VISUALISING THE INVISIBLE

Displaying Building Resource-use Benchmarks in a 3D City Visualisation

By

Alex Josephine Hills

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Building Science

Victoria University of Wellington

2014

Abstract

This thesis proposes a reinvention of the means of presenting statistical data about 3D urban environments. Conventional GIS use of 3D 'enhances' hard to understand 2D maps with even harder-to-understand histograms of data. The goal is to demonstrate the means by which data on energy and water-use in buildings can be used to enhance familiar 3D interactive city environments and be made accessible to the widest possible audience. Ultimately, resource benchmarks and other related publicly available information about the built environment could be presented in this highly accessible form. All information would be database driven, so automatically updateable. From this basic platform, applications that allow people to compare their own private records with public norms are easily constructed: a world where a building owner can compare their energy records with benchmarks for similar buildings and take action to improve if necessary, or to advertise accomplishments.

This study draws on data from the 'BEES' Building Energy End-use Study - a BRANZ research project documenting energy and water use in New Zealand commercial buildings. During the study a 'Websearch' survey was conducted, building a detailed picture of non-residential building stock in New Zealand with data collected on building typologies, characteristics and surroundings. A thorough research methodology was developed to ensure that high level data could be collected from 3,000 randomly selected buildings within the budget allocated for the project. The data was examined for quality, building characteristics and typology mix and a valuable layer of detail was added by inferring additional information from the basic Websearch dataset. Where sub-samples used in the BEES study were subject to refusal / survey participation rates, the level of potential bias in the mix of building typologies could be tested and allowances made. Energy and water use data collected for a random subset of the sample, could then be applied as benchmarks to the census of New Zealand commercial buildings.

In order to trial the communication of the benchmark results to the widest possible audience, an automated 3D city visualisation 'pilot' was generated of the Wellington Central Business District and a number of graphic tools were brought together to make the information publicly accessible and as useful as possible. The overall aim was to test the feasibility of applying this technique at a national level.

The research revealed three major recommendations: firstly, a national unique building identifier is required to ensure the accuracy of national building data and enable statistical results about the built environment to be accurately and reliably applied to real buildings; secondly, resource use data in 3D format is urgently required to improve the value of sustainable properties; lastly, creating a significant impact on building stock efficiency will depend upon the engagement of a wider audience. Developed further, this visualisation will enable construction professionals, building owners, developers and tenants to understand the built environment and implications of building design and typology on energy and water use.

Keywords

Buildings, Energy-use, Water-use, 3D City Visualisation, National Energy Survey, Built Environment, Building Stock Model, Building Resource-use, Augmented Reality, Data Visualisation, Building Typologies

Acknowledgements

Academic Supervisor – Michael Donn:
Michael Donn from the Centre for Building Performance Research (CBPR) has been a great source of encouragement, motivation and support during my last eight years at Victoria University both as a student and as a member of the academic staff. His previous experience researching buildings by trawling through large valuation databases has proved invaluable during the BEES project and his knowledge and passion about digital architecture is vast. This allowed me to push this topic area further than I imagined possible. Not many supervisors could have covered a topic that draws on research from such diverse areas. These spanned from Geographic Information Systems, Automated Digital Modelling, Big Data, Statistics, Visualisation, Building Stock Models, National Building Energy Surveys, Augmented Reality, Environmental Policy and Analytical Graphic Design.

BRANZ Supervisor – Nigel Isaacs:
A great deal of Nigel Isaacs' passion and enthusiasm has rubbed off on me during the BEES study over the last six years. I am particularly grateful for his support developing my statistical analysis skills which were particularly necessary due to the huge quantity of data this study produced. I immensely enjoyed all the animated discussions we had regarding this complex topic area and hope to have many more opportunities to work together again in the future.

The Building Energy End-use Study (BEES) team:
As well as recognising all the staff who worked on BEES, I would like to thank the building occupants, mangers, the staff of Property IQ NZ and the funders who supported this work over the years.
Further information on BEES is available from the BRANZ website: www.branz.co.nz

BRANZ:
Lynda Amitrano, Dr Michael Babylon, Dr Lee Bint, Nikki Buckett, Lisa Burrough, Michael Camilleri, Nigel Isaacs, Andrew Pollard, Johannes Roberti

CRESA:
Kay Saville-Smith, Ruth Fraser

Energy Solutions:
Rob Bishop

Centre For Building Performance Research:
Dr Michael Donn, Alex Hills (Author)

Consulting Statistician:
John Jowett

VUW Scholarship Recipients:
Shaan Cory, Claire Dykes, Anthony Gates, Quentin Heap, Alex Hills, Chi Yao (Henry) Hsu, Hayley Koerbin, James Thompson

Websearch Research Assistants:
Mina Abdel Nour, Shaan Cory, Claire Dykes, Koryn Evans, Anthony Gates, Alana Hamill, Chi Yao (Henry) Hsu, Nicola Jackson, R.D Jamal, Carrie Speirs, Chris S, Victoria Toner, Andrew Turner, Chris Van der Veen, Kate Walker, Kaiyi (Eugene) Zhang

Building Energy End-use Study (BEES) Funding:
I would like to acknowledge the Ministry of Business, Innovation and Employment, the Building Research Levy, and the Energy Efficiency and Conservation Authority, not only for providing the funding which allowed me to undertake this research, but also for supporting a project which is vital in ensuring New Zealand public have the information they need to reduce their carbon dioxide emissions.

School of Architecture and Design, Victoria University of Wellington
Victoria University of Wellington provided academic support and resources for the duration of the masters research and the four years prior to this during my work there, as a lecturer / research fellow in Sustainable Architecture. I would like to thank them for giving me the opportunity to attend conferences and meet so many interesting people.

TABLE OF CONTENTS

ABSTRACT **I**

Keywords **i**

Acknowledgements **ii**

Table of Figures **vi**

1.00: Visualising the Invisible **8**

1.01: Raising Value of Sustainable Developments and Upgrading the Building Stock 8

1.02: Reinventing the Means of Presenting Data about 3D Urban Environments 11

1.03: Representing Geo-coded Statistical Values 13

1.04: Augmented Reality and Virtual World Platforms 15

1.05: Automated Model Construction 16

1.06: Ensuring High Quality Data 17

2.00: Literature Review **19**

2.01: Building Energy End-use Study (BEES) Literature Review 19

2.01.01: Non Domestic Building Stock (NDBS) 20

2.01.02: Carbon Reduction in Buildings (CaRB) 21

2.01.03: Commercial Buildings Energy Consumption Survey (CBECS) 22

2.01.04: California Commercial End-Use Survey (CEUS) 22

2.01.05: Commercial and Institutional Building Energy Use Survey (CIBEUS) 23

2.01.06: The History of New Zealand Energy Surveys 23

2.01.07: International Surveys into Commercial Water-use 24

2.02: Research Methodology Review 25

2.02.01: BEES Research Design 25

2.02.02: Websearch Data Collection 25

2.02.03: Websearch Data Coding 26

2.02.04: Websearch Pilot 28

2.02.05: Measured Energy Data and Monitoring 28

2.02.06: Energy Data and Typology Mix 28

2.02.07: Benchmark Allocation in the Census 28

2.03: Visualisation Techniques 30

2.03.01: Automated Building of a City Visualisation 30

2.03.02: Precedents / Inspiration: Statistical Mash-ups 30

2.03.03: Graphic Considerations 32

2.03.04: Producing Benchmark Energy Average Colours 33

2.03.05: Colour Ranges –Appliance Labelling and EU Display Energy Certificates 34

2.03.06: Colour Interpretation and Colours for Impaired Vision 35

3.00: Research Methodology 36

3.01: Research Methodology Flowchart 36

3.02: BEES Census Data 38

3.03: Initial Corrections and Amalgamation 38

3.04: BEES Sample Selection 40

3.05: Data Collection Design – Team Meetings: 40

3.06: Street Surveys 41

3.07: Pilot for Method, Costings and Timing 41

3.08: Data Collection Review + Steering Group Meeting 42

3.09: Websearch Data Collection 43

3.10: Inferring New Data Sets 45

3.11: Data Cleaning Error Issues 46

3.12: Data Coding 47

3.13: Quality Checks for Bias Using Typology Mix 49

3.14: Floor Area / Storey Height Comparison : QV / Websearch Test 60

3.15: Attaching BEES Energy Use Indexes to Websearch Records 62

3.16: Websearch Criteria (Typology)Related to their EUI Averages 63

3.17: Attaching EECA and BEES Case-study End-use Percentages 66

3.18: Assigning Benchmark EUI Colour Ranges – Whole Building 66

3.19: Assigning Benchmark EUI Colour Ranges – Energy End-Uses and Water Use 67

3.20 Geo-coding Address information in Census 68

3.21: Automatically Constructing the 3D City Model 68

3.22: Algorithm Requirement – The Missing Link 70

3.23: Creating a Dashboard Interface 72

4.00: Graphic Representation 80

4.01: Layer Control Menu 81

4.01.01: Building Fill Colour 82

4.01.02: Building Outline Colour 82

4.01.03: Ground Texture 82

4.01.04: Objects 82

4.01.05: Zone 83

4.01.06: Highlight 83

4.01.07: Overlay 84

4.01.08: Elevation Mapping 85

4.02: Tools and Controls 86

4.02.01: Dialogue Box Properties Queries 86

4.02.02: Data Legend Display 86

4.02.03: Selection-Specific Controls 87

4.02.04: Data Entry Form 87

5.00: Data Quality and Privacy Issues 89

5.01: Coverage 89

5.02: Typology Bias 90

5.03: Analysis 91

5.04: Quality Management and Accuracy of the Data 94

5.05: Extension of the data 96

5.06: Energy and End-Use Data Combined with Websearch 98

5.07: Overcoming Data Privacy Issues 100

5.08: From Database to Visualisation 101

5.09: Sources of the Data 101

6.00: Conclusion 103

6.01: Accuracy of the Visualisations and Baseline Data 106

6.02: Open Data and the ‘Unique Building Identifier’ 107

6.03: Conducting a Building Stock Energy Visualisation Internationally 107

6.04: Gaps in the Data 108

6.05: Publicly Available Data Versus Confidentiality Agreements 109

6.06: Basing the Visualisation on High Quality Datasets 109

6.07: Scaling Up the Visualisation to National Level 111

7.00: Future Potential **113**

7.01: Future Applications of this Research 113

7.02: Testing the Success of the Visualisation Tool 114

8.00: Bibliography **116**

9.01: Appendix i: Websearch Data Sources 128

9.02: Appendix ii: Websearch Training Notes 128

9.03: Appendix iii: Criteria Benchmark Averages 128

9.04: Appendix iv: Mail-merge Websearch Checklist 128

Table of Figures

Figure 1 Display Energy Certificate (UK) 9
Figure 2 Geo-coded Display Energy Certificates - UK Public Buildings - (Volume of CO^2 Gas) 11
Figure 3 2D GIS Layers 13
Figure 4 Gapminder - Carbon Dioxide Emissions Versus Income per Person 14
Figure 5 "Project Glass" : Google Augmented Reality 15
Figure 6 Christchurch Earthquake Augmented Reality HitlabNZ 15
Figure 7 Automated 3D City Model from Excel Data using Generative Components Plug-in 16
Figure 8 3D Wellington - Earthquake Prone + Heritage / Non-Heritage Buildings 17
Figure 9 Websearch NZ Built Forms based on Steadman's Classification of Built Forms (NDBS) 27
Figure 10 LessEn Energy Map 30
Figure 11 Tweets + Flickr Uploads (Orange-Flickr, Blue-Twitter, White-Both) 31
Figure 12 Dashboard Scale 34
Figure 13 Methodology Flowchart Indicating the Research Boundaries and Authors of each Task 37
Figure 14 Initial Validation of Incorrect Census Buildings Before and After Correction 39
Figure 15 Size Strata Sample Allocation 40
Figure 16 Pilot Sample Tasks 41
Figure 17 Pilot Websearch Timing Task Breakdown Before and After Pilot Modifications 42
Figure 18 Records with Multiple Buildings Associated 43
Figure 19 Error Rates Found in Websearch Data 44
Figure 20 Websearch Data Collected + Inferred 45
Figure 21 Street View Coverage in NZ 46
Figure 22 Websearch Coverage: Aerial + StreetView 46
Figure 23 'Error' Codes – From 1,061 Websearch Records (35%) 48
Figure 24 BEES Climate Zones - New Zealand 48
Figure 25 Building Size Stratum Sample Comparison 51
Figure 26 Building New Use Type Sample Comparison 52
Figure 27 Building Use –mix Sample Comparison 53
Figure 28 Building Category / Site Density Sample Comparison 54
Figure 29 Building Decade Sample Comparison 55
Figure 30 Climate Sample Comparison 56
Figure 31 Building Size / Height Sample Comparison 57
Figure 32 Building Use / Size Sample Comparison 58
Figure 33 Floor Area Comparison QV and Websearch 60
Figure 34 Storey Height Comparison - Websearch v's Calculated QV Stories 61
Figure 35 Floor Plate Comparison QV v's Websearch 61
Figure 36 Comparison of Websearch Storey Height with Average Skylight Factor 62
Figure 37 Premises Per Building Illustrating Refusal Rate 63
Figure 38 Websearch Typologies Compared to Average Energy Intensity (EUI's) 64
Figure 39 Benchmark Averages (Bills) Versus Simulation Model Sample (Small sample sizes) 65
Figure 40 Benchmark Averages Applied to Wellington CBD Census 66
Figure 41 Energy + End-Use Benchmarks: Averages per Colour Range 67
Figure 42 Water Benchmarks 68
Figure 43 Map Marker Icon HTML Codes 68
Figure 44 Colouring the 3D Model Manually for the Wellington Pilot Visualisation 69
Figure 45 Colouring Algorithm - Illustrated Steps 71

Figure 46 Dashboard Interface Energy + Meshblock Zones 72
Figure 47 Meshblock Query: Average Energy and End-use + Typology Mix 73
Figure 48 District Plan with Building Energy 74
Figure 49 District Plan with Building Energy: Properties Query Legend 74
Figure 50 Lighting End-use with Tsunami Lines 75
Figure 51 Space Conditioning End-uses 76
Figure 52 Miscellaneous End-use Energy 77
Figure 53 Building Energy (Building Block) and Water Use (Outline) 78
Figure 54 Energy (Building Blocks) 'Mash-up' with Lighting End-use (Outlines) 79
Figure 55 StreetView Image of Energy Benchmark Visualisation 79
Figure 56 Layer Control Menu 81
Figure 57 Highlighted Meshblock Example 83
Figure 58 NESW Cross Hair Format: Overshadowing /Glazing Area/Inferred Insulation Levels 84
Figure 59 Spider Diagram Format: Overshadowing, Fabric + Age and Glazing Ratios 84
Figure 60 Overlaid Icon Infographics - Building Typologies 85
Figure 61 2D Traditional Noise Mapping 85
Figure 62 3D Noise Mapping onto Building Facades 85
Figure 63 Data Legend Display Types 86
Figure 64 Selection-Specific Control Panel 87
Figure 65 Time Slider for Longtiduinal Data (Net Zero, Stock Improvements, Energy Costs) 87
Figure 66 Data Entry Form – Building Level Query - Actual Energy Data from Building Owner 88
Figure 67 StreetView Limited Coverage Areas 89
Figure 68 Typical Errors and Assumptions: Present in 1,061 Building Records in the Websearch 91
Figure 69 Wellington Census - Energy, Lighting, Space Conditioning and Miscellaneous 93
Figure 70 2D GIS and 2D Extruded GIS Versions of the Pilot Visualisation 93
Figure 71 Websearch Workstation Screendump 94
Figure 72 Office Quality Grade A-D with Average Energy Results 97
Figure 73 BEES High and Low EUI End-use Mixes 98
Figure 74 EECA Energy End-Use Data 2008 99
Figure 75 Energy Average Benchmark EUI Map-markers Linked to Address Point Data 101
Figure 76 Wellington CBD Energy Benchmarks (Histograms, 2D Map, 3D Built Forms) 103
Figure 77 Energy + Meshblock (top) Lighting End-use + Tsunami Lines (bottom) 104
Figure 78 Wellington Space Conditioning versus Miscellaneous End-use Energy 105
Figure 79 Steadman Building Form Classification 110
Figure 80 Google StreetView Global Coverage 114

1.00: Visualising the Invisible

Is it possible to create an open source automated 3D city visualisation that makes normally-invisible building resource-use data visible? How can we be sure reliable energy and water-use benchmarks are applied consistently to a 3D city visualisation? What process should be employed collecting, analysing and inferring building characteristic data to ensure we base the visualisation on a database we can trust? Would a 3D visualisation improve accessibility to information that would otherwise remain invisible and unintelligible?

A quality visualisation enables you to see trends, patterns, and outliers that tell you about yourself and what surrounds you. The best visualisation evokes that moment of bliss when seeing something for the first time, knowing that what you see has been right in front of you, just slightly hidden" (Yau, 2013). Yau says "Data is more than numbers, and to visualise it, you must know what it represents" yet despite searching through numerous visualisations about the built environment, none discovered thus far map statistics about buildings onto actual building forms.

In order for a building resource-use visualisation to become a reality it will be necessary to resolve many issues. It will not be possible to provide solutions to all of these in this thesis alone, but each of the parameters involved in the process will be covered in the introductory chapter along with the potential benefits of such an interface. Benefits discussed include raising the value of sustainable buildings, encouraging efficiency upgrades, as well as raising public awareness of resource use. If successful the visualisation has the potential to reinvent the means of presenting data about the built environment, overcoming some of the current limitations of presenting geo-located statistical data in 2D. Future applications of this research are examined, including augmented reality and virtual world platforms along with an overview of tools enabling the automatic construction of the 3D city visualisation. The introductory chapter concludes with an overview of the data quality checks that may be necessary to ensure reliability of the benchmarks used and the correct allocation of those benchmarks.

Literature reviews in the second Chapter cover three areas – national building energy and water survey precedents, visualisation techniques and research methodology applicable to this thesis. The following sections will detail the specific research methodology and design that will enable the construction of a visualisation tool using publicly available data commonly available in many countries. The graphic options for the visualisation will be discussed, including examples from the Wellington Central Business District (CBD) visualisation composed as part of this thesis. The final section will detail a critical review of the baseline data and describe quality checks carried out in order to enable the highest quality benchmark energy and water results are assigned to the population.

1.01: Raising Value of Sustainable Developments and Upgrading the Building Stock

Energy, as a product, is invisible and often taken for granted by indifferent consumers. It rarely feels tangible except when absent (Darby, 2006)

Sarah Darby, a researcher at the Environmental Change Institute at Oxford University wrote in a 2006 report "The Effectiveness of Feedback on Energy Consumption" that a prime cause in energy wastage is that it is still "largely invisible to millions of users." And that "feedback on consumption is necessary for energy savings"

Throughout their life-cycle, buildings account for almost half of all greenhouse gas emissions (Johnson, 2007). Measuring, benchmarking and reporting energy and water consumption is the first step to reducing resource use (Clark, 2011). There is a global drive to make building energy and

water use data publicly available – particularly in the commercial sector (Malin & Roberts, 2012). Making this information accessible and legible, not to mention interesting, has the potential to impact energy and water use. Viewing the world through a 'lens' of resource-use data will make it feasible to envision the environmental performance of cities and neighbourhoods. It will be possible to glance at an area of the city and understand the underlying resource-use trends and provide designers, planners, landlords, tenants, energy companies and the general public with tools to work together to reduce consumption. Building owners and occupants will be able to compare their actual energy use to average benchmark energy intensities for their specific building typology.

In response to the European Union Energy Performance of Buildings Directive (European Committee for Standardization, 2007), public buildings over 1,000 m^2 in Europe are now required to publicly display energy certificates (Figure 1 ((Dept. Communities & Local Government, 2012))) and there are proposals to make this compulsory for all buildings. By 2016 landlords of the lowest performing buildings (F + G grade) will be required to install efficiency measures when requested by tenants. By 2018 these buildings will need to be upgraded in order to be legally leased (UK Government, 2011). Compulsory Building Energy Efficiency Certificates (BEEC) are also being introduced in Australia for office buildings over 2,000m^2 (Dept. Resources, Energy and Tourism, 2013). Making this kind of data publicly available places an extremely powerful tool in the hands of building owners, users and marketers. It has the potential to increase the market value of sustainable properties and discourage landlords / leaseholders from retaining or occupying buildings that perform poorly. There is currently no mandatory energy certification scheme proposed for New Zealand, so clear communication of benchmark building resource use data is crucial.

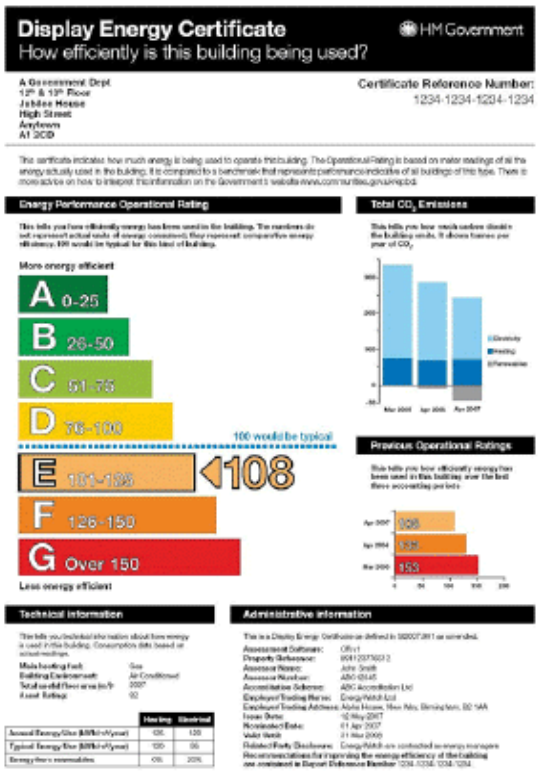


Figure 1 Display Energy Certificate (UK)

Many countries have introduced voluntary rating tools in order to improve the knowledge about the level of sustainability in the building stock. On one hand, it can be argued that the individual characteristics of place, such as the climate and type of building stock, necessitate an individual sustainability rating tool for each country, but the downside to this is that the rating tools for different countries are constructed on different parameters, which in turn has been a barrier to developing a global rating system (Reed, Bilos, Wilkinson, & Schulte, 2009). Building Research Establishment Energy Assessment Method (BREEAM) rated buildings account for 40% of all new buildings in the United Kingdom, however when it is realised that only 1%–2% of new stock is added to the total building stock each year, it is clear this accounts for a tiny proportion of the UK building stock (Reed et al., 2009) and it will be many decades before the entire stock is 'sustainable' through these means alone. To lower global carbon emissions from the built environment, it is crucial that the existing building stock is tackled urgently (UNEP, 2009). Concentrating on improving new buildings and making building regulations stricter will only result in incremental differences in emissions and resource depletion from the building industry. In addition, the lack of a comparable rating tool for

evaluating both existing and new properties means that the property market has not been able to value sustainable buildings appropriately (Warren-Myers, 2011). To-date building assessment tools such as BREEAM (UK, Holland), LEED (Canada, USA, China and India), and Green Star (Australia and New Zealand), have been voluntary and not mandatory schemes. While the 2003 EU Energy Performance Directive is compulsory, it currently only requires disclosure of energy performance rather than attainment of stringent performance targets (Reed et al., 2009).

This proposal attempts to develop a solution that will focus interest on the state of existing building stock - rewarding building owners who make efforts to improve their energy / water consumption and 'name and shame' the poorest performing building typologies, penalising these building owners through identification alone. If the visualisation tool that delivers this information is reliable, adaptable, enjoyable to use and freely accessible to all, it should be possible to improve the value of sustainable developments and encourage more owners to upgrade their properties to meet the needs of a more knowledgeable public. This would allow the market to drive sustainability – rather than relying only on regulation and government policy to effect change. An example of this change in action in New Zealand is the recent drive to upgrade Wellington buildings' seismic strength despite the lack of any legislation: "Fear is driving the upgrade and it is occurring now with no legislative requirement. There is no other choice – no tenant, no-hope" (Ian Cassels – Wellington Company (NZ Institute of Building, 2012)) "The brutal power of the market will determine - those that don't upgrade will go out of business" (Iona Pannett – Wellington City Council (NZ Institute of Building, 2012))

A 3D visualisation of building resource use benchmarks and building typology data expressed in 'heat map' colours on an open source tool such as Google Earth™ would bring this valuable information into the public domain. Architects and designers have the greatest need for this kind of information to be freely available. Enabling designers to visualise 'invisible' resource use alongside other building and site characteristics (communicated through the volume, grain and form of the city) may assist design processes. Architects and designers have been found to respond much more positively to 3D visual representation than to 2D (Ryan, 2007). This preference of 3D over 2D would logically extend to the publication of energy / water use data in graphs, reports and academic publications. Whilst there are numerous support tools available that encourage sustainable decision-making, these have limitations as they are dominated purely by the perceptions of the "expert" decision makers (e.g. building professionals). They focus on the technical design elements of a project, and they are typically two dimensional, where in reality the problems they are required to address are three-dimensional (volumetric) or even four-dimensional (temporal) (J. Isaacs, Gilmour, Blackwood, & Falconer, 2011).

In Europe, Display Energy Certificates act as an important catalyst to encourage behavioural change (Bull, Chang, & Fleming, 2012). Building owners and occupants need a tool to compare their own energy use to norms for their building typology. If a visualisation can begin to reveal this information effectively and dynamically to the public for all commercial buildings (rather than just large commercial buildings or offices as is the case with the European certificates) it follows that the property market will begin to feel the effects of this public knowledge. By 'toggling' visualisation building block colours from one statistic to another it will be possible to compare building typology information with benchmark resource use and reveal its distribution in certain areas of interest. Combining these comparative statistical visuals with dynamic longitudinal data and environmental / council zone overlays may enable governments to test out the effects of policy changes or planning legislation. This could include building height, overshadowing effects, or façade sunlight access.

1.02: Reinventing the Means of Presenting Data about 3D Urban Environments

The greatest value of a picture is when it forces us to notice what we never expected to see. (Tukey, 1977)

This thesis proposes a reinvention of the means of presenting data about 3D urban environments in a national building stock 3D visualisation. Conventional GIS use of 3D 'enhances' hard to understand 2D maps with even harder-to-understand histograms of data that communicate abstract statistical values along a "Z" axis (Figure 2 (Carbon Visuals, 2010)). The goal of this research is to demonstrate the means by which data on resource-use in buildings can be used to enhance familiar 3D interactive city environments and be made accessible to the widest possible audience.

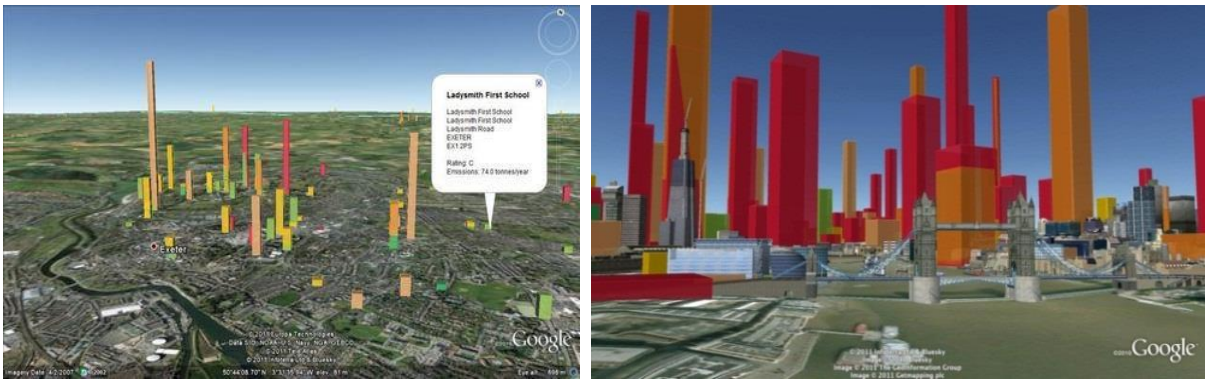


Figure 2 Geo-coded Display Energy Certificates - UK Public Buildings - (Volume of CO^2 Gas)

To test the potential of the idea, an annotated 3D visualisation of the Wellington Central Business District (CBD) has been constructed and coloured with average resource-use benchmarks from the BEES (Building Energy End-use Study) (Bishop, 2012b) - a BRANZ research project documenting energy and water use in New Zealand commercial buildings. By comparing typology data sourced from open data and internet searches (BEES Websearch survey) with measured energy results from the BEES energy bills and phone survey, it was possible to assign colours to building blocks in the visualisation (red indicating highest, through to dark green indicating lowest energy intensity). The BEES Websearch survey was originally designed, supervised and collated by the author and the building characteristics were rigorously quality checked and corrected, while being compiled into a single database. Confidential building resource use data from BEES could therefore be expressed as publicly available average energy benchmarks in the population. The visualisation identifies high volume – high energy versus low volume - low energy (and the inverse) using the 3D building volumes themselves as the mode of communication.

The overall aim for this pilot is to test the feasibility of applying this technique at a national level. Ultimately, benchmark resource-use data and other related publicly available information about cities could be presented in this form and would be database driven, so automatically updateable. From this basic platform, applications that permit people to compare their own private records with public norms are easily constructed: a world where a building owner can compare their energy records with benchmarks for similar buildings and take action to improve if necessary or to advertise accomplishments; or where the daylight factor of a shop in Wellington might be compared with a shop in Dunedin.

Building owners and leaseholders currently have very little information regarding typical energy and water consumption for similar building typologies. This means major equipment or supply faults triggering excess energy consumption have the potential to be overlooked. Resource supplies may

also be susceptible to exploitation or unintentional interception. There have been numerous cases of inadvertent (or fraudulent) theft of energy supplies by neighbouring properties, some of which were uncovered during the course of the BEES survey. With the gradual introduction of water metering and supply charges for commercial businesses in New Zealand this may also become a problem with water supplies in the future. Without openly accessible benchmarks on building resource use for comparison, this issue may become more commonplace with increased energy / water supply prices, but we would not be aware of this without access to the benchmarks in the first place.

During the BEES Websearch survey a world first approach was taken to visiting a stratified random sample of 3,042 commercial and industrial properties taken from the New Zealand government valuation database (QV, 2008) (Auckland Council, 2008). Google Earth™, Google Maps™ and Google StreetView™ were used to build a detailed picture of non-residential building stock in New Zealand with data collected on typologies, characteristics and surroundings. Data collected included floor area measurements, number of storeys, building height, glazing ratios per elevation, materials, form-type, building age, overshadowing, use-type and mix, as well as Google SketchUp® models, Google aerial and StreetView elevation images. The data collected were selected for two reasons: Firstly the BEES team deemed them to be potential indicators that contribute to energy and water use in buildings and secondly, the information could be collected from publicly available data within a reasonable time limit to keep research costs to approximately NZ$10-15 per building record.

The BEES team subsequently conducted detailed on-site measurements and telephone surveys on a subset of this Websearch sample. Energy and water bills have been gathered from a random sample to create a statistically valid picture of the end uses of energy and water in all NZ non-residential buildings. A previous BRANZ study HEEP "Household Energy End-use Project" collected equivalent data on NZ residential properties (N. Isaacs et al., 2005). This benchmark data from domestic and commercial buildings can theoretically be related to the publicly available typology and characteristic data and thus used to 'annotate' a 3D model of the whole country.

When the author began work on the BEES research project in 2008 it became apparent that one of the key requirements of the stakeholders was to make the resulting research publicly accessible. Results of building energy surveys around the world have had a long history of being reported in international journals and at conferences, but this is tantamount to preaching to the converted. If there is to be any hope of encouraging a significant impact on building resource use, it is necessary for this information to be communicated to a wide audience in an easily-digestible format. Only then will we have a chance to improve the perceived value of sustainable buildings, create the desire to upgrade the existing building stock and significantly reduce carbon emissions from the building sector.

3D building resource-use data has the potential to be made even more powerful when combined with other data sets (J. Isaacs et al., 2011). Examples include embodied energy, wind, earthquake risk, sun paths, weather, noise, fire, resource consent envelopes [recession planes and maximum building heights], traffic, heritage, solar access, council policies and potential savings models (net zero energy potential (Voss & Musall, 2011), energy saving and generation potential) to name just a few possibilities. Adding these extra layers of data and allowing quick switching between various data layers transforms a 3D visualisation into a valuable design tool.

1.03: Representing Geo-coded Statistical Values

2D GIS layers (Figure 3 (Land Surveyors United, 2013)) allow people to view, understand, question, interpret, and visualise 2D data in many ways that reveal relationships, patterns, and trends in the form of maps, reports, and charts. An example of this is communicating the effects of a train line / river / suburban boundary (a 2D linear object) upon statistical results. It is also possible to view two or more contrasting statistical results together and look for correlations. A notable example of this is Jane Jacobs' - property prices versus crime statistics discussed in her book "Life and Death of American Cities" (Jacobs, 1961). 2D GIS style mash-ups are abundant now, but this form of data is often obscure in presentation and fails to connect with or communicate to the average person in the street. Many existing examples of world research into building energy use have been collected and communicated using this 2D 'heat-map' approach.

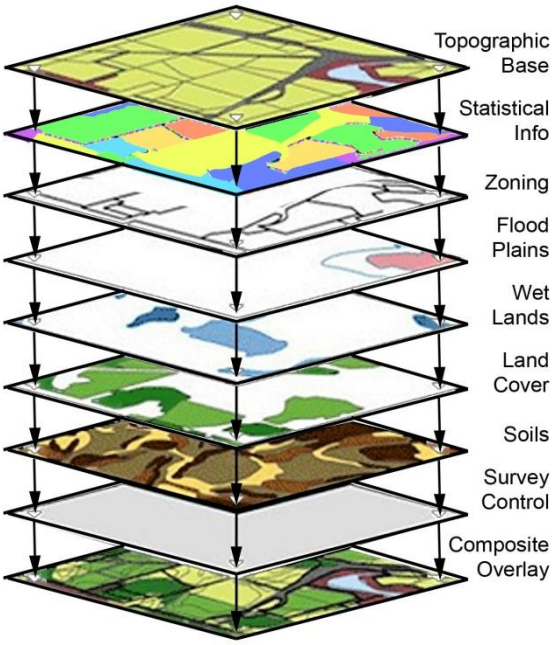


Figure 3 2D GIS Layers

This approach has been extended to include geo-located building footprint extrusions with Z-axis height or volumes related to statistical values (Figure 2 (Carbon Visuals, 2010)). These visuals powerfully demonstrate geographic relationships between data points when overlaid onto 3D platforms such as Google Earth™. They result in a 3D geo-located 'histogram' where the heights of buildings are scaled according to statistics, as well as viewing distance. The extruded blocks do not relate to the city grain, scale, or building forms and remain highly abstract. The main advantage with this extruded approach is that statistical data can be viewed at country / city scale. This may be ideal for indicating areas of interest - where results seem high or intriguing, however many of the relationships between buildings and their context are lost at this scale. It would be far more applicable if they were represented on real-scale building block forms, particularly when data relates to *all* buildings and not just sporadic building data points of a specific type. Considering the extrusions display information about the built environment, they are completely misleading as to the actual quantity of energy they are attempting to convey, as well as any relationship to the physical building the data is referring to. If a visualisation can make you aware of the physical building form, use, occupation and typology in relation to its energy intensity, then energy efficiency opportunities begin to reveal themselves by investigating the data and research efforts and policy change can be directed to the most urgent and effective areas.

There are plenty of visualisations around that communicate energy per square metre ($kWh/m^2/yr$) indices (Columbia Engineering School, 2011). These show the intensity of energy use as a function of land area (GIS maps) or at sporadic data points but not the actual energy use in relation to building form. If the primary goal is to communicate statistics about building resource use (quantity and intensity) then it follows that data should be communicated in a way that respects the urban form, grain and volume, providing a new way of looking at the city. This has far-reaching potential if the information can be revealed in-situ as a user travels around a city where data can augment ones view of the building stock.

A number of web visualisation tools have been developed in the last few years where users can select their choice of publicly available statistics and view them dynamically in 2D GIS format. Examples include the London Profiler (UCL, 2008) and the Berlin Atlas of Crime (Berlin Police, 2012) both of which were highly influential in the development of this subject. The ability to toggle from one 2D graphic to another animates the visualisation and helps the user understand relationships which may not be immediately apparent when viewing the same data in report, chart or book format.

There are many examples of geo-coded statistical representations breaking new ground graphically revealing patterns, trends and outliers that would otherwise remain invisible. The majority of these continue to represent statistics in a simple yet appealing 2D format; in the form of well-designed maps, plans or infographics (Yau, 2013).

A recent development in animated statistical visualisation initially developed by Hans Rosling's Gapminder Foundation in Sweden is the concept of a 'Trendalyzer' or animated motion chart (Figure 4 (Rosling, Rosling Ronnlund, & Rosling, 2005)). This technique conveys a fourth dimension hidden within data - It is a dynamic histogram exploring the relationship between the x and y axes whilst displaying each data point as a 'bubble' varying in size according to a third statistic component. The bubble colours can be linked to specific concepts or geographic areas which are clearly indicated on an adjacent map or key. The data changes over time using an animation timeline control which you can "play" hence introducing a fourth dimension (assuming longitudinal data is available). The time intervals conveyed could be microseconds, days, years or centuries depending on the research area being investigated. If you observe interesting behaviour in the motion or changing size of a "bubble" data point over time in the histogram you can hover your mouse pointer over it to identify the source of the data. If you wish to observe a particular data point throughout the time segment you can place a label on the point of interest to keep tabs on it. Likewise you can filter results so that you only display the data points of interest or zoom into an area of the chart to look at the data visualisation in detail.

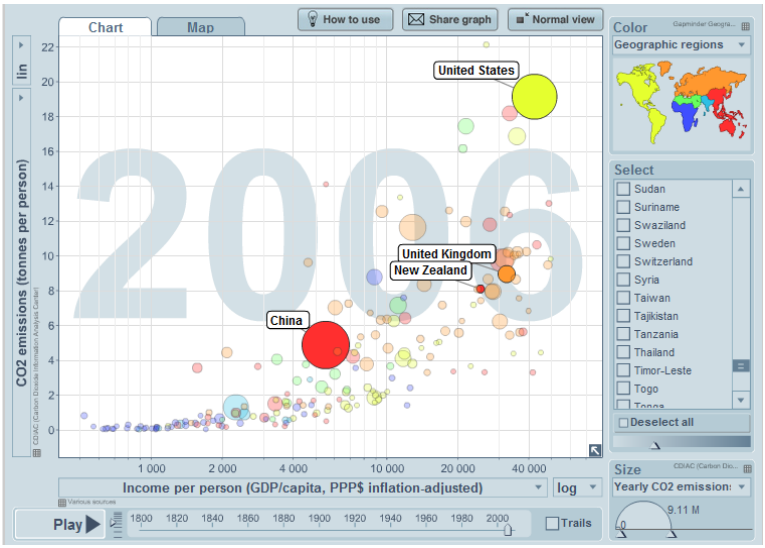


Figure 4 Gapminder - Carbon Dioxide Emissions Versus Income per Person

Statistics become increasingly powerful when viewed over time with high quality longitudinal background data. The application of this advanced visualisation style in a building resource-use visualisation could communicate:

- efficiency improvements made to the building stock over time
- sunlight access and thermal gains
- the effect of financial and energy payback periods
- the effects of rising costs of energy / water supply
- the level of difficulty in upgrading the building stock to Net Zero (Cory, 2015)
- how energy benchmarks change over time

A 4-dimensional (4D) city visualisation could push this idea further into the realm of noise, wind and energy simulations. For instance solar or thermal gain heat-maps could be rendered onto the building forms illustrating radiation or heat loss over the seasons (or the course of a day), helping building owners, local planners or designers site thermal mass within a building and solar panels or photovoltaics on the exterior for the greatest returns.

1.04: Augmented Reality and Virtual World Platforms

Augmented reality through a smart phone, tablet or Google Glasses (Figure 5 (Google, 2013c)) involves the use of 3D platforms such as Google Earth™'s virtual globe or Google StreetView®'s omnidirectional imagery, both of which can be geo-located to your exact position and orientation, thanks to GPS and tilt meters. There is a race to map the world with point cloud 3D visualisation to underlay augmented reality applications of the future accurately (Austen, 2013). This type of visualisation can be created by geo-referencing aerial or panoramic photos, known as 'photogrammetry' (Georg, 2001). It can also be made increasingly accurate with LiDAR (Cracknell, 2007) data which uses radar to pinpoint topography and built forms in three dimensional space accurately. This detailed information is translated into a digital surface, upon which aerial or elevation images can be 'draped' to create a highly realistic virtual model. Augmented reality can also exist as geo-located annotation graphics linked to physical forms and overlaid on real-time video footage on smart phones or tablets.

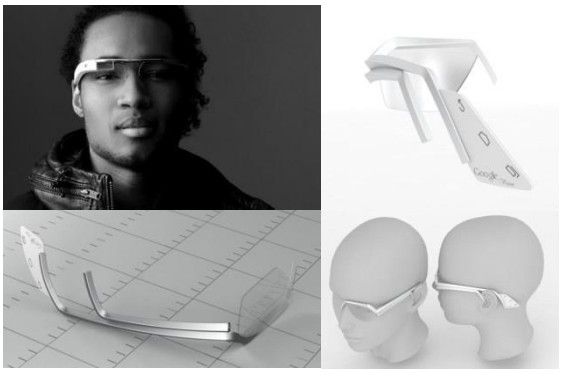


Figure 5 "Project Glass" : Google Augmented Reality

The technology already exists to include real-time video overlaid with historic city images or virtual objects, for example HitlabNZ's CityViewAR tool (Billinghurst, 2011) illustrates before and after visual comparisons of earthquake damage in Christchurch city (Figure 6). CityViewAR involved the creation of a textured 3D city model by triangulating data from a series of still photographs of the city as mapped building textures. These were brought to life as overlays on real-time video footage. This thesis seeks to explore how this technology might routinely be combined with building performance data, in the process making this data accessible as readily to non-specialists as to experts. This annotated virtual world is expected to become an internet of geo-coded objects providing a future platform for visualisations, information, art, graffiti, protest, advertising and commentary on physical real-world things (Open Geospatial Consortium, 2011). It has been suggested this may well become a 'litigation headache' when this technology takes hold in the future (Marks, 2012).

Figure 6 Christchurch Earthquake Augmented Reality HitlabNZ

Google™'s new augmented–reality game 'Ingress' sees players fighting for control of real landmarks or monuments dubbed 'portals' in their cities. It represents a huge step towards accurate augmented reality (AR's) in which real world objects are annotated with a virtual layer of information that is overlaid on a smartphone's or tablet's live camera footage. 'Ingress' runs as an android application

tied to the real world objects through GPS and tilt meters. Players contribute to the quality of the virtual model by geo-locating photographs from their phones which are later used to visualise the 3D virtual environment around them. Google™ is effectively automating the process of building a virtual world. A detailed record of where the players wander and the establishments they visit en-route is a data "gold mine" for Google™ to use to improve its location based services (Hodson, 2012). If this game goes viral it will help add to Google Earth™'s 3D workshop – a set of building 3D models that have been created in Google SketchUp™ with crowd-sourced 3D content that is carefully monitored by Google™ to ensure the best possible accuracy. If the visualisation proposed by this thesis capitalises on this kind of 'gaming' appeal the tool could have a higher chance of success (Section 2.03.03).

1.05: Automated Model Construction

It is already possible to generate basic 3D city models automatically from cadastral, topography and building elevation data (Ryan, 2007), but there are some limitations with this generation method especially where buildings have complex forms and high surface area. Although BEES research assistants built block models in Google SketchUp® during the Websearch survey, they have only been created for buildings in the sample and not for the population. These separate models have been geo-located using Google Earth™ which works seamlessly with Google SketchUp® and allows multiple model files to be loaded into a single 3D city visualisation with correct positioning and orientation. Even if all other buildings are constructed using an alternative method, the sources underlying the visualisation need to be distinguished and this would mean BEES random sample buildings would be identified. Confidentiality agreements undertaken by the BEES team will not allow this and thus an alternative method must be used to build a 3D model of New Zealand's commercial building stock.

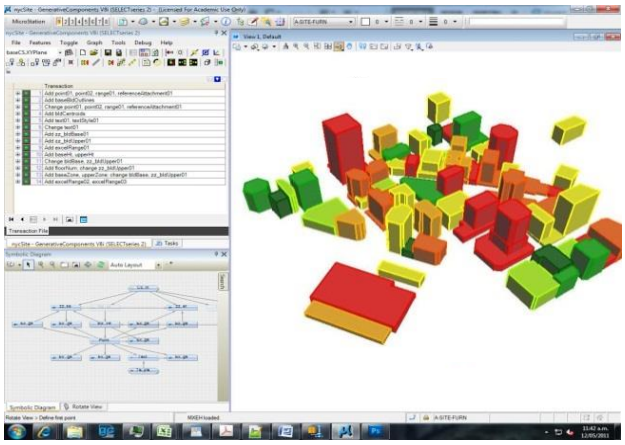


Figure 7 Automated 3D City Model from Excel Data using Generative Components Plug-in

Work at Victoria University Wellington's Centre for Building Performance Research indicates that tools such as Microstation's plug-in 'Generative Components' or Google Earth™ with Google SketchUp™ make it possible to automate the construction of building forms with parametric information contained in a database (Sullivan & Motley, 2011). This means we have the potential to construct and update building blocks, render surface colours and building outlines automatically according to statistics contained in a background database (Figure 7).

This automation is ideal in circumstances where new information on energy or water consumption may become publicly available or where building owners, wishing to upgrade benchmarks with actual energy data, may submit their energy bills and corresponding floor area for verification followed by inclusion in the visualisation. Indeed this may become a popular alternative as it would be a way of advertising their energy / water efficiency accomplishments. This is the mechanism by which the proposed visualisation could communicate efficiency improvements made to the building stock over time as discussed in Sections 1.03 and 4.02.04. This is how the NABERS building energy efficiency

certificates work in Australia with voluntary uptake in 75% of the commercial office market (NSW Office of Environment and Heritage, 2014).

Linking the verbose address-based data in the Websearch database to geo-located parametric 3D information in a visualisation model is demanding. ArcGIS "City Engine" (Esri CityEngine, 2013) has a "Spatial Connect" tool that can link Land Information New Zealand (LINZ) address point data (LINZ, 2011b) with adjacent cadastral land parcels (LINZ, 2011d), but it is only as accurate as the baseline data. Rows of data in a spreadsheet containing resource use benchmarks and building typology information could be linked to individual 3D building forms (LINZ, 2011a) and rendered automatically with a colour range according to statistics found in the database. The methodology in this thesis includes a manual and algorithm colouring method, that can be used to overcome some of the address issues (Section 3.22).

Figure 8 illustrates an automated 3D visualisation of central Wellington indicating heritage (red) / other buildings (green) along with earthquake prone buildings (dark red for heritage and blue for other buildings). This 3D model (Sullivan & Motley, 2011) and the author's own draft Wellington Visualisation model illustrated previously (Figure 7) were created with Microstation™ plug-in Generative Components (Bentley, 2013) using publicly available cadastral and building elevation data. For the purposes of the Wellington visualisation pilot, parametric data was manually matched to an spreadsheet containing statistics about the built environment, but the test model illustrated in Figure 7 demonstrates that automated modelling using a background database is a workable solution. This will become vital when automating the construction of national or regional 3D visualisations.

Figure 8 3D Wellington - Earthquake Prone + Heritage / Non-Heritage Buildings

1.06: Ensuring High Quality Data

Investigating the level of bias or correlation between the numerous BEES datasets, samples and sub-samples, ensures the most accurate assignment of the energy and water benchmarks to the census. Commercial buildings in the national and Wellington CBD (pilot visualisation) census, along with six BEES survey samples (Websearch, simulation models, monitored buildings, phone survey premises, water premises and energy bills premises) were compared for typology mix. The typology patterns present in each of the datasets were compared through the use of conditionally formatted (statistically coloured) frequency tables and histograms. Where a significant difference was observed, the buildings that differ were looked at in more detail (hybrid typologies) and compared with their average benchmarks.

The level of accuracy required for the purpose of this thesis will relate to the graphic output of the visualisation. This is associated with the number of easily detectable intervals in the chosen colour scales and the number of logical divisions of scale within infographics. Any improvement in the Websearch data accuracy also provides the BEES team with an improved instrument to extrapolate the energy and water results to census level. A detailed overview of the required level of accuracy for

this background data will be covered in Section 5.04. Further discussion on the graphic colour ranges and infographics proposed by the thesis are covered in Chapter 4: Graphic Representation. The literature review also discusses colour ranges and issues with colour interpretation in Sections 2.03.05 and 2.03.06.

The Websearch survey typology information is a rich set of data about the make-up of the commercial New Zealand building stock. Even without resource-use benchmarks, it has enormous potential when conducting research on the built environment and conveying the make-up of the building stock. A successful visualisation will require the highest quality typology data possible to generate coloured energy benchmarks assigned to the building forms. Using the rich characteristic data collected in the BEES Websearch survey and linking it to equivalent information contained in the NZ valuation database (floor area, building footprint, category of use and materials) makes it possible to extrapolate BEES building energy and water results up to census level. This real measured data hints at those building characteristics that have an influence on resource use, taking the research potential of this visualisation full circle.

Users should be able to compare benchmark energy results with their own energy consumption levels and with gradual improvements to the baseline data this automated visualisation can be continually improved. At worse this research will result in a detailed typology database of NZ buildings used to communicate information about buildings which will provide a valuable academic resource. At best, a visualisation tool will be proposed that uncovers the drivers of energy and water use – directing regulation to the highest consumers of resources and potentially improving the market value of efficient and sustainable buildings over time.

2.00: Literature Review

The visualisation proposed by this thesis required three diverse areas of research to be investigated; building stock energy surveys, statistical data collection methodology and visualisation techniques. The Literature Review begins with an exploration of New Zealand and international building energy survey precedents, research design and methodology (Sections 2.01.01-07) covering the broad topic of national energy surveys, which influenced the design and methodology of the BEES project.

The Websearch survey and thesis research design literature are discussed in Sections 2.02.01-07 including BEES research source data and papers applicable to this thesis. The literature review concludes with a review of visualisation techniques including dashboard design considerations, automated 3D modelling and statistical precedents.

2.01: Building Energy End-use Study (BEES) Literature Review

The Building Energy End-use Study (BEES) - a six year national study of energy and water use in commercial non-domestic buildings started in 2007. A major goal of BEES was to provide improved understanding of the drivers of energy and water use, as well as providing better baseline data (Bishop, Pollard, & Isaacs, 2012). As the name suggests, the goal of the research was also to uncover the end-uses of energy in commercial buildings. A literature review and annotated bibliography was prepared for the Building Energy End-use Study (BEES) (Cory, Hills, Isaacs, & Pyke, 2010). The review comprised New Zealand and international literature covering studies into energy and water use in non-residential buildings. The work was designed to identify and document all previous NZ studies into energy use as well as relevant international studies in order to support the development of the BEES research design. The literature review identified various research methodologies used, the type of equipment and monitoring devices employed, the social ramifications involved in energy consumption (behaviour, ownership, perceptions, comfort, energy determinants, spatial positioning within settlement areas), energy efficiency interventions, policy tools, GIS modelling and simulation and equivalent building water-use studies conducted around the world.

The equivalent studies that were reviewed in detail included: the United Kingdom's 'Non Domestic Building Stock' (NDBS) (Steadman, Bruhns, & Rickaby, 2000) and 'CaRB' (Carbon Reduction in Buildings) (Oreszczyn, Lomas, Ward, Guy, & Shipworth, 2013); USA's 'Commercial Buildings Energy Consumption Survey' (CBECS) (Energy Information Administration, 2013) and the 'California Commercial End-Use Survey' (CEUS) (Itron Inc., 2006); and lastly Canada's 'Commercial and Institutional Building Energy Use Survey' (CIBEUS) (Government of Canada, 2003). The NDBS study closely parallels the BEES study and is therefore covered in the greatest detail in the next Section 2.01.01. This is followed by an overview of each of the remaining four major international energy studies (Sections 2.01.02 – 2.01.05).

Previous New Zealand studies into building energy use include: 'Energy Demand in the Wellington Central Business District' (Baird, Donn, Porteous, & Runeson, 1978); 'Energy Performance of Buildings' (Baird, Donn, Pool, Brander, & Chan, 1984); The NZ Department of Statistics '1976/7 Inter-Industry Survey of the New Zealand Economy' (Peet, 1985); 'Lighting and Equipment Energy Survey 30 Wellington Buildings 1994/5' (M. Donn, Maule, & McAlister, 1995); 'Possible Energy Use Trends for New Zealand 2000/2010' (Energy Efficiency Conservation Authority & Centre for Advanced Engineering, 2000)(Energy Efficiency Conservation Authority & Centre for Advanced Engineering, 2000); 'Building Energy End-use Project' (BEEP) (Wilks, Donn, & Baird, 2003); and 'Housing Energy End-use Project' (HEEP) (N. Isaacs et al., 2005). An overview of the major New

Zealand energy studies is provided in Section 2.01.06, concluding with international water surveys in Section 2.01.07.

2.01.01: Non Domestic Building Stock (NDBS)

The first significant study on energy use in the United Kingdom's (UK) commercial buildings was the National Non-Domestic Building Stock (NDBS) project (Steadman, Bruhns, & Rickaby, 2000) sponsored by the UK Department of the Environment, Transport and the Regions (DETR) 1991 to 2001. The aim of the project was to provide support for government policy-making in carbon emissions reduction, in accordance with national commitments to reducing the impact of global warming. The report highlights the lack of information on the non-residential stock and introduced a national database of non-domestic buildings taken from United Kingdom valuation data. The quality of this data was checked and enhanced by carrying out a series 3,500 street surveys in wedge-shaped segments of central Manchester, Swindon, Tamworth and Bury St Edmunds. Unlike the BEES study, which began with a stratified random sample taken directly from the amalgamated NZ valuation records, the NDBS team were not satisfied they had a suitable sample frame to enable the random selection of buildings in the UK valuation database and instead, four wedge shaped city zones were chosen to include as many different types and sizes of buildings as possible. Like the targeted stage of the BEES project, interior surveys were carried out on a sub-sample of these buildings to obtain information about plant, equipment and the use of fuels. "Data obtained from the street surveys were entered into a geographical information system (GIS), over an Ordnance Survey map background, to create a system in which buildings are represented as assemblies of 'floor polygons' representing discrete areas of floor space on different levels. Data on occupants and construction were attached to each polygon, as well as information about storey heights and storey levels, allowing calculations to be made of the areas of walls. Buildings at the 3,500 addresses were represented by a total of around 14,000 floor polygons" (Steadman, Bruhns, & Rickaby, 2000). The BEES Websearch survey contained equivalent data to the NDBS street surveys however the output was a detailed database based on a random selection of valuation records rather than a series of GIS polygons with added survey data attached.

NDBS survey techniques were designed to provide breadth of coverage rather than great detail. Each building was inspected externally and the following characteristics were recorded: address; overall building form (including roof type and number of storeys); relationships to the site, the street system, and adjacent buildings; functional uses and names of occupiers; estimated age; and those details of fabric and construction (structural type, external wall and roof finishes, glazing type). Most of this information was recorded in the form of sketch plans and by means of `tick boxes' on proformas. Surveyors used their own `free form' descriptions for materials and activities, and classifications were made later. All buildings were also photographed with a 2 m measuring stick in each photograph in order to facilitate the later estimation of storey heights" (Brown, Rickaby, Bruhns, & Steadman, 2000). Information on limitations and advantages of this data collection method are discussed in the paper 'Surveys of Non-domestic Buildings in Four English Towns'. The NDBS street survey results were to be compared in due course with national valuation data on the non-domestic building stock and would serve as the basis for making inferences about built form and other characteristics not recorded in the rating data (Brown et al., 2000).

The NDBS involved the creation of complex building perimeters drawn for each floor level individually, thus ground-level carriageways, and setbacks or overhangs on upper levels, are all modelled in detail for four wedges of English towns (Holtier, Steadman, & Smith, 2000). This will constitute a significant difference in accuracy than that provided by the Websearch studies' rough

measurements and estimates from Google Earth / Street View images. In the BEES Websearch work the footprint is taken from an aerial photo and therefore it is not always possible to provide accurate footprints for each floor unless these are picked up from the StreetView elevation images. What is more likely to occur is a simple prismatic form extruded from the ground floor and then an allowance made for the setbacks / podiums etc. in the estimation of the total floor area of the building in the database and this area figure can then be compared against the floor area listed in the valuation database. The glazing ratio, total floor area and even ground floor footprint data are much rougher estimates in the Websearch. This is likely to be interpreted differently by each Websearch research assistant. In addition the GIS polygon method provides perimeter wall lengths, exposed wall and exact orientation whereas the Websearch data collection is more generalised with floor area and building height and approximate orientation noted down for each building in the sample. The BEES Websearch survey represented over 6.25% of New Zealand commercial building stock whereas the NDBS study, whilst more accurately recorded, represented only 0.2% of UK non-domestic building stock. In addition, the sample selection was restricted to four wedges of 'typical' UK towns rather than a random sample as was the case in the Websearch survey. It could therefore be argued that the high level data provided are of equal value overall. The detailed energy monitoring in NDBS survey allowed much more detailed analysis of complex building typologies / morphology where a range of energy data was available to infer energy benchmarks for buildings in all non-domestic use classes with reasonable sample sizes. There are a number of detailed papers that relate to NDBS that appear in the journal 'Environment and Planning B: Planning and Design'. Volume 27, Issues 1 and 5 have a specific focus on non-domestic building stock. The NDBS survey building form types (refer to 2.02.03) were closely mirrored in the NZ building stock with only one superfluous building form classification; the railway arch however, NZ building stock had a higher proportion of one storey 'string of single rooms' and 'single room' building forms than was uncovered in the NDBS study.

2.01.02: Carbon Reduction in Buildings (CaRB)

The second major study undertaken on energy use in the United Kingdom's non-domestic building stock was the Carbon Reduction in Buildings (CaRB) project (Oreszczyn et al., 2013). The big difference between the two UK projects is that CaRB also encompasses the domestic building stock. The CaRB project started in 2004 and the major aim of the research programme was to reduce carbon emissions from the UK building stock and create an innovative, public domain, socio-technical model of energy use in buildings applicable at national, regional, city and community level. The CaRB model was used to predict carbon emissions from the building stock and assess energy efficiency measures and their impact on emissions. (Carbon Reduction in Buildings (CaRB). 2009)

The 4 main research areas covered by CaRB were:

- Longitudinal studies of buildings: the analysis of energy consumption and carbon emission trends over time.
- Sociotechnical research: understanding both the social and technical dimensions of energy use and how these interact.
- Research on non-domestic buildings: constructing an energy and carbon emissions model of non-domestic buildings, particularly at the level of the UK building stock
- Constructing a community model: combining the energy and carbon emissions model for both domestic dwellings and non-domestic buildings"

Whilst the BEES study included sociotechnical research and aggregate energy consumption in commercial buildings, it did not include longitudinal data and excluded hotels, agriculture. heavy industry, health and educational buildings. When combined with the HEEP residential study (including all housing but not apartments or residential-commercial) the two studies become easier to compare however the large gaps in the BEES / HEEP combined studies would require further study to enable a proper comparison between UK and NZ resource use.

2.01.03: Commercial Buildings Energy Consumption Survey (CBECS)

A significant energy survey has been undertaken in the USA since 1979 known as the Commercial Buildings Energy Consumption Survey (CBECS) (Energy Information Administration, 2013). The latest survey was the eighth and was conducted in 2005. CBECS is currently conducted on a quadrennial basis. From 1979 to 1986, the survey was known as the Non-residential Buildings Energy Consumption Survey (NBECS). CBECS is a national sample survey of the commercial building stock which collects information on energy-related building characteristics, and their energy consumption and expenditures. Like the UK equivalent, the CBECS random sample covers approximately 0.2% of the non-domestic building stock. The data collection comprises a survey questionnaire of all commercial buildings in statistically selected geographic areas along with around one fifth obtained from lists of large buildings (hospitals, airports, and federal government buildings). Trained field staff walk or drive through these selected areas and record information about every commercial building in their sample over a four year period. In addition where survey response was insufficient actual energy use and billing information was provided by energy providers in USA through an online dataset thanks to the Energy Information Administration's mandatory data collection authority. In contrast BEES study scientists were only able to obtain energy data if buildings consented to the survey and signed an agreement, and this posed a significant disadvantage.

The commercial sector was found to encompass a vast range of building types: service businesses (retail, wholesale stores, hotel, motels, restaurants and hospitals) as well as certain buildings that would not be considered "commercial" in a traditional economic sense, such as public and private schools, correctional institutions, plus religious and fraternal organizations. Excluded from the sector were the goods-producing industries: manufacturing, agriculture, mining, forestry and fisheries, and construction" (Energy Information Administration (EIA). 2010).

The longitudinal work in this 35 year old study is a major strength and it covers a wider range of building types than the NZ studies, but the statistical sampling method used to gather the information is unlikely to be a highly accurate representation of the building stock. BEES research advantage was the small size of the country and the availability of a single dataset of buildings in the form of a taxation database. Given the vast size of the USA building stock this is no surprise, however the cooperation of the energy providers and the wide range of climate zones, neighbourhoods and built forms covered in the study provides a rich source of building energy use benchmarks.

2.01.04: California Commercial End-Use Survey (CEUS)

A study of commercial sector energy consumption in California was undertaken by the Californian state government to support the state's energy demand forecasting activities. The California Commercial End-Use Survey (CEUS) (Itron Inc., 2006) was a comprehensive study of commercial sector energy use, primarily designed to support the state's energy demand forecasting activities. A stratified random sample of 2,790 commercial facilities were collected from the service areas of Pacific Gas and Electric, San Diego Gas & Electric, Southern California Edison, Southern California Gas Company and the Sacramento Municipal Utility District. The sample was stratified by utility

service area, climate region, building type, and energy consumption level and covered approximately 17% of the state's commercial building stock making it one of the most comprehensive surveys undertaken in the world to date.

For each utility service area, floor stocks, fuel shares, electric and natural gas consumption, energy-use indices (EUIs), energy intensities, and 16-day hourly end-use load profiles were estimated for twelve common commercial building type categories" (Itron Inc and KEMA / ADM Assoc.James J. Hirsch & Assoc. 2006).

2.01.05: Commercial and Institutional Building Energy Use Survey (CIBEUS)

A significant project on commercial and institutional buildings was carried out in Canada and started in 2001. The project was called the Commercial and Institutional Building Energy Use Survey (CIBEUS) (Government of Canada, 2003) and was carried out by Statistics Canada on behalf of the Office of Energy Efficiency of Natural Resources Canada (NRCan). The surveys provide detailed information on the commercial sector which is used to assess how well Canada is fulfilling its commitment to reducing greenhouse gas emissions. (Office of Energy Efficiency of Natural Resources Canada (NRCan). 2000)

The main objective of this voluntary survey was to collect energy intensity information for the commercial and institutional buildings in Canada for a period of 6 months in 2000. More precisely, the survey involved collecting data from 5,000 buildings (0.4% of the total commercial building stock) establishing: Building characteristics; Occupancy characteristics; Energy efficiency characteristics; and Energy consumption. Target buildings were at least half designated commercial use with a floor area of at least 1,000 square feet ($93m^2$) located in Census Metropolitan Areas (CMAs) or Census Agglomerations (CAs) with population over 175,000. The CIBEUS dataset included churches, hotels, motels, educational and health buildings in addition to the commercial building use types covered in the BEES study.

2.01.06: The History of New Zealand Energy Surveys

'Energy Demand in the Wellington Central Business District' was undertaken as part of the Victoria University of Wellington, School of Architecture's course work in 1978 (Baird et al., 1978) which showed that the most important statistical determinants of commercial building energy use were the floor area and number of occupants, however, in order to make use of this knowledge on a national level it was necessary to have access to a comprehensive listing of all relevant floor space. Whilst this data was collected in the New Zealand census for residential buildings every five years, there was no listing of commercial non-residential buildings. Valuation data was the closest to this, but since it was used for financial and legal purposes, it did not record the number of structures per site, or valuations per building, so the only way of estimating the actual number of buildings was via a sub-sample survey. Other surveys were also conducted in Auckland (Shaw, Beca, Carter, Hollings and Ferner, & New Zealand Energy Research and Development Committee, 1979), Wellington (Baird et al., 1984) and Christchurch (R. W. Morris & Associates, 1985).

The NZ Department of Statistics 1976/7 Inter-Industry Survey of the New Zealand Economy resulted in a report describing the methodology, data sources and results of an input-output energy analysis (Peet, 1985); and Lighting and Equipment Energy Survey (M. Donn et al., 1995) with a focus on 30 Wellington commercial buildings. The results of these pilot studies created limited insight so a larger study was required to give more consistent findings.

The Centre for Advanced Engineering Canterbury produced a study on potential energy savings from the Energy Production, Transport, Communities, Households, Commercial and Industrial sectors. "This work did not attempt to create a detailed analysis of energy use within each sector but was more of a first pass at a new forecasting and analysis methodology assessing broad sector energy use patterns and trends" (Energy Efficiency Conservation Authority & Centre for Advanced Engineering, 2000). This report noted that subsequent phases of work would be needed to complete more detailed assessments of the identified energy and efficiency options, and their potential uptake rates. This led to the Building Energy End-use Project (BEEP) (Wilks et al., 2003) and this was an early precursor to the Housing Energy End-use Project (HEEP) (N. Isaacs et al., 2005) and the Building Energy End-use Study (BEES) specifically discussed in this thesis. The energy end-use breakdowns by sector provided by the Centre for Advanced Engineering Canterbury / Energy Efficiency & Conservation authority (EECA) along with the NZ energy data file, provided the underlying data for 'EECA - Energy End Use Database' (Energy Efficiency Conservation Authority, 2007). EECA end-use breakdowns by sector were used to communicate energy end-uses in the pilot visualisation of Wellington for this thesis. EECA end-use percentages have been be applied to building records in the census using detailed typology information from the BEES Websearch survey. Once the final BEES end-use data is released, the visualisation could be updated to represent the latest research.

2.01.07: International Surveys into Commercial Water-use

Due to water restrictions present in Australia and the United States - in particular California, the largest body of research on commercial water use has come out of these countries.

Australian studies include:

- 'Sustainable Water Management in Commercial Office Buildings' (Chanan, White, Howe, & Meenakshi, 2003);
- 'Water Efficiency Guide: Office and Public Buildings'(Quinn & Bannister, 2006);
- 'Benchmarking Best Practice for Water Use in the Commercial and Industrial Sector' (City West Water, Yarra Valley Water, & South East Water, 2006);
- 'Annual water Statistics 2005–06' (Dept. of Natural Resources and Water, 2006); and
- 'Scoping Study to Investigate Measures for Improving the Water Efficiency of Buildings' (GHD Pty Ltd. & Australian Greenhouse Office, 2006).

United States water studies include:

- 'Technology data characterizing water heating in commercial buildings: Application to end-use forecasting' (Sezgen & Koomey, 1995);
- 'Commercial and Institutional End-uses of Water' (Dziegielewski et al., 2000);
- 'Waste Not Want Not: The Potential for Urban Water Conservation in California' (California Department of Water, 2003);
- 'Demonstration of Water Conservation Opportunities in Urban Supermarkets' (Aquacraft, 2004);
- 'Evaluating the Benefits and Costs of Increased Water-Use Efficiency in Commercial Buildings' (Groves, Fischbach, Hickey, & RAND Corporation, 2007); and
- 'Commercial, Institutional and Industrial Water Use Survey Program' (Environmental Protection Agency, 2009).

Water surveys and studies in New Zealand are fairly limited in comparison. They include several BRANZ reports on water monitoring techniques and rainwater collection and two residential water surveys resulting in a 'public sector infrastructure toolbox' - Water End Use and Efficiency Project (WEEP) (Heinrich, 2007) which focussed on water used in the New Zealand tourism industry in 2008. Lee Bint completed her PhD on Benchmarking Water Use in Commercial Buildings as part of the BEES study in 2012 and a brief overview of the results are published in a BRANZ Build article (Bint, 2012).

At the time of writing the BEES water-use results are limited to only 34 commercial buildings within the BEES sample with a focus on buildings in the central Auckland area. The BEES Water-use Auckland Baseline Study (Roberti, 2012) illustrates some of the basic average water intensities related to building typology and use. Whilst these are not as comprehensive as the BEES energy survey results, these datasets could be used to illustrate average water-use benchmarks in the pilot visualisation using the same methodology described in this thesis for applying energy end-use benchmarks. As more data becomes available the automated nature of the visualisation will allow this further detail to be incorporated into the tool.

2.02: Research Methodology Review

2.02.01: BEES Research Design

One of the starting points for both the BEES and HEEP studies was the EECA 'Energy Efficiency Conservation Authority' - Energy End Use Database (Energy Efficiency Conservation Authority, 2007) see Section 2.01.06. EECA were one of the major stakeholders and funding sources for the BEES project and requested improved end-use information for the commercial and residential sectors to enhance the "top-down" energy information contained in their "End-use Database". They required data based on measurement and consumption rather than purely estimating and modelling. HEEP (N. Isaacs et al., 2005) covered the residential sector and BEES was designed to cover the non-domestic sector. Due to research budget limitations and the huge range of building typologies in the non-domestic sector (N. Isaacs, Jowett, Saville-Smith, & Hills, 2012), the BEES study was limited to commercial retail, commercial office and commercial mixed-use buildings (N. Isaacs et al., 2009). In order to pick up retail and office and retail uses incorporated into industrial sites, the "Industrial Service" and "Industrial Warehouse" sectors were also brought into the initial sample frame.

The criteria chosen for investigation in the BEES Websearch survey were checked against the details gathered in the NDBS street survey work (Brown et al., 2000) for an understanding of the issues around collecting data about buildings without interfering with them.

2.02.02: Websearch Data Collection

'Architectural Research Methods' illustrates a number of techniques in carrying out correlational and quantative research (Groat, 2002) including: coding to simplify verbose and disparate data entry; clarifying the correlational relationships; and establishing predictive relationships and calculating error rates. The Websearch design draws on this guidance within the constraints of the BEES team research requirements.

The Websearch data collection processes relied on Building Science students from Victoria University Wellington. They collected data on buildings with a little training and a standard procedure to ensure the underlying building typology data was as reliable as possible, and within the budget provided. Responding to the use of street surveys in the NDBS study (Brown et al., 2000) (which had a far higher budget than BEES) it was originally intended to send out a team of building

science students on street surveys collecting a number of criteria observed from outside the building. Unexpectedly at the time of the initial data collection pilot Google Maps StreetView became available for the majority of New Zealand, making street searches of limited sample areas (or wedges of several cities like the NDBS study) unnecessary. Instead research assistants could conduct a fairly comprehensive recording of a building through a "virtual site visit" without any safety issues or travel costs involved. In fact, for the same research budget as street surveys of just a few hundred buildings, the BEES study could now cover the entire random sample (3,042 buildings) for all of New Zealand.

Having a good quality dataset (both at Websearch sample and census level) meant relationships found between resource use and typology could be conveyed accurately. By combining the best of the valuation and Websearch survey results a good quality national data set was assured. Quality improvements were achieved by investigating:

- the difference between typology mixes in the samples and population
- frequency tables for typology distribution in the samples and population
- inferred typology data from the results of the initial Websearch survey
- the proportion of records with partial survey entry or building identification issues

2.02.03: Websearch Data Coding

Coding and restricted data entry lists (Groat, 2002) were required to make the Websearch survey data entry as efficient and accurate as possible (Section 3.12 and **Appendix i:** 'Websearch Data Sources'). The material options collected in the Websearch on building fabric, glazing and roof type were matched to standard material options available in "Open Studio" (US Department of Energy, 2013b) with the view that this may be convenient for any future energy simulation modelling and the belief that the material options more than adequately represented the variety of materials found in New Zealand. These materials were also linked to the material types collected in the valuation data (QV, 2008) so they could be compared at a later date for quality checks between the Websearch and Valuation datasets. Restricted data entry lists were also provided for office quality grade and built form codes.

Phil Steadman's paper "A Classification of Built Forms" (Steadman, Bruhns, Holtier, et al., 2000) identified the key built-form types (Figure 9) investigated in the NDBS study and these were re-interpreted to suit the New Zealand building stock and tested against the first sample during the pilot stage of the Websearch survey (refer to section 2.02.04). The only Steadman derived built form not regularly found within the NZ building stock was the railway arch built form. The GIS tools used in the NDBS survey (Holtier et al., 2000) were replaced with the quicker / easier option of Google Sketch-up software with which the Victoria University Wellington Building Science and Architecture Students were familiar. The Google Sketch-up software allowed the building model to be geo-located in Google Earth and used with the Open Studio" energy simulation modelling plug-in. The simple 3D building form was also then accessible for Energy-Plus for use in an energy simulation model.

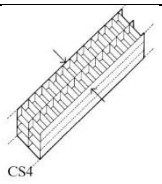
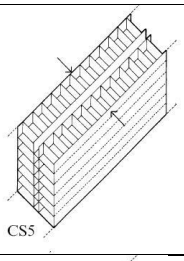
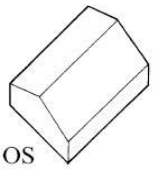
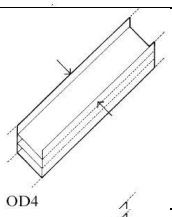
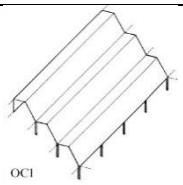
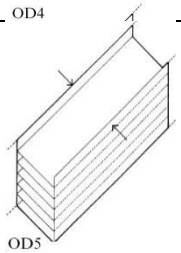
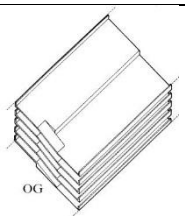
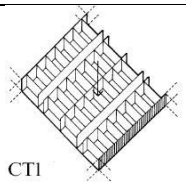
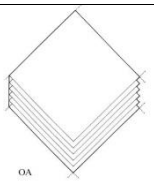
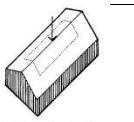
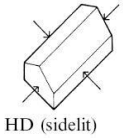
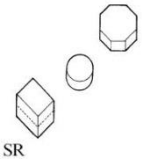
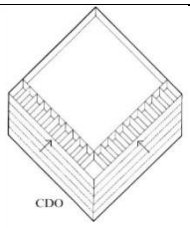
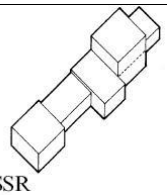
FORM DESCRIPTION	DIAGRAM	FORM DESCRIPTION	DIAGRAM
CS4 - Daylit (sidelit) cellular strip, 4 storeys or less		HA - Artificially lit hall	
CS5 - Daylit (sidelit) cellular strip, 5 storeys or more		OS - Open-plan space in a single shed	
OD4 - Daylit (sidelit) open-plan strip, 1 to 4 storeys		OC1 - Open-plan continuous single-storey space	
OD5 - Daylit (sidelit) open-plan strip, 5 storeys or more		OG - Open-plan car parking or trucking deck	
CT1 - Toplit cellular, single-storey		OA - Artificially lit open-plan multistorey space	
HD - Daylit hall, either sidelit or toplit (or both)	 	SR - Single-room form	
CDO4 - Daylit (Sidelit) Cellular strip with Open Plan Space 1-4 storeys CDO5 - Daylit (Sidelit) Cellular strip with Open Plan Space 5 storeys or more		SSR - String of single-room forms	

Figure 9 Websearch NZ Built Forms based on Steadman's Classification of Built Forms (NDBS)

2.02.04: Websearch Pilot

Measuring the User Experience (Tullis & Albert, 2008) provided guidelines for the pilot work on the Websearch data collection. This pilot identified the time taken to carry out and complete work on a valuation record in the sample in order for the BEES team to price the overall cost of research work for all 3,042 buildings. The analysis included an assessment of the accuracy and efficiency of each task proposed in the survey. The work was carried out by Building Science students who required a short training session. Section 3.07 and **Appendix ii**: 'Websearch Training Notes').

2.02.05: Measured Energy Data and Monitoring

After the Websearch survey was completed, energy bills and phone survey results were gathered by the BEES team from a subset of premises within these buildings. Appliance and lighting counts were noted with measured energy / CO_2 and water consumption monitored over a two week period (N. Isaacs et al., 2010). These research strategies were designed by the BEES team to produce average energy / water-use figures which could then be extrapolated from the measured buildings up to census (aggregate) level using valuation data. Due to the sampling method and the skewed distribution of the energy results, it was not possible to report quartiles and this led the author to the investigation of energy averages according to building typology. Since the Websearch data contains BEES resource consumption collected from physical measurements and energy bills taken from real buildings, this data has to be anonymised in any visualisation available for public access.

2.02.06: Energy Data and Typology Mix

Premise Energy Use Indices (EUI) ($kWh/m^2/yr$) were provided by the BEES team from billing information (including gas) gathered in the aggregate survey stage (Bishop et al., 2012). The results exhibited a high standard deviation and skewed distribution curve which relates to the vast range of building typologies, as well as the varying way buildings are used by occupants. Results were broadly averaged across all premises within a building using data from only the available premises averaged. No allowance was made for floor area differences between premises since this data was not available. Two high outliers were removed from the analysis due to inaccurate source data (Section 3.16).

2.02.07: Benchmark Allocation in the Census

Publicly accessible benchmarks must come from open data sources so everything communicated in the visualisation has to be derived from census level information rather than exposing measured results from the BEES survey which were subject to strict confidentiality agreements (N. Isaacs et al., 2009). Further coding work (Groat, 2002) was completed on the census of commercial buildings to assign building 'New use type' mix (Food, Office, Retail, Special, Hotel, Industrial), climate codes (Cory, 2015), rural / urban mix, form complexity and volume, providing the typology data used to colour buildings in the visualisation. In contrast, the benchmarks were gathered from the Websearch sample buildings alone. Benchmark averages with the lowest possible standard error rates and reasonable sample sizes (energy bills), were used as criteria to colour the census records (Sections 2.03.04 and 3.18 (Hogg & Tanis, 2009)). End-use data (lighting, space conditioning and miscellaneous) were assigned through EECA end-use percentages using the 'new use' code available at census level (Energy Efficiency Conservation Authority, 2007), Section 3.19.

Gaps in the energy data preventing complete coverage of the building stock (in other words those not covered by the BEES and HEEP (N. Isaacs et al., 2005) data sets) include heavy industrial plants, hospitals, apartments, commercial accommodation and education buildings, however many of these building types were looked at within case studies carried out by the BEES team. The construction of

the Wellington census pilot visualisation focuses on the central city and therefore the majority of buildings are represented in the BEES commercial census.

3D data is displayed using cadastral information layers and building elevation height data from Koordinates (Corkery & Coup, 2007), with concatenated addresses and corresponding colour map-marker code. The codes are set according to statistics from the Websearch on typology and energy use as well as further work on end-uses and BEES case study findings (Bishop, 2012b). This provides an automated 3D city model and a set of geo-located map markers, matched by address, directing the distribution of colour (resource use benchmarks). The challenge of automating the full procedure comes down to an issue of linking building forms in the visualisation to benchmark colours set by geo-coded map markers.

The Koordinates website (Corkery & Coup, 2007) is a repository for GIS information such as cadastral data, elevation and council plan data which is being extended all the time. "Whether you're a scientist, analyst, programmer or you simply want to explore the world around you, there are significant barriers to finding and using geographic data. The wide range of formats, projections and sources has made data inaccessible to many. Usually industry knowledge or specialist expertise is required to release the value of geographic data for you and your organisation…"

Whilst the address points (LINZ, 2011b) Land Information New Zealand (LINZ) maintain, allow the geo-coding of address data from the census, the address points are allocated by local councils for rateable properties principally to allocate voters to the correct electorate. The address point data is actively maintained, but incomplete and some locations are incorrect. Nevertheless it is by far the most comprehensive address database available.

"Address points only have a number and a key to a road centreline segment. …Addresses are not related to parcels and should not be a property key because they are not unique, consider a corner section. They do not define property boundaries. Think of addresses as the location of the letterbox marking the entrance to the property, not the building. The mapped point is generally located 15 metres from the centreline of the road at the entrance or at the neck of a rear section. Address ranges on a point are deprecated in the NZ address standard AS/NZS 4819:2003, a single number should be allocated to the principle entrance so the fire service can find it quickly and unambiguously." (LINZ, 2011b)

The disconnect in providing a fully-automated 3D visualisation sourced from a background database comes down to an issue of relating valuation data (and other databases) to physical built form accurately (US Department of Energy, 2013b). Working through the address issues raised, allows one to conceive a programer's set of logic instructions for a tool that will make the connect between addresses and 3D building polygons as accurately as possible. A proposed algorithm of logic instructions will be detailed in Section 3.22, which will correct some of the address errors for the purposes of the pilot visualisation.

The need to improve valuation and address point data vastly, to identify and accurately geo-locate building forms and boundaries has been identified (Stanford Business School, US Department of Energy, & White House, 2012) (N. Isaacs & Hills, 2013). A unique building identifier, (a building equivalent to a computer's "I.P address" (M. R. Donn, 2004)) would make it possible to link together information such as retrofit history, floor area, storey height, heritage status, earthquake status, energy usage and water use to name a few. This would enable valuable mash-ups and statistical visualisations

of separate datasets to occur on a national or local basis and remove critical barriers to the implementation of energy policies for the built environment. (Stanford Business School et al., 2012).

2.03: Visualisation Techniques

2.03.01: Automated Building of a City Visualisation

Clear visualisation of building statistics is vital for the building industry to move towards having a more efficient building stock. Rachel Ryan's masters thesis "Enhancing 3D Models with Urban Information" (Ryan, 2007) suggested that building professionals and members of the general public respond more favourably to 3D graphics than 2D charts, graphs and maps. She also demonstrated that building models can be constructed automatically using a combination of open data such as cadastral building footprints, roof elevation height and topography ground levels. This method has been utilised in producing the Wellington pilot visualisation.

Section 3.21 discusses the proposed methods used to colour buildings in the visualisation and details the use of Google Fusion Table (Google Apps, 2013a) for the geo-location of statistical results. The visualisation was created by manually matching building block forms to their corresponding colour code within Google Earth, however work carried out by Victoria University Wellington's Centre for Building Performance Research indicates that tools such as Microstation's plug-in 'Generative Components' or Google Earth™ with Google SketchUp™ also make it possible to automate the construction and colouring of building forms with parametric information and statistics contained in a database (Sullivan & Motley, 2011), Section 1.05. This would allow the final visualisation to be updated frequently with new background data as it arises.

2.03.02: Precedents / Inspiration: Statistical Mash-ups

The initial inspiration for this thesis draws from online statistical visualisation tools, where users can select their choice of publicly available statistics and view them dynamically in 2D GIS format. Examples include the London Profiler (UCL, 2008), New York Energy Map (Howard et al., 2012) (Columbia Engineering School, 2011), LessEn Energy Map (Figure 10 - (Urban Land Institute, 2009)) and the Berlin Atlas of Crime (Berlin Police, 2012). The ability to toggle from one 2D graphic to another animates the visualisation and helps the user understand relationships and correlations which may not be immediately apparent when viewing the same data in report, chart or book format. In the case of 'LessEn Energy Map' energy results are mapped as geo-

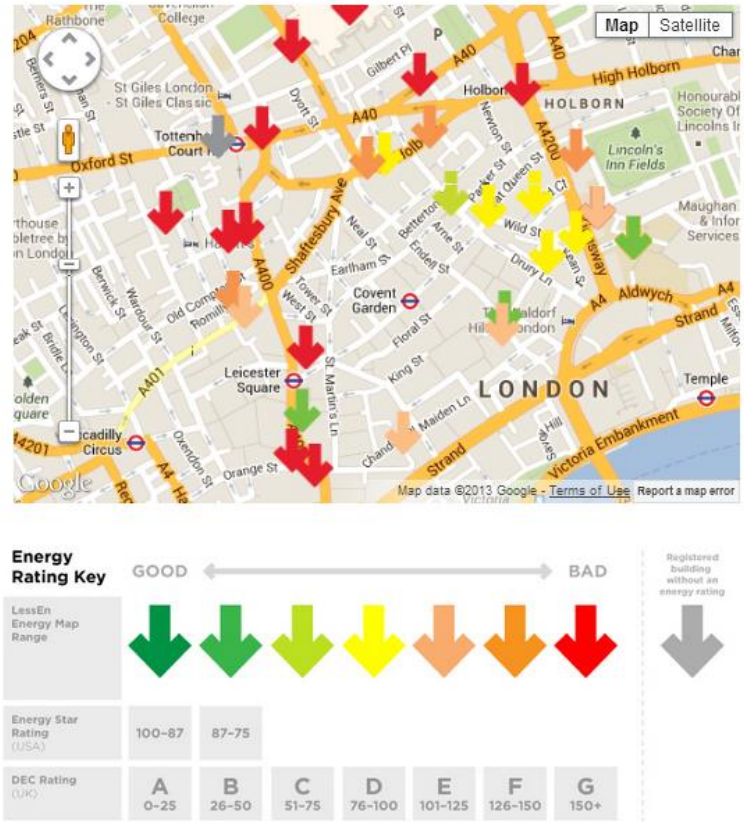


Figure 10 LessEn Energy Map

coded coloured arrows (Figure 10) for sporadic data points (those buildings with current energy certificates or those submitted to the model by users voluntarily). These match the colours used in European Display Energy Certificates and can be viewed in Google Maps StreetView on a desktop or using a mobile internet device in-situ. Section 1.03 introduces statistical visualisation precedents.

Understanding building energy in the context of a whole country has been explored in the USA's Buildings Performance Database (BPD) (U.S. Department of Energy, 2013). This enables users to perform statistical analysis on an anonymous dataset of tens of thousands of commercial and residential buildings from across the country. Users can compare performance trends among similar buildings to identify and prioritize cost-saving energy efficiency improvements and assess the range of likely savings from these improvements. This is a quality tool for investigating building energy policies and understanding aggregate / national building energy statistics, however the inability to view the information applied to local building stock means that the user is separated from the real meaning of the statistics as it applies to them. Although the tool is enjoyable and informative to use, energy as a concept remains intangible in its aggregate form, and its use, distribution within a city, or local neighbourhood remains completely unclear.

Figure 11 Tweets + Flickr Uploads (Orange-Flickr, Blue-Twitter, White-Both)

Visualisations have moved into the art of geo-coding Flickr image downloads (Ludicorp, 2004), reviews (Yelp, 2013), tweets (Costolo, 2014) and status updates to produce heat-maps overlaid onto a city, country or world map locating ideas, events and thoughts spatially. In the realm of visualising the invisible, it is now possible to infer a population's mood or opinions and plot this on a map (Austen, 2013). Figure 11 illustrates a snapshot of Twitter and Flickr updates in 2011 called 'See Something or Say Something' by Eric Fischer. Both Flickr and Twitter have accessible programming interfaces (API) which allow huge amounts of open data to be explored. Examples of visualisations pushing the boundaries of graphic representation can be found in the website Flowing Data (Yau, 2013a), StatsWorld – Interactive Maps of Open Data (StatSilk, 2013), or the plethora of examples contained in the book "Data Points: Visualization that Means Something" (Yau, 2013b). The mark of a quality visualisation "…helps you see what you otherwise would have been blind to if you looked only at the naked source". (Yau, 2013)

Mashblock (Prebble, 2010) is another illustration of the power of open data visualisation. It uses meshblock level data to breakdown the NZ census statistics into small chunks with pie charts, histograms and a map outline. This provides the general public with highly detailed information about neighbourhoods, their average mode of transport, religion, employment, age, language and income, using freely available data. This kind of strategy applied to groups of buildings would be valuable – communicating all the characteristics and energy information available from the

Websearch to meshblock zones. In theory if meshblock-level delivered energy became open data it would be possible to publish the energy delivered to these parcels of land (top-down) and allocate energy to the mix of buildings (bottom-up). This would be restricted in areas where there were only a handful of buildings per meshblock due to the potential to reveal sensitive data. This issue is also dealt with in Mashblock with the screening of meshblock level data in the centre of cities – where only the larger scale 'area unit' data can be investigated.

2.03.03: Graphic Considerations

Turkle's book "Simulation and its Discontents" (Turkle, 2009) talks of 3D digital space. It suggests that a digital model is not simply shown, it is "performed" and in the process observers are brought into a new relationship with the visual. "Simulations let you manipulate what was on the screen; more recently simulations encourage you to inhabit worlds or,…fall into them… From the earliest days, simulation seduced" (Turkle, 2009). The aim of the visualisation proposed by this thesis is to seduce the public with an invisible concept – building resource use. It provides a new way of thinking about the built environment, changing behaviour with regards to purchasing, occupying or leasing property that would in turn catalyse change in the market value of sustainable buildings. 3D visualisations are easily accessible and have this seductive power (Turkle, 2009).

Literature reviewed on visual and graphic representation styles and accessibility included Edward Tufte's "Beautiful Evidence" (Tufte, 2006) which stresses the importance of image as evidence and explanation, the principles of analytical design and warns of corruption in evidence presentations such as cherry picking and "chart junk". When asked about the art of presenting 3D digital cities in an interview in Cincinnati (Tufte, 2009), he argues "Don't ask how visualisations techniques can help display data. Ask how data can be best represented". In other words, understand and serve the data, not the technique or the software used. "Increase information resolution (the maximum useful bits per unit area and unit time)…Treat all problems as multi-variate, make wise comparisons, show causality and use whatever it takes to get the message across". In "Beautiful Evidence" Tufte also covers the shortfalls with the use of Excel charts and Powerpoint in presentation. Techniques used to combat these shortfalls include avoiding comparison between data with different axis scales and comparing 'apples' with 'apples' and clearly labelling points of interest and difference (Section 7.02).and these were taken into account in the resulting typology analysis and visualisation graphics conducted as part of this thesis.

'Understanding Playful User Experience Through Digital Games' (Korhonen, Montola, & Arrasvuori, 2009) acknowledges that the acceptance of a product depends on utilitarian and non-utilitarian properties and argues that "product design needs to improve the support of playful experiences in order to fit in with the users' multi-faceted needs". Korhonen et.al believe we live in an "experience economy" where consumers constantly seek new kinds of experiences from the products they use. Focussing on the effectiveness and efficiency of the user is often considered to be an adequate goal for the success of a product or a service but "positive emotions are essential for the sake of curiosity and the ability to learn new things" (Korhonen et al., 2009) . Research on playfulness (Sutton-Smith, 1997) listed below, is compared with visualisation design format proposed by this thesis (*italics in brackets*). Sutton-Smith suggests that 'play':

- goes hand-in-hand with learning *(this suggests a game platform may further encourage and inspire learning)*

- is where choices are dictated by luck or destiny, (*you can highlight or isolate certain building types or areas of interest raising questions about the pattern of energy use that would be invisible otherwise)*
- enforces the power status of the winning players *(advertising sustainable accomplishments or finding you have a workplace better than the benchmark introduces a level of competition)*
- is a means of "confirming maintaining or advancing the power and identity of the community of players" (*allowing users to see the change in the building stock occurring over time may give a sense of advancing power and identity to the users of the visualisation)*
- is imaginary as applied to creativity and "playful improvisation" in arts (*trying out new combinations of data together to find new meanings perhaps not envisaged by the author of the visualisation)*
- focusses on the enjoyment or fun aspect of the participating players (*3D visualisations of cities are still intriguing and novel to most people, flying around your local neighbourhood or place of work and virtually exploring urban space with and additional layer of information added should only serve to provide further enjoyment*)

Korhonen et.al believe these features of play all work together to improve the enjoyment of a product. Where play is combined with learning about an important, yet illusive concept, playfulness must surely be crucial to ensure that message gets across to the widest possible audience (Korhonen et al., 2009).

"Measuring the User Experience" (Tullis & Albert, 2008) looks specifically at the use of focus groups and usability questionnaires to test the effectiveness of tools. The book covers the analysis of user behaviour and attitudes, understanding and responding to the results of usability questionnaires, and some of the techniques it discusses were utilised in the Websearch survey design. Data collection task timing was measured before and after changes were made to the criteria collected - ensuring the average time taken per record was not unsustainable financially for the project (Section 3.07).

2.03.04: Producing Benchmark Energy Average Colours

Although the BEES research strategies were specifically designed to produce average-only energy results for extrapolation up to census level (Section 2.02.05), the visualisation proposed by this thesis required a set of seven energy / water benchmark ranges, in order to make use of internationally established colour scales for energy use. Providing just above and below average benchmarks would not have resulted in a range of energy benchmarks with which to compare real consumption data, so an alternative methodology was required to produce various benchmarks dependent on building typology and characteristics uncovered by open data. A rule of thumb (Hogg & Tanis, 2009) for minimum sample size when calculating a mean suggests that a combination of building characteristics with at least 25-30 energy results should provide a reasonably accurate average energy intensity estimate for that typology. To further quantify the standard error for each benchmark a simple calculation was made of the standard deviation divided by the square root of the number of results in the sample:

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Where 's' is the sample standard deviation (i.e., the sample-based estimate of the standard deviation of the population), and n is the size (number of observations) of the sample.

This was finally expressed as a percentage +/- error in relation to the original benchmark average value by multiplying the standard error by 1.96 (two standard deviations) and dividing the result by the original benchmark average result. By systematically working through all available public data in the Websearch and investigating energy intensity averages and standard error for each criteria, the author was able to allocate the most reliable of the resulting energy benchmarks to all buildings within the population. Since the city visualisation is to be automated, as more open data becomes available (where building owners submit real data about their property to advertise energy efficiency accomplishments) the database becomes increasingly detailed and refined. Within each particular building typology - the visualisation will attempt to present the average building energy use index (EUI) the typical end-use break down by building use-type (lighting, space conditioning and miscellaneous) as well as comparing average energy use indices across various climate zones throughout New Zealand. Rather than following a linear scale, the seven benchmark colours were initially set using 'natural breaks' between the valid benchmark averages when sorted in order from low energy to high energy averages.

The energy results provided by the BEES study have a heavy tailed distribution, indicating that many end uses have values far in excess of typical (median) or average (mean) usage. The classification of energy intensity into seven ranges (red through to dark green) have been calculated using natural breaks due to as shown in Figure 12. The precedent for using natural breaks to determine the best arrangement of values into different classes is known as the Jenks Optimisation Method (Jenks, 1967). The reported benchmark value for each of these ranges was provided by averaging available EUI results between each of the range breaks listed above with the yellow range reported as the average for all energy results (210 kWh/m^2/yr). For the pilot visualisation of Wellington the average values for the whole building record ranges (kWh/m^2/yr) were: red 314, orange 262, light orange 245, yellow 210, lime 174, turquoise 157, dark green 104 (Figure 12). Different values were calculated using the natural break method for the various end-use breakdowns according to the resultant values when end-use percentage breakdowns are applied.

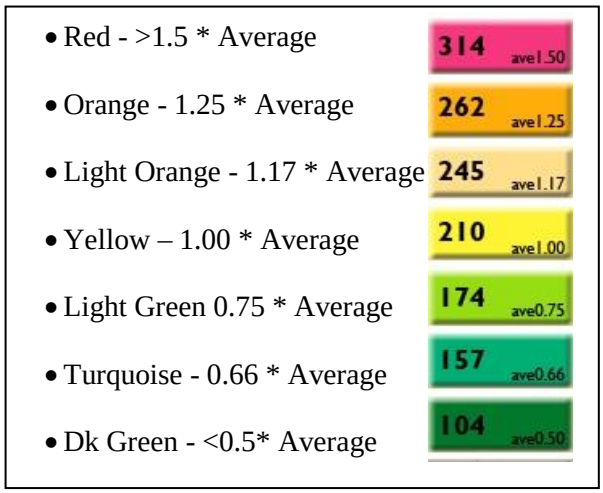


Figure 12 Dashboard Scale

2.03.05: Colour Ranges –Appliance Labelling and EU Display Energy Certificates

A range of seven colours was selected for the visualisation due to its history of use as measure of energy efficiency in Europe, first in home appliances and more recently in building energy performance certificates. The European Union (EU) Energy Label was established in 1992 under an EU Directive 92/75/EC (Council of the European Communities, 1992) to encourage businesses to market products and services that meet high standards of environmental performance and quality. Due to public familiarity with this system, the home appliance label, layout, colours and format were developed for use as a building energy performance certificate resulting in a new European Standard prEN 15217 (European Committee for Standardization, 2007) "Energy Performance of Buildings - Methods for Expressing Energy Performance and for Energy Certification of Buildings" enacted in 2007. The use of green (low/good/cool) powerfully suggests low energy use and environmentally-friendly and red (high / bad / hot) intuitively indicates high energy and poor environmental performance. It would also seem logical to relate benchmark energy use colour ranges to the

European Union's adopted Display Energy Certificate colour scales (Figure 1) where red is very high energy use and dark green is very low use since this is an established seven-colour system.

These seven-increment colour scales set the level of accuracy and hierarchy required for the energy and water benchmark ranges and the uneven number of stages provides a central range with which to express the average benchmark for the building stock.

2.03.06: Colour Interpretation and Colours for Impaired Vision

Whilst using a spectrum colour scale from red to green would seem to be a logical and intuitive colour range to communicate high / low energy for a person with normal vision, it is the worst-case scenario for those with any form of colour blindness. According to the paper "Task-based Color Scale Design" (Rheingans, 2000) a spectrum colour scale is only intuitive to those with a mental model of the progression of wavelengths of light and reading this scale can require some training time. For people with normal vision where the scale starts with red for high and uses violet for low values there is potential for confusion between the extreme values. The benchmark colours used for communication of energy results omitted the purple and blue opting for dark green for low values as adopted in energy certificates. The other issue is that the yellow result tends to be in the middle of the scale and this is very striking, tending to draw the eye to the average result rather than the extreme values which in the case of energy use, are presumably of primary interest. To respond to this issue in the pilot visualisation, the yellow average colour, was lightened within the colour range to reduce boldness. It is also possible to provide extra interactive tools in visualisations to improve accessibility to those with impaired vision / colour blindness as well as provide an alternative option for those who do not understand the progression of wavelengths of light. By clicking any colour in the legend scale, the buildings of a certain colour could be highlighted graphically. Alternatively a simple switch control could modify the overall scale to a bivariate hue-saturation scale to overcome most types of visual deficiencies.

The limitation of the use of a single colour hue with variable brightness is that there are less distinguishable display levels. Since the visualisation proposed by this thesis is a 3D model it is important that the colour scale used works best when viewing a 3D surface. Even basic colour renderings such as Google Earth use shading cues allowing the viewer to judge the 3D shape of representation objects. A colour scale that included brightness, would interfere with the brightness from rendered shading calculations (Rheingans, 2000). Whilst the colour block forms produced by Google Earth are not rendered with shading unless the sunlight setting is on, the basic building block colours are represented with different saturations of the same colour spectrum to communicate their 3D light and shade. Any colour range based solely on brightness / saturation rather than hue would be ineffective in communicating the 3D forms adequately. There is a whole field devoted to the study of haptic feedback for communication of volumetric data which would also make the information accessible to the blind (Institute of Electrical and Electronics Engineers, 2009). Further details about the graphics and presentation systems used follows in Chapter 4: 'Graphic Representation'.

3.00: Research Methodology

The resource-use benchmark data used in the 3D visualisation pilot of Wellington CBD has been drawn from the 'BEES' Building Energy End-use Study. This Energy Efficiency and Conservation Authority (EECA) (Energy Efficiency Conservation Authority, 2007) / Ministry for Business Innovation and Employment (MBIE) (MBIE, 2013) and Building Levy (Parliamentary Counsel Office, 1969) funded-project was managed from 2007 – 2013 by the Building Research Association New Zealand (BRANZ, 2013).

The following section provides an overview of an innovative and affordable method used to collect high level data about the built environment in the BEES project, designed and supervised by the author. It discusses methods used to identify and correct suspect building records in valuation data (used as a proxy for a New Zealand commercial building stock census) and search for possible bias in the mix of building typologies present in the various BEES sub-samples when compared with the census data. A discussion follows of the methods used to compare building typologies with confidential BEES energy bills results, arriving at a set of publicly accessible average energy benchmarks for all commercial buildings in the census. Finally a method for automatically assigning heat-map benchmark colours (expressed in kWh/m^2/yr) to a 3D city model of Wellington CBD is presented.

3.01: Research Methodology Flowchart

The aim of the visualisation proposed by this thesis is to seduce the public with an invisible concept - building resource use. This is intended to raise questions about the built environment, raise the value of sustainable properties and indicate potential relationships between neighbourhoods comparing resource use with building typology. The author's involvement in the BEES project included the research design and supervision of data collection eventually coined 'Websearch'. This was a survey for high level building typology data through internet research and street surveys by Building Science students from Victoria University Wellington. The BEES team embellished this database further with phone survey data and billing information collected from a sub-sample with a smaller sample monitored on site for actual energy use. The Websearch and subsequent BEES findings provided the author with access to a rich database of New Zealand commercial building typologies with energy results attached to a sub-sample. This was eventually mined to produce energy benchmarks for the pilot visualisation.

The research methodology flow chart on the next page (Figure 13) illustrates the research boundaries and sources of each task in the BEES project as a whole as well as the thesis visualisation and statistical analysis work flow. Further detail in each text box reveals sample sizes and percentages in terms of the total NZ building stock. Dotted lines / arrows indicate additional work that would be necessary to fully automate the visualisation, reduce error rates and remove gaps in the data in the interest of accurate representation of the entire building stock. The colours indicate the data sources, the author responsible for the work and topic areas involved in the completion of this thesis. All work flow items noted with a star in the following methodology chart were conducted by the author as part of this thesis.

Work Flow - BEES Research Project:		**Work Flow – Thesis Visualisation:**	
Orange -	BRANZ measured data (BEES)	Purple -	Quality checks / validation
Navy -	Valuation data (Quotable Value)	Grey -	Pilot study and review
Red -	Websearch (BEES) including statistical analysis for this thesis)	Blue -	Bias + typology mix checks
		Green-	Pilot visualisation work.

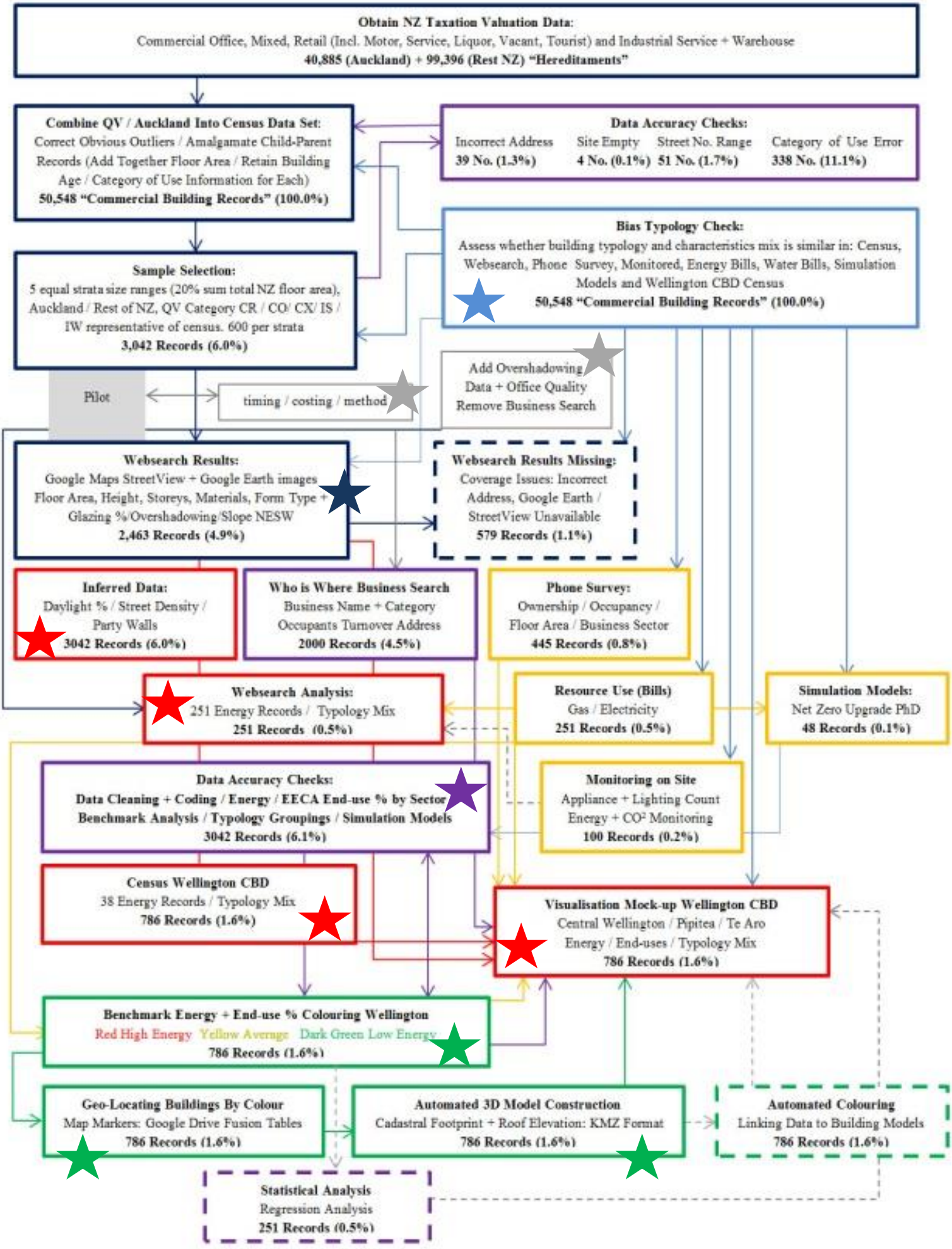


Figure 13 Methodology Flowchart Indicating the Research Boundaries and Authors of each Task

3.02: BEES Census Data

The BEES team identified a database of 50,548 New Zealand valuation records that form a census of all of the commercial, industrial service and warehouse buildings in New Zealand. The records were selected on the basis of their Quotable Value category code. Commercial: (Office, Mixed and Retail) and Industrial: (Service + Warehouse) categories which were all of interest to the BEES team. The retail group included several smaller retail-based categories: Commercial (Motor, Service, Liquor, Vacant, Tourist) so as not to miss any retail-based records. Every hereditament (a property subject to a tax rating assessment) containing any one of these categories was selected along with its "parent" and additional "child" records regardless of use category. The original valuations totalled 40,885 Auckland and 99,396 New Zealand 'hereditaments', however these were brought together into a database of 50,548 buildings using the address field as the key to amalgamate records into buildings. This method is subject to identification errors due to the poor linkages between addresses and buildings (Sections 2.02.07, 3.03 and 5.08).

3.03: Initial Corrections and Amalgamation

Before the random selection of census records could take place it was important to amalgamate two styles of data collection into a single database. One dataset covered Auckland (Auckland Council, 2008) and the other covered all remaining properties in New Zealand (QV, 2008). This meant totalling up the floor areas between parent and child records, combining categories of use and building age data from these child records into a single parent entry and noting any building age difference between the valuations to return a mixed age range. It also required the bringing together of two disparate data entry formats with differing definitions, rules and data collection methods. For example, in the Auckland database 'Year Built' field would be the exact year of construction entered into the database, whereas in the QV data field 'Building Age' would be rounded to the nearest decade including a large category known as 'b-1920' for any buildings built prior to 1920. The resulting data for the whole of New Zealand was adjusted to the lowest common denominator (decades from 1920 onwards) so that the information could analysed together.

An initial sub-sample of 3,121 incomplete or suspect records was investigated for accuracy by a team of Building Science research assistants from Victoria University Wellington, using Google Earth aerial photos and Google searches on addresses. The data was analysed for obvious floor area outliers (for example ft^2 instead of m^2) address errors and other data entry issues. This resulted in an initially 'cleaned' commercial building stock census numbering 50,548 buildings with a basic understanding of the proportion of records with floor area and address errors. To maintain search consistency the records were divided by locale so that any given area was surveyed by the same person and where possible using research assistants with local knowledge of a particular area around New Zealand.

The level of error present in the census was assessed and corrected where possible since this dataset would become the BEES sample frame. This was important since energy (and eventually water) results would be extrapolated up to census level compounding any floor area or use category errors in the process potentially increasing error rates accordingly. Figure 14 illustrates the error rate (6%) in the census of commercial buildings prior to validation work (left) and after (2%) the initial validation and data cleaning work was completed (right). It provides a breakdown of error type including address error, missing floor area, low floor area, largest floor area and incorrect category of use. Of the 3,121 (6.1%) buildings with suspect data in the commercial building census, 2,163 (69.3%) were corrected prior to the sample being drawn for the BEES project.

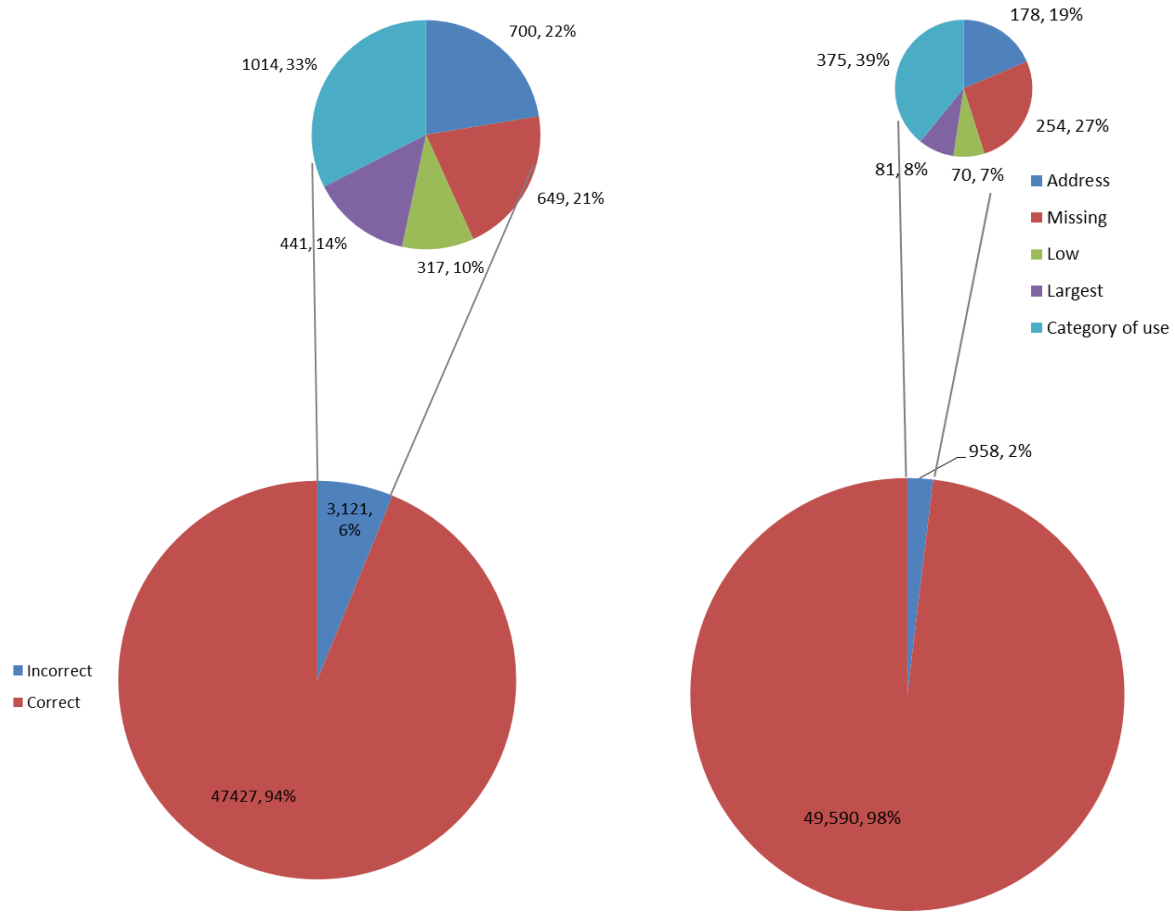


Figure 14 Initial Validation of Incorrect Census Buildings Before and After Correction

Accuracy in the census was also important for the purposes of the thesis visualisation. The mix of building typologies in the underlying census data would eventually be compared with each of the BEES derived sub-samples to ensure they were as representative as possible. Any sample selection or survey participation bias would have the potential to increase error rates in the resulting energy benchmarks. Being able to compare the mix of typologies (and eventually their applied energy benchmark averages) in the sample, sub-samples and census would later serve as a crosschecking device for accuracy refer to (Figures 25-32 and Sections 3.13-14).

3.04: BEES Sample Selection

A stratified random sample was selected by the BEES team from the census of the commercial New Zealand building stock which would eventually become the Websearch database. The strata were made up of five floor area, two location and five categories of building use groupings. The five floor area strata were created by totalling floor area for the whole of the commercial building stock and dividing this total area into five 20 percent groupings. These grouping boundaries set the floor area size ranges for each strata (Figure 15) (N. Isaacs et al., 2010).

Floor Area Strata	**1**	**2**	**3**	**4**	**5**	**Total**
Floor Area	*<650 m^2*	*650- 1499m^2*	*1,500-3,499 m^2*	*3,50- 8,999m^2*	*>9,000m^2*	
Approx. No. Builds	33,781	10,081	4,288	1,825	564	**50,548**
% of Buildings	67%	20%	8%	4%	1%	**100%**
Total Floor Area (million m^2)	9.9	9.6	9.5	9.6	9.8	**48.3**
% Total Floor Area	20%	20%	20%	20%	20%	**100%**

Figure 15 Size Strata Sample Allocation

Approximately 600 buildings were randomly selected from each size strata forming a database of 3,042 building records accounting for 6.0% of the commercial building stock by building count. Due to the uneven number of buildings by count in each size strata (600 records representing 33,781 census buildings in strata one, and only 564 records representing 564 census buildings in strata five) the random sample would only be able to be analysed in relation to aggregate building stock floor area. This means statistical analysis of building count would be meaningless without conversion rates applied for each size strata.

The remaining sample strata were based on location ('Auckland' (Auckland Council, 2008) valuations and 'RestofNZ' from the Quotable Value database (QV, 2008)) and category of use strata (Commercial Retail - CR / Commercial Office – CO / Commercial Mixed – CX / Industrial Service - IS / Industrial Warehouse – IW). Both location and category of use strata were proportionally represented in the sample according to their level of occurrence in the population.

3.05: Data Collection Design – Team Meetings:

Through a series of BEES team research design meetings conducted throughout 2008, a list of typology and site surroundings criteria were collated (Isaacs et al., 2009). The information collected were parameters that the BEES team collectively agreed may drive energy or water use in buildings. These criteria would allow the team to compare coarse typology and surroundings data against actual building energy results. It might also allow oft-quoted theories to be tested in the New Zealand context - for instance, that a high glazing ratio can decrease lighting energy use but increase the need for space conditioning. The high level (coarse) data fields the team initially agreed to collect were: floor area; storey height; glazing ratio; materials; form-type; building age; use-type; as well as elevation images and aerial photos from Google data.

One of the BEES research deliverables was making the information uncovered from this research publicly accessible, so it was important that the criteria collected were all details apparent from the exterior of the building and from other open data sources so that the results could eventually be used to communicate the results of the BEES research. This strategy mirrored the considerations taken in the UK's NDBS street surveys (Brown et al., 2000) Section 2.01.01.

3.06: Street Surveys

A building's size, shape, construction and occupation must all be identified before any sense can be made of its energy patterns. In most of the international examples of energy surveys previously reviewed (Section 2.01) site visits were conducted to assess building characteristics and suitability of the building for study. BEES initially began site visits of Websearch sample buildings and these visits were conducted from the street to record the constructional materials, size and orientation as well as addresses, building uses and business names associated with the building. The information was checked against valuation data and notes were made of any errors.

The methodology of data collection was reviewed not long after commencement of the site visits, as Google™ Maps StreetView® became available throughout most of New Zealand. An innovative method was introduced to collect high-level publicly available data about the built environment in a speedy and cost-effective way. Using StreetView, cheaper and safer 'virtual site visits' could take place instead of physical site visits. The data was collected from Google™ Earth aerial photographs and Google™ Maps StreetView® omni-directional images and a business name search was included for each record. Victoria University Wellington Building Science students were carefully briefed on the research work and regularly supervised throughout the process. **Appendix ii:** 'Websearch Training Notes' includes the information given to each researcher. Their work was checked and collated by the author into a master database over a number of university holiday periods. Not only did this resulting output require vast data cleaning and coding to be suitable for use in later statistical analysis work, but the various data fields collected could then be combined and studied further to infer characteristics that were not anticipated in the initial stages of the project (Section 3.10).

3.07: Pilot for Method, Costings and Timing

A pilot was required to assess timing and methods used in the data collection which would provide the BEES team with an estimate of the research budget required to complete the sample (approximately 1,000 - 1,500 person-hours). The data collection method was rigorously tested in a pilot survey utilising the same building science students from Victoria University Wellington involved in the earlier validation work.

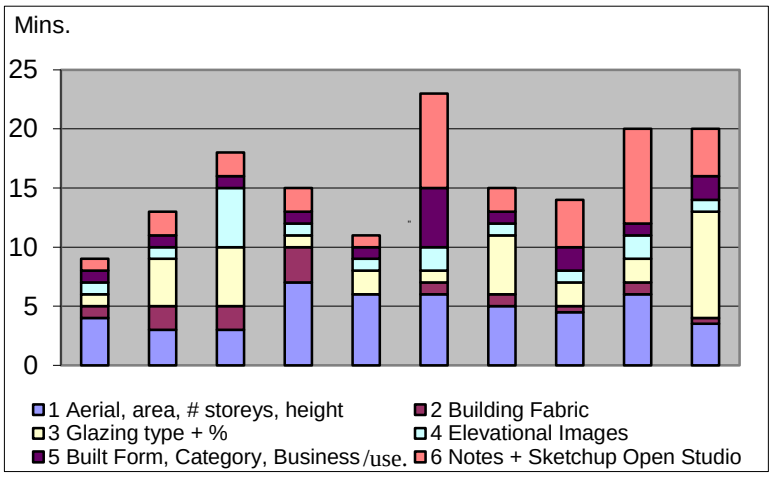


Figure 16 Pilot Sample Tasks

Of the 53 randomly selected pilot records completed 6% were missing Google Maps StreetView coverage but all had aerial photo coverage so most criteria were able to be collected for each record. Timing five research assistants completing the 53 pilot records over two afternoons the average time to complete a Websearch record was established as 15 minutes 45 seconds. Figure 16 illustrates ten buildings where each group of tasks was individually timed, partly to inform the layout and design of the data collection and partly to weigh up the time / money spent on the work versus the importance of the criteria selected. Further investigation of the breakdown in average task timing, before and after changes made to the data collection, are illustrated in Figure 17. The average task times were recorded for a sample of 10 buildings (two completed by each research assistant). This demonstrated an understanding of the likely costs of the project and enabled the case to be made for increasing the number of criteria collected. The modifications made after the initial timing test, were the inclusion of restricted data lists in Excel (which ensured consistent data entry); a slight change to the order of information collected (to improve efficiency and reducing switching between applications used); as well as the introduction of additional data fields, which together increased the average timing from 15 minutes up to 18 minutes. A discussion of the new data to be collected is covered in the next sub-section.

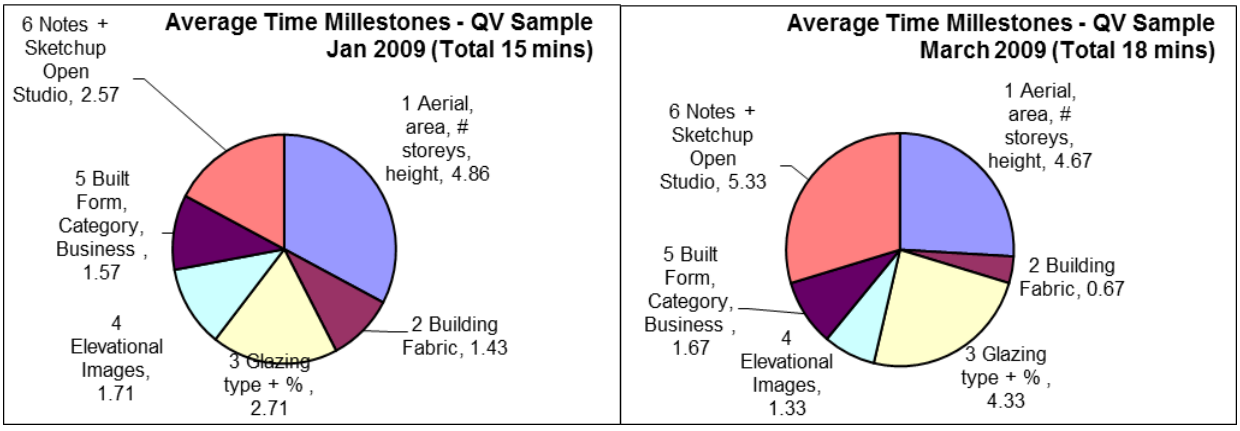


Figure 17 Pilot Websearch Timing Task Breakdown Before and After Pilot Modification

3.08: Data Collection Review + Steering Group Meeting

A cost-effective budget was established for high level data collection for the full Websearch sample (3,042 building records), and this led to modifications to the criteria collected. Overshadowing and site slope details were introduced which helped infer many more useful data fields in the final analysis detailed in the next sub-section. To further cut down wasted time to allow for the extended data collection, the business search was removed from the Websearch work flow, since this was to be provided by a separate business name and category search service known as 'Who is Where?'(E-ideas, 2013). This service maintains a list of New Zealand Businesses which is kept up to date by regular phone surveys.

The tests also introduced a simplification of the glazing frame data collection method where the window frame material only was noted, rather than including double and single glazing which was not easy to discern from Google StreetView photos. The Excel worksheet data fields were also reordered so that work gathered from aerial photos and StreetView were grouped together making data entry more efficient. This avoided the need to switch application screens so many times during the data entry process. A subsequent BEES steering group meeting reviewed the criteria further and introduced the collection of Office Quality grading ((Property Council of NZ, 2014) A, B, C, D) to tie in with research they were doing on building stock value and ownership.

3.09: Websearch Data Collection

The original property valuation information already included floor area, floor plate area, building age, building fabric and detailed category of use recorded when the record was first entered in the database. In addition, the integration of the 'WhoisWhere?' database generated business names and categories of use based on street address. The data was then enriched during the Websearch survey using Google Earth™, Google Street View™ and internet searches to correct the existing valuation data and provide further details on building and surrounding site characteristics.

Built form type was noted according to the Steadman-derived diagrams (Figure 9) which recorded whether the building was a single room, cellular space, open plan space or various combinations of the above. Glazing ratios were estimated to the nearest 10% for each building orientation visible in Google StreetView elevation imagery and images capturing as much of the building elevation in one screenshot as possible were linked to the database. Elevation image file names included a code indicating the closest orientation direction (N, E, S, W / NE, SE, SW, NW) for the façade and these, along with aerial photograph links from Google Earth, enabled the Websearch data to be quickly quality checked at a later stage. For each building record visible in StreetView an estimate was made of the number of storeys, the overall building height, wall materials (primary and secondary), surrounding building heights and an assessment of any change of level across the site.

Aerial photos were examined on Google Earth and a note was made of the likely roof material/s with primary and secondary fields available for mixed material roofs. The distance to surrounding buildings from each building orientation was recorded and the floor plate (footprint) area was estimated with the measurement tool in Google Earth. Lastly, a simple building energy performance simulation model was created for each building record and positioned geographically using Google SketchUp® (Google, 2013b) with the EnergyPlus™-ready (US Department of Energy, 2013a) plug-in 'OpenStudio™ (US Department of Energy, 2013b).

No. of Buildings Associated with Single Address	No. of Websearch Sample Buildings (Building Code Letter)		%
One Building	('A''s only)	2,787	91.6
Two Buildings	('A' - 'B')	165	5.4
Three Buildings	('A' - 'C')	47	1.5
Four Buildings	('A'- 'D')	24	0.8
Five Buildings	('A'- 'E')	9	0.3
Six Buildings	('A'- 'F')	7	0.2
Eight Buildings	('A'- 'H')	1	0.0
Ten Buildings	('A'- 'J')	1	0.0
Eleven Buildings	('A'- 'K')	1	0.0

Figure 18 Records with Multiple Buildings Associated

Records where numerous buildings were found at a single valuation address were dealt with by duplicating the original sample building record 'A' and recording the details for each additional building ('B' – 'K') discovered at the address - Figure18. The Websearch information was collected on all of these additional buildings by duplicating the valuation building record row and modifying the Websearch survey details accordingly, for each building where accessible from aerial or StreetView imagery. Where full information was not available research assistants were able to make educated assumptions and note them in the 'notes' column of the database for later analysis. **Appendix i:** 'Websearch Data Sources'.

The Websearch data collection work provided a deeper understanding of the typical errors found in the valuation data, and this was identified by the research assistants through the 'notes' field. This was crucial considering the data was being treated as a proxy for a census of the New Zealand commercial building stock, so it was important to establish any major errors in the dataset or shortcomings in the method used to amalgamate separate records into 'buildings'. This meant allowances could be made in manipulating the data when estimating aggregate energy for the commercial building stock as a whole, but also in relation to the benchmarks eventually applied to the pilot visualisation.

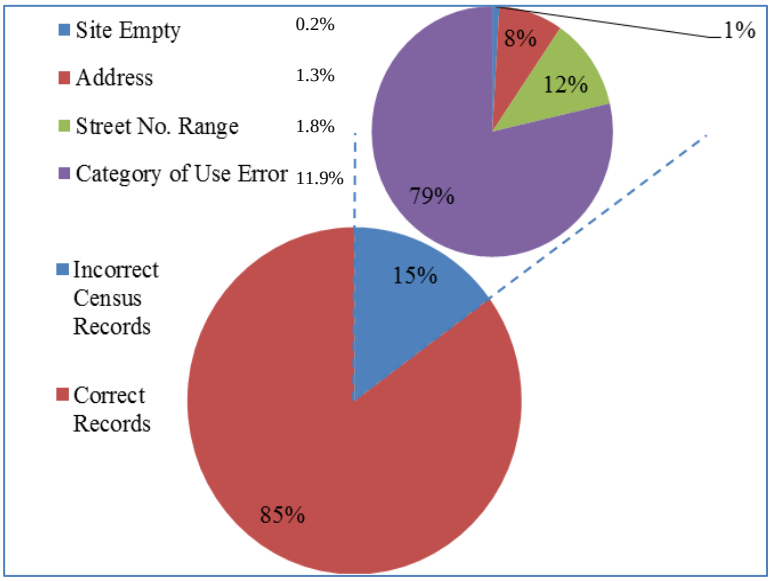


Figure 19 Error Rates Found in Websearch Data

Although the notes were entered as free text, which picked up all manner of issues and assumptions made in the data entry, the verbose fields were later searched for certain text strings and coded so that typical errors or issues with the data entry could be understood en-mass. These valuation errors, assumptions and missing buildings were identified over and above the floor area and address errors corrected during the census validation work (Section 3.03). Typical errors included finding empty sites, incorrect category of use codes, incorrect addresses or just a minor street number range description issue with the address. Although 15% of the building records are flagged with an error code, the majority of these errors relate to an incorrect category of use code (11.9%) as shown in Figure 19. This alerted the BEES team to the issue, which was resolved by cross-referencing to the business categories provided by the Who is Where? service.

When checking some of the records noted as having street number range errors (1.8%), it was found that some of the street numbers given in the valuation spreadsheet did not correspond exactly to the numbers given in Google Maps™. This error usually indicated buildings with a street number falling within the street number range, but not confirmed as an exact match. This was also true of buildings that had large frontages along more than one street. The address given by the valuation data may have been 'around the corner'; perhaps due to a business preferring to be identified on an adjacent street rather than its official postal address street number with only a small subset affected. These errors are considered minor and do not prevent data being collected.

The combination of aerial photos, Google Earth address placemarks and Google map street number labels allowed addresses to be identified correctly in the majority of cases and only 1.3% of the building records in the sample were found to have a completely incorrect address. In these cases data was still collected for a building best fitting the original details provided by the record, for instance by looking for a building with equivalent floor area in the adjacent area and matching the category of use / materials or any other details available in the source data.

These errors are considered to be a minor issue which would be unlikely to have a negative impact on results overall – especially when dealing with benchmark ranges rather than absolute quantities. Error

codes merely suggest further investigation or cross-checking may be required to ensure the results are accurate and relate to the correct building selected in the sample, if this data point is to be heavily relied upon in future research using the Websearch data.

3.10: Inferring New Data Sets

The basic Websearch data collection allowed further details to be inferred regarding the building and its surroundings: daylight factors for each elevation; average daylight factor for available facades (visible in StreetView); number and orientation of party walls; and significant ground slope. Average glazing ratios per building could be calculated and the compactness of the building form was estimated from area calculations.

An understanding of surface area and compactness could be estimated by comparing the calculated area value (floor plate multiplied by number of stories) with the estimated floor area as provided by the Websearch measurements. A simple factor division between the two area Figures provided a ratio of compactness and surface area with 0.9 to 1.1 being 'compact' with under 0.9 and over 1.1 both being associated with 'complexity'. The formulas used to infer extra information from the Websearch data are listed below, calculating:

- Daylight factor for visible elevations: =DEGREES(ATAN (*Average Distance to Obstruction/Average Height of Obstruction*))
- Average Daylight Factor: =AVERAGE (Daylight %North, Daylight %East, Daylight %South, Daylight %West)
- Glazing Ratio Average: =AVERAGE (*Daylight %North, Daylight %East, Daylight %South, Daylight %West*)
- Significant Ground Slope (assumes a square building footprint):=DEGREES(ATAN (*Average Ground Level Above Datum South – Average Ground Level Above Datum North) / Square Root of Floor Plate Area*)

A full list of the inferred data is illustrated in Figure 20.

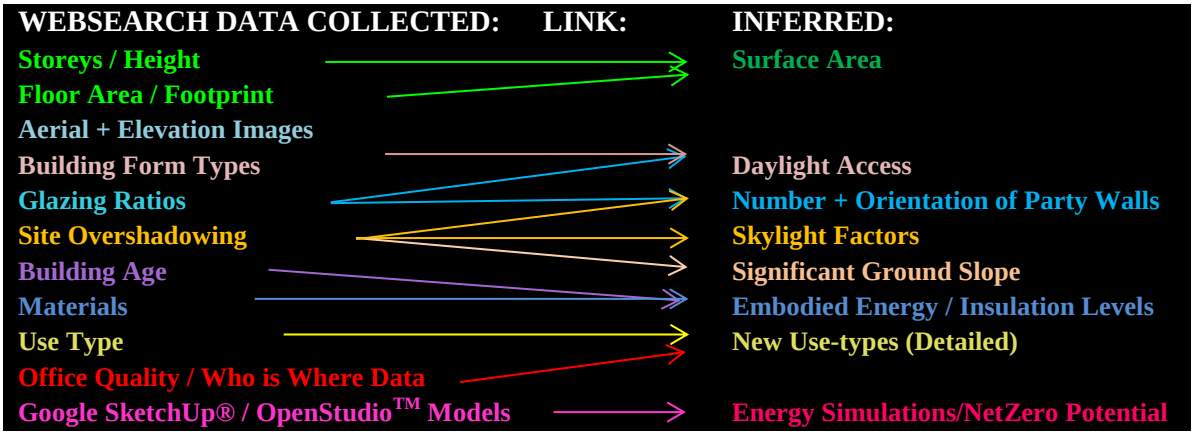


Figure 20 Websearch Data Collected + Inferred

The added information from the 'Who is Where?' business search was combined with 'detailed category' and 'improvement description' information in the valuation data. This formed a "New use type" which separated out F-"Food" R-"Retail" O-"Office" I-"Industrial", H-"Hotel", X-"Special" and lastly D-"Contains Dwelling". This included concatenated combinations of these codes to

various resolutions of detail from mixed versus single-use through to complex combinations numbering up to four 'New use type' codes – (Section 3.13).

Further research was carried out by the BEES team at a later stage regarding the assumption of insulation levels from building age and materials data (Cresswell-Wells, Donn, & Cory, 2012), embodied energy coefficients applied to the available Websearch data on materials (Berg, 2013) and net zero potential: the level of effort required to make each building net zero energy (Cory, 2015). Each of these research outputs have the potential for communication in a visualisation of the building stock, its characteristics and resource use (Section 5.05 and **Appendix i** 'Websearch. Data Sources')

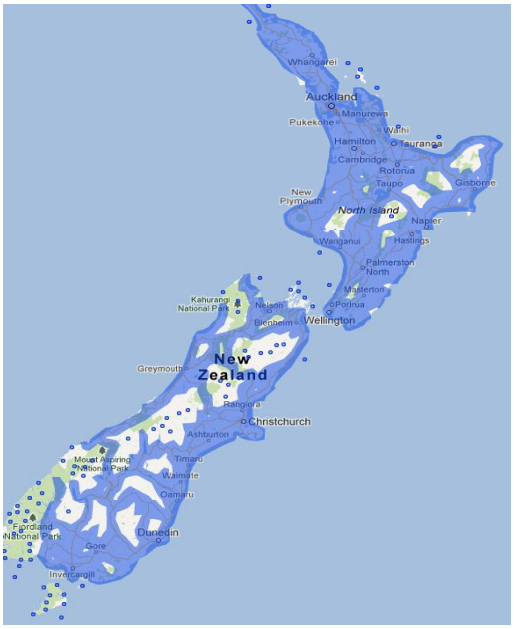


Figure 21 Street View Coverage in NZ

3.11: Data Cleaning Error Issues

As well as dealing with a proportion of buildings which did not have StreetView (Elevation) or Google Earth (aerial photo) coverage (Figures 21-22) there were inaccuracies found in the calculation of floor areas caused by the crudeness of the measuring tool in Google Earth. Graphs illustrating the comparison between original valuation floor areas and Websearch 'measured' floor areas can be found in Section 3.14 and this investigation helped to identify outliers which could be checked and corrected to improve the base data quality as much as possible.

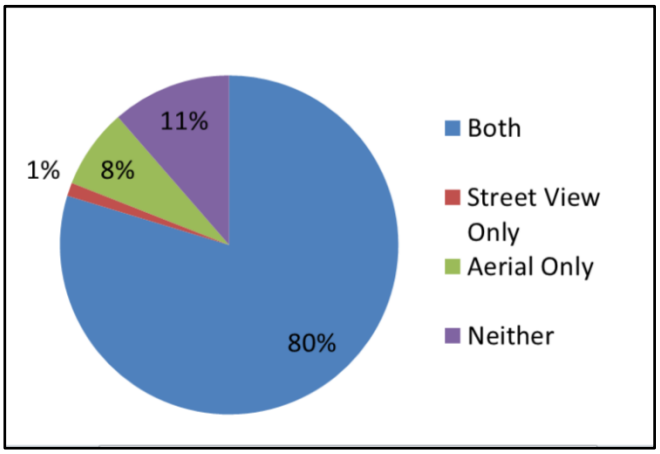


Figure 22 Websearch Coverage: Aerial + StreetView

Glazing percentages were found to be inconsistently entered, due to interpretation of an elevation and its glazed area as well as with the format of the data entry itself. Records with glazing ratios over 90% on any elevation were individually checked for accuracy by the author. The glazing results were initially rounded to the nearest 10% so that the results were as consistent as possible and could be eventually grouped into zero, low, medium and high glazing ratios and analysed against energy results. These wider groupings reduced the number of categories in the data, allowing energy bills to be compared against glazing ratio with larger sample sizes resulting in lower error rates and higher confidence in the resulting benchmarks (Section 3.16).

Thorough analysis of the differences between the original valuation building materials and the Websearch materials was conducted using conditionally formatted frequency tables. The tables clearly illustrated that there had been some misinterpretation between concrete and fibre cement wall materials due to the differences in building material mix in the dataset. Without carrying out further surveys it would be difficult to establish the level of error involved, but whilst this may present a problem for a visualisation communicating specific data on building typology the resource use

statistics are not affected by building material errors since it was not used to set resource use benchmarks.

Research assistant lethargy was also a problem with big complex buildings which created an initial bias in missing records in this typology which required a great deal of work to rectify. Changes in StreetView availability and aerial picture quality meant that returning to a record with missing data in the later stages in the project, returned full data which was not available initially. Over 150 records with large / complex sites or lack of web coverage (when first attempted) were added to the Websearch by the author to overcome these biases.

Researcher's initials were noted on all records attempted, which allowed any common error uncovered for one assistant to be filtered in Excel for fast correction. When multiple errors on glazing ratio and façade orientation in file names were attributed to a single research assistant it was possible to use the initials as a data filter to investigate quickly and correct the data accordingly.

Correcting the errors found at this stage allowed energy bill data to be compared against building typologies across the most complete dataset possible. In the example of the additional records completed by the author for large / complex buildings, this resulted in a far more accurate average energy benchmark for the typology than would have been achieved without the initial corrections, partly due to the increased sample size available for analysis and partly thanks to improvements made to the quality of high level data available for building typology analysis. This is just one positive example that justified the focus on data quality during this thesis, however many improvements made to the dataset (at the expense of considerable research time) will have had no impact whatsoever on the average energy benchmarks eventually used in the application of colours to the visualisation. **Appendix iii: '**Criteria Benchmark Averages' lists building typologies with the highest quality underlying open data and the lowest error rates. These represent the most important criteria requiring quality checking, however this list could not be established until the BEES team released the confidential Energy Bills data towards the end of the project.

3.12: Data Coding

The verbose 'notes' column was interpreted and simplified to enable further analysis. The notes described the reason for any 'error' or assumptions made affecting the data entry. These notes were coded into common error types using the text search filters in Excel. The notes picked up vacant buildings, vacant sites, address errors, Google map link issues, qv data errors, high surface area and podium buildings A list of the codes used are included in Figure 23.

Code	Errors	% of errors and #	
a	No Aerial View / Completely obscured /blurry / cloudy	4.5%	137
c	Car Park included / confusing old area figure	1.4%	43
d	Distant or Poor Quality Street View Access	5.5%	167
e	Data Entry Error / Mistype	1.0%	30
f	Storage Units Cool store or Greenhouse Noted	1.9%	59
q	QV error	6.0%	183
m	Websearch Data Entirely Missing	1.1%	32
o	Map reference / Street View Google Error	1.8%	54
p	Podium (Larger Ground / Lower Level Floor Area)	1.0%	29
r	Confusion over Building (37) / or St. # Range (421)	15.0%	457
s	Set Back upper floor or high surface area changes Area figures	2.3%	70
t	Outlier (QV data compared with Websearch) Websearch Correct	0.4%	13
u	New Area / Storey Height Assumed (estimate or StreetView only)	1.2%	36
v	Vacant Site / Under Construction / Empty Building	1.4%	43

Figure 23 'Error' Codes – From 1,061 Websearch Records (35%)

Typology mixtures were also coded as combinations of big > 1,000m^2 to small footprints <900m^2 and medium 900m^2>1,000m^2 with a variety of storey height ranges (small <1.5 storeys, medium 1.5>4.0 storeys and tall >4.0). Additional code columns were added to the data indicating small office, tall office, large-short mixed and retail. Materials were grouped into heavy-weight / light-weight and high / low cost codes and additional information such as climate zone (Figure 24 (Cory, 2015)), urban/rural mix, volume and decade were collated from the initial results. This coding work helped to clarify the correlational relationships between energy intensity and building typology mixes, establishing predictive relationships. Typology mixes present in the census and sample datasets, are illustrated in Figures 25-32. The energy average benchmarks applied to these groupings are listed in Figure 38 and **Appendix iii**: 'Criteria Benchmark Averages'.

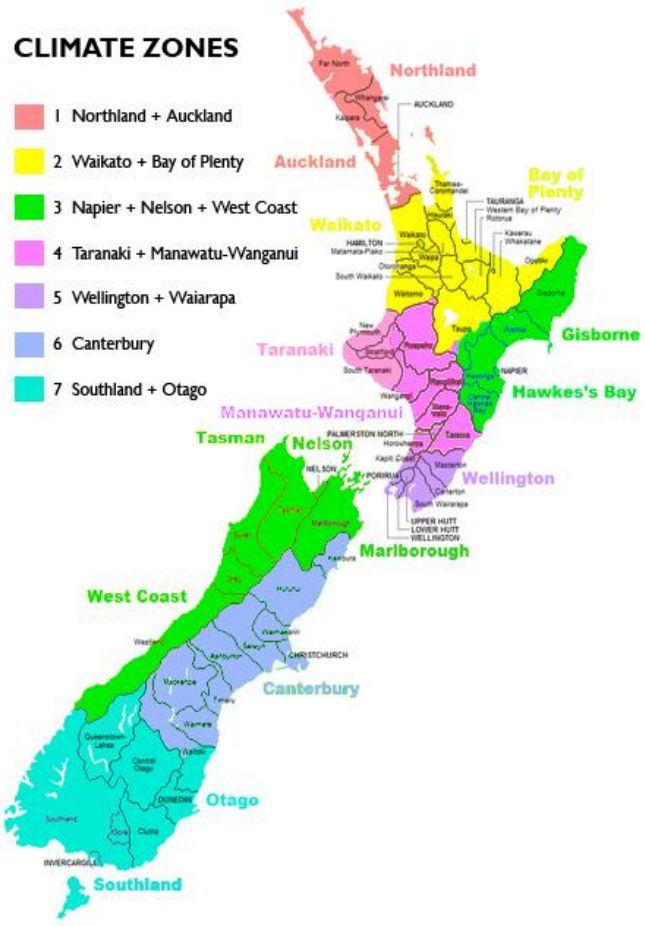


Figure 24 BEES Climate Zones - New Zealand

Frequency tables were produced with conditional formatting (colour scale applied to the numerical Excel data) showing the distributions of typologies present amongst the various samples by building count and total floor area. This highlighted outliers requiring further checking and correction (for instance small floor area buildings with a disproportionate number of storeys) or buildings with materials conflicts between the valuation data and the Websearch data. In these cases the buildings were

able to be isolated in the database and checked and corrected where necessary. Glazing ratios were simplified and material combinations were coded (heavy weight / lightweight and high cost / low cost) for energy benchmark analysis. This retained a reasonable energy bills sample (>25) when typologies were investigated for average energy benchmarks. Individual material categories would have limited value as a benchmark criteria due to small sample sizes (high error rates). Assembling small typologies into larger groups proved much more valuable for use in the Wellington CBD pilot visualisation.

The extended Websearch sample included all additional buildin*g*s identified and recorded at the street address in the valuation (3,468 buildings named 'A' through to 'K' derived from the initial 3,042 records). This dataset was compared with Websearch sample (3,042 building 'A''s only) in conditionally formatted (statistically coloured) frequency tables. Changes in typology mix were investigated for any effects perhaps due to their position in relation to the main road. Many of these 'additional buildings' were discovered in business parks or back-land sites without Google StreetView access, so much of the data was missing in additional buildings. Whilst not an exact science, the conditional formatting colours were examined for patterns, and the similarity of these between the Websearch building 'A's and the additional buildings suggested that the typology mix by floor area was consistent with the main changes observed in increased mixed industrial uses (+15%) and mixed food uses (7%), with equivalent reduction in single-use retail and office in terms of total building stock floor area.

In the case of the Wellington CBD census compared with the New Zealand census, a significantly different mix of typologies was found, as expected from such a purely urban central city area when compared with the entire country (Section 3.13).

3.13: Quality Checks for Bias Using Typology Mix

In order to ensure the visualisation and BEES research was based on the highest quality base data possible, the analysis of the varied mix of typologies present in the sample sets provided clues as to the level of bias caused by refusal / participation rates, by chance or due to selection method employed in the random samples. Six samples (Websearch, Phone Survey, Energy Bills, Monitored Buildings, Water Bills and Simulation Models (Cory, 2015)) and two census datasets (New Zealand, Wellington CBD) were produced by the BEES team, based on the original census valuation data. Each of the datasets were checked for bias by comparing the mix of typologies expressed in % of total floor area, enabling a clearer understanding of the make-up of the commercial building stock and identifying any differences that should be taken into consideration by the BEES team, for later statistical analysis work (Figures 25-32) .

Since the Energy Bills survey data was the only data available for analysis during the course of this thesis, the bias was not able to be compared and cross-checked across between the remaining sample sets as originally intended.

Both the Websearch and the simulation model samples (a selection of BEES building performance simulation models (Cory, 2015)) were stratified to provide robust statistics on building resource use with the minimum number of buildings drawn from each size strata to keep project costs down. This meant that strata 1 building records (less than 650 m^2 floor area) represented just 20% of the total building stock by total floor area, but 67% of the records by building count (Section 3.04). Direct analysis of the building record count within certain typology mixes, when compared to the census, was challenging. To be expressed accurately, conversion rates would need to be applied making

allowance for the imbalance of building count within each size strata. The mix could only really be understood in terms of total building stock floor area and not the chosen unit of study – the 'building'.

The other four samples used in the BEES survey were subject to refusal rates and had potential for bias (as well as the difficulties previously described regarding understanding building count statistics). These samples included:

- **phone survey** (collecting occupancy, ownership and floor area information),
- **monitoring** (where building energy was measured over a two week period),
- **energy bills survey** (including gas and electricity bills supplied by the building owner – and the only data source available for cross-checking during the course of this thesis),
- and the **water bills survey** (a small sample involving purely urban zones due to the high prevalence of water meters in these areas and a lack of available resources within the BEES research budget).

Lastly the data underlying the pilot visualisation was made up from a census of commercial buildings within the Wellington CBD including the 'suburbs' listed in the valuation as 'Pipitea', 'Wellington City' and 'Te Aro'. Understanding the differences in mix, between the national census, the energy bills sample and the Wellington CBD census was thought to be of benefit for the BEES statistical analysis as well as for the eventual realisation of this visualisation nationally. When eventually integrated with energy bills results from the BEES survey, the Wellington CBD buildings indicated a higher overall average energy benchmark than the overall average energy for the national population. This higher overall average energy result was used to check against the average of all energy benchmarks as applied to the Wellington CBD census. The difference between the typology mix in the two data sets indicated building types with a higher-than average energy intensity. This knowledge helped to define hybrid typology groupings to test against benchmarks to find energy intensity extremes. The standard error involved with the creation of benchmark averages could therefore be established for each typology and if desired, communicated in the visualisation as error bars.

Figures 25-32 on the following pages illustrate building record typology mix differences between each of the datasets listed below. These cover several criteria, most of which were used in the eventual assignment of energy benchmarks to the census data: building size stratum, use-type, use-mix, site density with valuation use type, building age, climate as well as complex size/height/use typologies are covered. Graph vertical-axes in the following figures are kept consistently scaled between pages to communicate the statistics to the reader as clearly as possible (Section 2.03.03).

- Commercial Census (50,548 Buildings):
- Phone Survey (445 Buildings /980 Premises)
- Energy Bills (251 Buildings /980 Premises)
- Water Bills (Bint, 2012) (34 Buildings)
- Websearch data (3,042 Buildings):
- Monitored (100 Buildings)
- Wellington CBD Census (786 Buildings)
- Simulation Models (Cory, 2015) (48 Buildings)

It was originally hoped that the focus on bias between samples and the analysis of the varied typology mixes would improve the accuracy of BEES energy statistics. These were expected to be directly provided by BEES statisticians for representation in the pilot visualisation. Instead these were created from first principles using typology energy intensity averages, for the purposes of the pilot visualisation.

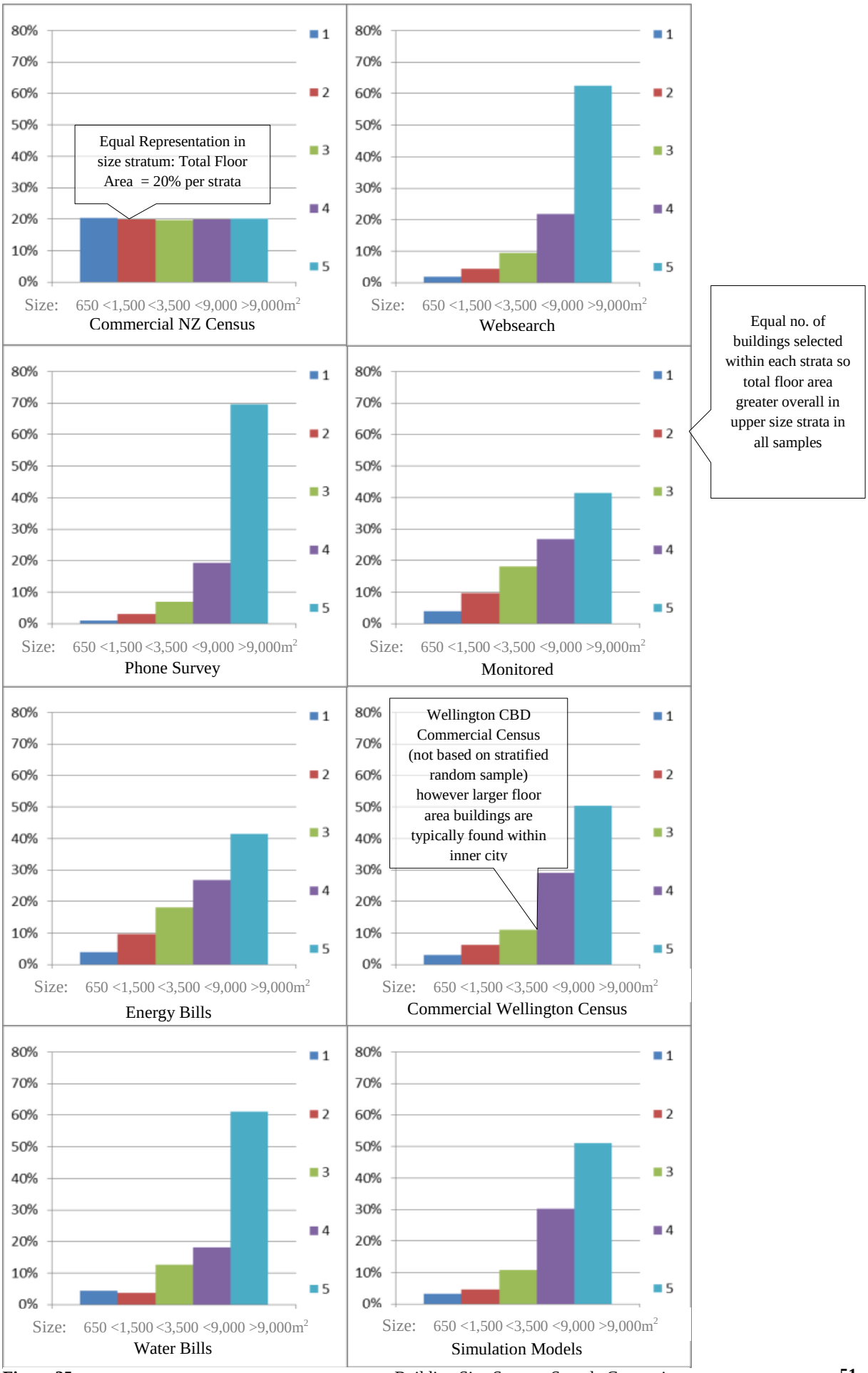


Figure 25 Building Size Stratum Sample Comparison

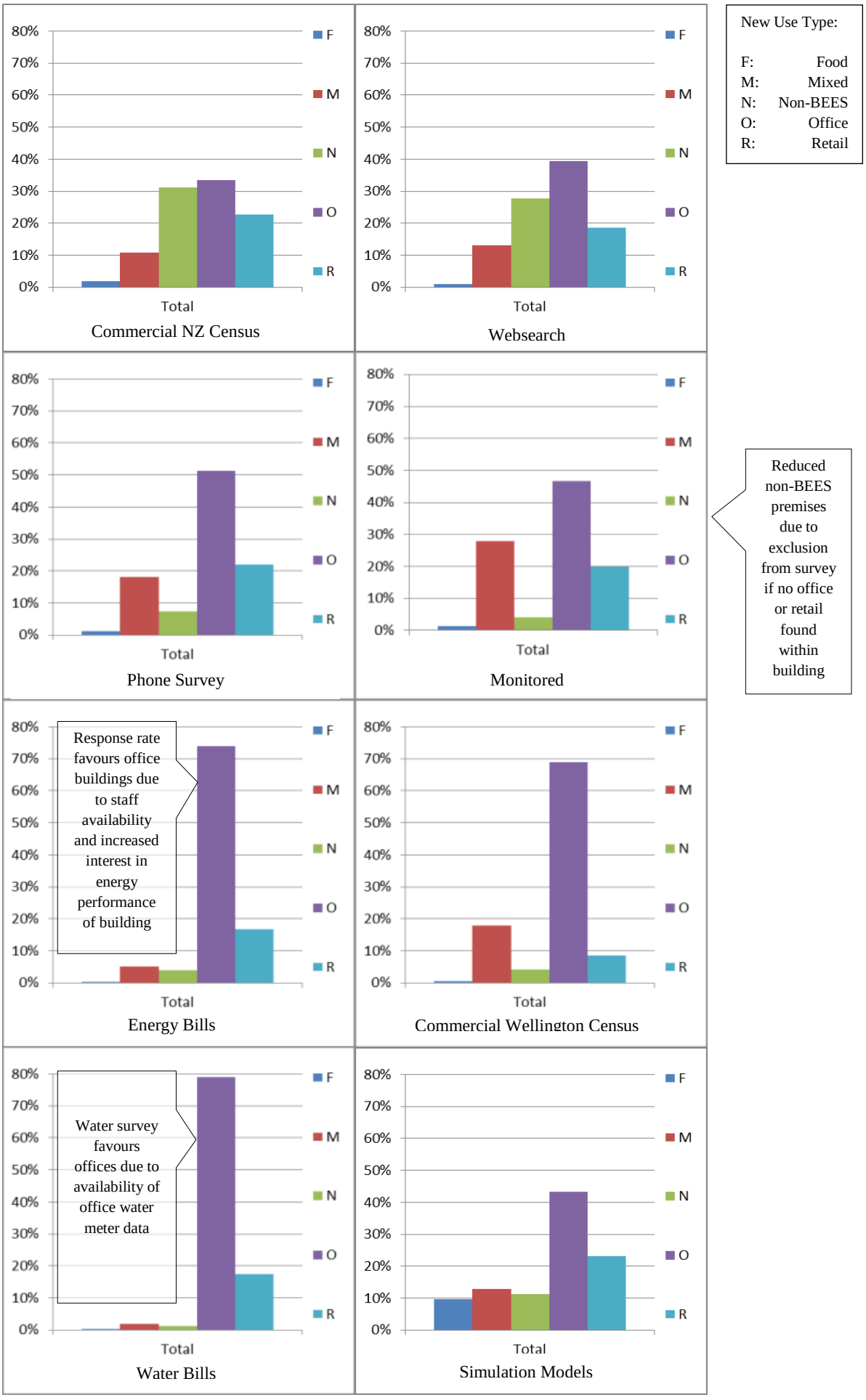


Figure 26 Building New Use Type Sample Comparison

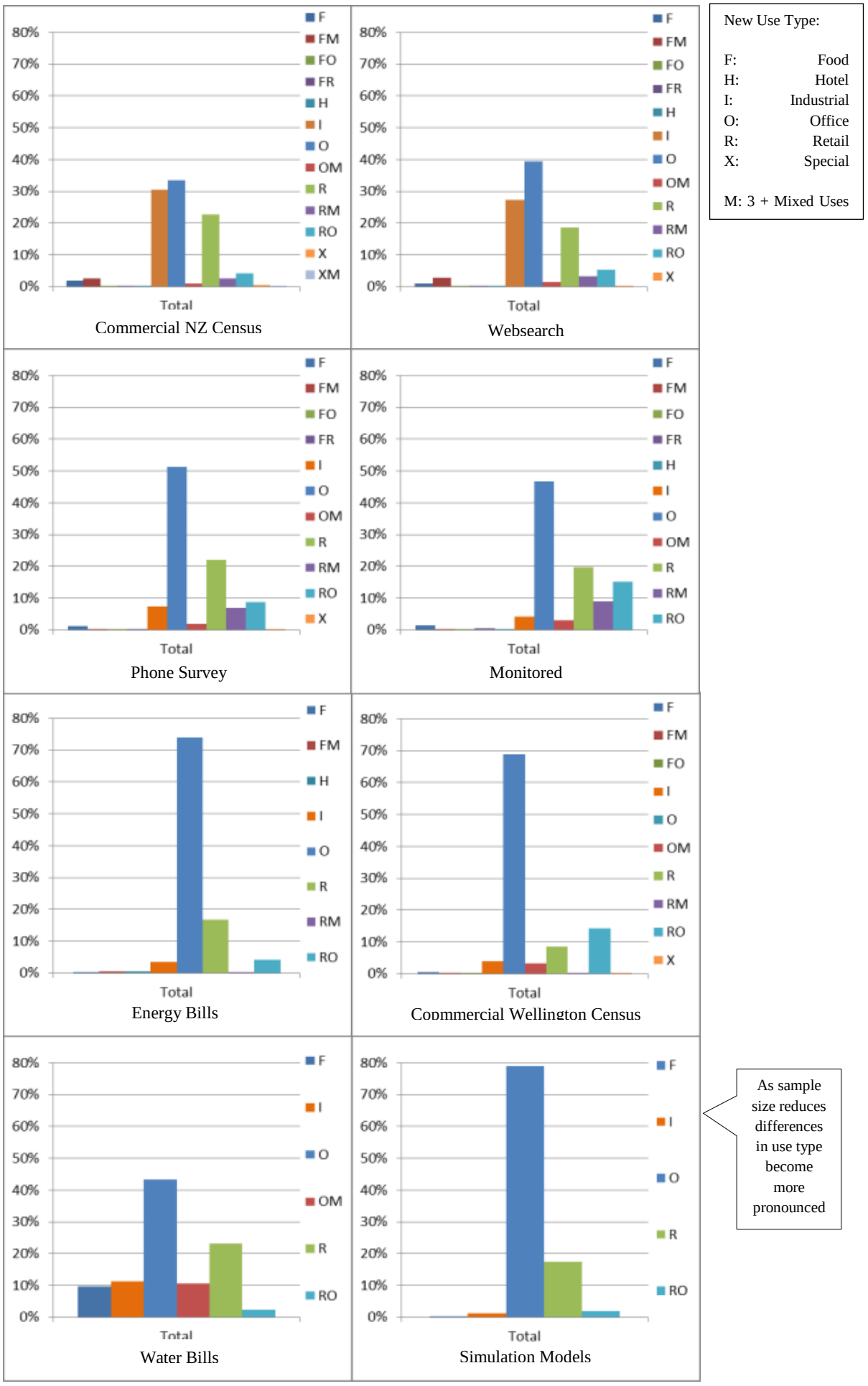


Figure 27 Building Use –mix Sample Comparison

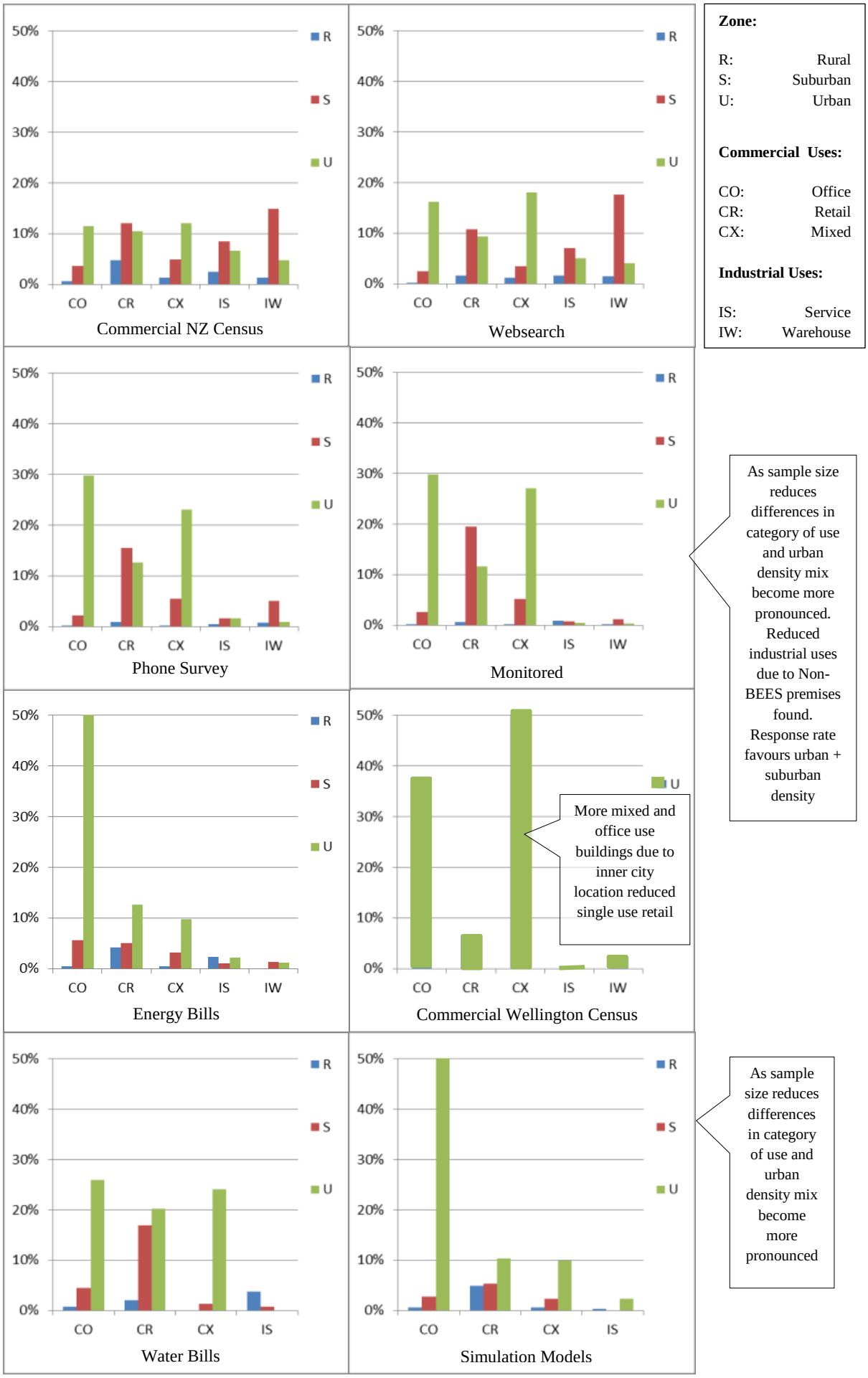


Figure 28 Building Category / Site Density Sample Comparison

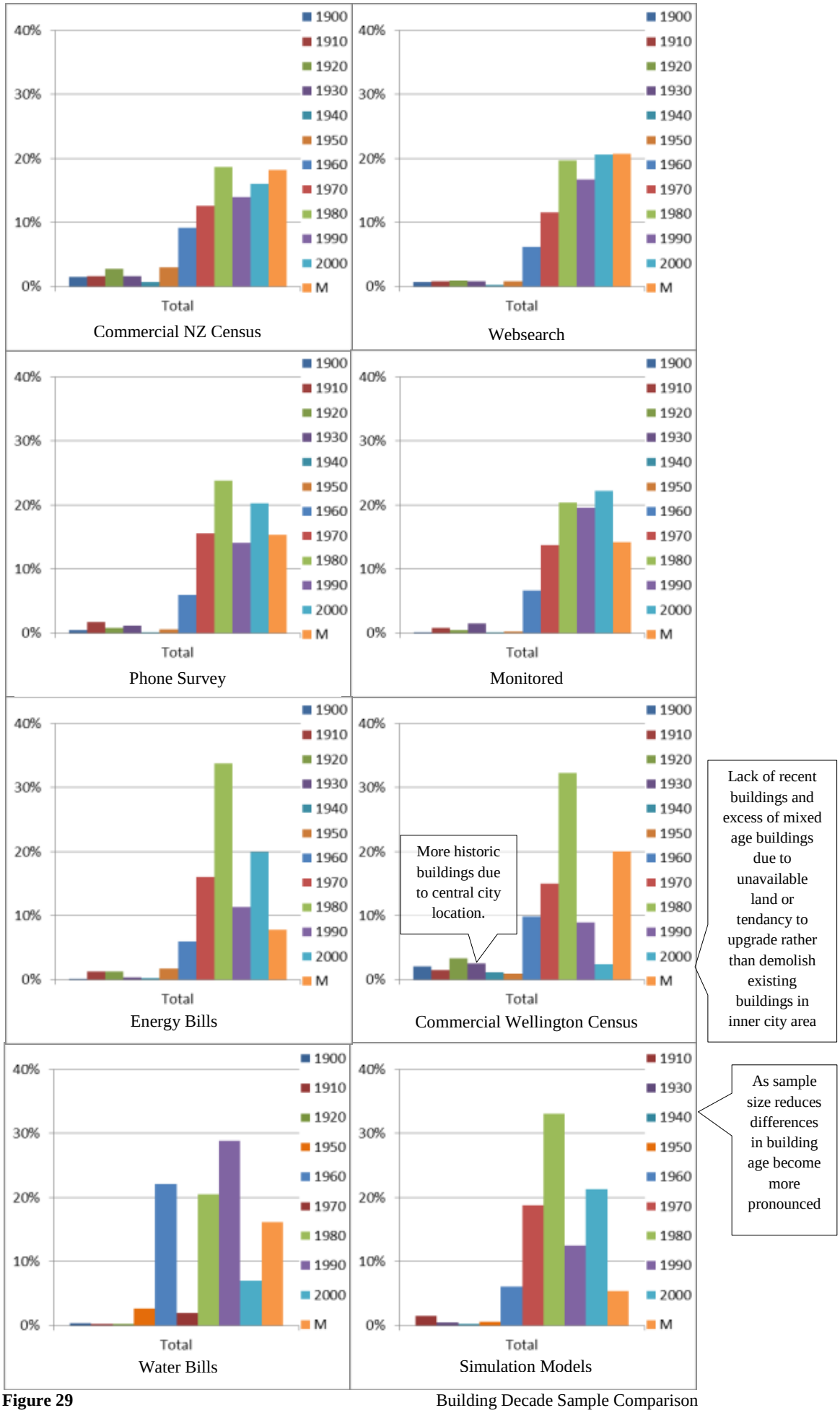


Figure 29 Building Decade Sample Comparison

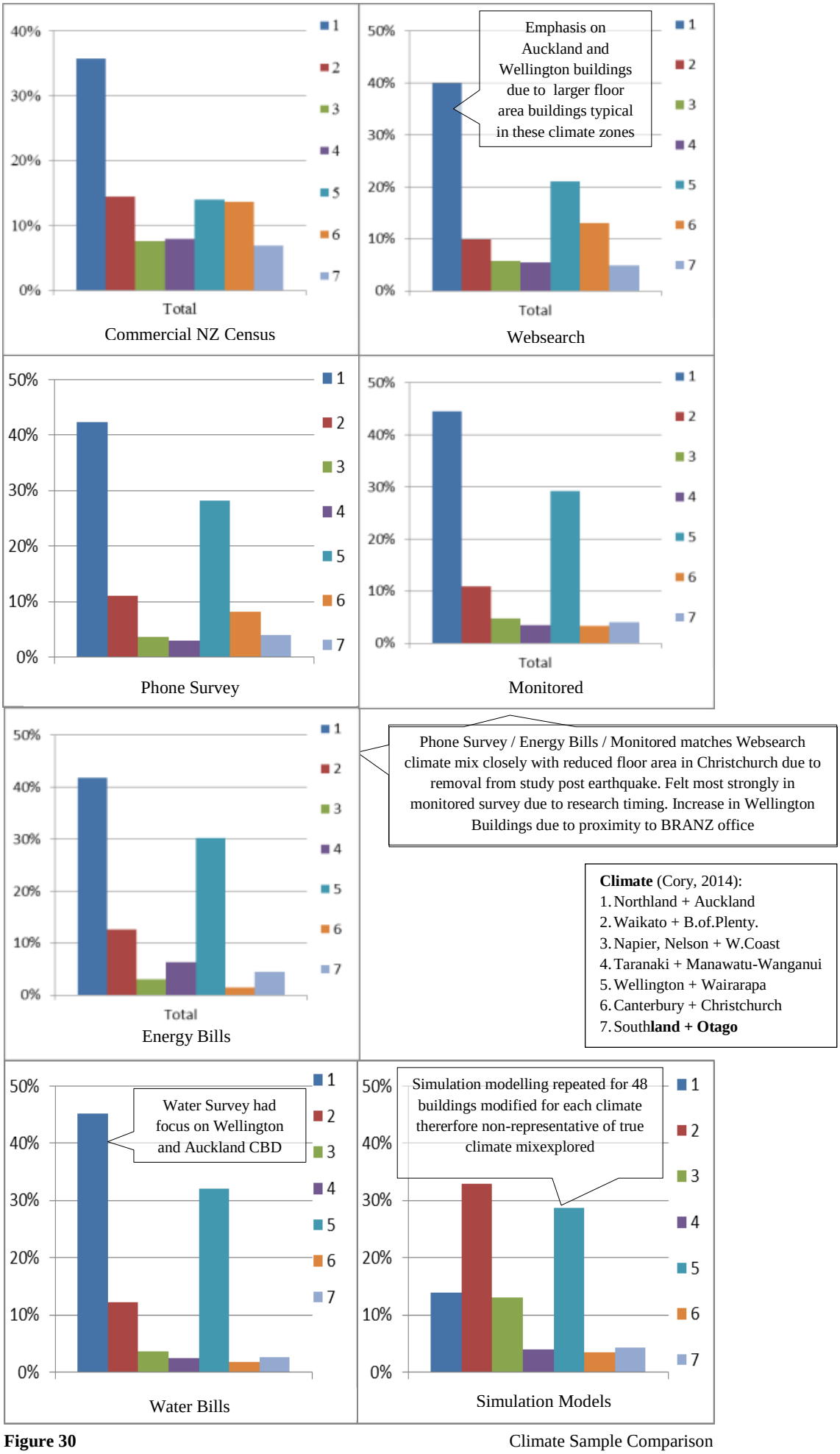


Figure 30 Climate Sample Comparison

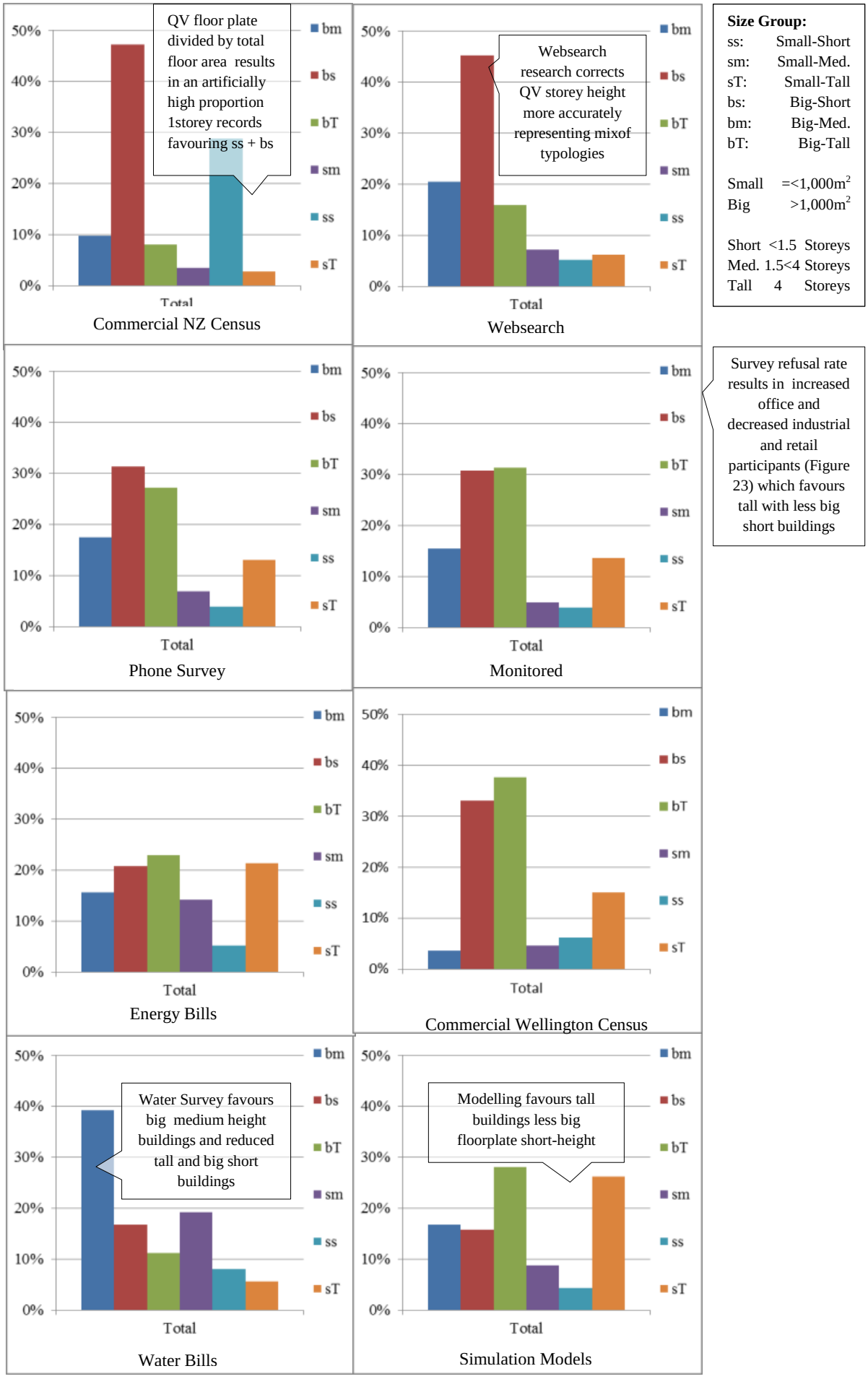


Figure 31 Building Size / Height Sample Comparison

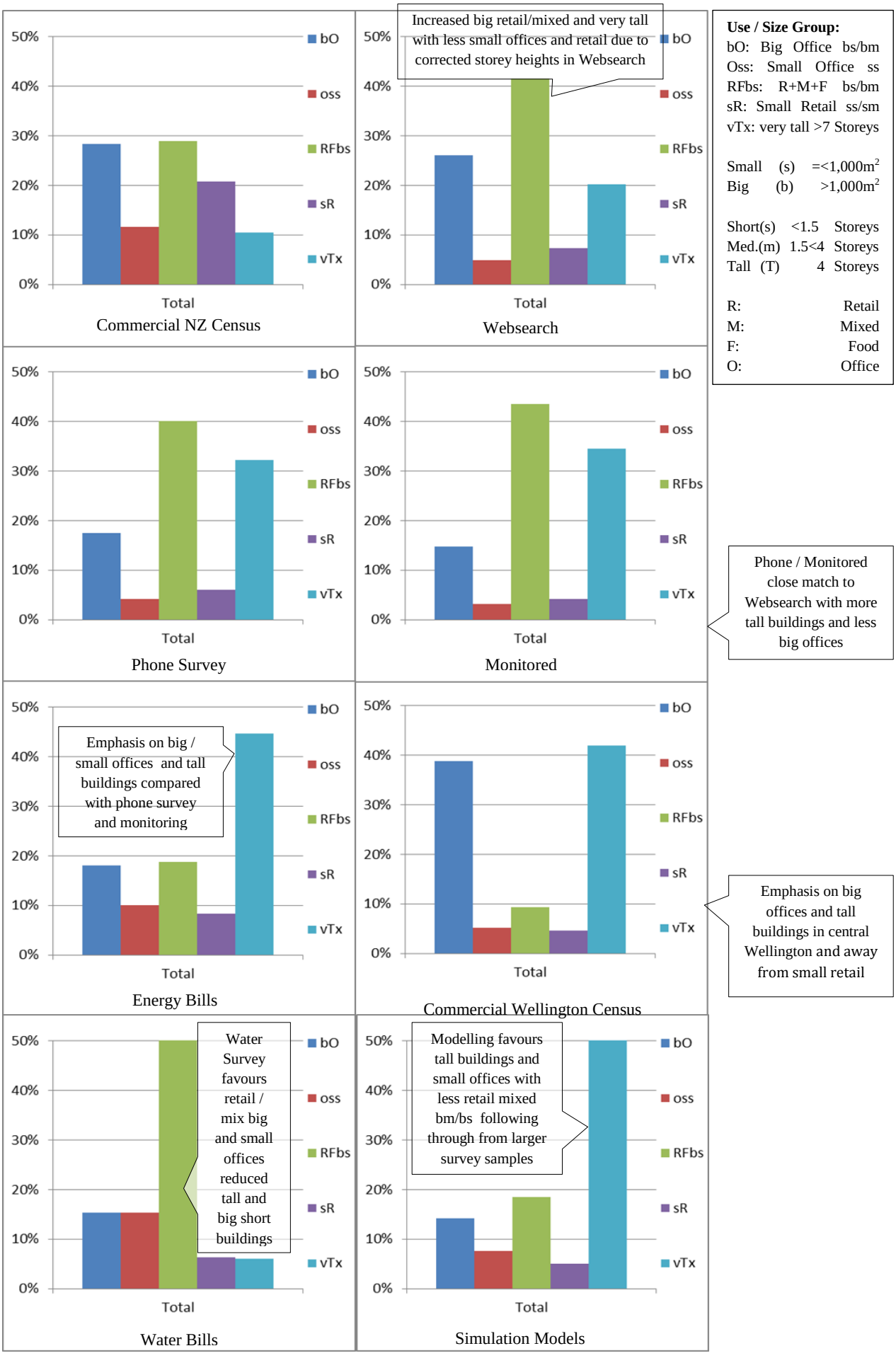


Figure 32 Building Use / Size Sample Comparison

There are subtle variations in the typology mix across the six samples when compared with the national census. Firstly, there was an emphasis on Wellington floor area in the monitored sample due to its proximity to the BRANZ office, which eliminated travel costs. This provided a larger number of Wellington buildings with bills data for comparison in the assignment of benchmarks and may have been beneficial for the pilot visualisation work, although the artificial focus meant these buildings were not randomly selected and this may have skewed the overall energy intensity average for the area. Secondly, survey uptake was more likely from floor areas related to offices with dedicated management / administration staff as opposed to retail + industrial premises, and evidence of this is illustrated in the typology mixes. Lastly the QV use types differ significantly to the Websearch 'new use types' provided by the Who is Where? business search in combination with the valuation detailed category information. QV categories of use were fraught with error (Section 3.09) as they were only assigned when the building was initially entered into the valuation database and never updated again. 'New use types' were based on a wider array of information sources, therefore they were assumed to be of a higher quality than the categories used in the initial Websearch sample selection strata.

The graphs in the previous illustrations indicate that the central Wellington commercial building census mix diverges considerably from the New Zealand commercial building census, most likely due to its location being a purely urban area. There was a larger proportion of historic and mixed age buildings, a disproportionate number of offices and retail-office buildings and a lack of industrial and single-use retail buildings. The detailed causes of these differences may be accounted for by high land value, a high number of large / mixed use buildings and the high density of commercial buildings in the inner city area generally. As a way of cross-checking the data, the confidential energy bill results available in the Wellington CBD area were averaged and found to be much higher (293 $kWh/m^2/yr$ for 38 buildings in Wellington) compared with the national average (204 $kWh/m^2/yr$ from 253 building records). This average was compared to the overall average of author-assigned benchmark values for the Wellington CBD census records and found to be within 18% of the average value (240 $kWh/m^2/yr$) provided by the confidential energy data. This proved a valuable cross-checking device, offering further reassurance that the Wellington buildings had a higher average energy intensity (as expected for an inner city area) than the norm for the national building stock. This could be attributed to the artificial BRANZ-focus on Wellington buildings, or genuinely because of higher EUI building uses being present within the centre of a city. By looking at the difference in typology mix between the census, energy bills and Wellington CBD building records, building types with more floor area in Wellington were thought to indicate those typologies with high energy intensity. Conversely, those typologies with lower floor area representation were assumed to have lower energy intensity typologies.

This high average for Wellington CBD buildings suggested that the extreme EUI building typologies (high: supermarkets, liquor stores, restaurants, take-aways and low: garden centres, activity centres and building supply shops - identified in the BEES case study work ((Bishop, 2012b) Section 3.17) may improve the -18% error accordingly. Where detailed building use information was available through valuation or Who is Where? the extreme EUI benchmarks could be additionally assigned to the Wellington CBD census for visualisation. Due to the likelihood that the extremely high EUI buildings would be present in the Wellington CBD area, including these case study EUIs would increase the overall average of Wellington CBD buildings closer to the level suggested by the cross-checking test (293 $kWh/m^2/yr$).

3.14: Floor Area / Storey Height Comparison : QV / Websearch Test

Floor areas (Figure 33), floor plate (Figure 34) and storey heights (Figure 35) from the Websearch were listed in order and compared to the values inferred or provided directly from the valuation (QV) data. These graphs identified outliers (>+/- 20%) which were then investigated individually by the author to confirm or correct as appropriate. The graphs show a clear relationship between the two values gathered using different methods with floor areas delivering the widest band of results. In 30% of the outliers investigated the disparity between floor area results appeared to be related to multiple buildings being found grouped within one valuation record. In 3% of outlier cases the floor area was out by a factor of 10, and the most likely explanation for this is that the valuation floor area was entered in square feet rather than square metres.

Websearch floor area estimates were related to the quality of Google aerial and Streetview imagery available for the property as well as the varying ability of research assistants to estimate the overall floor area accurately from the information provided. Buildings with high surface area such as podium buildings (1% building stock) , those with major setbacks (2%) and high surface area (8%) compound these floor area estimate errors further. This affects the visualisation output since the building volumes are automatically built by extruding polygon footprints to the top of the roof elevation height and therefore the volume of energy (and the building itself) is miscommunicated. This issue could be resolved by 'data-mining' open source cadastral data or by using emerging modelling methods (Section 1.05).

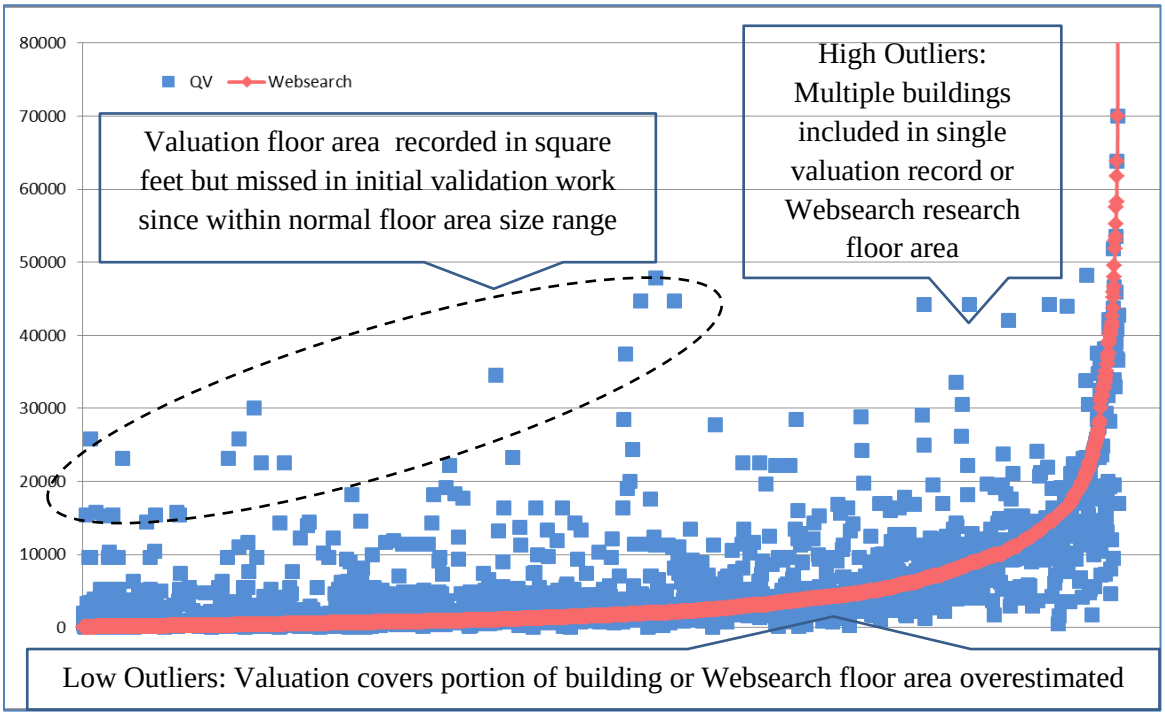


Figure 33 **Floor Area Comparison QV and Websearch**

The floor plate results (Figure 35) have less extreme outliers than the overall floor area results (Figure 33) suggesting floor plates provided by the QV data generally correlate with the measured floor plate in the Websearch data +/- 10%.

The reasons for any differences in floor plate result are accounted for by eaves overhangs, verandahs and roof canopies which would have not been measured in the QV data collection but would be hard to discern from aerial photos. The extreme outliers appear to be the result of data entry errors in the QV data (Figure 33 and 35).

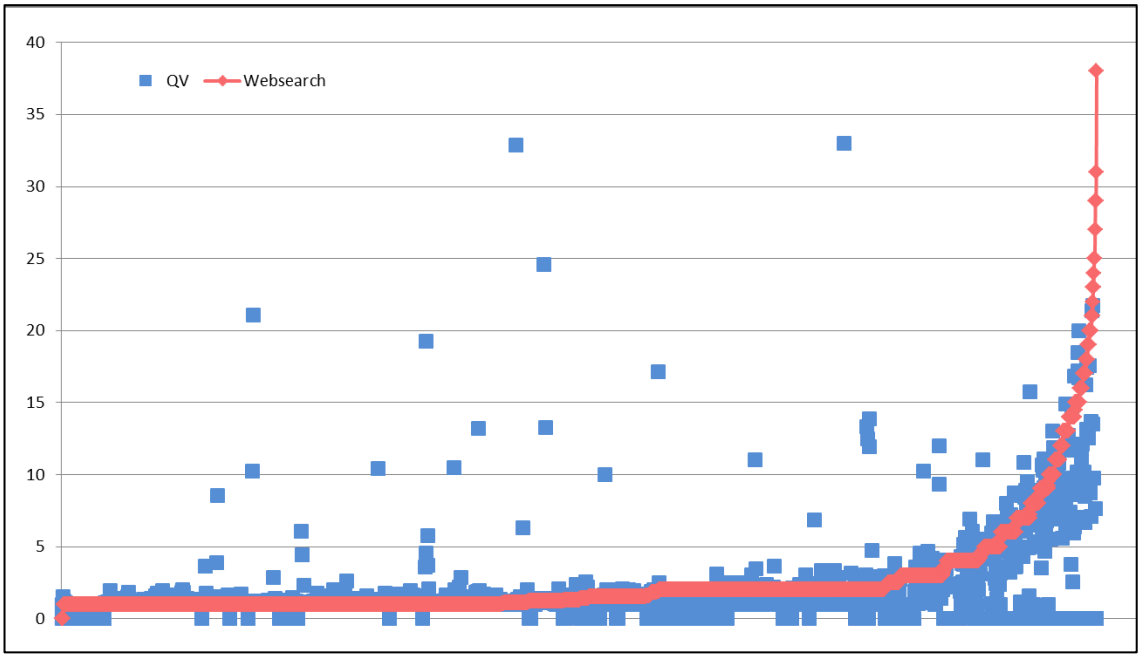


Figure 34 Storey Height Comparison - Websearch v's Calculated QV Stories

Storey heights (Figure 34) for the valuation data are calculated using the QV total floor area divided by the QV floor plate. The storey height results generally correlate very closely with the Websearch results and the QV results sitting on zero relate to missing QV floor plate data resulting from differing methods of measurement made between the various territorial authorities responsible for the data collection.

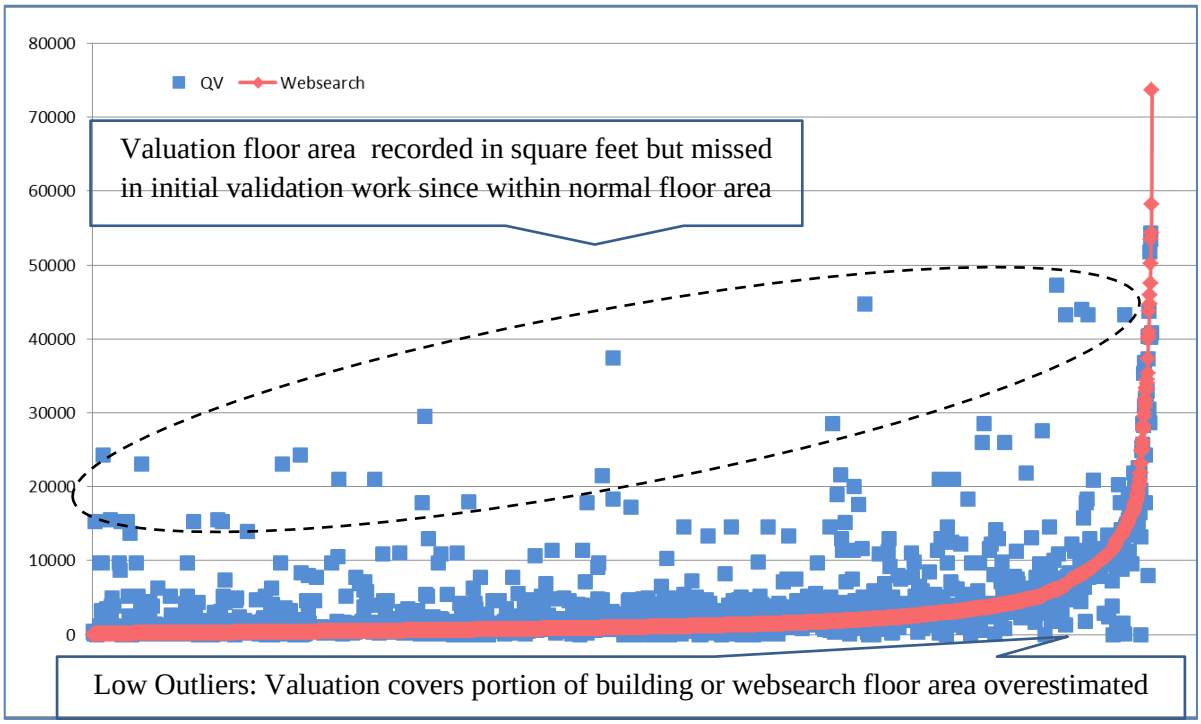


Figure 35 Floor Plate Comparison QV v's Websearch

The last graph (Figure 36) explores the relationship between the average skylight factor for a building compared to its Websearch number of storeys by ordering records in order of average overall skylight factor (blue line). As skylight factor increases, storey heights tend to increase indicating that taller buildings generally sit within sites with higher levels of overshadowing. This suggests that the site skylight factors might provide a good measure of site density. Site density could be linked to lighting end-use energy since the higher density sites create more overshadowing and are therefore likely to have higher benchmark lighting energy use than sites with low skylight factors overall. Further work using cadastral data and building elevation data would have the potential to improve the quality of the lighting end-use data as it relates to site density but this has not been carried out as part of this thesis, due to the need for specialist algorithms to extract the site overshadowing skylight factors automatically from parametric data.

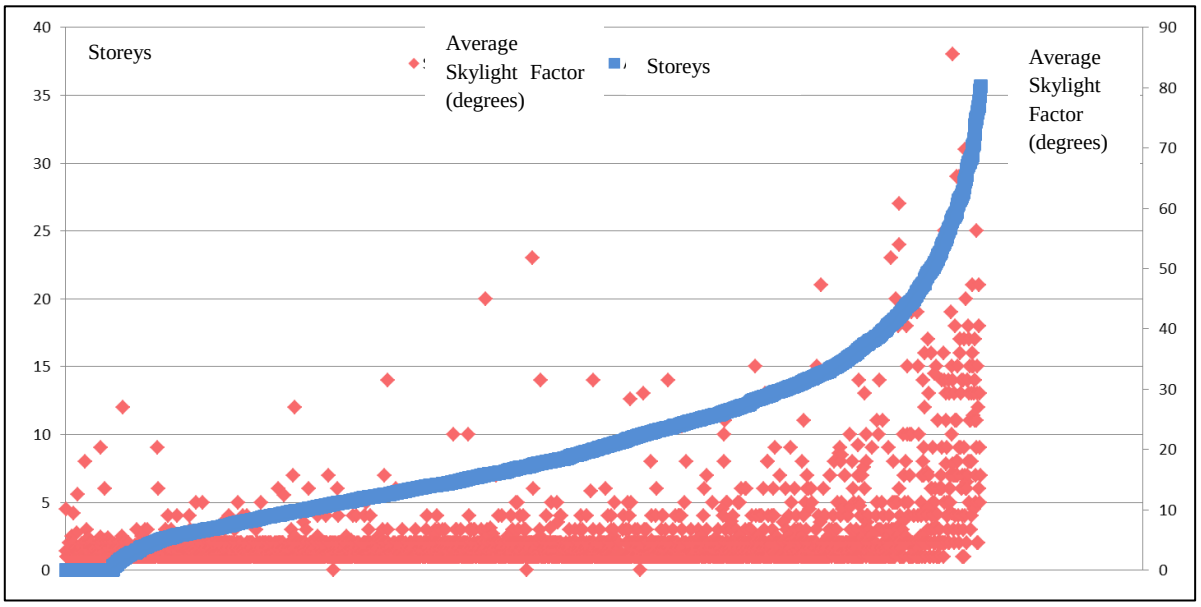


Figure 36 Comparison of Websearch Storey Height with Average Skylight Factor

3.15: Attaching BEES Energy Use Indexes to Websearch Records

Out of 980 premises returning energy bill results, 345 premises were deemed to have sufficient floor area and energy information to provide Energy Use Indices (EUI) in kWh/m^2/yr and these were provided to the author by the BEES team. These EUIs were matched to their corresponding building record in the Websearch (253 buildings). Some building records were associated with multiple surveyed premises and some of these also returned gas bill data EUIs (also expressed in kWh/m^2/yr for ease of comparative analysis). These multiple results for a single building were 'broadly' averaged with no account taken for the floor area associated with each premise or the mix of gas and electric included. This provided an EUI figure for each building within the energy bill sample set for analysis against common building typologies (assuming sample sizes remain above 26).

The number of premises per building with phone survey results versus those with energy results are indicated in Figure 37. Not all premises agreed to take part in the survey so it was impossible to calculate exact total building EUIs since the energy relationship between the premise and the whole building was not fully understood. The broad average between all known premise EUIs with certain characteristics in common, was thought to provide sufficient information to indicate a range of seven distinct energy use intensities. These spanned from very high (red) to very low (dark green)

The raw energy bill data was confidential so it was important for individual premise and whole building averaged energy results to remain anonymous. In order to be communicated to the public, the energy results needed to be applied aggregately with the aid of open data creating publicly available benchmarks. In addition, aggregate energy use data (and its distribution within larger groups of buildings) could be communicated in the visualisation. These larger groupings might include queried building typology groups, 'meshblocks' (the smallest geographic unit processed by Statistics New Zealand), 'area units' (aggregations of meshblocks) and lastly suburb-level data.

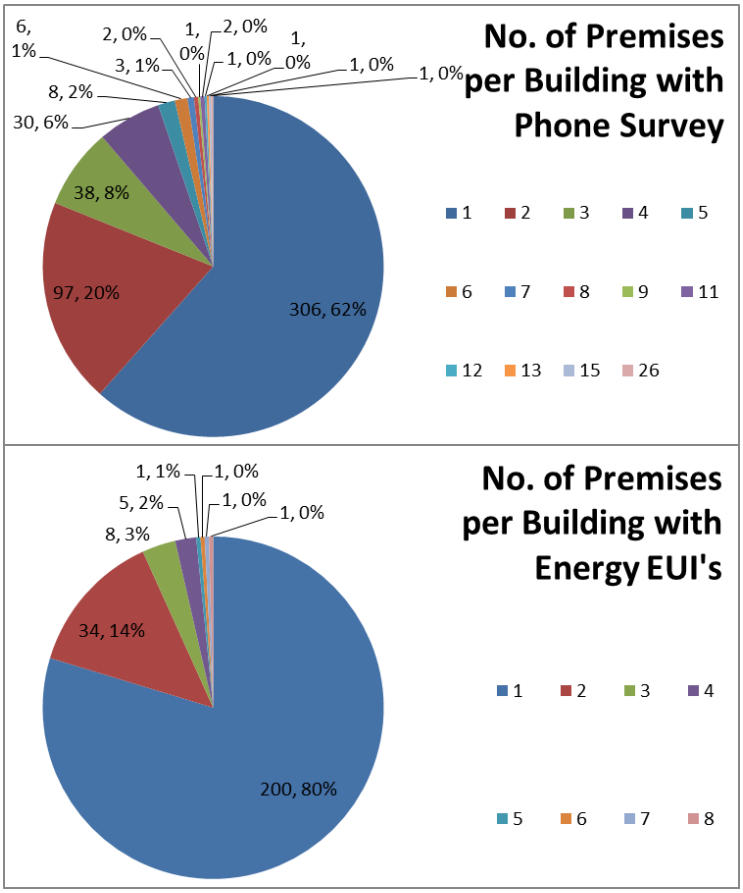


Figure 37 Premises Per Building Illustrating Refusal Rate

It would also theoretically be possible to include meshblock-level delivered energy data in the national visualisation as a comparison to the benchmark for a neighbourhood, but this would only be appropriate in meshblocks with more than a handful of buildings, otherwise sensitive energy consumption data would be revealed by default (Section 2.03.02).

3.16: Websearch Criteria (Typology) Related to their EUI Averages

In an attempt to establish the drivers of energy use from typology criteria collected during the Websearch work, every building characteristic or amalgamated hybrid typology (Section 3.13) was investigated for its average overall EUI (kWh/m^2/yr). **Appendix iii:** 'Criteria Benchmark Averages' list criteria and their corresponding energy average result in order from very high benchmarks through to very low benchmarks. Typologies with less than 20 energy bill results each were disregarded since sample sizes under 20 were not considered to yield a meaningful average result. Indeed many statistical rules of thumb, suggest a sample of 25-30 results is required (see Section 2.02.06). Limitations on the number of energy data points available meant this sub-sample size was difficult to achieve for many criteria available in both the Websearch and census database. A more thorough method of selecting the most reliable benchmark criteria averages involved calculating the standard deviation and standard error rate resulting in a +/-error % for each criteria benchmark average. In order to find typologies with extreme EUI averages the criteria with the highest and lowest averages and lowest standard error rate were combined with other high/low scoring criteria to form hybrid typologies. These hybrid use/size typologies were investigated further to find extreme energy averages with the lowest possible error rate, yet capable of illustrating the full range of benchmarks red (very high energy intensity) through to dark green (very low energy intensity) in the visualisation.

Two premise EUIs out of 253 (0.8% of sample with energy results attached) were omitted from the energy analysis due to an overall building energy average EUI over 2,000 kWh/m^2/yr. To put this

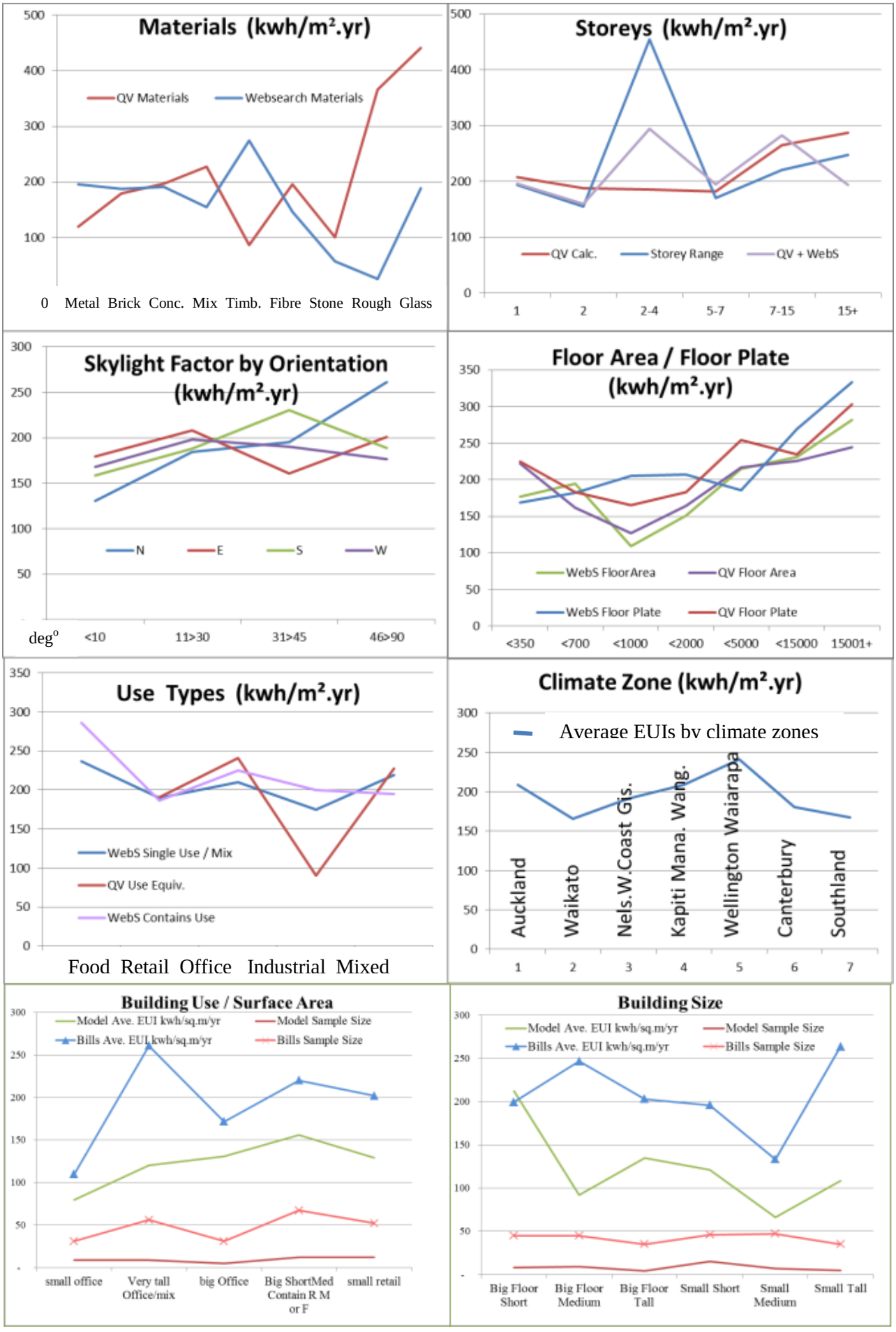


Figure 38 Websearch Typologies Compared to Average Energy Intensity (EUI's)

As floor area / floor plate area increases energy intensity generally rises, however a dip at 700<1,000m^2 present in three out of the four datasets suggested a typology group with a large sample size that might be combined with other low-EUI generating criteria to create the lowest possible EUI benchmark (displayed as dark green), yet retain a sample size between 26-67 and error rate less than +/- 31%. This balance between error rate and sample size was selected because it involved enough criteria related benchmarks to assign the seven colours to the entire census with minimum overlap between criteria groupings, yet remain within an acceptable level of error to be of use in differentiating a hierarchy of seven benchmark ranges.

The skylight factor by orientation graph in Figure 38 illustrates the relationship between skylight factor, orientation and energy intensity EUIs, and indicates the highest EUIs are found in buildings with high overshadowing on north facing facades and lowest in those with low overshadowing to the North. This suggests that energy use is affected by overshadowing and with further investigation, this could be taken into account when calculating lighting and end-use energy in future refinements of the visualisation.

Criteria benchmark averages using open data (available in the census, the 3D visualisation data or Who is Where? service) and with adequate sample sizes (26-67 energy bills results per criterion) and lowest possible error (< +/- 31%) were used in the application of seven hierarchical colours to the pilot visualisation. A full list of all criteria and their corresponding EUI averages can be found in **Appendix iii:** 'Criteria Benchmark Averages'. This reduced the selection of criteria benchmark averages that could be used to assign benchmarks in the pilot visualisation – these are listed in Figure 39 in order from very high (red) to very low (dark green):

		BEES SIMULATION MODELS						PREMISE ENERGY BILLS							
Field	Criterion – Websearch Field Investigated for EUI	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	Error % +/-	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	+/- error	Difference Models / Bills	Natural Breaks
ALL RECORDS	ALL	123	48	108	7	16	25%	204	253	214	16	13	13%	60%	
Wellington CBD	Te Aro, Piptea, CBD	79	4	43	2	22	53%	293	38	277	6	45	30%	27%	
QV Mats roof	conc	113	5	53	2	24	41%	268	29	228	5	42	31%	42%	4
Size Typology	Small Tall	109	5	59	2	26	48%	264	35	150	6	25	19%	41%	3
Use/ Surface Area	Very tall Office/mix	120	9	50	3	17	27%	261	56	191	7	26	19%	46%	13
Storey WebS.	10	94	5	31	2	14	29%	248	32	209	6	37	29%	38%	4
Floor Area QV	15000+	152	2	53	1	37	48%	244	34	204	6	35	28%	62%	3
Category Stratum	CR	167	17	152	4	37	43%	241	65	228	8	28	23%	69%	4
Contains Use:	F	88	4	70	2	35	78%	237	40	198	6	31	26%	37%	4
QV Mats	mix	151	12	153	3	44	57%	227	63	169	8	21	18%	66%	1
Secondary Mat	Present	108	8	89	3	31	57%	226	67	213	8	26	22%	48%	1
Mixed / Single use	Retail	112	10	114	3	36	63%	225	66	234	8	29	25%	50%	5
Use/ Surface Area	Bsh/med: R/M or F	156	12	177	3	51	64%	220	67	223	8	27	24%	71%	3
Floor Area QV	2000>5000	200	8	212	3	75	73%	217	58	246	8	32	29%	92%	7
Size Stratum	4	189	10	182	3	57	60%	210	53	175	7	24	22%	90%	8
Use/ Surface Area	small retail	129	12	80	3	23	35%	202	52	187	7	26	25%	64%	3
Size Typology	Big Floor Short	212	8	200	3	71	65%	199	45	200	7	30	29%	106%	3
Storey WebS.	4-8	115	9	67	3	22	38%	196	51	152	7	21	21%	59%	0
Size Typology	Small Short	121	15	75	4	19	31%	196	46	212	7	31	31%	62%	4
Sq Rt.Floor Plate	<=20m	118	12	81	3	23	39%	192	43	185	7	28	29%	61%	9
QV Floor Plate	350>700	120	9	61	3	20	33%	183	41	166	6	26	28%	66%	0
QV Floor Plate	1000>2000	112	11	67	3	20	35%	183	56	166	7	22	24%	61%	9
Size Stratum	3	98	10	75	3	24	48%	174	50	148	7	21	24%	56%	34
Volume	3000>5000	110	4	68	2	34	60%	140	42	107	6	16	23%	78%	7
Size Typology	Small Medium	66	7	54	3	20	60%	133	47	114	7	17	24%	50%	10
Volume	1500>3000	70	7	55	3	21	58%	123	26	89	5	18	28%	57%	9
Use/ Surface Area	small office	80	9	50	3	17	41%	110	31	96	6	17	31%	72%	-

Figure 39 Benchmark Averages (Bills) Versus Simulation Model Sample (Small sample sizes)

3.17: Attaching EECA and BEES Case-study End-use Percentages

The end-use breakdowns communicated in the pilot visualisation were sourced from the EECA End-Use Database selected from building related energy across all business sectors (Energy Efficiency Conservation Authority, 2007). The breakdown of end-uses indicated in the database generates a series of pie charts showing percentage mix of each end-use energy type by building use type. The percentage mixes for each end-use were applied to buildings in the Wellington CBD census according to 'New use type' fields which had been entered into the Websearch as well as the census during this course of this masters project. The new use types included: Food; Retail; Office; Hotel; Special; Industrial; and various combinations of these. This allowed the buildings in the visualisation to be viewed in Lighting, Space Conditioning and Miscellaneous Energy Benchmark modes, which when toggled with the Whole Building Energy to illustrate the different emphases of certain end-uses and their distribution in the city (Figure 48 and 50-52 Section 3.23). With eventual release of the 'measured' BEES end-use percentages, the end-use data in the national visualisation could be updated to the most recent information available, rather than utilising theoretical EECA end-use data from 2008.

Case study typical end-use breakdowns and EUI averages from the BEES study (Bishop, 2012b) build a more detailed picture of the breakdown of end-uses and also provide extreme whole-building energy intensities associated with certain building uses. The extremely low EUIs were identified in garden centres, activity centres, building supplies shops and warehouses. Extremely high EUIs were identified in supermarkets, restaurants, liquor stores and takeaways. The detailed building use information available in the valuation records and Who is Where? dataset identified specific use categories and business names such as those in the report by Bishop. In these cases the benchmark energy values applied to the census could be updated to reflect these extreme EUIs values. Due to the prevalence of high EUI uses in the Wellington CBD example (when compared to the low EUI uses) the inclusion of these extreme EUI values in the visualisation would increase the Wellington CBD overall energy benchmark average to a value more in line with the average suggested in the confidential Energy Bills data – thus reducing the difference previously uncovered in the cross-checking test (-18%) (Section 3.13) The EECA end-use percentages and BEES extremely high and low case study examples are illustrated in Figures 73 and 74 (Section 5.06).

3.18: Assigning Benchmark EUI Colour Ranges – Whole Building

For the purposes of colouring up the 3D building forms in the pilot visualisation, it was necessary to assign benchmark EUI averages to the Wellington CBD census. The range of benchmarks were listed in order from high to low EUIs (Section 3.16) and restricted to criteria available purely through open data sources (cadastral / valuation / business search). Criteria benchmark averages were used where sample size fell between 26-67 and the standard error was restricted to +/- 31%. Natural breaks (the differences between the criteria averages when placed in order from high to low (Section 02.03.04) were used

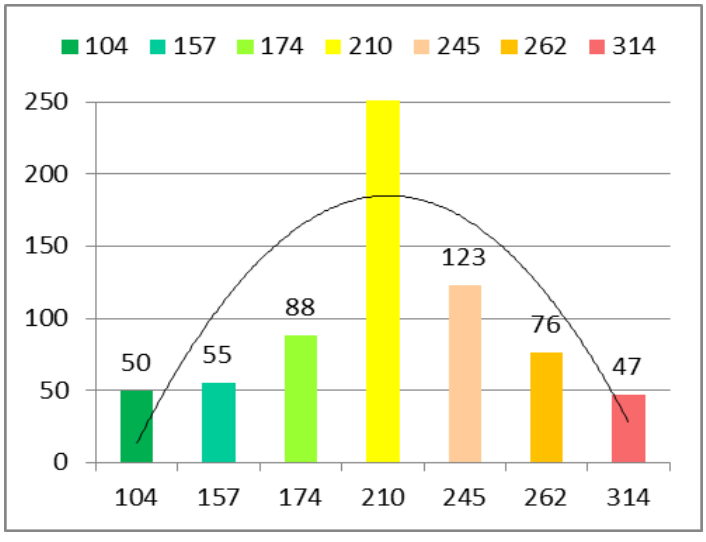


Figure 40 Benchmark Averages Applied to Wellington CBD Census

to set the boundaries of the colour ranges. The census records were coloured starting from the very highest (red), very lowest (dark green) and gradually working inwards (orange, turquoise, light orange, light green) towards the typologies associated with average benchmark EUI values (yellow). Any building records falling outside of these criteria benchmark types were left displayed as average (yellow). Rather than completely excluding typologies with a sample size over 67, these were broken down to produce hybrid use/size combinations as uncovered in the previous work (eg. small short office, small retail, large short mixed/retail/food, very tall). Energy averages and sample typology mix results are illustrated in Figures 31-2 and 38-9. The median benchmark buildings are heavily represented in the Wellington census (Figure 40) due to the order of colouring and the restriction of the error rate and sample size range. This could be combatted by increasing the acceptable error rate (currently <+/- 31%) and widening the sample size range (26-67), providing a longer list of criteria benchmark averages used to colour the Wellington census buildings. This results in fewer of the census records defaulting to a median benchmark, but without further expert statistical analysis there would be no way of knowing whether this would produce a more accurate visualisation. The average EUI *value* assigned to each colour was established from the average confidential energy bill result within the colour range in the Websearch CBD dataset. Yellow represented the average for the whole commercial building stock (210 kWh/m^2/yr) Figure 41.

3.19: Assigning Benchmark EUI Colour Ranges – Energy End-Uses and Water Use

Once each record in the Wellington CBD census (those records in the valuation with suburb recorded as Pipitea, Te Aro and Wellington City) had an allocated colour and its corresponding EUI average *value* allocated (Figure 41 – left column) it was possible to apply the pie-charts outlining the various EECA and BEES case study end-use breakdowns by percentage. These end-use percentages were attached to the Wellington census according to the 'New-use type' (Section 3.17). The average EUI figure assigned to each colour range was multiplied by the percentage breakdown given for lighting, space conditioning and miscellaneous energy use, producing three new columns of end-use statistics on each building. Unlike the whole building benchmark energy results, the building record 'New-use types' dictated the distribution of end-use percentages calculated. The colour range values were set for each end-use using the same process (natural breaks) as had been carried out with whole building energy but with different resulting EUI average values for the three end-use types communicated. These EUI end-use values do not necessarily add up to the corresponding whole building benchmark value due to the use of the 'Natural Breaks' method to assign colours, and the two different indicators used to assign the values (hybrid typology for whole building energy and 'New-use-type for end-uses). In retrospect adjusting the colour ranges so that these average results add up might have provided another cross-checking method to improve accuracy of the benchmarks.

Figure 41 Energy + End-Use Benchmarks: Averages per Colour Range

The exact process followed for applying whole building energy EUI averages could be used to produce the building water-use results from BEES. In this case, the average water result is 1.64 kl/m^2/year and the range spans from 12.8 times the average down to 0.06 times the average (Figure 42). The colour range key describing the water-use benchmarks would benefit from the use of a different colour range from energy benchmarks to provide clear differentiation. A colour range using a common denominator of dark green as low but using blue for high water use and red for high energy use would seem a practical choice to suit the topic being visualised. However, this theory would need to be tested further with focus groups and users and may prove a challenging colour range to be deciphered accurately (2:03:06).

Figure 42 Water Benchmarks

Further data from the Targeted Survey (BEES measured energy), the Aggregate Survey (BEES occupancy, ownership and use), embodied energy information (Berg, 2013) and net zero potential (Cory, 2015) (Section 3.10) could also be used to enhance the visualisation, using a similar methodology employed with the energy and end-use benchmarks.

3.20 Geo-coding Address information in Census

Now that the full set of Wellington Census buildings had benchmark statistics allocated for whole building energy and end-uses, it was necessary to geo-code this statistical data into Google Earth so that the colours could be presented in a 3D city visualisation format. The fields forming address, suburb and city were concatenated together to make a single comprehensive address field format for accurate geo-location. Map marker icon HTML codes (Figure 43) and statistical values could be applied to the energy and end-use figures according to the colour code allocated. The resulting database was uploaded to Fusion Tables (Google Apps, 2013a) and run through the 'geo-coding' menu option. Each coloured map-marker set was downloaded as separate KMZ files (Google Earth format) for whole building energy, lighting, space conditioning and miscellaneous end-uses. These files could be opened and viewed in Google Earth as a place layer, providing the distribution of colours across the Wellington CBD.

	red_blank		orange_blank		yel_blank		wht_blank
	ltblu_blank		grn_blank		blu_blank		

Figure 43 Map Marker Icon HTML Codes

According to the status report in Fusion Tables the addresses in the Wellington census were found to be 1% ambiguous. Ambiguous addresses result in map markers placed in the centre of the road nearest to the street number indicated address field, however they do not indicate a mailbox location as is the case of clear address matches. The address errors affecting 1% of the Wellington census closely match the address errors found in the Websearch results and the initial census error checks (Figure 19, Section 3.09). This further reinforces the accuracy of the data collected during the Websearch study.

3.21: Automatically Constructing the 3D City Model

With the map marker icons indicating energy and end-use colours distributed throughout central Wellington the next task was to create the basic building block models that would be coloured to

communicate the energy and end-use benchmarks. Using cadastral building footprint information along with building elevation height (LINZ, 2011c) (Corkery & Coup, 2007) downloaded as a KMZ file (Google Earth format) it is possible to automate the construction of an entire city model. The footprints are already set at correct elevation heights and can be extruded down to the ground with a single operation in Google Earth (Google Apps, 2013b). The operation involves right clicking on the KMZ file name in the 'Places' sidebar, then selecting 'properties', 'altitude', extend sides to ground.

Any operation carried out in Google Earth can be applied to multiple objects at once by grouping them in a new folder and by right clicking and selecting 'Properties' and 'sharing styles' in the 'style, colour' tab. For the purposes of achieving the Wellington census pilot visualisation the colouring was done manually by grouping the building footprints into folders according to their corresponding map-marker icon colour. Since completing a final visualisation of the entire country could not be coloured manually the next section covers instructions for a simple theoretical computer algorithm which could be used to fully-automate the operation. The building footprints were quickly located and grouped into folders for colouring using the Fusion Table map marker icons described in the previous sub-section.

In the manual method, all but the red map marker icons were temporarily switched off using the tick boxes in Google Earth 'Places' (Figure 44 - buildings not yet allocated a benchmark colour are shaded blue in this illustration). The nearest building footprint to each map marker was 'moved' into a sub-folder named 'red' and then once complete the folder was switched off to remove it from the remaining building forms. This task was repeated for each of the seven colours in the scale. These formed the basic block models for the statistical visualisation of whole building energy EUIs and the method was repeated in a separate KMZ file for each of the end-uses.

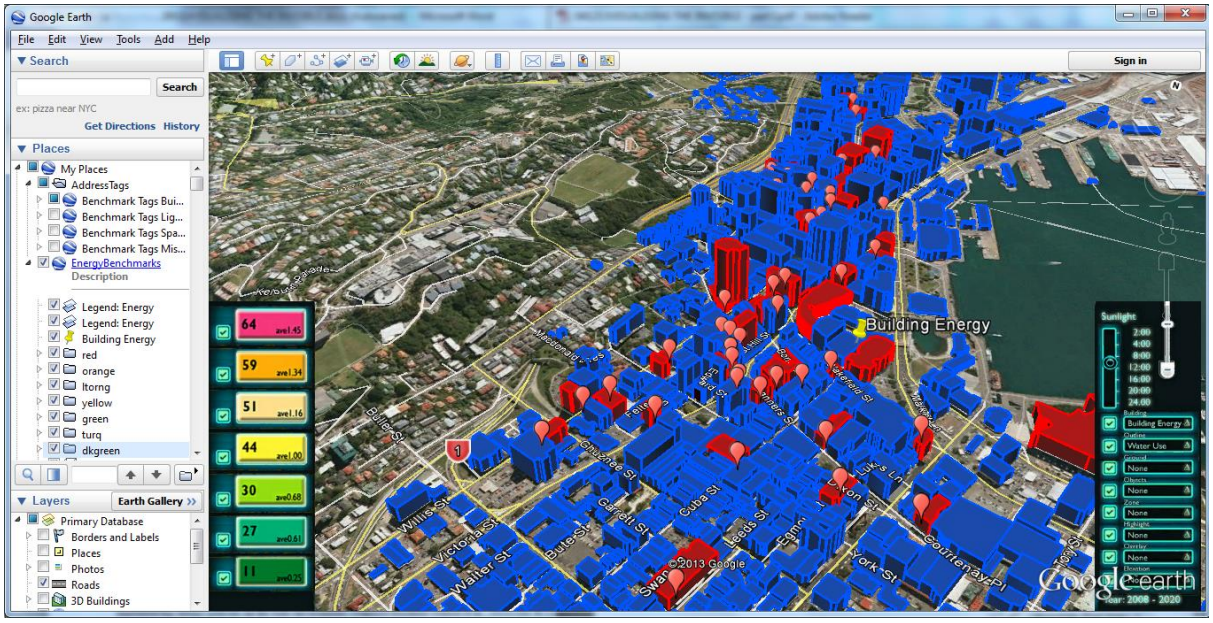
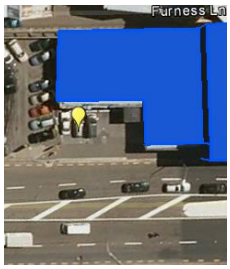
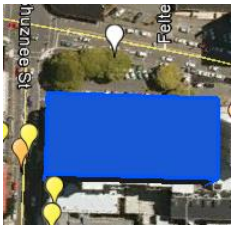


Figure 44 Colouring the 3D Model Manually for the Wellington Pilot Visualisation

There are several options for creating 3D automated city models however simple polygons imported into Google Earth and then modified to form building blocks ensure that the buildings can be viewed properly in Google Earth and Google Maps StreetView (including in situ on a mobile internet device) which insures the information remains accessible to the widest possible audience.

3.22: Algorithm Requirement – The Missing Link

The issues with fully-automating the construction of this visualisation lie with the connection between the address point data and the built forms which was overcome for the Wellington pilot visualisation by manually colouring the model within Google Earth (3.19). The disconnect between address and physical building is an issue for post deliveries, building research and the general public. Inaccuracies and inconsistencies in the address point data source have been covered in the literature review and in the previous section (2.02.07 and 3.21). The BEES resource use data, Websearch typology information, elevation height and meshblock-level statistics can only be geo-located approximately using street address information (text strings), whereas the building models are derived from accurately geo-located cadastral data, building footprint polygons and site parcel polygons sourced from 'Koordinates' (Corkery & Coup, 2007). To link the two in a national visualisation, it will be necessary to replace the manual colouring procedure with an alternative method using an algorithm (Figure 45):

	IF markers longitude / latitude coordinates (LINZ, 2011b) sit within a building footprint's coordinates (LINZ, 2011c). THEN Colour building footprint to match benchmark statistic colour code. Remove newly coloured buildings from selection set.
	IF >2 colour tags sit within a single polygon footprint (LINZ, 2011c) THEN colour building with highest energy colour from set (orange – in the example image), but highlight with code for checking / improving accuracy at later stage. Remove newly coloured buildings from selection set.
	IF marker coordinates sit within parcel boundaries (LINZ, 2011d) THEN colour closest polygon (LINZ, 2011c) found within parcel to match the marker colour code. (Check for validity measuring polygon volume is +/- 40% within limits of the volume given in the valuation data for footprint * elevation – IF NOT leave floor polygon/s in selection set. Remove newly coloured buildings from selection set.
	IF marker coordinates are within adjacent road area THEN match to nearest footprint (LINZ, 2011c) to marker. (Check for validity measuring polygon volume is +/- 40% within limits of the volume given in the valuation data for footprint * elevation height – IF NOT leave floor polygon/s in selection set. Remove newly coloured buildings from selection set.
	IF marker coordinates are in the centre line of the road and equidistant +/- 1m from >2 polygon footprints (LINZ, 2011c) THEN search for street address placemarks for street number +/- 2 (LINZ, 2011b) and colour *adjacent* polygon. (Check for validity measuring polygon volume is +/- 40% within limits of the volume given in the valuation data for footprint * elevation height) – IF NOT leave floor polygon/s in selection set. Remove newly coloured buildings from selection set.

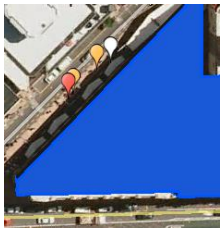
	IF >2 colour tags are to be allocated to a single polygon footprint (LINZ, 2011c) through any combination of previous instructions THEN colour building with highest energy colour from set, but highlight with code for checking / improving accuracy at later stage. Remove newly coloured buildings from selection set.
	IF there are polygon footprints remaining in the selection set, THEN apply colour from nearest colour tag assuming within the same or adjacent meshblock area (LINZ, 2011c) but highlight with code for checking / improving accuracy at later stage if address source data (LINZ, 2011b) is improved. Remove newly coloured buildings from selection set.
	Turn on all building polygons to display coloured model. IF any footprint polygons remain without a colour assigned, THEN allocate grey colour to highlight building/s sits outside of commercial census data set. (Focus future research on energy use for these building types.)

Figure 45 Colouring Algorithm - Illustrated Steps

The disconnect between address data and physical buildings has been previously identified as a barrier to carrying out quality research on the building stock and implementing energy policies in the built environment (Stanford Business School et al., 2012) (M. R. Donn, 2004). Whist the algorithm and manual colouring methods provide a short-term solution, for the purposes of this visualisation, identifying and geo-locating buildings accurately would mend the links between data and physical building forms and vastly improve the accuracy of future research on the building stock. A technology that generates a unique building identifier for each building could be used as a common identifier in any dataset with information on retrofit history, energy usage, typology and floor area. This would enable valuable mash-ups of separate datasets to occur on a national or local basis to a higher level of accuracy (Stanford Business School et al., 2012). In New Zealand the majority of commercial buildings (those with specified systems / compliance schedules) require a building warrant of fitness which must be renewed annually by a qualified building inspector. This means the assignment of a unique identifier with accurate geo-location could be carried out during the Building Warrant of Fitness inspection, simply by bringing a mobile internet device to the site and taking geo-located photographs. The data could be linked to any public information available through the compliance schedule / warrant document as well as existing valuation data, thus providing a far more accurate census dataset for future building stock research.

3.23: Creating a Dashboard Interface

Figure 46 Dashboard Interface Energy + Meshblock Zones

A dashboard interface was designed for use with a web-based Google Earth display. The functioning button and slider controls in the dashboard are activated with HTML code hyperlinks overlaying the Google Earth display. The function of the links is to switch on and off various combinations of KMZ files illustrating each of the datasets available for investigation. Korhonen recommends a playful user-interface to facilitate learning and inspire curiosity in a product (Korhonen et al., 2009). The dashboard has been designed to instigate a sense of playfulness, mimicking game design and neon light-like text giving a clue as to the focus of the visualisation – energy use. The selections light up when activated with drop-down menus providing a list of alternatives within each display medium (building outline, building form, ground plane, highlights and overlays). The black background behind the controls is faintly transparent to retain a visual connection to the 3D city forms to aid in navigation and zoom control. Various object-specific legends, keys and menus are displayed with right clicks on the data graphics. The following series of images (Figures 46-55) illustrate the pilot visualisation of Wellington CBD with building energy, energy end-use, water use and various combinations of these datasets. Overlaid council district plan, tsunami lines and examples of typology, energy and 'meshblock' queries are illustrated alongside the building resource use data. By toggling between each query it is possible to follow the 'heat' of high resource use as it navigates into different parts of the city area.

Figure 47 Meshblock Query: Average Energy and End-use + Typology Mix

Meshblock, Area unit and Territory zones of interest can be indicated (Figure 46) and highlighted (Figure 47). Building stock typology mix and energy / end-use benchmarks for the zone can be queried with a mouse right-click, and if data on delivered energy is eventually made public, there is potential for this to be compared with delivered energy data for the meshblock investigated.

For example by systematically querying meshblocks with a high number of red (high energy intensity) buildings and observing the mix of buildings present in each, a user of this visualisation may quickly uncover drivers of energy consumption.

Object based queries could be carried out on any level of data:

- Elevation (Orientation)
- Roof
- Building Block
- Outline
- Typology
- Meshblock
- Area Unit
- Territorial Authority
- Material types
- Building Use

If there is any risk of confidential material being disclosed due to the narrow level of investigation taken (such as querying a meshblock containing only one or two buildings) it would be possible to automatically move up a level to view the data at an aggregated level. A similar strategy to protect from disclosure of sensitive information. is used by Cameron Prebble in the Mashblock website (Prebble, 2010).

Figure 48 District Plan with Building Energy

Council information can be overlaid on the ground plane (Figure 48 includes a map of the district plan for Wellington) and legend information can accessed with a property query by right-clicking the ground plane image map (Figure 49). This would enable a user to investigate site information for a building or area of interest, which may be valuable for a designer or building owner wishing to upgrade or re-design their property.

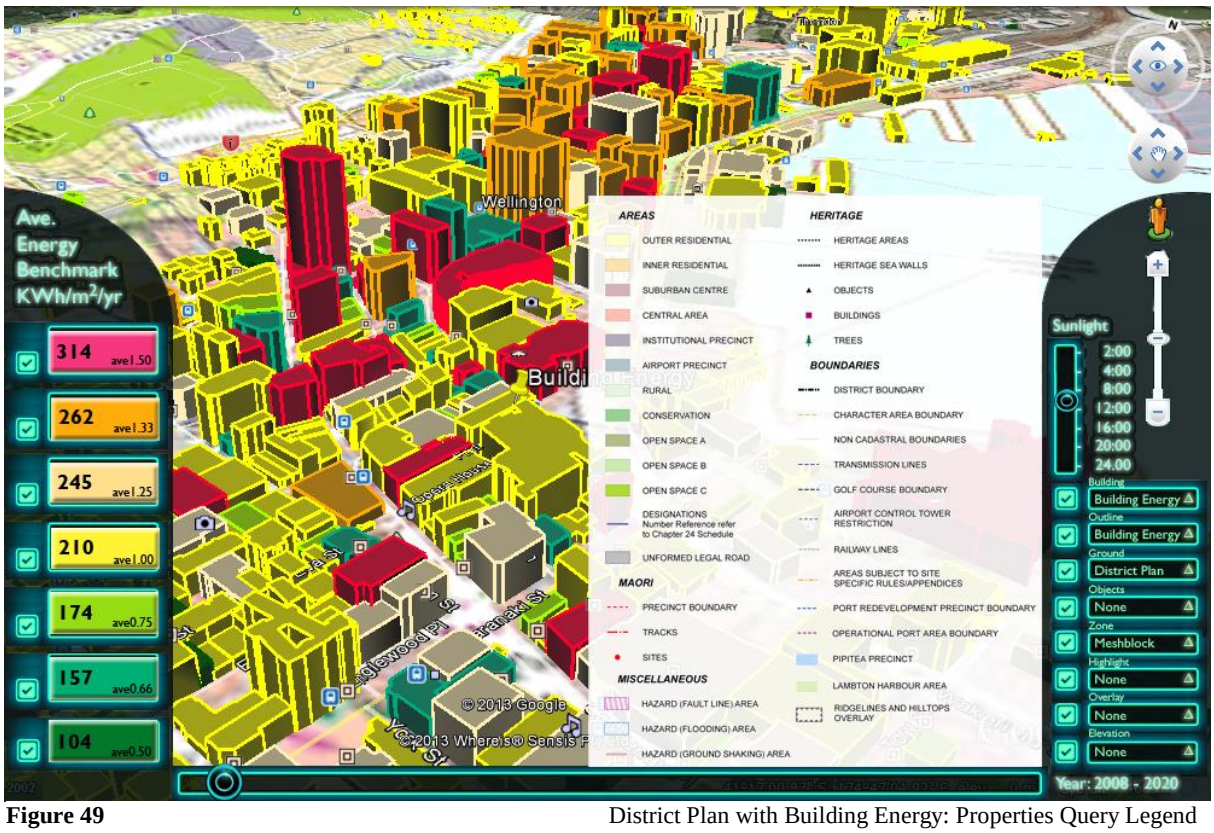


Figure 49 District Plan with Building Energy: Properties Query Legend

Figure 50 Lighting End-use with Tsunami Lines

In the example pictured above and at the top of the previous page (Figures 48 and 50), toggling between 'Building Energy' and 'Lighting End-use' makes it possible to witness the high energy (red) benchmark buildings gravitate from centre of town (Manners Mall), towards the business and retail district (Lambton Quay). The higher proportion of retail buildings found in the Lambton Quay area may be the cause of this change in concentration of high energy intensity buildings. Once this type of scenario is observed in the visualisation by the user, it may inspire increased interest in energy use generally, or perhaps lead to further research being conducted.

Figure 50 communicates Tsunami lines overlaid on the ground alongside building energy benchmark data. This ground plan map could just as easily communicate ground shakiness, view corridors, sun exposure and noise levels. By toggling the building forms on and off, it would be possible to view the tsunami lines without building forms blocking the view, whilst sill retaining an understanding of the relationship between the data displayed in each model snapshot. In another example of the power of this comparative information, Tsunami lines, could be overlaid with graphics illustrating the distribution of construction materials in the building stock. By displaying buildings with heavyweight materials versus lightweight materials (Section 3.12) it might be possible to provide information to the user on building survivability in the event of a tsunami.

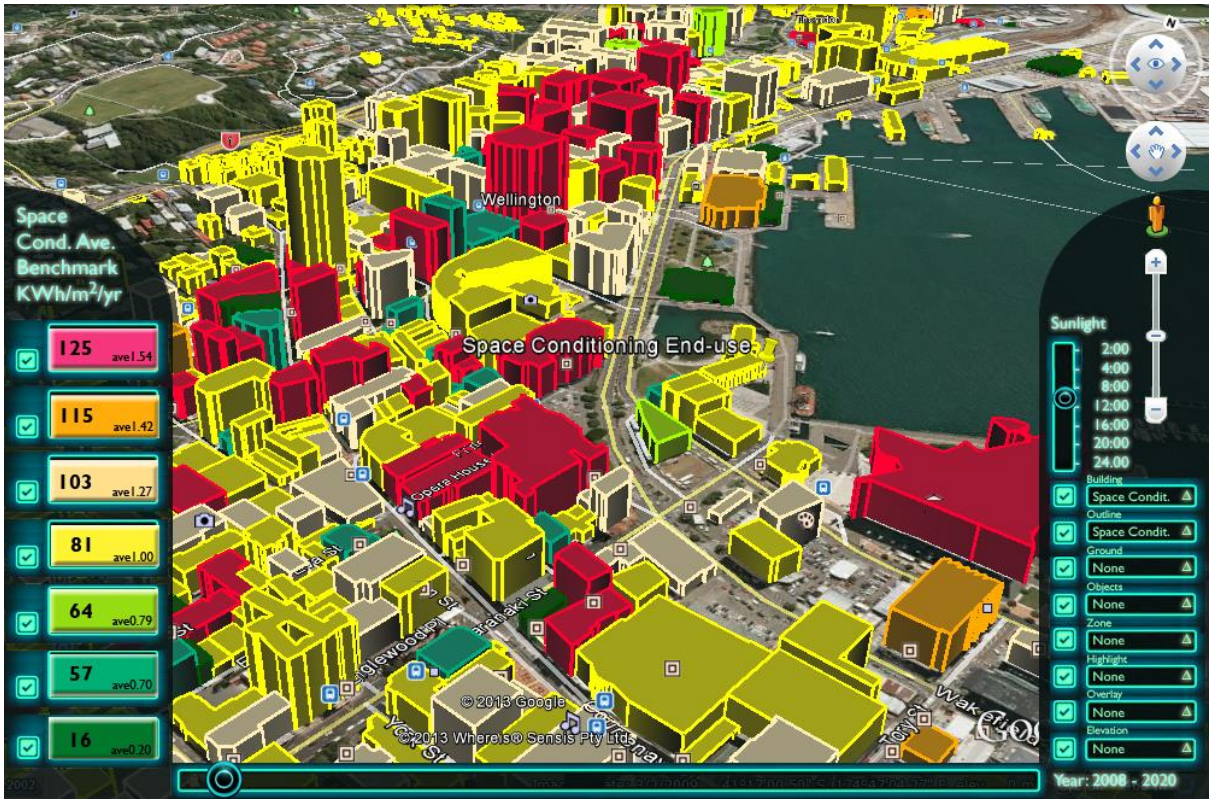


Figure 51 Space Conditioning End-uses

Switching to 'Space Conditioning end-uses' (Figure 51) the high energy outliers become sporadic around the city. This may prompt further investigation of the outlier buildings. A user may choose to query only the high energy intensity (red) buildings in this view above to see if there are any trends in the mix of typologies or building characteristics found in the selection set. A dialogue box containing statistics could be activated by right clicking the red button in the legend / key, as illustrated in the previous meshblock query example (Figure 47). The fact that energy intensity is described in 3D forms (rather than a 2D map) immediately prompts the user to take most notice of the buildings with high volume and consequently the highest overall energy use and the greatest potential for improvement.

Figure 52 Miscellaneous End-use Energy

A look at 'Miscellaneous end-use' (including refrigeration, cooking, equipment, lifts etc) in the same view suggests there are few low or high outliers in the centre of Wellington (Figure 52) which raises questions about the typical end-uses present in the inner city. This is perhaps due to building typologies found in the central city area being dominated by space conditioning and lighting energy end-uses, rather than heavy refrigeration and large scale cooking. It is also possible that this suggests the end-use breakdowns for miscellaneous energy are inaccurate since offices in tall buildings should have a proportion of energy devoted to lifts and these do not seem to be represented in the visualisation. If the visualisation had described these statistics in flat 2D heat maps, or by using the 2D GIS extrusions with Z-axis communicating the benchmarks it would be challenging to discern the volumes of energy involved, as well as understanding the urban grain and form of the buildings relating to the highest and lowest energy intensities.

Figure 53 Building Energy (Building Block) and Water Use (Outline)

Figure 53 illustrates the combination of two statistics displayed on each building form whilst retaining the form and urban grain of the city (note: unlike the energy benchmarks, the water-use colours used in the image above have not been based on real consumption data). This example illustrates water use and whole-building energy together. If a user is interested in high water use buildings the eye is immediately drawn to buildings displaying a dark blue outline, if interested in low water use the eye would instead be drawn to buildings with a green outline. Isolating these buildings in the visualisation by using a right click of the colour swatch in the key (turning off all buildings without these associated value ranges) would immediately present the user with a range of comparative energy benchmarks related to water use visualised in the walls of the building form (rather than outline). This would alert the user to any trends occurring between the two data sets in the process.

Due to the method Google Earth uses to display colours indicating a 3D form (light and shade) the main building colours appear as de-saturated versions of the colours indicated by the key which may confuse the user when attempting to discern values from the visualisation. Changing how these colours are displayed may reduce the appreciation of 3D afforded by the platform. It remains to be seen whether users of such a system would be satisfied with the combinations of colours and layering of information and whether the overall values are communicated effectively (Figure 53-54). This work would involve a series of focus groups as discussed in Section 2.03.03 and in Chapter 7.

Figure 54 Energy (Building Blocks) 'Mash-up' with Lighting End-use (Outlines)

The time slide-bar at the base of the dashboard is proposed for longitudinal data to be displayed with units ranging from hours, seasons or decades. This tool's inclusion depends on sufficient energy data either; made accessible by energy providers; through further research; or submitted by visualisation end-users. This time bar has the potential to animate statistics about energy pay back periods, the effects of rising cost of fuel, net zero building potential (Cory, 2015), solar gains, and sunlight access. The slide bar scale on the right hand side control displays units specific to the current selection. It allows adjustment of linear based settings such as hour, day, season, time, cost.

Creating this pilot in Google Earth KMZ file format, means that the visualisation can also be viewed in Google StreetView on a computer screen or in situ 'augmented reality style' as you walk around the city with a mobile internet device (Figure 55). By dragging the StreetView icon (indicated with dotted arrow line) onto the visualisation in an area of interest, it is possible to switch to a view from ground level. This provides a visualisation tool that acts as a 'lens' to view the world through effectively communicating the invisible *resource use* (Section 1.01). It provides an x-ray image uncovering meaning from the built environment forming a possible precursor to augmented reality display (Google, 2013c).

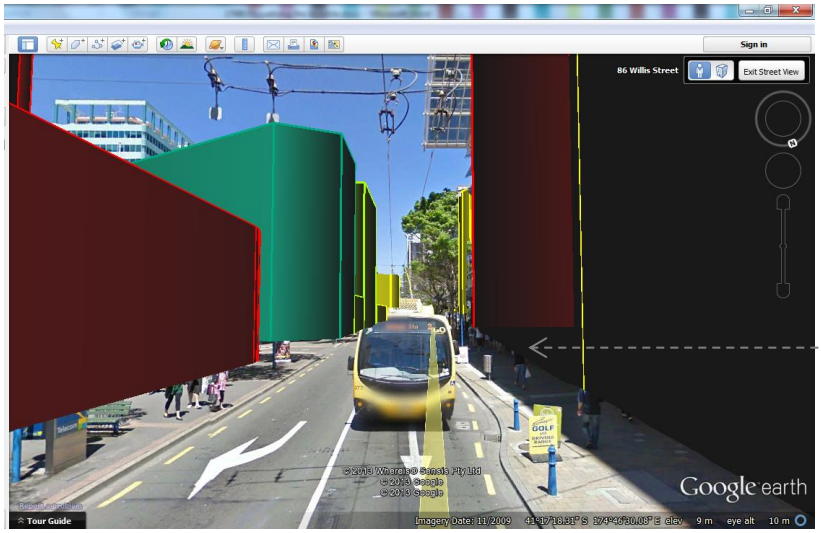


Figure 55 StreetView Image of Energy Benchmark Visualisation

4.00: Graphic Representation

The pilot visualisation constructed for this thesis is presented as a set of Google Earth KMZ model layers (a file containing geo-located place-marks, custom building labels and 3D model data). It is possible to view coloured block building forms with coloured outlines, overlaid on Google Earth satellite imagery mapped on topography, or even overlaid on Google StreetView® panorama images as you move through the virtual globe at street level. The layer options in the dashboard allow you to select a criteria for outline and another for building block colour communicating a variety of statistics (energy, water, end-uses, typology, net zero potential - refer Section 3.10 (Cory, 2015)) against each other. Ground conditions, council zones or tsunami levels can also be mapped as a geo-located topography texture image.

Heat loss, glazing ratios, solar access, wind and pollution levels could even be mapped in 3D as building textures. Imagine a façade coloured with solar gain information or noise level / pollution data forming a visible 3D "volume" of statistics which would be all the more legible in a dense city area. In this way it might be possible to visualise 3D statistics as a volume they relate to the built surfaces of the city facades, revealing a whole assortment of data that wouldn't be otherwise visible with 2D maps and GIS.

Being able to toggle quickly between different statistical datasets by clicking on and off layer options will make it easier for the visualisation to reveal relationships, patterns, and trends, potentially uncovering a link between two groups of statistics such as say - high energy use and curtain-walled buildings. This visualisation would assist leaseholders and property purchasers in their assessment of the value of the building stock and has the potential to uncover issues with design and efficiency upgrades and so would be even more valuable for designers and building professionals. Architects and designers have the potential to make the most of the 3D graphics but this technique also opens up vast information stores to the general public that would normally be found in research journals and international conferences.

The benefits of describing statistics about the built environment that are true to city form, volume and grain, are easy to imagine. Examining data at various levels of detail (elevation, building, meshblock, area unit and territorial authority) would provide a rich understanding of building stock makeup. As the quality of open data underlying the visualisation improves, further relationships between different floors / facades could be identified. Image-surface maps could be added to communicate information such as solar access, noise and wind levels, the tsunami evacuation zone, council zones, floodplains and earthquake hazard ground conditions. When viewed by a building professional engaged to employ efficiency measures, these previously invisible layers of information could quickly be taken into account in the design response. The visualisation allows a designer to directly compare and toggle between all publicly available information related to a particular site. This will undoubtedly reduce the time spent researching the project as well as reducing design fees for the building owner. From the placement of solar panels or photovoltaics by mapping simulated solar access on rooftops and facades to the siting of acoustic windows to cut out the worst of the noise pollution with one click of a mouse. The designer might combine information on glazing ratios with solar access for the best location for siting thermal mass within a building. The question then becomes - how much data can be overlaid before the visual becomes over-crowded, off-putting and confusing? This could be tested with the use of focus groups from various building industry professionals and lay persons to provide information with which to further refine the visualisation.

The colour ranges used will need to be tested for varied perception (Section 2.03.06). It is possible to convert the colour ranges to allow for common visual impairments but further research would need to be done to confirm they communicate high and low values as clearly as possible without eroding any of the 'light and shade' required to appreciate 3D forms. Tactile and haptic feedback (vibrations / sound) not only have the potential to make this data accessible to the blind but also the added feedback for able-sighted persons could provide a further dimension to any visualisation (Institute of Electrical and Electronics Engineers, 2009).

It will be important to use web-navigation tools to enable users to back track through their investigations – further enabling the toggling between statistical visualisations. When expanded to cover the whole country, care would have to be taken to ensure the data can be downloaded quickly in small 3D model sections focused around the field of view investigated by the user. If a user navigated to another area of the country, the graphics would need to be able to regenerate quickly in the new field of view. Google Earth and other 3D software tools originally intended for gaming such as 'City Engine' (Section 1.05) have algorithms which make the navigation between different levels of detail and views as streamlined as possible.

The 3D visualisation interface proposed as part of this thesis has the potential to provide a comparison between two sets of statistics using a combination of the building form and building outline colour. Buildings with a specific status or characteristic (heritage, earthquake risk, building use, age) are highlighted with a glowing halo. Right-click 'properties' queries on any layer activates legends, source data, benchmarks or typology mix depending on the object of interest. This makes it possible to view building characteristics alongside resource-use statistics - either by viewing the 3D visualisation as a KMZ "through" StreetView™ where photographs of buildings are also visible, or alternatively by attaching 'spider diagrams' or 'cross-hair' icons to each building model.

Building owners wishing to report 'actual' energy use indices for their building ideally would be able to submit their energy bills and floor area figures for consideration to the visualisation, just as Google has a 3D warehouse for Google Sketch-up models. This provides longitudinal data which can be viewed over time illustrating the (hopeful) gradual 'greening' of the building stock. This provides a 'gaming' dimension to the visual – promoting the playful to encourage wider appeal.

Combining this data with other publicly available information about the surroundings may influence design / purchase / lease decisions and has the potential to raise the value of sustainable buildings.

4.01: Layer Control Menu

The menus activating the various options for visualisation display are indicated on the dashboard with a brief description and an arrow which activates the dropdown sub-menu for each layer.

Each layer can be switched on and off using the toggle tick boxes (Figure 56) and the description of the criteria being displayed is visible in the adjacent text box.

Right-click 'properties' queries on objects in the visualisation provide legend, data source and aggregate typology information depending on

Figure 56 Layer Control Menu

what kind of object is being queried. The type of information displayed within each layer control is covered in the following sub-sections.

4.01.01: Building Fill Colour

The building fill colour menu will enable the building forms to be coloured with whole building energy use index, water use, building typology, net zero potential and end-use energy (lighting, space conditioning, miscellaneous). The units are energy - kWh/m^2/yr, water - kl/m^2/year and net zero potential - (defined as easy, moderate, aggressive according to the effort required to make the building net zero energy (Cory, 2015)). Other category based topics would have legends appropriate for the area being investigated Each of the colour ranges is identified with a colour legend scale on the dashboard with the average value per range clearly noted and content-specific depending on what is being investigated by the user. By isolating buildings of interest (and making all other polygon building forms invisible) it would be possible to illustrate benchmarks for a building with certain characteristics. Examples of the dashboard with a variety of building block visualisations of whole building energy through to various end-uses are illustrated in Figures 46-54 in Google Earth mode and Figure 55 in Google StreetView mode (Section 3.23).

4.01.02: Building Outline Colour

The building outline colour can be set to display any one of the criteria described for building fill colour – including highlighting buildings with chosen characteristics with a glowing outline. When conducting this part of the pilot visual it was soon realised that the building forms appear very dark without their bright outline colours making the comprehension of scale difficult. When attempting to comprehend water (outline) against building energy statistics (building form) in the pilot, it became challenging to distinguish the two vastly different colour scales. Whilst focus groups have not been conducted as part of this thesis, early indications suggest that the toggle function between views is more powerful than trying to comprehend two different statistics at once - the outline against a shady building form colour. A drop-down menu titled 'Outline' in the dashboard provides access to each data set. Examples of the dashboard with a variety of building outlines illustrating whole building energy benchmarks, end-use benchmarks, and water use benchmarks are illustrated in Figures 46-55, Section 3.23. A combination of building energy (building block) and water use (outline) and energy (building block) and lighting end-use (outline) benchmarks are also illustrated (Figures 53-54).

4.01.03: Ground Texture

Any image can be 'draped' over the topography in Google Earth. Overlay images providing extra information in the visualisation include:

- Ground Conditions
- Tsunami Lines
- Landslip Risk
- Noise
- Wind
- Solar Access
- Energy Land Intensity
- Council District Plan
- Wind Zone
- Temperature

A drop-down menu in the dashboard provides access to each data set and the right-click 'properties' queries provide legend and data source information as required. The district plan, meshblock outlines and Tsunami lines are illustrated on the ground plane in Figures 46-50, Section 3.23.

4.01.04: Objects

Live-linked 3D objects could be added to the visualisation to provide further information about the city. These might include shifting objects such as sun paths, traffic (density / average speed), public

transport (real-time location data), and wind (wind roses indicating speed and direction), as well as a number of static 3D objects such as recession planes, setbacks and maximum building height. The ability to display wind, traffic, heritage and earthquake prone buildings has been previously explored in work carried out by the Centre for Building Performance Research (Sullivan & Motley, 2011).

The two slider controls described in Section 4.02.03 have been designed to allow this information to be viewed over time as well as producing static images for a particular time.

An example of the potential of this layer might be to compare a new building design to the council's maximum building envelope just by loading in the building design model and recession plane objects at the same time. Further work using the building envelope volume could be useful for assessing likely sun penetration after development has taken place. A drop-down menu in the dashboard provides access to each data set and right-click 'properties' queries provide data sources or aggregate data as required.

4.01.05: Zone

A series of semi-transparent or wireframe layers made up of 2D polygons can be added to the visualisation and used to communicate elevation, meshblock, area unit and territorial authority perimeters. Querying these objects with a right click has the potential to link to zoned, aggregated information forming part of the 'properties' feature. The zone outline thickness must be scaled to at least 5 points width to be clearly discernible when viewed with building block data as this is already graphically very 'busy' (satellite imagery overlaid with coloured building blocks including outlines of their own). If the zone outline is appropriately scaled to view resolution it will be clearly visible at a variety of urban and local scales.

A drop-down menu in the dashboard provides access to each data set and right-click 'properties' queries provide data sources, typology mix, or aggregate data on the topic of investigation. Meshblock zones are indicated in pale turquoise in Figure 57.

4.01.06: Highlight

Highlights can be selected at any detail level (meshblock, area unit, territorial authority, building, and elevation) to identify any particular building typology or characteristic (Figure 58). The data sets able to be identified successfully with highlights include queries on:

- Historic Buildings
- Conservation Areas
- Earthquake Prone Buildings
- Building Use
- Building Age
- Building Material
- Roof Material
- Glazing Ratio
- Overshadowing

Figure 57 Highlighted Meshblock Example

The highlight potentially identifies a group of buildings, single façade or street block with a glowing boundary outline. Buildings not included in the highlight can be desaturated to assist the identification of the area of interest. A drop-down menu in the dashboard provides access to each data set and right-click 'properties' queries provide data sources or

aggregate data as required. The potential confusion of turquoise meshblock boundaries with the turquoise found in the building model as illustrated in Figure 58 requires further investigation and review, potentially when conducting a focus group on the visualisation (Section 7.02).

4.01.07: Overlay

The visualisation pilot clearly demonstrates the power of statistics being displayed in 3D rather than as 2D 'heat maps' or 2D extruded GIS formats. As demonstrated in Section 3.23, allowing the user to isolate groups of buildings with certain building characteristics, surrounding features, zoning, or particular benchmark values, inspires theories on the distribution of resource use to be investigated, as well as raising interest in the built environment generally.

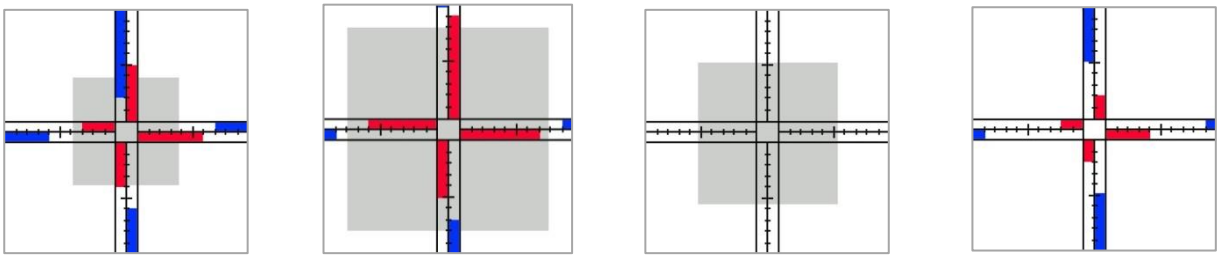


Figure 58 **NESW** Cross Hair Format: Overshadowing /Glazing Area/Inferred Insulation Levels

A number of graphic styles were considered for this icon design and layout including simple linear scale 'cross hairs' (Figure 59) and spider diagrams (Figure 60). The four cross-hair graphics indicate glazing ratios by orientation (red), overshadowing daylight factors by orientation (blue) and inferred insulation levels (grey) (Cresswell-Wells et al., 2012). The scale from the centre of the cross hairs to the outside edge of the icon runs consistently from 0-100% The four diagrams illustrate examples of: a) high overshadowing, medium glazing ratios and low insulation levels; b) low overshadowing, high glazing ratios and high insulation levels; c) no overshadowing, no glazing area with medium insulation levels; and d) low overshadowing, low glazing ratio and no insulation. A solid grey square for the full icon dimensions would indicate insulation levels up to current building code level.

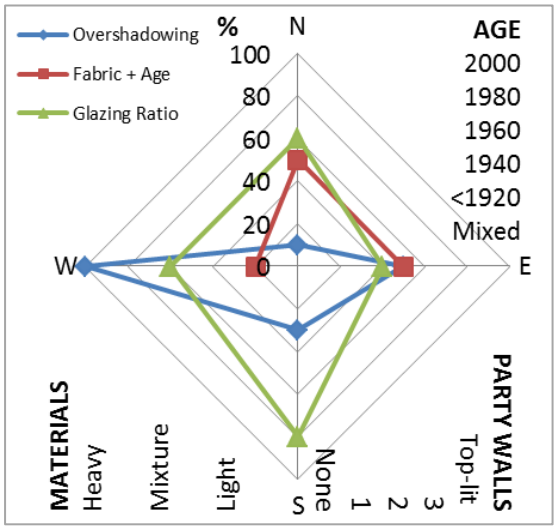


Figure 59 Spider Diagram Format: Overshadowing, Fabric + Age and Glazing Ratios

A spider diagram format (Figure 60) was also developed for describing more complex building characteristics. The orientation linear scales are from 0-100% as before, but the linear steps also stand for building age, materials and party walls and therefore many unit types need explanation, for the diagram to be read properly. The use of blue for overshadowing, green for glazing ratios and red for fabric and age details seem an intuitive choice for the display of complex typologies, but this graphic and the data choices used in the icon would need to be fully-tested using focus groups and user-surveys. The infographics could be 'attached' to buildings as view-display scaled icons or linked with a dotted line (Figure 61).

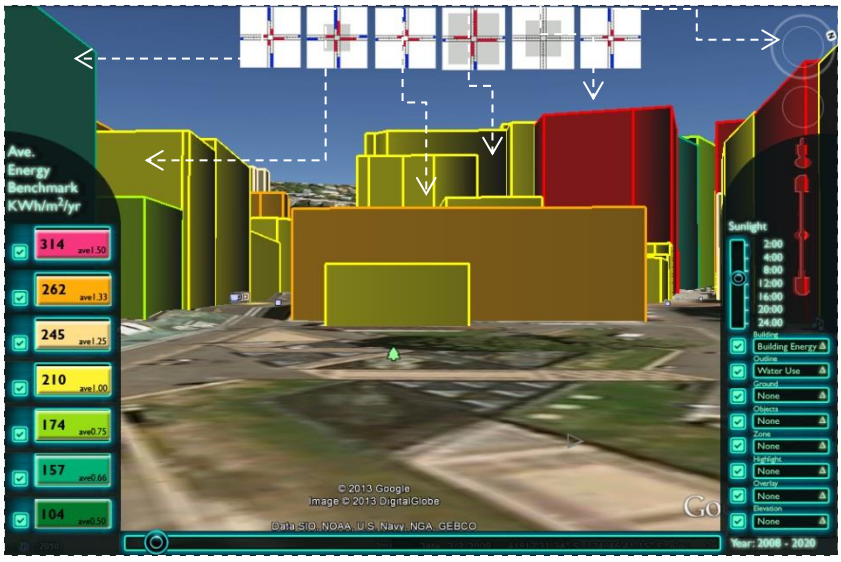


Figure 60 Overlaid Icon Infographics - Building Typologies

Overlaid data could further enhance the additional power already provided by the 3D display. By including icons identifying buildings by typology or characteristic mixes without removing the underlying benchmark statistics, many more opportunities for learning and research inspiration should be presented to the user. By activating icon info-graphics linked to each building in the immediate field of view it would be possible to identify typology combinations that tend to relate to high or low resource-use, further revealing more information about the building stock within the field of view than already provided by the 3D visualisation. Refer to Figure 61 for an example of the display of these icons from ground level.

4.01.08: Elevation Mapping

Any simulation rendering carried out on building block facades would need to be restricted to a small selection of the visualisation, much like a test rendering image is conducted within a certain boundary area, to cut down on the amount of time taken to render and view the information. The 3D appreciation of wind / heat / light as mapped colours projected onto building facades is much more comprehensive than a 2D approach (Figure 62-63).

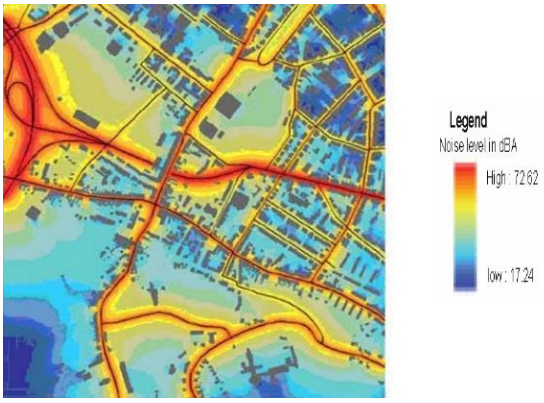


Figure 61 2D Traditional Noise Mapping

In the thesis 'a GIS-Approach to 3D Noise Modelling Using 3D City Models' Kurakula argues that 2D noise maps provide data on noise levels from only one fixed height however noise transmits in all directions and therefore pollution is of concern within and around buildings at various heights (Kurakula et al., 2007). This concern applies to noise, light, air quality and heat as well as the communication of business and population density 3D distribution. Kurakula uses 3D coloured contour lines to model the 3D volume statistics (Figure 63) which presents a graphic where the area of most interest requiring interpretation is likely to be the boundary between inside and outside of a building – therefore mapping the colours projected onto building facades seems to imply the volume and distribution of 3D data on 2D planes effectively.

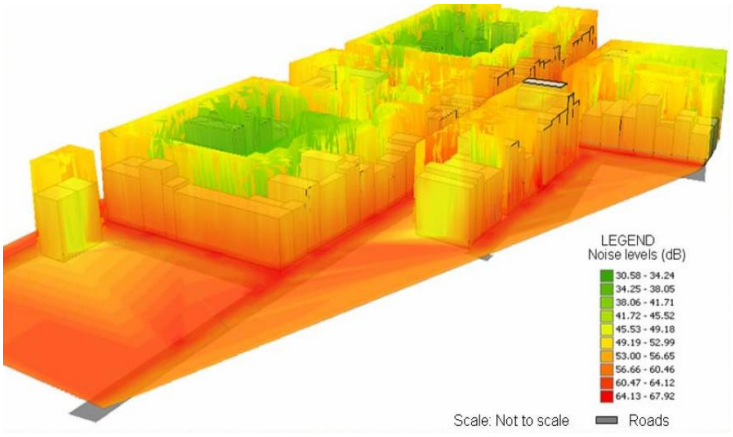


Figure 62 3D Noise Mapping onto Building Facades

Although the tool that would provide the simulation data has not been fully-investigated as part of this thesis, the sorts of data available for display might include:

- Solar Radiation
- Solar Access
- Air Quality
- Building Fabric Heat Loss (infra-red scans)
- Population Density (persons per m^2)
- Business Density (businesses per building)
- Party Walls
- Glazing Ratios

As with other menus in the dashboard a drop-down menu would enable access to each data set and right-click 'properties' queries would provide data sources or aggregate data as required.

4.02: Tools and Controls

4.02.01: Dialogue Box Properties Queries

All KMZ files displayed in Google Earth have the potential to display selection-specific data source information (text / html links / code) as well as images and graphs (typology mix, characteristics, legends and keys). A simple right-click with the mouse on the object of interest brings up associated data sets. The escape key (ESC) returns users to the previous screen. The pilot visualisation illustrates an example of the type of display a properties query might activate in Figure 49 Section 3.23.

4.02.02: Data Legend Display

The data legends used to communicate the colour scales in the visualisation are selection-specific. Where the building block and outline have matching settings, a single colour range legend will be illustrated. The details provided with each colour in the range include, the average value or description and if applicable, its relationship to the overall average for the commercial building stock (yellow in the case of energy) (Figure 63 left). If two different scales are needed (for different criteria selected for outline verses building block) the average value or description will be displayed along with its corresponding colour but with the relationship to the average omitted to save graphic space and reduce visual clutter (Figure 63 right).

The double legend colour scales will be displayed with the outline colour scale on the right hand side. It will clearly overlay the building block colour scale and will be highlighted with a glow around the boundary in the hope that this graphic implies its being the data range represented by the building outline.

Figure 63 Data Legend Display Types

4.02.03: Selection-Specific Controls

Additional selection-specific controls (Figure 64) will be required for certain layer control options, setting the time of day, year or decade or toggling on and off to view sunlight / sun path, or to select certain building typologies and characteristics as listed below:

- Historic Buildings
- Conservation Areas
- Earthquake Prone Buildings
- Building Use Types
- Building Age
- Building Material,
- Roof Material
- Glazing Ratio
- Overshadowing
- Form Type
- Building Use
- Storey Height
- Sunlight
- Time of Day
- Season
- Decade

Figure 64 Selection-Specific Control Panel

A time slider bar would allow any longitudinal data to be displayed over decades, hours or seasons. This animates the visualisation and may infer relationships changing over time depending on the criteria being investigated. The time slider mechanism has the potential to illustrate: building stock benchmarks changing over time; the effects of rising energy costs; and the rate of net zero building upgrades. Dynamic data regarding the environment can also be viewed over time including: sunlight; noise; wind; and air quality effects. Figure 65 illustrates the time slider bar from the pilot visual dashboard design layout.

Figure 65 Time Slider for Longtiduinal Data (Net Zero, Stock Improvements, Energy Costs)

4.02.04: Data Entry Form

The energy and end-use average benchmarks provide a base level of detail about the existing building stock from open data on typologies and characteristics present in each building. These allow owners to compare their actual energy intensity to benchmarks, which may alert them faulty equipment and appliances, or perhaps cases where a premise is unaware they have been paying for their neighbour's energy needs.

A powerful option with the potential to introduce longitudinal data into the visualisation, might be to include a system whereby building owners can advertise their accomplishments boasting of the energy efficiency upgrades they have made. General data from BEES combined with building characteristics provide the baseline data and public energy certificates and building owner-submitted records would allow comparison of benchmark with actual energy use. Whilst there are many attempts being made to introduce voluntary building energy certificates and rating systems internationally (Section 1.01) the advantages of making this information publicly accessible and accessible for all buildings has the potential to increase the market value of sustainable, efficient buildings. Owners would simply send proof of ownership, floor area and energy bills to be confirmed and entered into the visualisation for all to see. The time slide bar would be able to communicate the gradual changes in the building stock over time and building owners would be asked to update the information periodically since a change of ownership or behaviour could lead to very different energy use patterns.

In order to include accurate end-use figures it is important to confirm the exact uses, occupancy and equipment present in the building. Deciphering what happens within the building is key, because occupant behaviour is more important for accurate end-uses than building properties alone. A building

owner might also provide lists of appliances and equipment / occupancy / lighting although this vastly complicates any procedure and may turn owners off the idea altogether. If building owners are invited to submit real data about their buildings it is easy to imagine that the database will gradually be populated with more and more low energy and low water use buildings (since owners are unlikely to report resource-guzzling buildings).

The central administration that receives the actual building data from owners is envisioned much like Google's own 3D Warehouse, which supports the construction of simple 3D models for inclusion in Google Earth subject to audit. In 3D Warehouse the audit is mainly automated and is therefore a low cost service, however the ownership, use, energy bills and floor area information submitted to this visualisation is likely to require independent spot-checking to confirm accuracy (Google, 2013b). In the Europe and Australia public energy certificates provide an alternative regularly-audited source for this data stimulating positive competition between building owners in the process.

Figure 66 illustrates a building-level properties query. The properties listed include all open data available on the building being investigated. A hyperlink at the base of the list provides a link to a screen where the fields are able to be modified, just as entries are modified in a Wikipedia page. Values are edited by the owner and then supporting documents must be uploaded in support of the data. Property queries are available at every level of detail from elevation, building, meshblock. area unit and territory but only building level queries have the potential for modification since all other information comes from open data at an individual level, or in the case of energy delivered to a meshblock, at aggregate level. As an added incentive for providing actual energy or water use data from a building it would be possible to use the properties query to advertise businesses with information and opening hours added to the properties data, much like Google Maps.

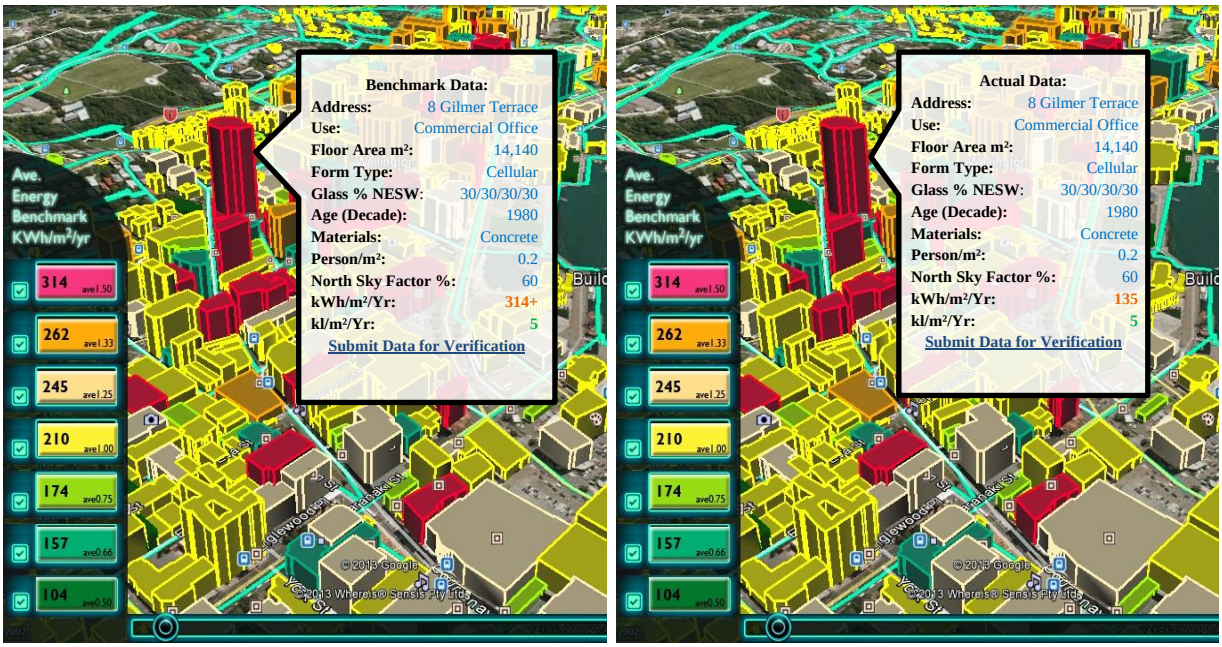


Figure 66 Data Entry Form – Building Level Query - Actual Energy Data from Building Owner

5.00: Data Quality and Privacy Issues

This section reviews the coverage, bias, analysis and quality management of the Websearch data and concludes with an overview of data privacy issues.

5.01: Coverage

Building records with coverage of both aerial, StreetView images (80%) as well as those with only one source (9%) and those with none (11%) are illustrated in Figure 22, Section 3.11.

A map of Google StreetView coverage in New Zealand is highlighted in Figure 21, Section 3.11, although due to the scale of the map some detail of the missing StreetView coverage is hard to discern. StreetView images were available in all major city centres, however Figure 67 lists New Zealand areas where StreetView coverage was found to be lacking.

Bay of Plenty – Tauranga, Mt Maunganui, Omanu Beach, Arataki, Whakatane
Gisborne – Whangara
Hawkes Bay – Wairoa, Flaxmere
Nelson – Beachville
West Coast – Greymouth, Hokitika
Canterbury – Christchurch, Harewood, Mcleans Is., Burwood, Airport
Marlborough – Blenheim, Lower Wairau
Northland – Whangarei, Kaipara, Mangawhare
Wellington – South Coast

Figure 67 StreetView Limited Coverage Areas

Between 2008-2012, additional StreetView imagery and new Google Earth aerial photos were uploaded regularly. By the time the Websearch data was being checked for errors and missing data towards the end of the project, the majority of building records abandoned by research assistants due to lack of data, were in fact found to be quite visible in later revisions.

Street View allows researchers to see the number of storeys in a building and often identify the business/building type, form, materials, glazing and overshadowing from surrounding buildings. The valuation data related to 2008 while the Google Earth Satellite Images and StreetView omnidirectional imagery studied were captured over a wide range of dates, so images were occasionally found to be out of date and inconsistent with aerial photos, contradicting the omnidirectional images. Research Assistants were directed to prioritise data in the StreetView images as they were generally considered the most recent and reliable information available. StreetView imagery limitations are listed as follows:

- Researchers were only able to identify buildings when they were on street fronts as it was impossible to see down drives or beyond natural features. For instance a number of industrial properties and other uses tend to be battle-axe (rear properties) blocked by street front buildings or features under different ownership and address.
- There are zones throughout New Zealand that have no StreetView, but these are decreasing every year. These are shown in Figure 67 above and they are generally less populated areas or the outskirts of towns. A map of StreetView coverage in the whole of New Zealand is illustrated in Figure 21, Section 3.11.
- Some business signage and building facades were unable to be seen from StreetView as they were too pixelated, with objects in the way, glare issues or not viewed at the right angle, however many of these records were resolved using further Google searches and refinement of the building uses.
- The addresses did not always match perfectly so assumptions were necessary.

- StreetView shows some buildings as vacant, under construction or absent, due to the time the StreetView panorama was taken. However this may differ to the current situation on the site (or indeed the valuation record which was based on 2008 building stock data).

5.02: Typology Bias

The potential bias in terms of typology mix was investigated in each of the sample datasets as well as the population. Typology groupings relating to climate, urban-rural mix, size, age, use and mix of uses were observed for differences using a series of standardised histograms. Figures 25-32 Section 3.13 illustrate typology charts with notes added discussing the possible causes for any differences between samples and census datasets. These Figures also indicate the mix of floor area within hybrid typology criteria used in the application of energy results to the visualisation.

The various data sets needed checking for consistency across the census, Websearch and sub-samples. A set of building typology graphs were produced in identical formats for each of the following samples to assess the comparison between building typology + characteristics mixes:

•	Population (Valuation Data - Commercial Buildings)	50,548 Buildings
•	StreetView (Websearch) Sample	3,042 Buildings
•	Phone Survey (Aggregate Study) * *from 980 Premises*	*445 Buildings
•	Monitored Energy Survey (Targeted Study)	100 Buildings
•	Energy Bills Survey (Aggregate Study) * *2 outliers removed*	*251 Buildings
•	Wellington CBD Population (Visualisation Selection Set)	786 Buildings
•	Water Bills Survey	34 Buildings
•	Simulation Model Sample (VUW-Centre for Building Performance Research)	48 Buildings

Conducting the bias analysis not only indicated reliability of the sampling methods used in the BEES study across a whole array of building typology criteria, but also provided a way of understanding how the energy bills information might differ from the national census, the Websearch and the Wellington CBD population. The latter two datasets were crucial since the energy bills information was eventually to be applied to the Websearch for the generation of benchmarks distributed amongst the Wellington CBD buildings. As discussed in Section 3.13, this indicated a prevalence of certain building typologies and lack of others which appeared to result in a much higher average energy intensity overall (293 kWh/m^2/yr), than that of the national data set (204 kWh/m^2/yr). This provided clues as to the building types which might be expected to exhibit higher or lower than average energy intensities, which led to the selection of hybrid typologies used in the creation of benchmarks.

In addition a series of frequency tables were produced with colour conditional formatting according to percentiles. Samples investigated for characteristic mix included: the 3,042 Websearch records (building ‘A’s only); those records with additional buildings ‘B’ ‘C’ ‘D” through to ‘K’ associated with a single street address; Wellington buildings as well as separating the non-BEES use buildings from the buildings with BEES uses. The characteristics investigated included materials across all of the size strata groupings and age, storey height, glazing ratio, overshadowing across all building use type mixes.

As the building stock typology make-up will gradually be changing over time, it will be important to maintain the quality of the data. The source valuation data and this analysis of typology mix captures a snapshot in time (2008) and conveniently for the BEES research work the majority of Google StreetView imagery was collected throughout New Zealand in 2008 and released digitally at the end of the year. It is worth noting that this does not apply to Google Earth aerial photos, which vary

dramatically in age depending on location. When updating future versions of the Websearch data, the ideal timing would be to coordinate this so it follows the next major update of Google StreetView imagery. Valuation data can be accessed at any chosen point and therefore should not dictate the timing of the research. The author would recommend, as a minimum, that the energy bills survey is updated within 10 years, and preferably extended to include non-commercial buildings. This new information would potentially be available for comparison with the 2008 dataset (along with any user-submitted data) to provide longitudinal data on energy intensity in buildings.

5.03: Analysis

The initial validation work on suspect records established the level of basic errors present in the valuation data. Analysis of the Websearch coverage suggests buildings with aerial-only numbered 8% of the sample; while StreetView-only were available in 1% of the sample. Both aerial and StreetView were available in 80% of the building records which meant that the majority of Websearch records would benefit from data collection from both sources. A list of common 'error' codes were established for assumptions made by the research assistants as well as reasons for no or partial information entered into the spreadsheet (Figure 23, Section 3.12). These 'error' and 'researcher assumption' codes were analysed for their frequency in the Websearch data collection. There were 1,061 building records (35%) with 'error' codes attached and many of these had more than one code, however most 'errors' merely stated certain peculiarities found in the record such as high building surface area, the inclusion of a car park / greenhouse or storage space. There were also a number of records stating that the QV data was incorrect and therefore updated in the Websearch data collection. These superficial 'errors' accounted for 14% of the building stock, or 426 buildings in all (Section 3.12) greatly reducing cause for concern regarding the reliability of the Websearch data.

Figure 68 illustrates the occurrence of each of the errors out of all those records registering an issue. Of the remaining, potentially more significant, errors (representing 21% of the population), lack of StreetView or aerial imagery and address or building identification issues dominate the picture. In addition there are vacant buildings / empty sites (affecting 1.4% of the population). Within the

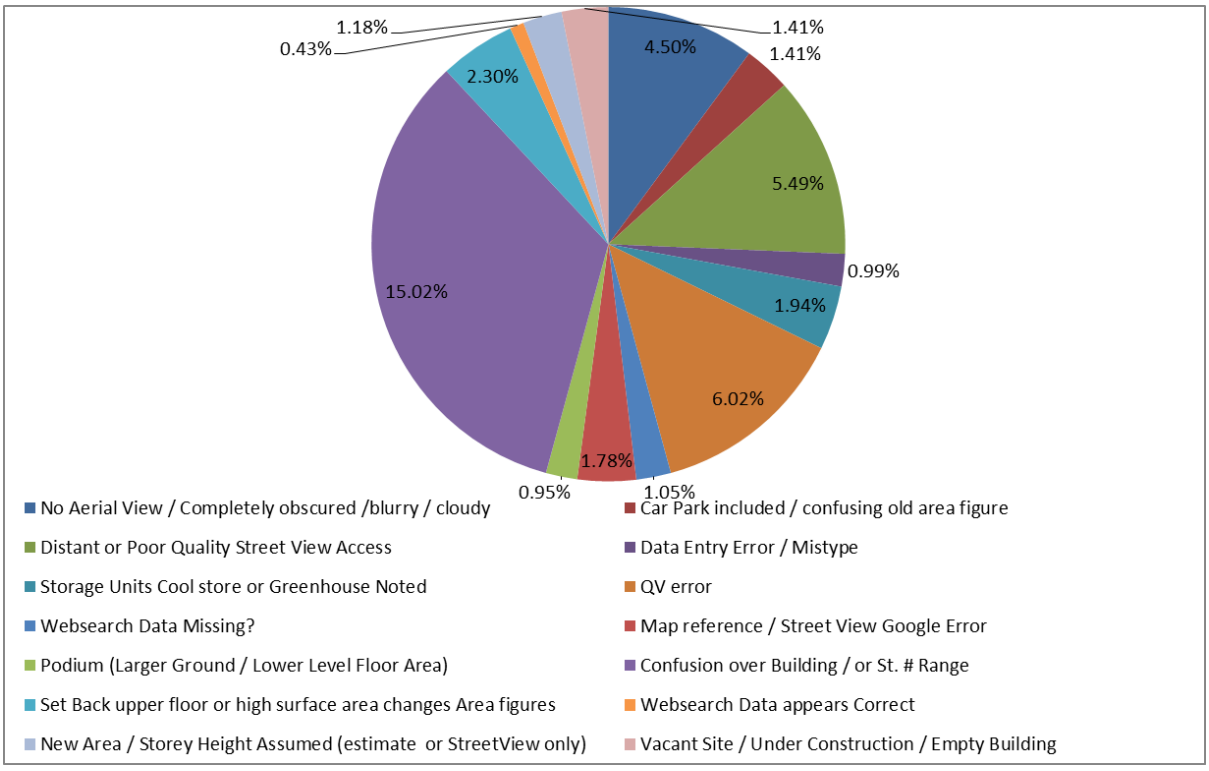


Figure 68 Typical Errors and Assumptions: Present in 1,061 Building Records in the Websearch

potentially significant errors, 458 buildings (15% of the population) were identified as having address errors, confusion over building location, or a very minor street range address identification note (where a building has a street no. range but is identified as just a single number within that range). It is unfortunate, that the coding for these diverse error types were not separated out since the street range address and Google Map placemark errors were often overcome by navigating down the StreetView 'road' or cross checking with internet searches. The notes column was used as a general dumping-ground of extra information so was extremely hard to analyse in more detail.

Assuming overall a +/-15% level of significant error in the Websearch data collection, this would affect the calculation of benchmarks in the Websearch database, but not necessarily translate to an equivalent sway in accuracy of the visualisation end-product. Benchmarks were applied to census typology data using cadastral building forms, resulting in an improved interpretation of volume and scale of the building stock, overcoming many of the errors affecting building size and height for example. Address errors accounted for 1.3% of the population, however this was combated with the manual (or automatic algorithm) colouring method used to assign benchmarks. The underlying valuation data exhibited an error rate of 4.7% (uncovered due to the initial validation work on the census). Due to the complex way these error rates interact it is impossible to quantify the exact level of error affecting the visualisation, but with improved use of open data, or the introduction of mandatory energy certificates / unique building identifiers (Section 2:02:08), accuracy could be increased significantly.

Analysis of typology mix across the commercial building stock was uncovered with pivot charts made of the various samples and census datasets. These provided an understanding of any bias present between the samples (Figures 25-32 Section 3.13), as well as conveying the composition of the stock which would be one of the proposed outputs of the visualisation, for instance, when querying certain zones or building types for a breakdown of the mix and distribution of resource use.

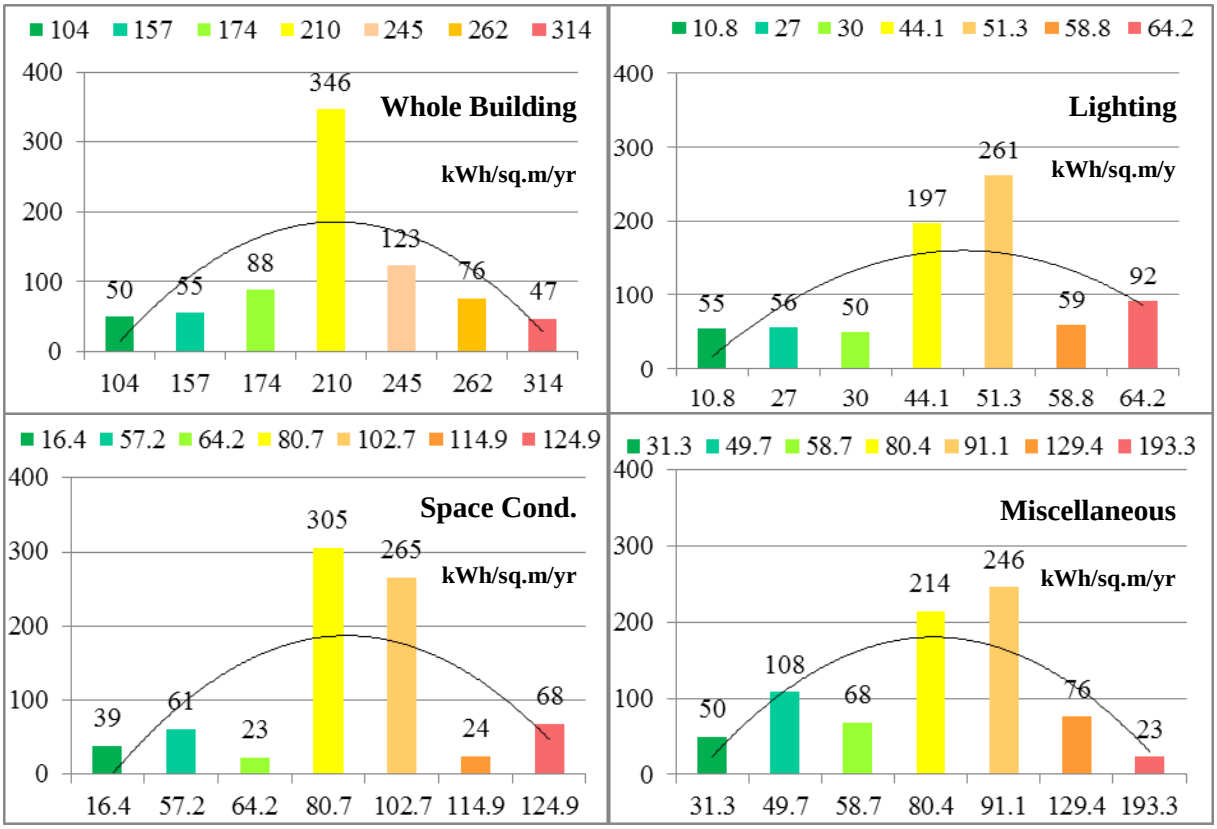


Figure 69 Wellington Census - Energy, Lighting, Space Conditioning and Miscellaneous

Further analysis was conducted on the mix of energy and end-use benchmarks as applied to the Wellington CBD census dataset (785 buildings). Figure 69 illustrates whole building and end-use energy benchmarks and their distribution in the Wellington CBD area (by building count), with red indicating very high energy intensity through to dark green indicating the lowest intensity buildings. The average EUI kWh/m^2/yr value calculated for each colour range is displayed on the horizontal axis in each graph.

Figure 70 2D GIS and 2D Extruded GIS Versions of the Pilot Visualisation

This information becomes mildly interesting from an energy distribution point of view when viewed in a 2D GIS heat map (Figure 70, left) and more intriguing but non-representative of building volume and form and consequently energy quantity, when displayed with extruded GIS where the Z-axis represents kWh/m^2/yr benchmarks (Figure 70 right). This extruded visualisation style focuses interest on the high benchmark buildings and the low benchmark buildings are practically invisible. This is partly a function of all buildings being visualised rather than sporadic data points, such as in the Carbon Visuals example (Figure 1, Section 1.01) since tall (red) buildings obscure short (dark green)

buildings. This focuses attention on the high energy intensity buildings and fails to communicate the volume of energy involved due to artificial building size. This also means the characteristics and surrounding features common with low energy intensity buildings may be completely overlooked.

5.04: Quality Management and Accuracy of the Data

The Websearch data collection timing pilot discussed in Section 3.07, provided a chance to amend the method, format and order of data collection. The screen and database layout was amended in consultation with the research assistants involved in this pilot. Particular attention was paid to layouts and systems employed by those with the highest quality and fastest data entry.

Figure 71 Websearch Workstation Screendump

Figure 71 indicates a typical working screen for the Websearch data collection with information gathered from Google Earth, Google Maps StreetView and other web searches, along with Microsoft Excel to enter the data points and Google Sketch-up to create the basic 3D model.

The accuracy of the floor area measurements was subject to the accuracy of the measurement tool in Google Earth. It was entirely dependent on the quality of the aerial photo measured. Area figures were generally assumed to be within +/-10%. Early on in the BEES research project a system of reporting on each individual building was developed by the author where a building record's aerial photo, elevation images and an extract of the important data collected would be summarised in a single page and used to confirm information or correct as necessary when conducting BEES site visits during the monitoring stage (Section 5.01, **Appendix iv:** 'Mail-merge Websearch Checklist'). Actual data recorded on site during the monitored surveys was not able to be compared with the Websearch results to establish the true level of error in the data collection, due to concerns this would raise the cost of the site visits, thus reducing the number of buildings that could be studied overall.

The accuracy of the floor and storey heights was also reliant on research assistant assessment. Building heights were gauged from StreetView images in relation to surrounding objects, passers-by and cars, along with a count of storeys using fenestration cues. Depth of field affected the accuracy of these estimates, however upper storey heights tend to be within a given range and whilst ground floor storey heights vary, they were clearly visible in StreetView imagery. Certain building materials, door and window heights signal scale in many circumstances, therefore the errors relating to building height are considered insignificant. The relationship between valuation (QV) and Websearch data is explored in Figures 33-35, Section 3.14, where Websearch floor areas, floor plates and storey heights are compared against QV valuation equivalents to assess the accuracy of the data sets involved.

The added data provided by including the overshadowing skylight factors for each façade visible in StreetView would allow energy use to be compared with the wider building environment. This takes into account the fact the lack of sunlight access the building is subjected to (or indeed an

understanding of the built-up nature of the site's location) which would be expected to exhibit a relationship with lighting energy end-uses as well as building useage. In Figure 36, Section 3.14, the average overshadowing ratio is compared with the storey height information illustrating that higher buildings tend to be found in areas with high overshadowing. This information could be provided by cadastral and building elevation data in future maintenance of this dataset and visualisation.

In the course of investigating and communicating the mix of typologies found in the Websearch data, in the form of coloured charts and histograms, outliers were highlighted and biases identified that were then able to be corrected. For instance, occurrences of low floor area buildings with a high number of storeys suggested an obvious error and these data points were investigated further by the author, gradually improving the quality of the dataset. Buildings found at incorrect addresses were individually investigated and for the most part these appeared to be correctly noted due to properties using an address 'round the corner' or with an entirely different street number range, however in two cases the buildings were incorrectly identified due to an error with Google Map links, so these were also revised with the correct details. One bias uncovered when reviewing buildings with aerial photo source data only was the lack of building height, elevation and surrounding data for 'battle axe' buildings (those down a long driveway, far away from public roads, or behind a row of buildings).

In particular a bias was found in the unfinished / incomplete Websearch data, caused by research assistants abandoning building records relating to large, complex buildings (or multiple buildings on one site). To remove this typology bias in the Websearch data the author tackled the 175 building records as quickly as possible to find key details that were missing from the Websearch dataset. Google Earth and StreetView images had improved since the records were initially abandoned and so basic data on 150 out of the 175 building records was added, partially removing this bias from the incomplete records. Floor area, height, glazing ratios, overshadowing and elevation information was provided for all 150 buildings. These buildings required further checking against council cadastral information to establish the most reliable building perimeters and parcel sizes possible.

Building fabric data entry was subject to certain misidentification especially between 'Concrete', 'Roughcast Render' and 'Fibre Cement' due to their similar appearance when viewed at distance. Examination of the average energy results returned against the various material options in the Websearch data reveal that materials have similar averages from different data sources (Figure 38, Section 3.16). Building records with conflicting materials noted in the valuation data when compared with the Websearch data were reviewed individually by the author and corrected accordingly. The inclusion of elevation image hyperlinks saved with the Websearch data made revisiting the data straight-forward Building material typologies were used in the allocation of benchmarks so it was important the material information was of the highest quality possible and with more experienced research assistants the error could be significantly reduced.

One of the major concerns with the accuracy of the creation of the benchmarks was the lack of StreetView access to battle-axe / rear properties (6%) and those records in areas without any StreetView or Google Earth access (11%). These may have swayed emphasis in the sample on buildings with a street frontage with a contributing effect on the creation of benchmarks.

Since the visualisation is to be automated for continual update and refinement, the errors encountered with this method of data collection could easily be overcome through open data sources. Floor area, elevation heights, overshadowing and site slope could be 'data-mined' from cadastral and elevation information found at the Koordinates website (Corkery & Coup, 2007)) rather than relying on partial

information in the Websearch. Materials, business names, building use, occupancy and ownership data could be sourced from a combination of valuation and business search data (such as the Who is Where? (E-ideas, 2013) service). Based on the Websearch experience, it is clear many additional data fields could be inferred from this base information and therefore the quality of data underlying the visualisation could be improved significantly. Energy bills surveys should be conducted every 5-10 years, preferably on the same buildings (if owners are willing to be included in such a survey again). The benchmarks could therefore be automatically re-assigned to the visualisation adding longitudinal data into the mix.

Another opportunity to improve the data used in the assignment of benchmarks would be to factor in more detailed information such as lighting, equipment and occupancy schedules already gathered from buildings during the BEES surveys but not available during the course of this thesis. Since equipment, occupant behaviour and attitudes change over the space of a decade, it would be important to revisit the monitored buildings to investigate changes in energy end-use and drivers of overall energy use within a 10 year period.

5.05: Extension of the data

A broad understanding of 'surface area' / 'building compactness ratio' was by comparing the floor plate multiplied by the number of stories, to the estimated floor area as provided by the research assistants' measurements (Section 3.10). Reduced floor area estimates were provided for setbacks and podium-type buildings and increased floor area estimates for buildings with visible evidence of sub-floor accommodation or where floor plate measurements were made of the 'tower' floor plate in a podium-building rather than the ground floor building perimeter. Assuming this gives an accurate picture of the building stock, 11% of built forms in the visualisation are over represented in volume by up to 55%. The 55% relates to a theoretical worst-case scenario of a 20 storey podium building with two 1,000m^2 footprint storeys and 18 storeys with floor area reduced to 50% of the ground footprint. The overall misinterpretation of volume in the visualisation is therefore estimated to be less than 6%.

The typology information that was available in both the valuation and the Websearch datasets provided the perfect opportunity to investigate the quality of the data. The typology mixes were compared using frequency tables with conditional formatting. These reveal similar patterns of typology mix regardless of the source of the data. Wherever possible the best of the results from the two sources were combined into single new data columns to provide the most continuous data set for analysis alongside the energy results. Materials, Floor Area, Storey Height and Floor plate from both valuation and Websearch sources were also compared against each other using energy results to test whether they result in similar average energy benchmarks, whether from the valuation or the Websearch datasets (Figure 38, Section 3.16.)

Building ages were entered for individual years in the Auckland valuation data, but only by decade in the Quotable Value database for the rest of the country. Building ages were therefore grouped to the lowest common denominator in a new field 'Building Decade'. The older buildings had been handled differently in the two valuation sources (before 1920 grouped together in New Zealand records, with individual years noted in the Auckland records) therefore the 'Building Decade' field grouped together buildings built prior to 1920 as one age type. Once energy results were compared with 'Building Decade' data the sample sizes for buildings prior to 1950 were too small, therefore building typologies using age required further consolidation to be meaningful. The manifestation of this in the visualisation would be to limit the user's ability to isolate certain building typologies when compared

with 'Building Decade', just as meshblock level delivered energy data may require limiting to Area-block when meshblocks contain too few buildings thus revealing sensitive data (Section 3.15).

The 'WhoisWhere?' (E-ideas, 2013) dataset provided a detailed understanding of business density, building use, occupant density and valuable building use information which, when combined with searchable terms in the valuation 'detailed use category' and 'improvement description', provided a detailed mix of uses in each building at census, Websearch and sub-sample level. This field was named 'New Use Type' and this information formed the basis for the end-use benchmark and hybrid typology allocation as well as the analysis of bias between various sample sets (Figures 25-32 Section 3.13). This business data provide a far higher quality and much more detailed data source than the valuation category of use information on which the BEES sample selection was based.

CBD office quality grading (Property Council of NZ, 2014) was added to the data collection part way through the Websearch survey, due to a research funding stakeholder request to allow comparison of office quality grading with energy use. Office quality data was sporadically entered with only 55% of records identified as CBD offices through 'use' codes having an office quality noted. Since the sample sizes with energy bills and office quality data available are low, it is difficult to make a clear connection between average energy benchmarks listed in Figure 72, except as a hierarchical comparison. The average energy intensity figure is 210 kWh/m^2/yr for all types of commercial building use (or 204 kWh/m^2/yr taking into account the two outliers that were removed), but when all the Websearch results with 'office quality' and energy bills data are measured for their average energy intensity the result is 157 kWh/m^2/yr based on a sample of 64 records, which is significantly lower than the overall average for the building stock. Office quality C has a very high average energy intensity benchmark (266 kWh/m^2/yr), however the sub-sample with energy bills results available is only four-strong, therefore the results are subject to a high standard error rate. Only the results for 'all office quality-graded buildings' and the D-grade provide a sample size large enough to be of use in any energy intensity analysis, but a trend is visible in the B and D results. With the availability of more thorough and reliable Property Council data on Office Quality Grading, it might be possible to include office-quality-derived energy benchmarks in future maintenance of benchmarks and update of the visualisation providing a much higher quality underlying data set.

Grade	Office Quality Grade Description	Energy Benchmark Average (kWh/m^2/yr)	Sample Sizes
A.	Landmark/ CBD / Naturally lit / Views/ Quality Entry/ UndercoverPk	141	13
B.	High Quality Space / Views / Quality Entry / Undercover Parking	125	19
C.	Good Quality Space / Reasonable Finish / Car Park Available	266	4
D.	Office Space: Lower / Poor Quality Finish	170	28
A-D	**All records with Office Quality Noted**	**157**	**64**

Figure 72 Office Quality Grade A-D with Average Energy Results

Data available in the Websearch has also been used in the Net Zero Energy Potential and energy simulation modelling work carried out at Victoria University Wellington for the BEES project. Using typology and characteristic data contained in the Websearch dataset, combined with HVAC, lighting and appliance schedules from the targeted survey monitoring work, it has been demonstrated that it is possible to calculate building energy use to +/- 5% accuracy of measured energy consumption (Cory, 2015). This data could be used in future updates of the visualisation to deliver the latest research available to the public in a heat-map coloured 3D format most users will understand. The list of EUI

benchmark averages derived from the energy bill survey (and used to colour up the pilot visualisation) are compared with 48 model simulation results in Section 3.16 and **Appendix iii:** 'Criteria Benchmark Averages'. The small sample sizes in the simulation models, plus the fact that gas energy has not been included, results in lower EUI averages, however where sample sizes are above 8 models, the results approximately align with the energy bill EUIs used in the pilot visualisation.

5.06: Energy and End-Use Data Combined with Websearch

Section 3.16 gives an overview of the way premises were combined into 'buildings' so that the energy bills results could be compared with Websearch and Population datasets. Energy benchmark averages were assigned to census data in seven colour range groups each allocated an average overall EUI value. This EUI value for each colour was further broken down according end-use percentages from the BEES high and low EUI case study buildings (Bishop, 2012a) (Figure 73) and the ECCA End-use Database EECA (Energy Efficiency Conservation Authority, 2007) (Figure 74).

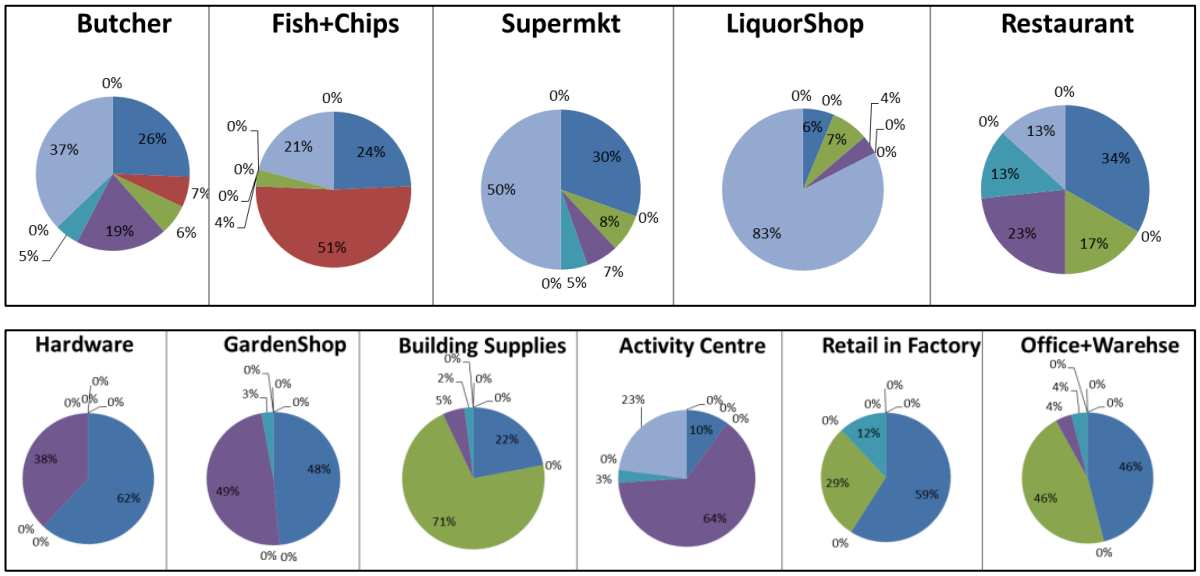


Figure 73 BEES High and Low EUI End-use Mixes

New Use Type categories had been entered automatically using data filters in the Wellington and population census datasets and this information was used to manipulate the whole building benchmark values already assigned for each colour. The original category of use information available in the valuation data was of very poor quality, and without the work done to provide detailed 'New Use Types' with the help of 'Who is Where?' data, the End-use benchmarks would have been impossible to apply with any confidence.

The EECA Energy-use Database came from a wide range of sources including EECA end-use surveys, census data, council databases, Energy Data File 2006/2007 and census data from Statistics New Zealand. End-use data is likely to change with the availability of more efficient equipment and alterations to the number of devices typical in certain building types. The percentage fields used to assign end-uses to the visualisation can be updated automatically with more accurate data from the BEES monitored buildings survey (lighting schedules, equipment schedules and floor area information) when performing maintenance on the underlying data. Fully comprehending the relationship between end-use breakdowns and building use will be an important step towards

accuracy in the visualisation, since the breakdowns rely on understanding what people do within buildings, not building typology in isolation.

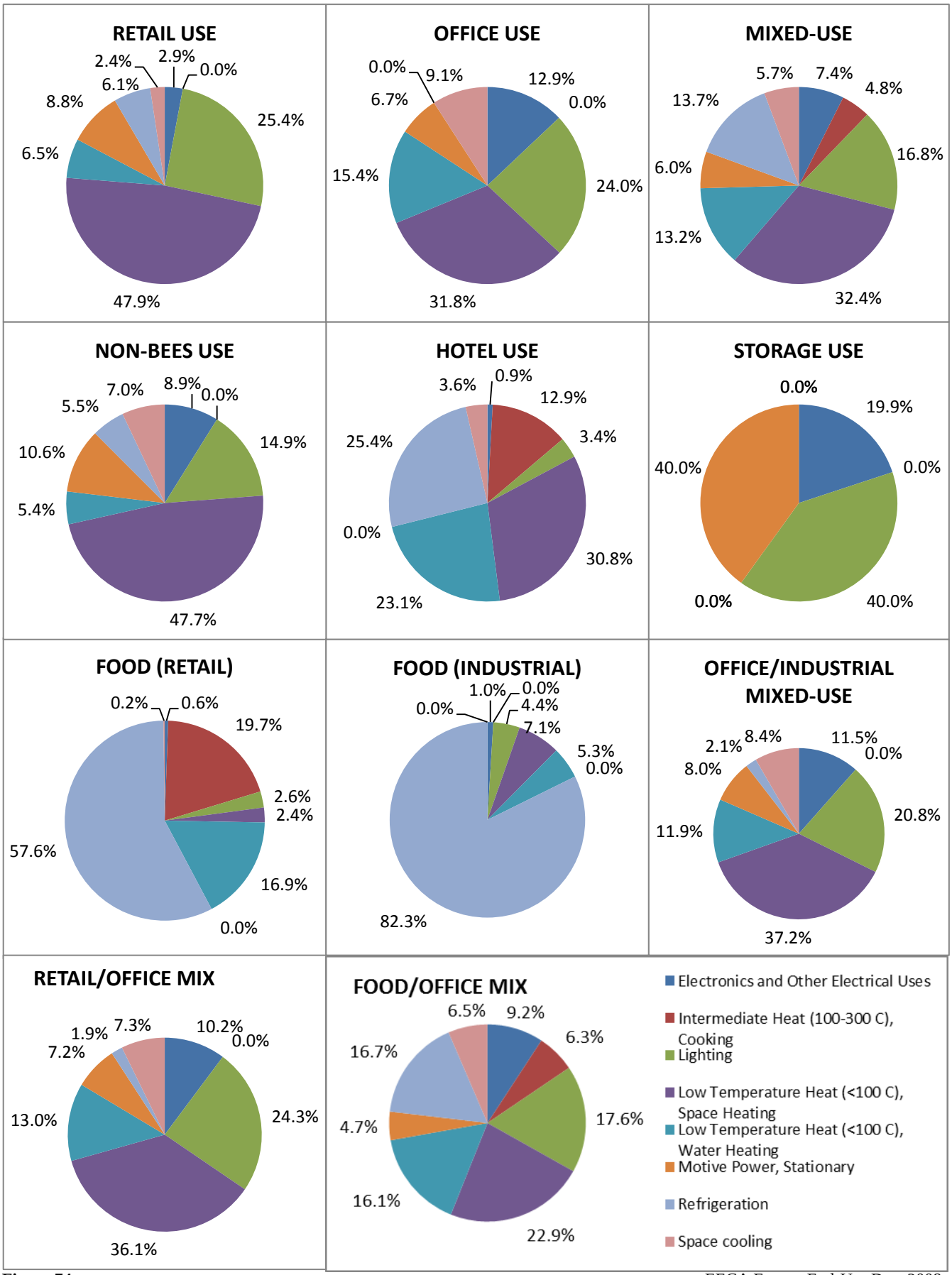


Figure 74 EECA Energy End-Use Data 2008

5.07: Overcoming Data Privacy Issues

Building typology and characteristics data gathered during the BEES project are to be made freely available; indeed this is the major aim of this research. This data will be combined with other publicly available statistics about buildings and the results of this work will be communicated through an open-source platform such as Google Earth™ or Google Maps StreetView™ encouraging the widest possible public access to the data. The sample of buildings studied in-depth around New Zealand during the BEES project will not be specifically identified or located to preserve confidentiality. Confidential energy bills data will be averaged according to building typology and typical end-use breakdowns, resulting in a set of resource-use benchmarks and typical end-uses for *all* buildings which can be made publicly available, keeping the confidential energy data pertaining to individual buildings out of the public realm.

The actual energy and water data for buildings may be submitted into the 3D visualisation by a building owner and this might need to be spot checked by a central administration service to confirm accuracy (Section 4:02:04). Obviously by submitting data to a publicly available visualisation the owner would be aware of the privacy issues but these would need to be clearly set out in some kind of disclaimer.

The valuation source data was originally purchased from the Quotable Value database in 2008 and this typology information allows extrapolation of details about the random sample to the census of non-residential buildings. Quotable Value data is publicly available, but it is also commercially sensitive since an organisation charges for the release of data. None of this purchased data will be directly communicated through the 3D model as only publicly available data collected during the Websearch survey will be identified or aggregate valuation data for, say, a meshblock area.

The benchmark figures quoted will be within ranges normally applicable to buildings according to use-type, form, size, height, building age, glazing ratios, and construction materials, as well as any other characteristics deemed to be relevant to energy / water use. The only data communicated in the visualisation that was originally sourced direct from the valuation data is building age (‘building decade’) however this information is available freely already through the QV database, but all other information contained in the valuation record requires payment to Property IQ for its use.

Visualisation of the mix of building typologies (for instance through property queries Section 4:02:01) is reliant on valuation data for its baseline information, however the energy benchmarks resulting from its analysis and applied to the population would not compromise Property IQ’s business practices. Property IQ and Auckland Council (as owners of the valuation data) would have to be consulted on the visualisation when made public to ensure that they are satisfied with the use of their data in such a format. With or without valuation data used as a baseline, it will be important to refresh the energy bills and end-use percentages at least every 10 years (in 2018 based on the date of valuation data used in the BEES project). Given there is a movement towards open data for all information (not compromising privacy) that has been paid for using public money, this would suggest that valuation data may be made public in the near future, vastly simplifying the issues raised regarding data privacy.

5.08: From Database to Visualisation

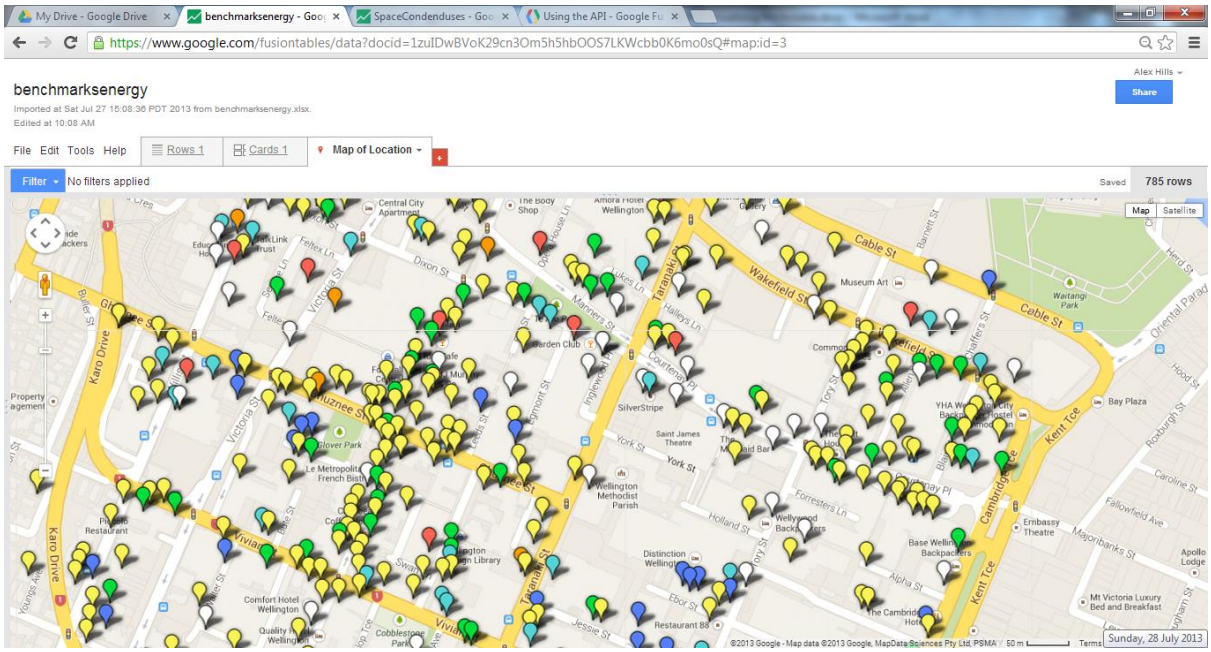


Figure 75 **Energy Average Benchmark EUI Map-markers Linked to Address Point Data**

This thesis proposes a set of logic instructions to automate the link between the database and the 3D model elements (Section 3.22). Google Fusion Tables (Google Apps, 2013a) provide a location for each data point using a geo-coding tool and during the process 1% of building addresses were found to be ambiguous (Figure 75). This verifies the discoveries made in the initial validation work where 2% of records were found to have incorrect address information, however this was reduced to 0.8% following the corrections made to the dataset (Figure 14 Section 3.03). Later analysis of the Websearch data entry (Figure 19 Section 3.09) indicated 1.3% of building records with an address error however many of these 'errors' were the result of an address simply being around the corner from the official postal address. The act of colouring the buildings (manually, or using the algorithm) rectifies the majority of these ambiguous addresses. Only those addresses without a street number or with an incorrect road name will remain ambiguous.

5.09: Sources of the Data

A detailed overview of the sources and method employed to collect each column field in the Websearch data, as well as inferred data gathered for the purposes of this thesis, is described in **Appendix i:** 'Websearch Data Sources'. Valuation data was purchased from Quotable Value (QV, 2008) and the Auckland Valuation database (Auckland Council, 2008). All hereditaments with a 'parent' 'orphan' or 'child' valuation with Commercial Office / Mixed / Retail (including Tourist, Vacant, Motor, Service Station and Liquor) or Industrial Service / Warehouse uses were purchased along with the other valuation records present in the building. Records were amalgamated into buildings to form a census dataset of 50,548 building records, (Section 3.03 details the process used).

The Websearch data collection relied heavily on the availability of Google StreetView™ omnidirectional imagery and Google Earth aerial photographs. 80% of the Websearch sample had both StreetView™ and aerial photos, and therefore represents the highest quality results, however 9% of the sample had at least one of these sources available, meaning only 11% of the data suffered from the absence of both plan and elevation data. The Websearch survey constituted a one-off check of

the variability of the publicly available data to enable a connection to be made between energy bills and typology data available in the population. If this visualisation is to succeed it will need to be an automated process using knowledge uncovered in the course of the Websearch survey.

The automation and maintenance of this visualisation requires:

- revised (or existing) benchmarks as confirmed with refreshed energy bills data (updated every 10 years – 2018+)
- revised end-use breakdown percentages from BEES assigned to each detailed use type (updated every 10 years 2018+)
- collating the cadastral and building elevation data into a 3D virtual model of the stock, updated automatically with any new release of open cadastral data (Section 3.21)
- new population valuation data to assign geo-located HTML colour code map markers (Section 3.19)
- the use of an algorithm (Section 3.22) to colour polygons in the 3D virtual model according to map markers in the vicinity.

Limitations in the accuracy of the 3D virtual model used to illustrate the energy graphics are related to the cadastral information publicaly available. Footprint polygons, extruded to the known building elevation height, result in many building forms having a high surface area due to various set-backs, or because they are podium buildings and therefore would over-represent the amount of energy consumed due to building volume being over estimated in the visualisation. A possible solution to overcome this error is by running a volume m^3 test (as described in Section 3.22 in the algorithm steps), thus highlighting the outliers caused by an over-estimate of volume or perhaps by mis-identification of the building in the course of the automation of the process.

Focus groups or user-surveys need to be conducted to assess the success of the visualisation graphics and their ease of use and comprehension, before a true understanding of the implementation issues can be reached (Section 7.02).

6.00: Conclusion

A major goal of this thesis was to derive a new method for presenting statistical data on the built environment, theoretically making it accessible to the widest possible audience (Sections 1.00-1.02). This concept is demonstrated in Figure 76 where the same statistical data (Wellington CBD energy and lighting end-use benchmarks) are expressed as histograms, 2D coloured 'heat-maps', extruded GIS and finally in 3D renderings – of built-form and city grain with colours revealing information about the built environment. Unlike the 'heat-map' and extruded 2D versions, the 3D city format allows the user to understand, question, interpret and visualise statistical information in many ways including some not necessarily anticipated by the author. This 3D format has a far higher chance of revealing relationships, patterns and trends that would remain invisible otherwise – particularly to lay persons (Section 1.01). The true context, grain and legibility of the city is maintained alongside the statistical information.

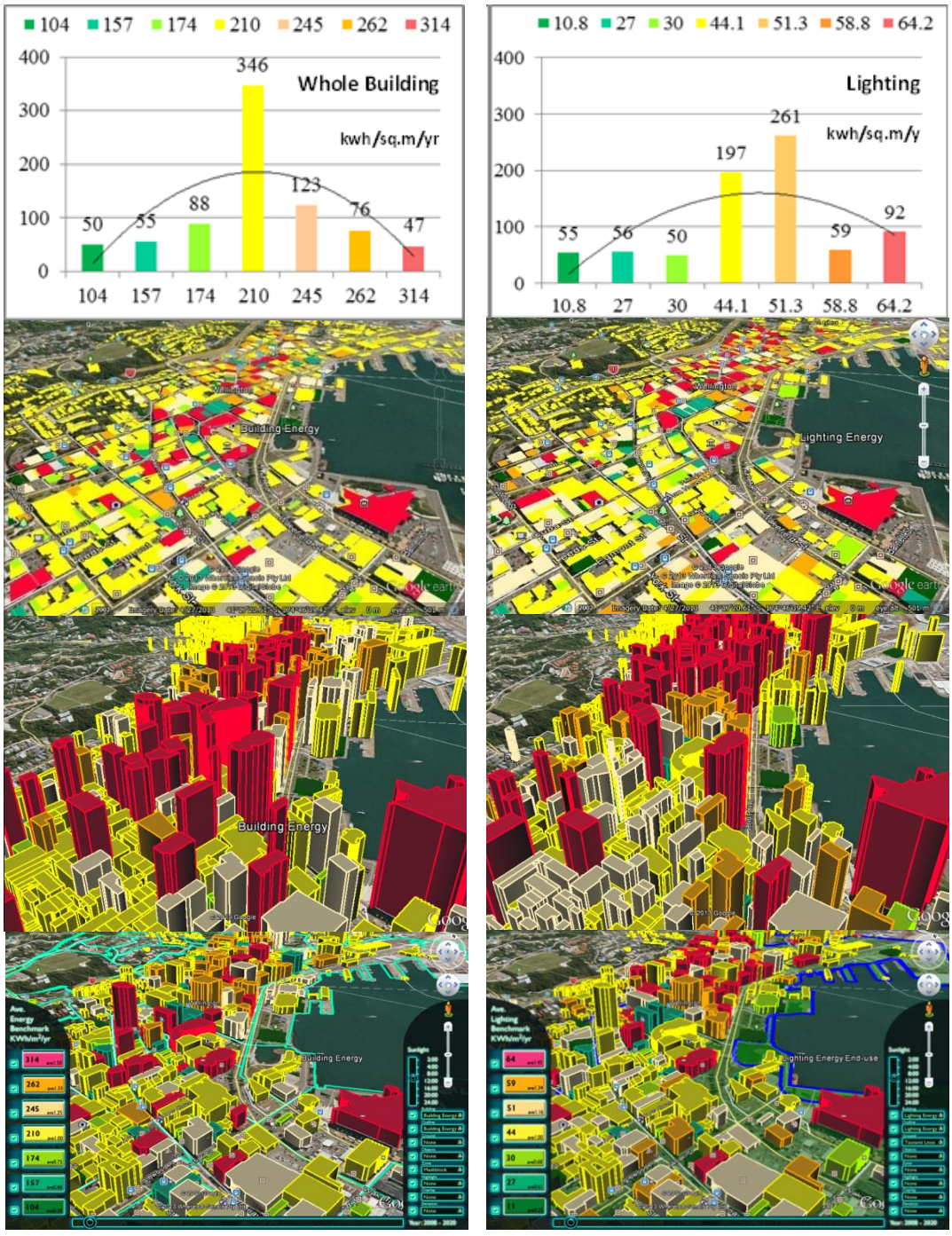


Figure 76 Wellington CBD Energy Benchmarks (Histograms, 2D Map, 3D Built Forms)

Observing the 3D visualisation results for whole building energy versus lighting end-uses successfully illustrates the major goal of this thesis (Section 3.23 and Figures 76-77). Buildings with the highest energy intensity are centred in the middle of town, but when the visualisation layer is switched to lighting end-use energy the highest energy areas gravitate towards the shopping district probably due to the concentration of retail.

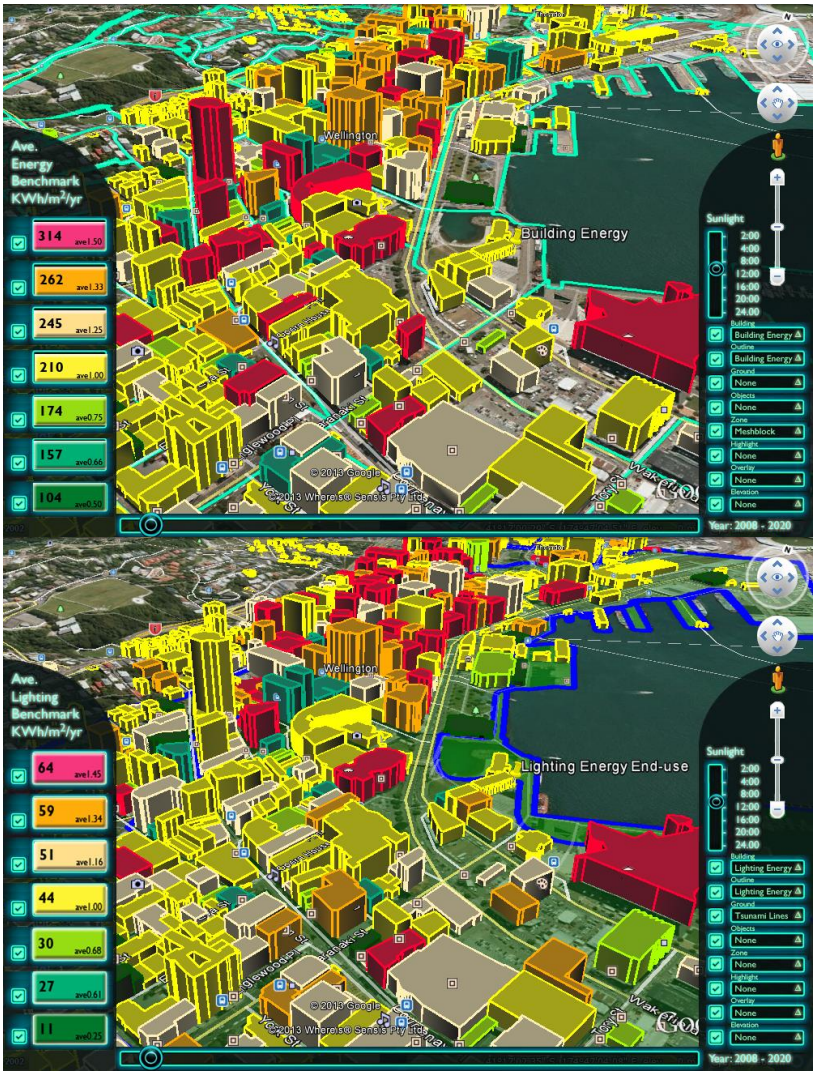


Figure 77 Energy + Meshblock (top) Lighting End-use + Tsunami Lines (bottom)

When switched to space conditioning the results become more sporadic with high areas focused on hotels, food halls and large complex surface-area office buildings (Figure 78) Finally the end-uses relating to miscellaneous energy, including refrigeration, cooking and other electrical appliances, result in few high and low outliers for the majority of buildings in the Wellington CBD when compared with building stock throughout the country. These visualisations raise many questions about the drivers of energy use and, at the very least, suggest areas where further research is required. When combined with the added functionality of identifying or isolating buildings with certain characteristics (Sections 4.01.06 and 4.02.01) this allows any area of interest to be investigated in further detail.

Since the quantity of energy consumed is implied by the volume of the 3D building form (and colour indicating energy intensity) the impact of floor area on energy use can be immediately focused to areas with the most efficiency potential. The visualisation dashboard has been designed to allow quick toggling between contrasting statistical results and the ability to display two result-types at the same time by comparing the building outline and fill colours (Sections 4.01.01-02, Figure 52-53). Lastly the visualisation proposes the display of these resource-use results in combination with council data (earthquake zones, heritage, district plan, potential building heights) and environmental simulation results such as solar gains, noise or pollution modelling. The latter are expressed in 3D using the building block facades as an indicator of volumetric statistical data. This quick switching between various datasets in a 3D visualisation draws the viewer's eye to statistics that change dramatically, and relationships between those statistics in an area of interest (such as low or high outliers). The added value of presenting these in three dimensions true to city grain and built form, is that the

location in question becomes more legible and information about building typology, overshadowing and surrounding characteristics becomes apparent whilst digesting the statistics of interest. The 2D GIS and extruded GIS completely distort the form of the city (Figures 76 and 78) and focus interest on buildings which may be small and insignificant in the context of the whole city. They are completely misleading as to the actual quantity of energy they are attempting to convey, as well as any relationship to the physical building to which the data refers. If a visualisation makes you aware of the physical building form, use, and typology in relation to its energy intensity, then energy efficiency opportunities begin to reveal themselves in the data and research efforts and policy change can be directed to the most urgent and effective areas (Sections 1.02-103).

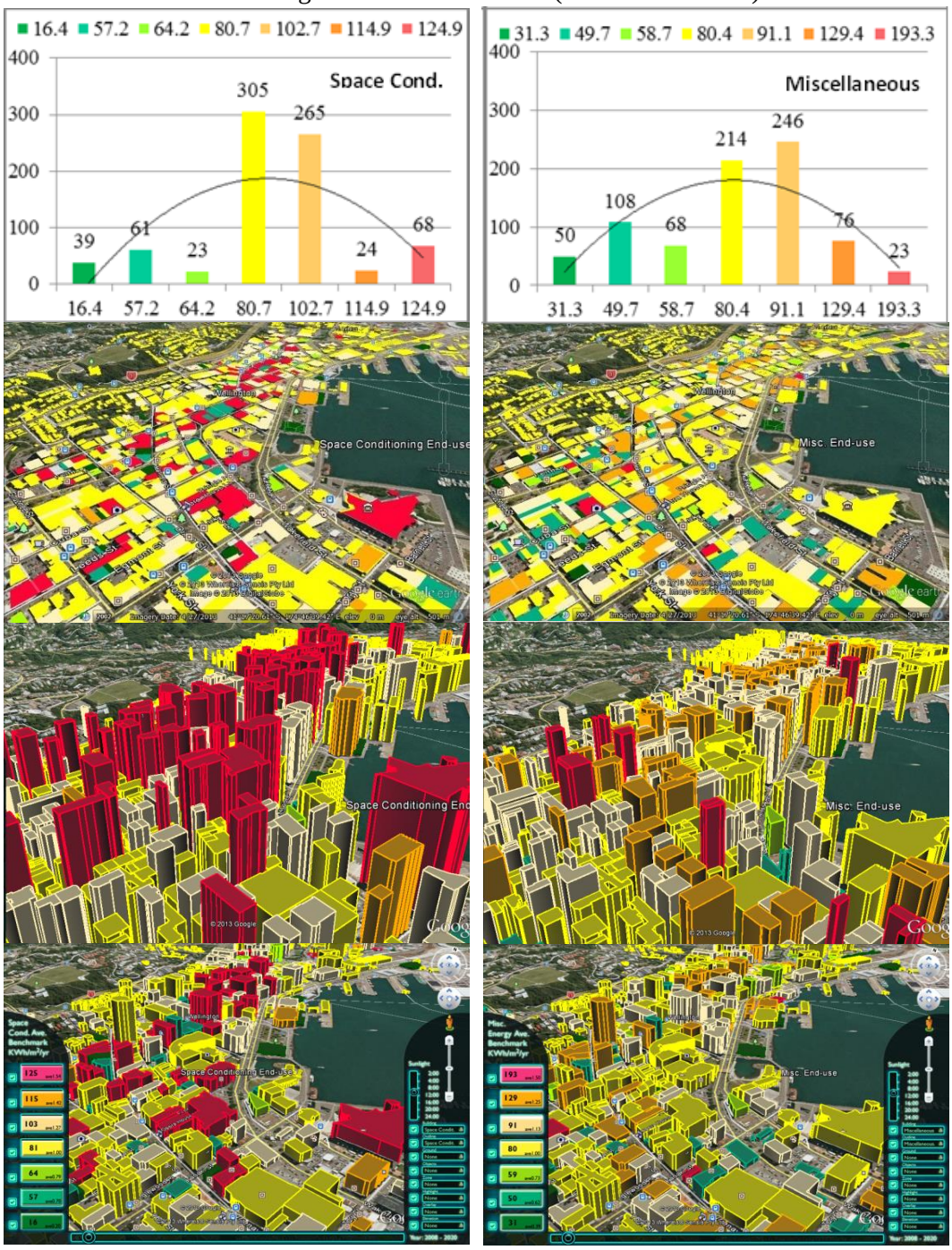


Figure 78 Wellington Space Conditioning versus Miscellaneous End-use Energy

This thesis brings together diverse statistical data on the commercial building stock, assembling this information through a three dimensional visualisation, whilst maintaining site context, built form and street grain. It demonstrates that a fully-automated, open-source visualisation is possible and has the potential to make invisible building resource-use-data visible to a wide audience (Section 1.01 and 1.04). Throughout the process, the background building stock typology data informing the visualisation has been laboriously tested and quality assured (Sections 3.13 and 5.04). The Websearch and visualisation output represents an unprecedented snapshot of the commercial building stock in New Zealand in 2008 (Section 3.13). The data should continue to raise questions and inform research about the built environment, and the methodology employed (Chapter 3) has the potential to be duplicated in other countries wishing to conduct building stock surveys. The visualisation style proposed will improve accessibility to information that would otherwise remain invisible and unintelligible.

6.01: Accuracy of the Visualisations and Baseline Data

Some issues with this method have been identified regarding the accuracy of physical building volumes in the 3D model, their link to address point data ((LINZ, 2011b) Section 2.02.07) and the quality of the samples used in the creation and assignment of benchmark averages.

Building volumes would be inaccurately represented in approximately 11% of the building stock in the visualisation due to the polygon footprint being extruded to roof elevation height with no account taken of podium buildings (1%), those with major set-backs (2%) and high surface area (8%). The worst-case scenario for overrepresentation of volume in the visualisation (based on a 20 storey podium building) would be an increase in total floor area of around 30% bringing the overall error down to less than 3%. Building volume was corrected manually in the pilot visualisation, however a long-term solution to this problem would be to automate the construction of the model using modelling techniques such as LiDAR and photogrammetry, as discussed in Sections 1.04-1.05. This solution presents its own challenges, as dissecting the detailed 3D topography, differentiating built-forms from their surroundings and linking these 'model chunks' to building records through address data is highly complex and worthy of being a thesis topic on its own.

When conducting quality research on the building stock it is necessary to link valuation datasets to physical built-forms and to visualise this data successfully it is necessary to link the valuation data to digital built forms (Section 5.08). Address fields (LINZ, 2011b) are prone to syntax errors as the postal address may be 'around the corner' from the building's main entrance and the complex variety of relationships between a valuation (hereditament) and a building make linking datasets fraught with difficulty. An attempt to correct the address data field in the valuation data was made during initial validation work on the census reducing the 2% error rate down to 0.8%. Despite this initial work on the population prior to sampling, the Websearch data collection revealed 1.3% of building records still had incorrect addresses listed. When subsequently geo-locating the resource-use data for the pilot visualisation using Google Drive 'Fusion Tables', 1% of buildings in Wellington CBD were found to have ambiguous addresses, confirming the accuracy of the Websearch findings (the reduction of error rate may be due to the focus of the pilot on an urban area). This resulted in a benchmark map-marker being placed in the centre of a road rather than at the mailbox location or clearly linked to a building parcel. Since the visualisation colouring was achieved manually in the pilot visualisation, it was possible to interpret the link between building and map-marker improving the accuracy of benchmark assignment. This would be an equal challenge for an automated algorithm (Section 3.22), however its ability to quickly distinguish buildings with volume or identification issues, would make correction of

these outliers possible. A better long-term solution would be the introduction of a unique building identifier system as discussed in Sections 3.22, 5.03 and 6.02.

In addition to the building identification (1-2% error) and 3D form issues (3.3% error), the creation and assignment of benchmarks were estimated to be subject to an error rate of around 15% in the Websearch dataset (Section 5.03). Due to the complex way these census and sub-sample error rates interact the percentages cannot simply be added together, but together they give a reasonable sense of the reliability of the baseline data. The successful integration of LiDAR technology and unique building identifiers would probably resolve the benchmark accuracy in the process. The automated nature of the visualisation means this data could be updated and communicated to the public as desired.

6.02: Open Data and the 'Unique Building Identifier'

The global movement towards open data has made more public sector information openly available than ever before. Valuation data has the potential for future use in research on the building stock. By linking valuation building records to other datasets it will be possible to add further sophistication to this research, as well as improve the accuracy of the base valuation dataset.

Valuation data must be vastly improved to identify and accurately geo-locate building form. This is a crucial step required to enable the highest quality research to be conducted on the building stock and make the best possible use of open data. A unique building identifier (a building-equivalent to a computer's 'I.P address') should be used as a common identifier, linking together information such as location, retrofit history, floor area, storey height, heritage status, earthquake status, energy usage and water use. In practice this may be assigned and linked to valuation data when carrying out a Building Warrant of Fitness annually as required by the 2004 Building Act (New Zealand Legislation, 2004). This would not only improve and verify the quality of the valuation data underpinning New Zealand's taxation system, but also enable valuable mash-ups of separate datasets to occur on a national or local basis (M. R. Donn, 2004) (Stanford Business School et al., 2012). This has the potential to remove two critical barriers to the implementation of energy policies in the built environment: firstly, energy policy can be directed to the biggest 'resource-guzzlers' if detailed building typology can be tied accurately to resource consumption. Secondly, by providing a enticing visual platform for building owners to advertise their efficiency accomplishments publicly or compare their energy use against national benchmark levels then it may be possible to raise the perceived value of sustainable developments (Section 1.01).

In the absence of a unique building identifier being available during future maintenance of the visualisation, the pilot visualisation methodology could be duplicated in combination with two modifications. These are an improved classification tree and multiple regression statistical analysis for the creation of benchmarks, and the use of an automated algorithm for colouring and highlighting suspect building forms for further manual checks. This will ensure the data is maintained as accurately as possible communicating a hierarchy of resource-use results, even if the average values used themselves remain within an error band of +/-15%.

6.03: Conducting a Building Stock Energy Visualisation Internationally

When conducting an energy survey of the building stock elsewhere, a census of building records must first be gathered from around the country. The most useful source for this type of high level data tends to be taxation databases, but these can be full of errors as they are collected by various regions / councils and by staff with varying abilities and different interpretations on collecting data. In

addition these records usually relate to 'hereditaments' or legal descriptions instead of 'buildings' (Sections 3.02 and 5.09). Once the most trustworthy database has been sourced a brief check must be carried out on the valuation data (Section 3.03). Quality checks can identify outliers / statistical anomalies and wherever possible corrections can be made prior to the formation of the random sample of buildings (Section 3.04).

In order to gather reliable statistics on building characteristics / typologies it was previously necessary to send out teams of researchers on site visits of a random selection of buildings around the country. This is an expensive form of data collection. It relies on researchers who have reasonable knowledge about buildings / construction and may involve risks to their health and safety. With the introduction of Google Earth™ and Google Street View™ the author has demonstrated it is possible for researchers to conduct building surveys from their desk (Section 3.06). This not only reduces health and safety fears for research staff, but has been achieved at a comparatively low cost of around NZ$15 for every building record investigated. This is a highly convenient method of data collection however there are issues regarding the completeness of the data, due to the misrepresentation of battle-axe (rear) properties, buildings situated away from the street-front or on private roads. The level of Google Earth and Google Maps StreetView coverage available will also vary significantly subject to location.

Resource-use data would need to be sourced from a sub-set of this random sample and the typologies' data compared to the building resource-use bills to generate benchmark averages. If end-use data is also required, a random sample of buildings must be surveyed for occupancy and equipment schedules and preferably resource-use monitored to confirm the accuracy of the bills data. Given the way the use and contents of buildings change over time, it would be important to refresh the data at maximum every 10 years.

6.04: Gaps in the Data

The BEES data (N. Isaacs et al., 2009) includes characteristic information on all commercial office, mixed and retail buildings along with industrial service / warehouses throughout New Zealand. There is also the potential for residential data to be included thanks to information from the earlier BRANZ study known as HEEP (Housing Energy End-use Project) (N. Isaacs et al., 2005). These two datasets cover the majority of NZ buildings, however there are gaps in the data including heavy industrial properties, agricultural buildings, schools, apartment buildings, hotels, prisons, hospitals and other residential care buildings.

As the valuation data has not been grouped into buildings for records outside of the BEES and HEEP studies it is not known how many special use buildings are excluded. These would have to be excluded in the visualisation (currently displayed as grey buildings in the pilot) unless further studies are to be conducted that cover specialised building uses, thus filling in the blanks. In the case of heavy industrial buildings the building energy and water use is overshadowed by resource use from industrial processes, so separating out the building related energy from the process related energy would present its own issues. Other more specialised buildings might be encompassed through case studies conducted by the various sectors involved (for instance schools monitored by the Ministry of Education). Hospitals, school and government buildings could be added to this data set by decree and the open data movement might result in making this data available for inclusion in the visualisation.

Ultimately, if all buildings in New Zealand had a benchmark resource-use allocated to them it would not only make the visualisation a more powerful tool, but it would be possible to compare 'top-down'

data such as meshblock level delivered energy and relate this to the mix of building types and building floor areas found within each meshblock. Whilst all of the data revealed in the visualisation is sourced from open data, special care would have to be taken to ensure that sensitive energy or water use results are not revealed through specific queries made on meshblock or building-level data. The pilot visualisation illustrates the success of communicating the commercial building benchmarks in a dense inner-city example. Whether these buildings alone warrant investigation in rural areas without the inclusion of all residential and 'special' buildings alongside them remains to be seen.

6.05: Publicly Available Data Versus Confidentiality Agreements

As discussed in Section 5.07, there are issues that surround the identification of BEES buildings, as well as the individual energy and water use information supplied by the BEES survey work, which was supplied in good faith with confidentiality agreements signed in the process. The visualisation proposed by this thesis has been designed so that it does not reveal sensitive or confidential data.

To get around the issue of acceptability of the data to certain industry groups (real estate agents and business owners that occupy or own low-rated building types) a tool is proposed that will allow owners to update benchmarks with 'actual' energy and water use data for their building. They will have the power to compare themselves to publicly available information in the form of benchmarks. This may spur them into action or alert them to a technical problem with their building, just by highlighting where their building sits within the average energy benchmarks for its typology. In addition, if a rival business posts a better-than-benchmark performance, this may encourage positive competition between building owners. If building prestige is eventually linked to lower resource-use intensity this may influence tenants' decisions and drive the market towards sustainability.

6.06: Basing the Visualisation on High Quality Datasets

A great deal of effort has been taken to make the database underlying the visualisation as accurate as possible within the budget confines of the BEES study. The literature review of equivalent energy studies around the world presents several methodologies used to gather high level data to be combined with measured energy results. The BEES study took advantage of new technologies such as Google Earth and Google Maps StreetView and other open data sources to build a comprehensive database of building typologies. These online data sources are 'blunt' tools which result in a rougher level of accuracy overall, however the sample size able to be investigated in the BEES study covers a far higher proportion of the national commercial building stock (6%) than equivalent international studies (Sections 2.01.01-2.01.05). For instance, the UK NDBS survey relied on street surveys of specially selected wedges of four UK towns and was based on a random sample representing just 0.2% of the total commercial building stock, and the Canadian version represented just 0.4% of the commercial building stock using multiple data sources to generate the sample frame. In the UK survey measuring sticks were used in the assessment of building heights and sizes, and most elevations would be accessible on foot (as opposed to those sourced through Google StreetView in the Websearch). This provided a more thorough database of information in the NDBS survey with higher reliability albeit on a smaller sample selection overall (Steadman, Bruhns, & Rickaby, 2000). The Canadian survey CIBEUS (Government of Canada, 2003) included detailed resource-use monitoring and building classification of the target population. This included all buildings with an area over $100m^2$, of which 50% or more are devoted to commercial or institutional activities, however the survey was restricted to dense urban areas and would be subject to survey participation rates (which have not been explicitly reported). This suggests that although the high level data collected in international studies through physical street surveys may be more accurate than is possible through online data sources, the relative size of the random sample surveyed may suggest the interpretation of results are prone to

error and duplication. We have performed quality checks of the Websearch data for accuracy, coverage, user-entry bias and sampling bias and where a problem was identified adjustments were made accordingly.

The automated nature of this visualisation provides a platform for incorporating improved data sets as they become available (either through the final output of the BEES study or through future efforts to model the building stock).

Improvements could certainly be made to the data collection method employed if this Websearch survey work was duplicated in another country. For instance a bias was found in the data where research assistants tended to over-estimate glazing area especially where there was a high glazing ratio, which meant the author spent considerable time correcting the data set. In reality providing glazing ratios to the nearest 10% wasted valuable research time when in fact all that was required was a basic differentiation between no glazing, very low, medium and high glazing ratio. This could apply equally to physical site visit surveys as it does to online data collection.

The built form codes adopted in the Websearch survey (Figure 79) were based on UK's NDBS street survey methodology (Steadman, Bruhns, Holtier, et al., 2000). The UK built form codes (which also included an additional railway arch not typical for New Zealand) inspired the BEES team to consider the recording of form and daylight penetration in high level data collection. However, since storey heights and building floor plate had already been collected in the Websearch survey process, the main need for a built form code revolved around the difference between open plan and cellular space and various combinations assuming these were discernible from StreetView imagery. The establishment of six complex form codes may therefore have been unnecessary. Instead a simpler way of collecting equivalent information would be to record the mixture of open plan versus cellular space.

In future, maintenance of the visualisation, open data available in New Zealand at Koordinates (Corkery & Coup, 2007) on building elevation (roof height above sea level) and cadastral footprint

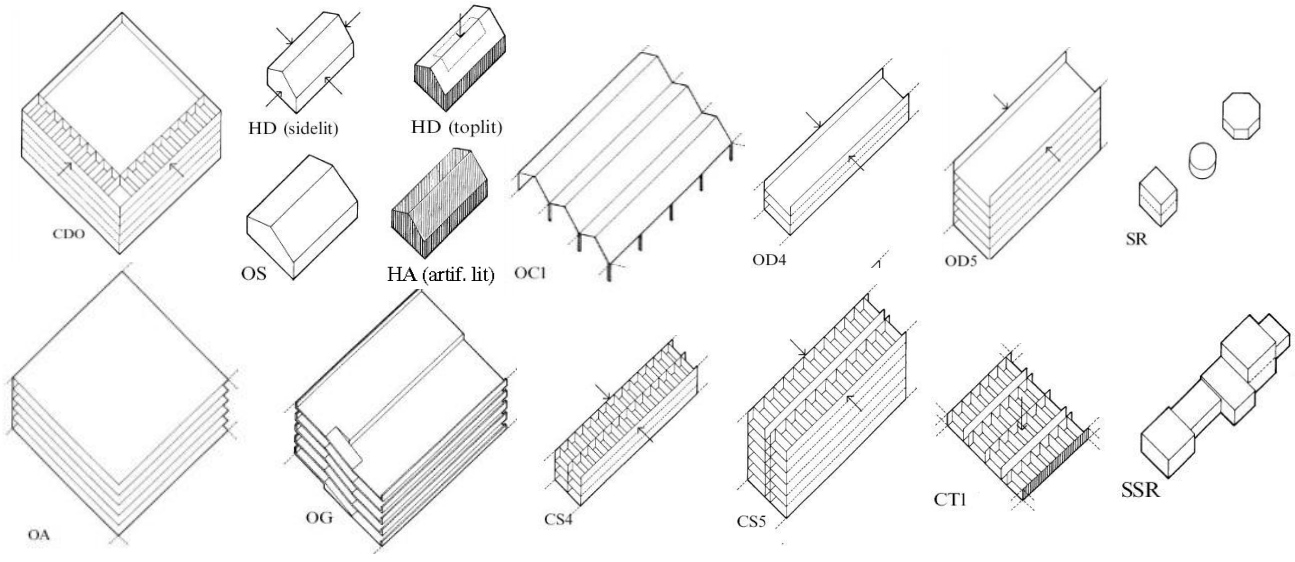


Figure 79 **Steadman Building Form Classification**

polygons, has the potential to be data-mined for detailed form and surroundings information. Rather than purely relying on research assistant assessment from photos (without the benefit of a measuring stick for relative scale), this information would provide better base-line data and free up research assistant time on more valuable data entry on use-type, materials and overshadowing as required,

depending on the quality of the source data available. Koordinates is a repository of publicly available geographic data which has parallels in other countries such as the Massachusetts Institute of Technology's (MIT) Geo-data Repository (Massachusetts Institute of Technology, 2014) and OSGeo 'Open Source Geospatial Foundation' (Open Source Geospatial Foundation, 2014).

6.07: Scaling Up the Visualisation to National Level

This thesis has demonstrated it is possible to assign energy benchmarks to the commercial building stock for the whole of New Zealand (Sections 3.18) within an error band of +/- 15%. Buildings were assigned a colour according to use, size, form and height information from open data sources **Appendix iii**: 'Criteria Benchmark Averages'. Whole building energy benchmark values are further split into end-use energy benchmarks using detailed information on building uses available through open data sources at census level (Section 3.19).

Whilst the colouring of building records has been demonstrated as straightforward, it will be necessary to automate the process of colouring the digital building models to scale up to a national level building stock visualisation. This will require a specialist algorithm (Section 3.22) to colour the automatically generated building forms (Section 3.21). Based on the Fusion Table assessment of 'ambiguity' in Wellington CBD address fields (1%) and the Websearch sample and census records noted to have an address error (1.3-2%), it is likely that around 1-2% of building records are not identified accurately in the baseline data. The algorithm has been designed to combat at least some of this error through the colouring process.

Whilst every effort has been made to base the visualisation data on the highest quality background data possible, the benchmark averages, along with the long-tailed distribution curve of energy bills data, result in a broad-brush approach to benchmarking the building stock in New Zealand. The aim of this work was to communicate resource-use benchmarks to the widest possible audience so that a building owner or occupant is able to compare their energy use against national averages for their building typology. Whilst the energy intensity benchmark average values are subject to an error of around +/- 15%, the seven colours representing the values, span from very high through to very low intensity. This hierarchy is communicated in the visualisation so it could be argued that this serves the purpose of raising interest in building resource use. The fact that this method of visualisation is automated means that the data can be updated as new open data sources or improved statistical analysis become available. In addition, the author proposes a method of adding crowd-sourced measured data to the visualisation as a means of promoting buildings where efficiency upgrades or good design result in much lower energy intensities and therefore are worthy of advertisement. This crowd-sourced data together with updates and maintenance of the visualisation over time, will provide longitudinal data which should improve the potential of the tool.

A resource-use model could revolutionise the building industry, either through its use as an academic tool to investigate resource use and building typology trends, or as a virtual notice board for resource use benchmarks or efficiency improvements made to the building stock. These visual tools are highly appropriate for display of data relating to buildings and cities. In heads-up augmented reality displays, they would allow access to information that may be useful when navigating a city, investigating a site for design purposes or obtaining further knowledge about properties / locations prior to purchase or lease. Developments in smart phone / tablet technology and augmented reality would be significantly enhanced with this kind of statistical representation. The simple short term vision for this broader potential is that widely communicating energy and water benchmarks for all

buildings will drive property markets towards more sustainable, energy-efficient and resilient buildings (Warren-Myers, 2011).

Creating a significant impact on building stock efficiency may also depend upon the engagement of a wider audience. Making this type of data publicly accessible places an extremely powerful tool in the hands of building owners and users. This in turn will create the opportunity to improve the perceived value of sustainable buildings and precipitate thought and discussion on which parts of the building stock benefit most from certain interventions. The improved knowledge would also allow the market to drive sustainability, rather than relying solely on good will, or regulation and government policy.

7.00: Future Potential

Energy is a notoriously difficult concept to visualise. Viewing the world through a 'lens' of resource-use data should make it possible to envision the environmental performance of cities and neighbourhoods. Augmented reality and virtual globes such as Google Earth and Maps StreetView are open source and thus ideal for the communication of statistical data overlaying city built forms to the widest possible audience. As of October 2011, Google Earth had been downloaded more than a billion times (Google Developers, 2011).

The 3D city visualisation makes it possible to glance at an area of the city and understand the underlying resource-use trends. These additional layers of information overlaid onto Google Earth™ will make it possible to construct automated building stock models at national, state, county or city level. If the dashboard interface proposed by this thesis is successfully tested for usability with focus-groups and surveys involving laymen and building professionals, whilst maintaining the enjoyable aspects of 'play' (Section 2.03.03), it should result in a successful end-product. It will allow users to understand the built environment and the implications of building design / typology on energy and water use.

With additional statistical work (multiple regression and characteristic tree analysis), improved cadastral / GIS-based building form data and maintenance updates of the energy bills and end-use breakdowns, the visualisation can be tweaked and vastly improved in quality over time.

7.01: Future Applications of this Research

The publicly available benchmarks illustrated in this visualisation pilot are unique to New Zealand and based upon a broad-brush analysis. The introduction of mandatory display energy certificates for all public and commercial buildings in New Zealand (as has been gradually established in UK and Europe and as a voluntary scheme in Australia) would be equally valuable. Initial research in the UK suggests that the display energy certificates are slowly improving the value of the most sustainable buildings (Bull et al., 2012). The poorest performing buildings will not be able to be legally leased in the near future and the certificates themselves will eventually apply to all building types and sizes, so the effects of this legislation will become far-reaching. If New Zealand certificates were introduced, they could be visualised alongside resource use benchmarks set by typology. Imagine the power of a view that showed whether your certificate placed you above, or below the average benchmark for your building type. In the absence of any equivalent system to be introduced for existing buildings in New Zealand, a visualisation that provides the next best alternative, detailed benchmarks according to building typology, has been demonstrated as a real possibility by this thesis.

There is potential for this study to be undertaken elsewhere in a similar manner. Property taxation databases like the NZ Quotable Value are widely available around the world albeit in various

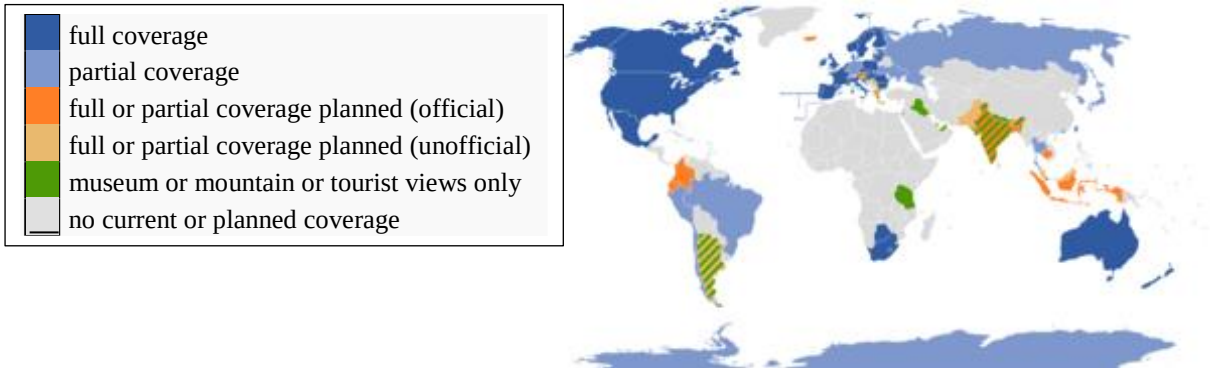


Figure 80 Google StreetView Global Coverage

configurations and of variable quality. Google Earth and Google Maps StreetView™ currently spans Japan, Europe, Australasia, USA, South Africa, Canada, Brazil and parts of Thailand with more and more countries being added all the time, Figure 80 (Google, 2013a).

7.02: Testing the Success of the Visualisation Tool

The visualisation proposed by this thesis must be tested with focus group user-surveys of lay persons as well as building professionals. Focus groups representing tenants, landlords and designers need to be consulted to investigate how much data can be communicated together before the 3D visualisation becomes over-crowded, off-putting and confusing. Each group should also be asked to comment on the type of information they would like to see available.

The success of the focus groups relies on some input from information technology specialists to make the visualisation function smoothly, such as the algorithm proposed in Section 3.22 and hyperlinks to make layers turn on and off within the Google Earth platform. This would establish whether or not the visualisation as proposed would achieve the wider goals as intended.

The literature reviewed regarding graphic representation styles and accessibility (Sections 2.03.03 and 2.03.06) stresses the importance of image as evidence and explanation displayed using analytical design, preferably without corruption or cherry-picking (Tufte, 2006).

"Don't ask how visualisation techniques can help display data. Ask how data can be best represented Increase information resolution (the maximum useful bits per unit area and unit time)…Treat all problems as multi-variate, make wise comparisons, show causality and use whatever it takes to get the message across" (Tufte, 2009).

The visualisation proposed by this thesis incorporates many of Tufte's analytical design principles, representing data on the built environment on the digital buildings themselves true to form and urban grain. The potential of this visualisation style to influence public opinion far outweighs the impact possible through charts, histograms, 2D heat-maps and extruded 2D GIS displaying precisely the same information (Figures 76-78 - Section 6.00).

Turning this visualisation into a game-style dashboard should make the information all the more enticing and enjoyable to access. Indeed, making data from the BEES study publicly available was one of the key deliverables requested from the stakeholders. Crowd-sourcing real data to display next

to benchmark data, along with the potential for the visualisation to be generated automatically, gives real capacity for quality improvement as new datasets are added to the mix.

Korhonen argues that "product design needs to improve the support of playful experiences in order to fit in with the users' multi-faceted needs… positive emotions are essential for the sake of curiosity and the ability to learn new things" (Korhonen et al., 2009) Section 2.03.03. Whilst not yet tested against a real focus group, the author has designed the visualisation to instigate a sense of playfulness, mimicking game design and neon light-like blue text hinting at the focus of the visualisation - recourse use. The control panel background is faintly transparent to retain a visual connection to the 3D city forms to aid in navigation and zoom control. Selecting any colour from the legend key, any typology or characteristic from the drop-down menus (or any group of buildings in the model), has the potential to isolate, query or highlight the statistics in an area of interest. There is capacity for a user to investigate their own line of enquiry within publicly available datasets with some sense of delight. This power to combine statistics at will, along with site information, new built environment research, and user-submitted data, should make resource-use publicly visible in an enjoyable way

Other graphic considerations raised in the literature review involve research on playfulness. (Sutton-Smith, 1997) argues that play: enforces the power status of the winning players; is a means of confirming, or advancing the power and identity of the community; is imaginary, creative or improvised; and focusses on the enjoyment or fun aspect of the participants.

A sense of competition is prompted by the visualisation through submitted resource-use data, advertising sustainable accomplishments (or perhaps finding you have a workplace better than the benchmark). Being able to observe change in the building stock over time through the display of longitudinal data might instil a sense of power and identity. Trying out new combinations of data together in 'mash-ups' raises questions and may inspire theories not envisaged by the author. The fact this comes packaged together in a 3D city (or building stock) visualisation with a set of dashboard controls would provide novelty value. Users could virtually explore urban space with an additional layer of annotated typology and resource-use data, ideally adding further enjoyment and raising awareness in the process. Where play is combined with learning about an important, yet illusive concept, playfulness must surely be crucial to ensure that message gets across to the widest possible audience.

The aim of the visualisation proposed by this thesis is to seduce the public with an invisible concept that of building resource use. It is designed to inspire a new way of thinking about the built environment, preferably changing behaviour with regards to purchasing, occupying or leasing property. This would potentially catalyse change in the market value of sustainable buildings. "3D visualisations are easily accessible …and have this seductive power" (Turkle, 2009).

8.00: Bibliography

Aquacraft. (2004). *Demonstration of Water Conservation Opportunities in Urban Supermarkets*. Metropolitan Water District of Southern California. Retrieved 15/05/15 from https://www.google.co.nz/search?q=Demonstration+of+Water+Conservation+Opportunities+in+Urban+Supermarkets&ie=utf-8&oe=utf-8&gws_rd=cr&ei=TB1VVeKeII338QXG04GYDw#.

Auckland Council. (2008). Auckland Rates and Valuation. Retrieved 17/05/13 from http://www.aucklandcouncil.govt.nz/EN/ratesbuildingproperty/ratesvaluations/ratespropertysearch/Pages/RatesSearch.aspx

Austen, K. (2013). Where in the World? *New Scientist*, (2900), 44–47.

Baird, G., Donn, M., Pool, F., Brander, W. D. S., & Chan, S. A. (1984). *Energy Performance of Buildings*. CRC Press, Boca Raton, Florida.

Baird, G., Donn, M., & Pool, F. (1978). *Energy demand in the Wellington central business district : final report* (No. 77). New Zealand Energy Research & Development Committee (NZERDC). *Stage 1 Report.*

Bentley. (2013). Generative Components. Retrieved 17/05/13 from http://www.bentley.com/en-US/Promo/Generative%20Components/default.htm

Berg, B. (2013). *Current Scholars - BRANZ* (Master of Building Science Thesis). Victoria University Wellington, Centre for Building Performance Research. Retrieved 22/08/13 from http://www.branz.co.nz/cms_display.php?sn=230&st=1&pg=11070.

Berlin Police. (2012). Berlin Atlas of Crime. Retrieved 17/05/13, from http://www.economicworldmap.net/berlincrime.html#%20imageanchor=

Billinghurst, M. (2011). HITLabNZ - CityViewAR. Retrieved 17/05/13 from http://www.hitlabnz.org/index.php/products/cityviewar

Bint, L. (2012). Benchmarking Water Use in Office Building. *BRANZ Build*, *125*.

Bishop, R. (2012a, July). Why Some Buildings Have Very High or Low EUI's. Retrieved 10/08/13, from

http://www.branz.co.nz/cms_show_download.php?id=10a45d57ad5c913a85c91e6afdb5ba7ea4634c48

Bishop, R., Pollard, A., & Isaacs, N. (2012). *BEES (Building Energy End-use Study) Year 5: Estimating Whole Building Energy Usage* (No. SR 277/4). BRANZ. Retrieved 18/05/13 from http://www.branz.co.nz/cms_show_download.php?id=e4b386ecdb291905f0523eaaf02c3e33a6d5d7e7.

BRANZ. (2013). BEES Publications and Presentations. Retrieved 22/08/13 from http://www.branz.co.nz/cms_display.php

Brown, F. E., Rickaby, P. A., Bruhns, H. R., & Steadman, P. (2000). Surveys of nondomestic buildings in four English towns. *Environment and Planning B: Planning and Design*, *27*(1), 11–24. http://doi.org/10.1068/b2571

Bull, R., Chang, N., & Fleming, P. (2012). The use of building energy certificates to reduce energy consumption in European public buildings. *Energy and Buildings*, *50*, 103–110. http://doi.org/10.1016/j.enbuild.2012.03.032

California Department of Water. (2003). *Waste Not Want Not: The Potential for Urban Water Conservation in California: Appendix C: Industrial and Commercial Water Use: Glossary, Data, and Methods of Analysis*. Oakland, California: Pacific Institute. Retrieved 20/05/13 from http://www.pacinst.org/reports/urban_usage/appendix_c.pdf.

Carbon Visuals. (2010). Carbon footprint of 40,000 UK public buildings. Retrieved 17/05/13, from http://www.carbonvisuals.com/work/google-earth-uk-public-buildings

Chanan, V., White, S., Howe, C., & Meenakshi, J. (2003). Sustainable Water Management in Commercial Office Building. In *Innovations in Water: Ozwater Convention and Exhibition*. Perth: Australian Water Association. Retrieved 20/05/13 from http://www.isf.uts.edu.au/publications/chananetal2003sustainableofficebuildings.pdf

City West Water, Yarra Valley Water, & South East Water. (2006). *Benchmarking Best Practice for Water Use in the Commercial and Industrial Sector*. Melbourne, Australia: Australian Water Association.

Clark, D. (2011, January 11). Carbon regulation for commercial property should be overhauled - UK-GBC – British Land. Retrieved 16/05/13 from http://www.britishland.com/responsibility/our-news/2011/24-01-2011.aspx

Columbia Engineering School. (2011). Model Created to Map Energy Use in NYC Buildings. Retrieved 10/08/13 from http://engineering.columbia.edu/model-created-map-energy-use-nyc-buildings

Corkery, E., & Coup, R. (2007). Koordinates - the one place for geodata. Retrieved 10/08/13 from http://koordinates.com/

Cory, S. (2015). *An Exploration of the Feasibility of Converting the New Zealand Commercial Building Stock to be Net Zero Energy* (Building Science PhD). Victoria University Wellington.

Cory, S., Hills, A., Isaacs, N., & Pyke, C. (2010). *Literature Review and Annotated Bibliography*. BRANZ: Victoria University of Wellington. Centre for Building Performance Research.

Costolo, D. (2014). Twitter. Retrieved 12/03/14 from https://twitter.com/

Council of the European Communities. Council Directive: on the indication by labelling and standard product information of the consumption of energy and other resources by household appliances, 92/75/EEC (1992). Retrieved 08/08/13 from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1992:297:0016:0019:EN:PDF

Cracknell, A. P. (2007). *Introduction to remote sensing* (2nd ed). Boca Raton, FL: CRC Press.

Cresswell-Wells, T., Donn, M., & Cory, S. (2012). *BEES Interim Report Year 5 - Christchurch Urban Form and Energy* (No. SR277/7). Retrieved 19/08/13 from http://www.branz.co.nz/cms_show_download.php?id=c23a619c2babfbedb4bfa5b59b2a162bd062203e. (19/08/13)

Darby, S. (2006). The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and Direct Displays*, *486*. Retrieved 16/05/13 from http://powerwatch.biz/site/wp-content/uploads/2012/02/smart-metering-report.pdf

Dept. Communities & Local Government. (2012, December). A Guide to Display Energy Certificates and Advisory Reports for Public Buildings. Retrieved 16/05/13 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/51164/A_guide_to_display_energy_certificates_and_advisory_reports_for_public_buildings.pdf

Dept. of Natural Resources and Water. (2006). Annual water statistics 2005–06. canberra, australia. Retrieved 19/05/13 from https://www.google.co.nz/search?q=Department+of+Natural+Resources+and+Water.+Annual+water+Statistics+2005%E2%80%9306.+Canberra%2C+Australia%3A+Department+of+Natural+Resources+and+Water%2C+2006.&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a&channel=fflb

Dept. Resources, Energy and Tourism. (2013, April 24). What is a BEEC? | CBD - Commercial Building Disclosure. Retrieved 17/05/13 from http://cbd.gov.au/get-and-use-a-rating/what-is-a-beec

Donn, M., Maule, L., & McAlister, S. (1995). *Lighting & equipment energy survey: 30 Wellington buildings 1994/5*. Victoria University of Wellington. Centre for Building Performance Research.

Donn, M. R. (2004). *Simulation of imagined realities: Environmental design decision support tools in architecture*. Retrieved 16/06/13 from http://researcharchive.vuw.ac.nz/handle/10063/136.

Dziegielewski, B., Kiefer, J., Opitz, E., Porter, G., Lantz, G., Aquacraft, & Nelson, J. O. (2000). *Commercial and Institutional End-uses of Water*. American Water Works Association Research Foundation. Retrieved 20/05/13 from ftp://ftpdpla.water.ca.gov/users/prop50/10024_BayArea/Attach.%208%20Supporting%20Docs/04%20New%20Business%20Guidebook/CIEUWS%20FINAL%20REPORT.pdf.

E-ideas. (2013). whoiswhere New Zealand leading B2B business lists. Retrieved 17/05/13 from http://www.whoiswhere.co.nz/about-us_wiw.html

Energy Efficiency Conservation Authority. (2007). EECA - Energy End Use Database [Database]. Retrieved 17/07/13 from http://enduse.eeca.govt.nz/default.aspx

Energy Efficiency Conservation Authority, & Centre for Advanced Engineering. (2000). *Possible Energy Use Trends for New Zealand 2000/2010*. Centre for Advance Engineering, University of Canterbury.

Energy Information Administration. (2013). Commercial Buildings Energy Consumption Survey (CBECS). Retrieved May 18, 2013, from http://www.eia.gov/consumption/commercial/about.cfm

Environmental Protection Agency. (2009). *Water Efficiency in the Commercial and Institutional Sector: Considerations for a WaterSense Program*. Retrieved 20/05/13 from http://www.epa.gov/WaterSense/docs/ci_whitepaper.pdf.

Esri CityEngine. (2013). *3D Modeling Software for Urban Environments*. Retrieved 18/05/13 from http://www.esri.com/software/cityengine

European Committee for Standardization. European Standard - Energy Performance of Buildings - Methods for Expressing Energy Performance and for Energy Certification of Buildings, Pub. L. No. ICS 91.140.10 (2007). Retrieved 08/0813 from file:///C:/Users/Architecture/Documents/BEES/BillbordassREFS/prEN15217%20from%20JCV%20Mar07.pdf

Georg, W. (2001, April 25). *Optical 3D-Metrology : Precise Shape Measurement with an extended Fringe Projection Method*. Retrieved 17/10/13 from http://archiv.ub.uni-heidelberg.de/volltextserver/1808/.

GHD Pty Ltd., & Australian Greenhouse Office. (2006). *Scoping Study to Investigate Measures for Improving the Water Efficiency of Buildings*. Canberra, Australia: Department of the Environment and Heritage.

Google. (2013a). *Google Maps with Street View*. Retrieved 24/08/13 from http://maps.google.co.nz/intl/en/help/maps/streetview/

Google. (2013b). Google SketchUp - Download. Retrieved 23/08/13 from http://google-sketchup.en.softonic.com/

Google. (2013c, February 28). Project Glass - Google+ Scrapbook Photos. Retrieved 17/05/13 from https://plus.google.com/+projectglass/posts

Google Apps. (2013a). Google Drive Fusion Tables. Retrieved 17/08/13 from http://www.google.com/drive/apps.html#fusiontables

Google Apps. (2013b). Google Earth. Retrieved 17/08//13 from http://www.google.com/earth/index.html

Google Developers. (2011, October 5). Official Blog: Google Earth downloaded more than one billion times. Retrieved 15/11/13 from http://googleblog.blogspot.co.nz/2011/10/google-earth-downloaded-more-than-one.html

Government of Canada, N. R. C. (2003). *Commercial and Institutional Building Energy Use Survey 2000, Summary Report*. Retrieved 18/05/13 from http://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/data_e/Cibeus2/scope_of_survey.cfm.

Groat, L. N. (2002). *Architectural research methods*. New York: J. Wiley.

Groves, D., Fischbach, J., Hickey, S., & RAND Corporation. (2007). *Evaluating the Benefits and Costs of Increased Water-Use Efficiency in Commercial Buildings*. Santa Monica, California, USA: Environment, Energy and Economic Development. Retrieved 20/05/13 from www.rand.org/content/dam/rand/pubs/technical_reports/2007/RAND_TR461.pdf.

Heinrich, M. (2007). Water End-use and Efficiency Project (WEEP) - A Case Study (Vol. CP 136 (2009)). Presented at the Sustainable Building 2007 NZ Transforming our Built Environment, Auckland, New Zealand: BRANZ. Retrieved 20/05/13 from http://www.branz.co.nz/cms_show_download.php?id=1007e839aa7e3b02ab6d012ca78d3c411f21098b

Hodson, H. (2012, December 1). Tech to the Streets. *New Scientist*, (2893), 19.

Hogg, R., & Tanis, E. (2009). *Probability and Statistical Inference Books a La Carte Edition.* (8th ed.). Pearson. Essex

Holtier, S., Steadman, J. P., & Smith, M. G. (2000). Three-dimensional representation of urban built form in a GIS. *Environment and Planning B: Planning and Design*, *27*(1), 51–72. http://doi.org/10.1068/bst5

Howard, B., Parshall, L., Thompson, J., Hammer, S., Dickinson, J., & Modi, V. (2012). Spatial distribution of urban building energy consumption by end use. *Energy and Buildings*, *45*, 141–151. http://doi.org/10.1016/j.enbuild.2011.10.061

Institute of Electrical and Electronics Engineers. (2009). World Haptics 2009: Third Joint EuroHaptics Conference and Symposium on *Haptic Interfaces for Virtual Environment and Teleoperator Systems*: Salt Lake city, Utah, USA,.

Isaacs, J., Gilmour, D., Blackwood, D., & Falconer, R. (2011). Immersive and Non immersive 3D Virtual City: Decision Support Tool for Urban Sustainability. *Journal of Information Technology in Construction*, *16*, 149–159.

Isaacs, N., Camilleri, M., French, L., Pollard, A., Saville-Smith, K., Fraser, R. Rossouw, P. & Jowett, J.. (2005). HEEP | Energy Use in New Zealand Households - Report on the Year 10 Analysis for the Household Energy End-use Project. Retrieved 17/05/13 from http://www.branz.co.nz/cms_show_download.php?id=b1ab61dd06f50e83e6a184b29b68a989472502ed

Isaacs, N., & Hills, A. (2013). Understanding the New Zealand Non-domestic Building Stock. *Building Research & Information*, *RBRI831274*. Retrieved 15/05/15 from http://www.tandfonline.com/doi/abs/10.1080/09613218.2013.831274?journalCode=rbri20

Isaacs, N., Jowett, J., Saville-Smith, K., & Hills, A. (2012). Understanding Energy and Water Use in New Zealand Non-Residential Buildings - an Ad Hoc Survey. Presented at *AMSTAT*. Retrieved 15/05/15 from http://www.amstat.org/meetings/ices/2012/papers/302177.pdf

Isaacs, N., Saville-Smith, K., Babylon, M., Bishop, R., Camilleri, M., Donn, M., Jowett, Moore, D. & J., Roberti, H. (2010). *Study Report: Building Energy End-use Study (BEES) Years 3* (No. SR 236). BRANZ. Retrieved 10/08/13 from

http://www.branz.co.nz/cms_show_download.php?id=95dad451acfda29895d90614873a2a45e5d6c25c

Isaacs, N., Saville-Smith, K., Bishop, R., Camilleri, M., Jowett, J., Hills, A., Moore, D., Babylon, M., Donn, M., Heinrich, M. & Roberti, H. (2009). *Study Report: Building Energy End-use Study (BEES) Years 1 & 2* (No. SR224). BRANZ. Retrieved 17/05/13 from http://www.branz.co.nz/cms_show_download.php?id=95dad451acfda29895d90614873a2a45e5d6c25c

Itron Inc. (2006). *California Commercial End-Use Survey* (No. CEC-400-20 06-0 05). California Energy Comission. Retrieved 18/05/13 from http://www.energy.ca.gov/2006publications/CEC-400-2006-005/CEC-400-2006-005.PDF

Jacobs, J. (1961). *The death and life of great American cities* (Modern Library ed). New York: Modern Library.

Jenks, G. (1967). The Data Model Concept in Statistical Mapping. *International Yearbook of Cartography*, *7*, 186–190.

Johnson, L. (2007, May 7). Building Design Leaders Collaborating on Carbon-Neutral Buildings by 2030. *Fast Company*. Retrieved 16/05/13 from http://www.fastcompany.com/678734/building-design-leaders-collaborating-carbon-neutral-buildings-2030

Korhonen, H., Montola, M., & Arrasvuori, J. (2009). Understanding Playful User Experience Through Digital Games. Presented at the *Designing Pleasurable Products and Interfaces*, Université de Technologie Compiègne. Retrieved 09/06/13 from http://books.google.co.nz/books?id=b2_iYgEACAAJ

Kurakula, V., Skidmore, A. K., Kluijver, H., Stoter, J., Dabrowska-Zielinska, K., & Kuffer, M. (2007). *A GIS-based approach for 3D noise modelling using 3D city models* (Master of Science). University of Southampton, Southampton, UK. Retrieved 22/05/13 from http://www.itc.nl/library/papers_2007/msc/gem/kurakula.pdf

Land Surveyors United. (2013). GIS. Retrieved 17/05/13 from http://landsurveyorsunited.com/group/gis

LINZ. (2011a, May 22). NZ Mainland Building Polygons (Topo, 1:50k) - National Topographic Office | New Zealand | GIS Data. Retrieved 17/05/13 from http://data.linz.govt.nz/layer/246-nz-mainland-building-polygons-topo-150k/

LINZ. (2011b, December). address - Ollivier & Co | New Zealand | GIS Map Data | Koordinates. Retrieved 17/05/13 from http://koordinates.com/#/layer/910-address/

LINZ. (2011c, December). NZ Building Footprints - Land Information New Zealand | New Zealand | GIS Map Data | Koordinates. Retrieved 17/05/13 from http://koordinates.com/#/layer/127-nz-building-footprints/

LINZ. (2011d, December). NZ Cadastral Parcel Polygons - Ollivier & Co | New Zealand | GIS Map Data | Koordinates. Retrieved 17/05/13 from http://koordinates.com/#/layer/1236-nz-cadastral-parcel-polygons/

Ludicorp. (2004). Welcome to Flickr - Photo Sharing. Retrieved 10/08/13 from http://www.flickr.com/

Malin, N., & Roberts, T. (2012, September 1). *Energy Reporting: It's the Law* [Environmental Building News]. Retrieved 16/05/13 from http://www.buildinggreen.com/auth/article.cfm/2012/7/30/Energy-Reporting-It-s-the-Law/

Marks, P. (2012, October 27). Paste Augmented-reality Video Graffiti on the Streets. *New Scientist*, (2888) Retrieved 15/05/15 from http://www.newscientist.com/article/mg21628886.000-paste-augmentedreality-video-graffiti-on-the-streets.html

Massachusetts Institute of Technology. (2014). MIT Geodata Repository - Geographic Information Systems (GIS). Retrieved 10/04/14 from http://libguides.mit.edu/Geodata

MBIE. (2013). HomePage | Ministry of Business, Innovation, and Employment. Retrieved 22/08/13 from http://www.mbie.govt.nz/

New Zealand Legislation. Building Act 2004 New Zealand Legislation, Pub. L. No. 2004 no. 72 (2004). Retrieved 03/12/13 from http://www.legislation.govt.nz/act/public/2004/0072/latest/DLM306036.html

NSW Office of Environment and Heritage. (2014, September 26). First 6 Star NABERS Energy Rating for Major Office Building Announced. Retrieved 14/05/13 from http://www.nabers.gov.au/public/WebPages/ContentStandard.aspx?module=10&template=3&include=mediarelease.htm&side=EventTertiary.htm

NZ Institute of Building. (2012, November 22). Central Chapter News - NZIOB. Retrieved 18/05/13 from http://www.nziob.org.nz/page/council-central-chapter

Open Geospatial Consortium. (2011, November). GeoMobile and Internet of Things | OGC(R). Retrieved 15/05/13 from http://www.opengeospatial.org/projects/initiatives/geomobilecd

Open Source Geospatial Foundation. (2014). OSGeo.org | Your Open Source Compass. Retrieved 10/04/14 from http://www.osgeo.org/

Oreszczyn, T., Lomas, K., Ward, I. C., Guy, S. C., & Shipworth, D. (2013). CaRB Project. Retrieved 18/05/13 from http://www.ucl.ac.uk/carb/

Parliamentary Counsel Office. Building Research Levy Act 1969 No 23 (as at 01 July 2013) (1969). Retrieved 23/08/13 from http://www.legislation.govt.nz/act/public/1969/0023/latest/DLM391231.html?search=ts_act_building+research&sr=1

Peet, N. J. (1985). *Direct and indirect energy requirements of the New Zealand economy 1976-77: An energy analysis of the 1976-77 interindustry survey.* (No. 96). New Zealand Energy Research and Development Committee, Auckland

Prebble, C. (2010). Mashblock - Stats on every NZ block. Retrieved 10/08/13 from http://mashblock.co.nz/

Property Council of NZ. (2014). Office and Retail grade matricies.xls - Quality Grading Matrix.pdf. Retrieved 04/03/13 from http://www.propertynz.co.nz/files/Research/Quality%20Grading%20Matrix.pdf

Quinn, R., & Bannister, P. (2006). *Water Efficiency Guide: Office and Public Buildings*. Canberra: Dept. of the Environment and Heritage.

QV. (2008). Quotable Value. Retrieved 17/05/13 from http://www.qv.co.nz/

Reed, R., Bilos, A., Wilkinson, S., & Schulte, K.-W. (2009). International comparison of sustainable rating tools. *The Journal of Sustainable Real Estate*, *1*(1), 1–22.

Rheingans, P. L. (2000). Task-based color scale design. In *28th AIPR Workshop: 3D Visualization for Data Exploration and Decision Making* (pp. 35–43). Retrieved 09/08/13 from http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=916058

Roberti, H. (2012, July). *BEES Water Use Auckland Baseline Study*. BRANZ. Retrieved 10/08/13 from http://www.branz.co.nz/cms_show_download.php?id=1d7d744215de2786f868227efae73f5a4fccb473

Rosling, O., Rosling Ronnlund, A., & Rosling, H. (2005, February 25). Gapminder: Unveiling the beauty of statistics for a fact based world view. Retrieved 17/05/13 from http://www.gapminder.org/

R. W. Morris & Associates. (1985). *Commercial sector energy use in Christchurch* (No. 115). New Zealand Energy Research and Development Committee, Auckland

Ryan, R. A. (2007). *Enhancing 3D Models with Urban Information: A Case Study Involving Local Authorities and Property Professionals in New Zealand: Quantifying the Benefit of 3D over Alternative 2D systems* (Master of Building Science Thesis). Victoria University Wellington. Retrieved 17/05/13 from http://researcharchive.vuw.ac.nz/handle/10063/1162

Sezgen, O., & Koomey, J. G. (1995). *Technology Data Characterizing Water Heating in Commercial Buildings: Application to End-Use Forecasting*. Report No. LBL-37398. Berkeley, CA: Lawrence Berkeley National Laboratory. Retrieved 20/05/13 from http://enduse.lbl.gov/Info/LBNL-37398.pdf

Shaw, R. A., Beca, Carter, Hollings and Ferner, & New Zealand Energy Research and Development Committee. (1979). *Greater Auckland Commercial Sector Energy Analysis*.

Stanford Business School, US Department of Energy, & White House. (2012, May). *Energy Data Initiative/Energy Data Jam | Open Energy Information*. Retrieved 10/08/13 from http://en.openei.org/wiki/Energy_Data_Initiative/Energy_Data_Jam

Steadman, P., Bruhns, H. R., Holtier, S., Gakovic, B., Rickaby, P. A., & Brown, F. E. (2000). A classification of built forms. *Environment and Planning B: Planning and Design*, *27*(1), 73–91. http://doi.org/10.1068/bst7

Steadman, P., Bruhns, H. R., & Rickaby, P. A. (2000). An introduction to the national Non-Domestic Building Stock database. *Environment and Planning B: Planning and Design*, *27*(1), 3–10. http://doi.org/10.1068/bst2

Sullivan, J., & Motley, K. (2011, June 10). *Communicating Information About the Urban Environment in 3D: Preliminary Study* (Building Science Honours Thesis). Victoria University Wellington.

Sutton-Smith, B. (1997). *The ambiguity of play*. Cambridge, Mass: Harvard University Press.

Tufte, E. R. (2006). *Beautiful evidence*. Cheshire, Conn.: Graphics Press.

Tufte, E. R. (2009, September 20). An Interview With Edward Tufte. Retrieved 11/08/13 from http://www.vizworld.com/2009/09/an-interview-with-edward-tufte/

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass: Addison-Wesley Pub. Co.

Tullis, T., & Albert, B. (2008). *Measuring the user experience collecting, analyzing, and presenting usability metrics*. Amsterdam; Boston: Elsevier/Morgan Kaufmann. Retrieved 09/06/13 from http://www.engineeringvillage.com/controller/servlet/OpenURL?genre=book&isbn=9780123735584

Turkle, S. (2009). *Simulation and its discontents*. Cambridge, Mass.: MIT Press. Retrieved 09/06/13 from http://site.ebrary.com/id/10292363

UCL. (2008, April 2). London Profiler. Retrieved 17/05/13 from http://www.londonprofiler.org/

UK Government. Energy Act (2011). Section 74(1) Retrieved 16/05/13 from http://www.legislation.gov.uk/ukpga/2011/16/contents/enacted *"The Green Deal" relates to restrictions placed on F+G rated buildings from 2016 and bans from 2018*

UNEP. (2009). *Buildings and Climate Change - Summary for Decision-Makers*. United Nations Environment Programme. Retrieved 17/05/13 from http://www.unep.org/sbci/resources/Publications.asp

Urban Land Institute. (2009). LessEN: Energy Map. Retrieved 10/08/13, from http://less-en.org/?page=LessEnEnergyMap

U.S. Department of Energy. (2013). DOE Buildings Performance Database. Retrieved 10/08/13 from https://bpd.lbl.gov/

US Department of Energy. (2013a). *EnergyPlus Energy Simulation Software*. Building Technologies Office: Retrieved 24/08/13 from http://apps1.eere.energy.gov/buildings/energyplus/

US Department of Energy. (2013b). *NREL: OpenStudio*. Retrieved 12/08/13 from http://openstudio.nrel.gov/

Voss, K., & Musall, E. (2011). *Net zero energy buildings: international comparison of carbon-neutral lifestyles.* Birkhauser Verlag AG.

Warren-Myers, G. (2011). Valuing Sustainability: A new challenge for the profession. *Australia and New Zealand Property Journal*, *3*(September), 157–164.

Wilks, A. E., Donn, M., & Baird, G. (2003). *BEEP: Building Energy End-use Project - energy end-uses in non-residential buildings* (EECA Energy Efficiency Conservation Authority). Centre for Building Performance Research, Victoria University Wellington.

Yau, N. (2013). *Data points: visualization that means something.* Indianapolis, IN, Wiley

Yelp. (2013). Yelp Wordmap | Yelp. Retrieved 10/08/13 from http://www.yelp.co.uk/wordmap/london/beer

WEBESEARCH DATABASE - SOURCES, COLLECTION METHODOLOGY, DATA CLEANING + ALTERNATIVE PUBLICALLY AVAILABLE SOURCES

WEBSEARCH FIELD COLUMN CODES: **DATA CLEANING:** **OPEN DATA:**

CODE:		NAME:	SOURCE:	BEES DATA HISTORY / CLEANING	MASTERS STUDY DATA CLEANING	METHOD UTILISED:	FREE CURRENTLY	PAYMENT REQ'D	EFFORT REQ'D	FUNCTION REQ'D FOR AUTOMATION	ALGORITHIM / IT TOOL REQ'D	PRIVATE DATA
	A	**SampleStrata**	QV data / Auckland Valuation data	Concatenate Columns B.C.D	n/a	n/a	Individual Floor Area / Cadastral Data	Bulk Floor Area Data	Admin Only	n/a	n/a	n/a
	B	**CatStratum**	QV data / Auckland Valuation data	Child amalgamated with parent, simplified category of use CM/CL/CS/CT/CV= CR	n/a	n/a	Individual Category	Bulk Category Data	Admin Only	n/a	n/a	n/a
	C	**RegionStrata**	QV data / Auckland Valuation data	Child amalgamated with parent, region stratum as Auckland + RestNZ	n/a	n/a	NZ Post Data		Admin Only	n/a	n/a	n/a
	D	**Size Stratum**	QV data / Auckland Valuation data	Total NZ Commercial Floor Area divided up into 20% creates size stratum 1-5	Adjusted after websearch into column JL "newSIZEstrata"	"TotalAreaBuilding" gives newSIZEstrata data	Individual Floor Area	Bulk Floor Area Data	Admin Only	n/a	n/a	n/a
	E	**Rank**	BRANZ data	Random sampling tools	n/a	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	F	**ID**	QV data / Auckland Valuation data	QPID(qv) + Parent / Property Key (Auckland) Duplicates removed in Sample Frame	n/a	n/a	Individual Code	Bulk ID codes	Admin Only	n/a	n/a	n/a
	G	**Floor Area**	QV data / Auckland Valuation data	Websearch research gives "totalAREAbuilding" child amalgamated with parent including total of area	n/a	n/a	Individual Floor Area	Bulk Floor Area Data	Admin Only	n/a	n/a	n/a
	H	**FloorPlate(Land Coverage)**	QV data Only	Websearch research gives "FloorPlate"	n/a	n/a	Individual Floor Area	Bulk Floor Area Data	Admin Only	n/a	n/a	n/a
	I	**ImprovementValue**	QV data / Auckland Valuation data	Child amalgamated with parent, unsure if value also amalgamated???	n/a	n/a	Individual Value	Bulk Improvement Values	Admin Only	n/a	n/a	Valuation Improvement Value only for private research (no publication of values)
	J	**Cat.Code**	QV data / Auckland Valuation data	Child amalgamated with parent, commercial category codes used CM/CL/CS/CT/CV/CR/CO/IW/IS	n/a	n/a			Admin Only	n/a	n/a	n/a
	K	**Code**	QV data / Auckland Valuation data	Child amalgamated with parent, commercial category descriptions used	n/a	n/a	Individual Category	Bulk Category Data	Admin Only	n/a	n/a	n/a
	L	**RegionalCouncil**	QV data / Auckland Valuation data	n/a	n/a	n/a	NZ Post Data	n/a	Admin Only	n/a	n/a	n/a
	M	**Address**	QV data / Auckland Valuation data	Auckland addresses concatenated from separate fields	n/a	n/a	NZ Post Data	n/a	Admin Only	n/a	n/a	n/a
	N	**Suburb/Town**	QV data / Auckland Valuation data	Auckland addresses concatenated from separate fields	n/a	n/a	NZ Post Data	n/a	Admin Only	n/a	n/a	n/a
	O	**Town/District**	QV data / Auckland Valuation data	Auckland addresses concatenated from separate fields	n/a	n/a	NZ Post Data	n/a	Admin Only	n/a	n/a	n/a
	P	**District**	QV data / Auckland Valuation data	Auckland addresses concatenated from separate fields	n/a	n/a	NZ Post Data	n/a	Admin Only	n/a	n/a	n/a
	Q	**BRaNZ 2**	BRANZ data	Random sampling tools	n/a	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	R	**BRaNZ 1**	BRANZ data	Random sampling tools	n/a	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	S	**X**	BRANZ data	Random sampling tools	n/a	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	T	**Slice**	BRANZ data	Random sampling tools	n/a	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	U	**Status**	BRANZ data	Random sampling tools	n/a	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	V	**BIZCount**	Finda Business Search	Returned Business Categories / Names	Business Numbers counted from FINDA Search	Check of no. of businesses returned in columns apn1-73 (EH-HB) Ignored in masters research due to high number of businesses	Finda Business Individual	Finda Business Search Bulk	Admin Only	n/a	n/a	n/a
	W	**ABN1**	Finda Business Search	Business Name 1. Returned by FINDA (no known reasons for order of businesses)	Business Name used to create new USE categories	These have not be ignored as category of use and business types likely to have remained even if business moved on...	Finda Business Individual	Finda Business Search Bulk	Admin Only	n/a	n/a	n/a
	X	**ABNcat1**	Finda Business Search	Detailed FINDA Business Category no. 1	Detailed Business Category used to create "newUSE1/2" categories.	These have not be ignored as category of use and business types likely to have remained even if business moved on...	Finda Business Individual	Finda Business Search Bulk	Admin Only	n/a	n/a	n/a
	Y	**ABN2**	Finda Business Search	Business Name 2. Returned by FINDA (no known reason for order of businesses)	Business Name used to create new USE categories	These have not be ignored as category of use and business types likely to have remained even if business moved on...	Finda Business Individual	Finda Business Search Bulk	Admin Only	n/a	n/a	n/a
	Z	**ABNcat2**	Finda Business Search	Detailed FINDA Business Category no. 2	Detailed Business Category used to create "newUSE1/2" categories.	These have not be ignored as category of use and business types likely to have remained even if business moved on...	Finda Business Individual	Finda Business Search Bulk	Admin Only	n/a	n/a	n/a
	AA	**Extra Buildings**	Websearch Database	Websearch Research	Count of No. of Buildings associated with address eg. "=COUNTIF($F$3:$F$3470,F3465)"	n/a	NZ Post Data	Who is Where $1 / record more	Admin Only	n/a	n/a	n/a
	AB	**New USE1**	Websearch Database	Websearch Research	algorithm - Search for Keywords / major business names or detailed FINDA category as indicated to provide code R(food, tavern, bar, restaurant,takeaway) O(offices,medical,professionals,administration) R(retail,shops,mall,plaza) in Columns W-Z, AM-AO, EH-HB	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm Appendix of Finda Categories	n/a	n/a
	AC	**New USE2**	Websearch Database	Websearch Research	algorithm - as for AB but secondary code indicated only if different from AB	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AD	**NewUSE3/Dwelling**	Websearch Database	Websearch Research	algorithm - as for AB but tertiary code indicated only if different from AB and if Dwelling indicated in QV codes (AM-AO)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AE	**F**	Websearch Database	Websearch Research	Building Contains New Use F (food, tavern, bar, restaurant,takeaway)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AF	**R**	Websearch Database	Websearch Research	Building Contains New Use R (retail, shop, supermarket, mall, plaza, butcher, point of sale warehouse)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AG	**O**	Websearch Database	Websearch Research	Building Contains New Use O (office, professional services, medical, government, admin., legal, headquarters, communications)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AH	**X**	Websearch Database	Websearch Research	Building Contains New Use X (special cinema, university, hospital, theatre, carehome, school, creche)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AI	**I**	Websearch Database	Websearch Research	Building Contains New Use I (Industrial, Storage, Warehouse, Factory, Manufacturing, Depot, Cool Stores)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AJ	**H**	Websearch Database	Websearch Research	Building Contains New Use H (hotel, cabin, campground, holiday park, passive outdoor, motel, hostel, clubrooms)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AK	**D**	Websearch Database	Websearch Research	Building Contains New Use D (dwelling, residential on site, flat, apartments, house)	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	Simple Find Text algorithm	n/a	n/a
	AL	**CODEconcatinate**	Websearch Database	Websearch Research	Concatenate codes AE-B	n/a	Finda Business / QV Individual	QV / FINDA data Purchased	Admin Only	n/a	n/a	n/a
	AM	**Category Detailed**	QV data / Auckland Valuation data	Child amalgamated with parent, Commercial Category Detailed Used	n/a	n/a	Individual Category	Bulk Category Data	Admin Only	n/a	n/a	n/a
	AN	**Improvement / Component Description**	QV data / Auckland Valuation data	Child amalgamated with parent, Commercial Improvement / Component Description	n/a	n/a	Individual Category	Bulk Category Data	Admin Only	n/a	n/a	n/a
	AO	**Landuse Type**	QV data / Auckland Valuation data	Child amalgamated with parent, Commercial Landuse Type used	n/a	n/a	Individual Category	Bulk Category Data	Admin Only	n/a	n/a	n/a
	AP	**Building Age**	QV data / Auckland Valuation data	Child amalgamated with parent, Commercial Building Age Used (Mixed indicated if diff. decades in associated records) Auckland data by year, QV data by decade	"BuildingDecade" Decade Before 1920, "B1920" and decade to bring together two different levels of detail in source data	n/a	Individual Age Data 4Free	Bulk Age Data	Admin Only	n/a	n/a	n/a
	AQ	**Building Fabric QV**	QV data Only	Child amalgamated with parent, Commercial Building Fabric Used	n/a	n/a	Individual Category	Bulk Category Data	Admin Only	n/a	n/a	n/a
	AR	**Roof Construction QV**	QV data Only	Child amalgamated with parent, Commercial RoofConstruction Used	n/a	n/a	Individual Category	Bulk Category Data	Admin Only	n/a	n/a	n/a
	AS	**BEESBuilding Record No.**	BRANZ data	Unique four digit ID no. for BEES buildings	Conversion required due to changes to BRANZ coding system	See Column CL for revised code if this affects building number so that premise data can be matched to websearch data	n/a	n/a	Admin Only	n/a	n/a	n/a
	AT	**Address Correction**	Websearch Database	Websearch Research	Deleted repetition of accurate address so that only incorrect addresses displayed	Identification of incorrect address in QV database (nearest building/s of that size used only where obvious mistake)	NZ Post Data		Admin Only	n/a	n/a	n/a
	AU	**ID Extra Buildings**	Websearch Database	Single Building "A" (3042) QV Record Websearch results in more than one building at the address then "B" - "K" copied rows of data	n/a	n/a	NZ Post Data	Who is Where $1 / record more	Admin Only	n/a	n/a	n/a
	AV	**Aerial HYPERLINK**	Websearch Database	Websearch Research	=Hyperlink Formula	Research uncovers the most reliable way to store hyperlinks in database is to use the Excel formula with adjacent cell as text file	n/a		Admin Only	n/a	n/a	n/a
	AW	**Aerial photo**	Google Earth (Websearch)	Websearch Research Google Earth	File Name Cleaning so that hyperlinks work reliably	Search for certain text strings in windows explorer and repair errors in database as well as file storage	Google Earth Free Assuming Non-Commercial use	n/a	n/a	n/a	download aerial photos: link to address field	n/a
	AX	**Floor Plate**	Google Earth (Websearch)	Websearch research	If QV data is complete and looks roughly correct then Floor Area divided by floor plate gives storey height	Trained Research Assistants Measure Floor Plate from Google Earth using Ruler tool (at least 10% error expected) Floor plate multiplied by storey no.s = "totalAREAbuilding"	Cadastral Data + Analysis to link to address		Admin Only	n/a	n/a	n/a
	AY	**# of storeys**	Google Maps StreetView (Websearch)	Websearch research Google Maps	Further work to infer building height where absent (from websearch or QV total area / floor plate)	Trained Research Assistants Estimate storeys from Google Maps StreetView elevation Images	Elevation Data / Total Floor Area divided by FloorPlate + Analysis to Link to Address		Admin Only	n/a	n/a	n/a
	AZ	**AVE Building Height (m)**	Google Maps StreetView (Websearch)	Websearch research Google Maps	n/a	Trained Research Assistants Estimate building height in metres from Google Maps StreetView elevation Images	Elevation Data / Total Floor Area divided by FloorPlate + Analysis to Link to Address Koordinates Elevation Data + Analysis to Link to Address		Admin Only	n/a	n/a	n/a
	BA	**Floor Area FORMULA**	Websearch Database	Websearch Research	Multiply Floor Plate by no.s of Storeys	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a

APPENDIX i - Websearch Data Sources

	BB	**Floor Area Estimate**	Websearch Database	Websearch Research	Where trained research assistants estimate building is a different total floor area than "FloorAreaFORMULA" provides	n/a	Cadastral Data + Analysis to link to address	n/a	Admin Only	n/a	n/a	n/a
	BC	**surface area**	Websearch Database	Websearch Research	Inferred Data: "FloorAreaFORMULA" divided by "FloorAreaEstimate" - creates information that suggests surface area or compactness ratio	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	BD	**GLAZING TOTAL RATIO AVE**	Websearch Database	Websearch Research	Formula Averaging all available Glazing Ratios	Average of estimated glazing ratios where available (BG,BH,BI,BJ)	n/a	n/a	Admin Only	n/a	n/a	n/a
	BE	**% Glazing - North**	Google Maps StreetView (Websearch)	Websearch research Google Maps	Search fo unusual figures simplified to nearest 10% to bring differing accuracies into line. Those with 90-100% glazing ratios investigated and modified accordingly due to over estimation	Average of estimated glazing ratio north elevation rounded to nearest 10%	n/a	Websearch Research Project	Research Assistants		Data mine StreetView / Cadastral Data for Adjoining Buildings	
	BF	**Elevation North HYPERLINK**	Websearch Database	Websearch Research	=Hyperlink Formula	Research uncovers the most reliable way to store hyperlinks in database is to use the Excel formula with adjacent cell as text file	n/a	Websearch Research Project	Admin Only	n/a	n/a	n/a
	BG	**Elevation North**	Google Maps StreetView (Websearch)	Websearch research Google Maps	n/a	StreetView Elevation Image closest to North Elevation (Naming protocol: "ID no."-N.jpg)	n/a	Websearch Research Project	Research Assistants	n/a	download elevation photos: link to address field	n/a
	BH	**% Glazing - East**	Google Maps StreetView (Websearch)	Websearch research Google Maps	Search fo unusual figures simplified to nearest 10% to bring differing accuracies into line. Those with 90-100% glazing ratios investigated and modified accordingly due to over estimation	Average of estimated glazing ratio east elevation rounded to nearest 10%	n/a	Websearch Research Project	Research Assistants		Data mine StreetView / Cadastral Data for Adjoining Buildings	
	BI	**Elevation East HYPERLINK**	Websearch Database	Websearch Research	=Hyperlink Formula	Research uncovers the most reliable way to store hyperlinks in database is to use the Excel formula with adjacent cell as text file	n/a	Websearch Research Project	Admin Only	n/a	n/a	n/a
	BJ	**Elevation East**	Google Maps StreetView (Websearch)	Websearch research Google Maps	n/a	StreetView Elevation Image closest to East Elevation (Naming protocol: "ID no."-E.jpg)	n/a	Websearch Research Project	Research Assistants	n/a	download elevation photos: link to address field	n/a
	BK	**% Glazing - South**	Google Maps StreetView (Websearch)	Websearch research Google Maps	Search fo unusual figures simplified to nearest 10% to bring differing accuracies into line. Those with 90-100% glazing ratios investigated and modified accordingly due to over estimation	Average of estimated glazing ratio south elevation rounded to nearest 10%	n/a	Websearch Research Project	Research Assistants		Data mine StreetView / Cadastral Data for Adjoining Buildings	
	BL	**Elevation - South HYPERLINK**	Websearch Database	Websearch Research	=Hyperlink Formula	Research uncovers the most reliable way to store hyperlinks in database is to use the Excel formula with adjacent cell as text file	n/a	Websearch Research Project	Admin Only	n/a	n/a	n/a
	BM	**Elevation - South**	Google Maps StreetView (Websearch)	Websearch research Google Maps	n/a	StreetView Elevation Image closest to South Elevation (Naming protocol: "ID no."-S.jpg)	n/a	Websearch Research Project	Research Assistants	n/a	download elevation photos: link to address field	n/a
	BN	**% Glazing - West**	Google Maps StreetView (Websearch)	Websearch research Google Maps	Search fo unusual figures simplified to nearest 10% to bring differing accuracies into line. Those with 90-100% glazing ratios investigated and modified accordingly due to over estimation	Average of estimated glazing ratio west elevation rounded to nearest 10%	n/a	Websearch Research Project	Research Assistants		Data mine StreetView / Cadastral Data for Adjoining Buildings	
	BO	**Elevation - West HYPERLINK**	Websearch Database	Websearch Research	=Hyperlink Formula	Research uncovers the most reliable way to store hyperlinks in database is to use the Excel formula with adjacent cell as text file	n/a	Websearch Research Project	Admin Only	n/a	n/a	n/a
	BP	**Elevation - West**	Google Maps StreetView (Websearch)	Websearch research Google Maps	n/a	StreetView Elevation Image closest to West Elevation (Naming protocol: "ID no."-W.jpg)	n/a	Websearch Research Project	Research Assistants	n/a	download elevation photos: link to address field	n/a
	BQ	**Business Name/s**	Google Maps StreetView (Websearch)	Websearch research Google Maps		Business Names confirmed or added from StreetView Imagery or Google searches on Address	n/a	Who is Where $1 / record	Admin Only	n/a	n/a	n/a
	BR	**Party Wall N**	Google Maps StreetView (Websearch)	Websearch Research	Overshadowing - (BW,BZ,CC,CF) distance to nearest building for each orientation gives location of party walls.	Binary Entry Inferred from Overshadowing Data Collected by Trained Websearch Assistants	Possible: Cadastral Data + Analysis to link to address	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	BS	**Party Wall E**	Google Maps StreetView (Websearch)	Websearch Research	Overshadowing - (BW,BZ,CC,CF) distance to nearest building for each orientation gives location of party walls.	Binary Entry Inferred from Overshadowing Data Collected by Trained Websearch Assistants	Possible: Cadastral Data + Analysis to link to address	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	BT	**Party Wall S**	Google Maps StreetView (Websearch)	Websearch Research	Overshadowing - (BW,BZ,CC,CF) distance to nearest building for each orientation gives location of party walls.	Binary Entry Inferred from Overshadowing Data Collected by Trained Websearch Assistants	Possible: Cadastral Data + Analysis to link to address	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	BU	**Party Wall W**	Google Maps StreetView (Websearch)	Websearch Research	Overshadowing - (BW,BZ,CC,CF) distance to nearest building for each orientation gives location of party walls.	Binary Entry Inferred from Overshadowing Data Collected by Trained Websearch Assistants	Possible: Cadastral Data + Analysis to link to address	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	BV	**Party Wall Total**	Google Maps StreetView (Websearch)	Websearch Research	Count party walls for each building gives 0-3 and n/a for those buildings without data	Count of Party Wall Data n/a for no data available	Possible: Cadastral Data + Analysis to link to address	Websearch Research Project	Admin Only	n/a	n/a	n/a
	BW	**Average Distance Nearest Building - N**	Google Earth (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google Earth measurment of average distance nearest building in North direction	Possible: Cadastral Data + Analysis to link to address	Websearch Research Project	Admin Only	Degrees Tan V Floor Plate: skylight factor - N (HY)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	BX	**Average Height Nearest Building - N**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average height nearest building in North direction				Degrees Tan V Floor Plate: skylight factor - N (HY)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	BY	**Average Ground Level From Base of Building - N**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average change in ground level from zero datum of lowest ground level elevation				Degrees Tan V Floor Plate: Site Slope Angle - N (HY)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	BZ	**Average Distance Nearest Building - E**	Google Earth (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google Earth measurment of average distance nearest building in East direction				Degrees Tan V Floor Plate: Overshadowing Angle - E (HZ)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CA	**Average Height Nearest Building - E**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average height nearest building in East direction				Degrees Tan V Floor Plate: Overshadowing Angle - E (HZ)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CB	**Average Ground Level From Base of Building - E**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average change in ground level from zero datum of lowest ground level elevation				Degrees Tan Tan V Floor Plate: Site Slope Angle - E (HZ)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CC	**Average Distance Nearest Building - S**	Google Earth (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google Earth measurment of average distance nearest building in South direction				Degrees Tan V Floor Plate: Overshadowing Angle - S (IA)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CD	**Average Height Nearest Building - S**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average height nearest building in South direction				Degrees Tan V Floor Plate: Overshadowing Angle - S (IA)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CE	**Average Ground Level From Base of Building - S**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average change in ground level from zero datum of lowest ground level elevation				Degrees Tan V Floor Plate: Site Slope Angle - S (IA)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CF	**Average Distance Nearest Building - W**	Google Earth (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google Earth measurment of average distance nearest building in West direction				Degrees Tan V Floor Plate: Overshadowing Angle - W (IB)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CG	**Average Height Nearest Building - W**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average height nearest building in West direction				Degrees Tan V Floor Plate: Overshadowing Angle - W (IB)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CH	**Average Ground Level From Base of Building - W**	Google Maps StreetView (Websearch)	Websearch Research	Introduction of overshadowing after pilot websearch stage	Google StreetView measurment of average change in ground level from zero datum of lowest ground level elevation				Degrees Tan V Floor Plate: Site Slope Angle - W (IB)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	CI	**Building Fabric**	Google Maps StreetView (Websearch)	Websearch Research	Categories chosen to match QV categories as much as practical	Primary material present in majority of façade	Council Archive Data Individual Only	Websearch Research Project	Research Assistants	n/a	Analysing colours / surfaces in StreetView Imagery	n/a
	CJ	**Building Fabric - Secondary**	Google Maps StreetView (Websearch)	Websearch Research	Categories chosen to match QV categories as much as practical	Secondary material present in majority of remaining façade	Council Archive Data Individual Only	Websearch Research Project	Research Assistants	n/a	Analysing colours / surfaces in StreetView Imagery	n/a
	CK	**Roof Construction**	Google Earth (Websearch)	Websearch Research	Categories chosen to match QV categories as much as practical	Primary material present in majority of roof surface	Council Archive Data Individual Only	Websearch Research Project	Research Assistants	n/a	Analysing colours / surfaces in StreetView Imagery	n/a
	CL	**Roof Construction - Secondary**	Google Earth (Websearch)	Websearch Research	Categories chosen to match QV categories as much as practical	Secondary material present in majority of remaining roof	Council Archive Data Individual Only	Websearch Research Project	Research Assistants	n/a	Analysing colours / surfaces in StreetView Imagery	n/a
	CM	**Glazing Type - Primary**	Google Maps StreetView (Websearch)	Websearch Research	Categories chosen to match QV categories as much as practical	Primary glazing frames present in majority of windows	Council Archive Data Individual Only	Websearch Research Project	Research Assistants	n/a	Analysing colours / surfaces in StreetView Imagery	n/a
	CN	**Glazing Type - Secondary**	Google Maps StreetView (Websearch)	Websearch Research	Categories chosen to match QV categories as much as practical	Secondary glazing frames present in majority of remaining windows	Council Archive Data Individual Only	Websearch Research Project	Research Assistants	n/a	Analysing colours / surfaces in StreetView Imagery	n/a
	CO	**Built Form - Primary**	Google Maps StreetView (Websearch)	Websearch Research	Categories chosen to coordinate with the UK system used in the Non-Domestic Building Stock Model	Trained Research Assistants use websearch info sheet to quickly identify form type		Websearch Research Project	Research Assistants	n/a	n/a	n/a
	CP	**Built Form - Secondary**	Google Maps StreetView (Websearch)	Websearch Research	Categories chosen to coordinate with the UK system used in the Non-Domestic Building Stock Model	Trained Research Assistants use websearch info sheet to quickly identify form type		Websearch Research Project	Research Assistants	n/a	n/a	n/a
	CQ	**BUILT FORM NEW**	Websearch Database	Websearch Research	Simplification of initial "Classification of Built Form - Steadman" codes due to added details on storey height etc. in database	n/a	n/a	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	CR	**BUILT FORM NEW SECONDARY**	Websearch Database	Websearch Research	Simplification of initial "Classification of Built Form - Steadman" codes due to added details on storey height etc. in database	n/a	n/a	Websearch Research Project	Research Assistants	n/a	n/a	n/a

	CS	**Sketchup HYPERLINK**	Websearch Database	Websearch Research	=Hyperlink Formula	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	CT	**Sketchup**	Google Earth + Maps StreetView (Websearch)	Websearch Research	Potential explored for automated modelling using Generative Components (Bentley Microstation Plugin) City Engine (Esri) and Excel Background Data - quality checks required.	Simple geo-located block model created using Google SketchUp with Open Studio plugin to create simple model files ready for Energy Plus Simulation Modelling (with application of glazing /	Cadastral Data Extruded to Elevation Height Data minus Ground Topography	n/a	Admin Only	n/a	n/a	n/a
	CU	**QUALITY FINISHES**	Google Maps StreetView (Websearch)	Websearch Research	Introduced following Steering Group Meeting		Council Archive Data Individual Only	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	CV	**Notes**	Websearch Database	Websearch Research	Coded to extract main reasons for errors / omissions	n/a	n/a	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	CW-EF	**StreetView Extras 1-18**	Google Maps StreetView (Websearch)	Websearch Research	n/a	n/a	n/a	Websearch Research Project	Research Assistants	n/a	download elevation photos: link to address field	n/a
	EG	**RA initials**	Websearch Database	Websearch Research	n/a	n/a	n/a	Websearch Research Project	Research Assistants	n/a	n/a	n/a
	EH-HB	**APN 1-73**	Finda Business Search	Business Name 1. Returned by FINDA (no known reasons for order of businesses)	Business Name used to create new USE categories.	These have not be ignored as category of use and business types likely to have remained even if business moved on...	Finda Business Individual	Who is Where $1 / record better than Finda Business Bulk Purchase which is out of date	Admin Only	n/a	n/a	n/a
	HC	**Aerial**	Websearch Database	Websearch Research	Analysis of blank fields to create coverage details	True False field indicating access to Aerial Pictures	n/a	n/a	Admin Only	n/a	n/a	n/a
	HD	**Street View**	Google Earth + Maps StreetView (Websearch)	Websearch Research	Analysis of blank fields to create coverage details	True False field indicating access to StreetView images	n/a	n/a	Admin Only	n/a	n/a	n/a
	HE	**Coverage**	Websearch Database	Websearch Research	Coverage code created	0 = zero aerial + street view, 0.5=street view only, 1.5=aerial only, 2=both available at time of websearch or during later	n/a	n/a	Admin Only	n/a	n/a	n/a
	HF	**NOTE CODE AH**	Websearch Database	Websearch Research	Coded to extract main reasons for errors / omissions for analysis of possible bias	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	HG	**Total Area Building**	Websearch Database / QV + Auckland Valuation Data	QV Data "G" or Websearch Research "BB"	Amalgamate best results from each source into new dataset	Google Earth Measurements plagued with inaccuracy - best to use Cadastral data for final tool	Cadastral + Elevation Data with Analysis	n/a	Admin Only	n/a	n/a	n/a
	HH	**New Category**	Websearch Database	Websearch Research	New Category amalgamated from CAT.Code (J) and ChangedCAT (HI)	n/a	n/a	n/a	Admin Only	n/a	n/a	n/a
	HI	**ChangedCAT**	Websearch Database	Websearch Research	Websearch research indicating change of use category from QV data	Google StreetView images or Google Searches for Business	n/a	Who is Where $1 / record	Admin Only	n/a	n/a	n/a
	HJ	**GLAZING N**	Websearch Database	Websearch Research	Binary code to indicate presence of Glazing Ratio Data for analysis of StreetView Coverage per building and to find average of glazing ratios available	eg. '=IF(BE3="",0,1)	n/a	n/a	n/a	Simple Database Formula	n/a	n/a
	HK	**GLAZING E**	Websearch Database	Websearch Research	Binary code to indicate presence of Glazing Ratio Data for analysis of StreetView Coverage per building and to find average of glazing ratios available	eg. '=IF(BE3="",0,1)	n/a	n/a	n/a	Simple Database Formula	n/a	n/a
	HL	**GLAZING S**	Websearch Database	Websearch Research	Binary code to indicate presence of Glazing Ratio Data for analysis of StreetView Coverage per building and to find average of glazing ratios available	eg. '=IF(BE3="",0,1)	n/a	n/a	n/a	Simple Database Formula	n/a	n/a
	HM	**GLAZING W**	Websearch Database	Websearch Research	Binary code to indicate presence of Glazing Ratio Data for analysis of StreetView Coverage per building and to find average of glazing ratios available	eg. '=IF(BE3="",0,1)	n/a	n/a	n/a	Simple Database Formula	n/a	n/a
	HN	**No. GLAZING DONE**	Websearch Database	Websearch Research	Street View Coverage per Building	No. of Glazing Ratios Collected (4 maximum) possible numeric indicator of site access and block size when combined with average overshadowing angles	n/a	n/a	n/a	Simple Database Formula	n/a	n/a
	HO	**BuildingDecade**	Websearch Data	Child amalgamated with parent, Commercial Building Age Used (Mixed indicated if diff. decades in associated records) Auckland data by year, QV data by decade	"BuildingDecade" Decade Before 1920, "B1920" and decade to bring together two different levels of detail in source data	n/a	n/a	Bulk Age Data	Admin Only	Simple Subsitution Algorithim	n/a	n/a
	HP	**ID duplicate**	Websearch Data	Websearch Research	Count of No. of ID's associated with record in database eg. "=COUNTIF($F$3:$F$3470,F3465)" - indicates extra building no. (1 for building A etc)	n/a	n/a	n/a	n/a	Simple Database Formula	n/a	n/a
	HQ	**RecordTotalBUILDINGSAREA**	Websearch Data	Websearch Research	Amalgamation of total building area for all extra buildings	eg. =SUMIF($F$3:$HG$3470,F3,$HG$3:$HG$3470)	n/a	n/a	n/a	Simple Database Formula	n/a	n/a
	HR	**MIXED/SINGLE USE**	Websearch Data	Websearch Research	Simplification of NewUSE Codes CodeConcatenate (AL) for various types of analysis energy related to typology	M = Mixed Use (including retail and/or office plus industrial or special), S=Single Use, N=100%NonBees Buildings (Iincluding	n/a	n/a	Admin Only	Simple Subsitution Algorithim	n/a	n/a
	HS	**CODEconcatSIMPLE**	Websearch Data	Websearch Research	Simplification of NewUSE Codes CodeConcatenate (AL) for various types of analysis energy related to typology	M = Mixed Use (including retail and/or office plus industrial or special), F=Food, R=Retail, O=Office, M=Mixed Use, N=100%NonBees Buildings (Iincluding Industrial / Special etc)	n/a	n/a	Admin Only	Simple Subsitution Algorithim	n/a	n/a
	HT	**CODEconcatLEFT2**	Websearch Data	Websearch Research	Simplification of NewUSE Codes CodeConcatenate (AL) for various types of analysis energy related to typology	eg. =LEFT($AL$3,2)	n/a	n/a	n/a	n/a	n/a	n/a
	HU	**STOREY RANGE**	Websearch Data	Websearch Research	Simplification of Storey Height into Range, 0,1,2,3,4,5,10+ for analysis	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	HV	**STOREY**	Websearch Data	Websearch Research	Storey Height Amalgamated - best of Storey Height Websearch + CalcStorey where available	Obvious errors corrected to allow for area mistakes by factor of 10 or by guessing from shadow on aerial photograph	Storey Height Collated a) from Websearch b) from QV Floor area / Floor Plate c) If major error obvious with b) then assumed to be 1 storey or out by factor of 10 (feet)	n/a	n/a	n/a	n/a	n/a
	HW	**CalcSTOREY**	Websearch Data	Websearch Research	Calculated Storey Height From QV data FloorArea/Floor Plate	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	HX	**STOREYdiff**	Websearch Data	Websearch Research	Calculate the difference between Websearch # of Storeys (AY) and CalcSTOREY (HW)	Useful for analysis of errors in database	n/a	n/a	n/a	n/a	n/a	n/a
	HY	**OVERshadN**	Websearch Data	Websearch Research	Calculation of Daylight Factor Inferred from Surroundings Data (BW)	Google Earth measurment of daylight factor for north elevation	Possible: Cadastral Data + Analysis to link to address	n/a	Admin Only	Degrees Tan V Floor Plate: Overshadowing Angle - N (HY)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	HZ	**OVERshadE**	Websearch Data	Websearch Research	Calculation of Daylight Factor Inferred from Surroundings Data (BW)	Google Earth measurment of daylight factor for east elevation	Possible: Cadastral Data + Analysis to link to address	n/a	Admin Only	Degrees Tan V Floor Plate: Overshadowing Angle - E (CA/BZ)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	IA	**OVERshaS**	Websearch Data	Websearch Research	Calculation of Daylight Factor Inferred from Surroundings Data (BW)	Google Earth measurment of daylight factor for south elevation	Possible: Cadastral Data + Analysis to link to address	n/a	Admin Only	Degrees Tan V Floor Plate: Overshadowing Angle - S(CF/CE)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	IB	**OVERshaW**	Websearch Data	Websearch Research	Calculation of Daylight Factor Inferred from Surroundings Data (BW)	Google Earth measurment of daylight factor for west elevation	Possible: Cadastral Data + Analysis to link to address	n/a	Admin Only	Degrees Tan V Floor Plate: Overshadowing Angle - W(CG/CF)	Data mine StreetView / Cadastral Data for Adjoining Buildings	n/a
	IC	**Approx. Building Length (Sqrt)**	Websearch Data	Websearch Research	Building Length determined by Square Root of Floor Plate	Floor plate from eg. =SQRT(HG$3)	Possible: Cadastral Data + Analysis to link to address	n/a	Admin Only	n/a	n/a	n/a
	ID	**Significant Ground Slope N/S**	Websearch Data	Websearch Research	Calculation of Site Slope where Significant	Use Surroundings data BY+CE divided by Approx. Building Length (IC)	n/a	n/a	n/a	Degrees Tan V Floor Plate: Significant Ground Slope - N/S ((BY+CE)/IC)	n/a	n/a
	IE	**Significant Ground Slope E/W**	Websearch Data	Websearch Research	Calculation of Site Slope where Significant	Use Surroundings data CB+CH divided by Approx. Building Length (IC)	n/a	n/a	n/a	Degrees Tan V Floor Plate: Significant Ground Slope - E/W((CB+CH)/IC)	n/a	n/a
	IF	**Summer 2011**	Websearch Data	Websearch Research	Prioritisation of Records to be completed by later websearch stage	Initial stages carried out by area but ran out of time so full sample was not completed in order therefore infill records were required to complete the data set.	n/a	n/a	n/a	n/a	n/a	n/a
	IG	**WIW Category Of Use**	Who is Where?	Websearch Research	VLookup Who is Where Data - Category of Use to improve new USE categories	Used to further refine / replace the detailed New Use codes	Free on individual basis through Google Search / StreetView Business signage, however more reliable and bulk available via Who is Where?	Who is Where $1 / record better than Finda Business Bulk Purchase which is out of date	Admin Only	n/a	n/a	n/a
	IH	**WIW Bizname**	Who is Where?	Websearch Research	VLookup Who is Where Data - Category of Use to improve new USE categories	Used to further refine / replace the detailed New Use codes	n/a	Who is Where $1 / record better than Finda Business Bulk Purchase which is out of date	Admin Only	n/a	n/a	n/a
	II		n/a	n/a	Working Columns	used initially to check data errors in storey height eg. small areas or size strata 1 with high storey heights NOW BLANK	n/a	n/a	n/a	n/a	n/a	n/a
	IJ	**Zone type**	Google Earth (Websearch)	Websearch Research	Google Earth Survey of Suburb Density	Use NZ Stats definition of primary and secondary urban area for "U" code, site suburb names on Google Earth to establish suburb street grain "S" and rural street grain (businesses fall on one or	NZ Post Data + Analysis List (See Appendix)	n/a	n/a	n/a	n/a	n/a
	IK	**Density**	Websearch Data	Websearch Research	Average of available daylight factors to give density	eg. =AVERAGE(HY3,HZ3,IA3,IB3]	n/a	n/a	n/a	n/a	n/a	n/a
	IL	**Typologies Size**	Websearch Research	Coding Small<900sq.m floor plate, Storey Tall = 5+ Medium 1.5-4 Short 0-1.5	Small medium height- sm,small Tall- sT, Small Short-SS, Big Short - Bs, Big Medium-Bm, Big Tall- BT	For benchmarking purposes	n/a	n/a	n/a	n/a	n/a	n/a
	IM	**Monitored**	Lee Bint VUWdataRequest.xls	BEES monitored buildings - energy monitored for 2 weeks min. survey of building and contents	awaiting results of energy intensity, appliance count and lighting counts for monitored building. May not be available prior to masters completion...	Assuming data was received in time for this visualisation.. the appliance counts and lighting counts would further improve the extrapolation potential to the census buildings... an investigation into the bias of typologies within the data set may also assist the		n/a	n/a	n/a	n/a	n/a

	IN	**NPI Discarded Premises**	Websearch Data	BEES phone survey buildings with billing information. Those with incomplete energy or bills are discarded by Nigel Isaacs.	Review dataset of discarded buildings to see if any bias present...	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	IO	**Typologies USE Type**	Websearch Research	n/a	Oss: Office small, RFbs: Contains Retail/Mix/Food, Big Short,OvTx: Very Tall Compact Form	Initially there were some results uinder 0.3 kwh/sq.m/year which would seem less than a light bulb but these have now been resolved. Energy Intensities over 2,000 kwh/sq.m/yr have been discarded from the analysis work		n/a	n/a	n/a	n/a	n/a	n/a
	IP	**phone survey**	Lee Bint VUWdataRequest.xls	BEES phone survey data available	Additional data available on occupancy ownership floor area and businesses within buildings through phone survey data set	An investigation into the bias of typologies within the data set may also assist the BEES team		n/a	n/a	n/a	n/a	n/a	n/a
	IQ	**bills**	Lee Bint VUWdataRequest.xls	BEES billing information available	Additional data available on energy / water bills on certain premises within building	An investigation into the bias of typologies within the data set may also assist the BEES team		n/a	n/a	n/a	n/a	n/a	n/a
	IR	**leeBINTcodeCHANGE**	Lee Bint VUWdataRequest.xls	BEES Code Change Information	Codes converted and changes noted on websearch database.	Code CHANGED notes where last four digits BEES Building No. remains as existing since bees building no. remains unaffected		n/a	n/a	n/a	n/a	n/a	n/a
	IS	**Electricity kWh/m^2/year**	Lee Bint VUWdataRequest.xls	BEES billing information available	Electricity energy intensity only given for first premise in list	An investigation into the bias of typologies within the data set may also assist the BEES team		n/a	n/a	n/a	n/a	n/a	n/a
	IT	**Gas kWh/m^2/year**	Lee Bint VUWdataRequest.xls	BEES billing information available	Gas energy intensity only given for first premise in list	An investigation into the bias of typologies within the data set may also assist the BEES team		n/a	n/a	n/a	n/a	n/a	n/a
	IU	**Water kL/m^3/year**	Lee Bint VUWdataRequest.xls	BEES billing information available	Water intensity only given for first premise in list - only 20 figures for water available within the bees stratified random sample	An investigation into the bias of typologies within the data set may also assist the BEES team		n/a	n/a	n/a	n/a	n/a	n/a
	IV	**Electricity AVE kWh/m^2/year**	Websearch Research	Premises Energy Provided by Lee Bint VUWdatarequest.xls	Average energy intensity calculated to provide figure for use in analysis and extrapolation	Average energy calculated by adding premise gas + electricity and then averaging the totals for each premise to give an average energy intensity per building in the websearch	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	IW	**AREA SAMPLE FRAME**	BRANZ data	Sample Frame.xls original file filtered from QV / Auckland data Michael Camilleri BRANZ (2008)	Refer records back to original sample frame floor area (quality assurance) so this matches BEES data	Vlookup of areas taken from original sample frame spreadsheet	n/a	n/a	Admin Only	n/a	n/a	n/a	n/a
	IX	**VOLUME Cu.m**	Websearch Research	n/a	Multiplication of Floor Plate * Storey height m (some calculated from QV etc)	used initially to indicate change of BEES building no. code. Now empty	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	IY	**Heavy Lightweight Mats WebS**	n/a	n/a	H for heavyweight, L for lightweight	used initially to indicate change of BEES monitored buildings vlookup Now empty	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	IZ	**Heavy Lightweight Mats QV**	n/a	n/a	H for heavyweight, L for lightweight	used initially to indicate change of BEES building no. code. Now	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JA	**Open plan v's Cellular**	n/a	n/a	Working Columns	used initially to indicate change of BEES building no. code. Now	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JB	**Colage Secondary timber/MetalP**	Websearch Research	n/a	Col code for collage of materials on façade (secondary = timber / MetalProfile)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JC	**climate zone**	Websearch Database	Websearch Research	Climate zone to match energy modelling zones in Shaan Cory's phd study.	Using regionl council boundary lines grouped into 7 zones.	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JD	**energy use per building**	Websearch Database	Websearch Research	Total Energy use per building Mw/yr	Average energy intensity multiplied by the building area in the sample frame spreadsheet.	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JE	**no. of occup. Min**	Who is Where?	Database purchased in November 2011 based on 2000 valuation addresses - includes all premises information available which, unlike FINDA data, is carefully checked by company to monitor turn over etc..	Occupncy no.s based on the sum of the minimum number of occupants quoted in ranges given in the Who is Where? Database for each premise within a building	Separate out occupant ranges into two cells, then use =sumif function for all premises to establish minimum no. of occupants	n/a	Who is Where $1 / record better than Finda Business Bulk Purchase which is out of date	Admin Only	n/a	n/a	n/a	n/a
	JF	**no. of occup. Max**	Who is Where?	Database purchased in November 2011 based on 2000 valuation addresses - includes all premises information available which, unlike FINDA data, is carefully checked by company to monitor turn over etc..	Occupncy no.s based on the sum of the maximum number of occupants quoted in ranges given in the Who is Where? Database for each premise within a building	Separate out occupant ranges into two cells (min. no. and max no.) then use =sumif function for all premises to establish maximum no. of occupants	n/a	Who is Where $1 / record better than Finda Business Bulk Purchase which is out of date	Admin Only	n/a	n/a	n/a	n/a
	JG	**no. of Biz wiw**	Who is Where?	Database purchased in November 2011 based on 2000 valuation addresses - includes all premises information available which, unlike FINDA data, is carefully checked by company to monitor turn over etc..	Premise no.s found within any given building address	n/a	n/a	Who is Where $1 / record better than Finda Business Bulk Purchase which is out of date	Admin Only	n/a	n/a	n/a	n/a
	JH	**CHANGEDresourceUSE**	Lee Bint VUWdataRequest.xls	BEES billing information changes	Various stages in changes to the energy results from the premises bills	x= initial change of energy intensity for each premise, xx= added premises and second change to energy intensities, xxx=third change - some from formula errors / analysis of code changes	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JI	**ERROR % INTENSITY LOW**	Lee Bint VUWdataRequest.xls	BEES billing information changes	Look at difference between premise average energy and lowest energy intensity premises within a building to establish whether it is reliable to use a simple average of premises in a building to give a building energy intensity (of if these should be weighted further to allow for difference in size of premise)	Highest EUI for building minus =AVERAGE(EUIPremise1+GasPremise1,EUIPremise2+GasPremise2,EUIPremise3+GasPremise3 etc.)	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JJ	**ERROR % INTENSITY high**	Lee Bint VUWdataRequest.xls	BEES billing information changes	Look at difference between premise average energy and highest energy intensity premises within a building to establish whether it is reliable to use a simple average of premises in a building to give a building energy intensity (of if these should be weighted further to allow for difference in size of premise)	=AVERAGE(EUIPremise1+GasPremise1,EUIPremise2+GasPremise2,EUIPremise3+GasPremise3 etc.) minus lowest EUI	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JK	**comment**	John Jowett alteration	Find replace to repair 100 records with mixed up qv data (QV or administration fields only - building age, strata code, bees building no, improvement discription, landuse type etc)	Subsequent work to ensure new use codes building decade data and related data columns are not affected by these errors in the data	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JL	**NewGroupingsPOSTwebS**	Websearch Database	Websearch Research	Revised SIZE that matches the coding for floor area strata for the building found by the websearch work.	Sort of records in order of floor area size in column BA "floor area estimate"	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	JN	**large short/v.tall retail and mixed NEW USE**	Websearch Research		**Rm** = retail short or Very Tall, and Mixed or Retail NEW USE			n/a	n/a	n/a	n/a	n/a	n/a
	JP	**Range WIW occupants / sq.m**	Who is Where? / Websearch					n/a	n/a	n/a	n/a	n/a	n/a

BEES (BUILDING ENERGY END-USE STUDY)

BRANZ / VUW RESEARCH PROJECT:
NOTES FOR RESEARCH ASSISTANTS

Supervisor: Alexandra Hills (Alexandra.hills@vuw.ac.nz / 021577135)

STRUCTURE OF THE DATABASE

ORIGINAL DATA:

The information contained in the worksheet columns comes from several sources:

F-P: Quotable Value Data (QV) "Parent" Records – Ownership Details / Value
X-AM: Business Suffix + Name / "Finda" Category Info from Data Market
AP-AU: Quotable Value Data (QV) Age, Building Fabric + Category Info

DATA COLLECTION:

We will be adding the following information to the database (in some cases information contained within the original data will be corrected):

AV-AZ: Research Study (Location / Aerial Photo / Building Floor Plate Area/s)
BA-BD: Research Study (Storeys / Height / Total Floor Area)
BE-BO: Research Study (Elevation Photos + Glazing % by orientation)
BQ-BR: Research Study (Correct Category / All Assoc. Business Names)
BS-CD: Research Study (Surrounding Buildings / Ground Levels)
CE-CM: Research Study (Building Fabric, Form + Sketch Up File)
CO-DX: Research Study (Notes / Extra Elevations)
DZ-EC: ID's not to be altered (Pilot # / Research Assistant ID / Unique OccuBuild ID)

BACKGROUND WORKSHEETS:

No changes should be made to these sheets they are for reference only.

DATA FROM AUCKLAND:

See "Slice 1-2AK" for supplementary information on building if you are getting really stuck on a record or if you think it will help clarify a grossly inaccurate floor area. Copy the ID number (Column F) and search for additional info from related "Child" QV records in this sheet which may give you additional information about each Floor and its area / use / category etc. "G" "1" "2" (See Columns R+S)

DATA FROM REST OF NEW ZEALAND:

See "Slice-1-2-QV" for supplementary information on building if you are getting really stuck on a record anywhere other than Auckland City. . Copy the ID number (Column F) and search for additional info from related "Child" QV records in this sheet which may give you additional information eg. Legal Description (Lot + DP number) which may help with council web searches if required.

DATA MARKET EXCESS INFO

Extra Businesses Listed Here. Some are major errors or duplications returned by Data Market.

DROPDOWN MENUS:

This worksheet describes the lists of possible categories / built fabric types / form types.

GETTING STARTED:

QV DATA:

The QV ID number (Column F) may represent a portion of a building, one building or several buildings. The information you are providing must relate to a *whole building*.

QV ID records have already been duplicated where we have sourced multiple business names associated with the QV record (Column Y). These are numbered "2" and above in "Business Suffix ID" (Column X). If the Business Suffix is "1" this row represents the first building "A". No Business Name should be associated with this row. (in Column Y)

If you come across a QV record that is related to more than one building - add new rows for each extra building on the site and duplicate the QV information in cells F-N so that each new row represents information about a separate building. Label column AX with "B" "C" etc.

IMPORTANT: **DO NOT ADD ANY COLUMNS** TO DATABASE.

"BUILDING" DEFINITION:
A "building" is defined by the outer building fabric / cladding envelope. Garages, shed's or other "ancillary" buildings under 30sq.m can be ignored on a site with multiple structures (but not where they are the only building on a site). Each building needs a separate letter code (Column AX).

USE CATEGORY CODE:
If the Category Code seems wrong and no commercial / industrial use can be found for any part of the building after checking then correct the category code in column BQ using one from your reference pages. The QV building categories we are looking at include **Commercial Mixed / Commercial Retail (including services, motor, tourist) / Commercial Office / Industrial Warehouse / Industrial Services.** We are not covering Residential buildings unless they have a commercial unit / mixed use component. We are not covering Education buildings (universities / schools / colleges) nor hospitals / medical centres etc. If the QV record does not conform to the **Commercial / Industrial** codes mentioned above just record the correct category code and calculate / confirm the floor area + building height. If the building appears vacant note this category code down. If the building has been demolished put 0 in the floor plate + Floor Area Estimate measurement and add this fact to the notes column (CO). In cases where there are mixed uses put down eg. CX Commercial Mixed and note down any further information in the notes column.

GENERAL POINTS:
- **If there are no Google Earth** aerial pictures available try the local council website eg. http://www.wellington.govt.nz/services/rates/search/search.html for property search. Failing this see if you can access elevation views on Street View. Once in a Street View Image "User Photos" may pop up in the right hand corner if there are any available related to that road.
- **If there are no Google Street View** images available (make sure you check roads all around the site) then see if there are any Google photos available in the area nearby indicated on Google Earth with a blue square.
- **If none of the above** suggestions work and you don't have any personal knowledge of the site then abandon the QV record and add a note to this effect in column (CO). We may purchase information (GIS / Terralink etc.) at a later stage for records like this.
- **Don't agonise over details**. Add information whenever possible even if you are not entirely convinced of your answer. If you have made a "bold" assumption then make sure this fact is noted in the NOTES (column CO).
- **Areas are only accurate to the nearest 10-20%** due to measurements on Google Earth. This is entirely acceptable don't spend too long trying to calculate.
- **Areas / heights** should be filled in to the nearest m. Decimal entries are not required.
- **Storey heights** should be to one decimal place to allow for complex building forms. Ie. If half the building area is 1 storey and the other half 2 storey = 1.5 storeys.

- **Still don't know?** Then insert a dash "-" in the cell/s to show you have tried to find an answer. This avoids records being mis-interpreted as "0" and shows what work has been done to date.
- **Don't type text with numbers into cells** expecting a figure. Any notes that may be required should be added to column CO
- **Can't find building?** If you are looking for a large commercial building but nothing appears associated with the address - be sure to look on the aerial photo as large commercial buildings are often easy to spot even with limited address information. Eg. Actual address is around the corner on a side street.

SUGGESTED APPLICATIONS:

- **Mozilla / Internet Explorer** (Yellow Pages / White Pages / Council Website) Add Favourites to speed up access to the sites you need (eg. the council website region you are dealing with.)
- **Google Earth** (Latest Version Please)
- **Google Maps + Street View** (Most handy when run inside Google Earth in the split screen)
- **Microsoft Office Picture Manager** (Set up a short cut to the Q drive folder you are working in)
- **Google Sketch Up with Open Studio Plug-in** (Latest Version Please)

HEIRACHY OF INFORMATION (RELIABILITY):

1. **Personal RECENT knowledge** of building / site (use my office phone if necessary)
2. **Council Website /** GIS High Resolution Pictures / Terralink (High Res Pics)
3. **Google Maps Street View** 2008 (collected between Jan – Oct 2008)
4. **Yellow Pages / White Pages** (Should be reasonable up to date)
5. **Google Earth** (satellite photos are generally much older)
6. **Google Web Searches** (you may be picking up old information)

<u>DATA ENTRY COLUMNS</u>

AV. Address Correction

Confirm street number or give alternate address where completely inaccurate. ie. Around the corner on a different street…NB* *If a range of street numbers are given and the address is a single number within this range do not waste time trying to find a single street number.*

AW. Aerial Photo (Hyperlink)

Maximise image on screen and save image from Google Earth file menu. Sometimes the default Google Earth marker provided is not in the right place, in such cases, a yellow pinpoint or shaded area was placed to help show exactly where the properties are located. Copy Paste the QV ID number (Column F) as the file name and save as a jpg (save in same folder as database). Enter hyperlink to that cell. (When selecting existing hyperlinks in excel hold down mouse button to avoid following hyperlink)

AX. ID Extra Building Code

The default building code is "A". Add new rows for each extra building associated with the property. Copy-Paste the QV data into the new rows. Label these "B" "C" etc. Fill in details about the each *whole building* in columns AV-DX. Ignore sheds / ancillary buildings / garages under about 30sq.m unless they are the only building at that property address.

AY.

AZ. Floor Plate (Land Coverage) sq.m

Measure building floor plate as it appears on Google Earth with the measuring tool. (check council website property search if necessary) Check Street View elevations to confirm building floor plate does not include large roof overhangs / carports / covered courtyards. The floor plate area may be given in some circumstances. 10-20% differences are to be expected.

BA. # Storeys (m) Explore all possible street view elevations around all sides of the building site and note storey heights orientation of each façade, and any variation in ground level around the building. The number of storeys should be given to 1 decimal place. Eg. If half the building is 2 storey and half is 3 storey the storey height is 2.5. (This figure is later multiplied by the floor plate to give a rough guide to total floor area for the building.

BB. Average Building Height (m)
Deduce approximate building height to nearest metre from building base. Allow for double height spaces where obvious on façade – compare to adjacent buildings / people / vehicles.

BC. Floor Area Calculation (Figure Generated by Excel)
This cell will automatically update with AY x AZ. Note how this compares to the QV floor area figure. If grossly inaccurate then copy paste the ID number into "Find" and check in the AK / QV worksheets for further information from "child" records. (This may uncover a basement space, unexpected floor levels or other useful details) Adjust storey height or Floor Plate if necessary.

BD. Floor Area Estimation
Fill in this cell if you need to include a known basement (see above) or if you can easily provide a better estimate for the total building area that can not be expressed by storey height x floor plate area.

BE. (BH, BK, BN,) Glazing % N, E, S,W
In street view and user photos etc. check all elevations. Chose closest elevation to each orientation N E S W. Note down glazing % to nearest 10%. Ensure you note 0% etc. where it is an obvious party wall / string of terrace buildings or a façade with no glazing at all. The % figure is approximate therefore assumptions can be made from oblique views or distant views. Check all around site on street view. Don't include roller doors in % ratio unless it is obvious that they conceal a shop window etc. Note in column CO if "bold" assumption is made. You do not have to use the % symbol.

BF. (BI, BL, BO) Elevation Photo Hyperlinks N, E, S, W
Chose closest elevation street view available for each orientation. Select full screen view icon and wait until text disappears. Screen Dump (Ctrl-Alt-PrtSc). Paste to Microsoft Office Picture Manager into the working folder. Save file with QV ID number, Building Letter Suffix "A" etc. and Orientation (N / E / S / W) etc. If close to 45% from either orientation then use the suffix SE / SW / NE / NW to clarify orientation of facade. Try to capture the whole building in one screen if possible (even an oblique view down a side street will do). If building is long and can not be viewed from a distant street view panorama or user photo then create additional images from left to right. Eg. "123456A-SE2.jpg" (Use capital letters for orientation) If building is tall and can not be viewed from a distant street view image or user photo then create additional images by looking up in panorama. Eg. "123456A-SEu.jpg" (use lower case "u"). Insert hyperlinks into cell as per aerial picture. Extra elevations hyperlinks are stored in any order in cells CP -DX

BQ. Correct Category Code
Check against all category related info as well as the business names given (see duplicate ID rows) Check whether additional business names / "finda" categories can be established from street view elevation images. Check using web search. Eg. Google / Yellow Pages etc. If category seems incorrect and/or building does not seem to contain any commercial / industrial uses at all then supply alternate category code using reference list (below). If only a small portion of the building appears to be commercial / industrial then mention this in NOTES (column CO). eg. "CX - Commercial Retail Assumed for Ground + First Floor with 20 Residential floors above" If building contains no commercial or industrial uses then abandon this record altogether + add note.

BR. Business Names
Viewed in street view or found in web search. If multiple businesses can be seen separate names in one cell with commas. You can copy and paste from Column Y if the same businesses have been found by Data Market. Check Data Market address (AB) as sometimes the original QV addresses

have been mis-interpreted (eg. 1 – 178 Lambton Quay is "unit 1" street number 178. where as Data Market has occasionally interpreted this as any street number between 1 and 178.)

BS. (BT-CD) Surrounding Site Dimensions N E S W (for day lighting and thermal modelling)

For each elevation orientation measure:

Average distance to nearest building,

Average height of adjacent buildings (Calculate average height of any daylight obstructions along the whole length of elevation eg. If half the length is 2 storey and half the length is 3 storey then assume eg. 7.5m height)

Average ground level from base of building (eg. This identifies large changes in ground level around a building like those on the terrace / Lambton quay – Lambton Quay elevation would be 0 m, Terrace elevation would be eg. 16 m and side elevations would be eg. 8 m each. Educated guesses are OK. Don't spend too long worrying about this it is just meant as a guide. Stick to the same orientations described for the purposes of glazing ratio.

CE. (CF – CL) Building Fabric

Primary = Majority of building fabric type

Secondary = Other secondary material / type.

More…? = Add comments to NOTES Column CO

You may find clues about building fabric in QV columns AP-AU. These can not be relied on as 100% accurate but it may help make an educated guess if a street view image is unclear.

All cells can be filled in using standard drop down lists of materials for consistency.

CF. Built Form

See diagrams in reference sheet. Best guess is ok. Bold assumptions as always should be noted in column CO. Primary = predominant form type. Secondary = extension or subordinate form type or indicate none dash "-" if building conforms to a single form class.

CM. Sketch Up File (using Open Studio Plug-in)

Set the Google Earth image so that the building is a reasonable size within the aerial photo, but including some of the surrounding sites to aid reference.

In Google Sketch Up click on the 'Get Current View' icon (right) this should bring the Google Earth aerial image into Google Sketch Up.

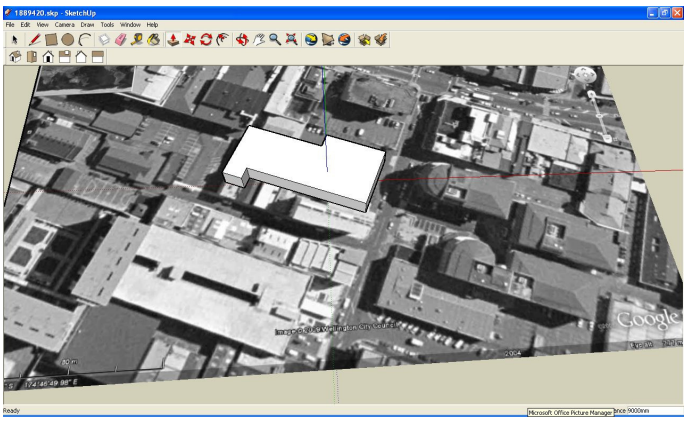
A zone is created using the Open Studio toolbar icons and then this zone is double clicked before building the SketchUp model in the normal way (right). The zone then expands to fit the 3D model built and the surfaces take on different colours (mustard for walls / burgundy for roofs) A secondary energy plus file with extension .idf is created when the SketchUp model is saved

(Open Studio can be downloaded free of charge at http://apps1.eere.energy.gov/buildings/energyplus/openstudio.cfm)

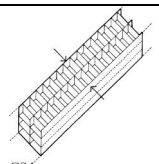
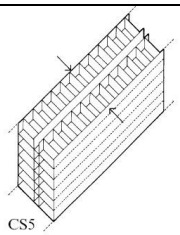
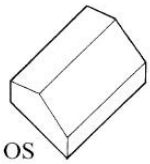
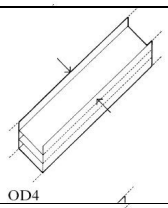
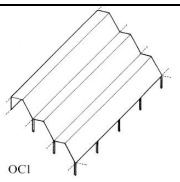
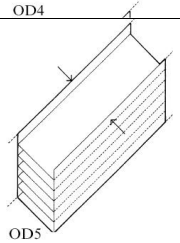
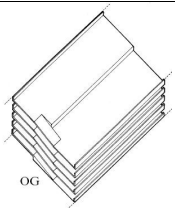
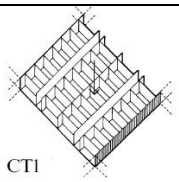
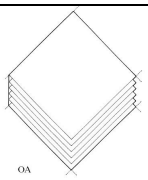
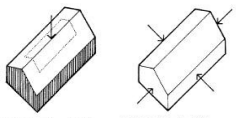
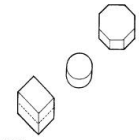
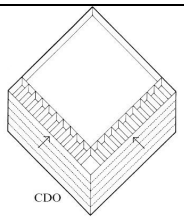
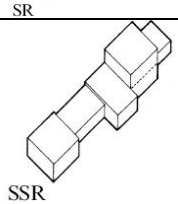
To check the location is correct place the model on Google Earth Click 'Place Model' (above right) This will need to be repeated every time the model needs to be replaced in Google Earth to update the view / location.

If the building has a striking roof form (defined as roof angles with pitch over 30 degrees) this is added on top of the simple block form with minimum 3D details to convey the overall envelope / volume.

The models are highly simplified and ignore details such as porches, façade profiles, curves and small projections. They may include courtyards and basic variety in storey heights. No glazing / openings are required.



REFERENCE SHEET:

FORM	DIAGRAM	FORM	DIAGRAM
]	 CS4	HA - Artificially lit hall	HA (artif. lit)
CS5 - Daylit (sidelit) cellular strip, 5 storeys or more	 CS5	OS - Open-plan space in a single shed	 OS
OD4 - Daylit (sidelit) open-plan strip, 1 to 4 storeys	 OD4	OC1 - Open-plan continuous single-storey space	 OC1
OD5 - Daylit (sidelit) open-plan strip, 5 storeys or more	 OD5	OG - Open-plan car parking or trucking deck	 OG
CT1 - Toplit cellular, single-storey	 CT1	OA - Artificially lit open-plan multistorey space	 OA
HD - Daylit hall, either sidelit or toplit (or both)	 HD (toplit) HD (sidelit)	SR - Single-room form	 SR
CDO4 - Daylit (Sidelit) Cellular strip with Open Plan Space 1-4 storeys CDO5 - Daylit (Sidelit) Cellular strip with Open Plan Space 5 storeys or more	 CDO	SSR - String of single-room forms	 SSR

CATEGORY CODES
CO Office type use
CR Retailing use
CX Numerous commercial uses
IW Industrial Warehouse (assoc. retail maybe incl.)
IS Industiral Service(interface w/ public)
CA Commercial Accomodation
CC Cinema Public Hall
CE Rest Homes
CL Liquor outlets bars
CM Motor Vehicle Sales Service
CP Parking Buildings
CS Service Stations
CT Tourist type attractions leisure
CV Vacant land typically for commercial
IE Educational buildings
IH Heavy Manufacturing
IL Light Manufacturing
IN Noxious industries (power gas fuel)
IV Vacant land typically for industrial
IX other industrial (mixed)
OH Health, Medical + Surgeries
OM Maori Sites
OP Passive Reserves
OR Religious uses
OS Sports grounds etc.
OU Utilities
OV Vacant land not obvious use
OX Other multiple
RB Bare land for subdivision into house sites
RC Converted dwelling now rental flats
RD Dwelling houses - detached or semi
RF Ownership home units (not dwelling houses)
RH Home + Income Additional Unit Associated
RR Rental Flats Purpose Built
RV Vacant Land typically for housing
OA Assembly Halls
OTHER

APPENDIX iii Criteria Benchmark Averages

		BEES SIMULATION MODELS						PREMISE ENERGY BILLS						
Field	Criterion Criteria Websearch	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	Error % +/-	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	+/- error	Difference Models / Bills
ALL RECORDS	ALL	123	48	108	7	16	25%	204	253	214	16	13	13%	60%
Wellington CBD	Te Aro, Piptea, CBD	79	4	43	2	22	53%	293	38	277	6	45	30%	27%
Storey Range	2-3	84	11	82	3	25	57%	305	80	148	9	17	11%	28%
QV Floor Plate	15000+	-	-					303	12	276	3	80	51%	
Contains Use:	H	52	2	55	1	39	148%	295	3	215	2	124	83%	17%
WS Floor Plate	5000>15000	10	1	-	1			293	27	251	5	48	32%	3%
Floor Area WebS.	15000+	152	2	53	1	37	48%	280	37	221	6	36	25%	54%
QV Mats roof	conc	113	5	53	2	24	41%	268	29	228	5	42	31%	42%
Office Quality	D	-						267	4	235	2	117	86%	
Size Typology	Small Tall	109	5	59	2	26	48%	264	35	150	6	25	19%	41%
Skylight Factor N	>=60'	66	4	55	2	27	82%	261	35	252	6	43	32%	25%
Use/ Surface Area	Very tall Office/mix	120	9	50	3	17	27%	261	56	191	7	26	19%	46%
Volume	10000>15000	72	3	59	2	34	93%	258	42	312	6	48	37%	28%
Volume	15000+	146	15	99	4	26	34%	255	107	253	10	24	19%	57%
QV Floor Plate	2000>5000	243	6	223	2	91	73%	254	23	272	5	57	44%	96%
Size Stratum	5	130	8	52	3	18	27%	254	76	244	9	28	22%	51%
Storey WebS.	10	94	5	31	2	14	29%	248	32	209	6	37	29%	38%
Storey Range	10	94	5	31	2	14	29%	248	32	209	6	37	29%	38%
Floor Area QV	350>700	133	6	60	2	24	36%	248	7	133	3	50	40%	54%
Glazing East %	<20%	139	15	95	4	25	35%	247	40	290	6	46	36%	56%
Size Typology	Big Floor Medium	92	9	81	3	27	57%	247	45	359	7	53	42%	37%
Floor Area QV	15000+	152	2	53	1	37	48%	244	34	204	6	35	28%	62%
Climate Zone	5	100	8	56	3	20	38%	241	59	239	8	31	25%	42%
WS Cat Code	Commercial Retail	177	14	162	4	43	48%	241	76	222	9	25	21%	73%
Category Stratum	CR	167	17	152	4	37	43%	241	65	228	8	28	23%	69%
Glazing Ave %	40<60%	114	6	58	2	24	41%	241	37	116	6	19	16%	47%
Glazing South%	<20%	132	23	140	5	29	44%	240	93	195	10	20	16%	55%
QV Mats roof	mix	174	2	21	1	15	16%	240	29	272	5	51	41%	73%
Contains Use:	F	88	4	70	2	35	78%	237	40	198	6	31	26%	37%
QV Floor Plate	5000>15000	94	1	-	1			235	20	225	4	50	42%	40%
WS Cat Code	Commercial Mixed	48	8	39	3	14	57%	233	65	218	8	27	23%	20%
Category Stratum	CX	84	10	73	3	23	54%	228	71	212	8	25	22%	37%
QV Mats	mix	151	12	153	3	44	57%	227	63	169	8	21	18%	66%
Secondary Mat	Present	108	8	89	3	31	57%	226	67	213	8	26	22%	48%
WS Roof Mat	trafficable	68	2	98	1	69	198%	226	33	283	6	49	43%	30%
Sq Rt.Floor Plate	35+	142	12	112	3	32	45%	226	133	227	12	20	17%	63%
Floor Area QV	5000>15000	112	12	49	3	14	25%	226	68	229	8	28	24%	50%
QV Floor Plate	<350	98	7	93	3	35	71%	225	22	214	5	46	40%	44%
Mixed / Single use	Retail	112	10	114	3	36	63%	225	66	234	8	29	25%	50%
#PartyWalls	1	96	6	50	2	21	42%	225	66	207	8	25	22%	43%
#PartyWalls	2-3	84	5	38	2	17	40%	224	46	251	7	37	32%	38%
Floor Area WebS.	5000>15000	105	10	61	3	19	36%	224	55	242	7	33	29%	47%
Glazing North %	20<40%	165	12	173	3	50	60%	224	36	228	6	38	33%	74%
Floor Area QV	<350	112	4	104	2	52	91%	222	16	255	4	64	56%	50%
Glazing Ave %	20<40%	102	14	73	4	20	38%	222	68	195	8	24	21%	46%
Glazing West%	>60%	30	2	28	1	20	130%	221	23	332	5	69	61%	14%
Glazing South%	50=<=60%	155	2	114	1	81	102%	220	29	170	5	32	28%	70%
Office Quality	n/a	122	40	117	6	18	30%	220	189	235	14	17	15%	55%
Ave Site Density	20<=40'	105	19	81	4	19	35%	220	83	240	9	26	24%	48%
Use/ Surface Area	Bsh/med: R/M or F	156	12	177	3	51	64%	220	67	223	8	27	24%	71%
Floor Area QV	2000>5000	200	8	212	3	75	73%	217	58	246	8	32	29%	92%
Glazing Ave %	<20%	146	15	165	4	43	57%	214	86	240	9	26	24%	68%
Ave Site Density	>40'	87	4	31	2	16	35%	213	36	202	6	34	31%	41%
Open Plan / Cellular	Cellular Space	150	16	158	4	39	51%	211	93	191	10	20	18%	71%
# glazing recorded	4	133	28	130	5	25	36%	211	129	226	11	20	19%	63%
Glazing East %	50=<=60%	62	4	23	2	12	37%	211	37	185	6	30	28%	29%
Glazing South%	20<=40%	47	1	-	1			211	20	208	4	47	43%	23%
Size Stratum	4	189	10	182	3	57	60%	210	53	175	7	24	22%	90%
Mixed / Single use	Single use	118	34	89	6	15	25%	210	153	220	12	18	17%	56%
Contains Use:	R	157	23	133	5	28	35%	210	132	204	11	18	17%	75%
Climate Zone	1	148	17	141	4	34	45%	209	94	245	10	25	24%	71%
Climate Zone	4	129	3	98	2	57	86%	208	13	243	4	67	63%	62%
Glazing North %	<20%	124	12	155	3	45	71%	208	44	182	7	27	26%	60%
WS Floor Plate	1000>2000	137	15	144	4	37	53%	207	64	244	8	30	29%	66%
WS Floor Plate	700>1000	92	7	59	3	22	47%	205	33	206	6	36	34%	45%
Glazing North %	40<60%	115	7	48	3	18	31%	203	35	146	6	25	24%	57%
Size Typology	Big Floor Tall	135	4	39	2	20	28%	203	35	206	6	35	34%	66%
# glazing recorded	0	121	9	76	3	25	41%	203	36	230	6	38	37%	60%
Glazing West%	20<=40%	103	8	69	3	24	47%	202	33	176	6	31	30%	51%
# glazing recorded	3	88	5	55	2	25	55%	202	63	214	8	27	26%	44%
Use/ Surface Area	small retail	129	12	80	3	23	35%	202	52	187	7	26	25%	64%
Skylight Factor E	>=60'	123	5	59	2	27	42%	201	35	149	6	25	24%	61%
Mixed / Single use	Non Bees	194	1	-	1			200	16	328	4	82	80%	97%
Size Typology	Big Floor Short	212	8	200	3	71	65%	199	45	200	7	30	29%	106%
Glazing East %	>60%	132	5	65	2	29	43%	198	26	286	5	56	55%	66%
WS Roof Mat	Metal Profile	135	28	128	5	24	35%	197	143	213	12	18	18%	69%
QV Mats	conc	122	3	98	2	56	90%	197	103	203	10	20	20%	62%
Secondary Mat	Not Present	126	40	112	6	18	27%	196	186	214	14	16	16%	64%
Storey WebS.	4-8	115	9	67	3	22	38%	196	51	152	7	21	21%	59%
Storey Range	4-8	115	9	67	3	22	38%	196	51	152	7	21	21%	59%
Size Typology	Small Short	121	15	75	4	19	31%	196	46	212	7	31	31%	62%

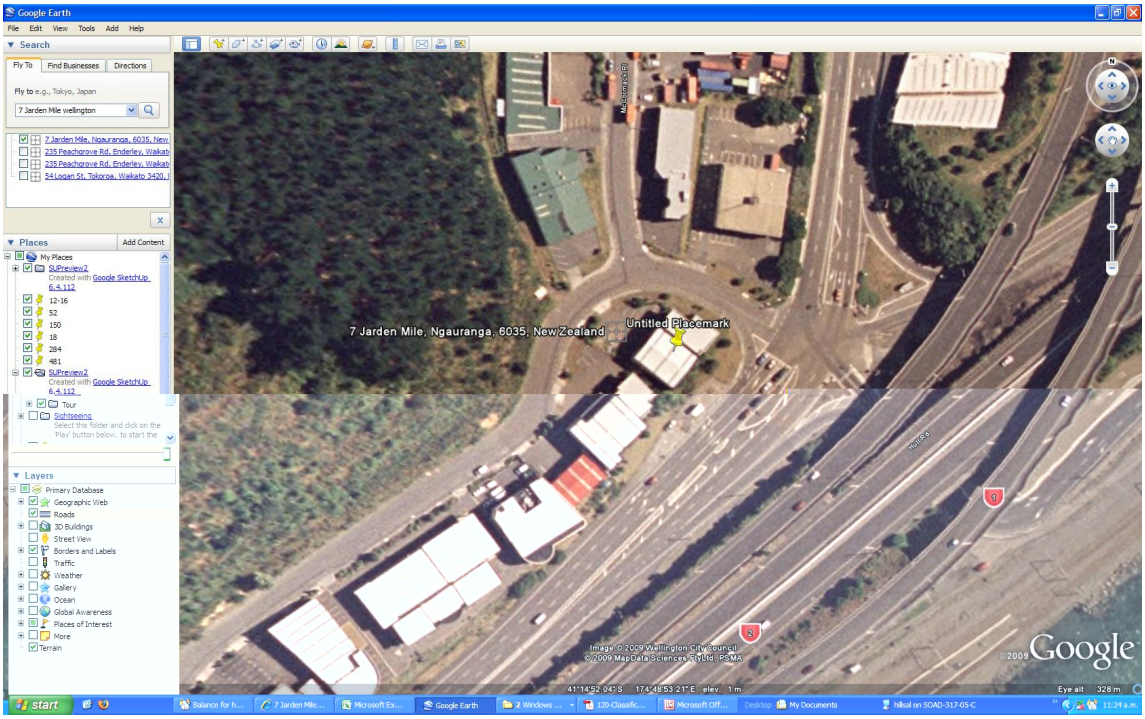
		BEES SIMULATION MODELS						PREMISE ENERGY BILLS						
Field	Criterion Criteria Websearch	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	Error % +/-	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	+/- error	Difference Models / Bills
ALL RECORDS	ALL	123	48	108	7	16	25%	204	253	214	16	13	13%	60%
Wellington CBD	Te Aro, Piptea, CBD	79	4	43	2	22	53%	293	38	277	6	45	30%	27%
Size Stratum	1	117	10	83	3	26	44%	195	34	215	6	37	37%	60%
Skylight Factor N	20<=40'	66	9	40	3	13	40%	195	22	139	5	30	30%	34%
Skylight Factor N	40<60'	139	4	106	2	53	75%	195	22	139	5	30	30%	71%
Mixed / Single use	Mixed use	141	11	161	3	49	68%	195	84	196	9	21	22%	72%
Storey Range	0-1	151	23	135	5	28	36%	194	89	205	9	22	22%	78%
Open Plan / Cellular	Open Plan	89	16	65	4	16	36%	193	53	223	7	31	31%	46%
Glazing East %	20<=40%	133	9	131	3	44	64%	192	41	152	6	24	24%	69%
WS Cat Code	Commercial Office	109	19	63	4	14	26%	192	81	218	9	24	25%	57%
Sq Rt.Floor Plate	<=20m	118	12	81	3	23	39%	192	43	185	7	28	29%	61%
Climate Zone	3	80	4	65	2	33	80%	191	19	118	4	27	28%	42%
Category Stratum	CO	103	20	59	4	13	25%	191	80	237	9	26	27%	54%
Volume	5000>10000	137	9	184	3	61	88%	190	21	190	5	41	43%	72%
Contains Use:	O	101	25	63	5	13	24%	190	133	169	12	15	15%	53%
QV Mats	steel	37	2	4	1	3	15%	189	122	200	11	18	19%	20%
QV Mats roof	steel	133	10	130	3	41	60%	189	122	200	11	18	19%	70%
Skylight Factor S	>=60'	101	3	19	2	11	21%	189	39	134	6	21	22%	54%
Glazing North %	>60%	127	12	60	3	17	27%	188	27	259	5	50	52%	67%
Skylight Factor S	10=>=20'	140	14	116	4	31	43%	188	58	179	8	23	25%	75%
#PartyWalls	0	139	30	126	5	23	32%	187	126	230	11	21	22%	74%
Mixed / Single use	Office	86	9	64	3	21	49%	186	82	201	9	22	23%	46%
WS Floor Plate	2000>5000	224	3	179	2	103	91%	185	26	251	5	49	52%	121%
QV Floor Plate	350>700	120	9	61	3	20	33%	183	41	166	6	26	28%	66%
Volume	<1500	131	10	81	3	26	38%	183	32	174	6	31	33%	72%
QV Floor Plate	1000>2000	112	11	67	3	20	35%	183	56	166	7	22	24%	61%
WS Floor Plate	350>700	91	11	60	3	18	39%	182	60	208	8	27	29%	50%
Glazing Ave %	>60%	122	7	68	3	26	41%	181	24	292	5	60	65%	67%
Climate Zone	6	416	1	-	1			181	14	176	4	47	51%	230%
Sq Rt.Floor Plate	20>35	116	23	122	5	25	43%	180	76	187	9	21	23%	64%
Skylight Factor E	<10'	167	14	160	4	43	50%	179	49	177	7	25	28%	93%
# glazing recorded	1-2	113	3	70	2	40	70%	178	25	190	5	38	42%	64%
Skylight Factor W	>=60'	102	5	47	2	21	40%	176	46	177	7	26	29%	58%
Contains Use:	I	595	1	-	1			175	32	341	6	60	68%	340%
Size Stratum	3	98	10	75	3	24	48%	174	50	148	7	21	24%	56%
Open Plan / Cellular	Complex / Mix	130	9	71	3	24	35%	173	59	254	8	33	37%	75%
Use/ Surface Area	big Office	131	5	107	2	48	72%	172	31	261	6	47	53%	76%
Glazing West%	50=<=60%	155	2	114	1	81	102%	171	27	124	5	24	27%	91%
Office Quality	B	139	4	48	2	24	34%	170	28	124	5	23	27%	82%
WS Floor Plate	<350	135	7	80	3	30	44%	169	31	169	6	30	35%	80%
Skylight Factor W	<10'	171	13	169	4	47	54%	168	52	158	7	22	25%	102%
Climate Zone	7	118	5	71	2	32	52%	167	16	123	4	31	36%	71%
Climate Zone	2	85	10	60	3	19	44%	165	38	202	6	33	39%	52%
QV Floor Plate	700>1000	64	5	56	2	25	77%	165	24	264	5	54	64%	39%
Floor Area QV	1000>2000	113	9	74	3	25	43%	164	34	190	6	33	39%	69%
Ave Site Density	<10'	233	8	186	3	66	55%	161	27	133	5	26	31%	145%
Glazing West%	<20%	154	17	155	4	38	48%	161	68	204	8	25	30%	95%
Skylight Factor E	40<60'	148	4	87	2	43	57%	160	24	119	5	24	30%	92%
Skylight Factor E	20<=40'	96	5	92	2	41	84%	160	24	119	5	24	30%	60%
Floor Area WebS.	1000>2000	177	8	183	3	65	72%	152	34	105	6	18	23%	117%
WS Roof Mat	non trafficable	112	9	60	3	20	35%	151	21	80	5	17	23%	74%
Size Stratum	2	91	10	79	3	25	54%	148	40	267	6	42	56%	62%
Glazing South%	>60%	114	8	49	3	17	30%	147	21	152	5	33	44%	77%
Ave Site Density	10=>=20'	79	10	46	3	14	36%	145	58	152	8	20	27%	54%
Office Quality	A	96	2	26	1	18	37%	142	13	131	4	36	50%	68%
Volume	3000>5000	110	4	68	2	34	60%	140	42	107	6	16	23%	78%
QV Mats roof	tile	216	1	-	1			135	3	72	2	42	60%	159%
Size Typology	Small Medium	66	7	54	3	20	60%	133	47	114	7	17	24%	50%
Skylight Factor N	<10'	209	8	194	3	68	64%	130	43	125	7	19	29%	160%
WS Roof Mat	tile / shingle	53	1	-	1			129	14	125	4	33	51%	41%
Floor Area QV	700>1000	69	7	63	3	24	67%	127	21	345	5	75	116%	54%
Office Quality	C	152	2	10	1	7	9%	126	19	59	4	14	21%	121%
Volume	1500>3000	70	7	55	3	21	58%	123	26	89	5	18	28%	57%
QV Mats roof	fibre cement	13	1	-	1			114	4	37	2	18	31%	11%
Use/ Surface Area	small office	80	9	50	3	17	41%	110	31	96	6	17	31%	72%
Floor Area WebS.	700>1000	75	8	63	3	22	58%	110	24	86	5	17	31%	69%
WS Cat Code	Industrial	117	2	109	1	77	129%	91	28	116	5	22	47%	128%
Category Stratum	IS	194	1	-	1			90	30	113	5	21	45%	214%
WS Cat Code	Other CP CA CV CM	152	4	72	2	36	46%	46	3	15	2	9	38%	335%

		BEES SIMULATION MODELS						PREMISE ENERGY BILLS						
Field	Criterion Criteria Websearch	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	Error % +/-	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	+/- error	Difference Models / Bills
ALL RECORDS	ALL	123	48	108	7	16	25%	204	253	214	16	13	13%	60%
Wellington CBD	Te Aro, Piptea, CBD	79	4	43	2	22	53%	293	38	277	6	45	30%	27%
Floor Area WebS.	15000+	152	2	53	1	37	48%	280	37	221	6	36	25%	54%
QV Mats roof	conc	113	5	53	2	24	41%	268	29	228	5	42	31%	42%
Size Typology	Small Tall	109	5	59	2	26	48%	264	35	150	6	25	19%	41%
Skylight Factor N	>=60'	66	4	55	2	27	82%	261	35	252	6	43	32%	25%
Use/ Surface Area	Very tall Office/mix	120	9	50	3	17	27%	261	56	191	7	26	19%	46%
Storey WebS.	10	94	5	31	2	14	29%	248	32	209	6	37	29%	38%
Storey Range	10	94	5	31	2	14	29%	248	32	209	6	37	29%	38%
Floor Area QV	15000+	152	2	53	1	37	48%	244	34	204	6	35	28%	62%
Climate Zone	5	100	8	56	3	20	38%	241	59	239	8	31	25%	42%
Category Stratum	CR	167	17	152	4	37	43%	241	65	228	8	28	23%	69%
Glazing Ave %	40<60%	114	6	58	2	24	41%	241	37	116	6	19	16%	47%
Contains Use:	F	88	4	70	2	35	78%	237	40	198	6	31	26%	37%
WS Cat Code	Commercial Mixed	48	8	39	3	14	57%	233	65	218	8	27	23%	20%
QV Mats	mix	151	12	153	3	44	57%	227	63	169	8	21	18%	66%
Secondary Mat	Present	108	8	89	3	31	57%	226	67	213	8	26	22%	48%
Mixed / Single use	Retail	112	10	114	3	36	63%	225	66	234	8	29	25%	50%
#PartyWalls	1	96	6	50	2	21	42%	225	66	207	8	25	22%	43%
Floor Area WebS.	5000>15000	105	10	61	3	19	36%	224	55	242	7	33	29%	47%
Glazing South%	50=<=60%	155	2	114	1	81	102%	220	29	170	5	32	28%	70%
Use/ Surface Area	Bsh/med: R/M or F	156	12	177	3	51	64%	220	67	223	8	27	24%	71%
Floor Area QV	2000>5000	200	8	212	3	75	73%	217	58	246	8	32	29%	92%
Ave Site Density	>40'	87	4	31	2	16	35%	213	36	202	6	34	31%	41%
Glazing East %	50=<=60%	62	4	23	2	12	37%	211	37	185	6	30	28%	29%
Size Stratum	4	189	10	182	3	57	60%	210	53	175	7	24	22%	90%
Glazing North %	<20%	124	12	155	3	45	71%	208	44	182	7	27	26%	60%
WS Floor Plate	1000>2000	137	15	144	4	37	53%	207	64	244	8	30	29%	66%
Glazing North %	40<60%	115	7	48	3	18	31%	203	35	146	6	25	24%	57%
Glazing West%	20<=40%	103	8	69	3	24	47%	202	33	176	6	31	30%	51%
Use/ Surface Area	small retail	129	12	80	3	23	35%	202	52	187	7	26	25%	64%
Skylight Factor E	>=60'	123	5	59	2	27	42%	201	35	149	6	25	24%	61%
Size Typology	Big Floor Short	212	8	200	3	71	65%	199	45	200	7	30	29%	106%
Storey WebS.	4-8	115	9	67	3	22	38%	196	51	152	7	21	21%	59%
Storey Range	4-8	115	9	67	3	22	38%	196	51	152	7	21	21%	59%
Size Typology	Small Short	121	15	75	4	19	31%	196	46	212	7	31	31%	62%
Skylight Factor N	20<=40'	66	9	40	3	13	40%	195	22	139	5	30	30%	34%
Skylight Factor N	40<60'	139	4	106	2	53	75%	195	22	139	5	30	30%	71%
Open Plan / Cellular	Open Plan	89	16	65	4	16	36%	193	53	223	7	31	31%	46%
Glazing East %	20<=40%	133	9	131	3	44	64%	192	41	152	6	24	24%	69%
Sq Rt.Floor Plate	<=20m	118	12	81	3	23	39%	192	43	185	7	28	29%	61%
Climate Zone	3	80	4	65	2	33	80%	191	19	118	4	27	28%	42%
Skylight Factor S	>=60'	101	3	19	2	11	21%	189	39	134	6	21	22%	54%
Skylight Factor S	10=>=20'	140	14	116	4	31	43%	188	58	179	8	23	25%	75%
QV Floor Plate	350>700	120	9	61	3	20	33%	183	41	166	6	26	28%	66%
QV Floor Plate	1000>2000	112	11	67	3	20	35%	183	56	166	7	22	24%	61%
WS Floor Plate	350>700	91	11	60	3	18	39%	182	60	208	8	27	29%	50%
Skylight Factor E	<10'	167	14	160	4	43	50%	179	49	177	7	25	28%	93%
Skylight Factor W	>=60'	102	5	47	2	21	40%	176	46	177	7	26	29%	58%
Size Stratum	3	98	10	75	3	24	48%	174	50	148	7	21	24%	56%
Glazing West%	50=<=60%	155	2	114	1	81	102%	171	27	124	5	24	27%	91%
Office Quality	B	139	4	48	2	24	34%	170	28	124	5	23	27%	82%
Skylight Factor W	<10'	171	13	169	4	47	54%	168	52	158	7	22	25%	102%
Ave Site Density	<10'	233	8	186	3	66	55%	161	27	133	5	26	31%	145%
Skylight Factor E	40<60'	148	4	87	2	43	57%	160	24	119	5	24	30%	92%
Skylight Factor E	20<=40'	96	5	92	2	41	84%	160	24	119	5	24	30%	60%
Floor Area WebS.	1000>2000	177	8	183	3	65	72%	152	34	105	6	18	23%	117%
WS Roof Mat	non trafficable	112	9	60	3	20	35%	151	21	80	5	17	23%	74%
Ave Site Density	10=>=20'	79	10	46	3	14	36%	145	58	152	8	20	27%	54%
Volume	3000>5000	110	4	68	2	34	60%	140	42	107	6	16	23%	78%
Size Typology	Small Medium	66	7	54	3	20	60%	133	47	114	7	17	24%	50%
Skylight Factor N	<10'	209	8	194	3	68	64%	130	43	125	7	19	29%	160%
Office Quality	C	152	2	10	1	7	9%	126	19	59	4	14	21%	121%
Volume	1500>3000	70	7	55	3	21	58%	123	26	89	5	18	28%	57%
QV Mats roof	fibre cement	13	1	-	1			114	4	37	2	18	31%	11%
Use/ Surface Area	small office	80	9	50	3	17	41%	110	31	96	6	17	31%	72%
Floor Area WebS.	700>1000	75	8	63	3	22	58%	110	24	86	5	17	31%	69%

		BEES SIMULATION MODELS						PREMISE ENERGY BILLS							
Field	Criterion Criteria Websearch	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	Error % +/-	Average Energy Intensity	Sample	Stdev	√ Sample#	Standard Error	+/- error	Difference Models / Bills	Natural Breaks
ALL RECORDS	ALL	123	48	108	7	16	25%	204	253	214	16	13	13%	60%	
Wellington CBD	Te Aro, Piptea, CBD	79	4	43	2	22	53%	293	38	277	6	45	30%	27%	
QV Mats roof	conc	113	5	53	2	24	41%	268	29	228	5	42	31%	42%	4
Size Typology	Small Tall	109	5	59	2	26	48%	264	35	150	6	25	19%	41%	3
Use/ Surface Area	Very tall Office/mix	120	9	50	3	17	27%	261	56	191	7	26	19%	46%	13
Storey WebS.	10	94	5	31	2	14	29%	248	32	209	6	37	29%	38%	4
Floor Area QV	15000+	152	2	53	1	37	48%	244	34	204	6	35	28%	62%	3
Category Stratum	CR	167	17	152	4	37	43%	241	65	228	8	28	23%	69%	4
Contains Use:	F	88	4	70	2	35	78%	237	40	198	6	31	26%	37%	4
QV Mats	mix	151	12	153	3	44	57%	227	63	169	8	21	18%	66%	1
Secondary Mat	Present	108	8	89	3	31	57%	226	67	213	8	26	22%	48%	1
Mixed / Single use	Retail	112	10	114	3	36	63%	225	66	234	8	29	25%	50%	5
Use/ Surface Area	Bsh/med: R/M or F	156	12	177	3	51	64%	220	67	223	8	27	24%	71%	3
Floor Area QV	2000>5000	200	8	212	3	75	73%	217	58	246	8	32	29%	92%	7
Size Stratum	4	189	10	182	3	57	60%	210	53	175	7	24	22%	90%	8
Use/ Surface Area	small retail	129	12	80	3	23	35%	202	52	187	7	26	25%	64%	3
Size Typology	Big Floor Short	212	8	200	3	71	65%	199	45	200	7	30	29%	106%	3
Storey WebS.	4-8	115	9	67	3	22	38%	196	51	152	7	21	21%	59%	0
Size Typology	Small Short	121	15	75	4	19	31%	196	46	212	7	31	31%	62%	4
Sq Rt.Floor Plate	<=20m	118	12	81	3	23	39%	192	43	185	7	28	29%	61%	9
QV Floor Plate	350>700	120	9	61	3	20	33%	183	41	166	6	26	28%	66%	0
QV Floor Plate	1000>2000	112	11	67	3	20	35%	183	56	166	7	22	24%	61%	9
Size Stratum	3	98	10	75	3	24	48%	174	50	148	7	21	24%	56%	34
Volume	3000>5000	110	4	68	2	34	60%	140	42	107	6	16	23%	78%	7
Size Typology	Small Medium	66	7	54	3	20	60%	133	47	114	7	17	24%	50%	10
Volume	1500>3000	70	7	55	3	21	58%	123	26	89	5	18	28%	57%	9
Use/ Surface Area	small office	80	9	50	3	17	41%	110	31	96	6	17	31%	72%	-

APPENDIX iv Mail Merge Websearch Checklist

«ID»*«ID_Extra_Buildings_Add_Row»* **WEB SEARCH DATA:**



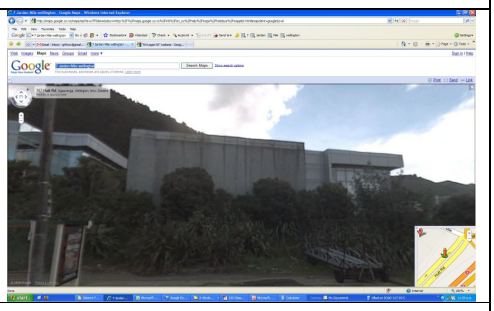
Existing QV Data:		**Additions / Corrections:**
QV Address:	*«Address», «SuburbTown», «TownDistrict»*	
Floor Area Total:	*«FloorArea»*	
Floor Plate Area (where avail.):	*«Floor_Plate_Land_Coverage»*	
Category Code:	*«CatCode»*	
Building Age:	*«Building_Age»*	
Building Fabric:	*«Building_Fabric_QV»*	
Roof Construction:	*«Roof_Construction_QV»*	

Data Market Info:		**Additions / Corrections:**
Business Name:	*«business_name»*	
"Finda" Category:	*«finda_category»*	
No. Employees:	*«employees»*	

Web Search Info Building

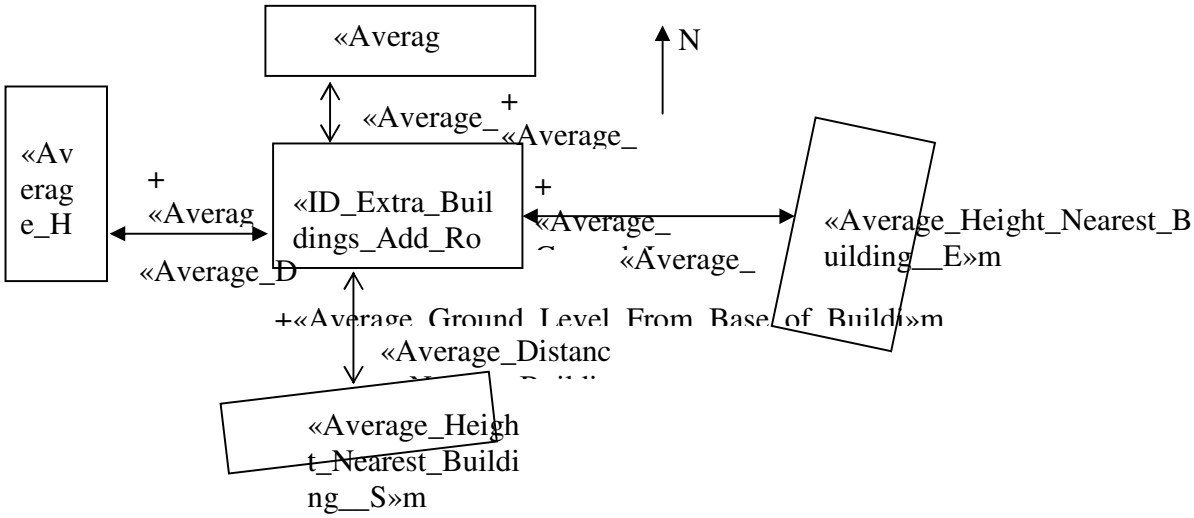
"*«ID_Extra_Buildings_Add_Row»«ID_Extra_Buildings_Add_Row»*": **Additions / Corrections:**

Building ID:	*«ID_Extra_Buildings_Add_Row»*	

Floor Plate Area:	*«Floor_Plate»*	
# Storeys:	*«M__of_storeys»*	
Average Building Height (m)	*«AVE_Building_Height_m»*	
Floor Area Calc:	*«Floor_Area_FORMULA»*	
Floor Area Est.:	*«Floor_Area_Estimate»*	
Category:	«Correct_Category_Code»	
Business Names:	«Business_Names»	
Glazing% - North	«M__Glazing__North»	
Elevation - North	«Elevation_North_HYPERLINK»	
Glazing% - East	«M__Glazing__East»	
Elevation – East	«Elevation_East_HYPERLINK»	
Glazing% - South	«M__Glazing__South»	
Elevation – South	«Elevation__South_HYPERLINK»	
Glazing% - West	«M__Glazing__West»	
Elevation – West	«Elevation__West_HYPERLINK»	
Building Fabric:	*«Building_Fabric»*	
Building Fabric Secondary:	*«Building_Fabric__Secondary»*	
Roof Construction:	*«Roof_Construction»*	
Roof Construction Secondary:	*«Roof_Construction__Secondary»*	

Glazing Type:	*«Glazing_Type__Primary»*	
Glazing Type Secondary:	*«Glazing_Type__Secondary»*	
Built Form:	*«Built_Form__Primary»*	
Built Form Secondary:	*«Built_Form__Secondary»*	
Notes:	*«Notes»*	
Extra Elevations :	*«Street_View_Extras_1»*	
	«Street_View_Extras_2»	
	«Street_View_Extras_3»	
	«Street_View_Extras_4»	
	«Street_View_Extras_5»	

Surrounding Building + Ground Heights (illustration only not to scale)



NB 'mail merge' document displays Websearch data field names which return values from the specific Websearch database when linked eg. distances and heights above (m)*