information would be required to tease out what differentiates the postcodes in the lowest energy usage in the upper left and lower left quadrants.

To see how building age may be related to the size of the different dwelling types and the number of bedrooms for the SEQ data, a joint biplot analysis of these three compositional variables was carried out. The first, third, fourth and fifth principal components for these three sets of variables are plotted in Figures 3-20 and 3-21. (The other principal components did not have any significant patterns so are not displayed). An examination of Figure 3-20 shows that the first principal component contrasts buildings built from 1952 to 2001, separate houses and four bedroom dwellings with buildings built before 1952 and to a lesser extent two bedroom dwellings. A one-way analysis of variance of the first principal component scores identified that the postcodes in the lowest quartile of energy usage had proportionately more buildings built prior to 1952 than the postcodes in the four highest quartiles of energy usage. In addition, the postcodes in the second lowest quartile of energy usage had proportionately more buildings built prior to 1952 than the postcodes in the highest quartile of energy usage. Further examination of Figure 3-20 reveals that the third principal component contrasts buildings built between 1987 and 2001 with those built between 2007 and 2011. The one-way analysis of variance of the third principal component scores demonstrated that the postcodes in the lowest quartile of energy usage had proportionately marginally more buildings built between 1987 and 2001 than the postcodes in the second lowest quartile of energy usage.

The plot in Figure 3-21 reveals that the fourth principal component contrasts two bedroom dwellings, semi-detached dwellings, two storey apartments and buildings built in the periods 1987 to 2001 and 2007 to 2011 with separate houses and buildings built before 1952. The one-way analysis of variance of the fourth principal component scores revealed that postcodes in the lowest quartile of energy usage had proportionately more semi-detached dwellings, two bedroom dwellings, two-storey apartments and buildings built in the periods 1987 to 2001 and 2007 to 2011 than the postcodes in the second, third and fifth quartiles of energy usage. The fourth principal component is a contrast between dwellings with three bedrooms and dwellings with four bedrooms (and to a lesser extent buildings built between 2002 and 2006). The one-way analysis of variance demonstrated that postcodes in the highest quartile of energy usage had proportionately more dwellings with four bedrooms (and to a lesser extent buildings built between 2002 and 2006) than the postcodes in the lowest quartile of energy usage.

In Figure 3-20 it is evident that there is a relationship between separate houses and four bedroom dwellings and to a lesser extent with homes built between 1952 and 1986. In addition, there appears to be an association between buildings built prior to 1952 and two bedroom dwellings. From Figure 3-21 there is some association between separate houses and buildings built prior to 1952. Further, there is some relationship between buildings built between 2007 and 2011 and those built between 1987 and 2001. Semi-detached dwellings and one bedroom dwellings also appear to be related in some way. The semi-detached and one bedroom dwellings also appear to be related to two- or more storey apartments and buildings built between 1952 and 1986, although the association may not be as strong as that between semi-detached dwellings and one bedroom dwellings. Semi-detached dwellings also appear to be relatively weakly related to two bedroom dwellings. There appears to

be some sort of association between buildings built between 2002 and 2006 and four bedroom dwellings.

A complicating factor for the analyses in this report is that both the Building Exposure and the Energex data sets have variables that are estimates. For example, the Energex data set includes ABS-derived estimates of the number of residences with different numbers of bedrooms, and the number of residences that fall into particular income bands. Similarly, the Building Exposure data has been collated from a number of sources [3] which means that the breakdown of the number of residences from different time periods are estimates and may not always be up to date. The analyses in this report can be improved by accounting for the fact that some of the variables are estimates. Methods exist to undertake regression analyses without compositional variables [9], but new methods would need to be developed when compositional variables are included in the regression analysis.

The data analysed in this report are highly aggregated both in space and time. It would be preferable to have individual residential data that contains information about the time the building was built, under which building code it was built, building material used, external and internal temperature, electricity tariffs for the residence and socio-economic variables for the household. This would increase the likelihood of identifying any differences in energy usage due to different building codes, while controlling for other variables such as temperature and socio-economics. The surveying work of EUDM will be a critical input into future analyses here.

It is worth noting, however, that the analysis of temporally fine-grained data can be computationally demanding. Given a sufficiently large data set (i.e. multiple years of smart meter data for households), the application of multi-core parallel processing techniques and the use of new algorithms that can efficiently invert large matrices via approximation (or by taking advantage of matrix sparseness) may be necessary.

Table 3-5 Summary of effect of different variables for the Australia-wide and South-East Queensland data sets

Variable	Description of effect
Income	For both SEQ and Australia-wide • Higher energy users tended to be in higher income brackets • Lower energy users tended to be in lower income brackets.
Floor area of building types	For both SEQ and Australia-wide • Separate houses tended to be associated with greater energy usage than semi-detached dwellings • Two-storey apartments tended to be associated with lower energy usage than separate houses.
Building age	Both SEQ and Australia-wide • Buildings built prior to 1952 tended to be associated with less energy usage than buildings built after 1951. SEQ • Buildings built between 2007 and 2011 tended to be associated with lower energy usage than buildings built between 2002 and 2006. Australia-wide • Buildings built between 2007 and 2011 tended to be associated with lower energy usage than buildings built between 1987 and 2001.
Wall material	SEQ • Houses with fibre cement walls generally used *less* energy than houses with timber or cavity and solid masonry walls. Australia-wide • Houses with fibre cement walls may have used *more* energy than houses with timber or cavity and solid masonry walls • Houses with timber or veneer masonry walls may use more energy than houses with cavity and solid masonry walls.
Roof material	SEQ • No obvious differences detected. Australia-wide • Houses with metal sheeting roofs may have used the least energy, while houses with imitation tiles may have used the most.
Number of bedrooms	SEQ

Variable	Description of effect
	• 1 or 2 bedroom houses tended to be associated with lower energy usage than houses with more bedrooms.
Building type	SEQ • Semi-detached residences tended to be associated with higher energy usage than flats and separate houses. • Flats tended to be associated with lower energy usage than separate houses.
Tenure type	Australia-wide • Private rent tenure tended to be associated with lower energy usage than other tenure • Own tenure tended to be associated with greater energy usage than other tenure.

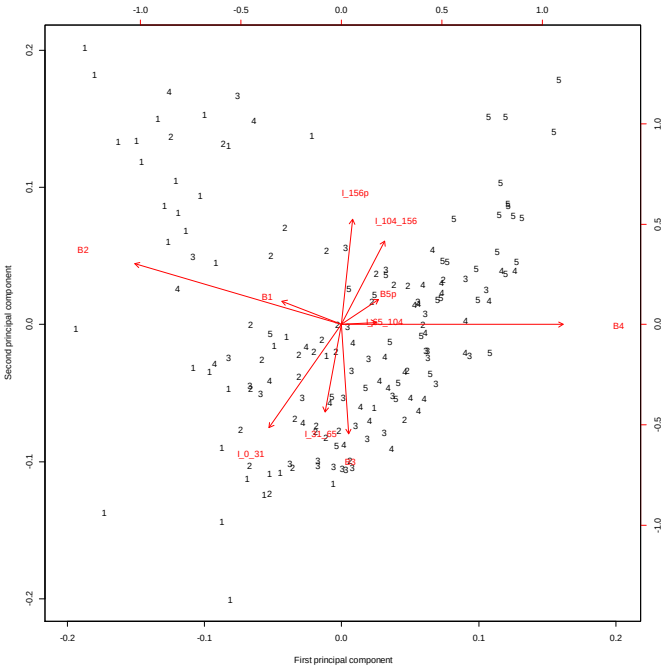

Figure 3-19 Biplot of first two components for number of bedrooms (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms) and income classes (I_0_31 = income <$31,000; I_31_65 = $31,000 ≤ Income < $65,000; I_65_104 = $65,000 ≤ Income < $104,000; I_104_156 = $104,000 ≤ Income < $156,000; and I156p = Income ≥ $156,000)

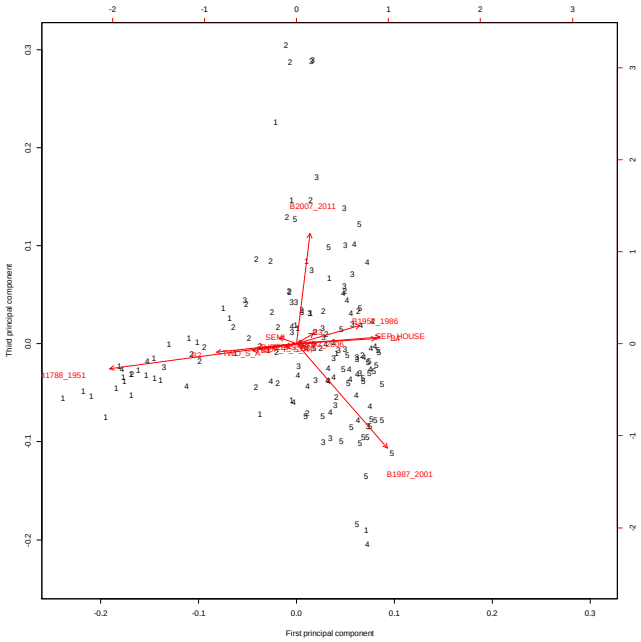

Figure 3-20 Biplot of the first and third components for number of bedrooms (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms), floor area for different dwelling types (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two-storey apartments; THREE_S_A = area for three-storey apartments; and FOUR_P_S_A = area for four- or more storey apartments), and building age classes (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011)

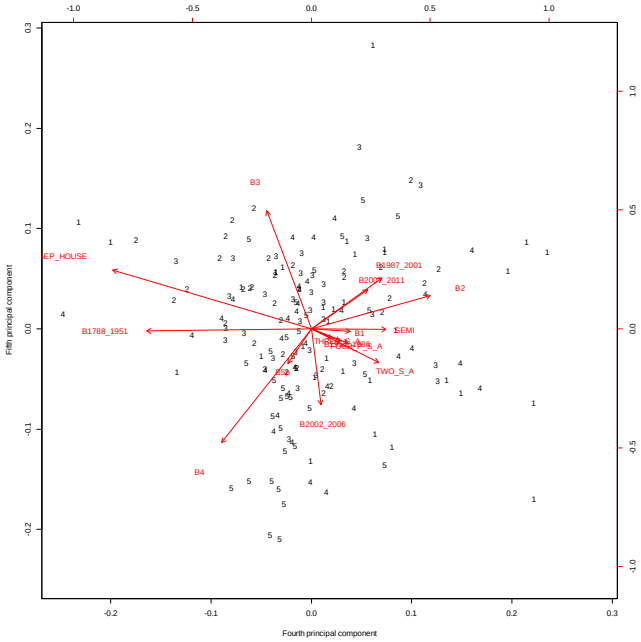

Figure 3-21 Biplot of the fourth and fifth components for number of bedrooms (B1 = 1 bedroom; B2 = 2 bedrooms; B3 = 3 bedrooms; B4 = 4 bedrooms; and B5+ = 5 or more bedrooms), floor area for different dwelling types (SEP_HOUSE = area for separate houses; SEMI = area for semi-detached dwellings; TWO_S_A = area for two-storey apartments; THREE_S_A = area for three-storey apartments; and FOUR_P_S_A = area for four- or more storey apartments), and building age classes (B1788_1951 = built between 1788 to 1951; B1952_1986 = built between 1952 and 1986; B1987_2001 = built between 1987to 2001; B2002_2006 = built between 2002 and 2006; and B2007_2011 = built between 2007 and 2011)

3.5 Summary & next steps

The analyses in this section have demonstrated that a number of variables including building material, characteristics and age, and socio-economic and climatic variables explain a significant amount of the variation in residential energy consumption. This was found to be the case for both the South-East Queensland and a national data set. Importantly, one data set was of continental spatial extent and the other covered a much smaller spatial extent but had more extensive temporal coverage. This suggests that these variables are related to residential energy consumption across different spatial and temporal scales.

The Australia-wide data set was collected annually and the South-East Queensland data set monthly. With modern smart meter technology, it is possible to collect data at much finer temporal scales, such as every 30 minutes. Having access to such data, plus socio-economic, climatic and building related variables – such as the time the building was built, under which building code it was built, building material used and electricity tariffs for the residence – may make it possible to identify how these variables may impact energy consumption under dynamic climatic and environmental conditions across a day and year. Much of this information is currently being collected as part of an extensive sampling program being conducted in Victoria and Western Australia. Future work could look at analysing this sampling data with a view to relating fine-grained temporal energy consumption data to building information and other variables.

3.6 References

[1] NCC (2016). National Construction Code: Complete Series

http://www.abcb.gov.au/Resources/Publications/NCC/NCC-2016-Complete-Series.

[2] Productivity Commission (2005). The Private Cost Effectiveness of Improving Energy Efficiency. Inquiry Report 36

[3] Geoscience Australia (2016). National Building Exposure System (NEXIS): Building Exposure Statistical Areas Level 1 (SA1) Aggregated Metadata. http: http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_82220

[4] Steadman, R. G. (1984). A universal scale of apparent temperature. *Journal of Climate and Applied Meteorology*, **33**, 1674 – 1687

[5] Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Monographs on statistics and applied probability. London: Chapman & Hall

[6] R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

[7] Pinheiro, J.C and Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer-Verlag.

[8] K. R. Gabriel (1971). The biplot graphical display of matrices with application to principal component analysis. *Biometrika* **58**, 453–467

[9] Carriquiry, A. L. (2001). Measurement Error Models. *International Encyclopedia of the Social & Behavioral Sciences*, Volume **14**, 9435 – 9442

4 Understanding household space heating behaviour – RBEE data set analysis and support vector machine modelling

4.1 Introduction

According to the Australian Government Department of the Environment and Energy, on average, 40% of the energy we use at home is for heating and cooling across Australia [1]. Hence, this section focuses on how people use heating within the household space in order to understand behaviours. For example, which combinations of indoor and outdoor temperatures result in people switching heating on, what influence does the time of day have on this behaviour, and what drives people to switch heating off (e.g. how long it has been on, time of day, or particular indoor or ambient conditions).

For domestic households, reverse-cycle air conditioners (or heat pumps) are the most energy efficient type of electric heater [2]. Here, we focus on households using a reverse-cycle heating system. In Section 4.2, we analyse the RBEE data set of space heating behaviour; in Section 4.3, we model the heating behaviour using a machine learning approach, and compare the modelling results with the RBEE data set; and in Section 4.4 we present a discussion of key findings.

4.2 RBEE data set analysis of space heating behaviour

4.2.1 Analysis method

The RBEE data set includes 111 households that use reverse-cycle AC for heating, with over three years of data available. These data sets were used to analyse heating behaviour. We first implemented a data pre-process. For each household, the following information was recorded: TA (ambient temperature), PAC (air-conditioner power), and TB (indoor temperature), with a 30-minute sample rate. To study heating behaviour, we chose data from targeted 'cold' days based on the maximum external temperature; that is, days with a maximum temperature lower than a threshold.

For the reverse-cycle AC, the 'cold day' heating behaviour was assumed to be related to key actions such as 'switch on' and 'switch off'. Similarly, as in [3], the switch on action was assumed to have occurred if the AC energy usage was below a nominal level for the preceding 30-minute interval, and above a nominal level for the interval in question. The switch off action was defined similarly. For each identified switch on and switch off action, the corresponding parameters were recorded; for example, time t (within 24 hour range), indoor temperature TB, ambient temperature TA and duration of the switch on period τ. Based on this information, we studied the statistical properties of the heating behaviour. The results are given below.

4.2.2 Results

First, we studied the relationship between the frequency of the switch on and off actions and the indoor and ambient temperatures.

Figure 4.1 is a histogram of the numbers of actions versus ambient temperature (top) and indoor temperature (bottom). In the top figure, the distributions of the *frequency* of AC switch on and off are similar. Conversely, the bottom figure shows a clear shift in the distribution of indoor temperature for switch on and switch off events. That is, on average the indoor temperature was approximately 5°C warmer at the time when the AC was switched off as compared with when it was switched on.

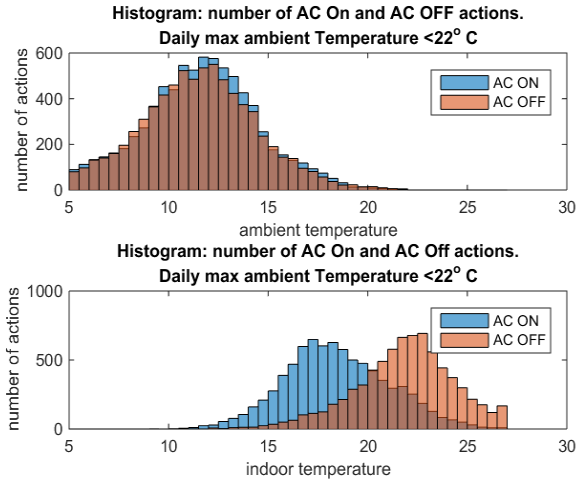

Figure 4-1 Histogram of the numbers of AC switching on and switching off actions versus ambient temperature (top) and indoor temperature (bottom)

Figure 4.2 shows a histogram of the numbers of actions versus time *t* within a 24-hour range [0, 24], where the blue bars represent the switching on action, and the orange bars represent the switching off action. Both switching on and switching off actions had two peak periods during the 24-hour range, one in the early morning and one in the late afternoon. In both peak periods, the switching off action was about 1.5 to 2 hours later than the switching on action. Also, most households tended to switch off the AC late at night, even though the indoor and ambient temperature decreases during the night.

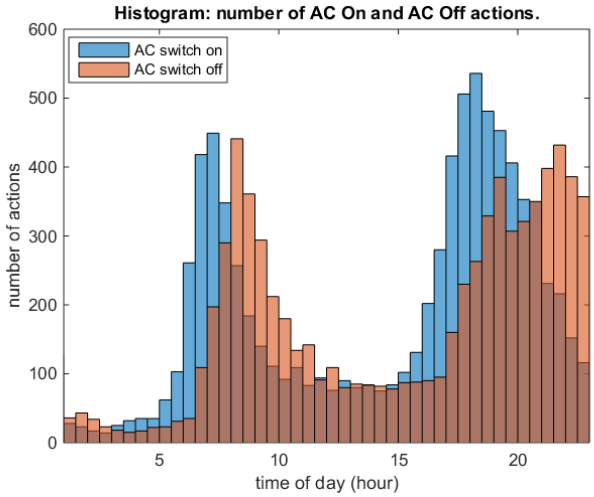

Figure 4-2 Histogram of the numbers of AC switching on and switching off actions versus time of the day

In addition to the *frequency* of the switch on and switch off actions, we are also interested in their respective *probabilities*. Since the amount of time that the AC is on and off are generally different,

equal action frequencies do not correspond to equal probabilities of the actions occurring. For the development of models for use in simulation tools, probabilities are typically most useful.

Thus we studied the switch on and switch off probabilities as a function of apparent ambient and apparent indoor temperatures. The *a priori* probability that the AC is switched on or switched off in any given 30-minute interval for any one randomly selected household is defined as the mean ratio of the frequency of the action to the total frequency of all possible actions for that state. The possible actions are either leaving the AC in the current state, or changing the state. We considered days where the maximum ambient temperature was less than 18°C.

The results are shown in Figure 4-3 along with polynomial fits through the data according to Eq. (1). The coefficients for these polynomials are given in Table 4-1.

$$p = \sum_{i=0..m} a_i T^i \tag{1}$$

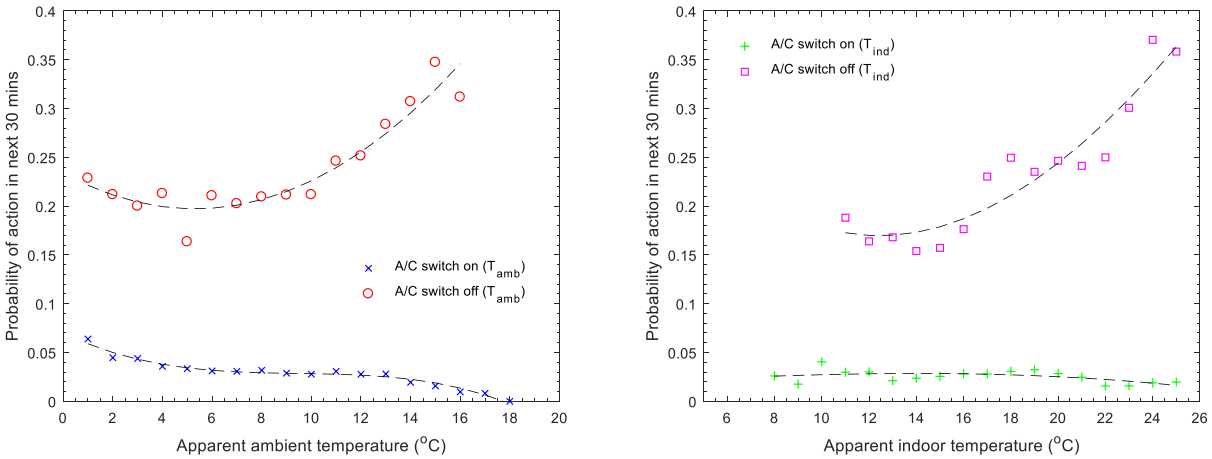

Figure 4-3 Probability of air-conditioner (A/C) 'switch on' action and 'switch off' actions as a function of apparent ambient temperature (left) and apparent indoor temperature (right) for space heating.

The AC switch on probability gradually increased as the apparent ambient temperature decreased. However, the switch on probability was largely independent of apparent indoor temperature. For switch off behaviour, there was also a trend of increasing probability as both indoor and ambient apparent temperatures were increased. However, particularly for ambient temperature, there is an indication that higher switch off probabilities may occur for very cold temperatures. This may be related to households switching off heaters before going to bed (as we will see next).

Table 4-1 Coefficients for polynomial fits of air-conditioner usage probability according to Eq. (1)

Event	Apparent ambient temperature fit coefficients	Range	Apparent indoor temperature fit coefficients	Range
Switch on	-4.4625E-5	2 to 18°C	-1.5809E-6	7 to 25°C
	1.2681E-3		-1.7432E-5	
	-1.2402E-2		1.4243E-3	
	7.0121E-2		1.6308E-2	
Switch off	1.3002E-3	2 to 18°C	-1.5250E-5	10 to 25°C
	-1.3800E-2		1.9857E-3	
	2.3382E-1		-4.2315E-2	
			4.1817E-1	

Air-conditioner switch on/off behaviour was also found to relate to the time of day. The switch on and switch off probabilities are shown in Figure 4-4 as functions of the time of day and apparent ambient temperature. Each line shows the probabilities for apparent ambient temperatures in a given range so that the influence of time of day can be separated from temperature.

There are clear peaks in switch on probability in the morning around 6:30 am and late afternoon around 5 pm. The afternoon peaks are generally higher than the morning peaks. The switch off probabilities show corresponding dips at the same periods in the morning and afternoon. The highest probabilities tend to occur at around 9 am and between 10 pm and 2 am, suggesting that occupants are turning off heating, perhaps before leaving the house and, either before going to bed, or perhaps shortly thereafter (for example using a timer). Additional data is required to test these assumptions.

Overall, the switch on probability variation with time of day is relatively consistent for different temperature levels; though with higher switch on probabilities for cooler temperatures. For the variation of switch off probabilities, there appears to be less variation with temperature for the peak switch off times, possibly suggesting that for switching off AC as these times (9 am and 10 pm to 2 am), routine is more of an indicator than the actual ambient temperature. Once again, further data and analysis is required to test these assumptions.

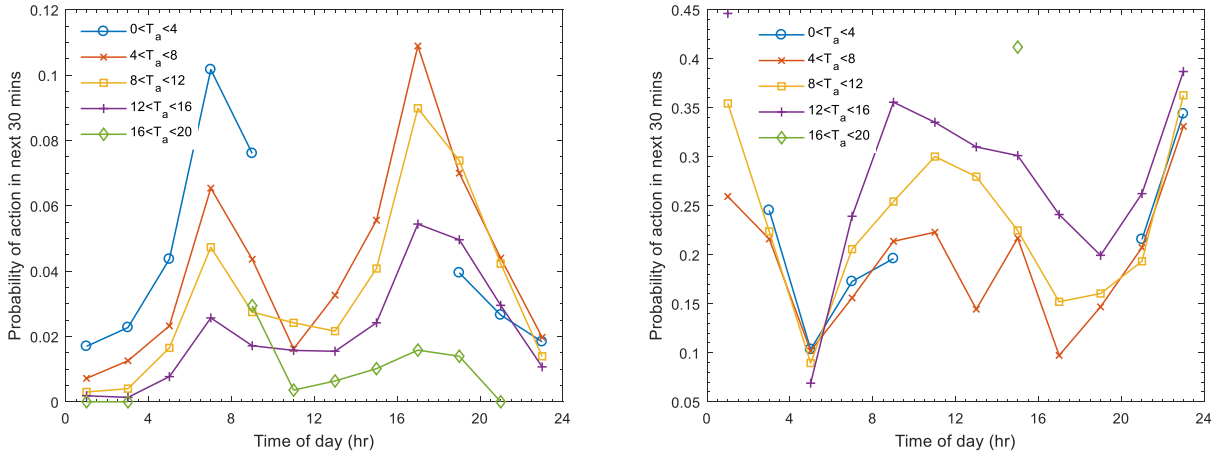

Figure 4-4 Probability of air-considitoner 'switch on' action (left) and 'switch off' action (right) as a function of time of day separated by apparent ambient temperature (T_a).

Finally, we studied the duration of AC running time, as shown in Figure 4.4, which presents the average cumulative probability that the AC has been switched off since the AC was switched on, as a function of hours. The figure shows that around 65% of air-conditioners operated for less than 2 hours before being switched off, and few air-conditioners continually operated for more than 6 hours.

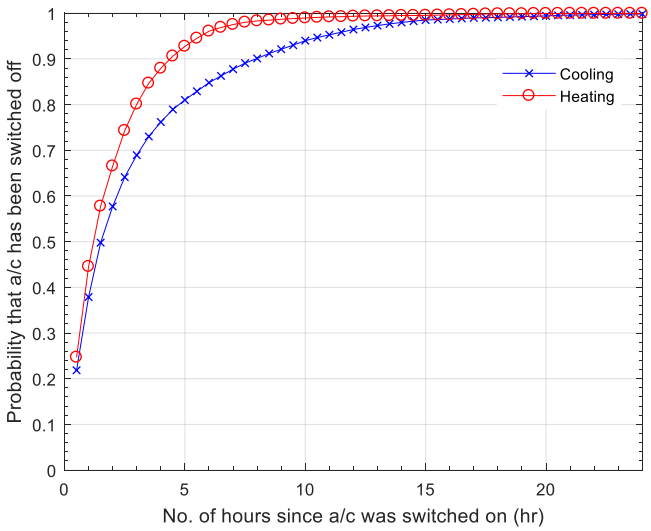

Figure 4-5 Probability function for AC switch off behaviour as a function of duration switched on

4.2.3 Comparison with space cooling behaviour

A comparison of our findings with those of Goldsworthy's study [3] of cooling behaviour using the same RBEE data set showed similarities and differences between the cooling and heating behaviours. Plots of the 'switch on' and 'switch off' probabilities for space cooling are repeated here in Figure 4-6 for reference.

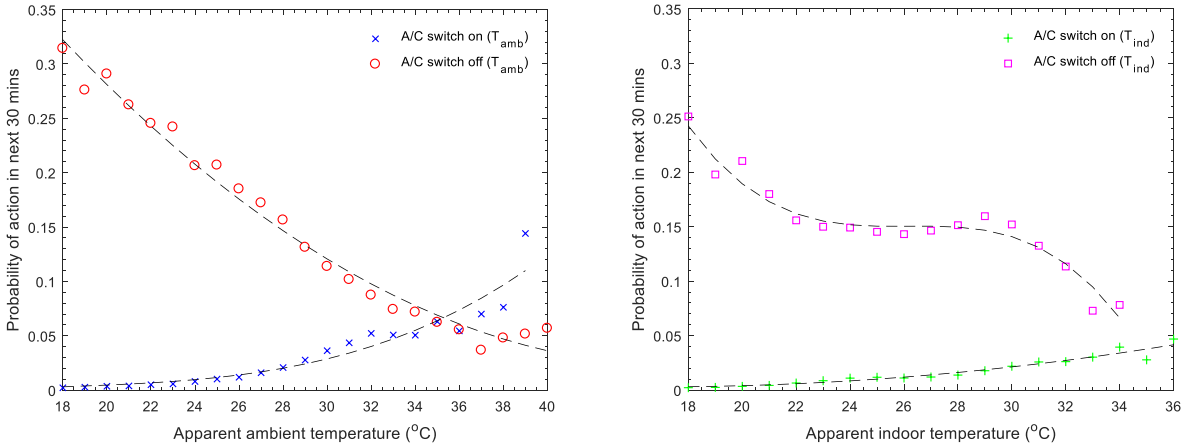

Figure 4-6 Probability of air-conditioner (A/C) 'switch on' action and 'switch off' actions as a function of apparent ambient temperature (left) and apparent indoor temperature (right) for space *cooling*.

For space cooling, there was a strong relationship for both the switch on and switch off probabilities for both apparent ambient and indoor temperatures, though particularly for the former. For space heating, the probabilities showed a significantly smaller temperature dependence.

There were also differences in the influence of the time of day on probabilities of actions. For the cooling behaviour, when the effect of temperature was removed, the probability of switching on AC was not strongly tied to the time of day. However, for space heating behaviour, there were clear peaks in switch on probability in the morning and evening, even accounting for the typically cooler ambient temperatures during these periods. For the switch off probability, the behaviour did show peaks in the afternoon and morning, though these were more spread out over several hours than for the heating behaviour, and the pattern over the day was more consistent when compared at equal ambient temperatures.

A comparison of the duration of AC running time for cooling and heating reveals that households tended to use heating for much shorter durations on average. For example, after 5 hours, 20% of air-conditioners were still on when operating in cooling, but only approximately 7% of air-conditioners operating in heating.

4.2.4 Summary

There is a clear need to develop improved residential air-conditioner usage models to understand and predict heating energy consumption in buildings. In order to model the heating behaviour, we studied the occupant initiated actions of switching AC on and off[3] and how these relate to the indoor and outdoor apparent temperatures, the time of day and the duration of AC heating. Along with quantitative data from which to build models, some other interesting findings were the presence of two peak switch on periods, and the much smaller variation of switch on probability with temperature as compared with space cooling behaviour.

[3] We have compared the recorded energy over the data measurement interval of 30 minutes to a threshold to decide whether the AC was on or off for that period. If the AC cycled on/off over a longer duration than 30 minutes, this would be incorrectly counted as a user controlled event. It is expected that the number of these instances would be small, as typically on/off compressor cycles are shorter than 30 minutes. Indeed, increasing the comparison interval to 1 hour did not significantly change the results, suggesting that short duration on/off cycles are correctly being discounted.

4.3 References

[1] *http://yourenergysavings.gov.au/energy/heating-cooling/understand-heating-cooling*

[2] *http://yourenergysavings.gov.au/energy/heating-cooling/understand-heating-cooling*

[3] Mark Goldsworthy, Towards a residential air-conditioner usage behaviour model for Australia, 2016

5 Linking building indoor and outdoor temperatures – generic predictive model

5.1 Introduction

As demonstrated in the preceding chapter, the internal building temperature is a key predictor of air-conditioner usage behaviour. Unfortunately, the measurement of internal temperatures is expensive and there is a genuine sparsity of relevant data in the field. With this in mind, this chapter explores whether internal temperatures can be estimated based only on a small set of more readily available variables.

Section 5.2 describes the development of generic thermal models for 140 RBEE homes {H}, using a data set of three key predictors (active AC, ambient air temperature and solar radiation) to estimate the indoor temperature. The impact of the built form of a home is then examined separately, exploring how thermal response (and generic thermal models) vary across the RBEES building types. Should a rich understanding of the relationship between the thermal models and built form be established, it represents the first step towards a complete methodology for estimating internal conditions from basic external observations.

5.2 Analysis method

This section describes the activities that have been carried out in line with the objective described in Section 5.1.

5.2.1. Development of a generic thermal model for each house in the data set

This activity included a series of tasks. First, from multiple years of half-hourly time series data, we selected time instances where active heating was unlikely. This selection was made by comparing the ambient temperature at different times of the day to a series of thresholds. The main aim was to avoid days were heating might be used (either by the measured air-conditioner or by other means such as gas heating). This process did not preclude days where no air-conditioning was used.

We then used a time series of predictor variables to estimate the current indoor temperature. The analysis involved regression using a series of values of each variable at the current time and at different times in the past. That is, we used the known values of the ambient air temperature, solar radiation, and AC power at the current and previous 30 minute intervals, along with the predicted values of the indoor temperature at preceding 30 minute intervals, to predict the indoor temperature for the current 30 minute interval. We refer to these previous values of a variable as the “lags” of that variable.

The regression process was used to find the optimal set of multiplying factors (i.e. weights) to apply to the lags so as to minimise the error between the model-predicted and actual indoor temperature.

This is shown mathematically in Eq. 1, where *TA* is air temperature, *PAC* is AC power, *G* is solar radiation and output *TB* is the modelled indoor temperature. The lag operator, *L*, is intended to include up to six lags (i.e. three hours) of each predictor only where the overall prediction error is reduced. The weights matrix, *W*, is the solution for each house. A mathematical technique known

as a 'feed forward delta rule' was used to draw matrix *W* in a linear fashion to avoid overfitting the model.

$$TB_t = \left(\sum_{i \in I} L^i TA_t + \sum_{j \in J} L^j PAC_t + \sum_{k \in K} L^k G_t \right) . W \; ; \; I, J, K \subset \{0 \dots 6\} \quad (1)$$

The analysis was repeated for all possible lag combinations (i.e. the inputs in Eq. 1) for each house. The optimal set of predictors was thus obtained for each house based on a minimised 99% confidence interval of the modelling error. Figure 5-1 shows the modelling error for two example houses.

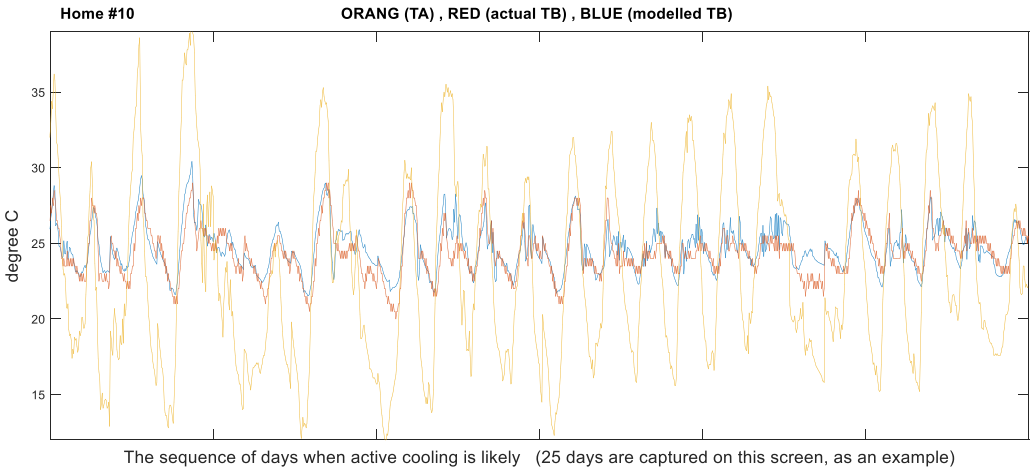

Figure 5-1 Modelling TB for house #10 (-0.9 to -0.7 °C).

The modelling error of each home is calculated as the mean absolute distance between actual and modelled TB. Under the assumption of normality of error, the 99% confidence intervals for the example house #10 is bound within -0.86 and -0.73 degrees. Median error for all 140 homes is bound within -0.06 and 0.01. The widest error range for all homes is -2.53 to 1.48 degrees.

5.2.2. Clustering houses by similarity of their thermal responses

To map thermal model to building type, we begin by grouping houses with similar thermal models. That is, we identified groupings of houses whose indoor temperature would respond in approximately the same way if the same ambient weather conditions and air-conditioner usage were present. Formally, home Y was said to be similar to home X if X's predictor variables applied to Y's thermal model returned X's response variable.

We expected that home Y, being similar to home X, would logically result in X being similar to Y – referred to as mutual adjacency. To check this, we calculated pairwise similarities for all homes. The resulting |H|x|H| adjacency matrix was closely symmetric, thus confirming our assumption.

The next step was to apply a 'swift' clustering to the adjacency matrix so that homes with similar thermal responses would be grouped into a unique cluster. We used a technique known as Principal component analysis (PCA) which effectively reduces and combines variables whilst still retaining most of the information. Figure 5-2 (a) shows the amount of information (%) captured by the inclusion of each additional Principle component. We can see that almost 81% of the information is captured by principle components 1 and 2.

The final step was to perform the actual clustering. This was carried out using a standard k-means clustering, based on the cosine similarity of the houses' scores. A good separation of house behaviours was obtained with 2 clusters. This can be seen in Figure 5-2 (b), which plots a measure known as silhouette width. A silhouette score of 1 indicates perfect clustering and -1 worst-case clustering. A score of 0.69 was achieved, which indicates good performance; that is a relatively clear separation of houses into either one or the other grouping.

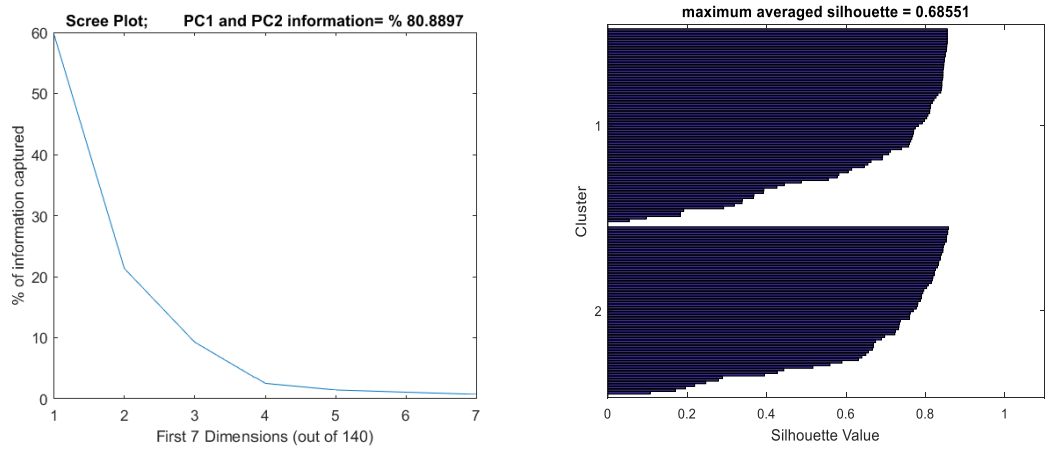

Figure 5-2 (a) Scree plot showing information captured on the first seven components, (b) maximised averaged silhouette width determining two clusters

The above clusters were based on the similarity of the houses' thermal *responses*. However, it is important to confirm that the underlying thermal models themselves are similar if they are within the same cluster and suitably distinct/different otherwise. To verify this, we first needed to obtain a representative thermal model per cluster. One way to achieve this is to take the average of the thermal models of all intra-cluster members for each cluster. Figure 5-3 shows averaged weights that link each input to the model output (*TB*) per cluster. The rows indicate a representative thermal model of the h=2 clusters, and the columns correspond to the inputs.

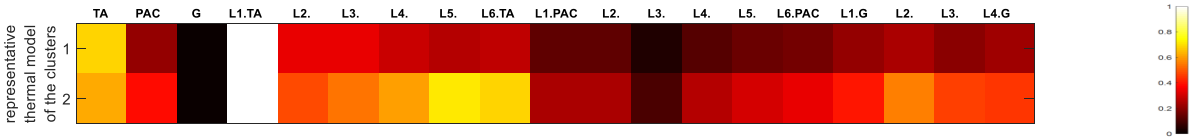

Figure 5-3 Sensitivity of the representative thermal models of the clusters to the input variables

It is apparent that both clusters were similar in terms of high sensitivity to *L1*, with *TA* being the dominant predictor, but showed a weak correlation with current solar radiation *G*. Cluster 2 was generally less sensitive to *TA* but more sensitive to *PAC*. Possibly, Cluster 2 is made up of houses with a better insulation or star rating. These conclusions are further discussed in Section 5.3.1.

5.2.3. Clustering houses by their high-level building characteristics

As seen elsewhere in this report, it is clear that building characteristics play a key part in shaping patterns of energy consumption. Hence, the grouping of homes by thermal response (Section 0) should, to some extent, match with the grouping of homes based on their building characteristics. About a dozen data fields related to building characteristics (see Appendix B) are available. We

compressed these features into three summary components, again using the PCA technique. We then selected two distinct clusters that each describe distinct building types (Figure 5-4).

We can now observe how these two clusters based on building characteristics compare with the two thermal model clusters produced in Section 5.2.2.

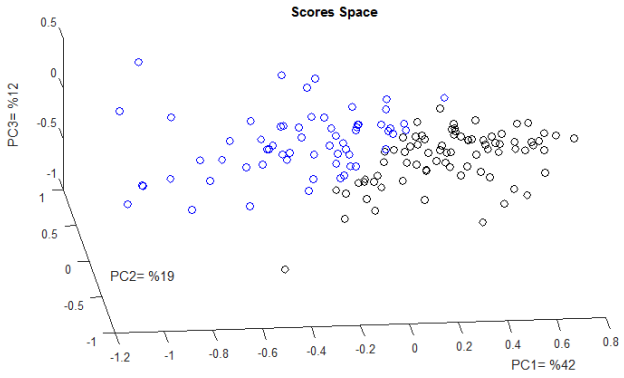

Figure 5-4 Buildings scores as split into two clusters

5.2.4. Matching the building structure and thermal model clusters

We now sought to understand how the thermal model clusters of Section 5.2.2 are related to the built-form clusters of Section 5.2.3. The aim was to understand if thermal models correspond to particular building characteristics. We do this by calculating the Jaccard similarity (J) index between pairs of clusters: one cluster based on its thermal model and another on its building characteristics. A high Jaccard score (nearer to 1) suggests high levels of overlap between the two clusters. That is, a high score means that most of the houses in the thermal model cluster and the built-form cluster are the same. A low score (nearer to 0) suggests low levels of overlap between the two clusters and that the houses in each cluster are different. If the thermal models can map well to built form, we would expect to see high Jaccard scores.

The results of the above analysis for 2, 3 and 4 clusters are given in Table 5-1. It can be seen that there is a clear relationship between thermal response clusters and built-form clusters for the two-cluster case (the left-to-right diagonal has markedly higher scores than the right-to-left diagonal). The relationship is present, but weaker, as the number of clusters grows. It is expected that improved matching between clusters may be possible through alternative approaches to the clustering process, such as drawing on a greater number of principle components and using an alternative distance metric for forming the clusters. As an example, we consider a cosine distance metric for building 8 clusters. Results in Figure 5-5 show good clustering performance, with coherent clusters indicated both visually and with silhouette scores. Table 5-2 shows that, using this new distance metric, correspondence between thermal and built-form clusters is good (though imperfect) even at this finer clustering granularity. Further work on improving performance for larger number of clusters is worthwhile, given that an increase in the number of clusters allows for more exact mapping of thermal models to a greater range of building types.

Table 5-1 Jaccard similarity for 2 clusters (left), 3 clusters (middle), and 4 clusters (right). Columns are built-form clusters. Rows are thermal response clusters. Highlighting shows the strongest correspondences between clusters.

	Cluster 1	Cluster 2

	Cluster 1	Cluster 2	Cluster 3

	Cluster 1	Cluster 2	Cluster 3	Cluster 4

0.45	0.20
0.28	0.44

0.28	0.22	0.14
0.12	0.29	0.15
0.27	0.07	0.28

0.25	0.13	0.07	0.15
0.20	0.22	0.08	0.08
0.05	0.21	0.22	0.06
0.06	0.16	0.15	0.18

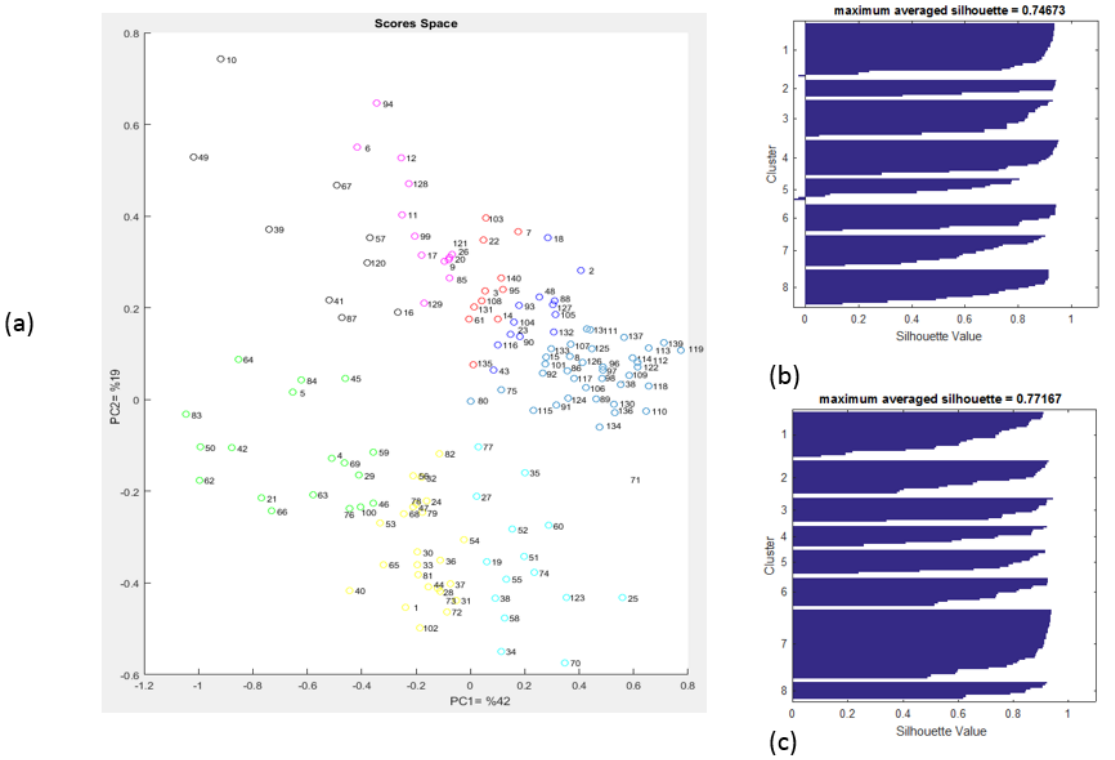

Figure 5-5 (a) 8 building clusters of homes' scores based on compression of 12 building characteristics, (b) silhouette widths for the 8 building clusters, (c) silhouette widths for h=8 thermal model clusters

Table 5-2 Jaccard similarity showing one-to-one matching of 8 clusters. Columns are built-form clusters. Rows are thermal response clusters. Highlighting shows the strongest correspondences between clusters.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8
Cluster 1	0.04	0.04	0.05	0.18	0.08	0.07	0.16	0.06
Cluster 2	0.06	0.08	0.10	0.00	0.00	0.04	0.00	0.13
Cluster 3	0.02	0.23	0.06	0.03	0.00	0.09	0.10	0.04
Cluster 4	0.00	0.06	0.03	0.03	0.23	0.03	0.14	0.04
Cluster 5	0.06	0.00	0.00	0.00	0.09	0.04	0.14	0.05
Cluster 6	0.26	0.06	0.04	0.00	0.04	0.03	0.04	0.00
Cluster 7	0.08	0.09	0.11	0.08	0.03	0.00	0.06	0.08
Cluster 8	0.16	0.00	0.07	0.03	0.00	0.17	0.10	0.00

5.3.1. Linking a thermal predictor p_n to a building characteristic c_m

In the preceding sub-section we produced 8 distinct clusters based on the building characteristics of houses in the RBEES set and 8 distinct clusters based on the thermal response of houses in the RBEES set. Using Jaccard similarity, we saw that there was almost always a built-form and thermal

cluster that had significant overlap in the RBEES houses they contained. Now, we can explore what those similarities mean with respect to understanding how building characteristics and indoor temperature response are related.

First, for two similar clusters, we consider just the homes that are in both clusters and disregard those homes that are different. Next, we can explore typical thermal models for each of the built-form clusters. This will give us a general model that explains how indoor temperature will respond for these building types. To build a general model for a given cluster, we simply look at all thermal models for the given cluster and take the mean of each thermal model parameter from Equation (1). Similarly, we can build a representative building type for the cluster by taking the mean of each building characteristic specified in Appendix B. We now have a general thermal model for each representative building type.

Figure 5-6 (a) shows the link between each of the 8 thermal models and the 19 predictor variables. Similarly Figure 5-6 (b) shows the link between the 8 building characteristic clusters and the 12 building predictor variables. The brighter colour expresses a greater importance (i.e. weight) for the particular variable. Matching of the variables that drive the building thermal response with those that define the building grouping by building characteristic is now possible. Table 5-3 gives a summary of values along rows of Figure 5-6 (a) and (b).

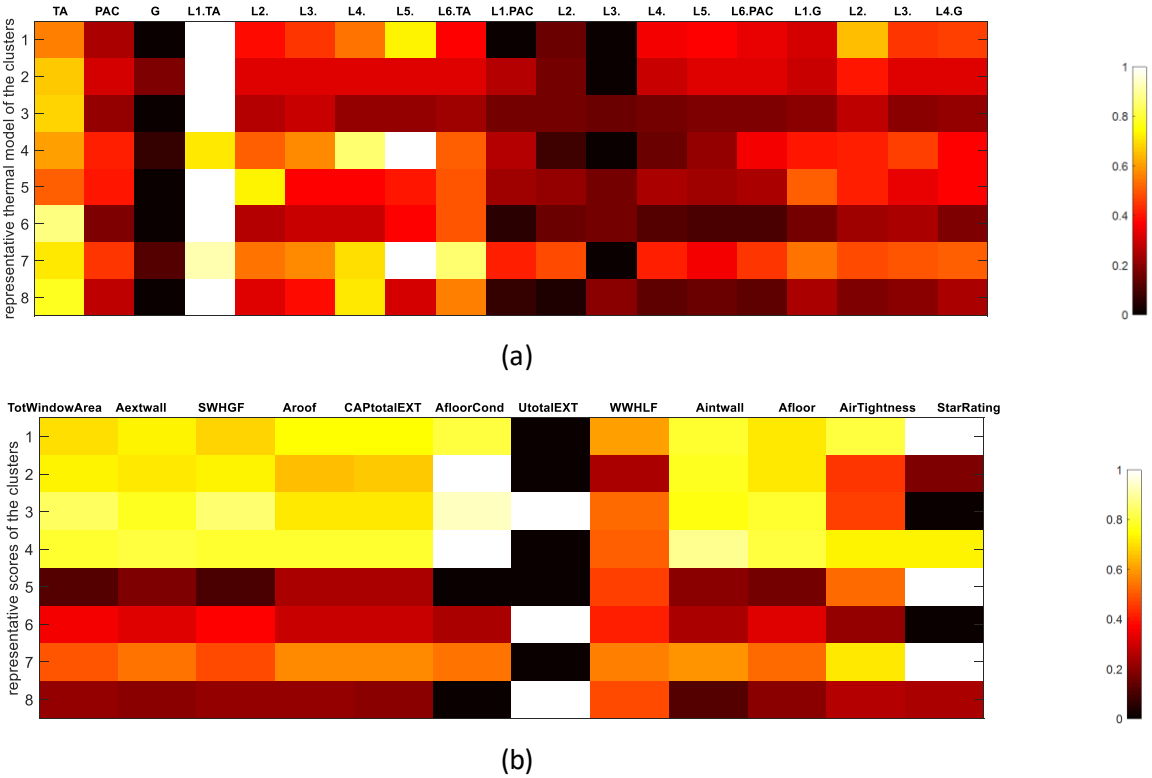

Figure 5-6 Building representative types matched to general thermal models. Rows in (a) correspond to rows in (b). See Appendix B for a description of each building parameter.

Table 5-3 Matching predictors with characteristics. See Appendix B for a description of each building parameter.

Matching cluster pairs	Strongest predictor	2nd best matching predictor	3rd best matching predictor	Worst matching predictor	Best matching characteristic	2nd best matching characteristic	3rd best matching characteristic	Worst matching characteristic
#1	L1.TA	L5.TA	L2.G	G	StarRating	AfloorCond	AirTightness	UtotalEXT
#2	L1.TA	TA	L2.G	L3.PAC	AfloorCond	Aintwall	SWHGF	UtotalEXT
#3	L1.TA	TA	L3.TA	G	UtotalEXT	AfloorCond	SWHGF	StarRating
#4	L5.TA	L4.TA	L1.TA	L3.PAC	AfloorCond	Aintwall	Aextwall	UtotalEXT
#5	L1.TA	L2.TA	TA	G	StarRating	AirTightness	WWHLF	UtotalEXT
#6	L1.TA	TA	L6.TA	G	UtotalEXT	WWHLF	SWHGF	StarRating
#7	L5.TA	L1.TA	L6.TA	L3.PAC	StarRating	AirTightness	Aintwall	UtotalEXT
#8	L1.TA	TA	L4.TA	G	UtotalEXT	WWHLF	AirTightness	AfloorCond

5.3 Summary of outputs, limitations and next steps

This chapter has developed a basic thermal model that estimates, for a given household, indoor temperature from ambient temperature, solar irradiance and air-conditioner use. Though useful, training the model requires historical data from the household to estimate a probable thermal response to conditions. To provide greater generalisation, the focus of much of this work has been on identifying how the properties of a thermal model relate to the properties of the building (and vice-versa). The goal is to develop a general thermal model that can be applied for specific types of built form. To this end, we developed two sets of clusters: one to describe a small set of distinct thermal models, and another to describe a small set of building types. Using Jaccard similarity, we then built a mapping methodology, which suggested a distinct class of thermal model for each of 8 different building types.

Though the process is promising, it is worth underlining that it is preliminary. The data set was small, including only 140 homes, while the intersection of clusters required for the mapping of thermal model to building type reduced this further still to only 41 homes. Clearly, further data is required to build upon this initial study.

The main limitation of the approach, though, is that a perfect mapping from built-form to thermal model is unlikely, since numerous confounding variables exist that are seldom captured in data sets of this type. For instance, thermal models are likely to be influenced by highly localised conditions, including cloud shading, wind gusts, air-conditioner de-ratings, etc. Similarly, the details of the built form are likely only ever to be approximations, as the building envelope invariably shifts over time due to renovations, degradation of building materials, etc. Most importantly, though, human behaviour will remain difficult to capture, including the propensity to open windows, use of window blinds, choice of clothing, etc. Still, in spite of this, this preliminary work shows that the development of at-least-indicative general thermal models based on basic building characteristics is within reach.

It is also worth noting that here we considered only cooling days. Future work could focus on heating days, allowing a comparison between the two different models and resultant clusters, and the formulation of a generalised solution for both cooling and heating. We also plan to revisit the development of the thermal model (Section 5.2.1) to see if alternative mathematical formulations can yield better results. To this end, the single layer feed forward model for generating weights W

(reflecting impact of drivers on indoor temperature) is planned to be replaced with a state of the art multi-layer deep learning tool with heuristic activation function and learning rule.

6 Conclusion

This report describes several separate yet closely related investigations into different aspects of building air-conditioning and occupant behaviour. The work is based on an analysis of existing data sets to deliver new insights. The problem of understanding both how occupants behave in relation to air-conditioner use, and also why they behave in the way that they do (i.e. the key motivations) is complex and nuanced. Occupant behaviour is also closely linked to the thermal response and characteristics of the building, and decoupling these two is not a trivial task. However, as shown here, it is possible to deliver valuable insights that can be used to make better informed decisions in the future. Given that building air-conditioning consumes the largest portion of the energy used in buildings, developing a detailed understanding of how air-conditioners are used is critically important, and the work here is a key step in that direction.

7 APPENDIX A

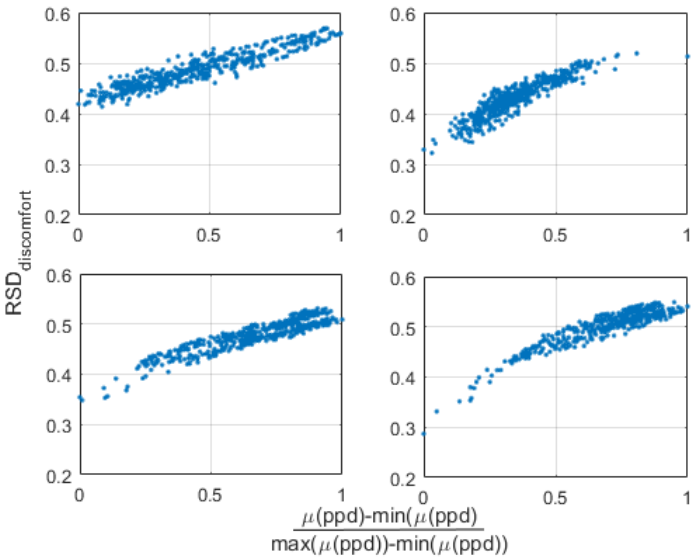

Figure 7-1 Comparison of relative scaled discomfort sensitivity versus normalised percentage discomfort for four climate zones. Clockwise from top left: Darwin, Brisbane, Hobart and Melbourne.

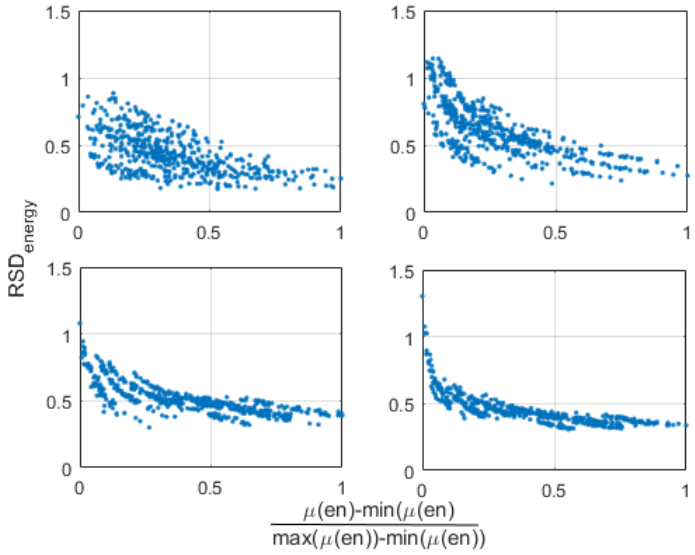

Figure 7-2 Comparison of relative scaled energy sensitivity versus normalised energy usage for four climate zones. Clockwise from top left: Darwin, Brisbane, Hobart and Melbourne.

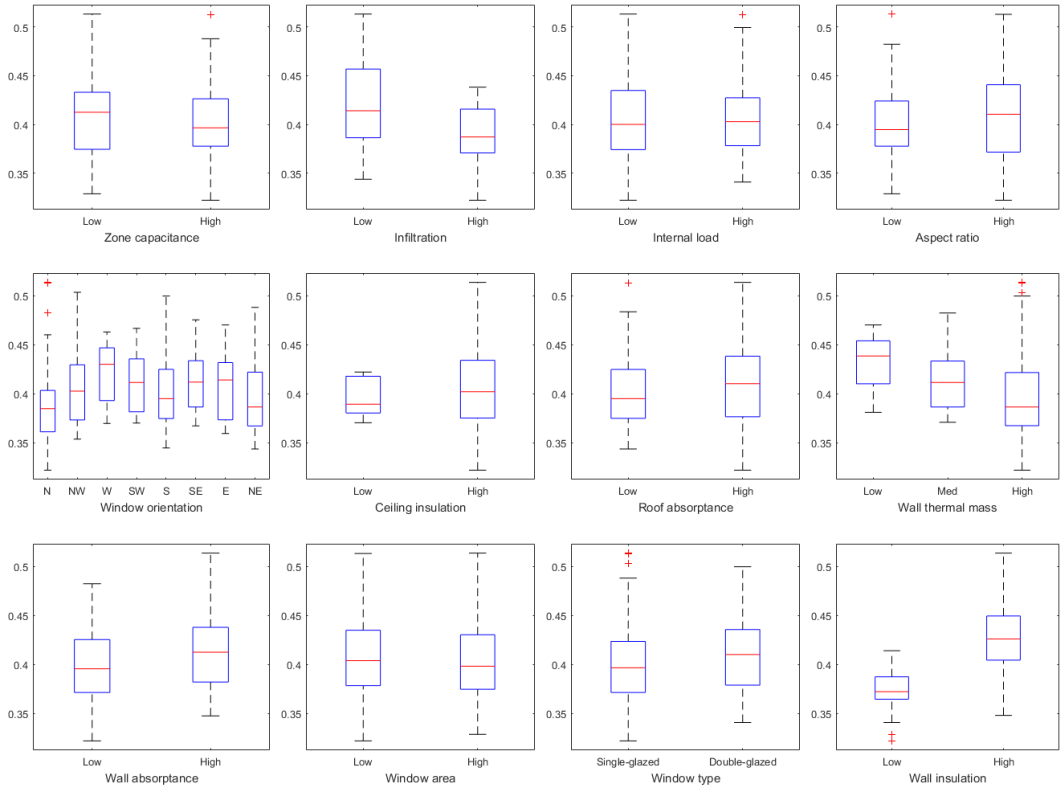

Figure 7-3 Sensitivity of discomfort (RSD parameter) to variations in occupant AC usage behaviour by building thermal parameters for Brisbane (subtropical climate) for buildings with the lowest 30% of energy consumption. Lower numbers indicate less sensitivity.

8 APPENDIX B

Table 8-1 High level building characteristics

Short Name	Units	Description
'TotWindowArea'	m^2	The total surface area of all the windows in the building
'Aextwall'	m^2	The total external wall surface area of the building
'SWHGF'	m^2	Summer window heat gain factor: this is a scale dependent measure calculated from the sum of the product of window area, summer exposure factor due to the window orientation and the window solar heat gain coefficient.
'Aroof'	m^2	The total roof surface area of the building
'CAPtotalEXT'	kJ/K	Total thermal mass of all external walls
'AfloorCond'	m^2	Total air-conditioned floor area
'UtotalEXT'	$W/m^2/K$	Area weighted heat transfer coefficient (U-value) of all external walls
'WWHLF'	$W/m^2/K$	Winter window heat loss factor: this is a relative measure calculated from the ratio of the sum of the window area and heat transfer coefficient to the sum of the area, solar heat gain coefficient and winter solar exposure factor
'Aintwall'	m^2	Total area of all internal walls
'Afloor'	m^2	Total floor area (includes multiple stories)
'AirTightness'	-	An indicator of the building air tightness encoded from a qualitative assessment of the building weather stripping on external windows and doors
'StarRating'	-	The NatHERS Star Rating of the building

CONTACT US

t 1300 363 400
+61 3 9545 2176
e enquiries@csiro.au
w www.csiro.au

AT CSIRO WE SHAPE THE FUTURE

We do this by using science to solve real issues. Our research makes a difference to industry, people and the planet.

As Australia's national science agency we have been pushing the edge of what's possible for over 85 years. Today we have more than 5,000 talented people working out of 50-plus centres in Australia and internationally. Our people work closely with industry and communities to leave a lasting legacy. Collectively, our innovation and excellence places us in the top ten applied research agencies in the world.

WE ASK, WE SEEK AND WE SOLVE

FOR FURTHER INFORMATION

Energy
Mark Goldsworthy
t +61 2 4960 6112
e mark.goldsworthy@csiro.au
w www.csiro.au