USING MULTIPLE LINEAR REGRESSION TO DISAGGREGATE ELECTRICITY CONSUMPTION FOR CLUSTER-METERED ACADEMIC BUILDINGS

By

Nicholas So,

Bachelor of Science,

The University of Western Ontario, London, Ontario

2010

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science

in the Program of

Building Science

Toronto, Ontario, Canada, 2014

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

ABSTRACT

Ryerson University does not have a means to gauge electricity consumption for half of their campus buildings. The installation of utility meters is outside of the University's budget, a situation that may be similar across other academic institutions. A multiple linear regression approach to estimating consumption for academic buildings is an ideal tool that balances performance and utility. Using 80 buildings from Ryerson University and the University of Toronto, significant building characteristics were identified (from a selection of 18 variables) that show a strong linear relationship with electricity consumption. Four equations were created to represent the diversity in size of academic buildings. Tested using cross-validation, the coefficient of variation of the RMSE for all models was 33%, with a range of error between 20% and 43%. The models were highly successful at modeling electricity consumption at Ryerson University with an average error of 14.8% for five building clusters. Using metered data from each cluster, raw estimates for individual buildings were adjusted to improve accuracy.

ACKNOWLEDGMENTS

Firstly, I'd like to offer my gratitude to my supervisor, Russell Richman, for presenting this research opportunity to me. In addition, his continued interest and support has been appreciated throughout this journey.

I'd like to thank staff members from Ryerson University and the University of Toronto for their efforts and patience in supplying the necessary data for analysis. Specifically, I'd like to thank Tonga Pham, John Fauquier, and Kevin Leong for their attentiveness in addressing my needs and concerns over the past two years.

Marcel Schweiker's influence on this project has been greatly understated. I'd like to specifically thank him for introducing and equipping me with the tools required to tackle my research questions. His dedication and insight into research methods pertaining to R were both invaluable and inspiring.

To my good friends and fellow classmates Deniz Ergun, Peta-Gaye Ebanks, Amanda Yip, and Joseph Earle: our friendship may have been initiated over our shared anxiety and stress levels about building science, but our relationship only grew because of our continued support for one another (and maybe our weakness to be lured by baby animal videos). Thank you for aiding me in my feeble attempts at a healthy work/life balance.

Lastly and most importantly, I'd like to thank my parents for their unwavering confidence in my ability to accomplish my goals. They were, and continue to be my foundation on which I am able to pursue my dreams and passions – a position I hope to be for them in the future.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

I. INTRODUCTION

Between 1990 and 2010, overall electricity consumption for commercial and institutional buildings has increased by 32.5%. While this demand is understandable with an increasing population and a shift to a more service-oriented economy, the greenhouse gas intensity (the amount of GHGs release per unit of energy produced) has only decreased by 5.8% over that same time period [1]. The reductions in air pollutants are thus a result of cleaner fuel sources and not a decrease in energy demand – something that needs to be addressed when striving to create sustainable long-term communities. Commercial buildings account for 37% of the Greater Toronto Area's (GTA) total electricity consumption, and one third of the greenhouse gases (GHGs) emitted [2, 3]. Therefore, focusing on the energy consumption in buildings is crucial when trying to reduce overall energy demand. In particular, attention should be given to existing buildings because they comprise 98% of Canada's (commercial) building stock – 78% of which have been built before 1989 [4, 5]. The level of impact buildings have on resource consumption is often difficult to quickly assess; the true scale of cities is often overlooked, but with a change of perspective (via aerial/satellite imagery), it is easier to accept how a city such as Toronto required just under 30 tWh in electricity alone in 2006 to power its buildings and homes [6]. While newly constructed buildings receive constant attention from the public for their sustainable and resource efficient designs, adjacent existing buildings are neglected, receiving little to no support for crucial retrofits throughout their lifetime. In Canada, only half of all commercial and institutional buildings have undergone any form of retrofitting in their lifespan [5]. A look at the Canadian Green Building Council's registered and certified LEED projects [7] will show the stark contrast between the number and area of the projects that pursue certification under New Construction (2,669) and those under Existing Buildings (353). The Toronto Green Standards, the City's guideline on sustainable building practices, fails to address the performance of existing buildings completely, opting to focus exclusively on the small fraction of buildings that are added each year to the roughly 3,000 non-residential buildings (nearly 2 million, including private households) already built within the city boundaries [8, 9].

Another reason to focus efforts on the existing building stock is because of their immense potential for reducing carbon emissions. Current efforts to reduce emissions are costly such as harvesting renewable energy and carbon sequestration. Not only are these exercised options expensive, but greater impacts can be made with reducing energy consumption [10]. The majority of the building sector's ability to mitigate emissions in 2030 has been estimated to come at little to no cost to owners, on a dollar per ton of CO_2 equivalent basis [11]. Figure 1 shows an underestimated value (omitting the effects from non-technological options, such as demand-side management, on further reduction in emissions) of the mitigation potential for buildings compared to other sectors. A large share of the mitigation potential from the building sector is categorized as a net-negative cost because of the relatively short payback period of many of the explored options (e.g. efficient lighting and appliance upgrades). As illustrated by Ürge-Vorsatz and Novikova [11], allowing existing buildings to merely exist without continuous renewal and upkeep is one of the largest opportunities missed for greater environmental stewardship and return on investment.

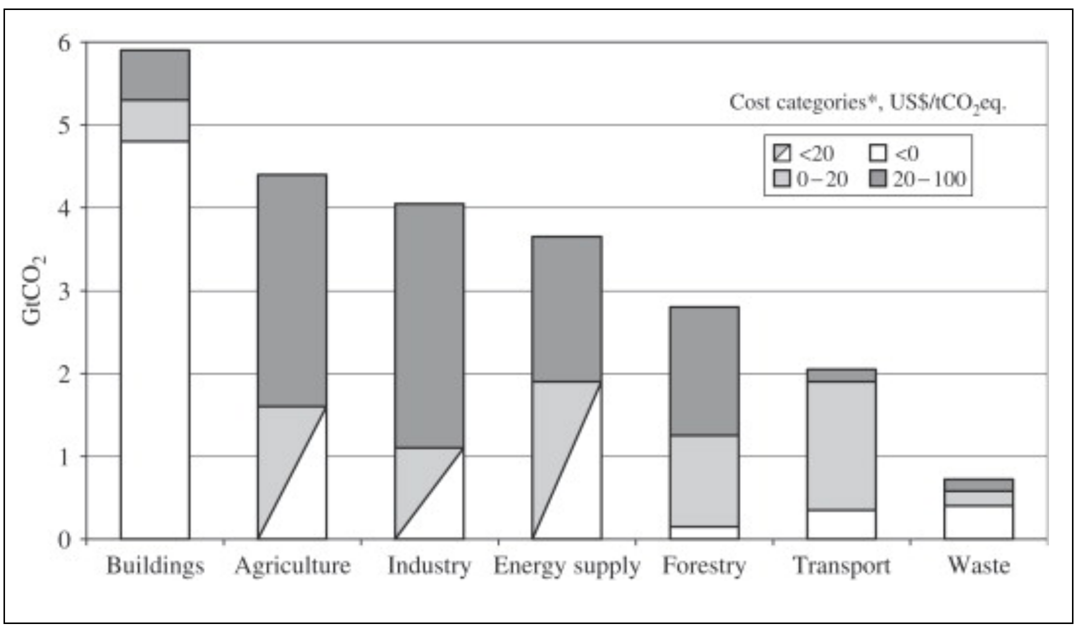


Figure 1 Mitigation potential for various sectors and their costs. [11]

Today, buildings in Ontario are constructed with smart meters in place, measuring and reporting consumption data that still needs to be manually read in many existing buildings.

This evolution of metering technology has substantially increased the data available, and subsequently, the understanding of usage patterns for electricity and other utilities within buildings – short and long term benefits of which are explored by the Energy Efficiency Office [6] and Briones et al. [12]. While Ontario has successfully installed smart meters for measuring electricity usage in most of its residential homes, other parts of the world are struggling with their widespread adoption [13]. Figure 2 shows the current penetration rates of smart meters around the world and their projected growth by 2016 [14]. While there are multiple ways to measure utility consumption for buildings, smart meters provide a reliable and easily assessable means to access current and historic consumption.

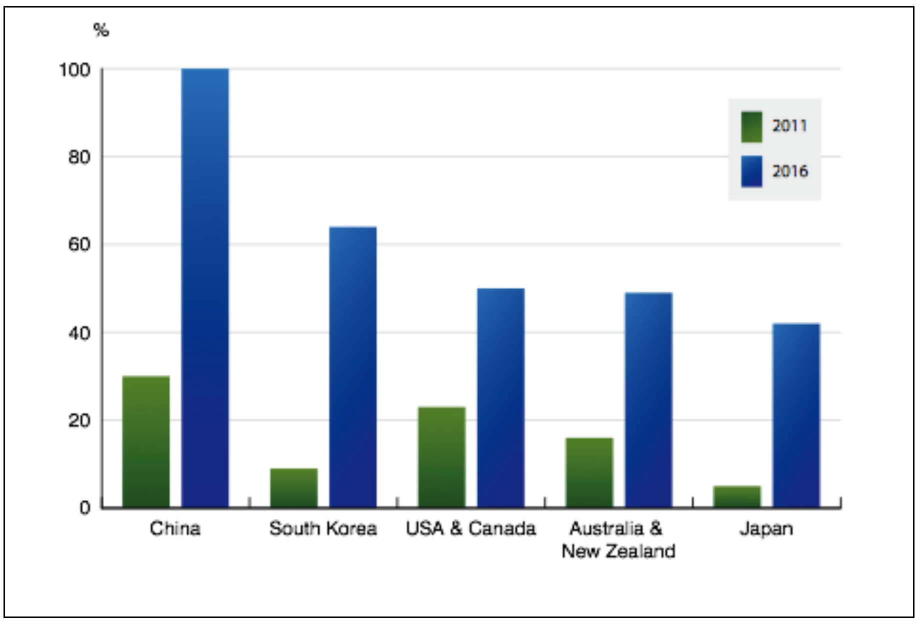


Figure 2 Penetration of smart meter technology in international markets. [14]

Reducing the consumption of electricity in buildings can be accomplished by various means. Control can be implemented through cost-demand management strategies where uniform or variable increases in utility costs throughout the day result in their reduced usage by customers. Alternatively, a mandate or legislature can be introduced to incorporate elements of LEED, Toronto Green Standard, or other local or international green standards into enforced provincial building codes – the Passive House standard is currently being used in building regulations in Spain, Belgium, and most commonly, in Germany [15]. These are only two of many options that are available to reduce electricity consumption. However, even with these readily available solutions, many buildings across Canada are constructed and operated without implementing such standards. In 2010, 55.2% of all commercial buildings surveyed across Canada had not implemented any energy efficiency features [5]. This proportion slightly decreased around the Great Lakes region (Quebec-Windsor corridor) to 48%. In addition, only 23% of surveyed buildings had an energy awareness program in place for building occupants. Both the perceived level of incentive by building owners and the lack of motivation from building occupants due to inadequate knowledge about energy issues reinforce existing operations and impedes change. Benchmarking buildings, another tool to assess electricity consumption within the city, can tackle both these critical issues. The benefits of benchmarking are twofold: (1) buildings of a certain type within a city can be compared to one another to determine which of them are performing poorly, and (2) buildings of a certain type and/or the whole stock of existing buildings can be compared to other cities to assess the potential for savings - particularly useful when establishing case studies. Benchmarking buildings is an important if not mandatory first step to addressing electricity consumption in the existing building stock.

Progress for compiling a comprehensive benchmark for North American cities is in its infancy stages. The United States has been working on benchmarking their existing building stock and publishing these results to further public awareness. At the national level, Energy Star, a joint program between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Energy (DOE), has been running a tool for several years called the "Portfolio Manager". It allows building operators and owners to monitor and assess their utility usages (electricity and water), create baselines for tracking facility improvements, gain recognition from the EPA and qualify for an Energy Star rating, and most importantly, it allows the (voluntary) sharing of data with other participants. To date, there are more than 40,000 participants with over 250,000 registered buildings in the United States [16]. The success of the platform has led to its adoption in Canada. Natural Resources Canada has been working closely with the EPA since 2011 to adapt its Portfolio Manager for Canadian use. The tool, which was initially released in

August 2013, offered benchmarking for office buildings and K-12 schools – it has since added hospital type buildings that can be benchmarked with the tool [17]. The Portfolio Manager is not without its faults; there is omission of university and college-type buildings in the current system. A separate program, the Sustainability Tracking, Assessment and Rating System (STARS), which is run by the Association for the Advancement of Sustainability in Higher Education, focuses on benchmarking university and college buildings in North America. Schools may choose to participate in a lengthier review and rating process under the STARS program or they may opt to report their utility data through the Campus Sustainability Data Collector. Both paths will result in the collection of utility data however only STARS participants will have access to the database. As of September 2014, 358 institutions have participated in the stars program with 35 of them being Canadian. A pilot has been started for 8 international schools to participate in STARS [18].

Tools and rating systems aside, laws and mandates have been enacted in recent years across North America requiring the mandatory disclosure of energy use, making city-wide benchmarks possible. The first two cities in the United States to introduce a disclosure law were Washington, DC and Austin in 2008; between 2009 and 2013, seven other cities joined the ranks of cities mandating energy disclosure. While the practice has yet to be standardized, the scope of buildings benchmarked often include commercial and multi-unit residential buildings that are greater than 1,000 m^2, at a minimum, and 23,000 m^2 at a maximum – depending on the city and building type [19]. In Canada, and specifically in Ontario, all public bodies have been required to report energy consumption for city-owned buildings since 2011 – during that year, just under 500 buildings were benchmarked in Toronto, none of which were academic buildings [20].

A brief profile of Toronto and Ryerson University is provided below which will give context to this thesis. Future researchers and users should find this helpful when applying the methods used here to other urban environments and climatic conditions.

A. Toronto

Over 80% of Canadians live in cities – cities that now consume more than 50% of all energy in Canada [6]. Toronto, Canada's most populous city, has 2.6 million inhabitants and a density of 945.4 people/km^2 in 2011 [9]. First established in the late 18th century, Toronto is a relatively young city located in the heavily industrialized Golden Horseshoe region of Southern Ontario. Due to its proximity to Lake Ontario, Toronto's climate is moderated with average daily temperatures peaking in July (22.2˚C), dipping in January (-4.2 ˚C) and averaging 9.2 ˚C for the entire year; annual precipitation is averaged at 834 mm for the region [21]. Climate normals (1971-2000) for the city can be seen in Appendix A1. Under the Köppen climate classification, the city is categorized as having a warm summer, humid continental climate, and is grouped with other cities such as Warsaw, Stockholm, Vienna, and Moscow [22]. However, unlike those cities mentioned, Toronto has one of the highest per capita electricity consumption of the developed world (Figure 3) at 10,000 kWh/capita in 2011 [23]. The City spent over $4.45 billion dollars on energy alone in 2005 – 60.4% of which was allocated to electricity costs [6]. The air emissions associated with generating electricity for commercial and industrial buildings in 2004 can be seen in Table I [24]. These emissions have contributed towards the increased occurrence of smog in the City from under 5 days in 1993 to more than 35 in 2007 [25]. The environmental effects of operating buildings are diverse and can often play a significant role in determining the quality of life in the urban areas they reside. However, because Toronto's building stock is one of the largest in Canada, it also has the greatest potential for energy savings in Canada if measures are implemented correctly.

Table I Air pollutants released in Toronto in 2004 from the generation of electricity

Air Pollutant	**Tonnes released in 2004**
Nitrogen Oxides (NOX)	7,010
Volatile Organic Compounds (VOCs)	99
Total Particulate Matter (TPM)	1,267
Carbon Monoxide (CO)	2,056
Sulfur Oxides (SOX)	16,709
Greenhouse Gases (GHGs)	4,913,000

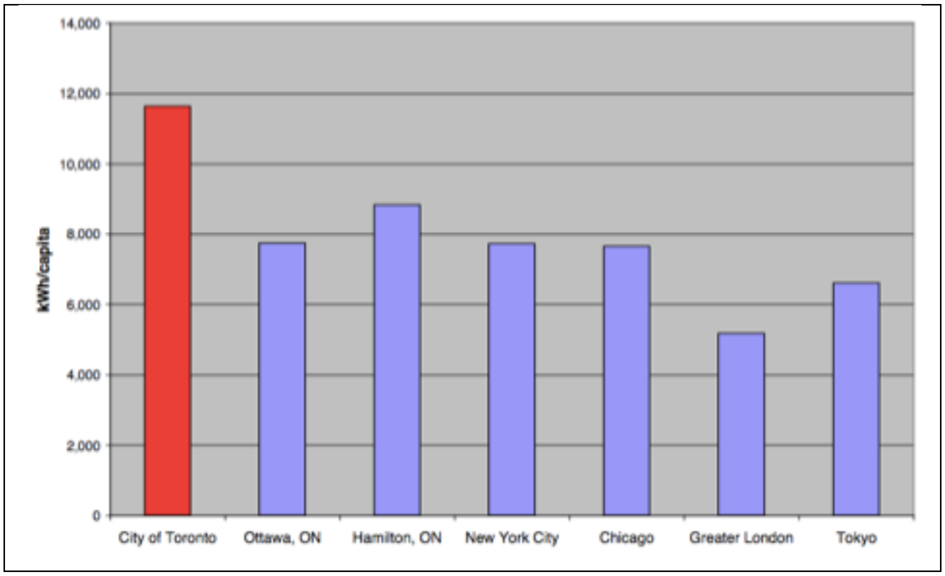


Figure 3 Total electricity use per capital in selected cities [23]

1) Ryerson University: Ryerson University is a publicly funded university located at the core of downtown Toronto. First established in 1948 as the Ryerson Institute of Technology, the relatively large Canadian university currently houses 38,000 undergraduate students, 2,300 graduate students, 1,700 administrative staff, and 780 faculty members [26]. Its single urban campus covers 8.5 ha and has a total of 30 buildings, providing an enclosed area of 269,384 m^2 (building specifications and details can be seen in Appendix A2). As of 2010, the average age of buildings on campus is 38.5, two years below the provincial average for university buildings [27]. In 2013, the university reported 52,383,677 kWh of electricity consumption for the year and emitted 12,847,605 Kg of GHG emissions relating to energy usage [28].

B. Statement of Problem

The emergence of building performance benchmarking tools is critical to the understanding and eventual reduction of electricity consumption in our communities. Unfortunately, before academic institutions such as Ryerson University can participate in such programs, an accurate measure for campus building utilities must exist for individual buildings. This missing link is not unique to academic institutions or even to properties in Ontario (Figure 4). The U.S. Energy

Information Administration has reported that approximately half of the sampled building respondents (3,600 individual buildings in 2003) could not provide the required energy usage data for the completion of the Commercial Buildings Energy Consumption Survey (CBECS) [29]. During a personal interview with Ryerson's Director of Campus Facilities and Sustainability, Tonga Pham, she explained that the installation of new utility meters and energy management systems place a heavy financial burden on building owners; with limited financial resources, many properties forgo this necessary service.

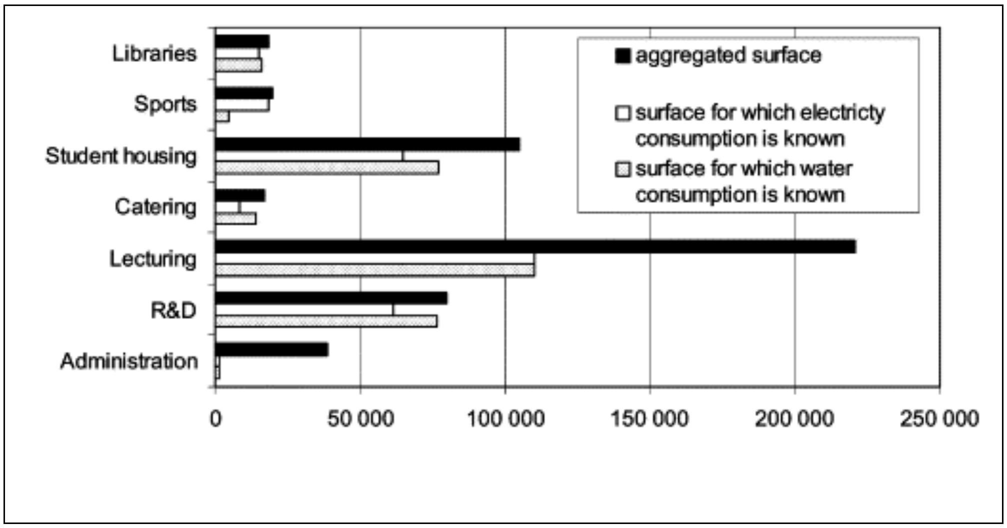


Figure 4 Area (m^2) of known utility consumption by activity at the University of Bordeaux. [30]

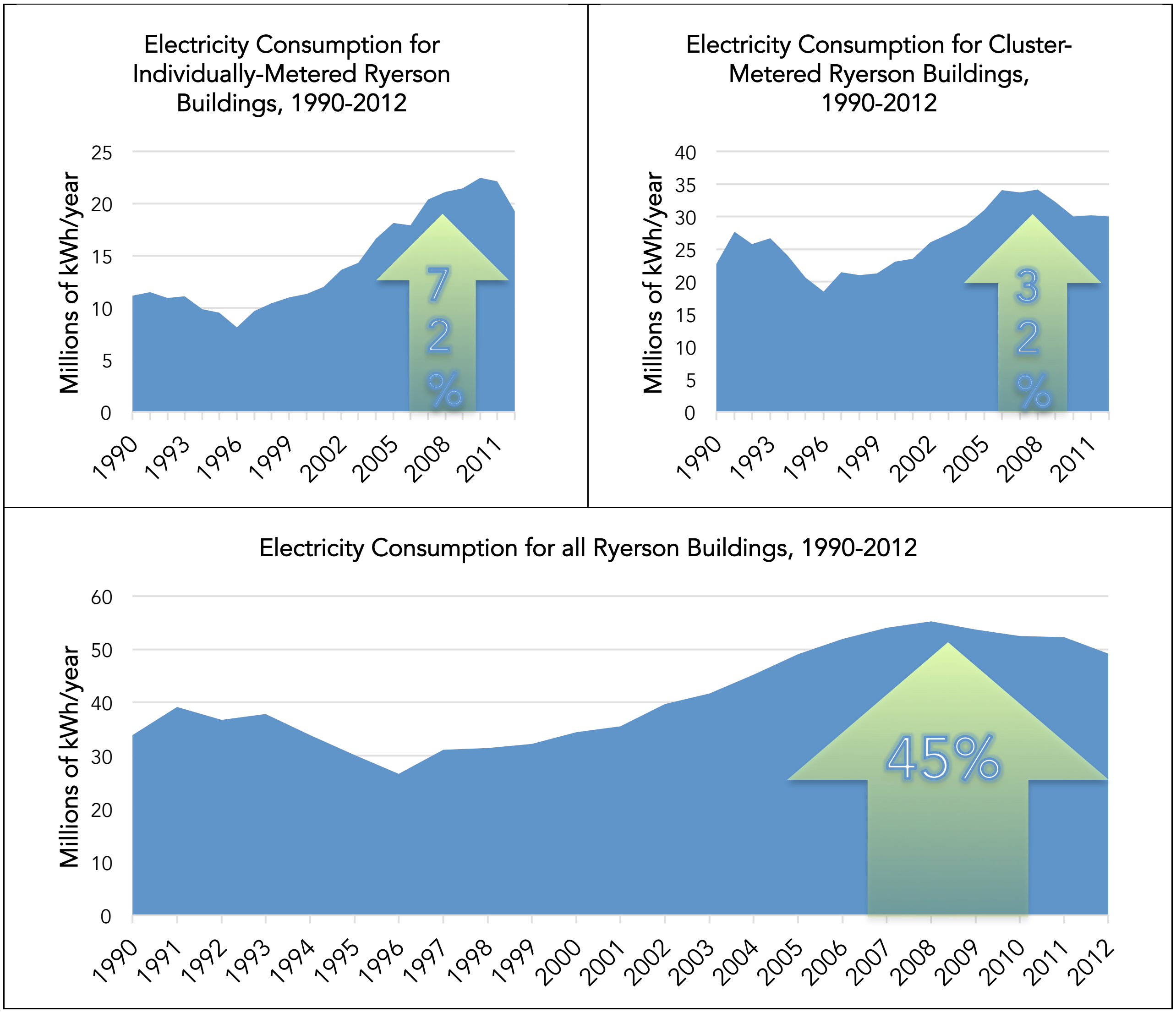


Figure 5 Historic electricity consumption for individually-metered, cluster-metered, and all buildings at Ryerson University.

Ryerson University currently does not have insight into the electricity consumption of its buildings on an individual basis. With increasing energy consumption (Figure 5) and associated costs for these utilities across Canada, greater focus on energy efficiency on campus is needed. When considering the effects of the addition and removal of buildings on campus since 1990, the trends differ to those shown in Figure 5; individual buildings use the same amount of electricity per unit area while buildings metering in a cluster and all Ryerson buildings see an increase of 17% and 8% respectively. Of the 32 buildings or spaces on campus, 14 share a meter with two or more buildings or spaces (i.e. energy consumption for these buildings is

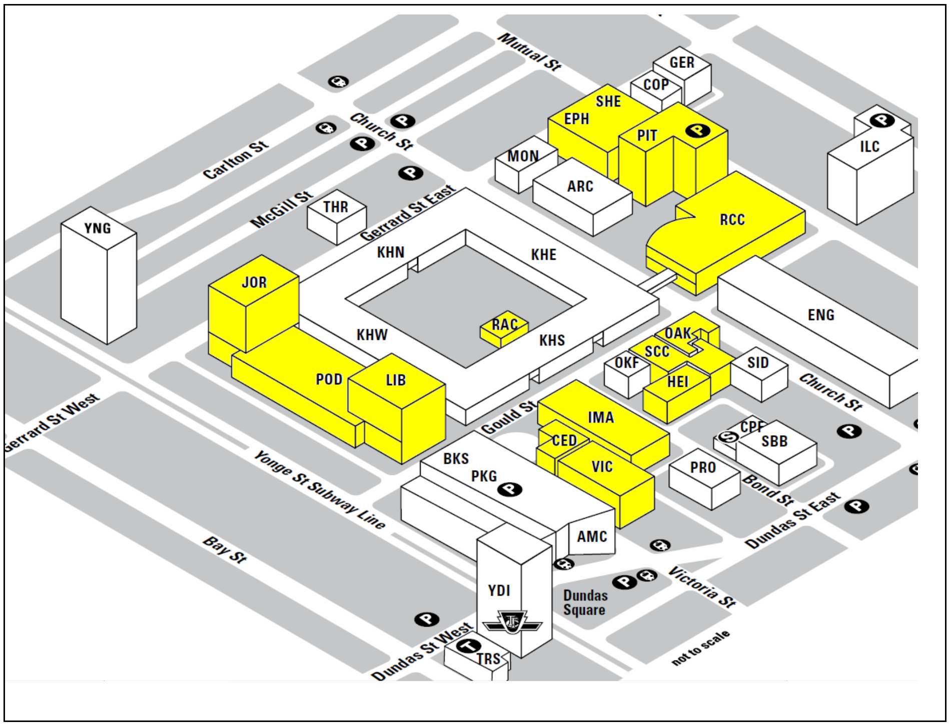


Figure 6 Ryerson University campus map. Highlighted buildings are those sharing an electricity meter with one or more buildings. Adapted from [31]

unknown on an individual level) which are highlighted on a campus map in Figure 6. According to Shiv Tangri, supervisor of Utility Management at Ryerson, where a more detailed level of reporting is required, a weak, unscientifically based estimation is made using an undisclosed method by Ryerson University. At the University of Massachusetts Amherst, a similar scenario is present where the number of utility meters is less than the number of buildings on campus. However, McCusker [32] was able to benchmark 84% of the total built area using available data in order to gauge energy performance for the university. In contrast, Ryerson's access to individual meters on campus represents 50% of the total built area – responsible for 41% of all electricity consumed on campus in 2012 (Figure 7). In order to provide the University with detailed and accurate data on the energy performance of its buildings, few options exist. The installation of energy meters for Ryerson buildings carry high upfront costs (i.e. 2 million dollars) and is outside the budget of the University at present, as communicated by Tonga Pham. A budget for installing meters campus-wide compiled by staff at the university can be seen in Appendix A3. Another option, building traditional energy models, is time consuming and costly for complex buildings which again, may not be an optimal solution for the University. A form of energy estimation based off of characteristic building variables is a viable option to gauge energy performance for Ryerson's campus buildings and should be fully explored.

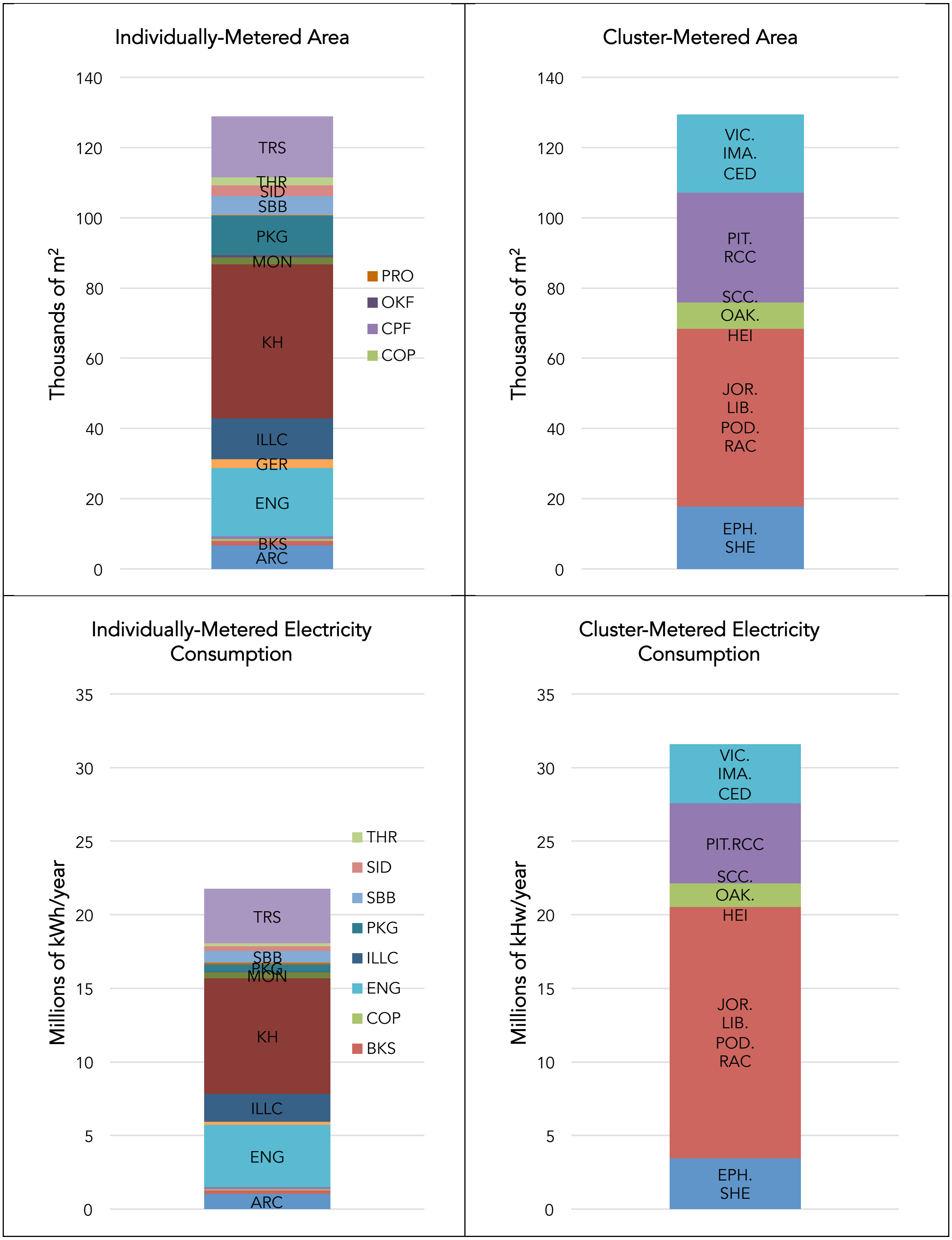


Figure 7 Internal area and electricity consumption in 2012 that is individually and cluster-metered, by building. Buildings with a small relative proportion are listed to the side.

C. Objectives

This thesis aims to establish a methodology for creating a series of equations used to estimate electricity consumption in academic buildings, based on variables relating to space usage and built form. The reason why electricity is the focus of this thesis is because it comprises the largest share of total energy consumption at the university (58% in 2012), and because it is the only sub-metered utility on campus that can provide the data necessary for the proposed methodology. These equations will address two prohibitive characteristics of other existing options: high costs and detailed knowledge of building science theory. The costs associated with implementing energy saving building features have long been a major hurdle to overcome in the construction industry. Those who do not wish to undertake such an expensive project often consider tools such as building energy models to assess their current and possible future consumption patterns in order to establish a case for action. However, when financial resources are limited, and future return on investment is obscure, the risks may be too great for building owners/managers of existing buildings to take. Closely related to costs is the issue of lack of knowledge. Energy modeling programs and building science knowledge is specific and takes years of experience to be able to make confident, informed decisions. This can intimidate decision makers, forcing them to back away from energy-saving retrofits. For owners who have access to larger budgets, this issue is remedied by hiring skilled professionals, however many cannot afford this luxury. While the use of an equation to gauge electricity consumption has the potential to minimize costs and be easy to understand, it must also be reliable and accurate. A delicate balance must be achieved between accessibility and accuracy to maximize the benefits to the user.

1) Research Questions: Using a database comprising of campus buildings from the University of Toronto and Ryerson University, annual electricity consumption will be correlated with selected variables to form a multiple linear regression that will accurately model consumption in large, multiuse buildings. In doing so, this work intends to address the following primary research question:

1. To what degree of accuracy, can the use of predictor variables be in estimating electricity consumption for academic buildings?

In doing this, the following secondary research questions will also be answered:

1. Does this approach offer a realistic alternative to installing meters for individual buildings for Ryerson University; what are the tradeoffs in cost and accuracy between the proposed method and existing alternatives?
2. Are there certain building typologies that work better or worse with this method? Do they pose real limitations when implemented?

D. Long-Term Goal & Vision

The full scope of this research program is large and is intended to be completed over several years. First and foremost, the end goal of this project will be to provide a tool for Ryerson Campus Facilities and Sustainability that will allow them to assess their current electricity consumption based on the variables required by the equation. Upfront cost savings aside, the results of this research will greatly increase the effectiveness of efforts made by Campus Facilities and Sustainability to reduce electricity consumption on campus. With greater insight into the specific electricity consumption for individual buildings, a benchmark can be created among campus spaces that can help identify underperforming buildings. Resources can then be diverted to those specific spaces/buildings in order to maximize the return on investment for campus sustainability initiatives as well as ensure all buildings meet the minimum performance levels. From there, a comprehensive energy sustainability plan can be created for the university, creating a strategic framework that will guide short and long-term development plans on campus, considering budgetary constraints and high priority issues.

E. Thesis Structure

The sections following the Introduction of this thesis include the Literature Review, Methodology, Model Development, Results, Discussion, and Conclusion. The Literature Review will introduce research relating to energy consumption in higher education institutions as well as outline existing work on methods of energy estimation. The section will conclude with a summary of commonly cited variables used when estimating electricity consumption and a critical analysis of existing literature. The Methodology will summarize the steps taken to establish working models while the Model Development section will discuss the chosen variables for analysis at length including reasons for their selection, sources and quality of data, and preparation of data for regression analysis. In addition to revealing the final building sample and their collected variables, the tools used to carry out the analysis and create equations will also be outlined. The Results section will focus on the top candidate models and their performance as tools to estimate electricity consumption. The effects of variations in the methodology are also shown and the process of arriving at the final models for application is explained. Lastly, the estimated electricity consumption for Ryerson's cluster-metered buildings is presented with a comparison to their metered cluster values. The discussion will explore model performance in greater detail with a comparison to Rahman's [33] simulated results for Ryerson buildings. A closer examination of the model variables as well as buildings excluded from the study will also be included. The Discussion will end with a focus on challenges, both ones experienced throughout the study and those that will face future researchers adopting similar methods. The thesis will be concluded by revisiting the goals and research questions of this project, and the performance of the detailed method. The direction of future work will also be touched upon in the closing sections.

Before proceeding any further, it is important to note that the terms "energy" and "electricity" are used interchangeably throughout this thesis. As this work is dealing exclusively with electricity, it should be assumed that references made to energy are synonymous to electricity – unless otherwise stated. This treatment of terms also extends to ratios where energy and electricity use intensities are used to refer to the same metric – the amount of normalized electricity consumed.

II. LITERATURE REVIEW

Research involving academic buildings, particularly post-secondary institutions, and their utility consumption patterns is limited throughout existing literature because they comprise a smaller proportion (compared to commercial and residential) of newly constructed and existing buildings. In addition, their complexity in space usage, occupant density, and plug loads from one building to another, make generalizing consumption into standard energy use intensity units difficult. Caeiro et al. [34] outlines various other factors that are unique to academic buildings which make estimation difficult. There are limited resources available specifically for university and college buildings; knowledge is often extracted from works on other building types (e.g. analyzing multi-use residential buildings to make deductions about dormitories). Potentially due to a lack of experience and knowledge in sustainability on campuses, many academic institutions have come together through a number of regional and national initiatives targeting issues of sustainability, including energy use and efficiency. Examples of such initiatives are the Higher Education Environmental Performance Improvement (HEEPI) and EcoCampus in the United Kingdom, the Canadian Alliance of College and University Sustainability Professionals (CUSP) in Canada, and the Association for the Advancement of Sustainability in Higher Education (AASHE) in North America. Innovative projects and data collection from university and college buildings has risen as a result of groups such as these, however peer-reviewed research articles on the subject is still limited.

Existing literature on energy consumption in buildings (including academic) can be grouped into one of three levels of research: assessment, benchmarking, and change. Ideally, each level should be built on top of one another, meaning for instance, that benchmarking should not occur for a building before assessment takes place (Figure 8). Researchers often either focus on one level of research (e.g. compiling a comprehensive national benchmarking system for grocery stores [35]), or they tackle more than one level in a project (e.g. measuring and benchmarking energy consumption in a grocery store and installing a more efficient HVAC system to lower the consumption below the national average [36]). Literature falling within the first two levels will be presented here because it relates most closely to the scope of this thesis.

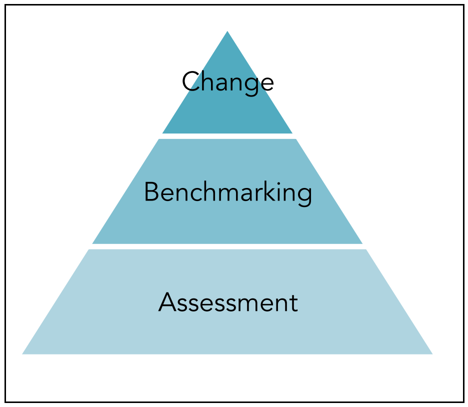


Figure 8 Suggested hierarchal nature of research in building energy consumption.

A. Electricity Consumption within Academic Buildings

This section will discuss electricity consumption in buildings that are found on university and college campuses. Much of the literature available and presented in this section focuses specifically on electricity, however, because energy consumption in all its forms is often studied cooperatively, research on other sources of fuel is also included. In addition, the works highlighted here focus mainly on studies completed on campus buildings – as opposed to inferring results from more commonly studied building typologies such as offices and multi-unit residential buildings. This will show the actual level of knowledge (or lack thereof) on energy consumption for higher education buildings without being distorted by works on space types that are similar to those found in universities and colleges. This distinction is important because existing literature has yet to prove whether this type of inference is accurate.

1) Normalization (Electricity Use Intensities): When measuring and benchmarking electricity consumption between buildings, the units that are most insightful are those that are attached to a key determinant of energy use. When overall consumption in kilowatt-hours is compared between two buildings, very little information is gained about the performance of the building – even when comparisons are made between similarly classified buildings (e.g. single detached houses, libraries, etc.). Instead, by linking consumption to a variable such as conditioned floor area, consumption can be normalized and compared on equal footing [29]. The most common

EUI metric that is used for buildings in general is floor area due to its simplicity and effectiveness in allowing comparisons to be made [37]. This is especially true when considering other forms of fuel, such as natural gas, due to its positive correlation with conditioned area; some studies have taken this further and have incorporated interior volume as the reporting metric [38]. Aside from variations on floor area, other metrics have been used consistently in existing literature; a common theme is focusing on the particular services offered by the building or space. For instance, occupants, such as students or staff, have been related to energy consumption in academic buildings, and the number of dishes prepared, used for kitchen spaces [30]. Ward et al. [39] studied 103 universities and 91 colleges to determine correlations with total energy consumption and certain indices (e.g. number of full time students, net internal area, age of buildings, etc.). It was found that the factors with the strongest correlations with energy consumption were gross interior floor area (r^2=0.86), net interior floor area (r^2=0.83), and number of full time research students (r^2=0.83). Climate is also an important metric used to normalize energy consumption as it defines the exterior conditions that a building's active systems must respond to. Normalizing to climate is especially important when comparisons are made between buildings in different climate zones. Because the models pursued in this thesis are customized for the Toronto-market, this type of normalization is insignificant (ie. the models will be created and tested on buildings within the city of Toronto). Instead, the quasi-standard electricity consumption unit, kwh/m^2/annum, will be used throughout this section so that comparisons between published results can be made.

Table II represents a summary of literature on EUIs for education-related spaces. Each entry specifies when the data was collected (or when the study was published), the location of the sampled buildings, the space classification that is linked to the measurement, and the size of the sample. Studies that measured total energy consumption and/or source/primary energy were omitted from Table II in order to promote uniformity and ease of assessment. Breaking down the table by geographic scope, buildings in North America use more electricity per unit area than those in Europe. A quick non-weighted average of the EUIs for academic-related spaces is approximately 210 kWh/m^2/annum for North America and 120 kWh/m^2/annum for

Europe. Adding office buildings (due to a lack of data for academic buildings) located in Asia, Figure 9 shows the ratios typical for each continent. The stark differences can be heavily attributed to the prevalence of air conditioning in buildings; in Europe, natural ventilation and high ceilings are common in older campus buildings while in Asia, hot and humid conditions require active cooling for occupant comfort. Baker and Steemers [40] points out that air conditioning can potentially account for 44% of energy demand in office buildings in the UK. The EUI reported for Asian countries may be inflated due to the sample comprising of multi-tenant office buildings in Lam et al.'s [41] study which are normally conditioned within a narrower margin than academic buildings. Geographic scope aside, Table II shows the variability that exists in the measurements for electricity use in academic buildings. Studies on buildings within the same continent vary considerably as well as within similar space classifications. Consequently, transferability of small sampled results is difficult between institutions.

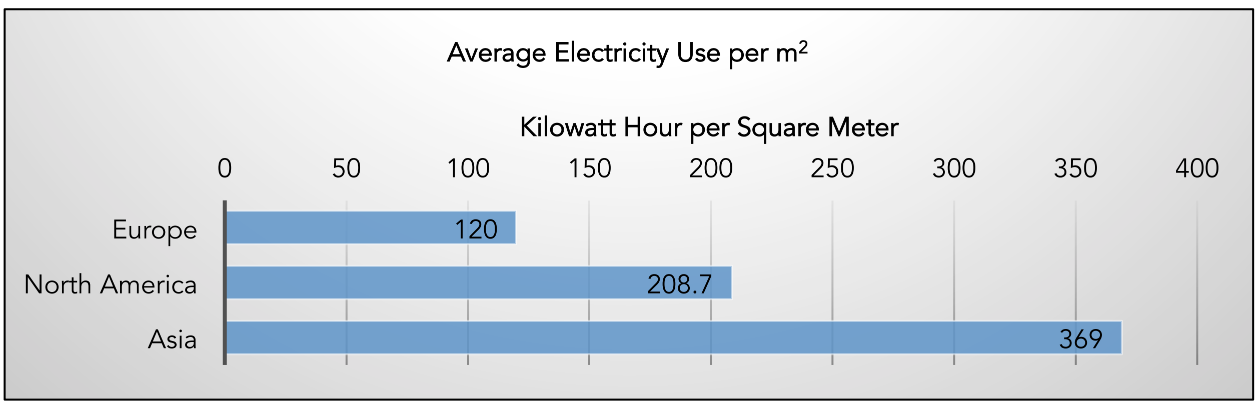


Figure 9 Average energy use intensity in higher education institutions in Europe and North America compared to the intensities of office spaces in Asia.

Table II List of studies and their reported EUIs for various university and college spaces around the world.

Year of Study	Location	Type of Space	EUI (kWh/m^2/ annum)	Sample Size			Reference
				Campuses	**Buildings**	**Area (m2)**	
2003	USA	Classrooms	204		88		[29]
		Dormitories	95		37		
		Libraries	213		36		
		Recreation	106		92		
		Office	186		976		
1999	Bordeaux	Administration	33			39,000	[42]
		Research and Development Labs	118			80,000	
		Lecture Halls	37			230,000	
		Restaurants	88			17,000	
		Dormitories	39			96,000	
		Sports Facilities	36			20,000	
		Libraries	25			18,000	
2001	Cergy	Student Residence and Eatery	48		n/a		[43]
1999	Krakow	Student Residence and Eatery	115		n/a		[43]
2000	Bucharest	Student Residence and Eatery	65		n/a		[44]
2000	Northern Ireland	University (Non Residential	39		39		[45]
		University (Residential)	60		19		
		Libraries	45		40		
		Cafeterias	42		44		
		Offices (Naturally Ventilated)	82		70		
2000	Canada	Education	111	40 (Colleges)	156		[46]
	British Columbia		119		n/a		
	Prairies		108				
	Ontario		106				

	Quebec		122				
	Atlantic		117				
2001	Anchorage	Elementary School	150			3,472	[47]
			111			5,723	
			89			5,723	
			107			5,723	
			121			6,166	
			111			5,767	
			102			5,631	
			102			3,017	
		Secondary School	131			4,265	
			146			2,534	
		High School	134			31,661	
2003	New Jersey	n/a	182		16 (estimated)		[48]
2000	Marseilles	n/a	67			120,000	[43]
2000	Rouen	n/a	30-750		32		[43]
1995	Denmark	n/a	79		n/a		[44]
2000	Cergy	n/a	32		4	30,000	[43]
1997	Finland	n/a	111		500		[49]
1998	Poland	n/a	204		3	12,549	[49]
2011	United Kingdom	n/a	290			26,700,000	[50]
2003	Ontario	University	167	37			[51]
	Atlantic		131	23			
	Quebec		200	22			
	Prairies		203	30			
	British Columbia/ Territories		158	11			
	Canada		176	123			
	Ontario	College	163	43			
	Atlantic		106	38			
	Quebec		152	78			
	Prairies		154	33			
	British Columbia/ Territories		151	36			
	Canada		152	228			

2011	Northern Ireland	Higher Education Institutions	95		328	621,858	[52][a]
	Scotland		127		1775	3,206,141	
	Wales		102		1108	1,494,439	
	England		123		> 12577	20,919,394	
2003	United Kingdom	Admin/ Support	90		22		[53]
		Sports Center	199		8		
		Libraries	186		3		
		Residences	57		37		
		Teaching	118		36		
		Medical Lab	325		15		
		Engineering Lab	130		24		
		Chemistry Lab	264		7		
		Computing	106		11		
2012	Massachusetts	Dining Hall	827		100 in total	833,919 in total	[32]
		Residential	249				
		Library	218				
		Recreation Centers	233				
		Academic (General)	249				
		Laboratories	615				
		Administrative (General)	357				
		Administrative (Health)	262				
2002-2012	USA	Chemistry Laboratory	415		28		[54]
2001-2012		Physical Laboratory	460		24		
2002-2012		Biology Laboratory	477		27		
2003-2012		Other Laboratory	424		64		
1993-2012		Chemical and Biological Laboratory	478		48		
1996		Fast Food	208		n/a		

	United Kingdom	Restaurant	144				Value for money initiative, no direct source
		Science Labs	165				
		Science Other	121				
		Arts	71				
		Residence Halls	93				
		Flats	49				
		Library (AC)	347				
		Library (Natural Vent.)	56				
		Students Union	165				
		Admin (AC)	144				
		Admin (Natural Vent)	47				
		Sports (Wet)	208				
		Sports (Dry)	83				
2006-2008	Colorado	High School	79	2	52210		[55]
Prior 1997	United Kingdom	Office	36		n/a		[56]
		Library	50				
		Catering	650				
		Sports	150				
		Lecture Halls	108				
		Laboratory	105				
		Teaching	22				

a: full details on each institution shown in Appendix B1

Benchmarking energy consumption for buildings is a powerful tool that not only gives an indication of current performance, but also informs researchers of historic and future trends in the sector. Databases allow for the assimilation and organization of potentially very large data sets, which is necessary when creating a reliable benchmark. Due to the resources required to track the number of buildings that exist, many of these databases are created and maintained by the government. These databases were first created for high impact building types such as industrial or residential buildings; these type of buildings have a proportionately large energy footprint in the building sector either due to their sheer numbers, in the case of residential, or their energy requirements for operation, in the case of industrial [57, 58]. Currently, the Energy

Use Data Handbook [1] is updated annually by the Government of Canada and provides consumption information on major energy consuming sectors in Canada. While these types of reports give insight on overall trends in terms of supply and demand of energy across the country, they lack the detail provided by [5, 29], documenting building characteristics and usage patterns alongside their energy demands. These programs are active but are less frequently updated and survey a sample of the entire building stock of that country. An example of such a program is the Commercial Buildings Energy Consumption Survey (CBECS), which covers buildings in the United States. CBECS was first conducted in 1979 and has historically been updated every three years; the most recent survey published was from data collected in 2003 and work is underway for its ninth iteration using 2013 survey data. To summarize, the Survey of Commercial and Institutional Energy Use [5] and CBECS [29] yields more insightful data on buildings and energy consumption but lacks the geographic breadth and inclusiveness of [1]. Both types of report provide valuable information on energy consumption within buildings, and their level of impact on a provincial and/or national scale. Certain pilot projects offer the benefits of both types of studies such as the Benchmarking Guide for School Facility Managers [59] however they are rare and are often not updated after the original publication. The benchmarking of schools in [59] is insightful and encompasses buildings across Canada however the focus of the study is on primary and secondary schools. Nevertheless, federal reports such as these are instrumental in conducting research in this field by supplying quality datasets to use and/or to improve upon.

There are a few papers that are especially relevant to understanding and estimating electricity consumption in academic spaces. While some of their findings are summarized in Table II, the remainder of this section will be dedicated to detailing the context, methodology and findings of such selected works. The objectives of Bonnet et al. [30] are very similar to those of this thesis in which they: (1) attempted to establish a methodology for auditing energy consumption in university buildings, (2) tested the methodology on real buildings at the University of Bordeaux, (3) gathered data about energy consumption to compile a database specifically tuned to university buildings, and (4) increased their understanding of the patterns and fluctuations in university buildings. Electricity use intensities for area and occupancy were calculated for campus buildings and were categorized by primary activity such as research, catering, and lecturing spaces. They found that libraries showed the smallest EUI of all spaces (25 kWh/m^2/annum), while (high tech) laboratories showed the greatest (123 kWh/m^2/annum). Subcategories were created for catering and research areas due to large variations in consumption making generalization difficult. Within catering, spaces were differentiated by the proportion of cooking space required to serve the area. Restaurants, for instance, required a larger kitchen area than cafeterias and as such, use more electricity overall per unit area. Furthermore, electricity was better correlated with the number of meals served in the space, however, the authors acknowledged the difficulties in accounting for this unit. Laboratory spaces were divided between high and low technology spaces based on their reliance on powered equipment. Unfortunately, even with the subgroups created, large variations persisted in labs and correlation with area was weak; the reported mean EUI of 117 kWh/m^2/annum for research areas was not statistically significant. The EUIs for all activity types were aggregated for all buildings and applied to the space usage breakdown for the University to determine what types of areas consumed the most and least electricity. The breakdown by area and total energy consumption of university spaces studied in their paper can be seen in Figure 10. The change in proportion between area and end-use is significant for catering, lecturing, and research and development spaces. Particular attention should be put on these spaces as they may be a key determinant for estimating electricity consumption for Ryerson buildings.

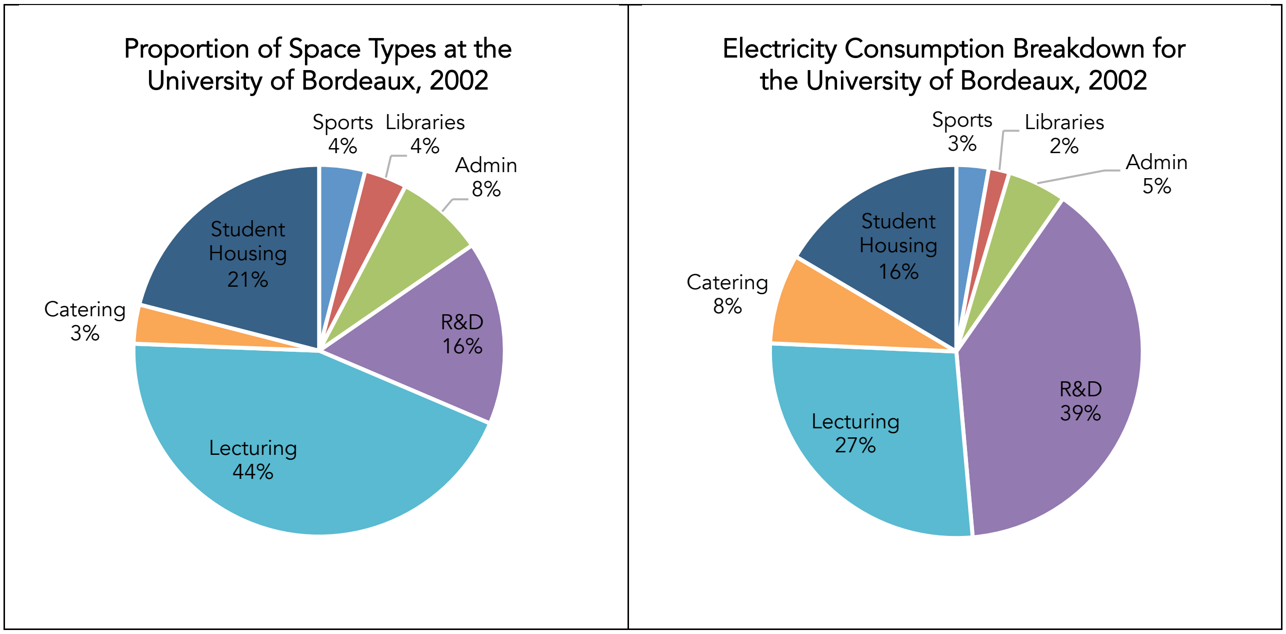


Figure 10 Space usage at the University of Bordeaux (right) and the estimated electricity end-use (left) calculated with aggregate EUIs. [30]

Another example of linking electricity consumption with a particular space type is outlined in [56] where estimation is based on average electricity benchmarks from United Kingdom higher education institutions. The breakdown of space types for the higher education sector can be seen in Figure 11 and is comparable to those found at the University of Bordeaux. The guide covers the assessment of electricity consumption as the first of a three-step process to benchmark and meet specific energy targets. Therefore, there are no published results within the guide nor are there any performance metrics for the assessment. Nevertheless, the concepts behind this method of relating consumption with specific activity spaces are very similar to the ones employed in this thesis and other published work.

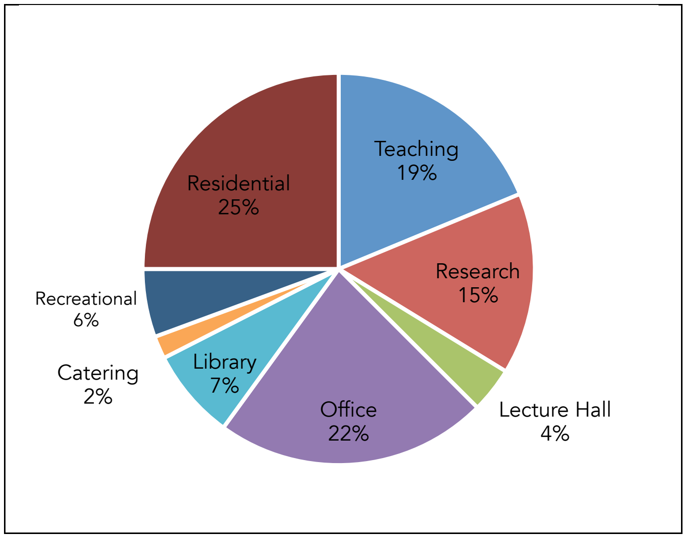


Figure 11 Space usage for the higher education sector in the United Kingdom. [56]

The challenges of generalizing electricity consumption for academic buildings is made evident with Elliott and Guggemos' [55] work on auditing energy use in high school buildings. The study attempted to identify the cause of variation in consumption between two seemingly similar high school buildings. One building held a LEED-Silver/Energy Star designation while the other did not – both had similar occupant sizes and overall construction. Through their audit and analysis, it was revealed that a combination of different plug loads and space usages was the root cause of the consumption difference. Specifically, the LEED-rated school had a higher density of computers in their comparatively smaller computer lab which offset the energy savings from using low wattage lighting throughout their building. Significant differences in EUI that were found between spaces in both schools included: computer labs (2.1x), trades classrooms (2.1x) gymnasiums (1.7x), kitchens (1.6x), and common areas (1.5x). While this variability is not unexpected with these space types, they remain a significant usage type in academic buildings, particularly in colleges and universities. Figure 12 shows the proportion of space usage and electricity consumption for the combined high schools. Similar to the University of Bordeaux, spaces which rely heavily on electrical equipment (e.g. kitchens, administrative, etc.) see an increase in share from area to electricity use. Interestingly however, computer labs in the high schools only see a modest increase in footprint from area to overall energy consumption. Elliott and Guggemos conclude by stating the weaknesses of whole building EUIs for multiuse buildings and recommends instead workspace and component EUIs for benchmarking and highlighting inefficiencies. While relating space usage types with certain energy profiles is one of the simplest approaches to generalizing electricity consumption, this study shows the effects of ignoring plug loads. However, [55] only compares and contrasts the energy profile between two buildings – it is expected that with a greater sample size, the differences will be diminished.

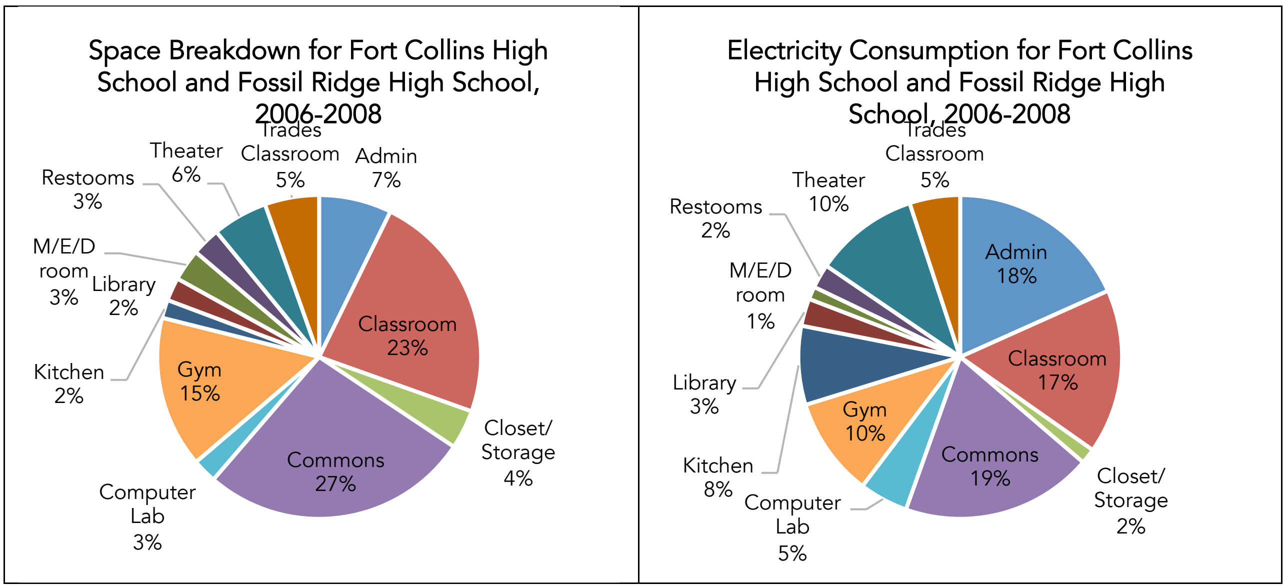


Figure 12 Space usage and electricity consumption for two high schools. [55]

B. Methods of Estimating Electricity Consumption in Buildings

As mentioned in the Problem Statement of the Introduction, the only solution available for quantifying the electricity consumption at Ryerson University, and potentially other institutions, is some form of model-based estimation. Aside from creating energy models for all cluster-metered buildings, there are many tools that can assist in forming a reliable estimate. Chung [37] details and compares several different mathematical models used to quantify building energy use. Zhao and Magoulès [60] covers a broader scope and includes several frontier prediction methods such as artificial neural networks and support vector machines. This Section will outline common methodologies used to estimate electricity or total energy consumption for all types of buildings.

Common to all methods, estimating energy consumption for a building relies on interpreting a series of variables. The complexity of the model is related to how many of these variables are considered, whether the correlation between energy consumption and the variables are linear or non-linear, and the inclusion or exclusion of interaction effects between variables. On one end of the spectrum, energy models are all-inclusive, allowing an entry to be made for all building variables that affect energy consumption; on the other end, simple normalization only accounts for one key determinant, such as area, to predict energy use. In-between the two, there are a host of alternatives that vary in the amount of variables considered, when forming an estimate. Successes have been found across the board making performance-based rankings difficult, but their approach and ease of use can be clearly defined [61].

Simulating energy use through computer models has been used extensively in the design/planning, maintenance/operation, and optimization/retrofit stages of the building lifecycle. This is due to its ability to predict detailed energy use from a flexible number of inputs for a particular building – a simple task for trained individuals. However, the main drawbacks to building and using energy models for prediction purposes is that the quality of inputs into the model are directly related to the outputs. If accurate physical measurements are not entered for the building, detailed characteristics are not attributed to building spaces, and materials and envelope assemblies are left unspecified, there is a high probability that the model will not be representative of the building. This is particularly troublesome for beginners as the use of default values is often detrimental to the model's accuracy. There are nearly 150 programs specific to energy modeling that are listed in the Building Energy Software Tools Directory compiled by the U.S. Department of Energy [62]. This level of diversity increases the difficulty for untrained professionals to seek and learn the proper tools they require; this partly explains why sourcing this work is often expensive and time consuming, especially for multiuse buildings [63]. While detailed models have a dedicated market, work on simplified models requiring much fewer user inputs has been published. These models, which focus on energy use stemming from the HVAC system rely on fewer and less complex inputs (i.e. weather and climate-related variables), and is detailed in Al-Homoud's [64] paper. Energy models have also been used in conjunction with statistical methods to narrow down variables of interest. Modeling programs provide a simple way of conducting sensitivity analysis on building variables by changing values between defined extremes to gauge the effects on energy consumption, which was employed in [65-68]. Depending on the perceived or known relationship between the independent variable and energy consumption, two values are simulated for a linear relationship and three or more values are simulated for nonlinear relationships. Once significant variables have been identified for a particular building or typology, they are analyzed and used in statistical models. This feature, however, is not exclusive to simulation programs as demonstrated by Hong et al. [38] where a sensitivity analysis for variables affecting electricity consumption and heat loss took place using artificial neural networks – the results of which can be seen in Figure 13. Energy modeling can also create large virtual sample sizes for training and/or testing other estimation methods that would not be possible with the current level of building surveys available. While there are certain inherent risks with using a virtual dataset to train/test a model, this option is invaluable for establishing early performance benchmarks.

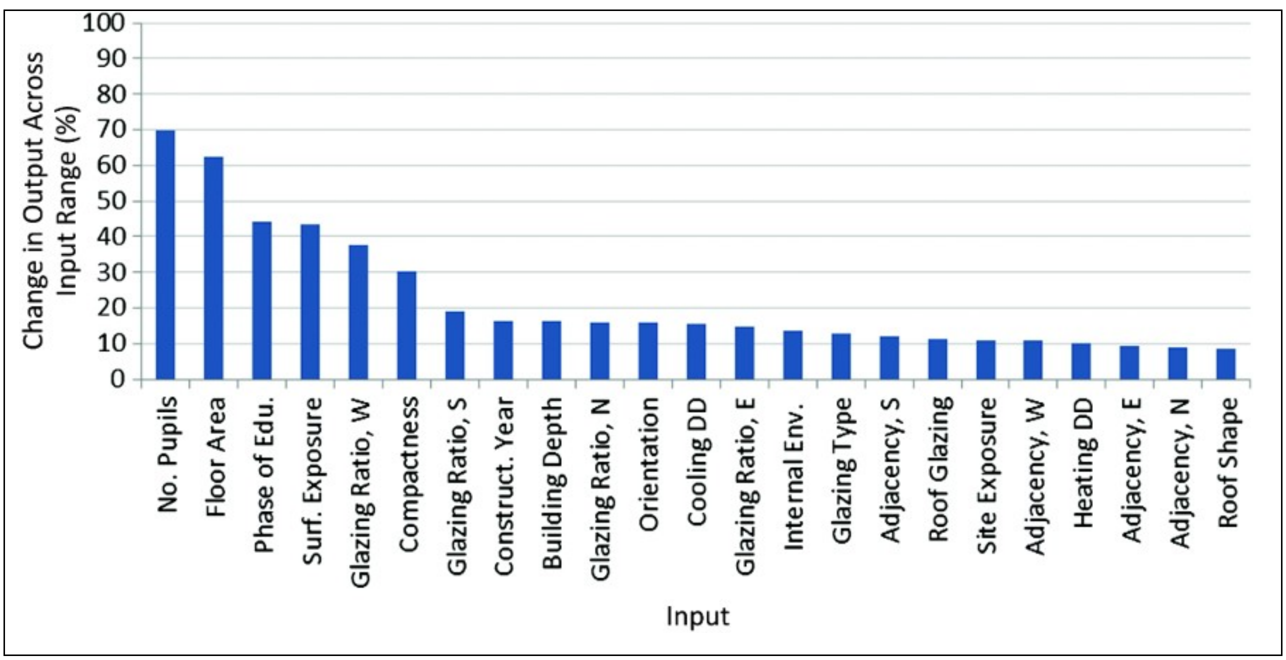


Figure 13 Sensitivity analysis (via artificial neural networks) carried out by [69] on variables affecting electricity consumption in primary and secondary schools in the United Kingdom.

Methodologies for estimating energy consumption have more recently stemmed from disciplines outside of statistics including pattern recognition (decision tree), and machine learning (artificial neural networks). In recent years, these methods have been tasked with estimating energy consumption for a variety of buildings. Decision trees are a way of categorizing data into groups in a visual manner. The method behind building or growing a decision tree is to find features or variables that can group individuals together to form subsets of the original sample. The node at the top of the tree, the root, represents the entire sample size; in a series of splitting actions (working down the tree), individuals are separated and categorized based on defining characteristics until all observations contained within a node are within an acceptable range of values, thus becoming a leaf of the tree [70]. Figure 14 is a general schematic of how decision trees are visually represented. Red lines represent buildings that don't meet the criteria and green lines represent the subset that do. Numbers contained within the brackets indicate buildings classified in each node and are used to calculate the average EUI for that leaf node. Depending on the application, decision trees can predict discrete (classification tree) or continuous (regression tree) values [71]. Within the context of buildings and energy consumption, EUIs would be classified based on significant variables such as glazing ratio and number of occupants. Important to note is that variables selected as significant are those that differentiate observations from one another and not on their correlation with energy consumption. For instance, if the entire building sample has a glazing ratio of 50%, it would not be a node on the tree because of its inability to split the sample despite its proven relationship with energy use. The greatest strengths with using decision trees are their ability to accurately categorize data without understanding the underlying relationships between dependent and independent variables. Also, their ability to provide a visual aid and/or logic statements that govern a particular sample increase their capabilities in communicating potentially complex relationships to the general public [72]. The effectiveness of using decision trees to estimate energy consumption is demonstrated in Yu et al. [72] who looked at predicting the EUI of residential buildings using ten inputs. The model was developed using a sample of 55 buildings and tested on 12 randomly selected buildings. The decision tree was able to correctly classify all but one of the buildings into their appropriate categories. However, the error rate between the reference and actual EUI for the buildings was apparent (average error = 25.6%). This means that despite the 91.7% accuracy of the decision tree in categorizing buildings, the variations in electricity consumption within each leaf node was considerable, taking away from the overall reliability of the method. A larger sample size would potentially address this shortcoming but further research is needed to consider this observation a major drawback. Decision trees also suffer when non-linear relationships exist between variables and when the tree is over fitted to the data it was built upon – a common issue with this method [73]. Despite these drawbacks, decision trees offer a realistic solution for estimating energy consumption for people with a lack of building knowledge [61].

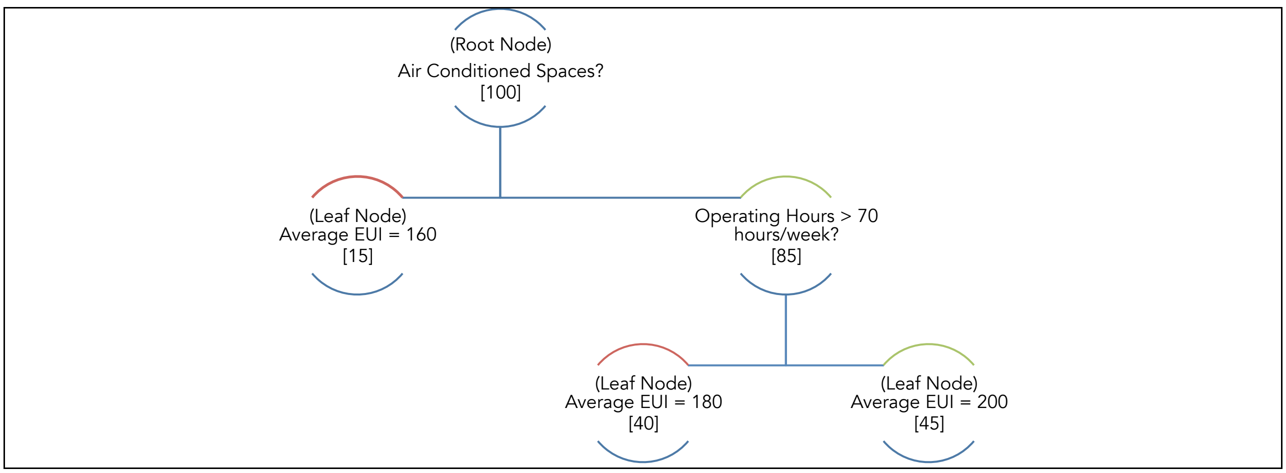


Figure 14 A hypothetical decision tree schematic using two predictor variables.

From a transparency standpoint, artificial neural networks (ANNs) operate in complete contrast to decision trees. They were originally modeled after the biological structure and properties of the central nervous system (i.e. brain) found in animals – hence being classified under machine learning. Today, they can be found in the backend processes powering speech and facial recognition in modern day electronics, among other things [61, 74]. ANNs are composed of a series of interconnected nodes (neurons) that are arranged in a series of layers. In its most basic form, ANNs have an input, hidden, and output layer (Figure 15). The input layer represents the body's five senses and is responsible for feeding all available data (simultaneously) to the brain, the hidden layer. The hidden layer contains a series of nodes with adaptive weights. During the learning phase, these connections and weights are adjusted to provide some defined optimal solution. In addition, the number of hidden layers can be increased to account for greater complexity among variables. The output of the ANN, an action or thought in animals, would be some measurement of energy consumption for research in this field. ANNs' strengths come from their ability to model complex, non-linear relationships between variables – an issue facing competing methods [61]. However, this strength comes at a costly price. Within a controlled environment, it is much easier to predict how a person will react to stimuli than it is to understand the motives behind their actions. They themselves may not be acutely aware of why they are doing something (e.g. natural instinct). This analogy explains why ANNs (along with support vector machines) belong to a group aptly named by the engineering community as "black box methods", for their inherent obscurity in variable weights [61, 75]. While this makes model building difficult to comprehend for beginners, users enjoy the benefits of being able to model complex and nonlinear relationships with simple inputs and outputs. The most commonly used model is a multilayer perception, a type of supervised network. This means that the model learns only when data on the expected outcome is available. Training occurs through a technique known as backpropogation in which the inputs are repeatedly entered into the model and the error (the difference between the output computed and the expected output) is then used to adjust the weights within the hidden layer to improve the accuracy. This is repeated until the model or network achieves an acceptable level of accuracy [74]. Performance-wise, ANNs are highly competitive when applied to buildings and energy consumption. Aside from comparative studies by Tso and Yau [61] and Hawkins et al. [76], the effectiveness of ANNs can be seen in Yalcintas and Ozturk's [77] work where a network was developed to model buildings using the 1999 CBECS database. Using eight input variables, the model outperformed multiple linear regressions created and tested using the same sample. In addition, the authors commented on the flexibility of the network, which could be applied to different climate zones after a period of training with the regional data – a feature not shared by the more structured multiple linear regression approach. Among primary and secondary schools Hawkins et al. [76] and Hong et al. [78] created neural networks that could predict electricity consumption within a 34% and 20.6% margin of error, respectively. ANNs have and will continue to play a role in quantifying energy consumption in a wide variety of buildings. Their ability to model complex relationships between variables continues to attract researchers in this field.

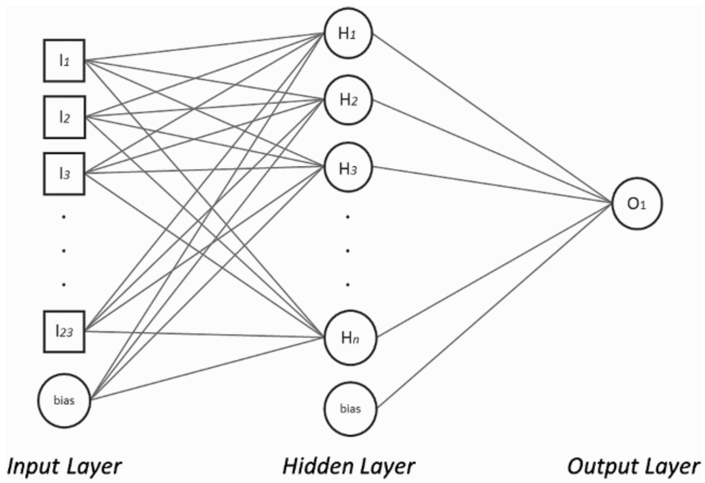


Figure 15 Visual representation of a general artificial neural network. [38]

1) *Multiple Regression Analysis:* Statistical regression models rely on identifying one or more variables that correlate well with energy consumption and are responsible for a significant portion of the overall building consumption. These variables are best identified by using historic data from buildings; alternatively, they can be identified by running sensitivity analysis via energy models (computer-based or other), as previously discussed. Regression models are similar to computer-based energy models in terms of how they process inputs and calculate energy consumption however there are no interactions between multiple equations as may be the case when running simulations. Statistical models exist to explain variations of a dependent variable with as few explanatory factors as possible. Regression models are capable of representing linear or non-linear relationships between variables, however, with building energy consumption, existing literature has heavily favored linear regression models for their simplicity and acceptable performance levels. Lam et al. [67] looked at correlating 28 parameters with energy consumption from 387 simulated buildings and found that eight of them fit better with a quadratic regression than a linear one; six of those parameters were related to the HVAC system of the building.

One of the earliest attempts at using multiple regression, a model comprising of more than one independent variable, to estimate energy consumption in buildings was Boonyatikarn [79] in 1982. His doctorate thesis explored the prediction of energy consumption in 50 institutional buildings in Michigan using building-related variables. His final model, which was successful in accounting for 93% of the variation observed in energy consumption, included the following ten variables: exhaust air rate, hours of HVAC operation, opaque wall area, type of refrigeration, percent of floor area air conditioned and cooling degree day, type of air handling unit, type of fuel used, volumetric flow rate divided by the power of the supply fan, shading in winter, and shading in summer. Over 30 years later, using statistical models has become one of the most commonly used methods of estimating energy consumption. Of the 23 papers surveyed by Chung [37], 12 papers used the ordinal least square (i.e. linear regression) method to estimate consumption, and four papers used other statistics-based methods; the remaining seven papers used various computer simulated models. Other early works employing multiple regression analysis include models on energy use on a military base [80], restaurant [81], recreation center [82], and residential buildings. Works by these authors are summarized in [63] which captures the level of knowledge and available tools for researchers tackling similar research problems as this thesis, more than two decades ago.

Developing a regression model is an iterative process that requires a thorough understanding of statistics. Fortunately, there are tools available to simplify the process by creating preliminary models that can be further developed. The most popular statistics software (SPSS, MATLAB, SAS, R, etc.) are all capable of running multiple regression analysis which is available through an add-on, if not offered as a core feature. These programs have the ability to automate the model and feature selection based on a variety of user-defined options. Relying heavily on an automated process to develop working models is often criticized as a thoughtless approach [83]. Thus it is important to use a priori knowledge of the relationships between dependent and independent variables to inform decisions on which correlations are meaningful [84]. Selected parameters can potentially require very specific and technical building information such as thermal performance of assemblies and shading coefficients, however, it is not a prerequisite for a well performing model, as is further discussed below. Examples of more recent studies using regression models to estimate energy consumption can be referenced in Table III (with further details shown in Appendix B2) as the majority of sourced studies are based on regression analysis. Further detail on this technique is discussed in the Model Development section including how collinearity among predictor variables, one of the main weaknesses of regressions analysis, is addressed.

C. Predictor Variables for Energy Consumption

Variables used to estimate electricity consumption in buildings are highly diverse, even among similar methods. Table III shows past literature on modeling electricity consumption in various types of buildings using one of the methods previously detailed (i.e. ANNs, decision tree, multiple linear regression). The Table includes the total number of variables, which averages seven for the included studies, and the number of variables in their models falling under the defined categories. Much more detail about the studies and their methods including their error rate, sample size, and target buildings can be seen in Appendix B2. Also referenced in the Appendix are the specific variables that defined each category such as outdoor air temperature and hours of rain for the site-specific/location category. There is a wide spread of variables selected for modeling especially when there are no preceding studies to build upon. Sensitivity analysis on model variables such as those in Figure 13 provides useful insight into their effectiveness as predictors but few studies take this extra step. Table III should be used to demonstrate this diversity with less emphasis placed on the frequency of certain categories being chosen. This is because Table III represents a small snapshot of reality and includes certain biases such as including multiple works from the same author(s) – leading to some variables being overrepresented. Given that there are no obvious trends or guidelines evident through the review of existing literature, variable selection will rest solely on the original guidelines previously established: that variables to be considered should have a proven impact on electricity consumption (i.e. high sensitivity) and that they require minimal effort to gather the necessary data for model creation and application.

Table III Summary of significant variables used to model electricity consumption in buildings with decision trees, linear regressions, and artificial neural networks.

Source	*Total Number Of Variables*	*Site-Specific/ Location*	*Lighting*	*Space Usage*	*Building & Equipment Ownership and Occupants*	*Physical Building Traits*	*HVAC Equipment*
[85]	2 or 3				3	3	1
[86]	2 or 3				4	1	1
[67]	12		1		2	2	6
[66]	17	1	1		3	6	1
[72]	10	1			2	5	
[61]	6				2		
	3				3		
	4				2		
	6				3		
	5				2		
	6				3		
[35]	2	1	1				
[87]	9		1		2	1	2
[79]	10	1			2	2	4
[88]	4	2			1		
[89]	8		1	1	2	3	
[90]	5			2		1	2
[68]	11		1		1	5	2
[91]	9		1		3		2
[65]	9		1		1	5	
Total		***6***	***8***	***3***	***41***	***34***	***19***

D. Critical Assessment of Current State of Research

Research papers presented and summarized in this section have shown varying levels of successes at quantifying and benchmarking energy consumption in buildings. However, these studies are often isolated from one another and lack a collaborative approach at tackling the main research questions. In addition, questionable practices, such as those listed below, should be corrected or addressed with urgency to promote cohesion and confidence among researchers and interested parties. This section will conclude with an assessment of weaknesses found in existing work and the foreseeable challenges facing future progress.

1) Weaknesses of Existing Literature: Sensitivity Analysis Based on Uncalibrated Building Models: A large proportion of work in understanding the sensitivity of certain variables on electricity consumption depends on constructing and executing energy simulations. The issue with relying on simulations is that their outputs may not be validated. Unlike cases where energy models are built using real buildings, calibrated using their utility bills, and applied to predict future consumption, the behavior of a base case model, used to test different variables, remains unchanged from one model to another. This would be akin to relying on outputs from an uncalibrated energy model. Similar to the process of calibrating an energy model, a base case model used for sensitivity analysis needs to be calibrated using actual building data to justify the outputs from using this method. Until such practices have taken place, results from exploring the effects of building variables using energy models cannot be entirely relied upon – their direction of effect (positive or negative) and their relative proportions to other variables may hold true, but their absolute effect on electricity use intensity may be inaccurate.

Shortage of Quantitative Studies on Energy Consumption and Building Variables: A focus on qualitative effects of building variables on energy use is abundant in literature but ones quantifying them are limited. This may be due to several reasons, all of which relate to the complexity and technicality of quantifying the relationship between specific building variables and electricity use. Without a discrete relationship, it is difficult to justify the need to prioritize certain variables over others.

Lack of Standardization: Energy use intensity, or more specifically "energy", can represent multiple measures and has yet to be standardized in current literature. EUI can refer to total building energy consumption, which includes all forms of energy supplied to the building (i.e. electricity, natural gas, and steam), or it can represent a singular form. Another dimension to consider is whether the consumption values represent site energy or primary energy – which takes into account energy lost during the generation and transmission of energy to the building. Electricity is particularly susceptible to transmission losses resulting in a larger difference between the measured site and primary energy value [92]. Most authors are aware that energy is a loose and generic term which needs defining however not all authors are clear in their attempts to distinguish the differences in the prefaces of their own work.

Compared to other indices (e.g. floor area, occupants, etc.), normalizing EUIs based on space types is very unstructured. While there seems to be a large variety of spaces that are linked with electricity consumption, a lack of effort in defining and differentiating these spaces is causing issues. Specifically, problems of knowledge transfer may mislead other researchers into benchmarking and comparing their data with incompatible studies. Olofsson et al. [93] mentioned inconsistencies in defining key terms in international research which led to complications when comparing works. Fortunately, disclaimers are more common in recent literature outlining the dangers of applying reported EUIs across the board.

Scoping Research Around the Limitations of Building Simulation Programs: Sensitivity analysis has traditionally involved using an energy model to simulate the effects of certain parameters on electricity consumption. There is a potential that following such a method limits the scope of analysis for researchers because their tested parameters are limited by the level of detail supported by the energy modeling program or model. In theory, the effects of using the existing method of sensitivity analysis are most likely negligible on the resultant model because the designers of such programs incorporate most, if not all, major variables that affect electricity consumption. However, it is unknown what effects using this methodology will have on how the problem is approached by researchers.

2) Challenges Facing Future Research: Complexity of the task: Variability in the form and use of buildings ultimately translates to a wide range of energy consumption values – even between similarly defined buildings (e.g. commercial, residential, etc.) As a result, it is difficult to isolate specific variables that are indicators of electricity consumption for one type of building. This in turn hinders further research into specific variables because published works do not collectively support the same variable for building types. This is a simplification of this issue as there are many other factors that are considered when selecting a model's predictor variables such as the availability of data and the scope of analysis.

Transferability: Attempts to create a large database of buildings where EUIs can be calculated with confidence requires a large amount of resources. Because of this, many researchers opt for smaller sample sizes to test relationships. This jeopardizes the ability for the EUIs calculated to be applied to a broader scope of buildings. As a result, the majority of current work that is available represents local snapshots of the consumption patterns for university buildings and not a robust universal benchmarking tool.

Motivation: A general lack of incentive to improve existing operations exists in the community. Quite simply, the cost of electricity in the United States has been close to an all time low in the past 50 years (Figures 16 & 17). Canada also benefits from some of the lowest costs of electricity due to their existing infrastructure in hydroelectric generation [94]. As a result, the motivations to reduce consumption are primarily environmentally based which are highly susceptible to changes in personal opinion and public opinion. This creates an inhospitable environment for researchers in the field due to fluctuations in support and resources provided by the government and other stakeholders [95].

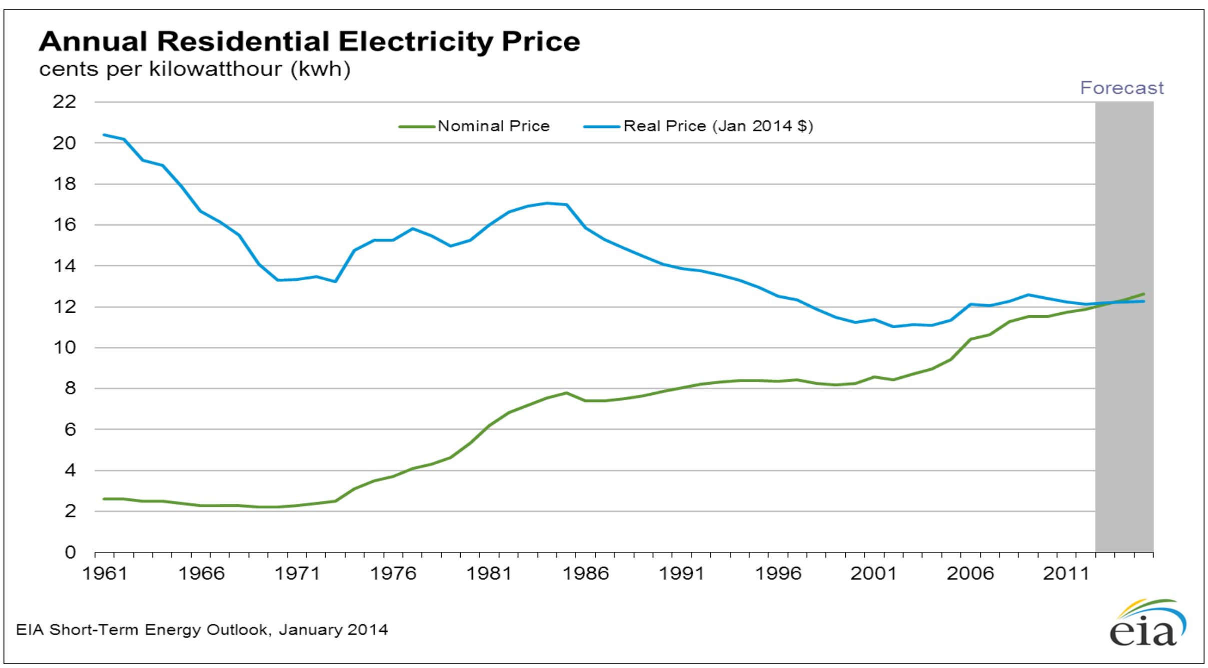


Figure 16 The historic and future electricity price for residential users in the United States, adjusted for inflation[96]

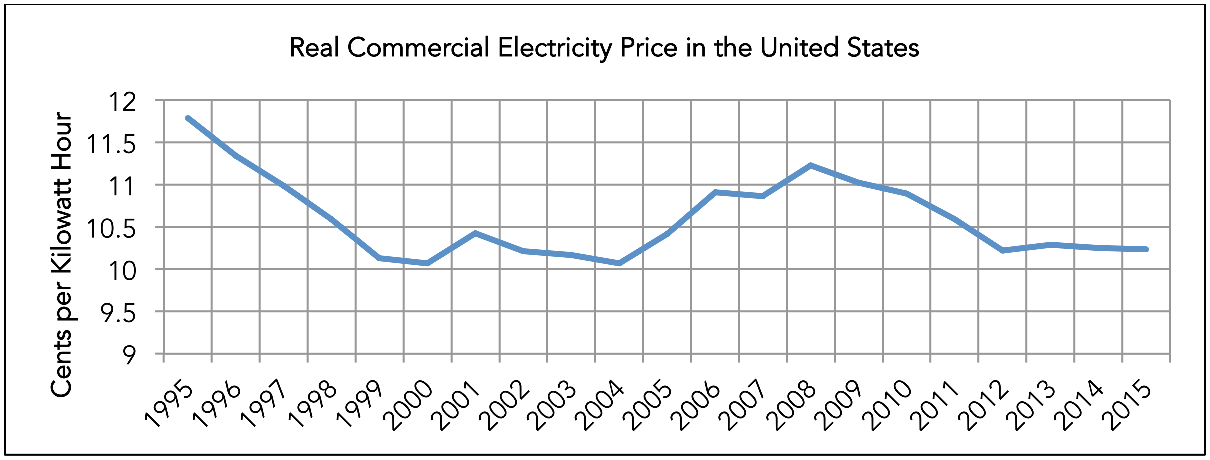


Figure 17 The historic and future electricity price for commercial users in the United States, adjusted for inflation [96]

Lack of Quality Data: There is a lack of data available to researchers, limiting their ability to explore potential relationships between building characteristics and electricity consumption. For instance, CBECS is used for the majority of papers originating from the United States and is referenced by international researchers as well. This is partly due to the comprehensiveness of the survey, but it is also an indication of the lack of sources for national data available to researchers. Most countries are not fortunate enough to have established such an extensive survey of existing buildings, Canada included [97]. In these instances, smaller samples must be manually surveyed by researchers which limits the applicability of the findings and ultimately the significance; what is true at a local scale may not me expressed at a national level and what variables may seem important in a few buildings may be much less so when sampling from a much larger population. The availability of such comprehensive surveys is necessary for the growth of research in this field. Without data, researchers must spend more time and resources to gather their own data, potentially deterring research altogether.

III. METHODOLOGY

Below is a basic summary of the approach used to disaggregate electricity consumption for Ryerson's cluster-metered buildings.

Gathering Building Sample	Ryerson University provided electricity consumption data for 23 utility meters on campus. 17 buildings were individually-metered.
	The University of Toronto supplied consumption data for 120 individually-metered buildings
Preparing the Building Sample	28 buildings were removed from the original sample from U of T and Ryerson on the basis of size, archetype, and electricity consumption.
	Subsets of the entire sample were created to reduce variability in consumption allowing for greater performance from models. Three, four, and five subsets were experimented with.
Collecting Data for Predictor Variables	Participating universities provided data on the number of above and below ground floors, space usage (defined by COU categories) and electricity consumption on a per building basis. The data for footprint shape, and shared external walls was gathered through a survey of satellite imagery and maps
	COU space categories were amalgamated into 13 variables. This re-categorization was based off of percieved energy use intensities for the original defined spaces.
Model Building	Multiple regressions were created for each subset. All possible combination of predictor variables were experimented with to determine the best models for predicting electricity consumption.
	The top five models from each subset with the lowest AICc score were averaged to form the final model coefficients. A weighted average was taken using the Akaike weights for each of the five candidate models.
Model Testing	Cross validation (leave-one-out) was used to test the candidate models before multimodel inference (averaging). The mean square error was calculated for each model and compared to the average electricity consumption to determine the predictive error or root mean square error.
Model Application	The final averaged models were applied to Ryerson's cluster-metered buildings. Depending on their size, one of four models was used to estimate electricity consumption.
	The raw estimates from the models were adjusted after a comparison with the metered usage. Buildings belonging to the same cluster were compared and adjusted proportionately, regardless of the model used for estimation.

IV. MODEL DEVELOPMENT

As shown in the literature review, there are many variables that have a proven relationship with electricity consumption in buildings. Therefore, estimating electricity consumption from variables can be flexible depending on a particular project's limitation. For Ryerson University, and potentially other academic institutions, a financial limitation exists which narrow down the number of variables that can be considered when trying to estimate electricity consumption. By favoring variables that are measured and collected on a regular basis – for purposes other than to estimate consumption – the time and costs associated with this method are minimized. The variables selected to relate with electricity consumption have been cited in exiting literature as being effective predictors but are simple to measure and report for untrained staff. This way of estimation is not to compete with existing methods, rather it is to compliment and provide a low-cost option for gauging consumption in a transparent way.

The variables studied in this thesis were narrowed down from an exhaustive list created during preliminary work. Table IV represents an early attempt to prioritize certain building characteristics according to their level of impact on electricity consumption. The Table was annotated with published data ([29, 65, 98]) to help make informed decisions on variable selection. Since this is an exhaustive list of variables, all components were considered regardless of the availability of data. Beside each indicator is a check signifying whether it affects a key load component in education buildings. Given the resources available, it was determined that the variables outlined in Figure 18 would provide a fair opportunity to test the performance of the methodology. Space usage and area-related variables form the majority of the model elements because of the detailed data available and also because of the great potential for space usage to indirectly affect electricity consumption; many indicators in Table IV are influenced to varying degrees by the primary activity for the space. The remaining variables include the number of above and below ground levels, a ratio for the unexposed building envelope, and a categorical variable representing six common footprint shapes. Greater details on these variables are disclosed in the following sections.

Table IV Exhaustive list of building variable categories that affect electricity consumption.

Measurable Indicators	Key Electrical Load Components			
	Ventilation (22%)	Cooling (21%)	Lighting (31%)	Plug Loads (9%)
Sources of Heat				
– Occupants [#/m^2]	✓	✓ 12		✓
– Equipment [# of Computers/Servers, Boilers, etc.]		✓ 22		✓
Equipment Efficiency [COP, %, etc.]	✓	✓	✓	✓
Heat Loss Through Envelope				
– Building Height				
– – Above-Grade Floors [#]		✓		
– – Below-Grade Floors [#]		✓		
– – Building Shape [Square, Rectangle, etc.]		✓		
– Thermal Resistance Properties				
– – Walls [RSI]		✓ 7		
– – Roof [RSI]		✓ 5		
– – Windows, conduction [U-Value]		✓ 7		
– Air Leakage [ACH]		✓ 5		
– Entryways [#]		✓		
– Window-Wall Ratio [%]		✓	✓	
Heat Gain Through Envelope				
– SHGC [#]		✓ 23		
– Building Orientation [N, S, E, W]			✓	
– Thermal Mass				
– Surface Albedo				
Conditioned Space (m^2)	✓	✓	✓	
Temperature Set-point (°C)		✓		
Space Usage				✓
– Ventilation Rate (L/s)	✓	✓		
– Lighting Density [W/m^2]		✓ 19	✓	
Elevators [hydraulic vs. mechanical. #]				✓
Time of Use [# of hours/day, # of days/week]	✓	✓	✓	✓

Notes:
- Numbers in Cooling column represent relative heat gain contributions during summer months in office buildings [98]
- Bold Entries are variables deemed significant; grey, insignificant, in Signor et al. [65]
- Electrical load components are from the US EIA for educational facilities [29]

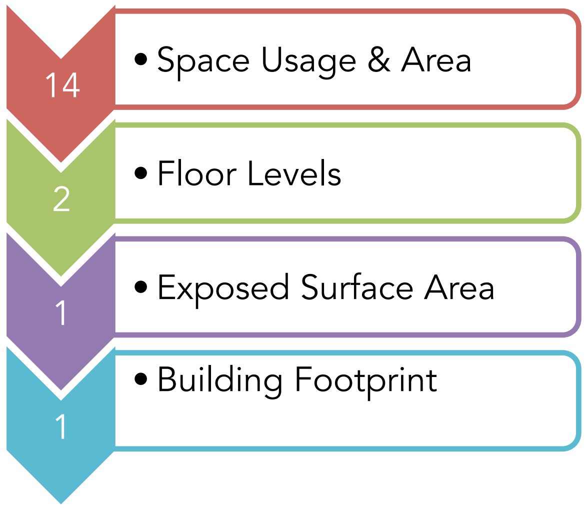


Figure 18 Number and type of variables analyzed for model creation

The form of the multiple linear regression is expressed in (1). In total, a maximum of 18 variables can be included in the regression. No climate-related variable is included in the regression because it is envisioned that coefficients will be updated annually, even if the model terms are the same from one year to the next. The following section will focus on gathering and priming data for the 18 variables, and the steps taken to create the resultant models which will eventually be used to estimate consumption for Ryerson buildings.

$$Y = b_0 + \sum_{i=1}^{18} b_i x_i \quad (1)$$

where

Y	is the annual electricity consumption;
b_i	is the coefficient of the variable;
x_i	is the value for the predictor variable;
b_0	is the error term or constant

A. Data Collection

1) Electricity Consumption Data: Measuring and verifying electricity consumption data for the buildings used in this thesis were outside of the scope of research. All consumption data for buildings used was provided by university staff from Ryerson University and the University of Toronto (U of T). Data in the form of monthly/annual kW and kWh for 23 meters on campus was provided by Ryerson's Campus Facilities and Sustainability [99]. The electricity meters represent 17 individual buildings and 14 buildings in five clusters – more information can be found in the Statement of Problem subsection and Figure 7. The supplied files contained user-entered summary sheets for the various meters on campus measuring electricity consumption. They were distinguished by utility meter and covered the academic fiscal years (May to April) between 1990 and 2009. An example of the consumption data provided can be found in Appendix C1.

A sample size of 17 buildings provides very little flexibility for statistical analysis. The sample set would be further reduced when preparing the training set due to the removal of unsuitable buildings/space types such as parking garages and non-traditional academic buildings. As a result, in order to increase the dataset to allow for more variations in built form and consumption patterns typical of university buildings, and to increase the chances of creating a reliable model, a request was sent to the University of Toronto to provide the necessary data in order for their buildings to be added to the Ryerson sample. U of T's 2013/2014 operating budget of $1.9 billion greatly exceeds Ryerson University's $500 million and is an indication of the size of the institution [100, 101]. Since weather and climate conditions are not being considered as a variable in the model (i.e. all buildings within Ryerson's campus will be exposed to very similar conditions due to its compactness), using U of T's buildings to supplement the sample size is ideal because the centers of each campus are roughly only 1.45 km apart [102], thereby reducing the climatic effects on energy consumption.

The University of Toronto supplied electricity consumption data for 120 buildings for the 2012 fiscal year. This number varies from year to year depending on new construction, demolition, and renovation activities taken place during that year. The consumption data supplied covers the years between 2005 and 2012, and is reported on an annual basis. In addition to providing the metered usage of a particular building, U of T has documented their estimates on the amount of energy gained or lost by receiving or providing chilled water to/from other buildings. At Ryerson University, two central chillers are responsible for supplying chilled water to service 77% of the total floor area on campus. Another 9% of the total floor area is serviced by Toronto's existing Deep Lake Water Cooling System, with the remaining area serviced by self cooling systems (i.e. non-centralized) [33]. The implications of this mixed system is that approximately 2/3 of all Ryerson buildings' electricity meters do not account for the energy spent on active cooling - these are accounted for in two buildings specifically, which result in inflated electricity readings. Between 1990 and 2012, the Library building cluster, which provides cooling for 10 other buildings, had an electricity intensity of 291 kWh/m^2/annum, a 53% increase over buildings relying on self cooling systems. To address this apparent issue, U of T has allocated a portion of energy consumption from buildings with central chillers to buildings being served chilled water. For the purposes of this thesis, the estimates for the consumption values with the chilled water allocation applied are used for U of T buildings.

To match the more recent years of data supplied by the University of Toronto, Ryerson supplied updated data for the months between 2010 and 2012. It was noted that there were differences in consumption for overlapping months in 2010 with the original dataset; in these cases, the more recent data replaced those months. It was determined that models would be created using the most recent data available instead of using the historic averages due to unique consumption patterns for campus buildings that were difficult to generalize (i.e. not a uniform increase or decrease of electricity consumption over time). Models for 2010 and 2011 academic years will also be pursued to determine the significance and sensitivity of parameters over time. However, only models created with 2012 data will be used for estimating consumption for Ryerson University buildings.

The sample of buildings from Ryerson University and the University of Toronto were refined by eliminating observations that were either suspect or did not adhere to certain characteristics typical of large academic institutions. First, outliers were identified for Ryerson's sample using the interquartile range calculated using historic data. Entries outside of 1.5 standard deviations from the upper and lower quartiles were flagged and omitted from further analysis, including the calculation of averages. Detailed historic data for utilities was not provided by U of T making it difficult for this type of outlier identification to be made with their buildings. Instead, outliers were rejected based on the coefficient of variance. This coefficient was first calculated for Ryerson buildings to determine the variance within the sample after outliers were removed. The maximum variance, 0.135, was found for consumption between 2010 and 2012. This variance was used as a guide to set an appropriate limit (0.15) for the coefficient of variance for U of T buildings. Buildings that exceeded the prescribed threshold were flagged for scrutiny; in some cases, the variance exhibited between 2010 and 2012 were spread across all three years while at other times, a clear distinction between one year and the other two existed. Depending on the situation, either the year with the suspect data was eliminated from the set or the entire building was removed. Second, the samples were scanned for suspect data entries potentially stemming from human error. This included entries that were repeated for several months but also included larger interruptions in reliable data including renovation work. For instance, Ryerson's bookstore was closed for almost 10 years for renovation and surrounding construction work leading to a significant drop in electricity consumption during those years. To preserve the accuracy of energy usage at the bookstore building, those entries were omitted because it was assumed that the building was closed to the public. In some cases, the months leading up to or after the closing or opening of a building is distorted which may be due to adjustments and optimization of building systems to meet the changes in the number of occupants. Often, these months are also identified during outlier identification, such as the opening of Ryerson's George Vasi Engieering and Computer Center, where three months were withheld from the final dataset. An example of outlier identification for Ryerson's individually metered buildings can be seen in Appendix C2. Third, select campus buildings have a very small footprint which are not typical for academic buildings. They often serve as sporadic administrative offices to be further developed when the need arises. In addition, smaller buildings are most susceptible to variations in energy use intensities. They are often difficult to model, regardless of the method of estimation, and therefore have been removed from the sample. For the purposes of this thesis, small buildings have been defined as ones with internal floor areas less than 1000 m^2. Related to this criterion are academic buildings in the form of houses. With the removal of smaller buildings, many houses have been removed from the sample, however larger houses can still persist. Since core academic buildings are typically not contained within houses, these were also removed. Parking garages, while not traditionally associated with academic buildings, exist in the sample because they are often not separately metered when constructed below a building. Because of this, freestanding garages were also left in the sample in hopes that the model will be able to associate car parks with low energy consumption. Figure 19 shows the proportion of eliminated buildings from the combined Ryerson and U of T building sample; the buildings eliminated under "Outliers" refers to a procedure yet to be detailed (refer to section C. Creating Subsets).

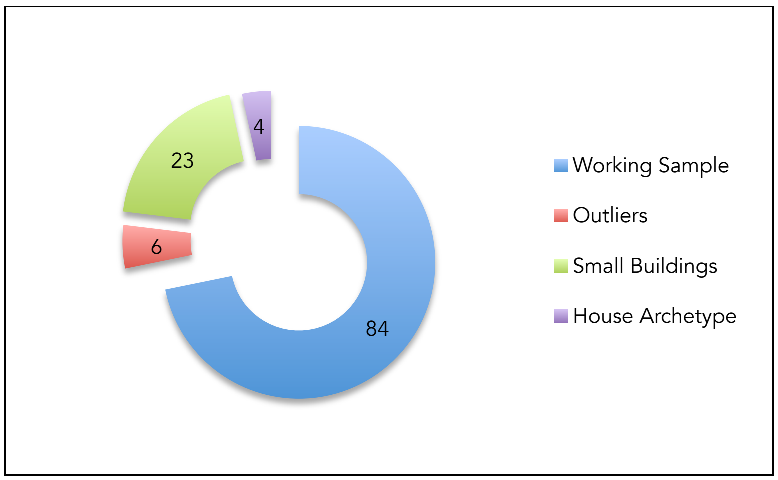


Figure 19 Number of buildings removed from the original sample for the 2012 model.

2) Council of Ontario University Survey Data: The Council of Ontario University (COU) represents 21 publicly funded institutions in the province of Ontario. As a member institution, the universities are required to measure and submit various operation and usage statistics on a regular basis. Included in these submissions are detailed area measurements for defined spaces in campus buildings. Examples of these space categories include central administrative office and related spaces and health service facilities which are both defined, along with the other 18 categories (a total of 47 subcategories) in [103].

The council of Ontario Universities first pursued a space formula and inventory classification system in 1967. Four years later, they established four task forces to examine questions relating to capital funding and the utilization of physical facilities within Ontario universities. One task force was dedicated to developing and testing a space utilization guide, two examined the space needs of education and health sciences, and the final task force studied the various aspects of building costs. In total, five reports (Building Block Series) were published from their collective findings which covered two major elements of the capital formula: space and cost. The eventual testing and subsequent revisions of the formulas were tested on five Ontario universities. In 1972, a subcommittee was appointed to continue the work on the Building Block Series and to receive comments and recommendations from Ontario universities. The space standards were reviewed and revised to reflect current teaching and research conditions in 1984 and in 1992, the first edition of the COU Building Blocks was published; periodic revisions of the standard have since taken place [103]. A comprehensive timeline of the development of the COU Space Standards can be seen in Appendix C3.

The original motives for developing a capital formula were to promote equity and objectivity, consistent standards, an opportunity for the provincial government to influence their financial obligations, and to provide an incentive for institutions to properly manage and allocate their resources. In order to accomplish these tasks a survey of physical facilities takes place on a triennial basis. Information on space inventory, student enrolment, full-time faculty and staff positions, lab contact hours, and library collection volumes is gathered between the months of January and April, sometimes with the help of students. This information is used to monitor changes in space needs, as well as to calculate capital funding requirements for the university system as a whole for the government. Outside the COU, the data is used by the Canadian University Reciprocal Insurance Exchange to determine the value of university facilities for insurance purposes. The Ontario Ministry of Training, Colleges, and Universities also uses the data to make informed decisions on allocating funds to each institution through the Facilities Renewal Program. Lastly, institutions themselves use the gathered data for a variety of reports including proposals and internal planning [103]. While the COU data is unique and extremely useful, it is not without faults. Specific to the inventory of space, the age of facilities is not considered despite it being recognized as having potentially significant impacts on efficiency and usability of spaces. Secondly, while the definitions of COU space categories are specific, there is room for interpretation between one administrator and another. Nevertheless, the opportunity to harness the gathered COU data to correlate with electricity consumption exists.

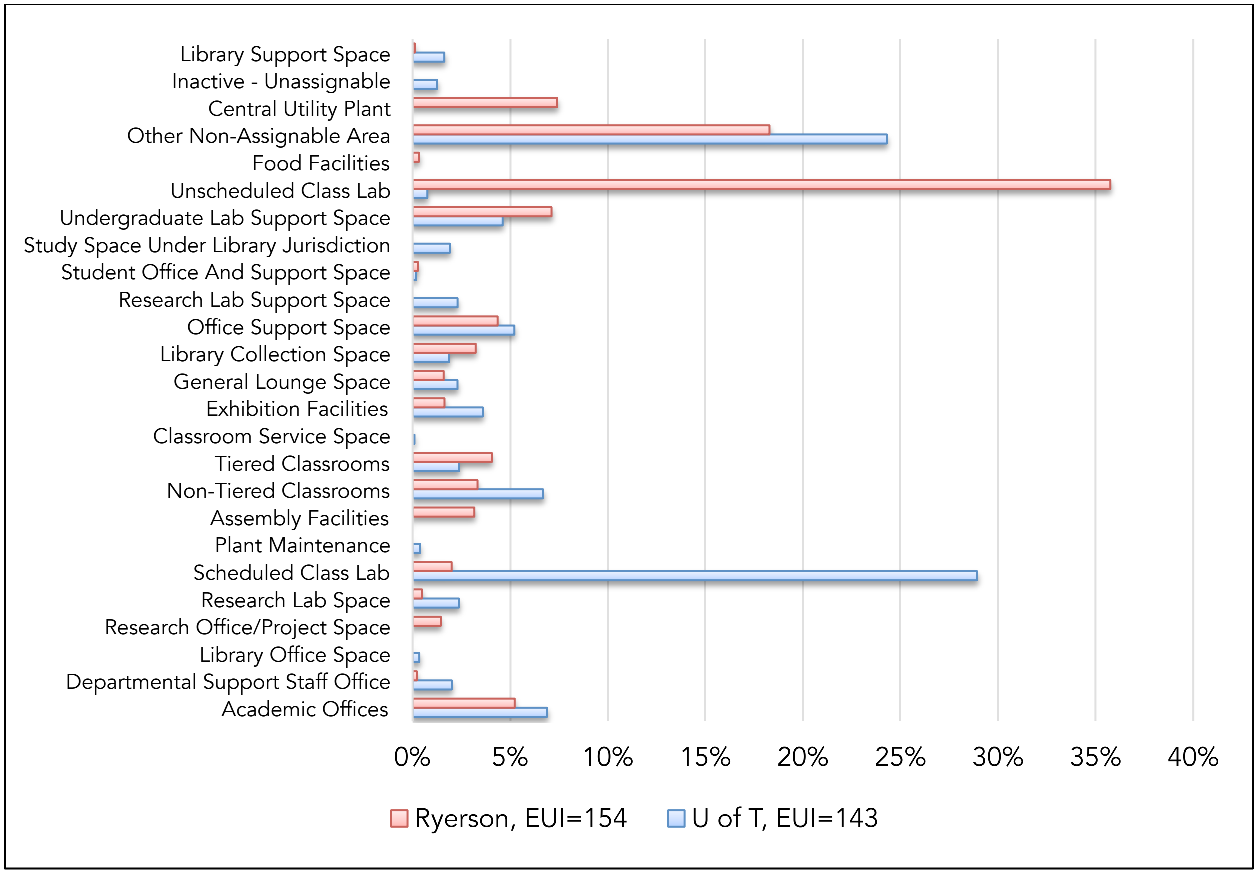


Figure 20 The proportion of spaces under COU space categories for architecture buildings at the University of Toronto and Ryerson University.

2012 data for COU space categories was gathered from Ryerson University [104] and U of T [105] from their respectful facilities and planning offices on interior areas defined within each category for every building on campus. At U of T, an official audit of spaces is taken every three years by upper level administrative staff for the department while at Ryerson, a less formalized approach is taken with ongoing updates to the VisonFM database throughout the year by a Facility Analyst. Figure 20 illustrates the space usage at architecture buildings from both universities. Typical non-specific education facilities at Ryerson University tend to dedicate roughly a third of their spaces each to offices, laboratories, and "other" spaces. More specialized or smaller buildings, and those housing centralized heating/cooling equipment, are excluded from this generalization. Spaces categorized under "other" include circulations spaces, washrooms, elevator shafts, and stairwells, among other things. Aggregating the COU categories based on perceived end-uses results in a simplified and easier to interpret diagram (Figure 21).

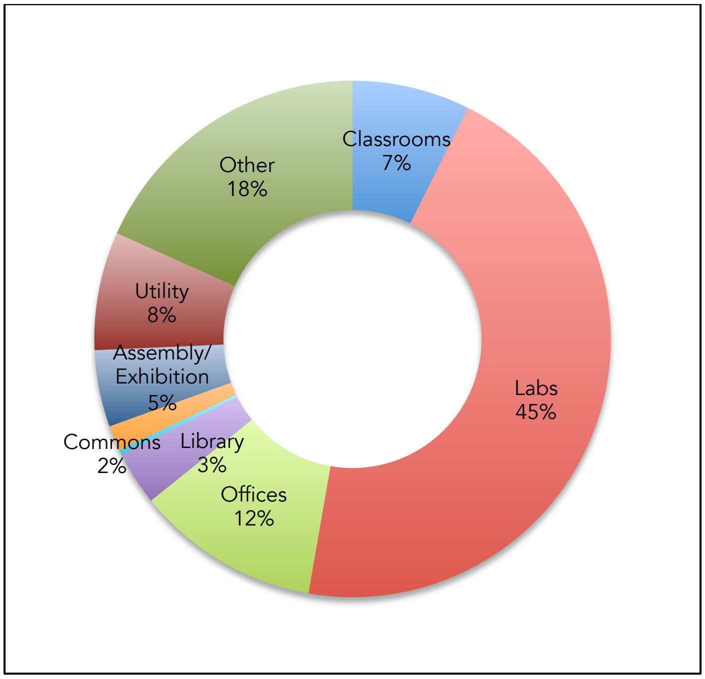


Figure 21 Areas attributed to primary end-uses based off of COU categories for Ryerson University's Architecture Building.

Table V COU space categories with a proportion of total area greater or equal to 3% across both Universities.

COU Space Category	Area (m^2)	Proportion of Total Area	Cumulative Proportion
Other Non-Assignable Area	389421.05	33%	33%
Research Lab Space	80970.08	7%	40%
Academic Offices	58624.81	5%	45%
Residence Living Space	56056.83	5%	50%
Scheduled Class Lab	49057.67	4%	54%
Office Support Space	39989.56	3%	57%
Parking Structures	36499.98	3%	60%
Library Collection Space	36418.98	3%	63%
Non-Tiered Classrooms	35497.32	3%	66%
Departmental Support Staff Office	32404.96	3%	69%

Of the 53 space categories that define Ryerson University and the University of Toronto's space usage, four have a proportion of total area of 5% or more, (Table V, a full list can be seen in Appendix C4). If the raw space categories were used to develop a model, many buildings would be excluded from the sample because they lack those specialized spaces that may be required to estimate consumption. The robustness of the model would suffer with the inclusion of highly specific space usage variables. For this reason, an attempt was made to group COU space categories together based on estimated energy use intensities (Figure 22 & Figure 23). For instance, daycare facilities and lounge spaces serve distinct functions however their energy demand is expected to be similar due to how the occupants are behaving. Another example of grouping similar categories is athletic activity and seating spaces; many of these spaces refer to arenas and courts where the seating and activity area are conditioned as one space. After a quick visual assessment of how university spaces at Ryerson were categorized, the COU data was simplified based on the following criteria: occupant density, hours of use, conditioning requirements, plug loads, and lighting intensity. This re-categorization of spaces is a necessary critical step that will undoubtedly affect the performance of the model. The risks assumed here are less than those of alternative pathways: either selecting a sample of space usage categories based on their perceived effects on electricity consumption or simply basing selection on a space's popularity on campus. The first option, creating models with a select few categories, requires an amount of justification that doesn't exist. A lack of precedent studies and knowledge for linking categories with their effects on electricity consumption makes it difficult to establish confidence when prioritizing certain categories over others. Further, how many categories should be selected from the set, and what balance of space categories with high and low energy demand should be considered? These unanswered questions further deterred progress down this path. The second option, using only the most common space usage categories, is illogical. If significant categories were arbitrarily defined as those with a proportion of 3% or more, that would result in 69% of all areas accounted for. Not only are there spaces that share similar EUIs within the selected set (Table V), but also, almost one third of spaces are not considered. Since it has been previously established that academic spaces suffer from highly variable EUIs, many of the eliminated spaces may have a very high impact on energy consumption in a few number of buildings. A self-directed reorganization of the COU data was superior to the other two options discussed for another crucial reason. Since there are numerous individuals interpreting the space definitions, there may be cases where staff members categorize similar areas differently – especially where multiple subcategories exist. For example, the definition for Category 4.2 Research Office/Project Space is very similar to Category 2.2 Unscheduled Class Laboratory Space despite the fact that they belong to different categories. By grouping categories based on expected EUI, the bias introduced by many human actors interpreting the COU space definitions is replaced solely by the author's own accountable and disclosed bias. The grouping of COU categories is one of many possible ways to extract information on electricity consumption tied to certain space types. It is evident that further work in this area of how best to utilize COU space usage data is needed to make better-informed decisions for future studies.

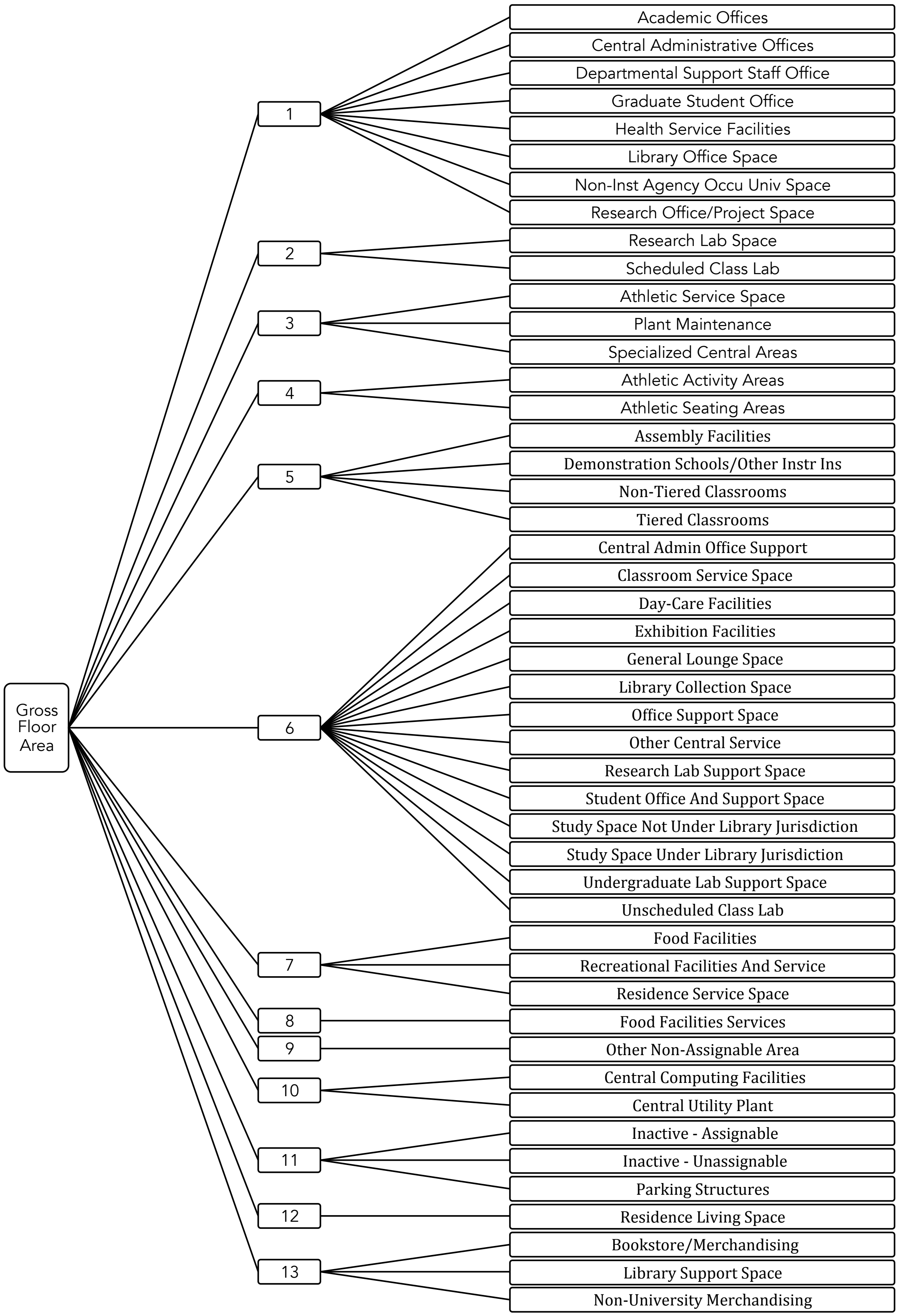


Figure 22 Process of combining COU space categories into 13 groups based on energy profile

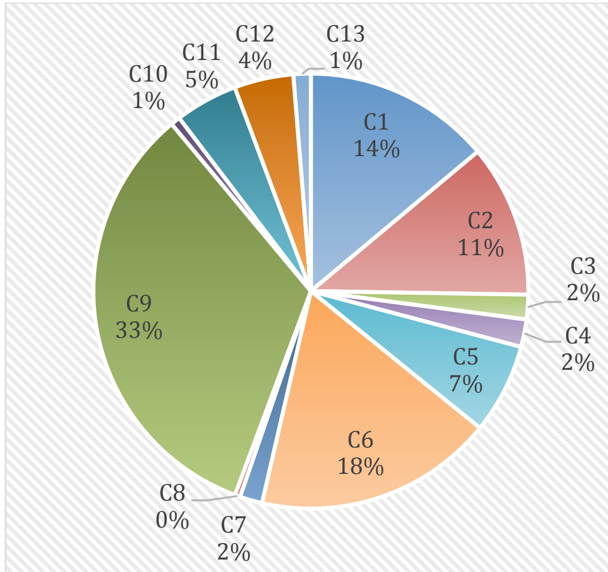


Figure 23 Proportion of area after amalgamation

3) Other Building Data: The data for the remaining variables that were collected – building footprint shape, the number of shared walls, and the number of above and below ground floors – were relatively less work intensive to prepare before analysis. The data on above and below ground floors was provided by Ryerson and U of T [104, 105]. No further differentiation was made between levels serving as an extension of the building and those serving as underground parking (which was only identified in one of the buildings) – the space usage variable partially accounts for that. The number of floors above ground level was obtained by taking the highest reported floor – ignoring floors where no assignable areas existed. In a few cases, the floors occupied by the universities only accounted for a fraction of all floors in a building. This scenario causes a misrepresentation of the actual number of floors the university occupies. For instance, if a university owns the 16th and 17th floor of a 20-floor building, the data entry for that building would read 17 above ground floors; 17 floors is nether the highest floor count for the building nor is it the number for floors occupied by the university. Fortunately, this issue exists only in a few buildings of the total sample and only for above ground floors; nine buildings from U of T and Ryerson exhibit occupation on fewer floors than the total count, and even less are found in the working sample. A potential solution to this problem is to split the two floor count variables into three: one for accounting for the number of floors the university occupies in a building and the other two for documenting the highest and lowest floor of the building, regardless of the tenant. Applied to the example above, the variables for floor count would be two, for the number of university occupied floors, and 20, for the maximum floor count above ground. Whether or not this will have a measurable impact on model performance (given the infrequency of this issue) should be a focus in future work.

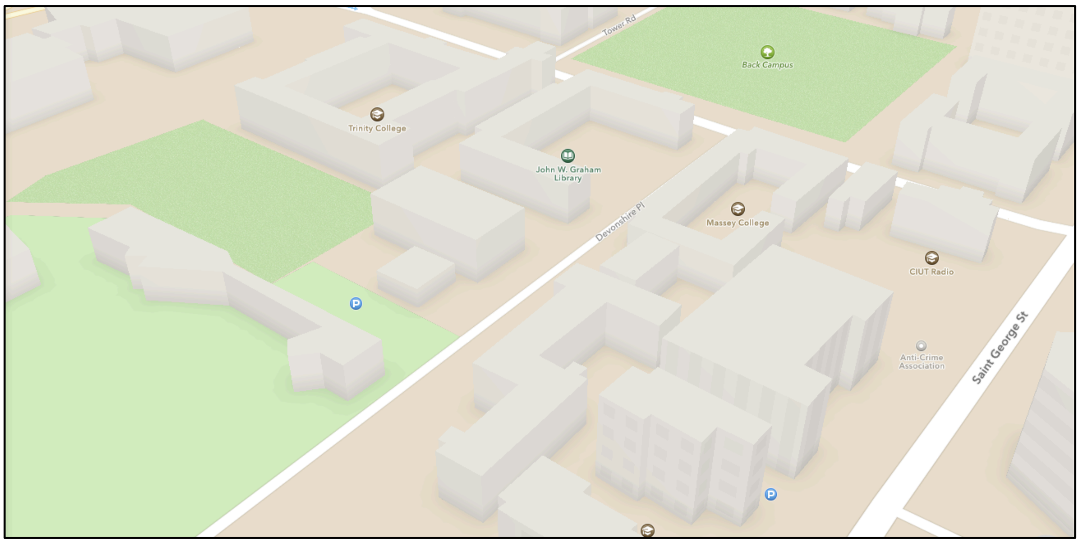


Figure 24 Screenshot of Apple's Maps program which was used to survey building geometry on both campuses.

Building footprint shape and the number of shared walls was measured by analyzing various cartographic data supplied by the universities [31, 106], Google [102] and Apple [107]. 3D satellite imagery and Google's Street View feature allowed the survey of building geometry and their surroundings to be completed efficiently (Figure 24). With respect to building shape, six footprint shapes were selected, for the purposes of this thesis, to categorize the academic buildings found within the working sample (Figure 25). In general, categorizing footprint shapes was a simple task with only a few buildings causing any conflict. Those buildings tended to have podiums where the tower portion of the building had a different floor shape than the base. Also some buildings consisted of multiple buildings which differed in shape. For these instances, a null value was used when no prevailing shape could be determined. This was done on a case-by-case basis as each instance presented a unique set of characteristics to consider.

A value ranging from one to six represented the footprint shapes; with this being the only categorical variable considered in the equation, the value is of no importance.

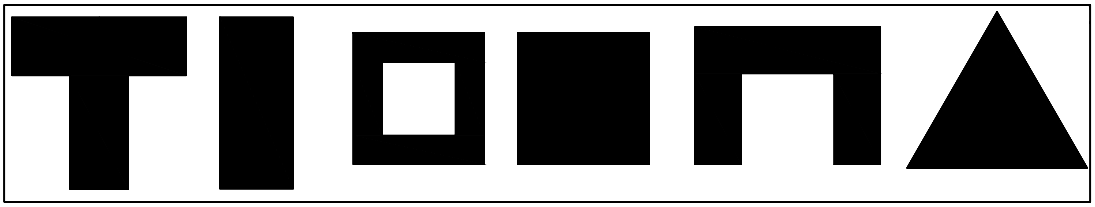


Figure 25 Six footprint shapes used to categorize academic buildings

The variable "shared walls" represents the number of surfaces (walls and roof) the building envelope comes into contact with, from neighboring buildings. It is a simple measure to account for the amount of exposed surface area, a factor that has been associated with energy calculations [65]. For the greater majority of buildings (89%), no shared walls were found. Those with shared surfaces were specifically for walls, with the exception of one Ryerson building – a bookstore whose roof was under an above-ground parking garage. A shielded roof affects heat loss and energy consumption for a building to a greater extent, when compared to a wall. Despite this, no differentiation was made in this variable between surface types because of the limited instances of a covered roof in the studied sample. Partial shared walls were accounted for by using fractions. A drawback for how this variable is measured is that it does not consider the relative area that is covered and uncovered. Potentially, this means that two buildings sharing the same number of shared walls may perform differently due to their surface area to volume ratio. Nevertheless, this tradeoff is made to promote simplicity in the methodology. An attempt was made to reduce the potential for an inflated shared wall variable by only counting surfaces where a significant surface area relative to the total was shared with another building. This added step may be difficult to instruct future users on.

B. R Programming

The statistical software used to create and test multiple linear regressions is R [108]. R is both the language and the environment for statistical computing and graphics. The term "environment" is intended to characterize it as a fully planned and coherent system, rather than an incremental accretion of very specific and inflexible tools, as is typically the case with other data analysis software [109]. It is popular among statisticians and data miners due to its flexibility and extensibility. It was first developed in 1993 by Ross Ihaka and Robert Gentleman and has grown substantially in recent years, both in popularity and in advancement. A more detailed history of R and its contributors can be seen in Ihaka [110].

Without a native graphical interface, R relies on a command line interface to interpret the language R, which is similar to MATLAB. The primary functions used to obtain the regression models will be detailed below which include: multi-model inference, cross-validation for generalized linear models, and stepwise variance inflation factor selection. Each function's script can be seen in Appendix C5.

1) Stepwise Variance Inflation Factor Selection: In order to build statistical models with significant predictor variables, collinearity among variables needs to be addressed. This is especially true for this thesis because data dredging was used and also because of the similarities in predictor variables (i.e. 14 of them dealing with floor area). By reducing collinearity, independent variables are distanced from one another resulting in stronger and truer relationships in the resultant models. Collinearity is represented by variance inflation factors (VIF) where a higher number equates to stronger collinearity among variables. VIF is calculated by taking the inverse of the coefficient of determination (r^2) for the regression between one predictor variable, j, against all others (2).

$$VIF_j = \frac{1}{1-R_j^2} \tag{2}$$

The VIFs are then used to screen out variables with high collinearity. The custom script from Beck [111] allows for the stepwise selection of variables based on VIF (i.e. VIF is calculated each time a variable is removed from the set). Since the VIFs change significantly with each iteration, this ensures that all remaining predictor variables are under the defined threshold.

2) Multi-Model Inference (Dredge): As eluded to previously, data dredging was the method used to create candidate models. In R, the dredge function creates all possible combinations of explanatory variables and ranks them by a defined information criterion – in this thesis, Akaike Information Criterion is used (AIC). The number of models evaluated in Dredge is exponentially related to the number of predictor variables; with the removal of collinear variables, the number of models fitted in each subset (refer to section C. Creating Subsets) did not exceed 262,144. AICc denotes AIC adjusted for smaller sample sizes – a greater penalty is assigned for extra model parameters compared to AIC. Both measures are interpreted the same however due to the small sample size in this thesis, AICc is used. AIC is a metric used to evaluate competing models by considering the goodness of fit as well as the complexity. As opposed to r^2 or adjusted r^2, a measure of explained variance in the response variable by predictor variables, AIC seeks to identify parsimonious models [73, 112]. AIC is a relative value and cannot be compared to models from other subsets or other studies. The difference in AIC between models dictates the amount of information gained or lost; a delta between the lowest AIC and a competing model's is substantial if under two. Information can still be gained from models with a delta of 10 but with much less support [73]. When there are several models that have a delta of less than two, it is possible to average the models, a process known as multi-model inference [113]. A drawback of relying on dredge results to formulate a model is that its coefficients will be specific to the sample used for training. Using AIC will guarantee the best fitting model, not the most universal model. This is acceptable for our applications, but for research towards a model applicable to a wider scope of academic institutions, dredge may not be appropriate.

3) Cross-Validation for Generalized Linear Models: In order to validate the models identified through Dredge, they need to be tested on a new set of data that was not used to construct the model. Because the sample size was limited in each subset, the group was not further divided into a training and test set which is what is commonly done when data is not scarce. Instead, Leave-one-out Cross-validation [114, 115], where one observation from the subset is omitted during model building and used to validate the subsequent model, was used to gauge performance – comparing the predicted and actual electricity consumption for that building. Examples of studies dealing with energy consumption in buildings using cross-validation can be seen in [116] and [117]. The removed observation was then reintegrated into the subset and used in the next iteration of model building and testing – with a new building being used for validation. This was repeated until all buildings in the subset had been "left out" once (i.e. used to test), and an associated error rate had been calculated. In total, a subset containing 20 buildings had 20 unique errors that were then averaged for each model in the form of a mean square error (MSE). A general schematic of this method can be seen in Figure 26.

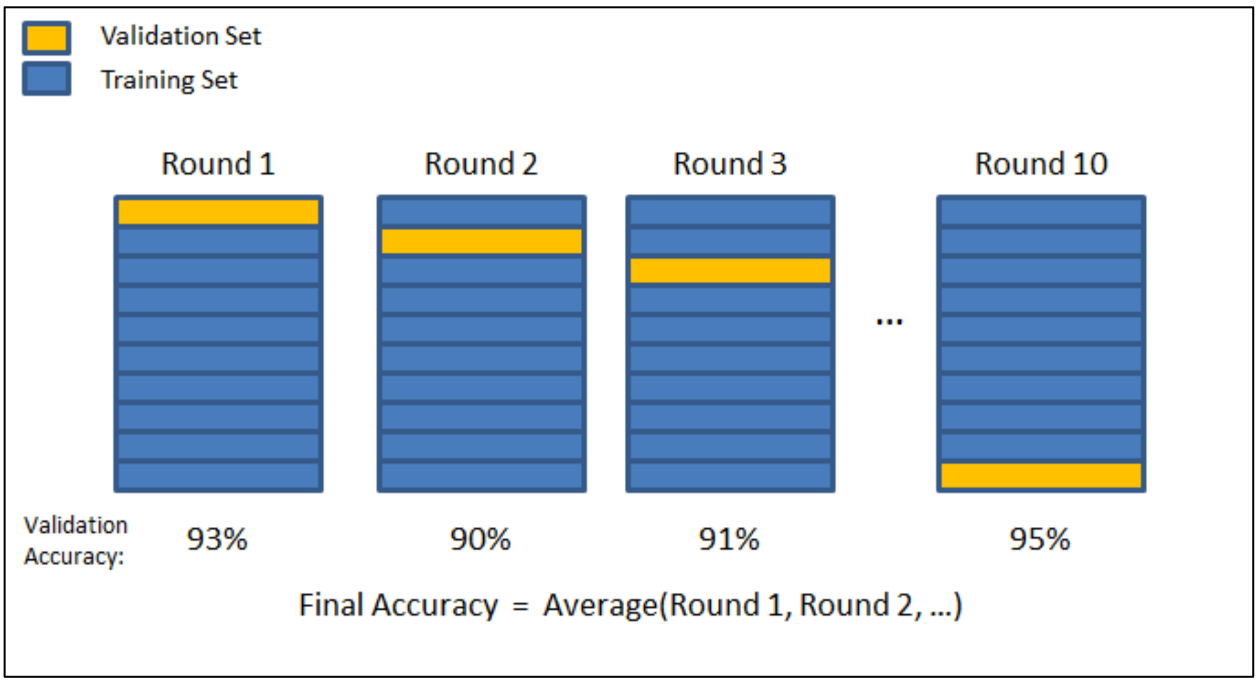


Figure 26 Example of leave-one-out cross validation with 10 samples. [118]

C. Creating Subsets

Breaking down the sample into subsets was required in order for a variety of building types and sizes to be considered. Even after eliminating buildings with an interior area below 1000 m^2, removing houses from the sample, and omitting outlier data, there was a great deal of variation in area, electricity consumption, and energy use intensity, as shown in Table VI. Creating a static model to represent this level of diversity would likely result in poor and unreliable performance. In order to address this, subsets of the sample set were created so that multiple static models could be created to represent each group. There were many options available on how to divide the sample such as by archetype, or construction date however those would require additional resources for future users to collect. Instead, the total interior floor area was used as the boundary condition to separate the buildings. Since many of the variables in the model are derived from floor area, it is important to divide buildings in such a way that their EUIs are distinct from one group to another. Figure 27 shows the relationship between two possible indicators and EUI: interior floor area and whole building electricity consumption. Selecting a factor that yields a strong correlation with EUI is ideal but since both factors show a weak relationship (r^2 = 0.36 for electricity consumption and 0.19 for floor area), the criteria shifts to one that spreads EUI values the most. Floor area is superior in this aspect which is partially why it was chosen to define subsets. The other reason behind its selection was the difficulty for buildings with unknown consumption levels to be placed in one subset; determining which model is most applicable to estimating consumption is a critical step in this methodology. The boundaries were created by using interquartile ranges of the building interior floor areas. This ensured that the number of individuals in each subset are similar, if not equal. University campuses tend to have a few exceedingly large buildings that offset the rest of the buildings (Figure 28) – if equal ranges in area are used to divide the sample, the number of individuals in each group would be highly skewed. The number of subsets created was arbitrarily set at four however performance with three and five sets were created for performance tests.

Table VI Descriptive statistics of the building sample from Ryerson and University of Toronto after the elimination of small buildings, houses, and outliers.

	Electricity Consumption (kWh/Annum)	Area (m^2)	EUI (kWh/m^2/annum)
Min	109,400	1029	26
Max	21,459,710	82,688	564
Range	21,350,310	81,659	538
Mean	2,561,098	10,772	238
Median	1,106,517	6,951	204
Standard Deviation	4,266,670	12,507	105

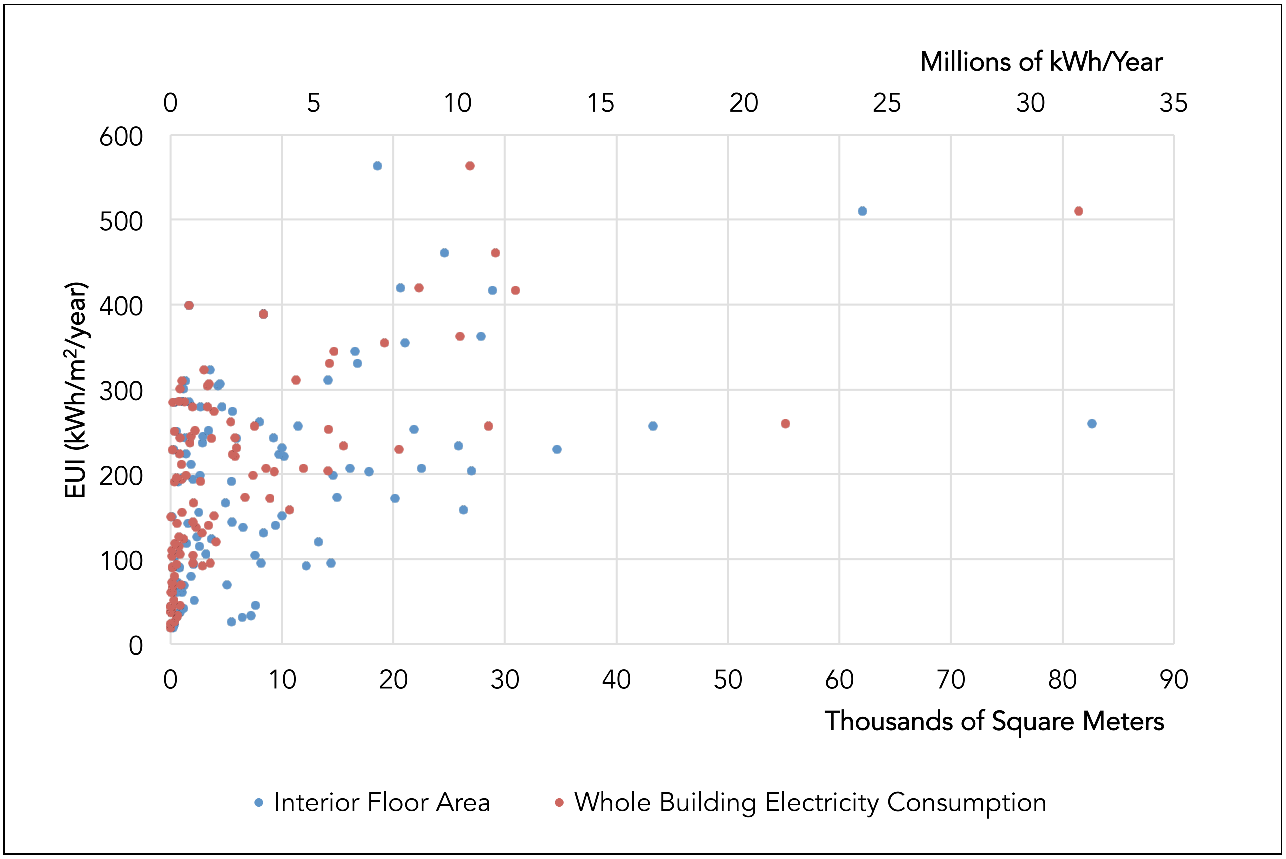


Figure 27 Correlation between EUI and floor area, and EUI and electricity consumption for sample buildings.

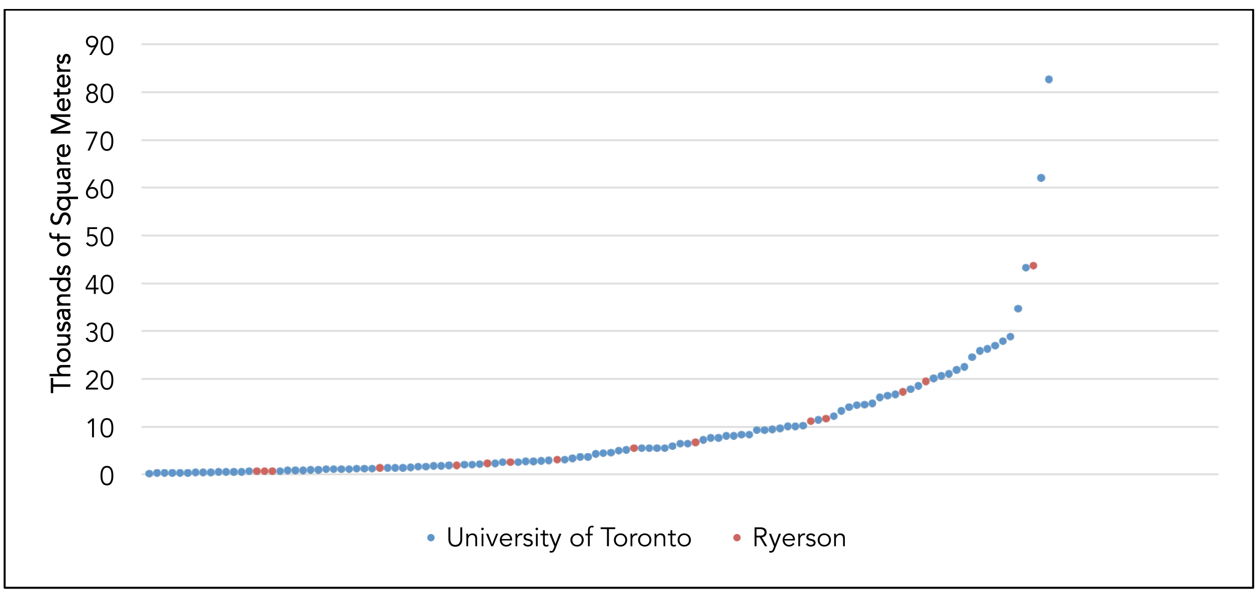


Figure 28 Ryerson and University of Toronto buildings sorted by area.

Within each subset, an additional test for outliers was done to ensure that variance within each subset was acceptable. Using the standard deviation method, as outlined in the Electricity Consumption Data subsection, buildings with a consumption value outside of 2 standard deviations away from the mean of the subset were withheld. This resulted in four buildings being removed from the 2012 and 2010 years, and three from 2011. The areas used to define each subset are shown in Figure 29; Table VII displays descriptive statistics for each of these defined subsets.

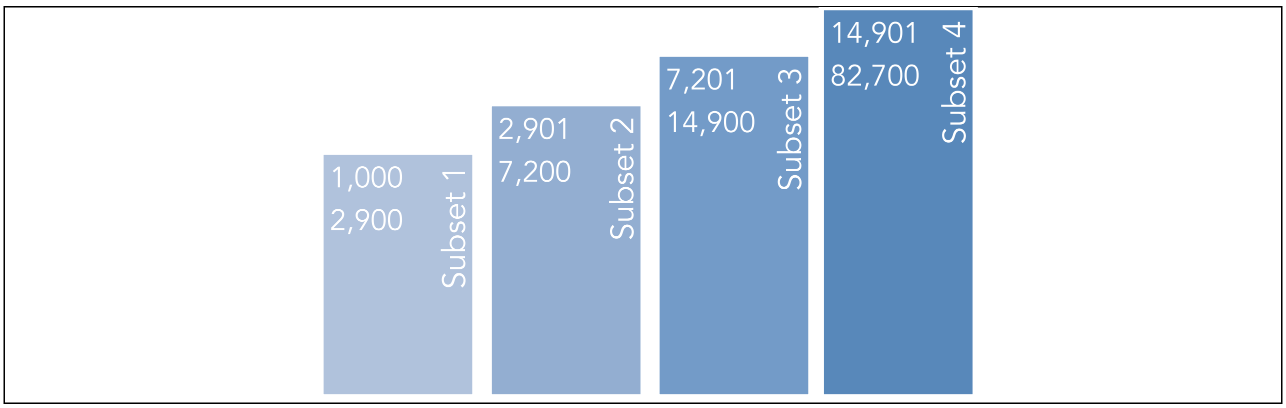


Figure 29 Minimum and maximum areas defining each of the four subsets.

Table VII Descriptive statistics for Ryerson and University of Toronto buildings in four subsets (n=80)

Subset	1			2		
	Electricity Consumption (kWh/annum)	**Area (m^2)**	**Average EUI (kWh/m^2/ annum)**	**Electricity Consumption (kWh/annum)**	**Area (m^2)**	**Average EUI (kWh/m^2/ annum)**
Min	109400	1029		141602	2961	
Max	681737	2880		1522560	6683	
Range	572337	1851		1380958	3722	
Mean	352335	1875	188	839897	4809	175
Median	319650	1832		839171	4998	
Standard Deviation	148204.5	599		430184.9	1192	
Subset	3			4		
Min	241920	7219		2584200	14950	
Max	3240728	14577		21459710	82688	
Range	2998808	7358		18875510	67738	
Mean	1634703	10133	161	7417457	26269	282
Median	1555273	9860		5865227	21414	
Standard Deviation	881249.4	2240		4488290	14971	

D. Building Subset Details

Data for each of the defined subsets used to create and test the models is displayed in Tables VIII – XI. As discussed in the previous Section, subsets were created to reduce the variance observed with electricity consumption in academic buildings and to allow for more accurate models to be created. The subsets are defined by interior floor area which will also determine which of the four equations are used to estimate consumption for a particular building. All area-related variables (c1-c12 & Interior Floor Area) are measured in square meters and “2012 Electricity Consumption”, in kWh.

Table VIII Subset 1 predictor variables and electricity consumption

Building Count	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	Interior Floor Area	Footprint Shape	Above Ground Floors	Below Ground Floors	Shared Wall	2012 Electricity Consumption
1	110	0	535	0	0	26	0	0	357	0	0	0	0	1029	5	2	0	0	294850
2	0	0	0	0	0	697	0	0	375	0	0	0	0	1072	5	2	1	0	209600
3	125	330	0	0	28	292	0	0	356	0	0	0	0	1130	4	3	1	0	339120
4	0	0	0	0	0	0	0	0	250	0	0	0	1052	1302	4	2	0	3	231177
5	443	0	0	0	116	244	0	0	509	0	0	0	0	1313	5	4	1	0	406800
6	44	0	9	0	0	167	523	211	326	0	0	75	0	1355	5	3	1	0	329200
7	613	0	0	0	0	366	0	0	403	0	0	0	0	1382	4	5	0	0	310100
8	317	371	0	0	100	361	0	0	427	0	0	0	0	1576	5	3	1	0	224540
9	113	0	8	0	293	154	374	15	700	0	0	0	0	1657	6	3	0	0	661080
10	460	447	0	0	0	363	0	0	430	0	0	0	0	1699	5	3	0	0	485120
11	446	593	0	0	0	216	0	0	637	72	0	0	0	1964	5	4	1	0	401344
12	955	0	0	0	0	380	0	0	678	0	0	0	0	2013	4	3	1	0	391360
13	251	26	0	0	929	280	0	0	626	0	0	0	0	2112	3	3	1	0	109400
14	308	883	0	0	0	497	0	0	506	78	0	0	0	2272	3	3	0	0	236742
15	623	0	5	0	338	522	0	0	848	0	0	0	21	2358	3	3	1	0	297851
16	708	0	164	0	125	502	0	0	900	0	157	0	0	2556	4	5	1	0	396000
17	495	221	0	0	0	815	0	0	609	7	248	0	171	2565	5	4	1	0	211834
18	389	0	115	0	0	1002	20	0	1023	33	11	0	0	2592	5	6	1	0	298040
19	679	317	21	0	169	567	0	0	912	0	9	0	0	2674	2	3	0	0	530800
20	483	0	0	0	553	610	111	52	1065	0	0	0	6	2880	3	3	1	0	681737

Table IX Subset 2 predictor variables and electricity consumption

Building Count	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	Interior Floor Area	Footprint Shape	Above Ground Floors	Below Ground Floors	Shared Wall	2012 Electricity Consumption
1	470	818	0	0	485	379	0	0	727	0	82	0	0	2961	4	4	1	0	723494
2	246	1435	0	0	0	653	0	0	672	54	0	0	0	3060	4	3	0	0.5	301817
3	1145	229	0	0	387	287	0	0	1011	110	0	0	0	3169	5	4	0	0	336000
4	1410	0	23	0	81	634	0	0	1276	0	0	0	0	3424	2	3	1	0	860558
5	597	900	29	0	151	593	0	0	1044	0	263	0	0	3576	2	3	1	0	1153962
6	687	0	114	0	213	707	0	0	1222	209	537	0	0	3689	5	5	1	0	457916
7	1627	212	68	0	0	734	0	0	1645	0	0	0	0	4285	5	4	1	0	1304471
8	791	0	11	0	456	1217	42	9	1880	0	0	0	0	4406	2	3	1	0	1350680
9	1088	1422	33	0	344	344	0	0	1387	0	0	0	0	4618	5	4	0	1	1291026
10	1522	0	0	0	0	2001	0	0	1346	0	0	0	59	4928	2	3	1	0	817783
11	360	0	1045	0	0	854	0	0	1626	0	1184	0	0	5069	4	3	1	0	352100
12	327	1529	0	0	0	1838	0	0	1244	34	0	0	479	5451	4	4	1	1	787256
13	1134	1129	20	0	164	1152	43	0	1725	0	91	0	0	5457	4	5	1	0	1048717
14	729	363	26	0	872	927	0	0	1845	0	664	0	35	5461	2	4	0	0	141602
15	512	1735	21	0	503	1267	0	0	1347	0	69	0	90	5545	2	5	1	0	795000
16	1413	60	70	0	733	906	0	0	2106	0	0	0	271	5558	5	7	1	0	1522560
17	1159	1169	0	0	98	1157	0	0	1824	0	490	0	0	5896	3	5	2	0	1427445
18	0	0	19	0	0	394	540	405	1807	0	0	3280	0	6445	2	3	1	0	200200
19	75	0	1375	3194	0	3	0	0	1852	0	0	0	0	6500	5	2	0	0	896091
20	460	166	0	0	705	3605	22	0	1222	495	0	0	8	6683	5	4	0	0	1029254

Table X Subset 3 predictor variables and electricity consumption

Building Count	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	Interior Floor Area	Footprint Shape	Above Ground Floors	Below Ground Floors	Shared Wall	2012 Electricity Consumption
1	316	888	277	0	177	216	0	0	2476	0	2869	0	0	7219	2	3	0	0	241920
2	1009	0	0	0	0	3588	0	0	2915	0	101	0	0	7612	5	4	0	0	792182
3	0	0	17	0	0	0	1275	26	1931	0	0	4382	0	7631	1	3	1	0	346240
4	1310	1568	259	0	349	1396	0	0	2718	0	394	0	0	7993	2	4	2	0	2093025
5	2533	0	484	0	279	470	0	0	2516	0	1832	0	0	8114	2	6	2	0	774000
6	1828	0	10	0	1492	1122	0	0	3834	16	11	0	9	8321	5	2	1	1	1089504
7	1467	2553	0	0	603	1377	0	0	2338	0	0	0	0	8337	3	4	1	0	3240728
8	1433	2889	25	0	406	766	31	0	3729	0	0	0	0	9279	3	4	2	0	2255508
9	1117	0	24	0	887	3104	210	1	3313	0	502	0	287	9447	3	4	0	0	1323000
10	2160	116	0	0	158	1472	15	0	3056	0	62	0	2678	9716	6	3	0	0	2171462
11	0	0	22	0	0	0	698	0	2529	0	0	6639	117	10004	2	7	0	0	1514970
12	4887	0	41	0	171	1513	0	0	2692	0	714	0	0	10018	3	10	1	0	2312893
13	2989	0	1834	0	31	1947	0	0	2922	80	391	0	0	10193	5	10	1	0	2258450
14	157	0	0	0	0	0	0	0	254	21	10747	0	0	11178	5	4	0	0	483024
15	2754	2183	99	0	1169	1612	27	78	3490	0	13	0	0	11424	4	10	1	0	2934000
16	139	0	94	0	139	689	1472	306	1629	71	2607	4529	0	11675	3	11	2	0	1872826
17	1353	1206	14	0	2713	2743	0	0	4079	0	0	0	88	12196	5	3	2	0	1123530
18	1763	77	106	0	1874	3763	28	13	5094	0	368	187	18	13290	1	3	1	0	1595575
19	1174	428	305	584	6559	502	0	0	4889	0	0	0	0	14441	2	3	1	0	1377900
20	465	0	19	0	538	594	993	0	4833	0	0	7039	97	14577	4	17	1	0	2893328

Table XI Subset 4 predictor variables and electricity consumption

Building Count	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	Interior Floor Area	Footprint Shape	Above Ground Floors	Below Ground Floors	Shared Wall	2012 Electricity Consumption
1	495	0	1318	3196	1500	1473	1441	852	4185	0	232	122	135	14950	1	3	2	0	2584200
2	5570	408	103	0	844	2333	36	0	5509	0	1251	0	0	16055	5	7	1	0	3323914
3	2984	2864	25	0	1263	4137	390	175	4621	0	0	0	103	16560	4	4	1	1	5706302
4	2911	1084	0	0	4604	2583	68	6	5915	174	0	0	0	17344	4	9	0	0	3670595
5	618	0	90	0	0	9309	135	24	6424	0	78	0	1100	17778	5	2	3	0	3618075
6	688	4943	1474	0	0	2838	0	0	8582	0	0	0	0	18525	5	13	1	0.5	10441613
7	2023	5209	93	0	2714	2301	0	17	5305	1764	0	0	0	19427	5	5	2	0	4244977
8	343	0	32	0	0	4	646	0	5038	0	5269	8807	0	20139	1	10	2	0	3459651
9	967	6217	689	0	694	4927	0	0	6412	3	734	0	0	20643	5	7	1	0	8656200
10	3945	7559	0	0	872	2041	0	0	6590	0	0	0	0	21007	2	4	1	0	7451414
11	2650	4621	80	0	970	5156	248	0	6549	0	1519	0	28	21821	4	5	1	0	5519853
12	2465	9446	0	0	842	3945	0	0	7812	0	0	0	67	24577	3	7	1	0	11339730
13	1234	0	99	0	897	2008	3991	786	9407	0	45	7333	37	25836	3	5	1	0	6024152
14	6746	2202	47	0	3162	3198	503	175	9764	0	482	0	43	26320	3	6	2	0	4149488
15	1698	674	3212	12439	414	659	73	0	7388	0	367	0	74	26998	5	3	1	1	5501120
16	5212	5137	51	0	1472	4462	36	0	9237	655	1574	0	29	27863	5	16	2	0	10100100
17	3411	6562	65	0	993	6244	0	0	11238	0	258	0	90	28861	1	5	1	0	12025508
18	8585	448	333	0	5742	7596	26	87	10818	151	514	0	404	34705	5	12	2	0	7963436
19	6477	4472	112	0	3162	4745	63	40	15023	0	9060	0	135	43288	2	8	5	0.5	11109106
20	7706	198	455	0	1421	33528	517	290	31305	621	853	0	5795	82688	6	14	2	0	21459710

A great deal of variation can be seen in the re-categorized COU categories (Figure 22) which is to be expected. Specialized space variables such as c12 (living spaces), and c4 (athletic spaces) are almost nonexistent in smaller buildings. c9, mainly representing circulation spaces, is closely related to the total interior floor area of the building and can be found in all buildings. The majority of buildings across all subsets do not share a surface with one or more buildings and have at least one below ground level.

IV. RESULTS

A. Candidate Models

The results from Dredge provided an exhaustive list of models for each subset. The list of models with an AICc delta of 7 or less can be seen for each subset in Appendix D1 [73]. The models used for further analysis were the top five models ranked by AICc for each subset (Table XII). The number of variables chosen for the models ranged from one to five with the larger buildings generally utilizing more variables. One reason for this may be that larger buildings have a value for more space categories than smaller ones. Another interesting observation is that space categories 5, 6, 7, and 11 consistently have negative coefficients while categories 2, 3, and 9 have positive ones. Aside from the obvious linkages that can be made to the energy demands of each unique space (which cannot completely explain the observations), the negative and positive coefficients can be attributed to the expected hours of occupation for those spaces. Using the results from Davis and Nutter [119], who assigned occupancy factors for common university spaces, a common factor among space categories assigned a negative coefficient were their scheduled usage. For instance, classroom, their support spaces, and student eateries all handle occupants at regular intervals during the week. This is in contrast to the spaces with positive coefficients such as research laboratories and plant/maintenance areas which may handle the same number of occupants but dispersed over many hours of the week. This deduction is only a possibility of the true mechanisms at work; further analysis is required to make any further conclusions.

The weights of space-related variables are much less than above and below ground floors and the number of shared walls – this is only an indication of the sensitivity of these variables rather than their importance over one another. Variables with distinct units cannot be compared with one another in any meaningful way. Below ground floors seem to be more significant in modeling electricity usage for buildings smaller than 6,683 m^2, while buildings larger than that rely on above ground floors. The number of surfaces shared with another building is found to be significant in determining the electricity consumed but it is not commonly relied upon.

Table XII Top five candidate models from each building subset ranked by AICc.

	Model Rank	Intercept	Below Ground Floors	Above Ground Floors	c2	c3	c5	c6	c7	c9	c11	c13	Shared Wall	adjR^2	df	logLik	AICc	delta	weight
Subset 1	1	192,118	NA				NA	-301		472				0.393	4	-260.998	532.664	0.000	0.119
	2	172,489	NA				-209	-381		605				0.475	5	-259.547	533.379	0.715	0.083
	3	217,164	-81,289				NA	-253		486				0.456	5	-259.913	534.112	1.448	0.058
	4	218,262	-109,137				NA	NA		343				0.342	4	-261.806	534.278	1.614	0.053
	5	182,611	NA				NA	NA		284				0.222	3	-263.484	534.469	1.805	0.048
Subset 2	Model Rank	Intercept	Below Ground Floors	Above Ground Floors	c2	c3	c5	c6	c7	c9	c11	c13	Shared Wall	adjR^2	df	logLik	AICc	delta	weight
	1	61,000	335,355		NA				-2,023	506	-804	NA	NA	0.691	6	-275.574	569.609	0.000	0.497
	2	47,560	360,826		NA				-2,146	531	-866	-591	NA	0.714	7	-274.773	572.879	3.270	0.097
	3	-1,385	350,205		NA				-1,995	525	-780	NA	143,835	0.700	7	-275.250	573.834	4.225	0.060
	4	160,613	NA		NA				-1,884	596	-697	NA	NA	0.524	5	-279.876	574.038	4.429	0.054
	5	-1,751	324,610		50				-1,977	533	-789	NA	NA	0.695	7	-275.429	574.191	4.582	0.050
Subset 3	Model Rank	Intercept	Below Ground Floors	Above Ground Floors	c2	c3	c5	c6	c7	c9	c11	c13	Shared Wall	adjR^2	df	logLik	AICc	delta	weight
	1	312,421		156,541	571						NA	501	NA	0.717	5	-289.020	592.326	0.000	0.206
	2	465,415		150,550	522						-78	447	NA	0.761	6	-287.339	593.140	0.814	0.137
	3	669,342		136,998	465						-96	NA	NA	0.677	5	-290.358	595.002	2.676	0.054
	4	506,712		142,533	518						NA	NA	NA	0.609	4	-292.264	595.195	2.868	0.049
	5	215,987		165,161	597						NA	528	538,483	0.734	6	-288.409	595.280	2.954	0.047
Subset 4	Model Rank	Intercept	Below Ground Floors	Above Ground Floors	c2	c3	c5	c6	c7	c9	c11	c13	Shared Wall	adjR^2	df	logLik	AICc	delta	weight
	1	-780,896		187,814	575	NA	-398			641			NA	0.947	6	-304.776	628.013	0.000	0.117
	2	-629,345		NA	619	NA	NA			692			NA	0.917	4	-309.307	629.281	1.268	0.062
	3	-1,232,911		149,613	607	NA	NA			642			NA	0.930	5	-307.599	629.483	1.471	0.056
	4	-1,141,608		201,956	585	NA	-365			641			1,189,720	0.955	7	-303.304	629.941	1.929	0.044
	5	-1,704,834		165,017	641	628	NA			641			NA	0.941	6	-305.835	630.131	2.119	0.040

Surprisingly, no models utilized the data on building footprint to estimate consumption. The only conclusion that can be drawn from this is that its effects on electricity use are relatively low compared to the other studied variables. Lastly, models for subset 4, containing buildings greater than 14,950 m^2, all have intercepts below zero. It is difficult to infer too much information about the predictor variables and electricity consumption because causality has not been established in this body of work, nor inferred from existing literature. These models merely represent correlations between select predictor variables and electricity consumption.

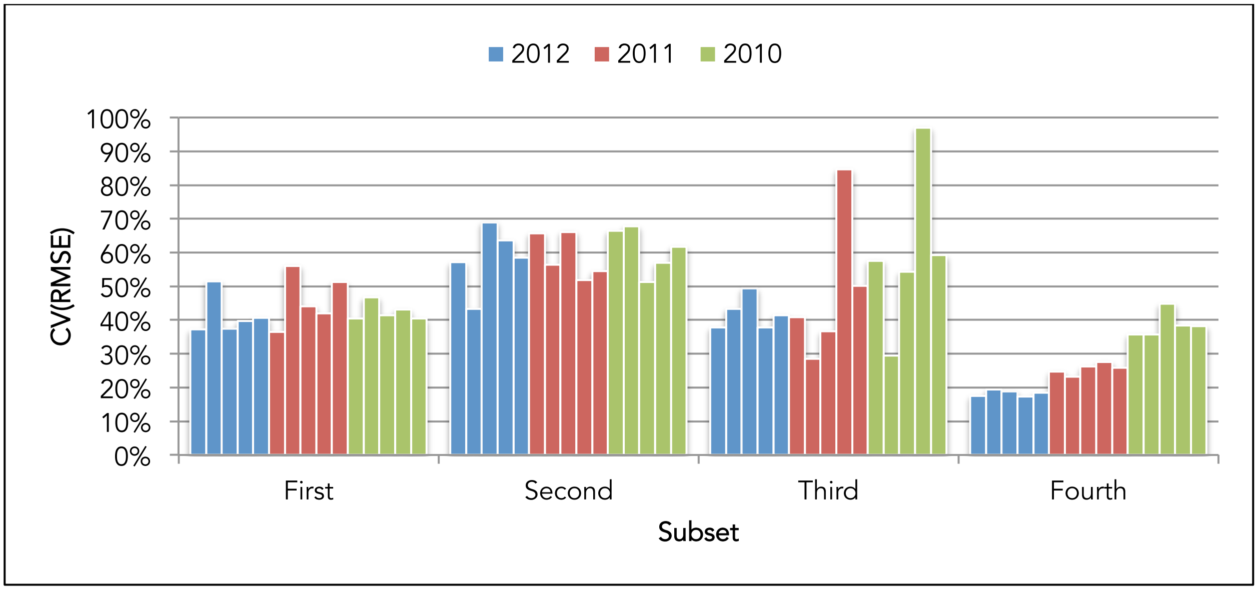


Figure 30 Performance of top models for all subsets using consumption data from 2010 to 2012. The larger the error, the less accurately the model has predicted consumption for buildings in that subset.

Figure 30 shows the coefficient of variation of the root-mean-square-error, CV(RMSE), for the top models for each subset across three years of data, calculated through cross-validation. Focusing on 2012 data, performance is similar between subsets 1 and 3; the second subset had the worst performing models and the fourth, the best. A general decrease in accuracy can be seen for earlier years which is an interesting trend because predictor variables were not updated for 2010 and 2011 models (due to a lack of data). This implies that using data that does not correspond to a particular year's consumption data potentially hinders the performance of top models. As Figure 30 represents average errors, it is difficult to conclude that using a particular model will result in a specific error for all buildings within the subset.

Additionally, models with a lower average reported error have the possibility of being less accurate than other models when used on certain buildings. Instead, average errors show which models perform better (more accurate) when considering all buildings within each subset. Because of this fact, it is sometimes beneficial for models within a similar error range to be combined and averaged in hopes that performance gains from multiple models can be incorporated into the final model – a practice which is further explained in subsection C. Application of Models on Ryerson's Individually-Metered Buildings.

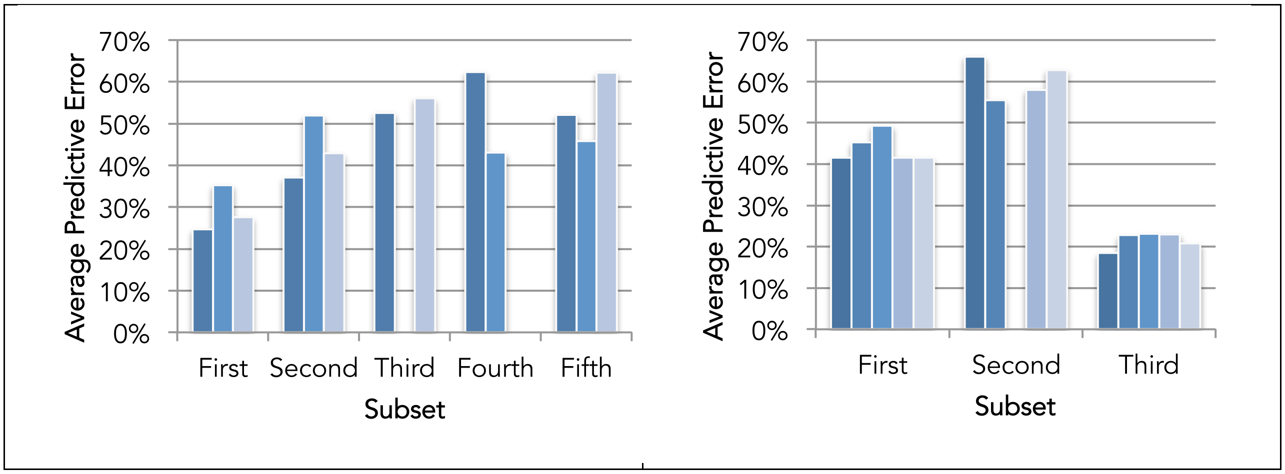


Figure 31 Model performance created using three and five subsets. Models resulting in an error rate above 100% are omitted.

As mentioned in Model Development, the use of four subsets was initially set arbitrarily. The process of creating and testing models was repeated for buildings divided into three and five subsets to see their effects on performance (Figure 31). With five subsets, the models created in the first set perform better than the original case with four subsets. Also, performance from the third subset onwards does not seem to improve as they do with the original case. Performance from three subsets mirrors the trend found with four subsets: an error rate around 40% for smaller buildings, which increases in the second subset, before finally settling around 20% for large buildings. Figure 31 indicates that performance from models seems to improve when a finer scale is used for smaller buildings. In addition, little impact on performance is made with the inclusion of more buildings at the larger end of the floor area spectrum.

Potentially, this means that the methods used here would benefit most from using smaller area intervals to define subset boundaries for small buildings, and larger area intervals for large buildings.

B. Application of Models on Ryerson's Individually-Metered Buildings

In order to reduce model selection bias in situations where there is a small sample size relative to the number of variables, multi-model inference or model averaging is used which promotes confidence in the results [83]. Top candidate models presented in Table XII are averaged based on their Akaike weights, which represents the relative likelihood of that model being true over all others. Coefficients for the top five models in each subset are averaged and shown in the first row after each header in Table XIII. Below the model coefficients are the values for the variables called upon. For instance, Oakham house (OAK) has one below ground floor, 1,686 m^2 of floor space categorized as c6, and 1,120 m^2 in c9. The sum of the intercept with the products of all terms results in the initial 2012 electricity consumption for that building.

Table XIII Final averaged models applied to each cluster-metered building at Ryerson. Displayed below the model's coefficients are the values for the building variables. Different clusters are identified by colour.

Subset 1	Intercept	Below Ground Floors	c6	c9				Estimated EC2012	Adjusted Estimated EC2012
	200,657	*-20,842*	*-181*	*418*					
HEI		1.00	419.12	847.44				457,714	**517,217**
OAK		1.00	1,686.48	1,120.37				341,809	**386,244**
Subset 2	Intercept	Below Ground Floors	c11	c13	c7	c9		Estimated EC2012	Adjusted Estimated EC2012
	47,560	*360,826*	*-866*	*-591*	*-2,146*	*531*			
SHE		-	-	-	-	2,025.42		1,123,653	**932,632**
RAC		2.00	-	-	11.60	867.03		1,204,971	**1,711,059**
SCC		1.00	-	4.99	-	104.09		460,738	**520,634**
CED		1.00	-	-	-	1,102.40		994,084	**795,267**
Subset 3	Intercept	Above Ground Floors	c11	c13	c2	Shared Wall		Estimated EC2012	Adjusted Estimated EC2012
	371,566	*154,028*	*-24*	*431*	*552*	*57,673*			
EPH		4.00	36.70	-	3,359.00	3.00		3,015,035	**2,502,479**
JOR		15.00	-	-	38.35	0.33		2,722,204	**3,865,529**
RCC		3.00	10.40	-	2,436.45	-		2,179,090	**2,222,672**
IMA		3.00	1,190.05	-	884.06	-		1,292,839	**1,163,555**
VIC		9.00	842.10	-	695.19	1.00		2,178,872	**1,960,985**
Subset 4	Intercept	Above Ground Floors	c2	c3	c5	c9	Shared Wall	Estimated EC2012	Adjusted Estimated EC2012
	-997,899	*143,818*	*599*	*80*	*-196*	*651*	*165,686*		
LIB		11.00	327.44	940.21	321.28	5,519.85	0.50	4,468,527	**6,345,308**
POD		4.00	542.70	25.59	764.21	5,479.28	2.00	3,653,314	**5,187,706**
PIT		14.00	-	-	-	3,291.66	-	3,158,541	**3,221,712**

The adjusted estimate for each building is determined by comparing the estimated electricity consumption for each cluster with the meter readings in 2012. Depending on the difference between the two values, a constant multiplier is applied to all buildings in that cluster in order to increase the chances of accurate results. Important to note is that buildings from each cluster may be modeled using different equations. Figure 32 shows the raw estimates for cluster-metered Ryerson buildings and their 2012 electricity meter readings. Accuracy of the estimates ranged from an underestimate of 29% (JOR.LIB.POD.RAC) to an overestimate of 20% (EPH.SHE) with an average error rate across all five clusters at 14.8%. Electricity estimates for the JOR.LIB.POD.RAC cluster were adjusted the most (increase of 42%); on average however, the clusters were corrected by 16.8%. The relatively large difference observed between estimated and actual electricity consumption for the JOR.LIB.POD.RAC cluster may be attributed to the location of a large centralized cooling system that provides cooling for 10 other campus buildings, which was previously discussed in the Model Development section. EPH.SHE's consumption may have been overestimated due to the fact that they are considered by the model to be two separate buildings (with significant shared surfaces) when in fact SHE is more of an extension of EPH. Lastly, the differences seen in the SCC.OAK.HEI cluster is questionable due to a lack of updated data for the meter; the value shown in Figure 32 represents electricity consumption that has been lowered by 11% (average decrease of all buildings from 2008 and 2012 – 12% for all cluster-metered buildings) from its last reported consumption in 2008.

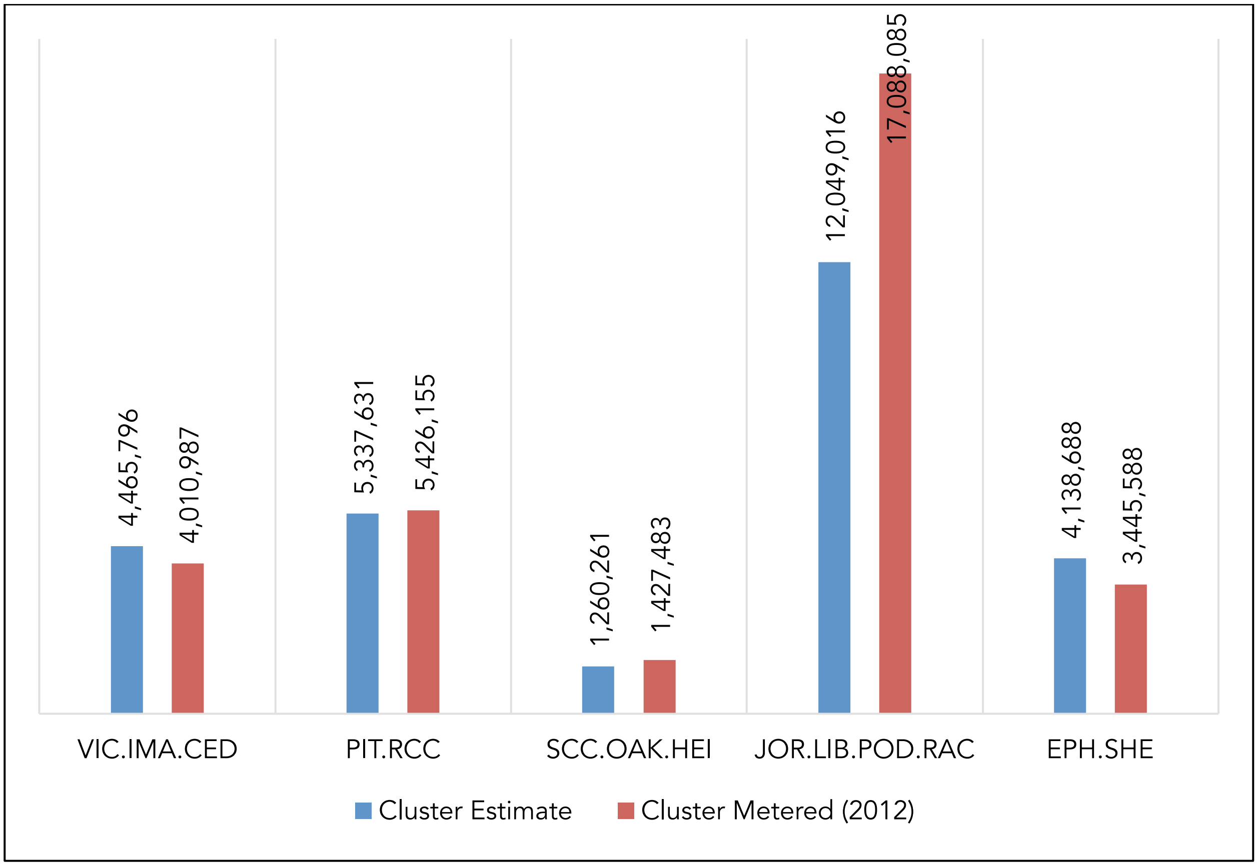


Figure 32 Raw electric consumption estimates (kWh) for Ryerson's clusters and their actual metered values in 2012

The results shown in this section positively reflect the strength of multiple linear regression in estimating electricity consumption in large and complex buildings. The next section will provide a more critical analysis on the strengths and weaknesses of the prescribed method and the future applicability of the models.

V. DISCUSSION

A. Model Performance

Figure 32 previously summarizes the accuracy of the estimated electricity consumption for each of Ryerson's building clusters. While this provides some insight into the accuracy of the developed models, their level of performance can be misleading. The reported errors can be cancelled from multiple buildings resulting in a false sense of accuracy for individual buildings within the cluster. In order to combat this, a comparison of results between this project is made with Rahman's [33] work, which was completed in 2010. Rahman studied the feasibility of using a Heat Recovery Ventilation system and expanding Toronto's Deep Lake Water Cooling system to cover Ryerson's buildings. In order to quantify the benefits of reduced energy consumption and GHG emissions in these scenarios, an energy audit was completed for 16 buildings (86% of Ryerson's total floor area), followed by energy simulations based on 2006 data using Carrier HAP (Hourly Analysis Program). The results from simulations for buildings that were covered by this thesis and Rahman's are shown in Figure 33.

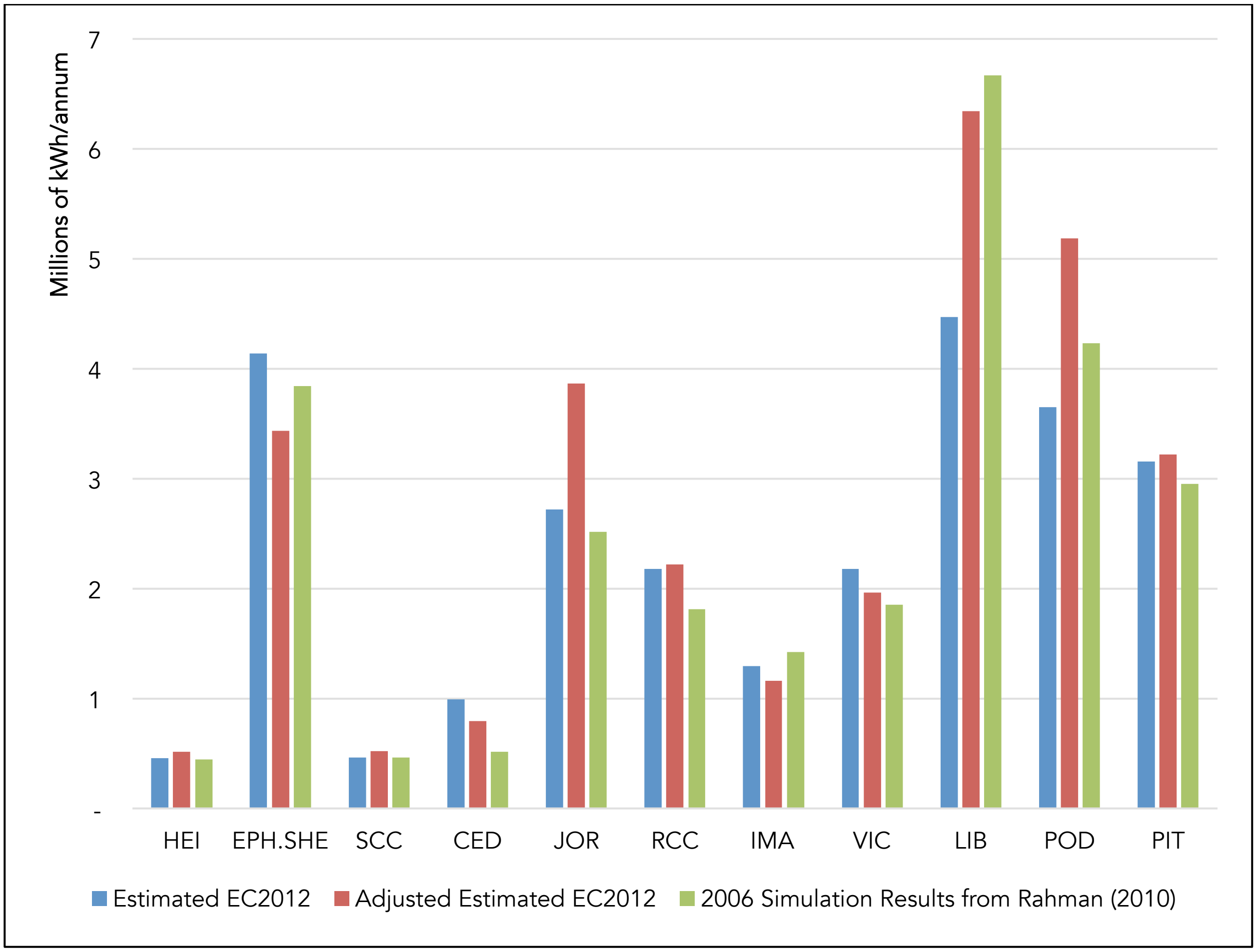
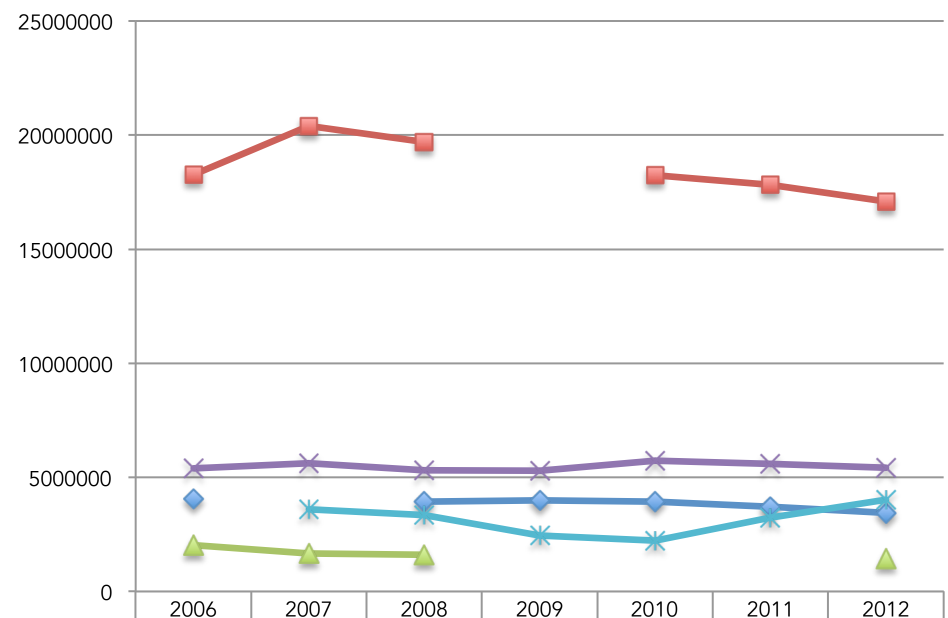


Figure 33 Comparison of estimated electricity consumption from this study with the simulated results from [33].

In most instances, the adjustments made to the original estimates for individual buildings further distances the results from Rahman's. Only four of eleven buildings see a decrease in the difference between the simulated and model-estimated values. The reason why Rahman's results can be relied upon as a benchmark is due to its reported accuracy. Comparing the performance on two individual buildings (whose consumption is known), his results for the Engineering Building (ENG) and the School of Interior Design (SID) yielded an error of 1.2% and 2.3%, respectively. Using the averaged models created in this thesis, the error for those same buildings were 36% (ENG) and 34% (SID). While the models may be more accurate, their year of estimation was six years previous to those of this study making direct comparisons difficult. Figure 34 displays the short-term trend seen for each building cluster. As mentioned previously, the SCC.OAK.HEI cluster's last fully metered year of consumption was in 2008 – it is therefore excluded in making generalizations about consumption during 2006 and 2012. In total, electricity consumption for the remaining four clusters saw a decrease of 7% between 2008 and 2012 (complete data was not available for 2006 due to missing or outlier data). On a cluster-basis however, EPH.SHE and JOR.LIB.POD.RAC both experienced a decline during this period of 15% and 7% while VIC.IMA.CED saw an increase of 11%. A marginal 1% increase in consumption was found at the PIT.RCC cluster. From this we can tentatively conclude that that Rahman's results in Figure 34 is expected to vary from those of this thesis by a factor of 11% to -15% from usage patterns alone. Again, it is difficult to ascertain the fluctuations found within individual clusters meaning that the differences in estimated consumption from this study and Rahman [33] may be due to an actual increase or decrease in electricity consumption between 2006 and 2012, or may originate from low precision between both methods.



	2006	2007	2008	2009	2010	2011	2012
EPH.SHE	4057549		3928862	3994928	3,943,569	3,709,375	3,445,588
JOR.LIB.POD.RAC	18276049	20399666	19697876		18,248,639	17,831,577	17,088,085
SCC.OAK.HEI	2,042,267	1,675,334	1,603,913				1,427,483
PIT.RCC	5397903	5607650	5321522	5293557	5,722,330	5,581,427	5,426,155
VIC.IMA.CED		3599652	3359113	2456452	2,228,852	3,234,204	4,010,987

Figure 34 Electricity consumption trends for meters representing multiple buildings over the analysis period

Comparisons can be made with measured consumption data on a cluster-basis (Figure 35). As expected, energy simulations run by Rahman [33] yielded more accurate results than the multiple regressions created in this thesis. Of the five clusters, one of them (PIT.RCC) is more accurately estimated using regression analysis. Comparing the average error between the two methods, the Carrier HAP models were roughly 3 times as accurate than the models used here (4.4% vs 14.8%). As mentioned in the Results section, the error associated with EPH.SHE and JOR.LIB.POD.RAC clusters may be inflated due to unique circumstances not found elsewhere in the building sample. This comparison highlights the performance gains from using energy simulations over statistical models. However, as earlier mentioned, the methods introduced in this thesis are not to compete with existing methods, such as energy simulation. Instead, it is to offer another solution for scenarios where the need for information gain is to be maximized with as few resources available. Consideration has yet to be given to the methods behind each approach. Arguably, the methods in [33] are more work intensive and undoubtedly more technical in nature when compared to this thesis. The detailed inputs required for energy simulations are extensive and may or may not be transferrable to other buildings; each building's performance is predicted with a unique model whereas using statistical models, only four equations are required to model the majority of academic buildings. Ultimately, it is left for the user to decide whether the difference in accuracy between the two methods is significant. For most planning applications, the accuracy of the regression approach may be suitable for prioritizing certain buildings over others. In cases where a detail breakdown of energy end uses is needed, such as when selecting which energy efficiency measure will be most effective, simulations may be the best choice. These conclusions can only be made confidently with more comparisons of different methods on the same building sample – an area with much needed growth. With numerous modeling programs to choose from, the error results may vary considerable with other programs. Carrier HAP is not a common choice among academics but its modeling abilities are comparable to current standards (DOE 2.1, eQUEST) [120].

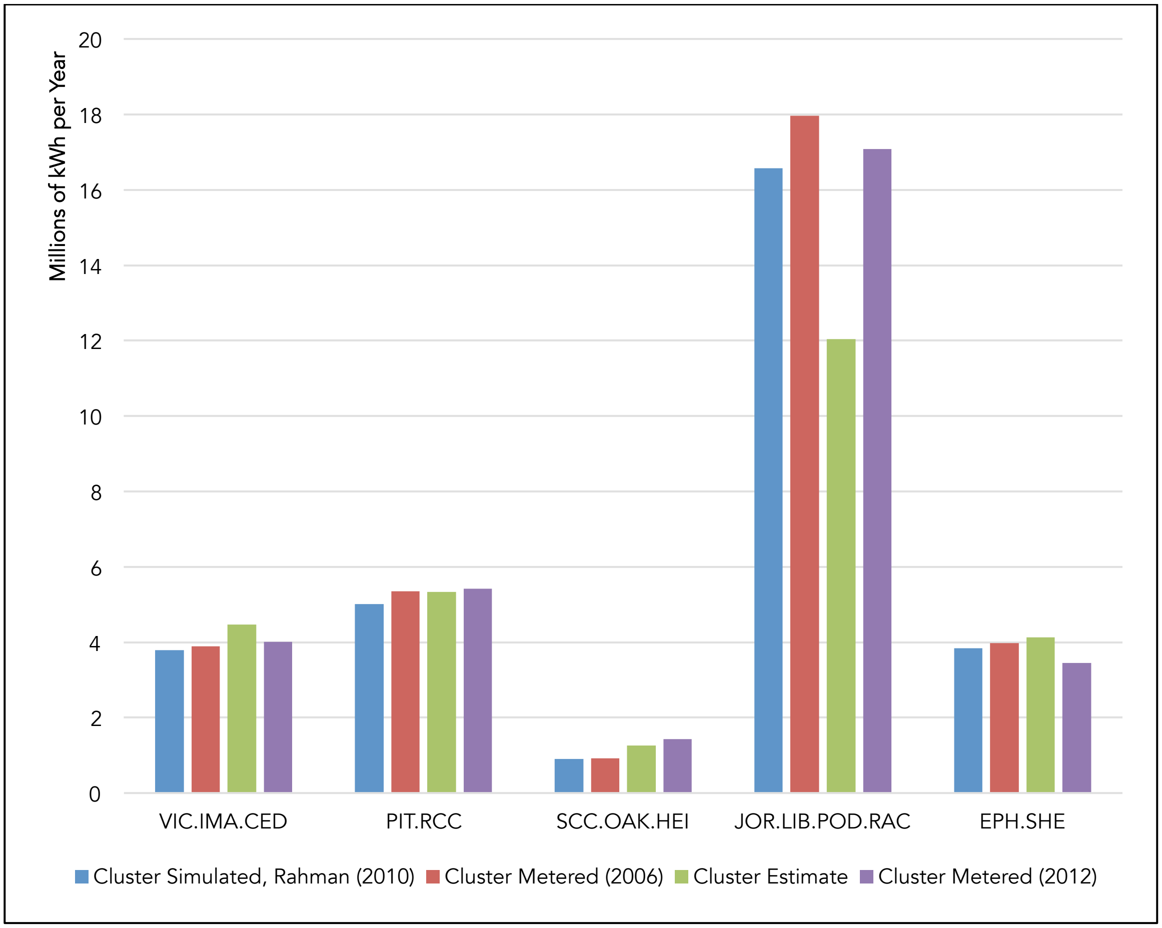


Figure 35 Estimates for Ryerson clusters from this study and [33] compared to actual metered readings

B. Significance of Model Variables

Of the 18 variables selected for analysis, only 11 of those were deemed significant when estimating electricity consumption in academic buildings. A summary of the finalized linear models is shown in Table XIV. This section will focus on the 11 variables and determine whether or not the variations in their weightings can be explained with building science principles.

Table XIV Model variables for each subset

Subset (Range of included Floor Areas in m2)	Intercept	Below Ground Floors	Above Ground Floors	c2 Research Labs	c3 Athletic Service Space, Animal Holding Areas, Plant Maintenance	c5 Assembly Facilities, Classrooms	c6 Day Care Facilities, Lounges, Library Stacks, Exhibition Spaces, Support Spaces for Classrooms, Labs, and Offices, etc.	c7 Residence Service Space, Dining Areas, Recreational Areas	c9 Janitorial Closets, Mechanical and Electrical Spaces, Public Washrooms, Circulation Spaces, Loading Docks, Service Tunnels	c11 Inactive Spaces, Parking Areas	c13 Merchandising Areas, Library Support Spaces	Shared Wall
1 (1000-2900)	200,657	-20,842					-181		418			
2 (2901-7200)	47,560	360,826						-2,146	531	-866	-591	
3 (7201-14900)	371,566		154,028	552						-24	431	57,673
4 (14901-82700)	-997,899		143,818	599	80	-196			651			165,686

In order to better understand the model coefficients, a closer look at the properties of the buildings within each subset is needed. Since there were only 20 buildings in each subset, the observations should not be extended to the entire building population or even to all academic buildings. They should instead be treated as nuances, unique to the building sample used (U of T and Ryerson University buildings). Figures 36 – 38 compare the values for each variable across all subsets. Focusing on floor levels, it is evident that larger buildings tend to have an extended upper range of floor levels compared to smaller buildings; subsets 1 and 2 peak at three or four above ground floors. These trends are much more subdued for below ground floors with only the largest buildings having three or more levels (most likely due to large underground parking structures). The equations for the first two subsets use below ground floors while the last two subsets use above ground. This indicates that the electricity consumption for buildings under 6,683 m^2 is influenced more by the existence of a basement than multiple above ground floors; small building mostly either have zero or one below ground level while they range from two to seven for above ground. For the most part, the selection of this variable for smaller building makes sense however the weightings assigned to each model are questionable. A small or negative coefficient in model 1 is expected because small buildings tend to use their basement levels as storage/service spaces. In model 2 however, the coefficient is much larger – this may be to offset the large negative coefficients for other variables in the model, namely c7, and the relatively low intercept. For subsets 3 and 4, the large variance in above ground floors seems to impact electricity consumption more than basement levels. While it may appear that the weighting of the variable is consistent across both models, the difference in the intercept signifies that model 3 relies on above ground floors significantly more than model 4.

The relatively large coefficients assigned to above and below ground variables may cause problems when estimating consumption for smaller buildings. There are instances with model/subset two where the total building consumption is less than the coefficient assigned to below ground floors. In total, 25% of the subset suffers from this, which may explain why its performance is the weakest among the four models (Figure 30). Despite this fact, the remaining space usage coefficients are almost all negative and quite large which may offset the model's sensitivity to below ground floors.

All in all, the variations observed with the number of floor levels seem to be governed more by the other variables of the models than any building science principle. However, the potential lack of impact the number of floors of a building has on its electricity use intensity was brought up in the Model Development section when discussing the collection of data. Nevertheless, building floor levels are statistically significant variables in the presented models.

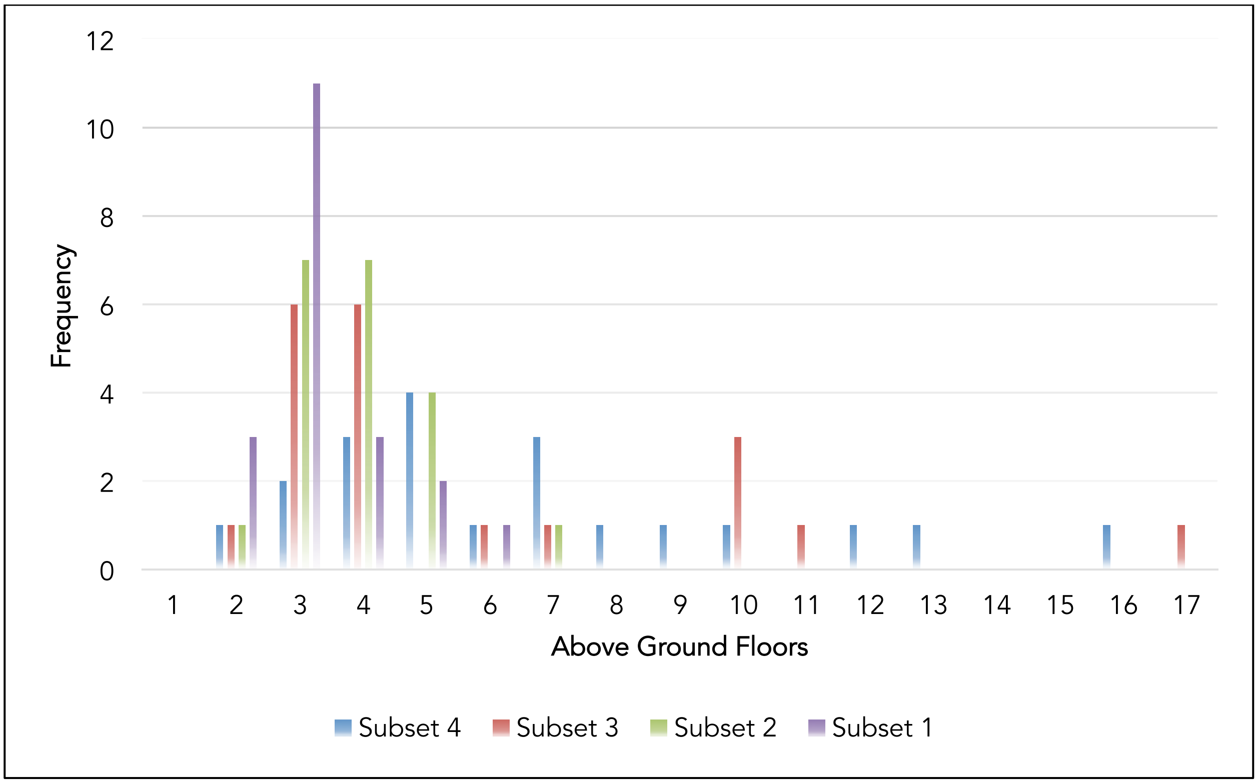


Figure 36 Number of above ground floors for buildings in each subset

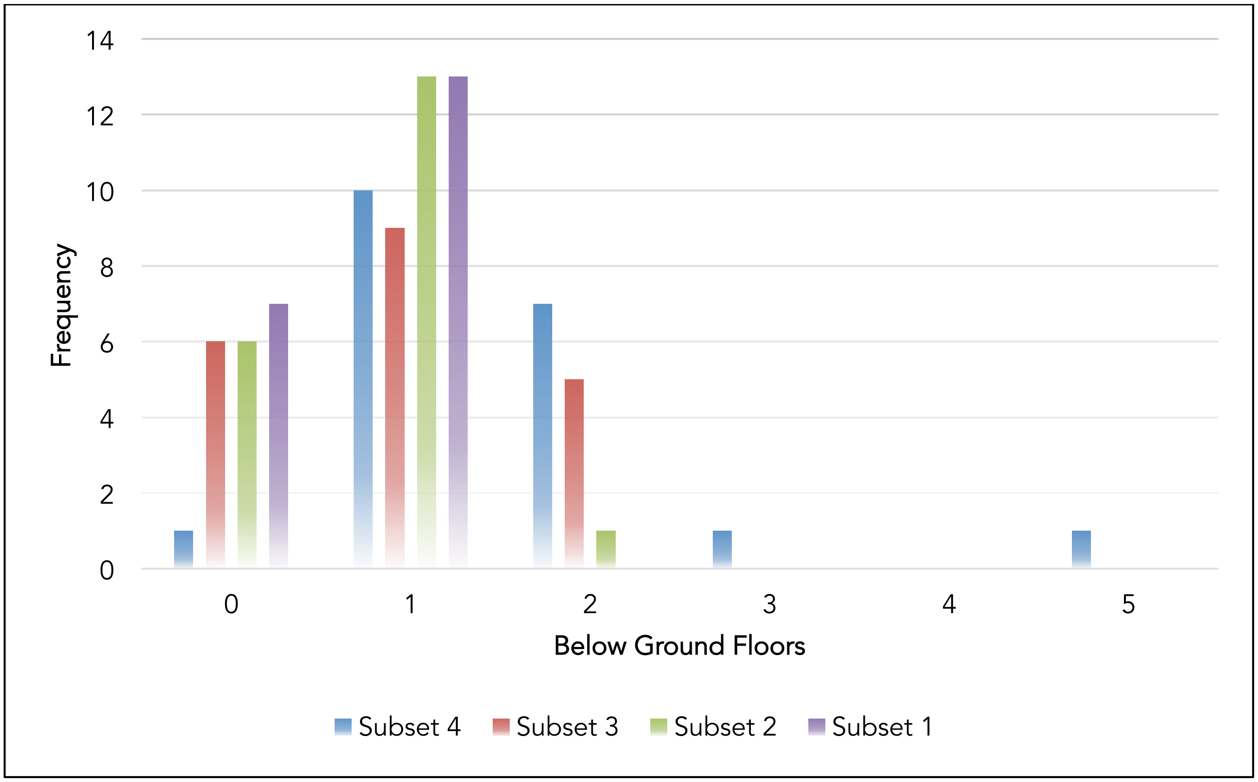


Figure 37 Number of below ground floors for buildings in each subset

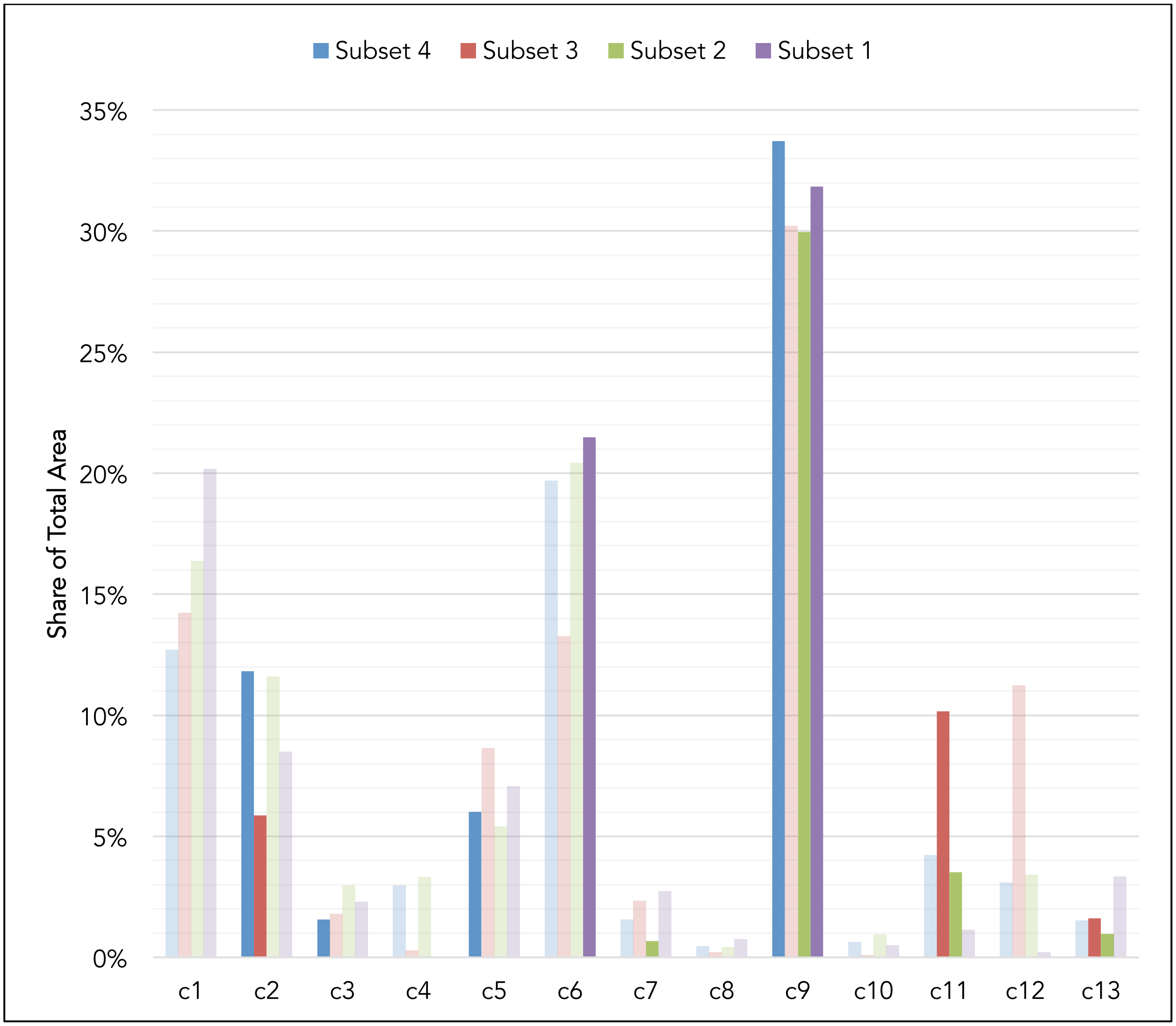


Figure 38 Comparison of space category ratios for buildings in each subset. Categories not included in the final models are semi-transparent.

Theoretically, as the number of shared surfaces increases for a building, its energy consumption should decrease because heat gain/loss through that surface is reduced. Because it is electricity consumption that is being predicted, the effect of shared surfaces may be minimized since no buildings in the sample rely on electric heating. Looking at Table XIV, models 3 and 4 rely on this variable to predict electricity consumption in academic buildings. However, the data for this variable is highly sporadic with only 5% of buildings showing any shared surfaces in subsets 1 and 3, 15% in subset 2, and 20% in subset 4. It can be argued that larger buildings tend to have one or more shared surfaces as oppose to their smaller counterpoints; despite the building property leading to reduced heat loss, it may be an indication of an oversized building which may consume more energy. The uncommon nature of this building property, combined with the positive coefficients assigned to this variable suggests that it functions as a minor correctional term rather than one that is heavily relied on to estimate energy consumption.

This critical analysis of model terms should not be used to detract from the performance of the presented models. While many of the predictor variables and their coefficients are not supported by common building science theory, their ability to predict electricity consumption in academic buildings is proven. In other words, the selected variables are statistically significant with energy consumption but their assigned weights cannot be used as a basis for building science theory without further analysis.

C. Commonalities Between Omitted Academic Buildings

Eliminating buildings with an area smaller than 1000 m^2 was unavoidable due to extreme variations in electricity consumption in published studies as well as those found in the building sample. Stemming from the same logic, a blanket approach to removing houses from the sample is also used. Electricity consumption in houses is assumed to be noticeably different and is not included in the definition of a typical multi-use academic facility. A less blunt approach could be taken, considering other building properties that may lead to high variation in electricity consumption, but that level of research is outside the scope of this project.

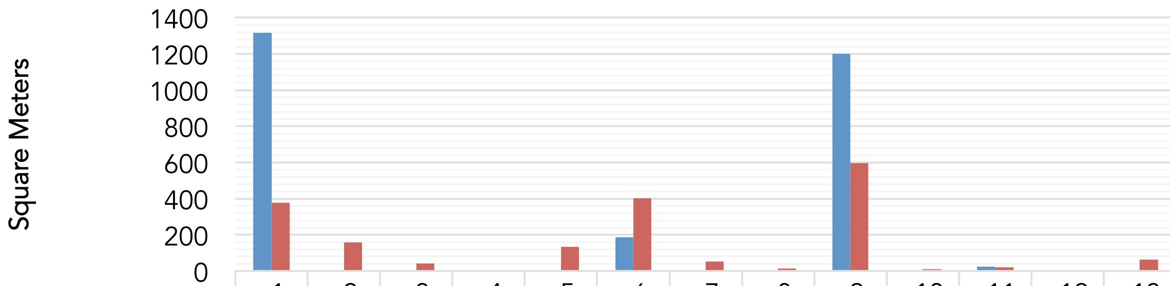
Outliers were identified in two stages. Buildings were first identified as outliers when their 2010-2012 electricity consumption exceeded their historic usage. This could have been caused by a number of reasons as discussed in Model Development. The buildings were eliminated from the sample because of their flux in consumption which a static model would struggle with. The second stage of outlier identification occurred after establishing the subsets for model building. Here, buildings were eliminated if their consumption was significantly higher (>2x) than the average within the subset. It is these buildings that are of interest in this section to determine if there are identifiable trends that will indicate whether a building will be able to be modeled using the regressions created in this thesis.

Table XV Buildings that consumed significantly more electricity than other buildings in the subset

Building Name	Institution	Year of Construction	Subset	
Fields Institute	Toronto	1995	1	
Aerospace Building	Toronto	1959	2	
Leslie L. Dan Pharmacy Building	Toronto	2004	3	

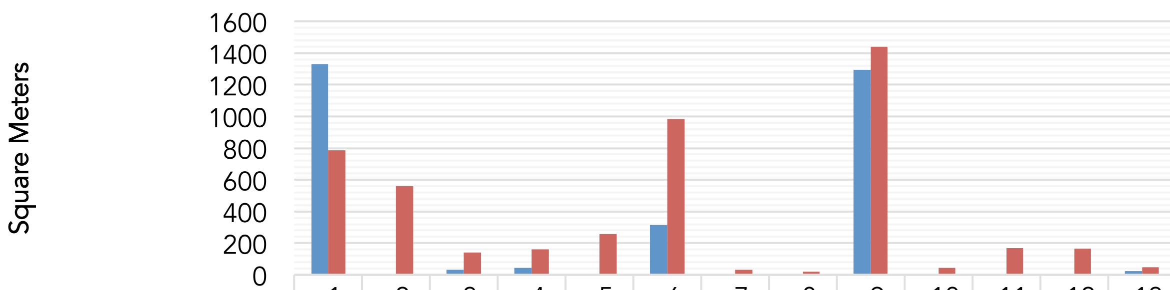
Medical Science Building	Toronto	1969	4	

Table XV lists details about the four buildings that were eliminated at the second stage of outlier identification for the 2012 model year. All buildings removed from the sample are buildings from the University of Toronto. They share very few commonalities in construction type and year of construction but they manage to consume significantly more electricity than other buildings. A closer look at the study variables reveals major differences in space usage (Figures 39 – 42). Across all buildings, significant offsets from the subset average can be seen in c2, c3, c8, and c12. When factoring in the direction of the offset, only c12 (residential spaces) is consistently below the subset average while the other differences are varied. It is important to note that buildings within the first two subsets measured zero for many of the space categories (8 and 7 for subsets 1 and 2). The significant differences (<33% or >300%) in the two smaller buildings were almost always smaller than the average; the differences in the larger two buildings were more balanced. Aside from space usage, not many differences were found between outliers and their respectful subsets. Other than the first subset, the outliers were not the largest buildings, nor did they contain major differences in the number of shared walls and below ground levels. Greater variability was observed with above ground floors but nothing set these buildings apart from the majority. Looking at a building characteristic that wasn't directly modeled in this thesis, there is a definite possibility that plug loads for specialty equipment (specifically linked to laboratory/research spaces) are the cause of the increased energy consumption in these outlier buildings. If equipment that draws large amounts of energy is installed in lab spaces, an increase in electricity use intensity for these spaces will occur. A large increase in intensity will offset the relatively low amounts of lab-dedicated spaces in some of these outlier buildings leading to a greater amount of energy consumed. By definition, lab spaces, regardless of equipment type, are categorized as one in this thesis; in future iterations it will be beneficial to adopt an approach similar to Bonnet et al. [30] (i.e. incorporating plug loads to help distinguish lab spaces).



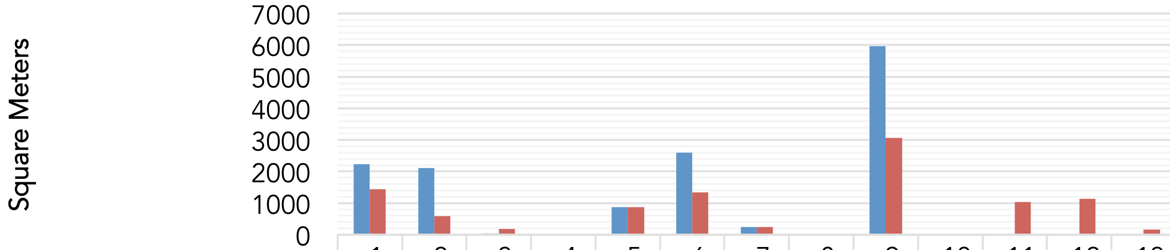
	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13
Fields Institute	1315	0	3	0	0	186	0	0	1198	0	23	0	0
Average	378	159	43	0	133	403	51	14	597	9	21	4	63

Figure 39 Area for defined space categories for the Fields Institute (U of T)



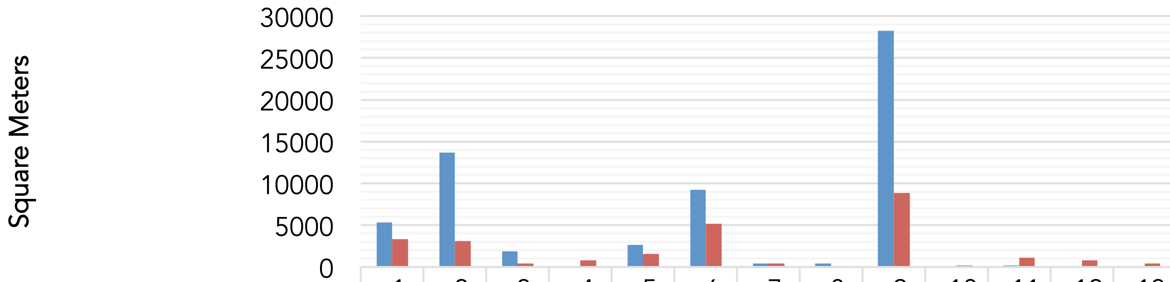
	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13
Aerospace Building	1329	0	33	43	0	315	0	0	1296	0	0	0	24
Average	788	558	143	160	260	982	32	21	1440	45	169	164	47

Figure 40 Area for defined space categories for the Aerospace Buildings (U of T, Off-campus)



	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13
Leslie L. Dan Pharmacy Building	2225	2118	33	0	877	2604	251	3	5972	0	23	0	0
Average	1443	595	181	29	877	1344	237	21	3062	9	1030	1139	165

Figure 41 Area for defined space categories for Leslie L. Dan Pharmacy Building (U of T)



	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13
Medical Science Building	5320	13699	1859	0	2608	9248	397	427	28241	65	212	0	0
Average	3336	3102	414	782	1578	5174	409	123	8856	168	1112	813	402

Figure 42 Area for defined space categories for the Medical Science Building (U of T)

Drawing conclusions from these differences in building characteristics is very difficult due to the pre-existing variability found in academic buildings – a reason why they are seldom studied. For instance Figure 43 shows the data for the c3 category (athletic service space/plant maintenance) for buildings in the fourth subset. C3 was graphed because it exhibited the most variability when comparing the outlier buildings to the averages. However, when looking at the sample, a great deal of variation still exists. On the other end of the spectrum, c9 (Circulation spaces) offsets were not deemed significant when comparing outliers to their subset averages. Figure 44 shows the data for the c9 category for the fourth offset. Indeed the variance is minimized in this case but the existence of a building with a greater offset from the average means that this type of analysis will not be effective in identifying buildings with high EUIs and thus difficult to model with a linear regression.

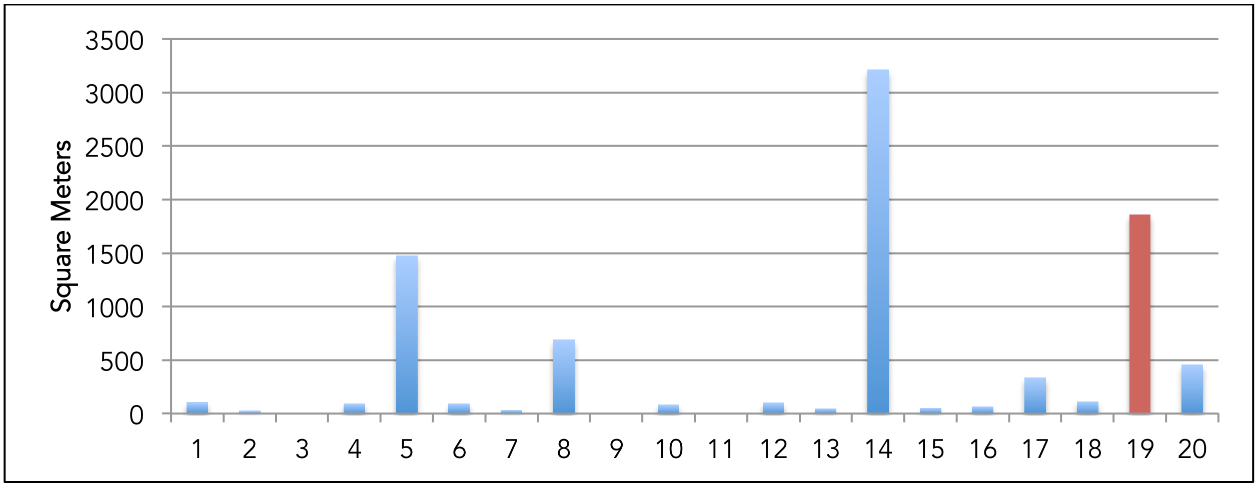


Figure 43 The natural variations found with athletic service/plant maintenance spaces for buildings in the fourth subset. The value for the outlier building is shown in red.

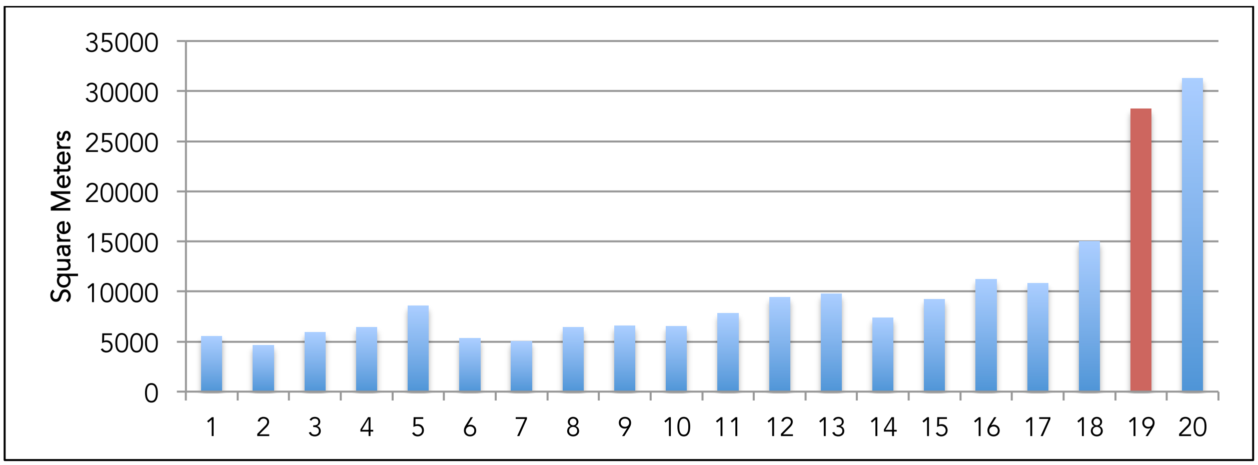


Figure 44 The natural variations found with circulation spaces for buildings in the fourth subset. The value for the outlier building is shown in red.

Identifying certain building characteristics that cause high electricity usage intensities for the four omitted buildings can possibly be done on a case-by-case basis. Justifications can be made for these anomalies by overanalyzing specific details. However, formulating general rules to apply towards other sample sizes cannot be made because of the few instances of outlier buildings that exist. Even with greater application of the models to other institutions, trends can only be identified if the same variables are collected and used to formulate the models. Instead, focus should be placed on attempting to model the highly diverse electricity consumption found in buildings with areas less than 1000 m^2. These buildings represent nearly

20% of U of T and Ryerson's building sample, much more than the four omitted buildings due to very building-specific properties.

D. Challenges to Implementing Regression Models in Other Settings

The performance of the multiple linear regression models created and used on Ryerson's cluster-metered buildings shows the strength of statistics when coupled with basic building science and occupancy principles. However, there are some obstacles that need to be overcome before the successes seen in this thesis can be repeated for academic institutions in Ontario and around the world.

Starting with the variables of the model, quality COU space data is required to maximize performance. As seen when using space usage data that was a year or two behind that of the measured consumption year, the models' accuracy suffered. Not only is it important to have updated data, the accuracy of the data is critical. Differences in data collection between Ryerson and U of T were detailed in the Model Development section along with issues of interpreting space definitions. These issues will continue to exist and influence model behavior as long as no formal collection method is cemented among participating universities. The Council of Ontario Universities space standards are common within the province of Ontario but receive much less attention nationally and internationally. Other provincial organizations linking universities exist within Canada including the Research Universities' Council of British Columbia, the Association of Atlantic Universities and the Bureau de Coopération Interuniversitaire. Among these groups, there are instances of adapting the COU space classifications [121] to accomplish similar goals however evidence of progress towards a uniform space standard across is limited. What this means is that equations used in this thesis will only be compatible with universities that tabulate their space usages based on the COU building blocks. Academic institutions using other definitions will benefit from the methods detailed in this thesis but will need to generate customized regressions to make use of the variables available. Climate and architecture play a critical role in determining the applicability of the algorithms developed in this thesis to other international markets. As displayed in Figure 9 in the Literature Review, the electricity use intensity varies greatly from place to place due to a host of reasons including: perception of comfort, societal norms, climate, and architecture. The equations used in this thesis won't apply to these markets but, as mentioned above, the methodology prescribed will be applicable to develop customized models. Variables analyzed for model building may be similar to the ones studies in this thesis but they should be arrived at only after considering the built forms found within the academic institution(s). In other words, narrowing down model parameters should be done with a keen sense of context.

The next challenge to building reliable linear models to predict electricity consumption involves building diversity. As buildings become more unique from one another – both in their design and function – their energy profile is expected to change. While this issue has not been prominent in this thesis project, it has the potential to develop in future iterations as universities invest more into the aesthetics of their facilities to attract academics and students [122]. Also, historically preserved buildings with new additions will become more commonplace as space constraints in urban campuses become a greater issue. The building sample used in this thesis contained such unique buildings mentioned above with no issue so it is difficult to gauge how effective the methods used here will be in the future. With that being said, no regression model will be future-proof for all academic buildings – significant predictor variables will change with time. This is because electricity consumption within buildings is highly dependent on variables that are in a state of flux. How these variables change in comparison to the others and how electricity consumption responds for that given year can result in completely different models being created. Gallachóir [97] analyzed the relationship between electricity consumption at University College Cork and its student population. It was found that historic trends dating back to the early 1980s were similar across both metrics. From 1994 and onwards, electricity consumption outpaced the growth seen in the student population. This was later attributed to the influx of personal computers on campus, increasing the plug loads. In addition, it was noticed that the energy trends for different space types were heavily influenced by the space's reliance on technology and equipment. For instance, research spaces and laboratories experienced a higher rate of change in energy intensity compared to dormitories, offices, and libraries.

From a practical standpoint, as buildings adopt new technologies or components that affect how energy is consumed within them, it is important to consider these changes by re-evaluating model variables to improve accuracy. For example, if a substantial share of buildings within a sample utilizes high performance windows, it may be worthwhile to collect the necessary data to include that variable in the analysis process. The frequency for model re-evaluation is dependent on the pace at which universities introduce new buildings into their existing building stock either through new construction or renovation, relative to the size of the sample used for model creation. In addition, it may be prompted with poor performance results from existing models. Ultimately, the institutions that this thesis is targeting is a niche population that is expected to decrease in number with time. It is hoped that the costs associated with metering utility usage on an individual building basis the methods prescribed in this thesis are used as a last resort for new academic buildings as metering and sub-metering becomes the new standard for construction.

Gathering a large enough sample of buildings to construct and validate the models is a hurdle not unique to multiple linear regressions. This issue is compounded by the lack of published material on minimal or optimal sample sizes for various modeling techniques – it is generally recommended to include as large a sample as possible to reflect reality. An aspect of this challenge that relates to the methods used in this thesis is the prevalence of small buildings (under 1000 m^2) in samples. If the target and/or sample buildings are within the range of areas that experience high variability in electricity consumption, it may be increasingly difficult to find accuracy in multiple linear regression models. The definition of what constitutes a "small building" may change from one study to the next but it is expected that this handicap will persist while dealing with the studied variables in this thesis.

E. *Suggested Areas for Future Research*

1) Model Variables: The variables considered for analysis in this thesis were selected for their balance in their relationship with electricity consumption and ease of use to measure. While they have been proven to predict consumption in academic buildings, there is a great deal of refinement that can occur that may improve performance in future iterations of the models. One of the greatest areas of contention in the methods of this thesis is the reclassification of COU space categories. The results of this thesis are due to the way the space usages are categorized, however there are a multitude of alternatives to explore. Minor changes in groupings or elimination of certain categories may potentially have profound effects on model performance. Work should be dedicated specifically on how best to harness the existing COU area database to represent energy consumption in university buildings. As discussed in the Model Development section, the variables accounting for above and below ground floors and shared building surfaces are flawed. In this thesis, the flawed approach was embraced for simplicity purposes. Future work should focus on how best to represent these building variables in models. Ultimately, the choice for how to represent floor height and exposed surface area should be left for the users of the model as their priorities shift.

2) Defining Subsets: Subsets were defined by floor area because it was a simple guideline to follow with no room for interpretation. However, creating subsets on the basis of many other factors should be explored in future work. For instance, subsets can be created with any of the existing variables analyzed in this thesis, such as COU categories, or they can be created from a new variable, such as construction type. Expanding work in this area will definitely increase the time and effort needed to create subsets for future models, especially if they are created on new variables requiring additional measurement. Another area worth revisiting was previously eluded to in the Model Development section. Reworking subset boundaries so that more models are created for smaller buildings should be pursued. Because of the difficulty in modeling smaller buildings, models should cover narrower area-intervals in hopes of capturing the unique characteristics of those buildings. This adjustment can be flipped for larger buildings, with less models dedicated to them. It is hoped that by addressing these characteristics in the sample, overall prediction error for all buildings can be reduced in future iterations.

3) Multimodel Inference: Selecting the "best" model for application is another area that is highly contentious, however, unlike reclassifying COU space categories, the reasons behind this are due to external conditions. There are mixed signals from the statistical community as to how best to carry out model inference. The procedure used in this thesis was pieced together from multiple sources with several of them recommending model averaging. With that being said, there are multiple ways of selecting and/or averaging candidate models, some of which may result in a superior model than the one presented in this thesis. Again, future work should examine if there is an ideal method of model inference from dredge outputs.

4) Stepwise Versus Hierarchical Regression: An alternative model to pursue in future work is hierarchical regression [123] which differs from the stepwise regressions used in this thesis. While both approaches strive to identify the "best" set of predictor variables, they differ in key ways which will affect the resultant model. The main difference between both model types is that stepwise relies more on an automated approach, where the computer program tests the model's performance in an iterative fashion and through a series of algorithms, determines whether to keep or discard the predictor variable. Hierarchical regressions on the other hand are more reliant on the researcher's theory on relationships than the computer program's algorithm. This type of regression is best suited when groups of variables exist which exhibit some level of collinearity. These groups are introduced into the model and the variance calculated is used to determine their significance. This approach is best suited for researchers who have a sense of the key determinants of electricity consumption in their buildings and are pursuing a more universal model that applies to buildings outside of their sample.

VI. CONCLUSION

There is a niche group of properties where buildings are metered in a cluster – due to cost saving initiatives or pre-existing metering configurations – and owners/managers are unable to invest the required amount to upgrade the infrastructure (i.e. individual utility meters/submeters for each building). As a result, plans to measure and implement sustainability measures for buildings can be impeded by a lack of information. This thesis looked at a potential solution for Ryerson University to gain reliable electricity consumption data for all its cluster-metered buildings. This method is not targeted towards replacing the installation of individual meters (which provides real, up-to-date, data), rather it is a low-cost, temporary option for the realization of consumption patterns for building whose usage is concealed when measured as a group – a necessary first step before planning, implementing, and measuring sustainability goals can take place.

The method behind creating the four regression models was identical, save for the different training and testing samples used. As predicted, the MSE decreased for models created and tested on the last two subsets (containing larger campus buildings) versus the first two. In addition, it was found that model performance would likely increase with greater segregation of buildings in the first subsets, and that performance would not be affected with the amalgamation of larger buildings from the last two subsets.

Multimodel inference or model averaging of the top candidate models ranked by AICc was used to obtain the final models for each subset to be applied on Ryerson's buildings. The estimates from the four models were within a reasonable range (i.e. CV(RMSE) = 14.8%) from the measured consumption for cluster-metered buildings. These estimates, for each building, were proportionately adjusted based on the differenced observed between the estimated and measured value for each cluster. It is uncertain whether the adjusted values prove to be more accurate since the comparison with Rahman's [33] simulated 2006 results for the same buildings are mixed. Despite being roughly three times less accurate than Rahman's energy models, the time and resources invested building and validating models seem to favour multiple linear regression, especially when estimating consumption for numerous buildings.

For the most part, the coefficients assigned to the variables used by the linear regressions are not supported by building science principles. It may be suggested that feature and model selection is based off of performance metrics, rather than significant relationships, however more work is needed to make this conclusion. Academic buildings were removed from the sample for a number of reasons (archetype, size, outlier, etc.) however only those whose consumption significantly differed from the subset averages were analyzed for similarities. Not only were there no measured similarities between the outlier buildings, but their properties were not always at the extreme end of the spectrum. This either suggests that another factor, not accounted for in this thesis (e.g. lab energy intensity, as suggested in the Discussion), is responsible for their above-average electricity consumption, or that buildings simply cannot be pre-screened for compatibility with this method. In either case, more data on outlier buildings (i.e. repeat applications on other building samples) will help identify if the model suffers from limitations yet to be identified.

This thesis detailed a method of estimation for academic buildings based on multiple linear regression. It is a reliable tool for Ryerson University to use for estimating electricity consumption in their cluster-metered buildings that have yet to be upgraded. Ryerson University aside, this thesis is useful for other universities or colleges where consumption data is lacking for whatever reason. It is expected that similar successes with this method will be had with modeling other building types, based on past literature. Ultimately, this method should offer another option for trained or untrained individuals working towards the measurement and publication of energy consumption for all buildings.

REFERENCES

[1] Office of Energy Efficiency, "Energy Use Data Handbook," Natural Resources Canada, Ottawa, 2013.

[2] Greening Greater Toronto. (2009, Jan 15). *Commercial Building Energy Initiative*. Available: http://www.greeninggreatertoronto.ca/initiatives/commercial.asp

[3] G. R. Newsham, "New Energy Codes and Their Impact on Lighting Design," *Proceedings, Institute for Research in Construction, National Research Council Canada,* vol. 19, pp. 3-11, 1994.

[4] Business Development Bank Of Canada. (2011, Jan 15). *Financing options for making buildings greener*. Available: http://www.bdc.ca/en/advice_centre/articles/Pages/financing_green_buildings.aspx - .T9tAyxc1_4T

[5] Office of Energy Efficiency, "Survey of Commercial and Institutional Energy Use - Buildings 2009," Natural Resources Canada, Ottawa, 2012.

[6] Energy Efficiency Office, "Energy Efficiency & Beyond on Toronto's Sustainable Energy Plan," ed, 2007.

[7] Canada Green Building Council. LEED Project Profiles [Online]. Available: http://www.cagbc.org/leed/projectprofile_EN.aspx

[8] TOBuilt. (2009, Jan 15). *Reports*. Available: http://www.tobuilt.ca/Reports.html

[9] Statistics Canada. (2012, Jan 2). *Focus on Geography Series, 2011 Census*. Available: http://www12.statcan.gc.ca/census-recensement/2011/as-sa/fogs-spg/Facts-cma-eng.cfm?LANG=Eng&GK=CMA&GC=535

[10] Institute for building efficiency. (2010, Jan 16). *Why Focus on Existing Buildings?* Available: http://www.institutebe.com/Existing-Building-Retrofits/Why-Focus-On-Existing-Buildings.aspx

[11] D. Ürge-Vorsatz and A. Novikova, "Potentials and costs of carbon dioxide mitigation in the world's buildings," *Energy Policy,* vol. 36, pp. 642-661, 2008.

[12] J. Briones, J. Greenwood, and S. Vasta, "The Market Impace of Accessible Energy Data," MaRS Discovery District, 2012.

[13] IESO, "Smart Metering Entity (SME) Time-of-Use Mandate Progress Report Through June 30, 2013," 2013.

[14] T. Ryberg, "Smart Metering in North America and Asia-Pacific," ed: Berg Insight, 2011.

[15] International Passive House Association. (n.d., Aug 1). *Passive House legislation*. Available: http://www.passivehouse-international.org/index.php?page_id=176

[16] U.S. Environmental Protection Agency, "Portfolio Manager Upgrade Project," ed, n.d.

[17] Natural Resources Canada. (2014, Sept 7). *Energy Star Portfolio Manager Project Timeline*. Available: http://www.nrcan.gc.ca/energy/efficiency/buildings/energy-benchmarking/faq/15373

[18] Association for the Advancement of Sustainability in Higher Education. (2014, Sept 7). *STARS Dashboard*. Available: https://stars.aashe.org/institutions/data-displays/dashboard/

[19] K. Palmer and M. Walls, "Can Benchmarking and Disclosure Laws Provide Incentives for Energy Efficiency Improvements in Commercial Buildings?," 2014.

[20] City of Toronto, "2011 Annual Energy Consumption & Greenhouse Gas (GHG) Emissions," E. W. M. Office, Ed., ed. Toronto, 2013.

[21] Environment Canada. (2013, Nov 1). *Canadian Climate Normals 1971-2000 Station Data*. Available: http://www.climate.weather.gc.ca/climate_normals/results_e.html?stnID=5051&lang=e&dCode=1&province=ONT&provBut=Go&month1=0&month2=12

[22] M. Kottek, J. Grieser, C. Beck, B. Rudolf, and F. Rubel, "World map of the Koppen-Geiger climate classification updated," *Meteorologische Zeitschrift*, vol. 15, pp. 259-264, 2006.

[23] B. Purcell. (2013, Jan 23). *Nailed it! Toronto exceeds Kyoto target*. Available: http://www.toatmosphericfund.ca/2013/04/24/toronto-exceeds-kyoto-target/

[24] ICF International, "Greenhouse Gases and Air Pollutants in the City of Toronto," 2007.

[25] Toronto Environmental Alliance, "Toronto Smog Report Card 2007," Toronto, 2007.

[26] Ryerson University. (2013, Jan 2). *Ryerson University at a Glance*. Available: http://www.ryerson.ca/news/media/quickfacts/

[27] Council of Ontario Universities, "Ontario Universities' Facilities Condition Assessment Program," Toronto, Ontario 1718-0260, 1718-0279, 2010.

[28] Ryerson University. (n.d., Jan 23, 2014). Energy and GHG Report 2013. Available: http://www.ryerson.ca/content/dam/campusfacilities/docs/Ryerson University GHG Report-2013-07.pdf

[29] U.S. Energy Information Administration. (n.d., Feb 3). *2003 CBECS Survey Data*. Available: http://www.eia.gov/consumption/commercial/data/2003/

[30] J.-F. Bonnet, C. Devel, P. Faucher, and J. Roturier, "Analysis of electricity and water end-uses in university campuses: case-study of the University of Bordeaux in the framework of the Ecocampus European Collaboration," *Journal of Cleaner Production*, vol. 10, pp. 13-24, 2002.

[31] Ryerson University, "Ryerson Campus Map," ed, 2013.

[32] K. McCusker, "Measuring, Managing and Visualizing Building Energy Consumption & Carbon Emissions," Master of Science, Environmental Conservation, University of Massachusetts Amherst, 2013.

[33] Z. Rahman, "Energy Analysis Using Carrier HAP Program and Deep Lake Water Cooling (DLWC) Feasibility Analysis for Ryerson University," Master of Engineering, Mechanical Engineering, Ryerson University, Toronto, 2010.

[34] J. Caeiro, W. Gang, and J. Glusing, "Identifying the determinants of energy use in Texas A&M University campus at Kingsville," in *Reston, VA: ASCEProceedings of the 2011 Architectural Engineering National Conference, March 30-April 2, 2011, Oakland, CA| d 20110000*, 2011.

[35] C. NORÉN and J. PYRKO, "An analysis of electricity consumption and load de-mand in Swedish grocery stores," *ECEEE Summer Study 1999*, 1999.

[36] S. Tassou, Y. Ge, A. Hadawey, and D. Marriott, "Energy consumption and conservation in food retailing," *Applied Thermal Engineering*, vol. 31, pp. 147-156, 2011.

[37] W. Chung, "Review of building energy-use performance benchmarking methodologies," *Applied Energy,* vol. 88, pp. 1470-1479, 2011.
[38] S.-M. Hong, G. Paterson, D. Mumovic, and P. Steadman, "Improved benchmarking comparability for energy consumption in schools," *Building Research & Information,* vol. 42, pp. 47-61, 2014.
[39] I. Ward, A. Ogbonna, and H. Altan, "Sector review of UK higher education energy consumption," *Energy Policy,* vol. 36, pp. 2939-2949, 2008.
[40] N. Baker and K. Steemers, *Energy and environment in architecture: a technical design guide*: Taylor & Francis, 1999.
[41] J. C. Lam, R. Y. Chan, C. Tsang, and D. H. Li, "Electricity use characteristics of purpose-built office buildings in subtropical climates," *Energy Conversion and Management,* vol. 45, pp. 829-844, 2004.
[42] P. Faucher, "Le Programme ECOCAMPUS, étape préalable à la prise en compte de la problématique du développement durable dans la gestion des Universités et Centres de recherche – Analyse méthodologique et technologique fondée sur l'étude de cas de l'Université Bordeaux 1," University of Bordeaux, 1999.
[43] J. F. Bonnet, C. Devel, P. Lagière, P. Faucher, and J. Roturier, "Environmental management policies in the Higher Education (HE) sector: quantified aspects from 5 case-studies in France under the ECOCAMPUS initiative," in *European Council for an Energy Efficient Economy Summer Study*, 2001.
[44] J. Roturier, J. F. Bonnet, P. Faucher, J. S. Nørgård, E. Loot, S. Gabriel*, et al.*, "Ecocampus: a "Practive-What-You-Preach" European Collaboration : Final Report of the ECOCAMPUS work group," 2000.
[45] P. Jones, R. Turner, D. Browne, and P. Illingworth, "Energy benchmarks for public sector buildings in Northern Ireland," in *Proceedings of CIBSE National Conference, Dublin*, 2000.
[46] Office of Energy Efficiency, "Commercial And Institutional Building Energy Use Survey 2000," Government of Canada, Ottawa, 2002.
[47] S. Chow, A. R. Ganji, B. Hackett, P. Parkin, and A. Fetters, "Energy assessment of selected schools in Anchorage School District," in *26th World Energy Engineering Congress, Atlanta, GA*, 2003, pp. 12-14.
[48] P. Jansson, J. Blanck, P. Giordano, D. Johnson, and S. Ross, "Undergraduate Research on Sustainability: Campus Energy Analysis and Building Energy Audits," in *American Society for Engineering Education Annual Conference and Exposition*, 2004.
[49] P. Faucher, A. d. Almeida, E. Apostolidou, D. Berbecaru, D. Bernard, T. Laine*, et al.*, "Implementing Energy-Efficient and Environment-Safe Programs in the Management of European University Campuses and Research Laboratories : Case-Studies from the ECOCAMPUS Collaboration," in *Proceedings from the 1998 ACEEE Summer-Study on Energy Efficiency in Buildings : Energy Efficiency in a Competitive Environment*, Pacific Grove, California, 1998.
[50] G. Griffith, "Estates: The foundations of Higher Education Higher Education Estate Statistics 2011/2012," 2013.

[51] Office of Energy Efficiency, "Consumption of Energy Survey for Universities, Colleges and Hospitals, 2003," Government of Canada, 2005.
[52] Higher Education Statistics Agency, "Estates Management Record - Environmental," ed, 2011.
[53] Higher Education Environmental Performance Improvement. (2006, Results of the HEEPI HE Building Energy Benchmarking Initiative 2003-4.
[54] International Institute for Sustainable Laboratories. Labs21 Energy Benchmarking Tool [Online]. Available: http://labs21benchmarking.lbl.gov
[55] J. Elliott and A. Guggemos, "Validating electric use intensity in multi-use buildings," *Journal of Facilities Management,* vol. 9, pp. 52-63, 2011.
[56] Building Research Energy Conservation Support Unit, "Energy efficiency in further and higher education," 1997.
[57] U.S. Department of Energy, "Energy Efficiency Trends in Residential and Commercial Buildings," 2008.
[58] A. T. Zhou, Vestal; Harris, Jeffrey; Bostrom, Paul, "Promoting Energy-Efficient Buildings in the Industrial Sector," 2011.
[59] Office of Energy Efficiency, "Benchmarking Guide for School Facility Managers," Government of Canada, Ottawa, 2001.
[60] H.-X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews,* vol. 16, pp. 3586-3592, 2012.
[61] G. K. Tso and K. K. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy,* vol. 32, pp. 1761-1768, 2007.
[62] Energy Efficiency & Renewable Energy. (2011, April 8). *Building Energy Software Tools Directory.* Available: http://apps1.eere.energy.gov/buildings/tools_directory/subjects.cfm/pagename=subjects/pagename_menu=whole_building_analysis/pagename_submenu=energy_simulation
[63] J. MacDonald and D. Wasserman, *Investigation of metered data analysis methods for commercial and related buildings*: Oak Ridge National Laboratory, 1989.
[64] M. S. Al-Homoud, "Computer-aided building energy analysis techniques," *Building and Environment,* vol. 36, pp. 421-433, 2001.
[65] R. Signor, F. S. Westphal, and R. Lamberts, "Regression analysis of electric energy consumption and architectural variables of conditioned commercial buildings in 14 Brazilian cities," in *Seventh International IBPSA Conference, Rio de Janeiro, Brazil*, 2001, pp. 1373-1379.
[66] F. S. Westphal and R. Lamberts, "Regression analysis of electric energy consumption of commercial buildings in Brazil," in *Proceedings of Building Simulation*, 2007, pp. 1543-1550.
[67] J. C. Lam, S. Hui, and A. L. Chan, "Regression analysis of high-rise fully air-conditioned office buildings," *Energy and Buildings,* vol. 26, pp. 189-197, 1997.

[68] S. Kajl, M. Roberge, L. Lamarche, and P. Malinowski, "Evaluation of building energy consumption based on fuzzy logic and neural networks applications," in *Proc of CLIMA*, 2000, p. 264.
[69] A. T. Chan and V. C. Yeung, "Implementing building energy codes in Hong Kong: energy savings, environmental impacts and cost," *Energy and Buildings,* vol. 37, pp. 631-642, 2005.
[70] O. Z. Maimon and L. Rokach, *Data mining and knowledge discovery handbook* vol. 1: Springer, 2005.
[71] A. M. A. Ghoson, "Decision tree induction & clustering techniques in SAS Enterprise Miner, SPSS Clementine, and IBM Intelligent Miner–a comparative analysis," *International Journal of Management & Information Systems (IJMIS),* vol. 14, 2011.
[72] Z. Yu, F. Haghighat, B. Fung, and H. Yoshino, "A decision tree method for building energy demand modeling," *Energy and Buildings,* vol. 42, pp. 1637-1646, 2010.
[73] K. P. Burnham and D. R. Anderson, *Model selection and multimodel inference: a practical information-theoretic approach*: Springer, 2002.
[74] Mu Sigma. (n.d., May 3). *NEURAL NETWORK*. Available: http://www.mu-sigma.com/analytics/thought_leadership/cafe-cerebral-neural-network.html
[75] B. Lantz, *Machine Learning with R*: Packt Publishing Ltd, 2013.
[76] D. Hawkins, S. Hong, R. Raslan, D. Mumovic, and S. Hanna, "Determinants of energy use in UK higher education buildings using statistical and artificial neural network methods," *International Journal of Sustainable Built Environment,* vol. 1, pp. 50-63, 2012.
[77] M. Yalcintas and U. Aytun Ozturk, "An energy benchmarking model based on artificial neural network method utilizing US Commercial Buildings Energy Consumption Survey (CBECS) database," *International journal of energy research,* vol. 31, pp. 412-421, 2007.
[78] S.-M. Hong, G. Paterson, E. Burman, P. Steadman, and D. Mumovic, "A comparative study of benchmarking approaches for non-domestic buildings: Part 1–Top-down approach," *International Journal of Sustainable Built Environment,* 2014.
[79] S. Boonyatikarn, "Method for developing energy budgets and energy design guidelines for institutional buildings," 1982.
[80] N. Leslie, G. Aveta, and B. Sliwinski, "Regression based process energy analysis system," *ASHRAE Trans.;(United States),* vol. 92, 1986.
[81] R. Mazzucchi, "The project on restaurant energy performance end-use monitoring and analysis," *ASHRAE Trans.;(United States),* vol. 92, 1986.
[82] J. S. Haberl and D. E. Claridge, "An expert system for building energy consumption analysis: prototype results," *ASHRAE transactions,* vol. 93, pp. 979-998, 1987.
[83] P. M. Lukacs, K. P. Burnham, and D. R. Anderson, "Model selection bias and Freedman's paradox," *Annals of the Institute of Statistical Mathematics,* vol. 62, pp. 117-125, 2010.
[84] P. F. Doherty, G. C. White, and K. P. Burnham, "Comparison of model building and selection strategies," *Journal of Ornithology,* vol. 152, pp. 317-323, 2012.
[85] T. Sharp, "Benchmarking energy use in schools," in *Proceedings of the ACEEE 1998 Summer Study on Energy Efficiency in Buildings*, 1998.

[86] T. Sharp, "Energy benchmarking in commercial office buildings," in *Proceedings of the ACEEE*, 1996, pp. 321-329.
[87] W. Chung, Y. Hui, and Y. M. Lam, "Benchmarking the energy efficiency of commercial buildings," *Applied Energy,* vol. 83, pp. 1-14, 2006.
[88] S. E. Lee and P. Rajagopalan, "Building energy efficiency labeling programme in Singapore," *Energy policy,* vol. 36, pp. 3982-3992, 2008.
[89] S. Lee, Y. Park, and C. Kim, "Investigating the Set of Parameters Influencing Building Energy Consumption," in *Reston, VA: ASCEProceedings of the 2011 International Conference on Sustainable Design and Construction| d 20120000*, 2012.
[90] C. Stoy, S. Pollalis, and D. Fiala, "Estimating buildings' energy consumption and energy costs in early project phases," *Facilities,* vol. 27, pp. 187-201, 2009.
[91] E. Azar and C. C. Menassa, "A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings," *Energy and Buildings,* vol. 55, pp. 841-853, 2012.
[92] J. U. Straube, Kohta, "BSD-151: Understanding Primary/Source and Site Energy," 2010.
[93] T. Olofsson, A. Meier, and R. Lamberts, "Rating the energy performance of buildings," *The International Journal of Low Energy and Sustainable Buildings,* vol. 3, 2004.
[94] Natural Resources Canada. (2013, March 7). *About Electricity*.
[95] E. Chung, "Federal science hobbled by cuts and policies, poll says," ed: CBC, 2014.
[96] U.S. Energy Information Administration, "Average Retail Prices of Electricity, 1960-2011," in *Annual Energy Review*, ed, 2012.
[97] B. Ó Gallachóir, M. Keane, E. Morrissey, and J. O'Donnell, "Using indicators to profile energy consumption and to inform energy policy in a university—A case study in Ireland," *Energy and buildings,* vol. 39, pp. 913-922, 2007.
[98] APS, "Envelope," 2001.
[99] S. Tangri, "Master Ryerson Energy Use," Ryerson University, Ed., V1 ed, 2012.
[100] University of Toronto, "Long Range Budget Guidelines 2013-14 to 2017-18 and Budget Report 2013-14," 2013.
[101] Ryerson University, "2013-14 Budget Priorities and Expensitures," 2013.
[102] Google Maps, "University of Toronto," ed, 2014.
[103] Council of Ontario Universities, "Council of Ontario Universities' Building Blocks," ed. Toronto, Ontario, 2010.
[104] J. Fauquier, "Room Share Area," Ryerson University, Ed., ed, 2012.
[105] B. Armstrong, "Bldg Summary by COU Category," University of Toronto, Ed., ed, 2013.
[106] University of Toronto. (2009). *UofT Campus Map*. Available: http://www.osm.utoronto.ca/map/f?p=110:1:1071464023124501:::::
[107] TomTom, "University of Toronto," ed, 2013.
[108] R Core Team, "R: A Language and Environment for Statistical Computing," ed. Vienna, Austria: R Foundation for Statistical Computing, 2014.
[109] (n.d., May 27). *What is R?* Available: http://www.r-project.org
[110] R. Ihaka, "R: Past and future history," *COMPUTING SCIENCE AND STATISTICS,* pp. 392-396, 1998.

[111] M. W. Beck, "Collinearity and stepwise VIF selection," in *R is my friend* vol. 2014, ed, 2013.
[112] H. Bozdogan, "Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions," *Psychometrika,* vol. 52, pp. 345-370, 1987.
[113] M. J. Mazerolle, "Appendix 1: Making sense out of Akaike’s Information Criterion (AIC): its use and interpretation in model selection and inference from ecological data," *Mouvements et Reproduction des Amphibiens en Tourbières Perturbées,* pp. 174-190, 2004.
[114] J. Starkweather, "Cross Validation techniques in R: A brief overview of some methods, packages, and functions for assessing prediction models," 2011.
[115] D. M. Hawkins, S. C. Basak, and D. Mills, "Assessing model fit by cross-validation," *Journal of chemical information and computer sciences,* vol. 43, pp. 579-586, 2003.
[116] S. Karatasou, M. Santamouris, and V. Geros, "Modeling and predicting building's energy use with artificial neural networks: Methods and results," *Energy and Buildings,* vol. 38, pp. 949-958, 2006.
[117] J. Yang, H. Rivard, and R. Zmeureanu, "On-line building energy prediction using adaptive artificial neural networks," *Energy and buildings,* vol. 37, pp. 1250-1259, 2005.
[118] C. McCormick, "K-FOLD CROSS-VALIDATION, WITH MATLAB CODE," F. CV, Ed., ed, 2013.
[119] J. A. Davis III and D. W. Nutter, "Occupancy diversity factors for common university building types," *Energy and buildings,* vol. 42, pp. 1543-1551, 2010.
[120] U.S. Department of Energy. (2011, July 10). *HAP*. Available: http://apps1.eere.energy.gov/buildings/tools_directory/software.cfm/ID=385/pagename_submenu=/pagename_menu=/pagename=
[121] Province of British Columbia, "B.C. Universities Space Manual," M. o. A. Education, Ed., ed, 2003.
[122] L. Dormehl. (2013, July 8). *It's all in the planning*. Available: http://universitybusiness.co.uk/Featured-Content/its_all_in_the_planning
[123] M. Lewis, "Stepwise versus Hierarchical Regression: Pros and Cons," *Online Submission,* 2007.
[124] B. O'Meara. (n.d., Aug 13). *Model Selection Using the Akaike Information Criterion (AIC)*. Available: http://www.brianomeara.info/tutorials/aic
[125] M. Nielsen, "If correlation doesn’t imply causation, then what does?," in *DDI* vol. 2014, ed, 2012.
[126] J. Frost, "Regression Analysis: How Do I Interpret R-squared and Assess the Goodness-of-Fit?," in *The Minitab Blog* vol. 2014, ed, 2013.
[127] J. Frost, "Multiple Regression Analysis: Use Adjusted R-Squared and Predicted R-Squared to Include the Correct Number of Variables," in *The Minitab Blog* vol. 2014, ed, 2013.
[128] J. Sauro. (n.d., Aug 12). *Fundamentals of Statistics 1: Basic Concepts :: The Standard Deviation and Coefficient of Variation*. Available: http://www.usablestats.com/lessons/sdcv

[129] T. Baguley, "Multicollinearity and collinearity (in multiple regression) - a tutorial," in *Psychological Statistics* vol. 2014, ed, 2013.
[130] Explorable.com. (2009, Aug 13). *Statistical Correlation*. Available: https://explorable.com/statistical-correlation
[131] C. McCormick, "K-FOLD CROSS-VALIDATION, WITH MATLAB CODE," in *CHRIS MCCORMICK* vol. 2014, ed, 2013.
[132] J. Schneider. (1997, Aug 12). *Cross Validation*. Available: http://www.cs.cmu.edu/~schneide/tut5/node42.html
[133] S. Pandey and C. L. Bright, "What Are Degrees of Freedom?," *SOCIAL WORK RESEARCH-NEW YORK-*, vol. 32, p. 119, 2008.
[134] D. C. E. Sonte, Jon. (2006, Aug 12). *Stats Tutorial - Degrees of Freedom*. Available: http://www.chem.utoronto.ca/coursenotes/analsci/StatsTutorial/DegFree.html
[135] D. A. Freedman and D. A. Freedman, "A note on screening regression equations," *the american statistician*, vol. 37, pp. 152-155, 1983.
[136] StatTrek.com. (n.d., Aug 13). *Interquartile Range*. Available: http://stattrek.com/statistics/dictionary.aspx?definition=interquartile_range
[137] Statistics.com. (n.d., Aug 13). *Likelihood Function:*
. Available: http://www.statistics.com/glossary&term_id=257
[138] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting*, vol. 22, pp. 679-688, 2006.
[139] P. Flom. (2012, Aug 12). *What's the difference between variance and standard deviation?* Available: http://stats.stackexchange.com/questions/35123/whats-the-difference-between-variance-and-standard-deviation
[140] Science Buddies. (n.d., Aug 12). *Variance & Standard Deviation*. Available: http://www.sciencebuddies.org/science-fair-projects/project_data_analysis_variance_std_deviation.shtml

APPENDICES

APPENDIX A1 – TORONTO CLIMATE NORMALS

TEMP.	JAN	F	M	A	M	J	J	A	S	O	N	D	YEAR
DAILY AVG	-4.2	-3.2	1.3	7.6	14.2	19.2	22.2	21.3	17	10.6	4.8	-0.9	9.2
SD	2.7	2.5	2	1.5	1.9	1.4	1.2	1.2	1.1	1.5	1.4	2.5	0.8
DAILY MAX	-1.1	-0.2	4.6	11.3	18.5	23.5	26.4	25.3	20.7	13.8	7.4	1.8	12.7
DAILY MIN	-7.3	-6.3	-2	3.8	9.9	14.8	17.9	17.3	13.2	7.3	2.2	-3.7	5.6
PRECIP.													
RAIN (MM)	29.1	26.2	42	63.2	73.3	71.5	67.5	79.6	83.4	64.7	67.3	41.9	709.8
SNOW (CM)	38.2	26.6	22	6	0	0	0	0	0	0.1	8.1	32.2	133.1
DAYS WITH >= 5 MM	4	3.2	4.3	4.7	4.7	4.5	4.1	4.4	4.7	4.2	5	4.5	52.3
DEGREE DAYS, ABOVE 18	0	0	0.1	1.4	16.8	62.6	132.3	109	35.4	1	0	0	358.7
TOTAL HOURS OF BRIGHT SUNSHINE	88.3	110.3	156.3	185.4	229.1	256.2	276.2	241.3	188	148.4	83.6	74.7	2037.6
GLOBAL INCIDENT RADIATIO N (MJ/M2)	153.7	227.2	378.8	478.2	599.5	652	677	578	418.1	283.5	146.7	120.6	4713.5

APPENDIX A2 – RYERSON BUILDING SPECIFICATIONS & DETAILS

Each building profile contains notes detailing the address, the year of original construction, the major academic departments housed, the average EUI, and the data integrity, which represents the number of measurement points that were used to measure the EUI.

Energy trends for each building show the annual consumption and average monthly consumption plotted with Toronto's HDD and CDD, and the months of September to May highlighted (typical academic year). Years representing unreliable data are transparent.

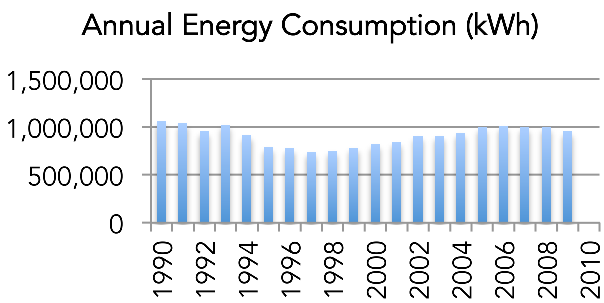
Sources of data:

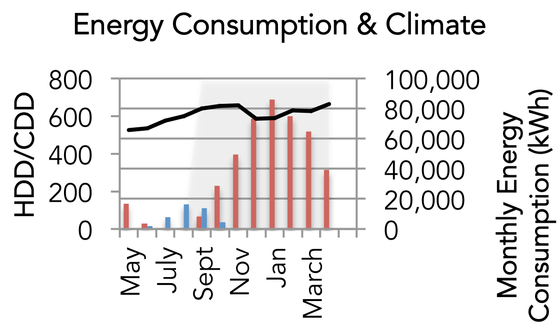
- Ryerson University
- Various Periodicals
- Urbandb
- Architecture/designer websites
- Wikipedia
- City of Toronto construction date map

Architecture Building

- 325 Church Street
- Built in 1981
- Department of Architectural Science
- 240/240 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 136.06 l 142.46

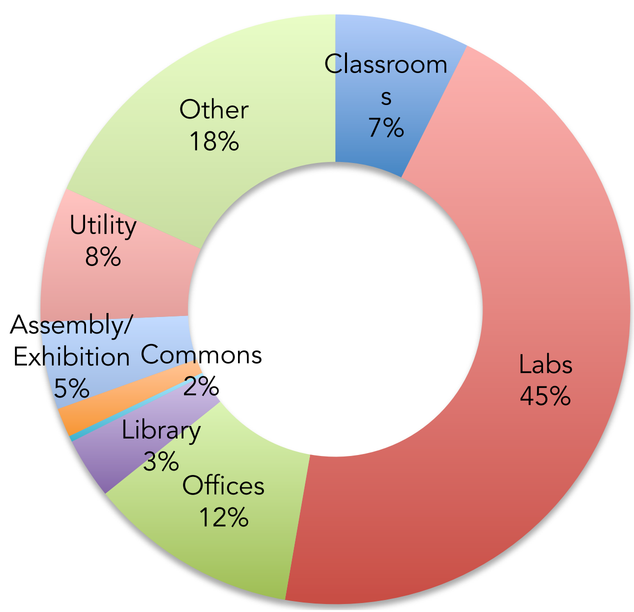
Energy Trends





Space Usage

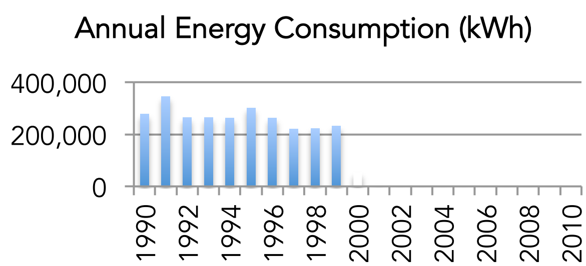
Usage	Area
Labs	3,031
Other	1,222
Offices	769
Utility	495
Classrooms	493
Assembly/Exhibition	321
Library	224
Commons	106
Food	22
Total	6,683

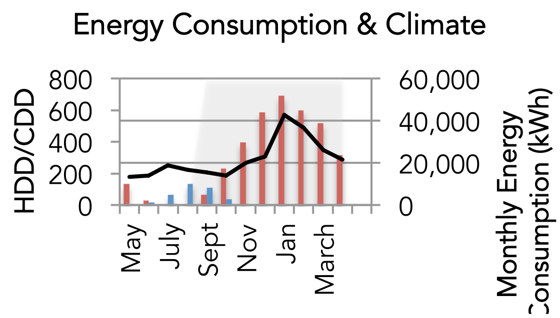


Bookstore

- 17 Gould Street
- Built in 1988
- No elevator
- 120/120 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 203.55 l 1.02

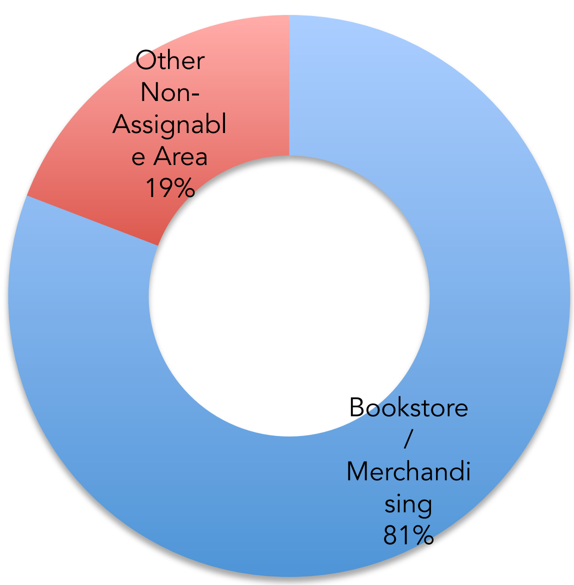
Energy Trends





Space Usage

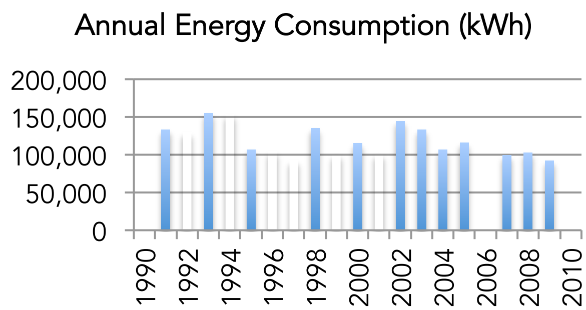
Usage	Area
Bookstore	1,052
Other	250
Total	1,302

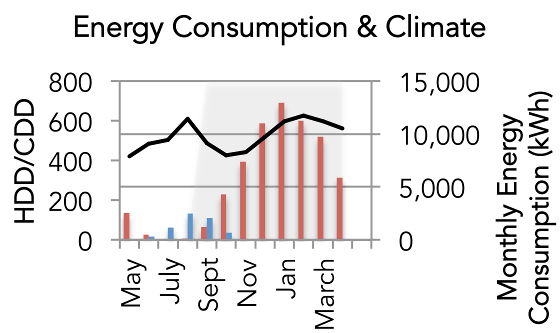


Cooperative Education

- 101 Gerrard Street East
- Built in 1950 (unverified)
- Originally housed Ryerson Theater School, now Office of Co-operative Education and Internship
- No elevator
- 208/216 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 187.33 l 144.59

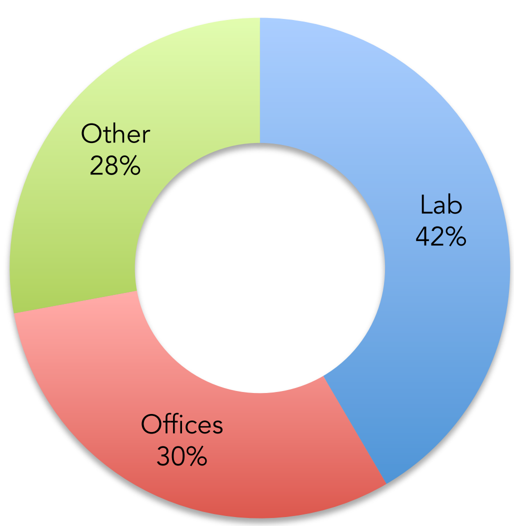
Energy Trends





Space Usage

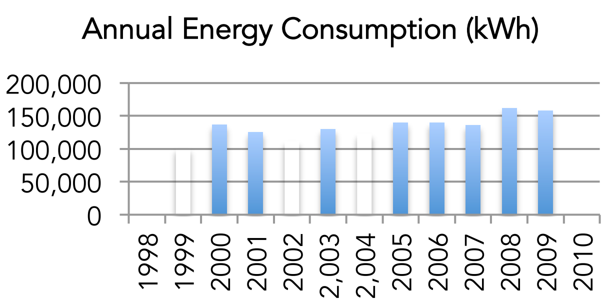
Usage	Area
Lab	265
Offices	195
Other	178
Total	638

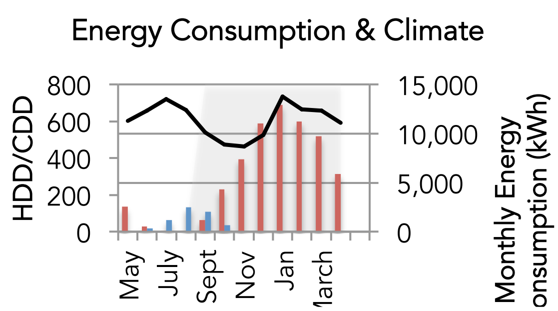


Campus Planning & Facilities

- 111 Bond Street
- Built in 1960 (unverified)
- Administrative offices for campus planning and facilities
- No elevator
- 128/132 Data Integrity
- EUI [kWh/m^2/year] (avg|2009) = 212.90 | 238.71

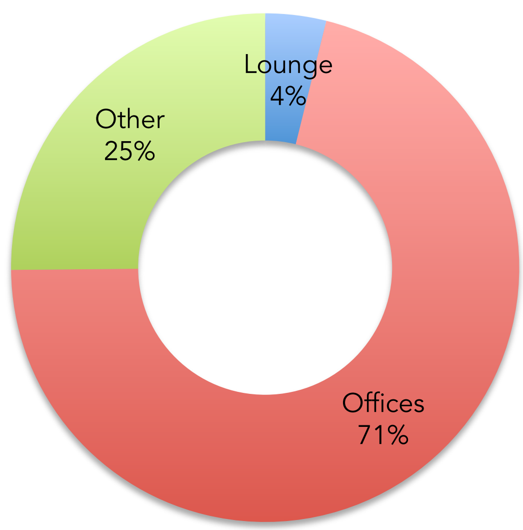
Energy Trends





Space Usage

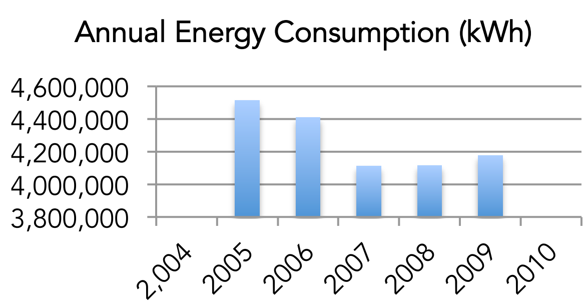
Usage	Area
Offices	470
Other	166
Lounge	26
Total	662

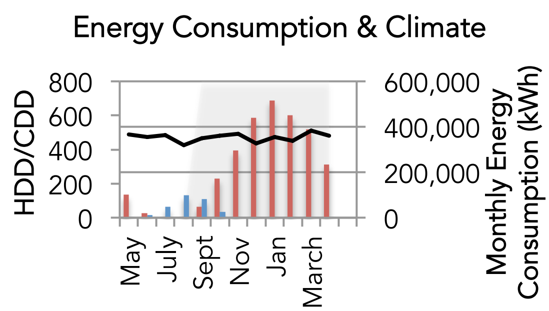


George Vari Engineering & Computing Center

- 245 Church Street
- Built in 2004
- Houses departments of Electrical Engineering, Computer Engineering, Computer Science and Aerospace Engineering, and four major Civil Engineering labs
- 60/60 Data Integrity
- EUI [kWh/m^2/year] (avg|2009) = 219.54 | 214.97

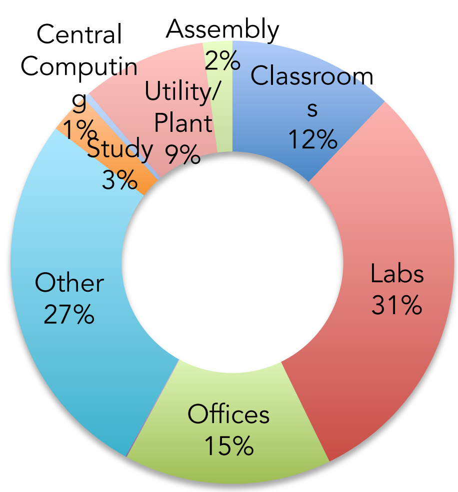
Energy Trends





Space Usage

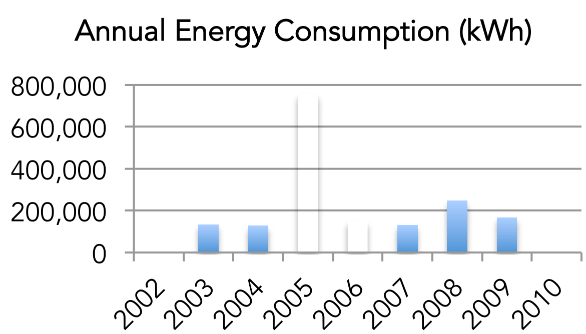
Usage	Area
Labs	5,969
Other	5,305
Offices	2,939
Classrooms	2,345
Utility/Plant	1,758
Study	577
Assembly	419
Central Computing	99
Food	17
Total	19,428

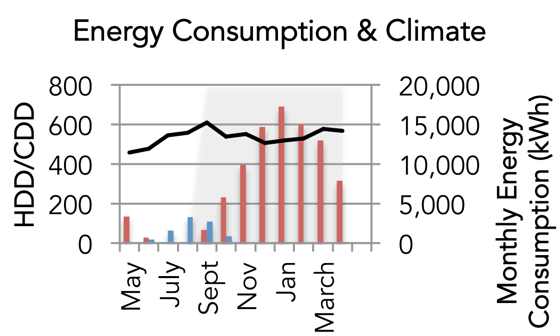


Research & Graduate Studies

- 111 Gerrard Street East
- Built in early 1950's; acquired by Ryerson in 2001
- No elevators
- 81/84 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 62.81 l 65.07

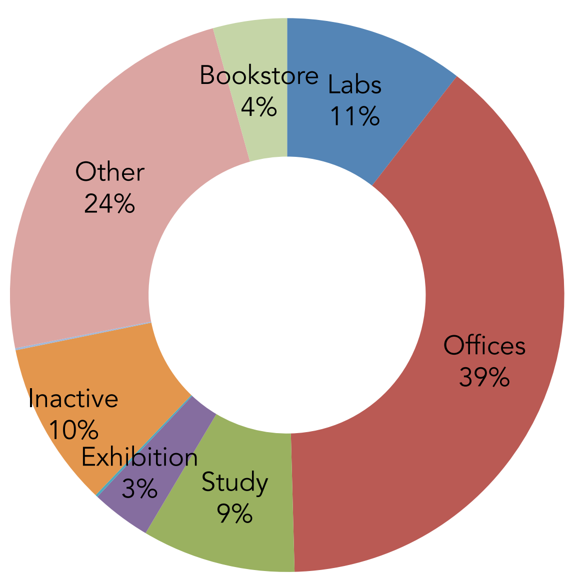
Energy Trends





Space Usage

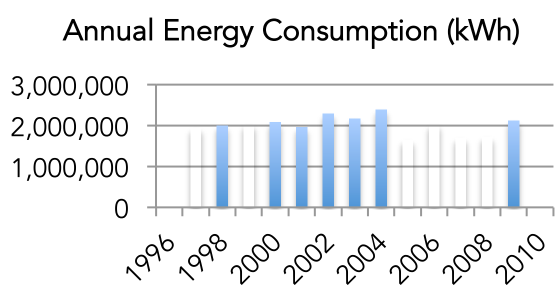
Usage	Area
Offices	1,002
Other	609
Labs	269
Inactive	248
Study	230
Bookstore	111
Exhibition	89
Central Computing	4
Utility	3
Total	2,565

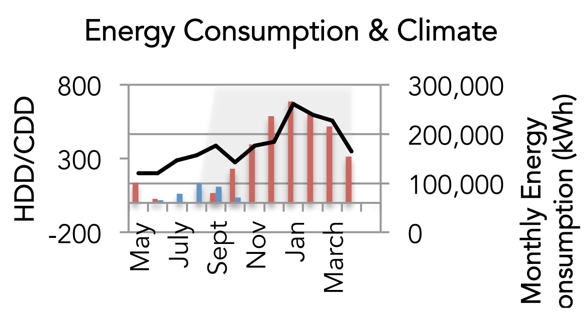


International Living Learning Center

- 133 Mutual Street
- Built in 1987
- Former Ibis Hotel; purchased by ryerson in 1993
- 252-room residence
- 146/156 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 183.19 l 181.05

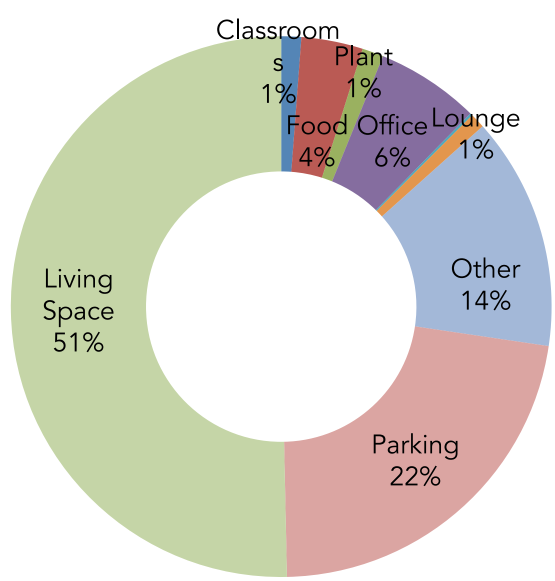
Energy Trends





Space Usage

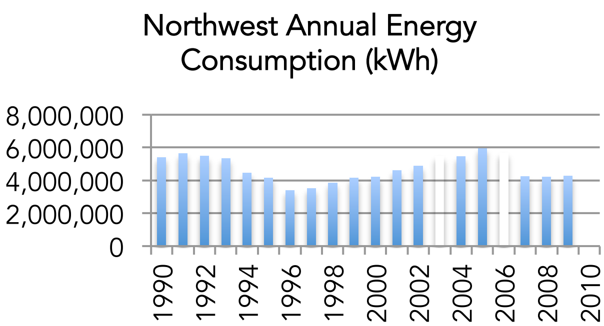
Usage	Area
Living Space	5,877
Parking	2,607
Other	1,629
Office	723
Food	430
Plant	149
Classrooms	139
Lounge	105
Central Computing	16
Total	11,675

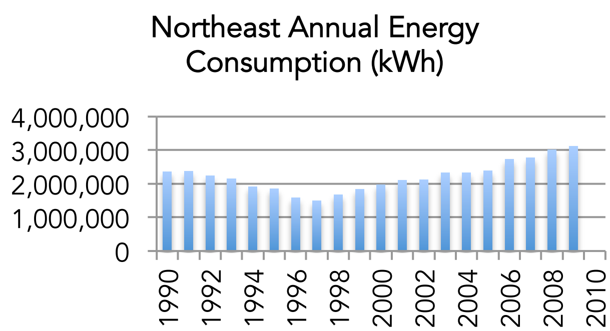


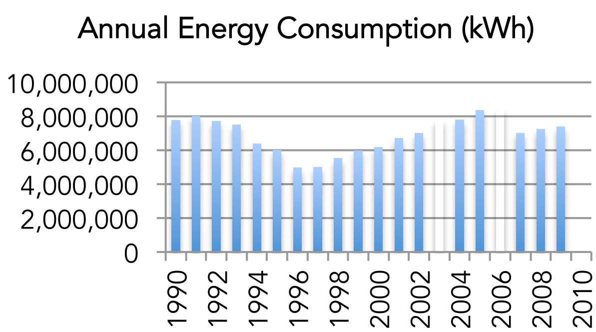
Kerr Hall

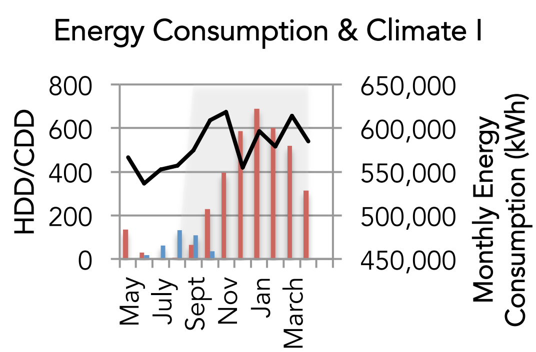
- (N) 43 Gerrard Street East, (E) 340 Church Street, (S) 50 Gould Street, (W) 379 Victoria Street
- North building houses Ryerson Theater
- Built from 1960-1969; opened in 1963
- 238/240 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 155.63 l 168.99

Energy Trends

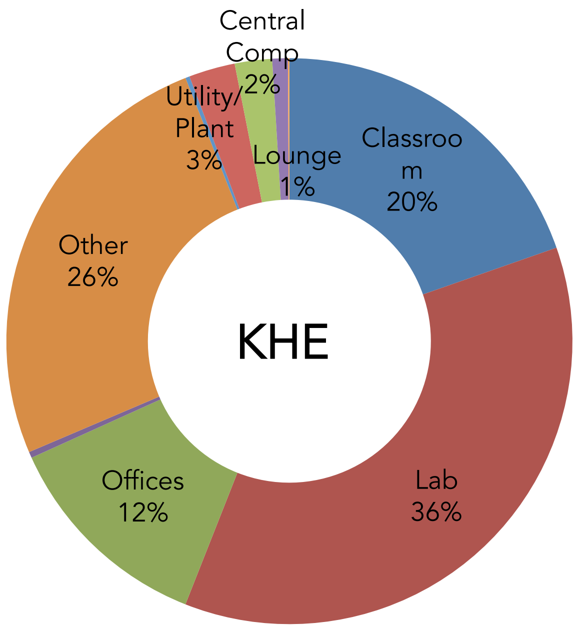


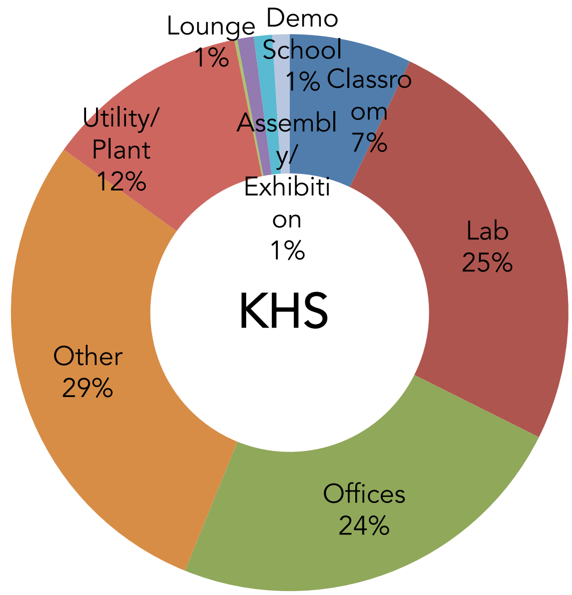


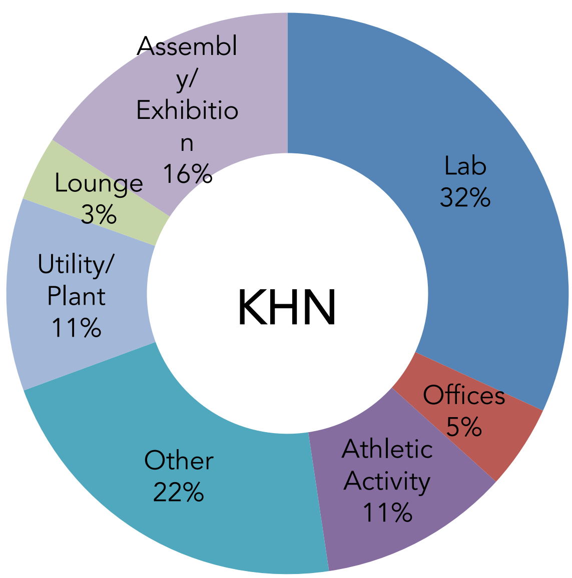


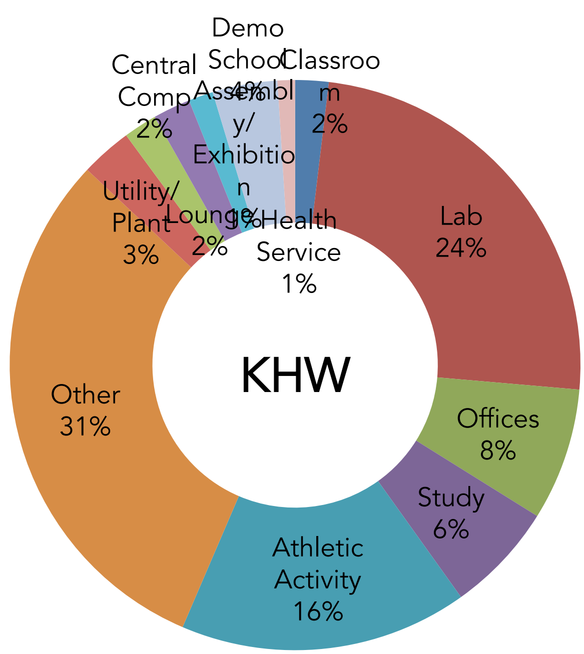


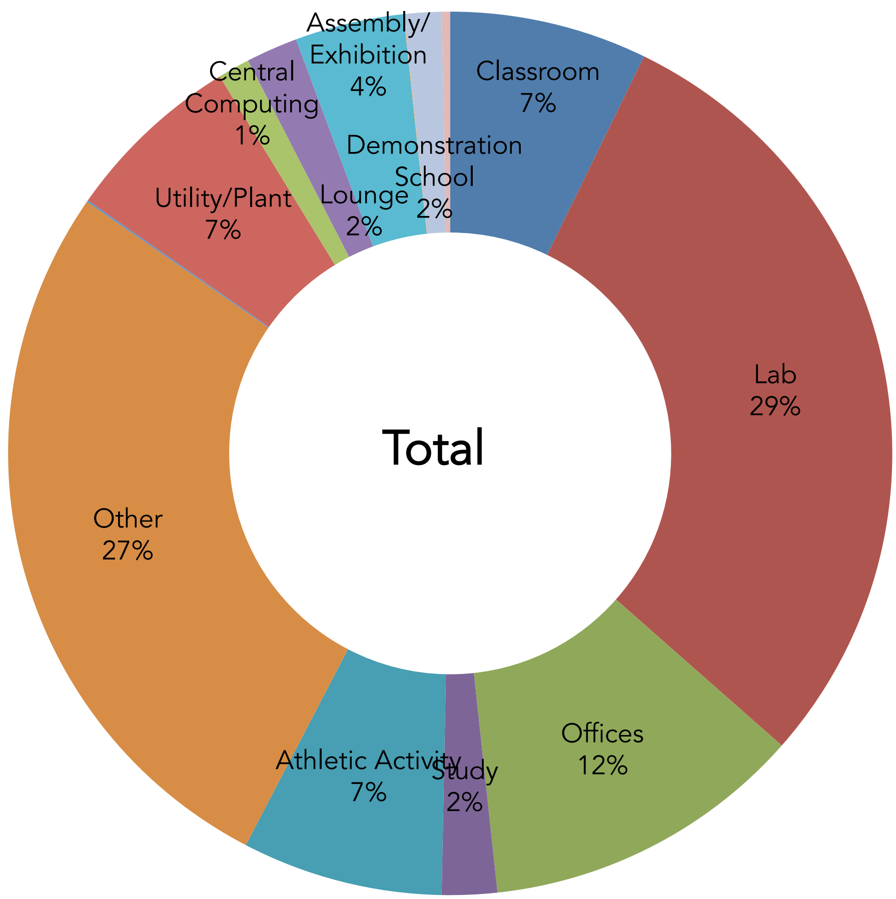
Space Usage


KHE
Central Comp 2%
Utility/ Plant 3%
Lounge 1%
Classroom 20%
Other 26%
Offices 12%
Lab 36%


KHS
Lounge 1%
Demo School 1%
Classroom 7%
Utility/ Plant 12%
Assembly/ Exhibition 1%
Lab 25%
Other 29%
Offices 24%


KHN
Assembly/ Exhibition 16%
Lounge 3%
Utility/ Plant 11%
Lab 32%
Offices 5%
Athletic Activity 11%
Other 22%


KHW
Demo School 4%
Classroom 2%
Central Comp 2%
Assembly/ Exhibition 1%
Utility/ Plant 3%
Lounge 2%
Health Service 1%
Lab 24%
Other 31%
Offices 8%
Study 6%
Athletic Activity 16%

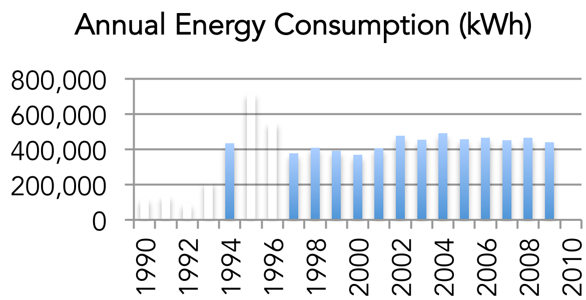


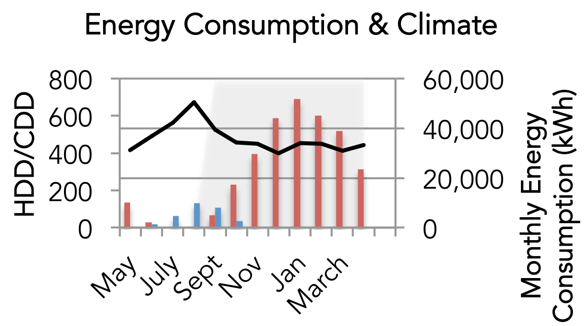
Usage	**KHN**	**KHE**	**KHS**	**KHW**	**Total**
Lab	2,935	4,096	2,469	3,304	12,804
Other	2,003	2,865	2,815	4,125	11,809
Classroom	0	2,213	693	255	3,162
Utility/Plant	1,024	298	1,158	392	2,872
Assembly/Exhibition	1,454	0	103	188	1,745
Study	0	39	0	843	882
Lounge	343	100	93	287	824
Demonstration School	0	0	97	497	594
Central Computing	0	240	17	256	513
Health Service	0	0	0	133	133
Food	0	27	0	0	27
Inactive	0	9	0	0	9
Total	7,759	9,887	7,445	10,280	35,374

Civil Engineering Building

- 341 Church Street
- Constructed in 1929
- Houses the Department of Civil Engineering
- 189/192 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 221.01 l 223.83

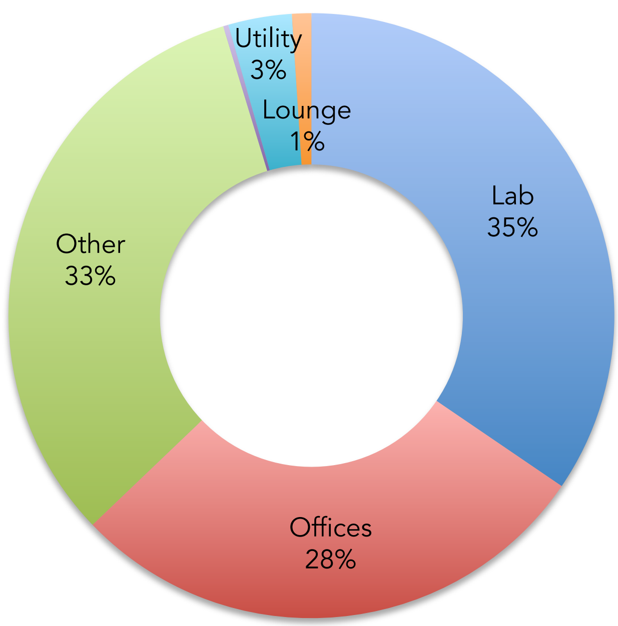
Energy Trends





Space Usage

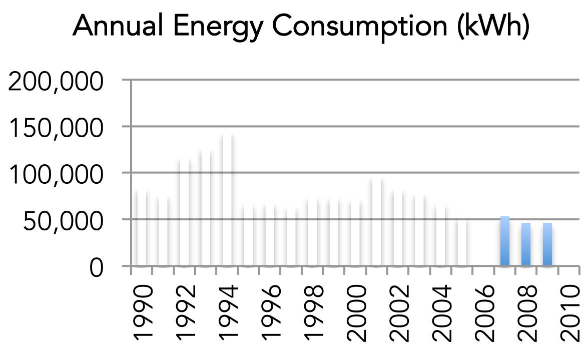
Usage	Area
Lab	678
Other	637
Offices	556
Utility	67
Lounge	20
Central Computing	5
Total	1963

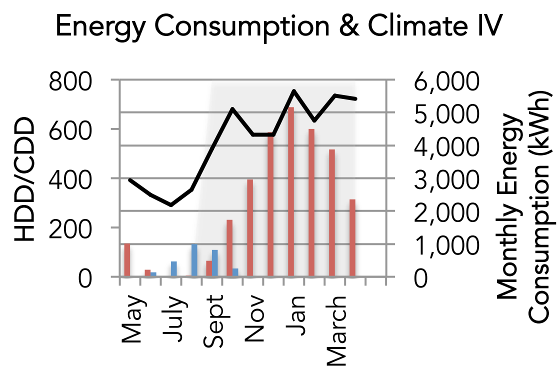


O'Keefe House

- 137 Bond Street
- Built in 1880; renovated in 1889 and 2004
- 33-room student residence
- No elevator
- 36/36 Data Integrity
- EUI [kWh/m^2/year] (avg|2009) = 70.46| 67.11

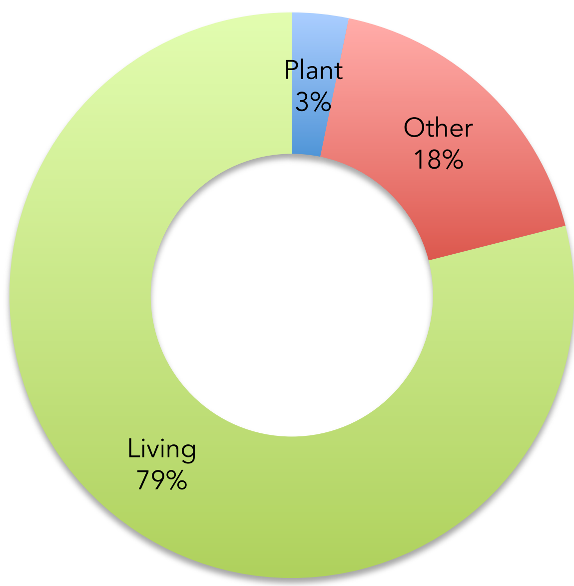
Energy Trends





Space Usage

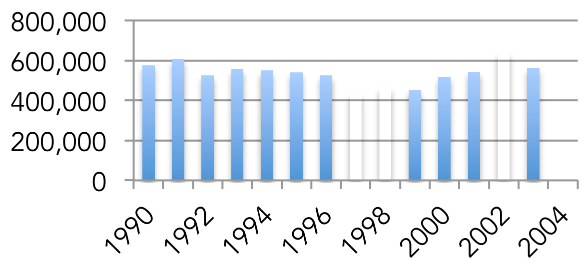
Usage	**Area**
Living	541
Other	122
Plant	22
Total	685

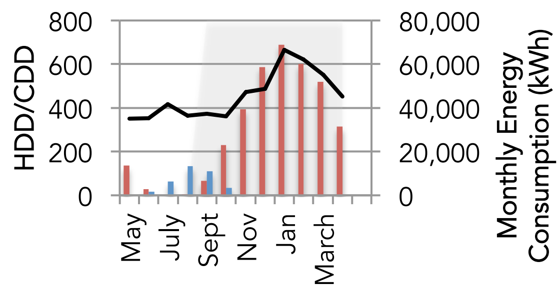


Oakham House

- 63 Gould Street
- Built in 1848, acquired by Ryerson in 1958
- Used for non-academic, cultural and recreation activity
- 164/168 Data Integrity
- EUI [kWh/m^2/year] (avg|2003) = 315.31 | 328.22

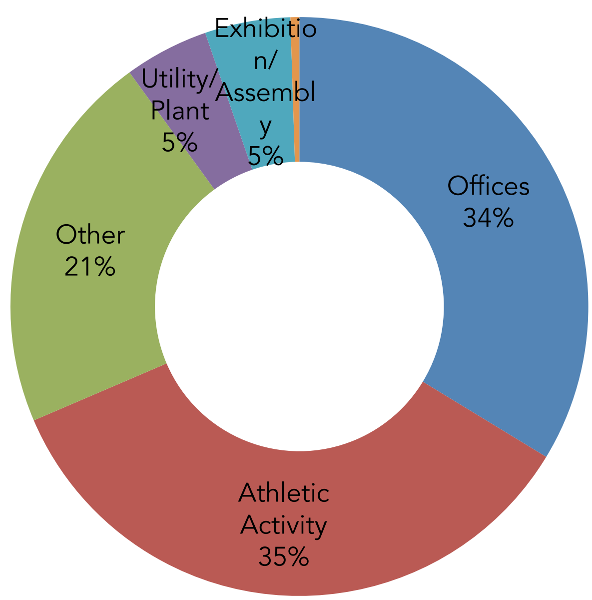
Energy Trends





Space Usage

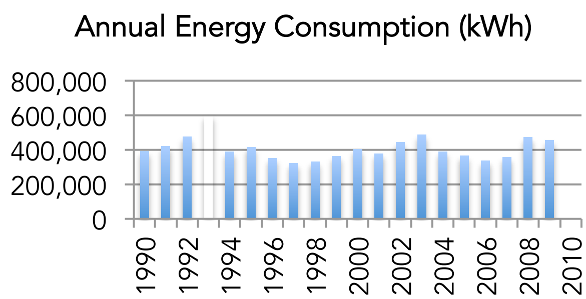
Usage	Area
Athletic Activity	597
Offices	577
Other	367
Exhibition/Assembly	82
Utility/Plant	81
Food	9
Total	1713

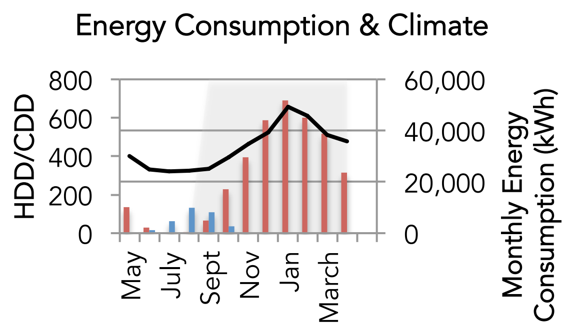


Parking Garage

- 300 Victoria Street
- Built in 1988
- No elevator
- 230/240 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 35.31 l 40.71

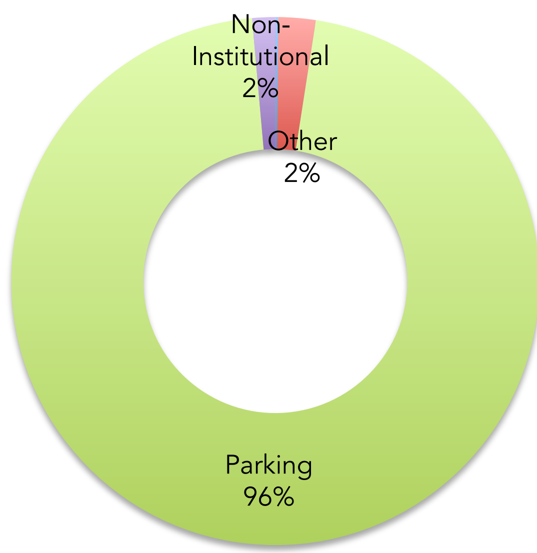
Energy Trends





Space Usage

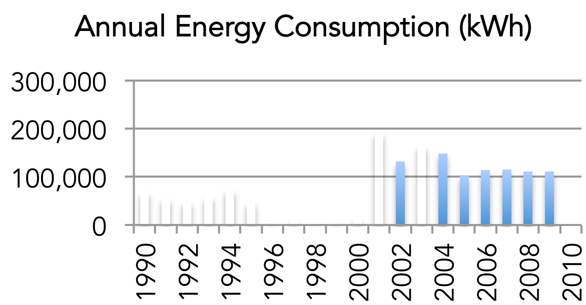
Usage	Area
Parking	10,747
Other	254
Non-Institutional	157
Plant	21
Total	11,179

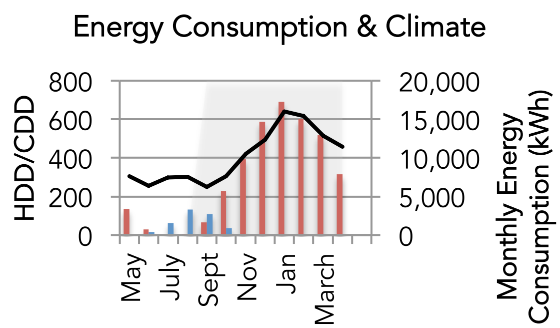


Projects Office

- 112 Bond Street
- Built in 1860 (unverified) or 1901-1930 (construction date map); acquired by Ryerson in 1966
- No elevators
- 93/96 Data Integrity
- EUI [kWh/m^2/year] (avg|2009) = 464.28 | 433.43

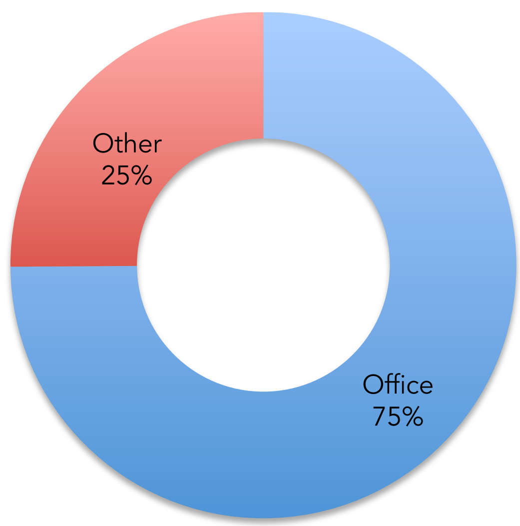
Energy Trends





Space Usage

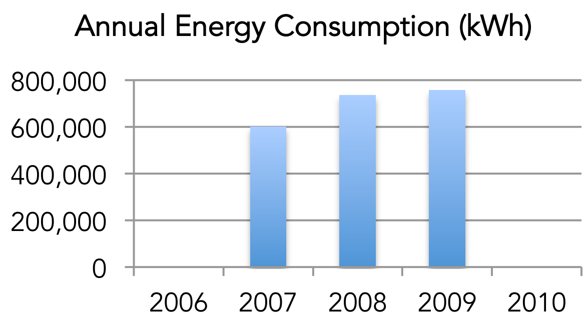
Usage	Area
Office	192
Other	64
Total	256

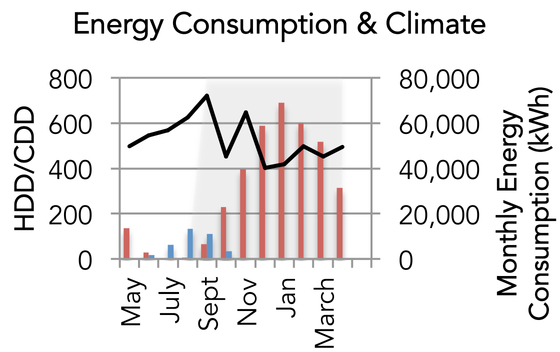


South Bond Building

- 105 Bond Street
- Built between 1976 and 2003; Acquired by Ryerson in 1966
- No elevators
- 36/36 Data Integrity
- EUI [kWh/m^2/year] (avg|2009) = 127.99 | 138.77

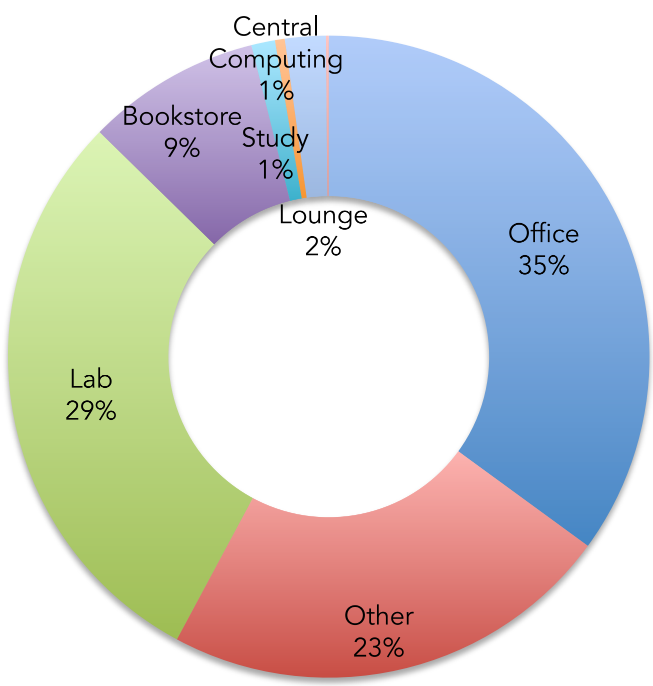
Energy Trends





Space Usage

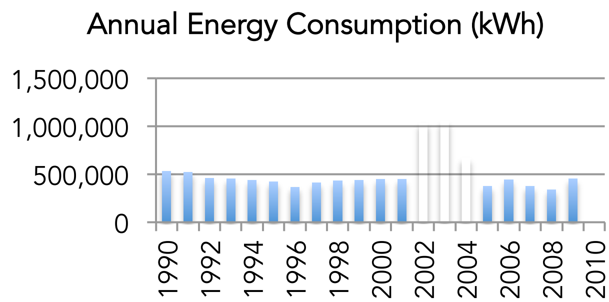
Usage	Area
Office	1,910
Lab	1,607
Other	1,244
Bookstore	479
Lounge	112
Study	65
Central Computing	27
Plant	8
Total	5,452

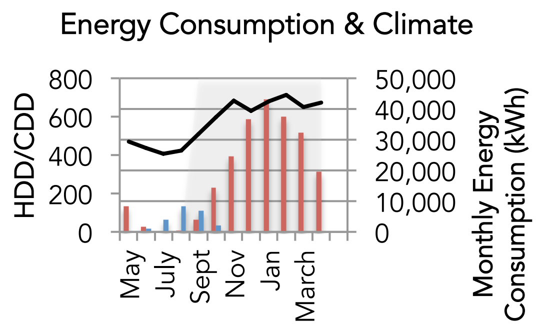


School of Interior Design

- 302 Church Street
- Built in the 1800s; acquired by Ryerson in the mid 1970s
- 28/235 (12%) data omitted
- 204/204 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 141.32 l 148.59

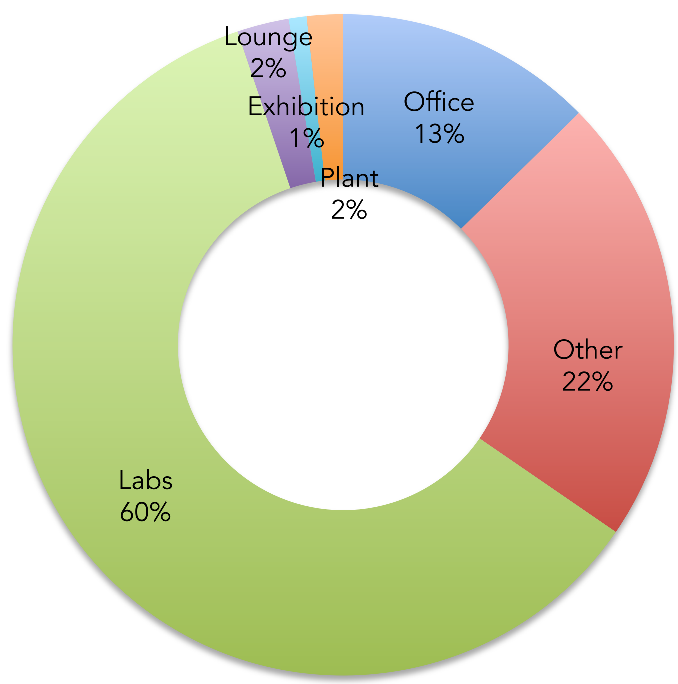
Energy Trends





Space Usage

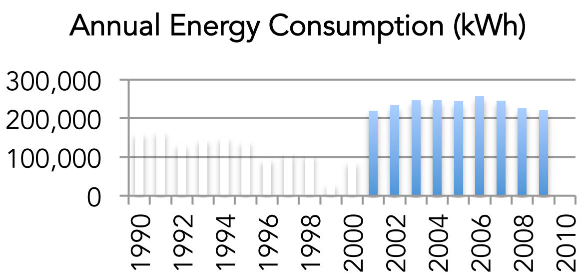
Usage	Area
Labs	1,844
Other	672
Office	386
Lounge	77
Plant	54
Exhibition	27
Total	3,060

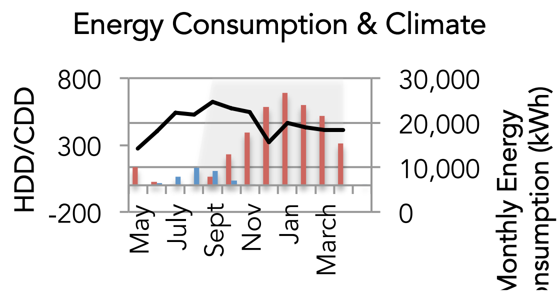


Theater School

- 46 Gerrard Street East
- Built in 1885
- Houses faculty offices
- No elevators
- 108/108 Data Integrity
- EUI [kWh/m^2/year] (avg|2009) = 104.40 | 96.75

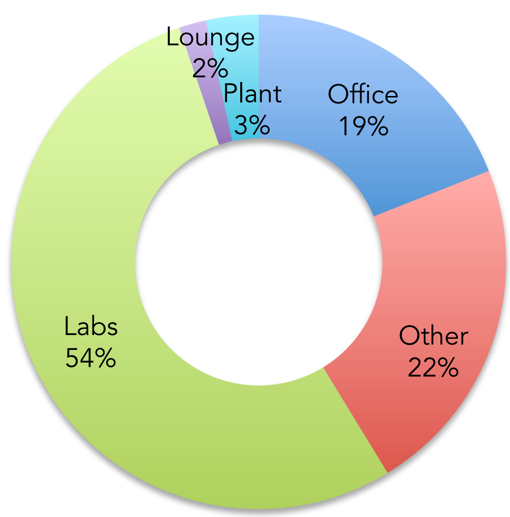
Energy Trends





Space Usage

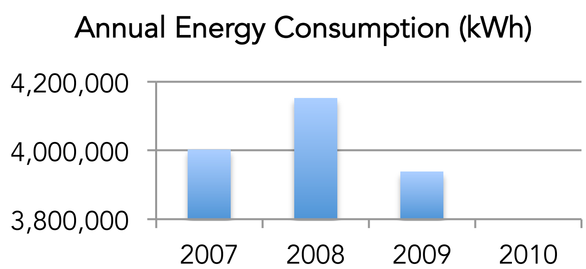
Usage	Area
Labs	1,216
Other	506
Office	432
Plant	78
Lounge	40
Total	2,272

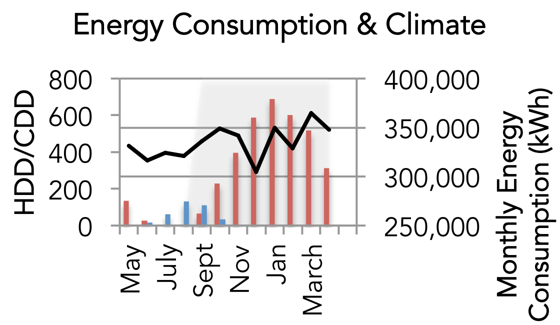


Ted Rogers School of Management

- 575 Bay Street
- Built in 2006
- 108/108 Data Integrity
- EUI [kWh/m^2/year] (avg|2009) = 232.29 | 227.05

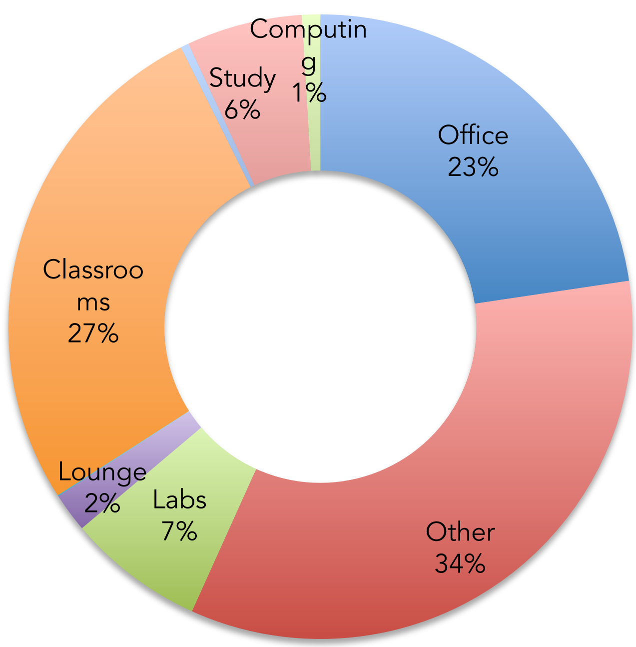
Energy Trends





Space Usage

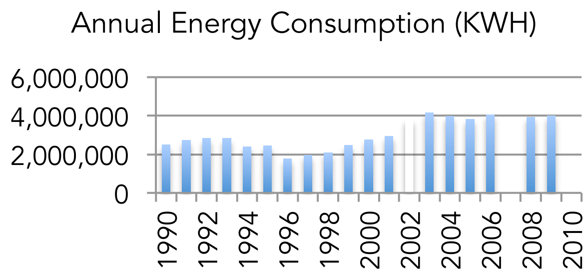
Usage	Area
Other	5,915
Classrooms	4,632
Office	3,929
Labs	1,229
Study	1,040
Lounge	352
Computing	165
Food	74
Plant	9
Total	17,345

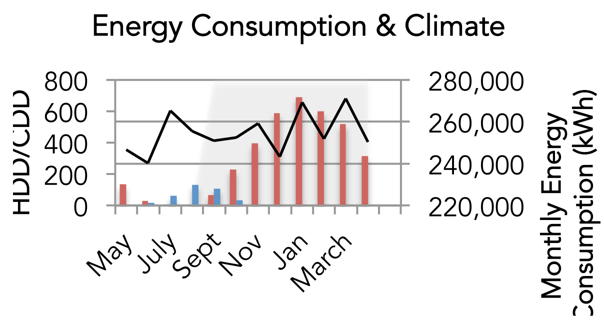


(1)Eric Palin Hall & (2)Sally Horsfall Eaton Centre for Studies in Community Health

- (1) 87 Gerrard Street East; (2) 99 Gerrard Street East
- (2) 2002 addition to EPH
- 215/216 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 168 l 225

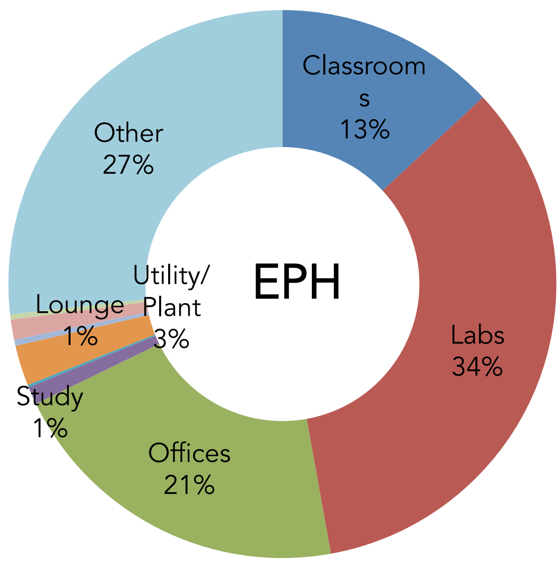
Energy Trends

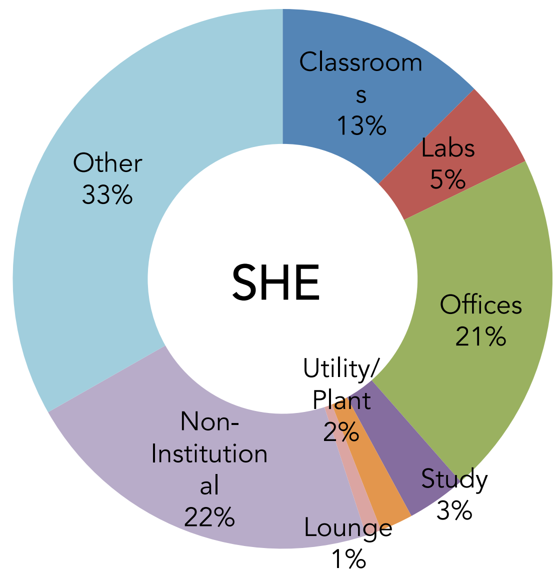


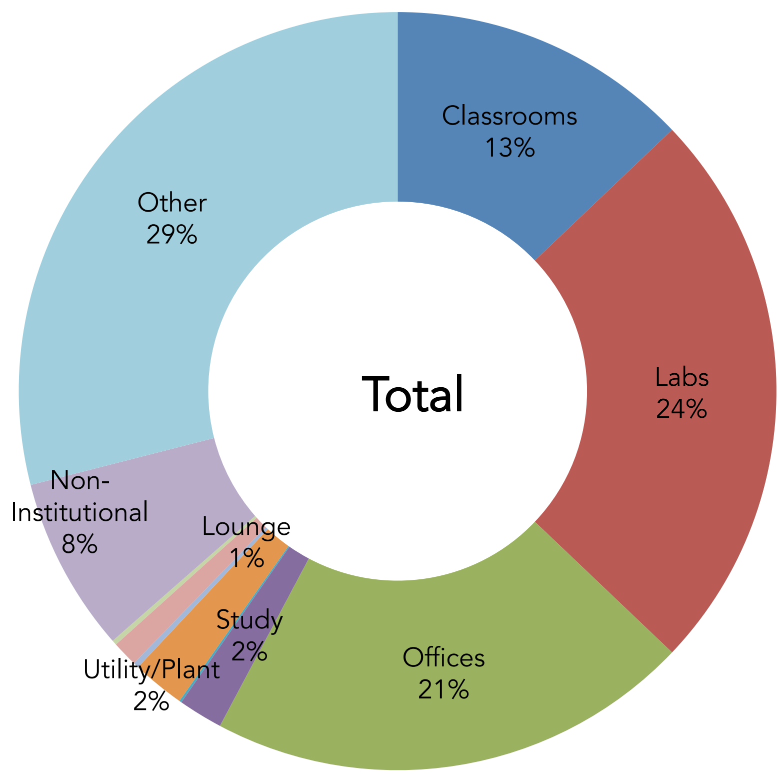


Space Usage

Usage	EPH	SHE	Area Total
Other	33,689	21,801	55,490
Labs	42,915	3,437	46,352
Offices	25,968	13,599	39,566
Classrooms	16,442	8,247	24,689
Non-Institutional	0	14,289	14,289
Utility/Plant	3,000	1,298	4,299
Study	1,318	2,327	3,645
Lounge	1,477	613	2,089
Central Computing	460	0	460
Inactive	395	0	395
Food	185	0	185
Total	125,849	65,611	191,459


EPH
Classroom
s
13%
Labs
34%
Offices
21%
Study
1%
Utility/
Plant
3%
Lounge
1%
Other
27%

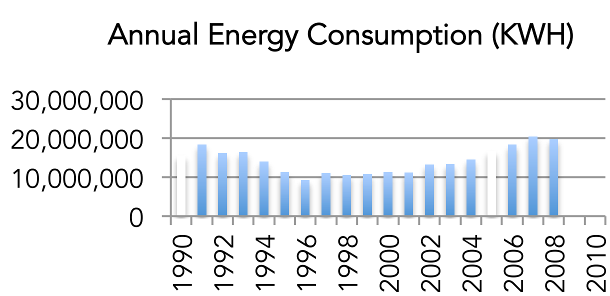

SHE
Classroom
s
13%
Labs
5%
Offices
21%
Study
3%
Utility/
Plant
2%
Lounge
1%
Non-
Institution
al
22%
Other
33%

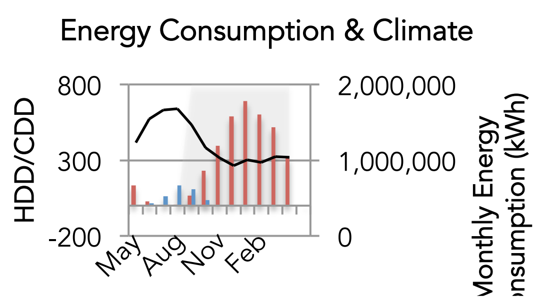

Total
Classrooms
13%
Labs
24%
Offices
21%
Study
2%
Utility/Plant
2%
Lounge
1%
Non-
Institutional
8%
Other
29%

(1)Library Building, (2)Jorgenson Hall, (3)Podium, & (4)Recreation and Athletics Center

- (1) 350 Victoria Street; (2) 380 Victoria Street; (3) 350 Victoria Street; (4) 40 Gould Street
- Built in (1) 1974; (2) 1971; (4) 1987 (2) home to the Faculty of Arts and administrative offices; (3) serves as an above walkway linking the Library Building and Jorgenson Hall. Includes a large cafeteria, and administrative offices for student services; (4) Facilities include: cardio room, fitness center, 6 gyms, 4 squash courts, 2 studios, a 2 yard six-lane pool, and change rooms with sauna. Walls are cast-in-place reinforced concrete
- Number of storeys: (1) 11; (2) 12; (3) 3; (4) 2 (underground)
- 214/216 Data Integrity
- EUI [kWh/m²/year] (avgl2008) = 278 l 389

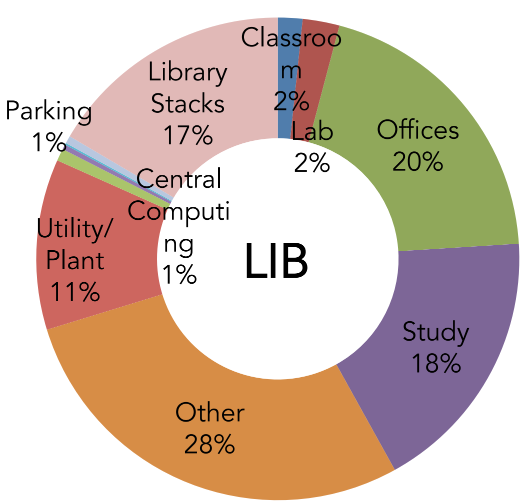
Energy Trends

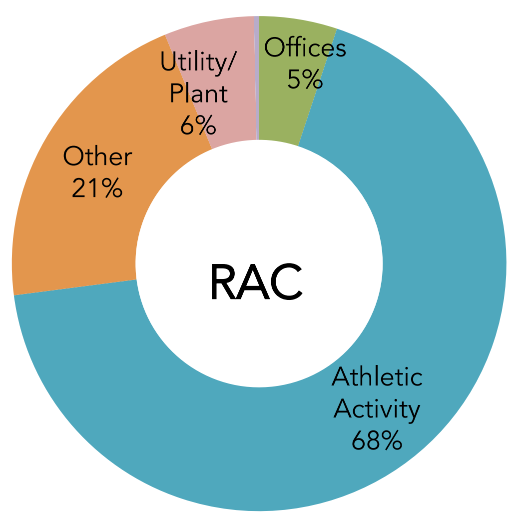


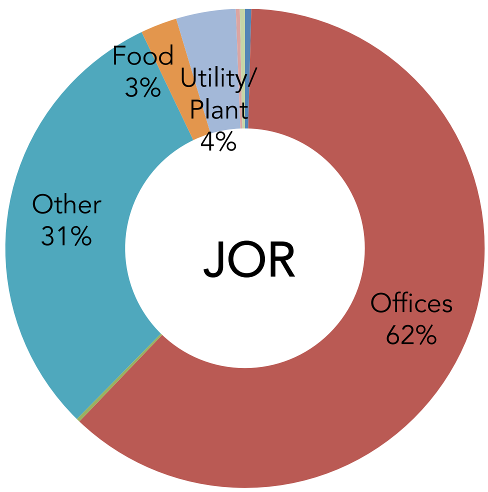


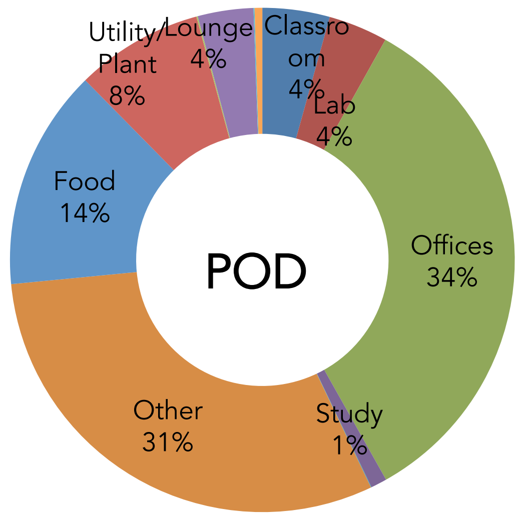
Space Usage

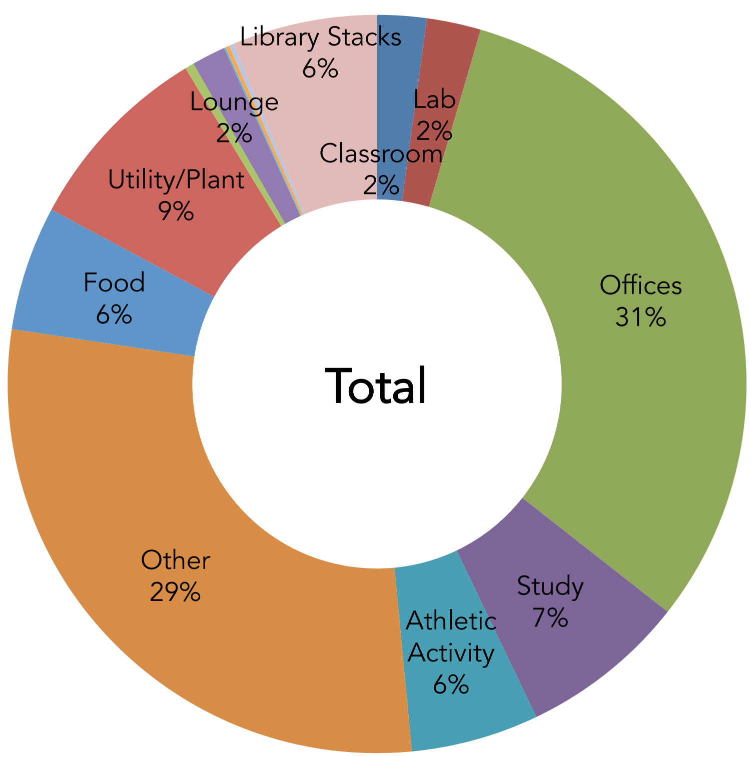
Usage	JOR	LIB	POD	RAC	Area Total
Offices	60,208	41,736	65,133	2,285	169,362
Other	29,778	59,415	58,978	9,333	157,504
Utility/Plant	3,925	23,993	15,769	2,637	46,323
Study	228	37,667	2,000	0	39,895
Library Stacks	0	34,680	0	0	34,680
Athletic Activity	0	0	74	30,403	30,477
Food	2,400	0	27,329	0	29,729
Lab	413	5,177	7,340	0	12,930
Classroom	0	3,458	8,313	0	11,771
Lounge	335	696	6,876	149	8,056
Central Computing	261	1,736	168	0	2,165
Parking	0	1,082	0	0	1,082
Ex-University Merchandising	0	0	952	0	952
Inactive	0	240	133	0	373
Total	97,548	209,880	193,065	44,807	545,299






JOR
Offices
62%
Other
31%
Food
3%
Utility/
Plant
4%

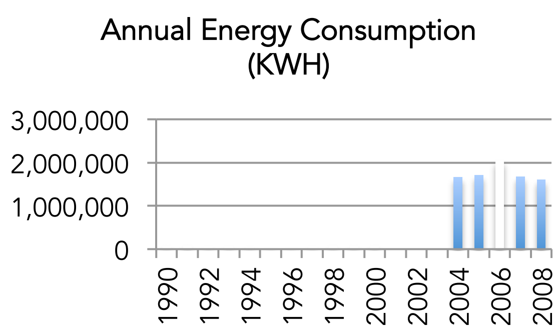

POD
Classro
om
4%
Lab
4%
Offices
34%
Study
1%
Other
31%
Food
14%
Utility/
Plant
8%
Lounge
4%

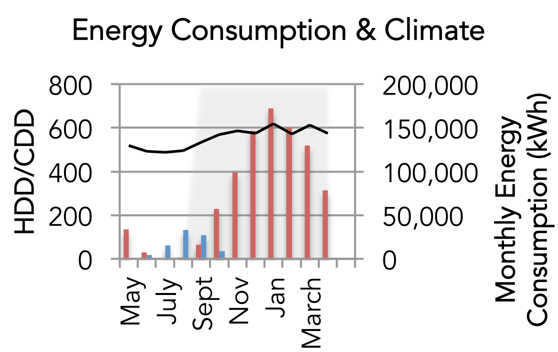

Total
Library Stacks
6%
Lab
2%
Classroom
2%
Offices
31%
Study
7%
Athletic
Activity
6%
Other
29%
Food
6%
Utility/Plant
9%
Lounge
2%

(1)Student Campus Center, (2)Oakham House, & (3)Heidelberg Centre-School of Graphics Communications Management

- Located at (1) 55 Gould Street; (2) 63 Gould Street; (3) 125 Bond Street
- Built in (1) 2005; (2) Built in 1848, acquired by Ryerson in 1958 (3) 2002
- (1) 3 storeys and houses meeting rooms for student organizations, a café and restaurant; (2) Used for non-academic, cultural and recreation activity; (3) 4 storeys containing computer and printing labs, lecture rooms, offices, and lounges
- 59/60 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 222 l 215

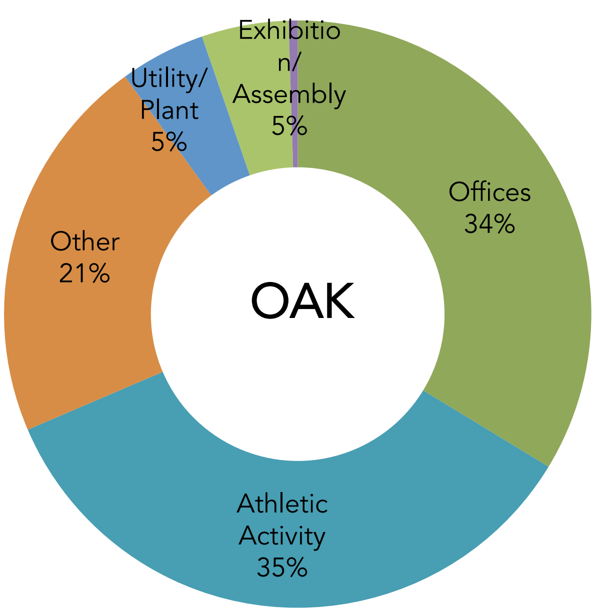
Energy Trends

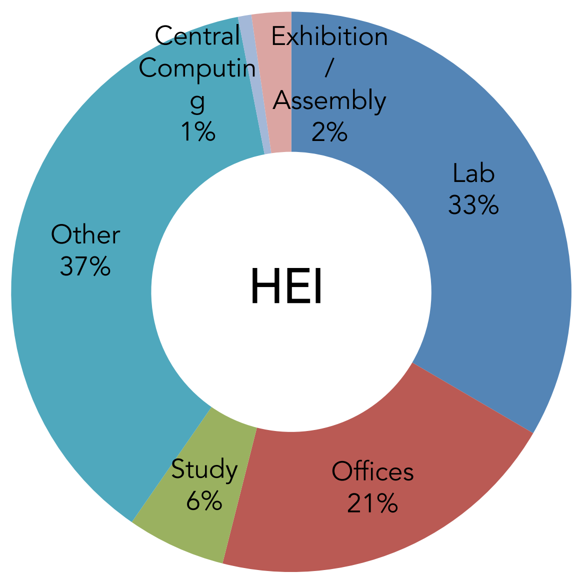


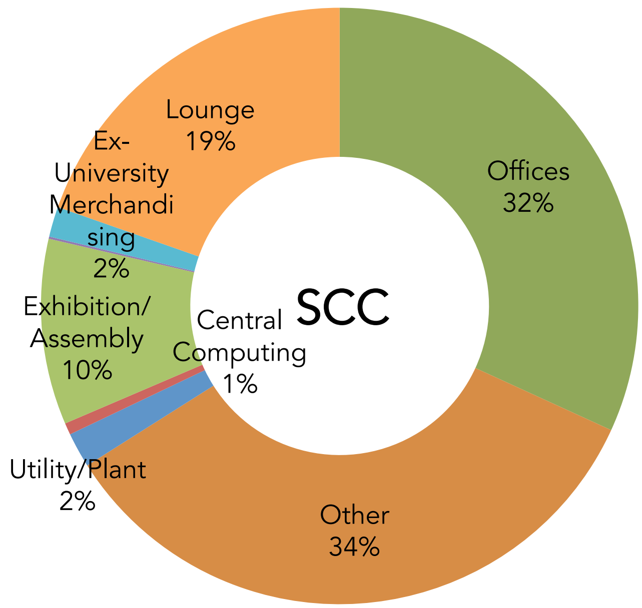


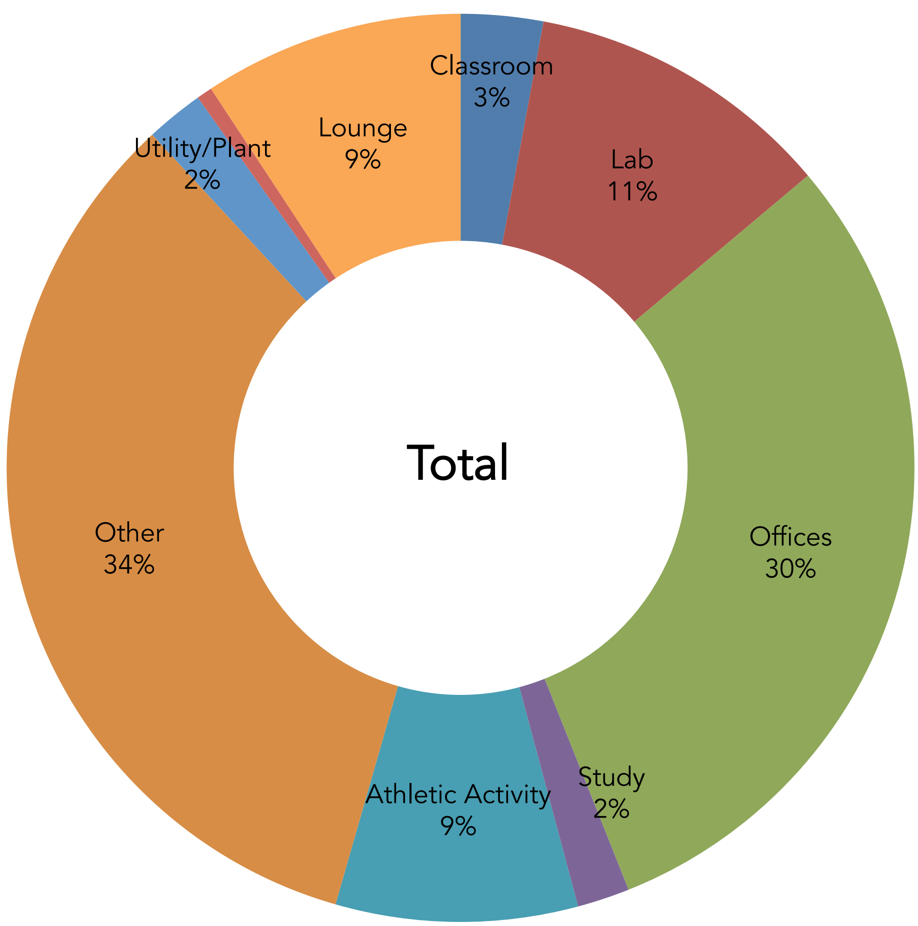
Space Usage

Usage	HEI	OAK	SCC	Area Total
Other	9,122	3,953	12,060	25,135
Offices	5,021	6,215	11,231	22,468
Lab	8,172	0	0	8,172
Lounge	0	0	6,922	6,922
Athletic Activity	0	6,430	0	6,430
Classroom	2,186	0	0	2,186
Utility/Plant	0	871	670	1,540
Study	1,393	0	0	1,393
Central Computing	190	0	225	415
Exhibition/Assembly	558	883	3,552	0
Food	0	92	42	0
Ex-University Merchandising	0	578	0	0
Total	26,642	19,022	34,702	74,661

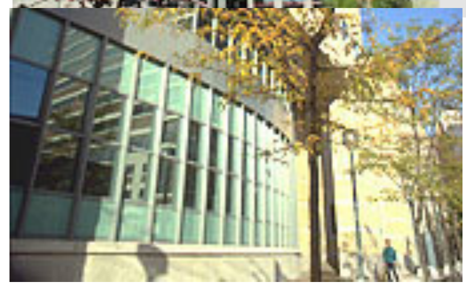





SCC
Offices
32%
Other
34%
Utility/Plant
2%
Central
Computing
1%
Exhibition/
Assembly
10%
Ex-
University
Merchandi
sing
2%
Lounge
19%

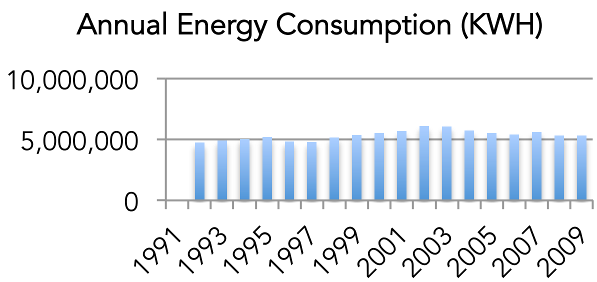

Total
Classroom
3%
Lab
11%
Offices
30%
Study
2%
Athletic Activity
9%
Other
34%
Utility/Plant
2%
Lounge
9%

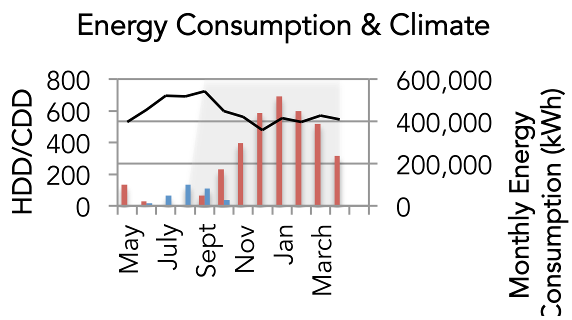
(1)Pitman Hall & (2)Rogers Communications Center



- Located at (1) 160 Mutual Street; (2) 80 Gould Street
- (1) Built in 1991; (2) first used in 1992
- (1) largest residence on campus with 14 floors and 565 rooms. The building contains a cafeteria, offices and study spaces in the lower floors; (2) 4 tv studios, 4 radio production suites
- 216/216 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 171 l 169

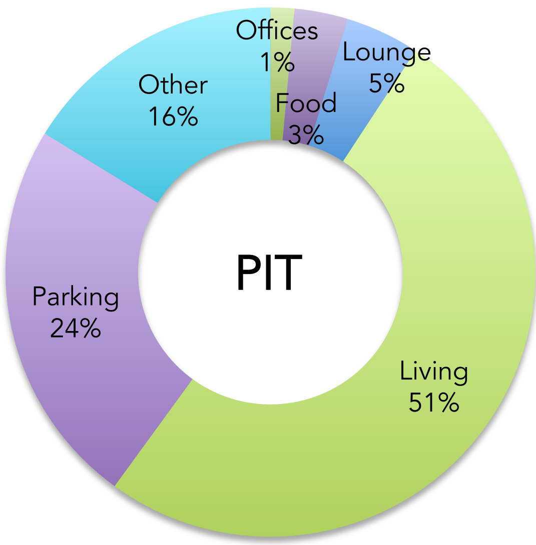
Energy Trends

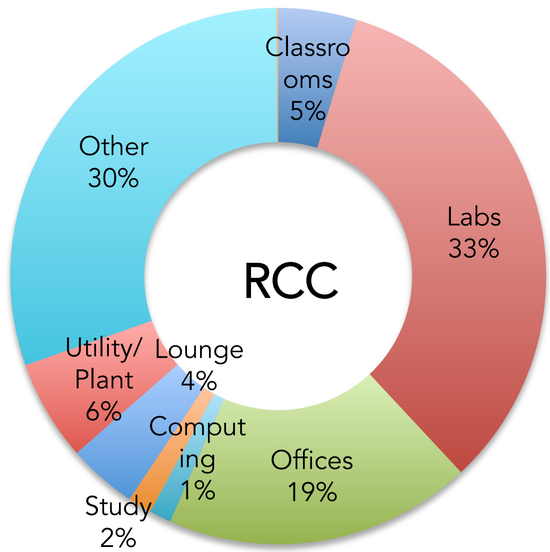


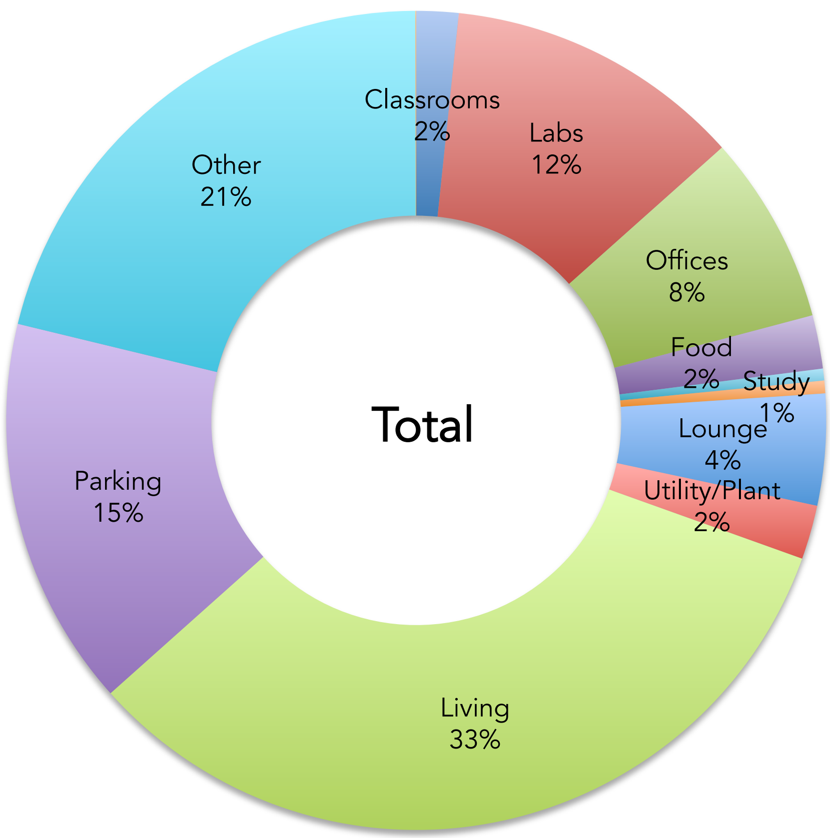


Space Usage

Usage	PIT	RCC	Area Total
Living	110,890	0	110,890
Other	35,431	35,941	71,372
Parking	51,881	0	51,881
Labs	0	39,550	39,550
Offices	3,170	22,014	25,184
Lounge	9,864	4,953	14,817
Utility/Plant	0	7,201	7,201
Food	7,111	0	7,111
Classrooms	0	5,595	5,595
Study	0	1,694	1,694
Computing	0	1,577	1,577
Inactive	0	112	112
Total	218,347	118,637	336,984


PIT
Offices
1%
Lounge
5%
Food
3%
Other
16%
Parking
24%
Living
51%

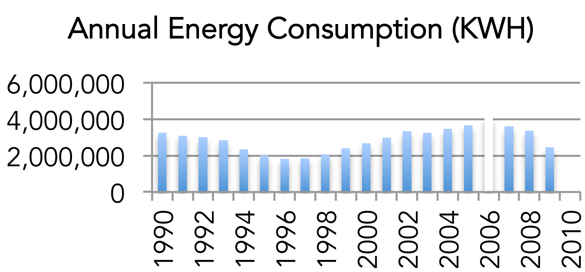

RCC
Classro
oms
5%
Labs
33%
Other
30%
Utility/
Plant
6%
Lounge
4%
Comput
ing
1%
Offices
19%
Study
2%

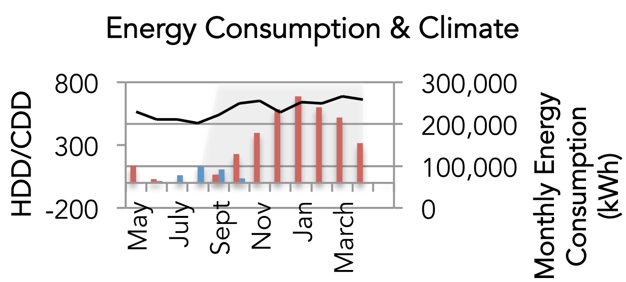

Total
Classrooms
2%
Labs
12%
Other
21%
Offices
8%
Food
2%
Study
1%
Lounge
4%
Utility/Plant
2%
Parking
15%
Living
33%

(1)Ryerson Image Center, (2)Victoria Building, & (3)Heaslip House

- Located at (1) 122 Bond Street; (2) 285 Victoria Street; (3) 297 Victoria Street
- (1) built in the 1960s, renovations from 2008 to 2012; (3) opened in 2005 incorporating historic façade of O'Keefe House and class and copper sheet panels
- (1) 3 floors; (2) 8 floors; (3) 7 floors
- 227/228 Data Integrity
- EUI [kWh/m^2/year] (avgl2009) = 127 l 111

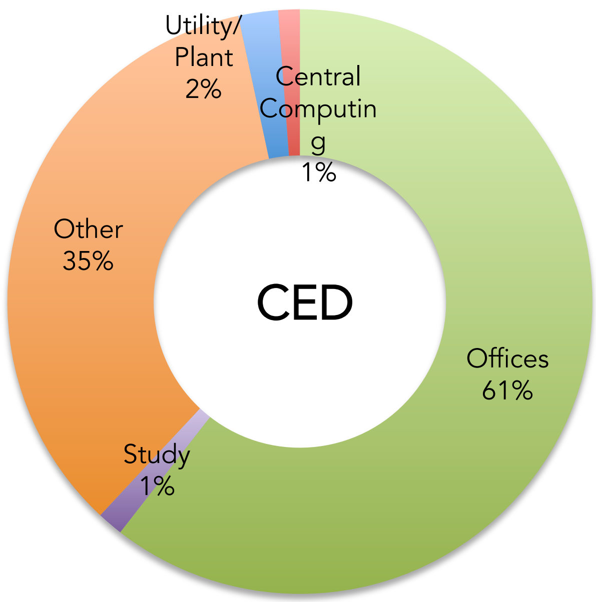
Energy Trends

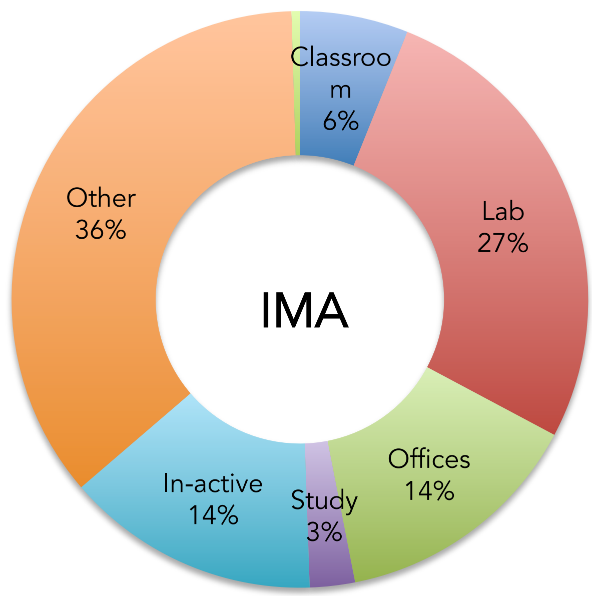


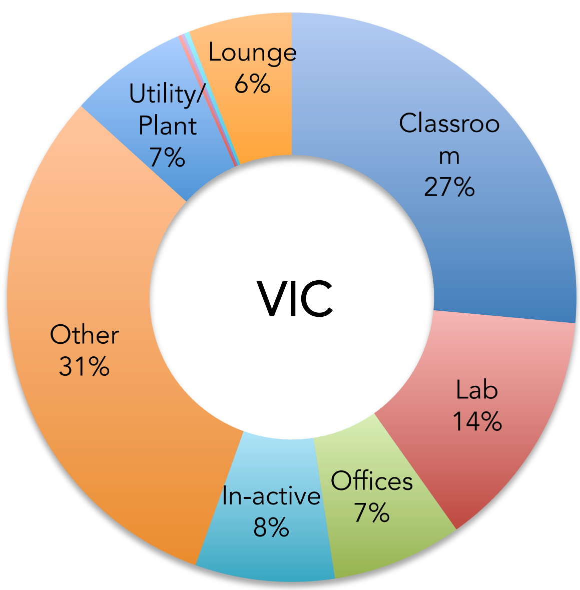


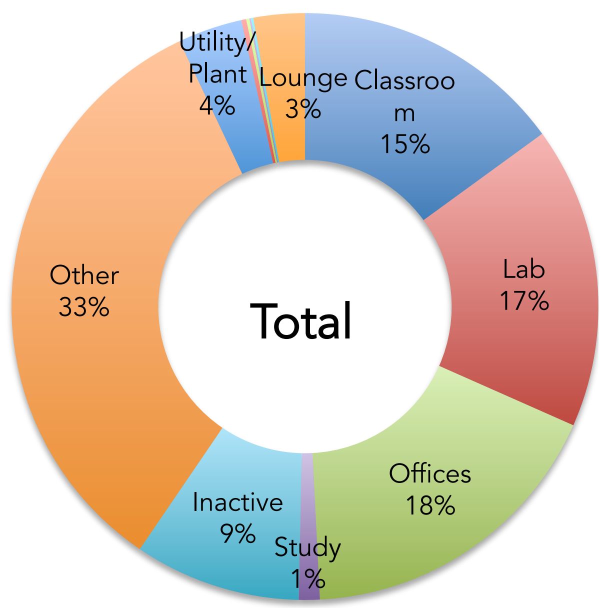
Space Usage

Usage	CED	IMA	VIC	Area Total
Other	11,866	32,214	35,738	79,818
Offices	20,717	12,713	8,450	41,880
Lab	0	24,011	15,706	39,717
Classroom	0	5,467	30,276	35,744
In-active	0	12,810	9,064	21,874
Utility/Plant	751	0	7,701	8,453
Lounge	0	0	6,725	6,725
Study	491	2,270	0	2,761
Central Computing	406	0	236	642
Exhibition/Assembly	0	418	0	418
Non-institutional Agency	0	0	331	331
Food	0	0	191	191
Total	34,231	89,903	114,418	238,554


CED
Utility/ Plant 2%
Central Computing 1%
Other 35%
Offices 61%
Study 1%


IMA
Classroom 6%
Other 36%
Lab 27%
In-active 14%
Study 3%
Offices 14%


VIC
Lounge 6%
Utility/ Plant 7%
Classroom 27%
Other 31%
Lab 14%
In-active 8%
Offices 7%


Total
Utility/ Plant 4%
Lounge 3%
Classroom 15%
Other 33%
Lab 17%
Offices 18%
Inactive 9%
Study 1%

APPENDIX A3 – BUDGET FOR SOURCING AND INSTALLING METERING EQUIPMENT FOR RYERSON BUILDINGS

Metering by Facility

Phase 1

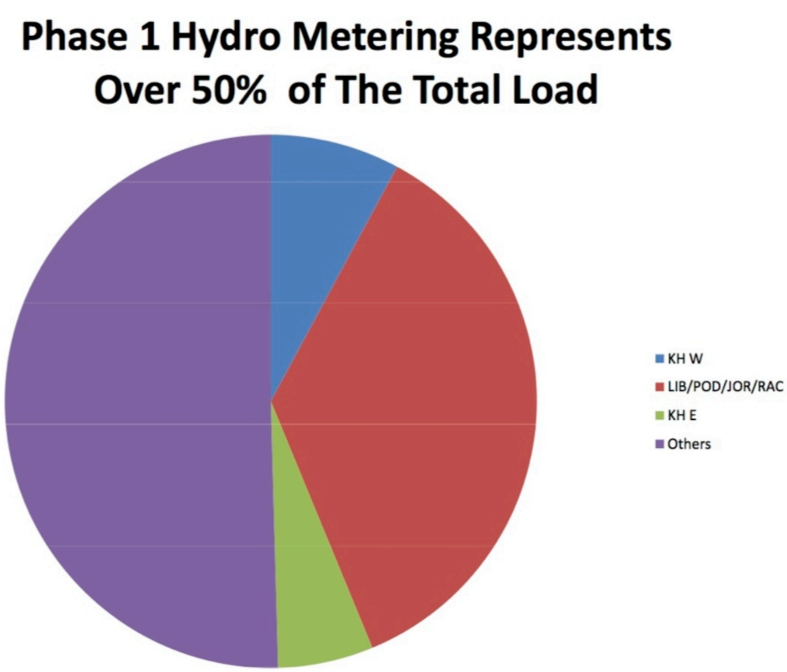
- Library (All)
- Podium (All)
- Jorgenson Hall (All)
- Kerr Hall (Hydro, Dom. Water, No steam)

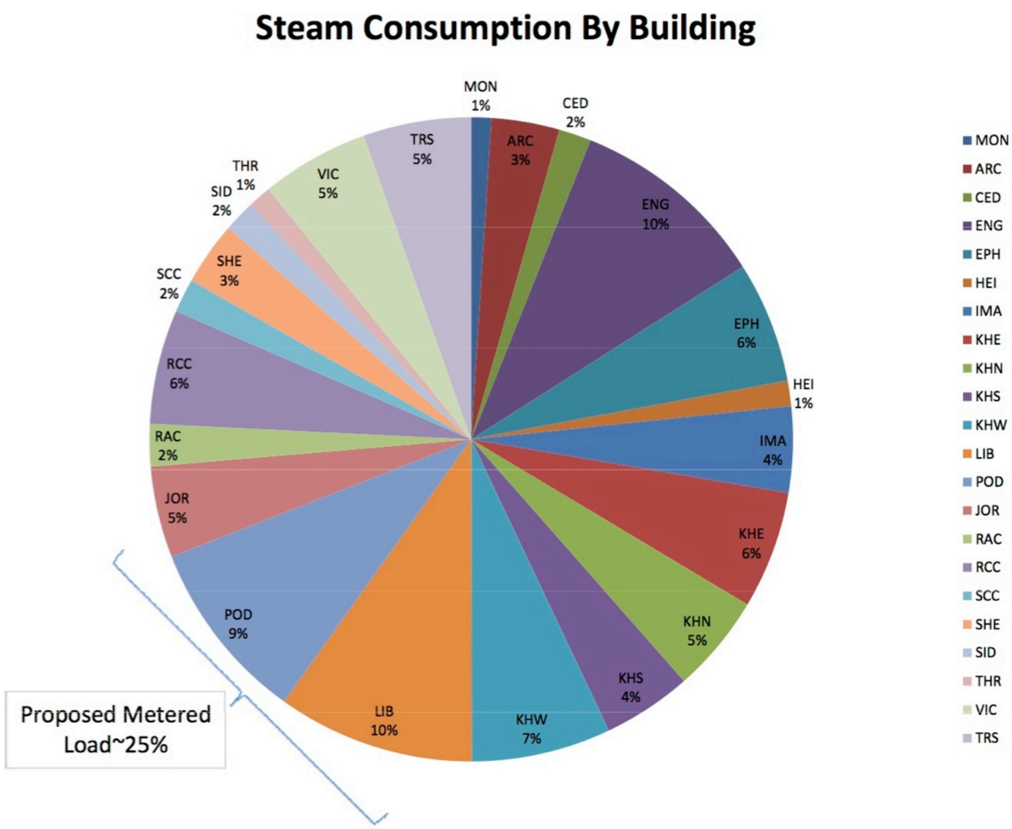
Phase 2

- ENG
- TRS
- IMA
- EPH/SHE
- RAC
- VIC
- PIT
- ILLC
- SCC
- ARC
- SBB
- MON
- SID
- THR

Phase 3

- GER
- COP
- OKF
- BKS
- 110 Bond St.
- CPF





Campus-Wide Budget Costs			
Enterprise Level System Software Package and Modules	$100,000	$150,000	$250,000
Field Level Software and Hardware	$30,860	$10,000	$40,860
Contingencies		$46,000	$46,000
Design and Energy Management Fees		$114,000	$114,000
Taxes (HST)			$58,612
Estimated Measure Cost			$509,472

Building Specific Budget Costs - Phase 1

Cost Component	Supply	Install	Total
Electrical Meters	$9,300	$18,600	$27,900
Steam Meters	$2,600	$12,580	$15,180
Chilled Water Meters	$6,000	$10,135	$16,135
Domestic Water Meters	$3,200	$12,165	$15,365
Communication Gateways	$7,200		$7,200
Junction Boxes	$705		$705
DPT Transmitters	$1,000	$2,200	$3,200
Building Display Screens	$1,750	$3,250	$5,000
Contingencies	$13,900		$13,900
Design and Energy Management Fees	$35,000		$35,000
Taxes (HST)			$18,146
Portion - Campus-Wide Budget Costs			$20,379
Estimated Measure Cost			$178,110

Phase 1 - KH, LIB/POD, JOR

Cost Component	Total
Electrical Meters (LIB/POD- 1 Meter; JOR - 1 meter; KH - 2 meters)	$111,600
Steam Meters (3x2)	$91,080
Chilled Water Meters (KHX2; LIBx1; JORx1)	$64,540
Domestic Water Meters (KHX4; LIBx1;JORX1)	$92,190
Communication Gateways (KHx1; LIBx1;JORX1)	$46,095
Junction Boxes (KHx1; LIBx1;JORX1)	$2,115
DPT Transmitters (KHx6;LIbx2; JORx2)	$32,000
Building Display Screens (KHx1; LIBX1; JORX1)	$15,000
Contingencies	$41,700
Design and Energy Management Fees	$105,000
Taxes (HST)	$78,172
Portion of Campus-Wide Budget Costs (KHx1;LIBx1; JORx1)	$61,137
Estimated Measure Cost	$740,628

APPENDIX B1 – 2010/2011 ESTATES MANAGEMENT RECORDS FOR UK HIGHER EDUCATION INSTITUTIONS

INSTITUTION	GROSS INTERNAL AREA TOTAL HEI (M2)	TOTAL HEI - ELECTRICITY (KWH)	EUI	AVERAGE EUI
ENGLAND				
ANGLIA RUSKIN UNIVERSITY	108851	11821783.93	108.6051936	
ASTON UNIVERSITY	108748	16874212	155.1680215	
BATH SPA UNIVERSITY	40023	3581771	89.49281663	
THE UNIVERSITY OF BATH	216297	27384265	126.6049229	
UNIVERSITY OF BEDFORDSHIRE	90267	10988525	121.7335793	
BIRKBECK COLLEGE(#3)	42411	8062380.2	190.1011577	
BIRMINGHAM CITY UNIVERSITY	164221	16639748	101.325336	
THE UNIVERSITY OF BIRMINGHAM	483090	66667503	138.0022418	
UNIVERSITY COLLEGE BIRMINGHAM	59795	6514335.6	108.944487	
BISHOP GROSSETESTE UNIVERSITY COLLEGE LINCOLN	18549.5	1210751.3	65.2713712	
THE UNIVERSITY OF BOLTON(#6)	46096.745	..		
THE ARTS UNIVERSITY COLLEGE AT BOURNEMOUTH	21452	2127817	99.18967928	
BOURNEMOUTH UNIVERSITY	100646	12419816.06	123.4009902	
THE UNIVERSITY OF BRADFORD	127332.26	14763659	115.9459433	
THE UNIVERSITY OF BRIGHTON	179478	12975436	72.29541225	
THE UNIVERSITY OF BRISTOL	429527	61390512	142.925851	
BRUNEL UNIVERSITY	231973.19	24389286	105.1383826	
BUCKINGHAMSHIRE NEW UNIVERSITY	54571.58	6694687	122.6771701	
THE UNIVERSITY OF CAMBRIDGE	632394	116137222	183.6469385	
THE INSTITUTE OF CANCER RESEARCH(#3)	29095	13516778	464.573913	
CANTERBURY CHRIST CHURCH UNIVERSITY	124140	11638679	93.7544627	
THE UNIVERSITY OF CENTRAL LANCASHIRE	172791.09	21132026	122.2981231	
CENTRAL SCHOOL OF SPEECH AND DRAMA(#3)	8792.4	857470	97.523998	
UNIVERSITY OF CHESTER	115065	9106327	79.14072046	
THE UNIVERSITY OF CHICHESTER	51374	3885624	75.63405614	
THE CITY UNIVERSITY	118990	13203791	110.9655517	
CONSERVATOIRE FOR DANCE AND DRAMA	32199	2968725	92.1992919	
COURTAULD INSTITUTE OF ART(#3)	9370	1356343	144.7537887	
COVENTRY UNIVERSITY	154229	16855303	109.2875075	
CRANFIELD UNIVERSITY	192676	17930867	93.06227553	

UNIVERSITY FOR THE CREATIVE ARTS	74090	7064329	95.34794169
UNIVERSITY OF CUMBRIA	102653.5	8059858.95	78.51518896
DE MONTFORT UNIVERSITY	157713	13764452	87.27531656
UNIVERSITY OF DERBY	126472	13129265	103.8116342
UNIVERSITY OF DURHAM	343416	39304324	114.4510564
THE UNIVERSITY OF EAST ANGLIA(#5)	230327	34543620	149.9764248
THE UNIVERSITY OF EAST LONDON	106863	6453468	60.39010696
EDGE HILL UNIVERSITY	87541	6980377	79.73837402
THE UNIVERSITY OF ESSEX(#5)	220056	21436860	97.41547606
THE UNIVERSITY OF EXETER	229563	31060937	135.304631
UNIVERSITY COLLEGE FALMOUTH	52727	4788769	90.8219508
UNIVERSITY OF GLOUCESTERSHIRE	79584	4815344	60.50643345
GOLDSMITHS COLLEGE(#3)	90288	6347814	70.30628655
THE UNIVERSITY OF GREENWICH	127272.2	15023843	118.0449698
GUILDHALL SCHOOL OF MUSIC AND DRAMA	22613.63	2256919	99.80348135
HARPER ADAMS UNIVERSITY COLLEGE	43749.93	3331064.41	76.13873691
UNIVERSITY OF HERTFORDSHIRE	219625	32986205	150.1933068
HEYTHROP COLLEGE(#3)	5517	682439	123.6974805
THE UNIVERSITY OF HUDDERSFIELD	112979	11788377	104.3413112
THE UNIVERSITY OF HULL	228297	22144562	96.99891808
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE	470186	111703001	237.5719417
INSTITUTE OF EDUCATION(#3)	397658	5281914	13.28255436
THE UNIVERSITY OF KEELE	160895.12	12404090	77.09425867
THE UNIVERSITY OF KENT	230443.04	19850435	86.14031042
KING'S COLLEGE LONDON(#3)	381515	66989062	175.5869677
KINGSTON UNIVERSITY	161944	19042970	117.5898459
THE UNIVERSITY OF LANCASTER	212841	27690653	130.1001828
LEEDS COLLEGE OF MUSIC	8530	876054	102.7026964
LEEDS METROPOLITAN UNIVERSITY	210310.07	23795799	113.1462654
THE UNIVERSITY OF LEEDS	598301	83741761	139.9659385
LEEDS TRINITY UNIVERSITY COLLEGE	36456	2455308	67.34990125
THE UNIVERSITY OF LEICESTER	288776.32	41191118	142.6402206
THE UNIVERSITY OF LINCOLN	101804.08	11040542	108.4489148
LIVERPOOL HOPE UNIVERSITY	78019	5914469	75.80805958
LIVERPOOL JOHN MOORES UNIVERSITY	164553.16	15512831	94.27245882
THE LIVERPOOL INSTITUTE FOR PERFORMING ARTS	11027	1319646	119.6740727
THE UNIVERSITY OF LIVERPOOL	437254	58270764	133.2652509

UNIVERSITY OF THE ARTS, LONDON	234987	12769609	54.34176784
LONDON BUSINESS SCHOOL(#3)	31530	5833213	185.0051697
UNIVERSITY OF LONDON (INSTITUTES AND ACTIVITIES)(#3)(#7)	165715	30802636	185.8771747
LONDON METROPOLITAN UNIVERSITY	151169	16861441	111.5403357
LONDON SOUTH BANK UNIVERSITY	125453	15329000	122.1891864
LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE(#3)	195344	20736531	106.1539182
LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE(#3)	24042.9	6728040	279.8347953
LOUGHBOROUGH UNIVERSITY	282919.06	27884101	98.55858068
THE MANCHESTER METROPOLITAN UNIVERSITY	264371	25104365	94.95884571
THE UNIVERSITY OF MANCHESTER	822246.07	107121755	130.2794369
MIDDLESEX UNIVERSITY	91287.72	11353213	124.3673629
THE UNIVERSITY OF NEWCASTLE-UPON-TYNE	460671	56624346	122.9171057
NEWMAN UNIVERSITY COLLEGE	17292	1316064	76.10825815
THE UNIVERSITY OF NORTHAMPTON	108457	9124451	84.12966429
THE UNIVERSITY OF NORTHUMBRIA AT NEWCASTLE	216364.57	25753582	119.0286469
NORWICH UNIVERSITY COLLEGE OF THE ARTS	15345	934144	60.876116
THE UNIVERSITY OF NOTTINGHAM	584256	78258000	133.9447092
THE NOTTINGHAM TRENT UNIVERSITY	199458	26366684.39	132.1916613
THE OPEN UNIVERSITY(#7)	152742	21961566	143.7821032
OXFORD BROOKES UNIVERSITY	195483	16221316	82.98069909
THE UNIVERSITY OF OXFORD	576913	113710120	197.100984
UNIVERSITY COLLEGE PLYMOUTH ST MARK AND ST JOHN	34036.6	3452783	101.4432405
THE UNIVERSITY OF PLYMOUTH	117374.46	15094407	128.6004383
THE UNIVERSITY OF PORTSMOUTH	208651.08	23049682	110.4699865
QUEEN MARY AND WESTFIELD COLLEGE(#3)	215500	36805304	170.7902738
RAVENSBOURNE	13926	2882453	206.9835559
THE UNIVERSITY OF READING	309421.14	30567293	98.78863804
ROEHAMPTON UNIVERSITY	95619.44	8489850	88.78790756
ROSE BRUFORD COLLEGE	8389	833088	99.30718798
ROYAL ACADEMY OF MUSIC(#3)	12665.99	1798768	142.0155866
ROYAL AGRICULTURAL COLLEGE	23681.8	2055648	86.80286127
ROYAL COLLEGE OF ART	24380	2535270	103.9897457

ROYAL COLLEGE OF MUSIC	17792	1541957.5	86.66577675	
ROYAL HOLLOWAY AND BEDFORD NEW COLLEGE(#3)	153099.41	15993719	104.4662354	
ROYAL NORTHERN COLLEGE OF MUSIC	17834.26	2365710	132.6497427	
THE ROYAL VETERINARY COLLEGE(#3)	59859	9789008	163.5344393	
ST GEORGE'S HOSPITAL MEDICAL SCHOOL(#3)	70944.9	13800619	194.5258785	
ST MARY'S UNIVERSITY COLLEGE, TWICKENHAM	48556	3272093	67.3880262	
THE UNIVERSITY OF SALFORD	207182.48	21036698.28	101.5370522	
THE SCHOOL OF ORIENTAL AND AFRICAN STUDIES(#3)	30491.78	3948422	129.491358	
THE SCHOOL OF PHARMACY(#3)	14763	2822604.14	191.1944822	
SHEFFIELD HALLAM UNIVERSITY	162270.41	19942663	122.8977175	
THE UNIVERSITY OF SHEFFIELD	457658.22	57594043	125.8450968	
SOUTHAMPTON SOLENT UNIVERSITY	121748	11294652	92.77073956	
THE UNIVERSITY OF SOUTHAMPTON	413140	34545824	83.61771796	
STAFFORDSHIRE UNIVERSITY	137387	17219028	125.3322949	
UNIVERSITY CAMPUS SUFFOLK(#5)	33456	2717189	81.2167922	
THE UNIVERSITY OF SUNDERLAND	135387.73	13115330	96.87236798	
THE UNIVERSITY OF SURREY	259718.89	26860991	103.4233243	
THE UNIVERSITY OF SUSSEX	206820	26691994	129.0590562	
TEESSIDE UNIVERSITY(#2)	133108.5	13660210	102.6246258	
TRINITY LABAN CONSERVATOIRE OF MUSIC AND DANCE	20640.6	2437773	118.1057237	
UNIVERSITY COLLEGE LONDON(#3)	454936	98738999	217.0393176	
THE UNIVERSITY OF WARWICK	456060.297	64138604.3	140.6362376	
UNIVERSITY OF THE WEST OF ENGLAND, BRISTOL	228400	20258303	88.69659807	
THE UNIVERSITY OF WEST LONDON(#2)	64254	3899440	60.68789492	
THE UNIVERSITY OF WESTMINSTER	163472	18022491.7	110.2481875	
THE UNIVERSITY OF WINCHESTER	58499	3797723	64.91945161	
THE UNIVERSITY OF WOLVERHAMPTON	169095	15567511	92.06369792	
THE UNIVERSITY OF WORCESTER	70781	5231233	73.90730563	
WRITTLE COLLEGE	40604	2349063	57.85299478	
YORK ST JOHN UNIVERSITY	69158.45	6575036	95.07205555	
THE UNIVERSITY OF YORK	313902	35700647	113.7318239	116.5395504
WALES				
ABERYSTWYTH UNIVERSITY	190242	18704151	98.31767433	
BANGOR UNIVERSITY	193572	15641339	80.80372678	

CARDIFF UNIVERSITY	417550	49971187	119.6771333	
CARDIFF METROPOLITAN UNIVERSITY(#2)	91982	7729787	84.03586571	
UNIVERSITY OF GLAMORGAN	133154	13735369	103.1540096	
GLYNDŴR UNIVERSITY	57463	6039023	105.0941127	
THE UNIVERSITY OF WALES, NEWPORT	55683.69	5473936	98.30411742	
SWANSEA METROPOLITAN UNIVERSITY	52958	3137686	59.24857434	
SWANSEA UNIVERSITY	199420.77	21116733.36	105.8903411	
UNIVERSITY OF WALES TRINITY SAINT DAVID(#1)(#2)	65255	3945634	60.46485327	91.49904086
SCOTLAND				
THE UNIVERSITY OF ABERDEEN	247287.263	22460027.63	90.82565498	
UNIVERSITY OF ABERTAY DUNDEE	52841.5	4772945	90.32569098	
THE UNIVERSITY OF DUNDEE	218563.3	29457938	134.7798921	
EDINBURGH COLLEGE OF ART	31496.294	2066886	65.62314919	
EDINBURGH NAPIER UNIVERSITY	101267	10932186	107.9540818	
THE UNIVERSITY OF EDINBURGH	726368	106256073	146.2840778	
GLASGOW CALEDONIAN UNIVERSITY	116319.6	12567478	108.0426515	
GLASGOW SCHOOL OF ART	40574	2270647	55.96310445	
THE UNIVERSITY OF GLASGOW	375149.01	64614906	172.2379755	
HERIOT-WATT UNIVERSITY	173681.56	20067796	115.543619	
QUEEN MARGARET UNIVERSITY, EDINBURGH	45729	4606175	100.7276564	
THE ROBERT GORDON UNIVERSITY	98618.5	10663982.5	108.133692	
THE UNIVERSITY OF ST ANDREWS	249857.08	28592961	114.4372655	
SCOTTISH AGRICULTURAL COLLEGE	87387.1	5856288	67.01547482	
THE UNIVERSITY OF STIRLING	147018.27	18787289	127.7888048	
THE UNIVERSITY OF STRATHCLYDE	337884	39070525	115.6329539	
THE UNIVERSITY OF THE WEST OF SCOTLAND	128550	9847154	76.60174251	105.7598522
NORTHERN IRELAND				
THE QUEEN'S UNIVERSITY OF BELFAST	320006	37069897	115.8412561	
ST MARY'S UNIVERSITY COLLEGE	14585	811397	55.63229345	
STRANMILLIS UNIVERSITY COLLEGE	35415.02	1346441	38.0189253	
UNIVERSITY OF ULSTER	243929.53	21811675	89.41793558	74.72760261
ALL OF UK				**112.783877**

APPENDIX B2 – EXPANDED DETAILS ON STUDIES FROM TABLE III

Author	Year (Published)	Location	Data Source	Sample size (real or simulated)		Method of Estimation			
				Training	Validation	Decision Tree	Multivariate Linear Regression	Artificial Neural Network	Other
Sharp, T. R.	1998	United States	CBECS 1992	21-83 (real)			x		
Sharp, T. R.	1996	United States	CBECS 1992	76-254 (real)			x		
Lam, J. C.	1997	Hong Kong	Survey/Energy Model	387 (simulated)	20 (Simulated)		x		
Signor, R.	2001	Brazil (14 cities)	Energy Model	1024 (simulated)	10 (simulated)		x		
Westphal, F. S.	2007	Brazil (3 cities)	Energy Model	23040 + 792 (simulated)			x		
Yu, Z.	2010	Japan (6 districts)	Reseach committee on Investigation on Energy Consumption of Residential Buildings and Architectural Institute of Japan Comprehensive survey	55 (real)	12	x (C4.5 algorithm)			
Tso, G. K. F.	2005	Hong Kong	physical and questionnaire-based survey	1166 (summer), 1000 (winter); real		x (Winter) x (Summer)	x (Winter) x (Summer)	x (Winter) x (Summer)	
Nóren, C.	1999	Sweden (South)	Survey	21 (real)	4		x		
Chung, W.	2006	Hong Kong		30 (real)					
Boonyatikarn, S.	1982	Michigan		50 (real)			x		
Lee, W.-S.	2008	Taiwan	Utility Bills	47 (real)			x		
Lee, S.	2012	United States	CBECS 2003	5215 (real)				x	
Stoy, C.	2009	Switzerland	Physical Survey	80	5		x		
Kajl, S.	2000	Montreal		700 (simulated)	52 (simulated)			x	
Azar, E.	2012	United States	2003 CBECS	520 (real) for 30 models					Sensitivity analysis from running 990 simulations

Author	Energy Definition	Model Fit/Measure of Error				Building Type	Number of Variables
		RMSE	r2	MAPE	Other		
Sharp, T. R.	Electricity		0.35-0.89			School Buildings larger than 1000 ft2	2 or 3
Sharp, T. R.	Electricity		0.74-0.88			Office buildings larger than 1000 ft2	2 or 3
Lam, J. C.	Electricity		0.988			Office Buildings - 40 storeys - Fully Air Conditioned with VAV system - 35 x 35 m footprint - 44% WWR	12
Signor, R.			> 0.986 for each city		avg < 2% difference between modelled and predicted values	Office buildings with air conditioning	9 (variables combined to result in linear relationship with consumption)
Westphal, F. S.	Electricity		0.973			(1) Typical commercial office building - 27 x 7.5 m footprint - 5 floors - 1001 m2 (874 m2 conditioned) (2) warehouse - 2500 m2 area - square footprint - 1 floor - fully air conditioned	17
Yu, Z.	Electricity, gas, liquefied petroleum gas, kerosene (kWh)		0.92			Residential buildings	10
Tso, G. K. F.	Electricity	44.397 (kWh)				Domestic households with an average monthly consumption greater than	6
		39.363 (kWh)					3
		45.184 (kWh)					4
		39.424 (kWh)					6
		44.142 (kWh)					5
		39.527 (kWh)					6
Nóren, C.	Electricity	10-20%				Chain Grocery Stores	2
Chung, W.	Electirity		0.7082			Grocery Stores	9
Boonyatikarn, S.	?		0.94			Institutional Buildings	10
Lee, W.-S.	Electricity *(kWh)		0.84			Government Office Buildings	4
Lee, S.	Electricity	5.41				Commercial	8
Stoy, C.	Electricity?		0.73	14%	7-17% (validation)	Offices	5
Kajl, S.	Electricity and Gas, separated	63060 (kwh electricity consumed by 52 simulated buildings); 20160 (kwh electricity consumed for cooling by 52 simulated buildings)			(CV) 1 for electrical energy and 4 for electrical cooling energy	Offices	11
Azar, E.	Total energy					Offices	9

Author	Site-Specific/Location			Lighting		Space Usage	
	Building Azimuth	Outdoor Air Temperature	Hours of Rain	Lighing Load	Lighting Control	% area cooled	Space Usage
Sharp, T. R.							
Sharp, T. R.							
Lam, J. C.				x			
Signor, R.				x			
Westphal, F. S.	x			x			
Yu, Z.		x					
Tso, G. K. F.							
Nóren, C.		x			x		
Chung, W.				x			
Boonyatikarn, S.		x (CDD)					
Lee, W.-S.		x	x				
Lee, S.				x		x	
Stoy, C.						x (and % ventillated)	x (+ data processing center)
Kajl, S.				x			
Azar, E.				x			

Author	Building & Equipment Ownership and Occupants					
	Specific Plug Loads	Non-electric Energy Use	HVAC Operation Responsibility	Occupant Density	Owner Occupancy	Operating Hourrs
Sharp, T. R.	x (walk-in coolers)	x	x			
Sharp, T. R.	x (Personal computers)			x	x	x
Lam, J. C.	x			x		
Signor, R.	x					
Westphal, F. S.	x			x		x
Yu, Z.		x (heat, hot water, and kitchen equipment)		x		
				x	x (electric water heater and rangehood)	
			x	x	x	
				x	x (electric water heater and rangehood)	
			x	x	x (electric water heater, ventilation fan, clothes dryer, rangehood)	
				x	x (electric water heater and rangehood, ventilation fan)	
Tso, G. K. F.			x	x	x (electric water heater, washing machine, clothes dryer, rangehood)	
Nóren, C.						
Chung, W.				x		x
Boonyatikarn, S.		x				x
Lee, W.-S.				x		
Lee, S.	x (computers)					x
Stoy, C.						
Kajl, S.				x		
Azar, E.	x	x (hot water)				x

Author	Physical Building Characteristics															
	Gross Floor Area	Year of Construction	Window Shading Coefficient	WWR	Roof Construction	projection factor for window overhangs	Solar Heat Gain Coefficient	façade construction	Air Infiltration	Frame construction	multi-unit or single unit (house)	opaque wall area	Number of Floors	Skylight Atriums	Number of Elevators	Building Footprint
Sharp, T. R.	x	x			x											
Sharp, T. R.	x															
Lam, J. C.			x	x												
Signor, R.				x	x (transmittance /absorptance)	x	x	x (transmittance /absorptance)								
Westphal, F. S.				x	x (thermal capactiy, transmittance, solar absorptance)	x	x	x (thermal capactiy, transmittance, solar absorptance)	x							
Yu, Z.					x (heat loss coefficient)			x (heat loss coefficient)	x	x	x					
Tso, G. K. F.																
Nóren, C.																
Chung, W.		x														
Boonyatikarn, S.			x									x				
Lee, W.-S.																
Lee, S.				x									x	x		
Stoy, C.															x	
Kajl, S.						x	x	x					x			x
Azar, E.																

Author	HVAC Equipment									
	Boiler Efficiency	HVAC/AHU Technology	outdoor airflow	fan efficiency	Chilled Water supply water temp	Chiller COP	Cooling Thermostat Setpoint	Interior Setpoint	Presence of economizer/chiller	Cooling System
Sharp, T. R.										x
Sharp, T. R.									x	
Lam, J. C.			x	x	x	x	x	x		
Signor, R.										
Westphal, F. S.						x				
Yu, Z.										
Tso, G. K. F.										
Nóren, C.										
Chung, W.						x		x		
Boonyatikarn, S.		x	x	x						x
Lee, W.-S.										
Lee, S.										
Stoy, C.										
Kajl, S.	x		x							
Azar, E.							x	x		

APPENDIX C1 – EXAMPLE OF UTILITY DATA PROVIDED BY RYERSON UNIVERSITY

KW

GERRARD ST.

Fiscal Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
2010	31.00	29.00	39.00	30.00	40.00	26.00	32.00					
2009	26.00	26.00	31.00	29.00	28.00	28.00	24.00	36.00	26.00	31.00	36.00	25.00
2008	26.00	31.00	34.00	32.00	35.00	19.00	30.00	34.00	38.00	29.00	31.00	25.00
2007	17.00	29.00	36.00	38.00	36.00	32.00	32.00	33.00	38.00	32.00	25.00	26.00
2006	18.00											
2005	24.00	31.00	36.00	36.00	24.00	22.00	19.00	22.00	25.00	22.00	24.00	19.00
2004	29.00	19.00	19.00	18.00	19.00	24.00	14.00	24.00	26.00	12.00	31.00	24.00
2003	29.00	24.00	24.00	26.00	24.00	26.00	24.00	31.00	31.00	29.00	28.00	29.00
2002	25.00	24.00	31.00	34.00	24.00	22.00	29.00	36.00	38.00	31.00	29.00	24.00
2001	19.00	22.00	26.00	21.00	28.00	17.00	14.00	17.00	19.00	19.00	24.00	24.00
2000	26.00	24.00	22.00	24.00	21.00	28.00	18.00	16.00	29.00	26.00	31.00	29.00
1999	19.00	31.00	26.00	29.00	24.00	19.00	19.00	24.00	37.00	24.00	26.00	24.00
1998	17.00	26.00	26.00	29.00	29.00	29.00	29.00	22.00	24.00	17.00	24.00	24.00
1997	17.00	26.00	26.00	26.00	29.00	19.00	26.00	14.00	14.00	17.00	24.00	24.00
1996	26.00	31.00	36.00	26.00	29.00	19.00	19.00	38.00	34.00	34.00	24.00	17.00
1995	31.00	36.00	36.00	41.00	26.00	24.00	24.00	38.00	34.00	34.00	31.00	26.00
1994	38.00	38.00	38.00	58.00	46.00	46.00	38.00	38.00	48.00	43.00	36.00	36.00
1993	34.00	34.00	34.00	50.00	55.00	48.00	48.00	48.00	48.00	48.00	58.00	38.00
1992	36.00	48.00	48.00	53.00	48.00	36.00	36.00	38.00	43.00	55.00	43.00	43.00
1991	43.00	46.00	48.00	47.00	43.00	38.00	38.00	38.00	38.00	34.00	36.00	36.00
1990		50.00	48.00	38.00	41.00	41.00	41.00	41.00	46.00	46.00	41.00	38.00

kWh

101 GERRARD ST.

Fiscal Year	Total Of KWH	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
2010	56,538.00	5,520.00	5,040.00	9,840.00	7,380.00	12,680.00	7,860.00	8,218.00					
2009	92,280.00	7,320.00	7,200.00	7,200.00	7,440.00	6,480.00	7,440.00	7,440.00	9,600.00	7,560.00	9,480.00	7,080.00	8,040.00
2008	102,600.00	7,320.00	7,200.00	9,000.00	8,280.00	8,160.00	9,120.00	7,800.00	10,560.00	10,680.00	9,000.00	7,440.00	8,040.00
2007	99,493.00	6,360.00	7,080.00	6,720.00	10,320.00	6,600.00	6,000.00	8,280.00	9,720.00	11,580.00	11,704.00	9,089.00	6,040.00
2006	6,120.00	6,120.00											
2005	116,100.00	9,600.00	8,420.00	12,000.00	9,600.00	7,280.00	7,200.00	9,600.00	11,400.00	9,120.00	17,000.00	7,680.00	7,200.00
2004	106,472.00	11,206.00	6,848.00	8,093.00	7,471.00	6,226.00	6,226.00	7,346.00	7,595.00	12,600.00	9,131.00	11,528.00	12,202.00
2003	133,228.00	13,696.00	8,093.00	11,829.00	9,961.00	9,587.00	9,463.00	9,587.00	13,447.00	12,700.00	11,206.00	11,829.00	11,829.00
2002	144,580.00	9,360.00	9,712.00	10,459.00	11,331.00	9,587.00	7,720.00	12,576.00	13,323.00	16,311.00	16,809.00	14,692.00	12,700.00
2001	100,800.00	9,360.00	11,160.00	3,120.00	9,960.00	6,480.00	6,360.00	6,360.00	6,480.00	9,000.00	10,680.00	11,880.00	9,960.00
2000	115,080.00	8,040.00	8,760.00	7,560.00	7,080.00	6,840.00	6,480.00	7,560.00	7,080.00	15,360.00	12,720.00	15,120.00	12,480.00
1999	99,960.00	5,760.00	7,920.00	10,080.00	12,480.00	120.00	6,240.00	6,840.00	8,640.00	8,040.00	8,040.00	14,520.00	11,280.00
1998	135,360.00	4,800.00	13,560.00	10,080.00	14,880.00	14,880.00	14,880.00	12,480.00	11,880.00	9,720.00	6,480.00	6,000.00	15,720.00
1997	92,520.00	4,800.00	10,080.00	10,080.00	6,240.00	6,120.00	3,480.00	18,720.00	5,640.00	5,640.00	6,480.00	6,000.00	9,240.00
1996	104,280.00	9,720.00	12,120.00	8,040.00	12,360.00	6,120.00	3,480.00	4,920.00	9,120.00	9,480.00	14,040.00	10,080.00	4,800.00
1995	106,320.00	5,640.00	8,040.00	8,040.00	13,920.00	7,560.00	6,000.00	6,600.00	9,120.00	9,480.00	14,040.00	8,160.00	9,720.00
1994	153,240.00	9,480.00	9,480.00	9,480.00	35,160.00	13,200.00	10,320.00	9,120.00	9,120.00	17,040.00	11,880.00	9,360.00	9,600.00
1993	155,040.00	7,200.00	7,200.00	7,200.00	24,960.00	14,760.00	10,800.00	10,800.00	10,680.00	10,680.00	10,680.00	21,240.00	18,840.00
1992	129,960.00	8,280.00	9,600.00	9,600.00	11,280.00	11,280.00	8,400.00	8,400.00	14,520.00	14,640.00	16,200.00	10,320.00	7,440.00
1991	133,200.00	8,280.00	9,840.00	11,880.00	11,160.00	11,880.00	7,320.00	7,320.00	7,320.00	7,320.00	34,320.00	8,280.00	8,280.00
1990	136,440.00		14,520.00	11,880.00	11,520.00	8,400.00	6,480.00	6,480.00	11,760.00	16,080.00	16,080.00	22,080.00	11,160.00

APPENDIX C2 – EXAMPLE OF OUTLIER IDENTIFICATION IDENTIFIED WITHIN UTILITY DATA

Example of data rejection for campus buildings based on renovation work/space closures and outlier months (1.5xSD). Red cells indicate omitted data points and orange, missing data due to inactivity.

Cooperative Education (101 GERRARD ST. E.)

Fiscal Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	Annual Energy Consumption (KWH)
1990		14,520	11,880	11,520	8,400	6,480	6,480	11,760	16,080	16,080	22,080	11,160	
1991	8,280	9,840	11,880	11,160	11,880	7,320	7,320	7,320	7,320	34,320	8,280	8,280	
1992	8,280	9,600	9,600	11,280	11,280	8,400	8,400	14,520	14,640	16,200	10,320	7,440	129,960
1993	7,200	7,200	7,200	24,960	14,760	10,800	10,800	10,680	10,680	10,680	21,240	18,840	
1994	9,480	9,480	9,480	35,160	13,200	10,320	9,120	9,120	17,040	11,880	9,360	9,600	
1995	5,640	8,040	8,040	13,920	7,560	6,000	6,600	9,120	9,480	14,040	8,160	9,720	106,320
1996	9,720	12,120	8,040	12,360	6,120	3,480	4,920	9,120	9,480	14,040	10,080	4,800	104,280
1997	4,800	10,080	10,080	6,240	6,120	3,480	18,720	5,640	5,640	6,480	6,000	9,240	
1998	4,800	13,560	10,080	14,880	14,880	14,880	12,480	11,880	9,720	6,480	6,000	15,720	
1999	5,760	7,920	10,080	12,480	120	6,240	6,840	8,640	8,040	8,040	14,520	11,280	99,960
2000	8,040	8,760	7,560	7,080	6,840	6,480	7,560	7,080	15,360	12,720	15,120	12,480	115,080
2001	9,360	11,160	3,120	9,960	6,480	6,360	6,360	6,480	9,000	10,680	11,880	9,960	
2002	9,360	9,712	10,459	11,331	9,587	7,720	12,576	13,323	16,311	16,809	14,692	12,700	144,580
2003	13,696	8,093	11,829	9,961	9,587	9,463	9,587	13,447	12,700	11,206	11,829	11,829	133,227
2004	11,206	6,848	8,093	7,471	6,226	6,226	7,346	7,595	12,600	9,131	11,528	12,202	106,472
2005	9,600	8,420	12,000	9,600	7,280	7,200	9,600	11,400	9,120	17,000	7,680	7,200	116,100
2006	6,120												
2007	6,360	7,080	6,720	10,320	6,600	6,000	8,280	9,720	11,580	11,704	9,089	6,040	99,493
2008	7,320	7,200	9,000	8,280	8,160	9,120	7,800	10,560	10,680	9,000	7,440	8,040	102,600
2009	7,320	7,200	7,200	7,440	6,480	7,440	7,440	9,600	7,560	9,480	7,080	8,040	92,280
2010	5,728	5,230	10,210	7,657	12,949	8,156	8,218	9,401	10,646	9,774	9,463	6,288	103,720
2011	5,852	6,288	8,031	8,040	5,977	7,284	8,218	6,724	8,156	9,774	7,222	6,910	88,476
2012	6,412	6,724	8,467	7,346	6,599	6,226	8,591	9,338	13,136	8,342	9,899	6,226	97,306
Average Monthly	**7,742**	**8,598**	**9,330**	**9,916**	**8,504**	**7,152**	**8,311**	**9,658**	**11,135**	**11,407**	**9,782**	**9,727**	**109,324**

Dec. 1978 (cont'd)

Highlights:

- Ten input measures were reduced to five: FTE undergraduates, FTE graduates, FTE faculty, Lab contact hours and Library volumes.
- All ***fourteen*** universities were included in the survey rather than just the original five.
- Separate sub-categories for *Instructional Office* were introduced.
- Variable space factors for *Laboratories* were adopted to reflect the needs of different disciplines.

July 1981 **Published: Inventory of Physical Facilities of Ontario Universities. 1980-81 (#81-7)**

July 1984 **Published: Inventory of Physical Facilities of Ontario Universities. 1983-84 (#84-5)**

Dec. 1984 Task Force to Review COU Space Standards was struck.

1986-1987 The COU space formula began to be used directly for the allocation of provincial funds. The Ontario Ministry of Training, Colleges and Universities (MTCU) now uses space entitlement as generated by the formula to determine the allocation of repair and renovation funds under the Facilities Renewal Program.

Nov. 1987 The COU Standing Committee on Space Standards and Reporting was established. Future consideration was to be given to the question of eligibility of federated and affiliated institutions for capital funding.

Jan. 1988 **Published: Building Blocks: Volume VII: The Final Report of the Task Force to Review COU Space Standards.**

1986-87 triennial survey published (#88-1).

Jan. 1988 (cont'd)

Highlights:

- Space factors were converted to metric units.
- Five input measures were increased to eight: added FTE non-faculty research appointments; FTE non-academic office staff; and total assignable space.
- Instructional labs: some changes were made to the discipline groups.
- Research: non-faculty researchers weighted at 50%; discipline group changes.
- Office: non-academic office staff included as a new input measure.
- Plant Maintenance appears as a separate category.
- Changes to formula for Library and Athletics.

1988 The Committee on Space Standards and Reporting was created. It replaced the Task Force on COU Space Standards.

Feb. 1990 A Background Paper Prepared for the Joint Working Group: **University Space Standards in North America.**

Feb. 1990 Report of the Joint Working Group on University Space Utilization: **Classroom Space Utilization in Ontario Universities,1988-89.**

The review of classroom space and its utilization was carried out to satisfy a concern expressed by the "Ontario Ministry" that classrooms were underutilised and that the COU space guidelines were too generous. Data from a sample of Ontario Universities showed that the universities were meeting or exceeding the hours of use per week and percentage of seats occupied in the classroom space factor, and the average area for the different furniture types currently in use was also consistent with the assumption in the formula. The survey also showed that extensive use during the evening was occurring in many institutions.

Jan. 1991 **Published: Inventory of Physical Facilities of Ontario Universities, 1989-90 (#91-1)**

March 1991 Federated and affiliated Colleges to be included in the triennial survey.

1992	**Published: COU Building Blocks-Users Guide (first edition)**
Dec. 1993	**Published: Inventory of Physical Facilities of Ontario Universities, 1992-93 (#93-5)**
July 1997	**Published: Inventory of Physical Facilities of Ontario Universities, 1995-96 (#330)**
1997	The Committee reviewed the library volume equivalencies used in the generation of stack space entitlement. With the assistance of OCUL, the Committee adopted the methodology used by Canadian Association of Research Libraries (CARL).
1999	Provided background information and scenarios for PriceWaterhouseCoopers to examine the impact of the anticipated increase of students on the existing physical facilities of Ontario universities.
1999	A Joint Working Group was established consisting of representatives from OAPPA (physical plant administrators), CSAO (senior administrative officers) and the Committee on Space Standards to develop a facilities condition assessment program.
2000	Provided input to the Investing in Students Task Force. The Task Force was established to advise the Minister on ways to ensure that public funds supporting postsecondary education are directed at providing the highest quality of education while ensuring access for students, affordability and accountability. The Inventory report was listed as an example of collaborative practices in facilities, maintenance and equipment.
April 2000	**Published: Inventory of Physical Facilities of Ontario Universities, 1998-99 (#670)** Commencing with the 1998-99 report, the FTE enrolment figures include ineligible students (for the purpose of provincial funding) for both graduate and undergraduate students.
October 2000	The Committee conducted a survey to determine the extent to which the current stock of classrooms had the capability of handling year 2000 technology. The Committee was concerned that universities may not be equipped to handle the increase of students.

January 2003 **Published: Inventory of Physical Facilities of Ontario Universities, 2001-02 (#734)**

Dec 2005 **Published: Inventory of Physical Facilities of Ontario Universities, 2004-05 (#779)**

2007 The Committee on Space Standards and Reporting worked with representatives from OCUL (Ontario Council of University Libraries) to review and update library stack, study and support space standards.

Feb 2010 **Published: Inventory of Physical Facilities of Ontario Universities, 2007-08 (#816)**

APPENDIX C4 – COMPLETE LIST OF AREAS DEFINED UNDER COU SPACE CATEGORIES FROM BOTH UNIVERSITIES

COU SPACE CATEGORY	AREA (M2)	RELATIVE SHARE OF TOTAL
OTHER NON-ASSIGNABLE AREA	389421.05	33%
RESEARCH LAB SPACE	80970.08	7%
ACADEMIC OFFICES	58624.81	5%
RESIDENCE LIVING SPACE	56056.83	5%
SCHEDULED CLASS LAB	49057.67	4%
OFFICE SUPPORT SPACE	39989.56	3%
PARKING STRUCTURES	36499.98	3%
LIBRARY COLLECTION SPACE	36418.98	3%
NON-TIERED CLASSROOMS	35497.32	3%
DEPARTMENTAL SUPPORT STAFF OFFICE	32404.96	3%
TIERED CLASSROOMS	28468.89	2%
RESEARCH LAB SUPPORT SPACE	27984.92	2%
STUDY SPACE UNDER LIBRARY JURISDICTION	26093.75	2%
CENTRAL ADMINISTRATIVE OFFICES	24174.09	2%
GRADUATE STUDENT OFFICE	23944.9	2%
ATHLETIC ACTIVITY AREAS	21841.33	2%
INACTIVE - ASSIGNABLE	15623.87	1%
GENERAL LOUNGE SPACE	14222.31	1%
UNDERGRADUATE LAB SUPPORT SPACE	13228.39	1%
CENTRAL ADMIN OFFICE SUPPORT	13142.37	1%
UNSCHEDULED CLASS LAB	12258.49	1%
RESIDENCE SERVICE SPACE	10902.7	1%
STUDY SPACE NOT UNDER LIBRARY JURISDICTION	10759.86	1%
RESEARCH OFFICE/PROJECT SPACE	10747	1%
CENTRAL UTILITY PLANT	10619.93	1%
PLANT MAINTENANCE	9332.36	1%
LIBRARY SUPPORT SPACE	9293.38	1%
ASSEMBLY FACILITIES	8744.2	1%
STUDENT OFFICE AND SUPPORT SPACE	8465.62	1%
LIBRARY OFFICE SPACE	8105.4	1%
ATHLETIC SERVICE SPACE	7921.72	1%
FOOD FACILITIES	7081.38	1%
NON-INST AGENCY OCCU UNIV SPACE	6762.55	1%
DEMONSTRATION SCHOOLS/OTH INSTR INS	6689.45	1%
CLASSROOM SERVICE SPACE	6016.25	1%
FOOD FACILITIES SERVICES	5528.3	0%

BOOKSTORE/MERCHANDISING	4837.88	0%
INACTIVE - UNASSIGNABLE	4654.59	0%
SPECIALIZED CENTRAL AREAS	3797.08	0%
NON-INSTITUTIONAL AGENCIES	3520.73	0%
RECREATIONAL FACILITIES AND SERVICE	3401.45	0%
CENTRAL COMPUTING FACILITIES	3254.57	0%
OTHER CENTRAL SERVICE	2581.46	0%
EXHIBITION FACILITIES	2409.67	0%
ATHLETIC SEATING AREAS	2064.32	0%
HEALTH SERVICE FACILITIES	1701.61	0%
DAY-CARE FACILITIES	1222.09	0%
INACTIVE ASSIGNABLE	649.82	0%
DEMONSTRATION SCHOOLS	594.37	0%
NON-UNIVERSITY MERCHANDISING	271.02	0%
EX-UNIV. MERCHANDISING FACILITY	142.16	0%
RESEARCH OFFICE	132.06	0%

This function is a wrapper for the VIF function in the fmsb package

```
vif_func<-function(in_frame,thresh=10,trace=T){

 require(fmsb)

 if(class(in_frame) != 'data.frame') in_frame<-data.frame(in_frame)

 vif_init<-NULL
 for(val in names(in_frame)){
    form_in<-formula(paste(val,' ~ .'))
    vif_init<-rbind(vif_init,c(val,VIF(lm(form_in,data=in_frame))))
    }
 vif_max<-max(as.numeric(vif_init[,2]))

 if(vif_max < thresh){
  if(trace==T){
     prmatrix(vif_init,collab=c('var','vif'),rowlab=rep('',nrow(vif_init)),quote=F)
     cat('\n')
     cat(paste('All variables have VIF < ', thresh,', max VIF ',round(vif_max,2), sep=''),'\n\n')
     }
  return(names(in_frame))
  }
 else{

  in_dat<-in_frame

  while(vif_max >= thresh){

    vif_vals<-NULL

    for(val in names(in_dat)){
     form_in<-formula(paste(val,' ~ .'))
     vif_add<-VIF(lm(form_in,data=in_dat))
     vif_vals<-rbind(vif_vals,c(val,vif_add))
     }
    max_row<-which(vif_vals[,2] == max(as.numeric(vif_vals[,2])))[1]

    vif_max<-as.numeric(vif_vals[max_row,2])

    if(vif_max<thresh) break

    if(trace==T){
     prmatrix(vif_vals,collab=c('var','vif'),rowlab=rep('',nrow(vif_vals)),quote=F)
```

```
    cat('\n')
    cat('removed: ',vif_vals[max_row,1],vif_max,'\n\n')
    flush.console()
    }

   in_dat<-in_dat[,!names(in_dat) %in% vif_vals[max_row,1]]

   }

  return(names(in_dat))

  }

 }

lm.dat <- read.csv("2012l.csv", header=T, sep=",")

form.in<-paste('ec2012 ~',paste(names(lm.dat)[-15],collapse='+'))
mod1<-lm(form.in,data=lm.dat)
summary(mod1)

keep.dat<-vif_func(in_frame=as.matrix(lm.dat[-15]),thresh=5,trace=T)
form.in<-paste('ec2012 ~',paste(keep.dat,collapse='+'))
mod2<-lm(form.in,data=lm.dat)
summary(mod2)
```

The Dredge script below shows the global model after collinear variables have been removed.

```
fit2 <- glm(ec2012 ~ c2 + c3 + c5 + c6 + c9 + c11 + c12 + factor(Footprint.Shape) +
Above.Ground.Floors + Below.Ground.Floors + Shared.Wall)
options(na.action = "na.fail")
dredge(fit2, extra = c("adjR^2"))
```

Below is an example of the script used to carry out Leave-one-out Cross-validation. Notice that the global model is a more refined (i.e. less predictor variables) version of the one used in the Dredge script.

```
bestfit2c <- glm(ec2012 ~ c6 + c9 + Below.Ground.Floors)
cv.glm(dfspace2, bestfit2c)$delta
```

APPENDIX D1 – 2012 CANDIDATE MODELS WITH AN AICC OF SEVEN OR LESS (DREDGE OUTPUT)

Subset 1

Intercept	Above Ground Floors	Below Ground Floors	c11	c12	c2	c3	c5	c6	c9	Shared Wall	df	logLik	AICc	delta	weight
192118	NA	NA	NA	NA	NA	NA	NA	-301	472	NA	4	-260.998	532.664	0.000	0.119
172489	NA	NA	NA	NA	NA	NA	-209	-381	605	NA	5	-259.547	533.379	0.715	0.083
217164	NA	-81289	NA	NA	NA	NA	NA	-253	486	NA	5	-259.913	534.112	1.448	0.058
218262	NA	-109137	NA	NA	NA	NA	NA	NA	343	NA	4	-261.806	534.278	1.614	0.053
182611	NA	NA	NA	NA	NA	NA	NA	NA	284	NA	3	-263.484	534.469	1.805	0.048
202421	NA	NA	NA	NA	NA	-175	NA	-324	482	NA	5	-260.659	535.604	2.940	0.027
185079	NA	NA	NA	NA	NA	-279	-250	-433	648	NA	6	-258.578	535.617	2.954	0.027
215560	NA	NA	NA	NA	NA	NA	NA	-324	456	-32762	5	-260.688	535.663	2.999	0.027
188819	NA	NA	-258	NA	NA	NA	NA	-272	467	NA	5	-260.828	535.941	3.278	0.023
177235	NA	NA	-585	NA	NA	NA	NA	NA	314	NA	4	-262.670	536.006	3.342	0.022
221737	-13953	NA	NA	NA	NA	NA	NA	-282	488	NA	5	-260.885	536.057	3.393	0.022
195080	NA	-63549	NA	NA	NA	NA	-177	-332	596	NA	6	-258.832	536.125	3.461	0.021
182495	NA	NA	NA	NA	40	NA	NA	-308	482	NA	5	-260.926	536.137	3.473	0.021
184692	NA	NA	NA	510	NA	NA	NA	-298	479	NA	5	-260.947	536.181	3.517	0.021
237024	-32341	NA	NA	NA	NA	NA	-252	-355	670	NA	6	-258.896	536.254	3.591	0.020
199276	NA	NA	NA	NA	NA	NA	-218	-412	593	-38705	6	-259.045	536.551	3.888	0.017
253080	-32546	NA	NA	NA	NA	NA	NA	NA	349	NA	4	-262.947	536.561	3.897	0.017
352335	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2	-265.993	536.691	4.028	0.016
198288	NA	-124551	NA	174 2	NA	NA	NA	NA	383	NA	5	-261.277	536.839	4.176	0.015
210489	NA	-97805	-443	NA	NA	NA	NA	NA	360	NA	5	-261.281	536.847	4.184	0.015
167146	NA	NA	-341	NA	NA	NA	-219	-347	605	NA	6	-259.204	536.869	4.205	0.015
234853	NA	-92934	NA	NA	NA	-239	NA	-277	503	NA	6	-259.216	536.893	4.230	0.014
254283	NA	-93752	NA	NA	NA	NA	NA	-279	466	-46509	6	-259.232	536.925	4.261	0.014
269612	-24798	-101965	NA	NA	NA	NA	NA	NA	389	NA	5	-261.448	537.181	4.518	0.012
174034	NA	NA	NA	NA	NA	NA	-81	NA	317	NA	4	-263.302	537.270	4.606	0.012
230460	NA	-118941	NA	NA	NA	-164	NA	NA	345	NA	5	-261.533	537.351	4.687	0.011
242731	NA	-119188	NA	NA	NA	NA	NA	NA	321	-30567	5	-261.559	537.404	4.741	0.011
170964	NA	NA	NA	811	NA	NA	NA	NA	299	NA	4	-263.383	537.432	4.769	0.011
167500	NA	NA	NA	354	NA	NA	-207	-378	609	NA	6	-259.518	537.498	4.834	0.011
170510	NA	NA	NA	NA	9	NA	-207	-382	606	NA	6	-259.542	537.546	4.883	0.010
185938	NA	NA	NA	NA	NA	-61	NA	NA	283	NA	4	-263.452	537.570	4.906	0.010
188740	NA	NA	NA	NA	NA	NA	NA	NA	276	-8842	4	-263.466	537.599	4.935	0.010
202603	NA	-94318	NA	127 6	NA	NA	NA	-237	506	NA	6	-259.582	537.626	4.962	0.010
182173	NA	NA	NA	NA	2	NA	NA	NA	285	NA	4	-263.484	537.635	4.972	0.010
211481	NA	-106287	NA	NA	NA	NA	-55	NA	364	NA	5	-261.707	537.699	5.035	0.010
229521	NA	-113294	NA	NA	-40	NA	NA	NA	340	NA	5	-261.739	537.763	5.100	0.009

213906	NA	-78611	-191	NA	NA	NA	NA	-233	482	NA	6	-259.811	538.084	5.420	0.008
239364	-10693	-79666	NA	NA	NA	NA	NA	-240	498	NA	6	-259.840	538.141	5.477	0.008
216226	NA	-80848	NA	NA	3	NA	NA	-254	487	NA	6	-259.913	538.287	5.623	0.007
213442	NA	-74895	NA	NA	NA	-317	-218	-382	643	NA	7	-257.484	538.300	5.637	0.007
392838	NA	-62313	NA	NA	NA	NA	NA	NA	NA	NA	3	-265.560	538.620	5.957	0.006
222007	NA	NA	NA	NA	NA	-327	-269	-483	640	-50196	7	-257.661	538.655	5.992	0.006
163051	NA	NA	-683	NA	NA	NA	-125	NA	369	NA	5	-262.212	538.710	6.046	0.006
358711	NA	NA	NA	NA	NA	NA	NA	NA	NA	-42511	3	-265.616	538.731	6.067	0.006
256604	-35423	NA	NA	NA	NA	-298	-300	-408	722	NA	7	-257.714	538.762	6.098	0.006
232266	NA	NA	NA	NA	NA	-206	NA	-356	466	-39122	6	-260.210	538.881	6.217	0.005
267605	-46970	NA	NA	NA	NA	NA	-157	NA	440	NA	5	-262.310	538.906	6.242	0.005
359720	NA	NA	-347	NA	NA	NA	NA	NA	NA	NA	3	-265.767	539.034	6.370	0.005
227031	-22573	NA	-485	NA	NA	NA	NA	NA	354	NA	5	-262.418	539.123	6.459	0.005
233358	NA	-76056	NA	NA	NA	NA	-182	-361	579	-48884	7	-257.987	539.308	6.644	0.004
345337	NA	NA	NA	NA	NA	NA	53	NA	NA	NA	3	-265.920	539.340	6.676	0.004
355895	NA	NA	NA	NA	NA	-83	NA	NA	NA	NA	3	-265.945	539.389	6.726	0.004
358389	NA	NA	NA	NA	-38	NA	NA	NA	NA	NA	3	-265.951	539.401	6.738	0.004
363733	NA	NA	NA	NA	NA	NA	NA	-28	NA	NA	3	-265.969	539.438	6.775	0.004
166569	NA	NA	-579	746	NA	NA	NA	NA	327	NA	5	-262.576	539.438	6.775	0.004
332067	6050	NA	NA	NA	NA	NA	NA	NA	NA	NA	3	-265.975	539.450	6.786	0.004
353552	NA	NA	NA	-324	NA	NA	NA	NA	NA	NA	3	-265.979	539.459	6.795	0.004
253626	-16955	NA	NA	NA	NA	NA	NA	-304	475	-35659	6	-260.518	539.498	6.834	0.004
285819	NA	-111470	NA	NA	NA	-299	NA	-316	482	-58328	7	-258.089	539.511	6.847	0.004
231477	-13715	NA	NA	NA	NA	-174	NA	-305	498	NA	6	-260.546	539.553	6.890	0.004
211366	NA	NA	-233	NA	NA	NA	NA	-297	453	-31055	6	-260.546	539.554	6.891	0.004
198860	NA	NA	-205	NA	NA	-159	NA	-298	477	NA	6	-260.551	539.564	6.901	0.004
183919	NA	NA	-586	NA	NA	NA	NA	NA	305	-9662	5	-262.646	539.577	6.914	0.004
279814	-37630	NA	NA	NA	NA	NA	-271	-388	666	-46578	7	-258.129	539.592	6.929	0.004
179502	NA	NA	-579	NA	NA	-40	NA	NA	313	NA	5	-262.654	539.593	6.930	0.004
176973	NA	NA	-585	NA	1	NA	NA	NA	314	NA	5	-262.670	539.625	6.961	0.004

Subset 2

Intercept	Below Ground Floors	c11	c13	c2	c4	c5	c6	c7	c9	Shared Wall	df	logLik	AICc	delta	weight
61000	335355	-804	NA	NA	NA	NA	NA	-2023	506	NA	6	-275.574	569.609	0.000	0.497
47560	360826	-866	-591	NA	NA	NA	NA	-2146	531	NA	7	-274.773	572.879	3.270	0.097
-1385	350205	-780	NA	NA	NA	NA	NA	-1995	525	143835	7	-275.250	573.834	4.225	0.060
160613	NA	-697	NA	NA	NA	NA	NA	-1884	596	NA	5	-279.876	574.038	4.429	0.054
-1751	324610	-789	NA	50	NA	NA	NA	-1977	533	NA	7	-275.429	574.191	4.582	0.050
42823	313762	-818	NA	NA	-44	NA	NA	-2059	537	NA	7	-275.440	574.214	4.605	0.050
28487	331976	-793	NA	NA	NA	NA	35	-1985	504	NA	7	-275.442	574.218	4.609	0.050
49473	350521	-803	NA	NA	NA	85	NA	-1978	490	NA	7	-275.492	574.317	4.708	0.047
-143854	429855	-863	-1214	NA	NA	NA	NA	-2195	614	408675	8	-272.717	574.524	4.915	0.043
706227	403475	-693	NA	NA	NA	NA	NA	-1603	NA	NA	5	-280.368	575.022	5.413	0.033
84623	NA	-761	NA	NA	-137	NA	NA	-2025	674	NA	6	-278.864	576.190	6.581	0.019

Subset 3

Intercept	Above Ground Floors	Below Ground Floors	c11	c13	c2	c3	c4	c7	c8	Shared Wall	Total Area	df	logLik	AICc	delta	weight
312421	156541	NA	NA	501	571	NA	NA	NA	NA	NA	NA	5	-289.020	592.326	0.000	0.206
465415	150550	NA	-78	447	522	NA	NA	NA	NA	NA	NA	6	-287.339	593.140	0.814	0.137
669342	136998	NA	-96	NA	465	NA	NA	NA	NA	NA	NA	5	-290.358	595.002	2.676	0.054
506712	142533	NA	NA	NA	518	NA	NA	NA	NA	NA	NA	4	-292.264	595.195	2.868	0.049
215987	165161	NA	NA	528	597	NA	NA	NA	NA	538483	NA	6	-288.409	595.280	2.954	0.047
-171165	143081	NA	NA	494	580	NA	NA	NA	NA	NA	55	6	-288.448	595.357	3.031	0.045
323556	169979	NA	NA	486	535	NA	NA	-274	NA	NA	NA	6	-288.459	595.380	3.054	0.045
260850	160897	NA	NA	516	578	NA	662	NA	NA	NA	NA	6	-288.689	595.839	3.513	0.036
488869	165931	NA	-84	427	477	NA	NA	-322	NA	NA	NA	7	-286.416	596.165	3.839	0.030
296466	153926	NA	NA	511	578	136	NA	NA	NA	NA	NA	6	-288.888	596.237	3.911	0.029
301724	161633	NA	NA	497	570	NA	NA	NA	-821	NA	NA	6	-288.891	596.243	3.917	0.029
-53140	135799	NA	-81	439	530	NA	NA	NA	NA	NA	59	7	-286.539	596.412	4.086	0.027
291124	155424	33484	NA	511	561	NA	NA	NA	NA	NA	NA	6	-288.998	596.458	4.132	0.026
685751	155827	NA	-101	NA	415	NA	NA	-378	NA	NA	NA	6	-289.404	597.270	4.944	0.017
382973	157273	NA	-72	472	545	NA	NA	NA	NA	388552	NA	7	-286.981	597.295	4.969	0.017
513231	159134	NA	NA	NA	478	NA	NA	-328	NA	NA	NA	5	-291.677	597.640	5.314	0.014
422565	153842	NA	-75	461	529	NA	457	NA	NA	NA	NA	7	-287.158	597.649	5.323	0.014
102896	121279	NA	-98	NA	475	NA	NA	NA	NA	NA	64	6	-289.665	597.792	5.466	0.013
-20510	128121	NA	NA	NA	529	NA	NA	NA	NA	NA	60	5	-291.779	597.843	5.517	0.013
451250	148768	NA	-77	456	528	98	NA	NA	NA	NA	NA	7	-287.259	597.851	5.525	0.013
456605	153327	NA	-77	447	522	NA	NA	NA	-427	NA	NA	7	-287.299	597.931	5.605	0.013
487099	151375	-28739	-80	437	529	NA	NA	NA	NA	NA	NA	7	-287.321	597.975	5.649	0.012
753079	142613	-141751	-103	NA	507	NA	NA	NA	NA	NA	NA	6	-289.988	598.437	6.111	0.010
452138	147527	NA	NA	NA	533	NA	NA	NA	NA	342312	NA	5	-292.086	598.457	6.131	0.010
492272	148764	NA	NA	NA	517	NA	NA	NA	-986	NA	NA	5	-292.129	598.543	6.217	0.009
-348492	150952	NA	NA	524	611	NA	NA	NA	NA	610089	63	7	-287.620	598.574	6.248	0.009
549593	146070	-83672	NA	NA	545	NA	NA	NA	NA	NA	NA	5	-292.156	598.597	6.271	0.009
478344	144976	NA	NA	NA	522	NA	413	NA	NA	NA	NA	5	-292.171	598.627	6.301	0.009
502416	141467	NA	NA	NA	520	50	NA	NA	NA	NA	NA	5	-292.251	598.788	6.462	0.008
637392	139678	NA	-94	NA	474	NA	NA	NA	NA	173773	NA	6	-290.304	599.070	6.744	0.007
659084	140149	NA	-94	NA	465	NA	NA	NA	-481	NA	NA	6	-290.320	599.102	6.776	0.007
654338	138177	NA	-95	NA	467	NA	187	NA	NA	NA	NA	6	-290.336	599.133	6.807	0.007
233491	176834	NA	NA	512	562	NA	NA	-251	NA	497771	NA	7	-287.913	599.159	6.833	0.007
144085	171287	NA	NA	549	609	NA	786	NA	NA	597938	NA	7	-287.916	599.165	6.839	0.007
667886	136706	NA	-96	NA	466	14	NA	NA	NA	NA	NA	6	-290.357	599.175	6.849	0.007
-133376	156504	NA	NA	481	546	NA	NA	-258	NA	NA	52	7	-287.923	599.180	6.854	0.007

Subset 4

Intercept	Above Ground Floors	Below Ground Floors	c10	c11	c12	c2	c3	c5	c8	c9	Shared Wall	df	logLik	AICc	delta	weight
-780896	187814	NA	NA	NA	NA	575	NA	-398	NA	641	NA	6	-304.776	628.013	0.000	0.117
-629345	NA	NA	NA	NA	NA	619	NA	NA	NA	692	NA	4	-309.307	629.281	1.268	0.062
-1232911	149613	NA	NA	NA	NA	607	NA	NA	NA	642	NA	5	-307.599	629.483	1.471	0.056
-1141608	201956	NA	NA	NA	NA	585	NA	-365	NA	641	1189720	7	-303.304	629.941	1.929	0.044
-1704834	165017	NA	NA	NA	NA	641	628	NA	NA	641	NA	6	-305.835	630.131	2.119	0.040
-163416	NA	NA	NA	NA	NA	598	NA	-306	NA	701	NA	5	-307.931	630.147	2.135	0.040
-1613806	170024	NA	NA	NA	NA	615	NA	NA	NA	641	1402331	6	-306.024	630.509	2.497	0.033
-1193503	193094	NA	NA	NA	NA	605	456	-336	NA	641	NA	7	-303.663	630.659	2.646	0.031
-986171	NA	NA	NA	NA	NA	651	547	NA	NA	695	NA	5	-308.202	630.689	2.677	0.031
-119961	NA	-415440	NA	NA	NA	607	NA	NA	NA	714	NA	5	-308.238	630.761	2.748	0.030
-1306112	168420	NA	-999	NA	NA	623	NA	NA	NA	648	NA	6	-306.330	631.121	3.108	0.025
-870105	NA	NA	NA	NA	NA	628	NA	NA	NA	697	1130876	5	-308.454	631.193	3.181	0.024
-762835	199444	NA	NA	-135	NA	573	NA	-392	NA	647	NA	7	-303.946	631.226	3.213	0.023
-883201	197163	NA	-720	NA	NA	590	NA	-354	NA	646	NA	7	-303.963	631.260	3.247	0.023
-423470	177791	-293738	NA	NA	NA	569	NA	-369	NA	659	NA	7	-303.969	631.271	3.258	0.023
-627252	NA	NA	-784	NA	NA	633	NA	NA	NA	702	NA	5	-308.653	631.591	3.579	0.019
-737100	140321	-373797	NA	NA	NA	596	NA	NA	NA	665	NA	6	-306.582	631.626	3.613	0.019
-571605	NA	NA	NA	-106	NA	618	NA	NA	NA	699	NA	5	-308.981	632.248	4.236	0.014
-1207278	162444	NA	NA	-142	NA	604	NA	NA	NA	647	NA	6	-306.906	632.274	4.261	0.014
-1021425	202949	NA	NA	NA	NA	604	NA	-395	799	635	NA	7	-304.495	632.322	4.310	0.014
-593094	198770	NA	NA	NA	-72	549	NA	-439	NA	634	NA	7	-304.551	632.436	4.423	0.013
238446	NA	-363796	NA	NA	NA	589	NA	-277	NA	719	NA	6	-307.011	632.483	4.470	0.012
-1735920	180763	NA	-897	NA	NA	653	582	NA	NA	647	NA	7	-304.631	632.595	4.582	0.012
-778897	NA	NA	NA	NA	64	639	NA	NA	NA	696	NA	5	-309.176	632.638	4.626	0.012
-1106333	160949	-392020	NA	NA	NA	604	NA	NA	NA	665	1448386	7	-304.697	632.727	4.714	0.011
-1173107	218469	NA	NA	-167	NA	584	NA	-353	NA	647	1367377	8	-301.821	632.732	4.719	0.011
-350350	NA	-435579	NA	NA	NA	615	NA	NA	NA	720	1198148	6	-307.163	632.787	4.775	0.011
-1631285	188796	NA	NA	-179	NA	613	NA	NA	NA	648	1585174	7	-304.739	632.811	4.798	0.011
-653777	NA	NA	NA	NA	NA	623	NA	NA	111	691	NA	5	-309.303	632.892	4.880	0.010
-409709	NA	NA	NA	NA	NA	607	NA	-274	NA	704	932538	6	-307.288	633.037	5.025	0.009
-518535	NA	NA	NA	NA	NA	625	410	-248	NA	702	NA	6	-307.289	633.039	5.027	0.009
-498857	NA	-360015	NA	NA	NA	636	477	NA	NA	714	NA	6	-307.334	633.129	5.117	0.009
-1486701	165987	NA	NA	NA	NA	637	NA	NA	852	635	NA	6	-307.358	633.178	5.166	0.009
-1630515	183552	NA	-848	NA	NA	628	NA	NA	NA	647	1235116	7	-304.992	633.317	5.304	0.008
-1285798	187987	NA	-1169	-182	NA	622	NA	NA	NA	656	NA	7	-305.052	633.438	5.425	0.008
-1254281	155907	-303544	NA	NA	NA	629	564	NA	NA	660	NA	7	-305.065	633.464	5.452	0.008
-772372	191791	-317407	NA	NA	NA	579	NA	-333	NA	660	1245763	8	-302.200	633.491	5.478	0.008
-1296545	146453	NA	NA	NA	33	617	NA	NA	NA	645	NA	6	-307.559	633.580	5.568	0.007
-217410	NA	NA	-543	NA	NA	610	NA	-269	NA	707	NA	6	-307.598	633.657	5.644	0.007

-124071	NA	NA	NA	-94	NA	597	NA	-298	NA	707	NA	6	-307.638	633.737	5.725	0.007
-148905	NA	-390301	-704	NA	NA	620	NA	NA	NA	721	NA	6	-307.656	633.774	5.761	0.007
-957714	NA	NA	-681	NA	NA	660	506	NA	NA	704	NA	6	-307.660	633.782	5.770	0.007
-1324879	NA	NA	NA	NA	119	693	640	NA	NA	704	NA	6	-307.722	633.906	5.894	0.006
-1649555	174006	NA	NA	-112	NA	636	581	NA	NA	645	NA	7	-305.335	634.004	5.992	0.006
-2062050	198404	NA	NA	NA	NA	664	NA	NA	1339	630	1584202	7	-305.343	634.019	6.006	0.006
-1781260	172262	NA	NA	NA	NA	635	428	NA	NA	641	834878	7	-305.405	634.144	6.132	0.005
-855038	158554	-335601	-919	NA	NA	612	NA	NA	NA	668	NA	7	-305.413	634.159	6.147	0.005
-1561422	227144	NA	NA	NA	NA	630	NA	-356	1222	631	1360529	8	-302.558	634.207	6.194	0.005
-823055	NA	NA	NA	-130	NA	627	NA	NA	NA	706	1242162	6	-307.925	634.311	6.299	0.005
-149523	NA	NA	NA	NA	NA	596	NA	-306	-58	701	NA	6	-307.930	634.321	6.308	0.005
-158391	NA	NA	NA	NA	-2	597	NA	-307	NA	701	NA	6	-307.931	634.323	6.310	0.005
-1930150	157954	NA	NA	NA	90	674	695	NA	NA	649	NA	7	-305.495	634.323	6.311	0.005
-884870	213642	NA	-888	-166	NA	591	NA	-337	NA	653	NA	8	-302.618	634.327	6.314	0.005
-1981468	182701	NA	NA	NA	NA	674	635	NA	911	633	NA	7	-305.505	634.344	6.331	0.005
-1197418	208628	NA	-605	NA	NA	597	NA	-331	NA	645	1090313	8	-302.656	634.404	6.391	0.005
-837692	NA	NA	-647	NA	NA	638	NA	NA	NA	704	986750	6	-307.987	634.435	6.423	0.005
-920996	NA	NA	NA	-78	NA	648	512	NA	NA	701	NA	6	-308.014	634.489	6.477	0.005
-1018138	NA	NA	NA	NA	NA	647	402	NA	NA	697	594740	6	-308.030	634.521	6.509	0.005
-265153	NA	-412896	NA	NA	61	626	NA	NA	NA	717	NA	6	-308.106	634.674	6.662	0.004
-1138333	NA	NA	NA	NA	101	660	NA	NA	NA	704	1277972	6	-308.108	634.677	6.665	0.004
-555313	NA	NA	-888	-132	NA	633	NA	NA	NA	712	NA	6	-308.124	634.710	6.698	0.004
-1277639	201848	NA	-687	NA	NA	619	441	-297	NA	645	NA	8	-302.835	634.762	6.749	0.004
-37172	NA	-509431	NA	60	NA	605	NA	NA	NA	715	NA	6	-308.183	634.827	6.814	0.004
-1007598	NA	NA	NA	NA	NA	654	547	NA	98	695	NA	6	-308.199	634.859	6.846	0.004
-127108	NA	-415135	NA	NA	NA	608	NA	NA	31	714	NA	6	-308.237	634.936	6.924	0.004
-1781534	164401	NA	NA	NA	73	639	NA	NA	NA	648	1498825	7	-305.802	634.936	6.924	0.004

GLOSSARY

Akaike Information Criterion	A metric used to guide model selection. In theory, the chosen model is often one that minimizes the Kullback-Leibler distance between the model and the "truth". In practice, the criterion looks for models with a good fit with minimal number of parameters.
– **AICc**	A modified version of AIC applicable to small samples. An increased penalty for additional parameters is used. [73, 124]
– **Akaike Weight**	The relative likelihood of the model, given the data. Normalized to sum to 1 among candidate models. [73]
Causation	A cause and effect relationship between variables. Separate from correlation, causation among variables can only be inferred from a well design randomized controlled experiment. [125]
Coefficient of determination (R^2)	A metric used to convey the goodness of fit for a model. For a regression, the values range from 0 to 1 and measure how close the data is to the fitted line. [126]
– **Adjusted R^2**	A modified version of R^2 that takes into account the number of predictor variables in the model. The value increases only when the additional variable(s) improves the model more than would be expected by chance. The adjusted R^2 is always lower or equal to the R^2 for that model. [127]
Coefficient of Variation	A measure of dispersion calculated by normalizing the standard deviation of a sample with its mean. [128]
Collinearity	An undesirable property among predictor variables where excessive correlation exists. This correlation has no negative impacts on the performance of the regression model but is problematic when the effects of individual predictor variables are of interest. [129]
Correlation	A function of the relationship between variables. [130]

– **Coefficient of Correlation** A measure of correlation that indicates the strength and direction of the relationship between variables. Values range from 1 (positive relationship) to -1 (negative relationship) with weaker relationships having a coefficient closer to 0. [130]

Cross Validation A process that allows for model validation by splitting the data into a training set and testing set. [131]

– **K-Fold** Data is randomly divided into *k* equal sized subsets. Cross validation is then run *k* number of rounds. During each round, 1 subset will be used for validations and the remaining ones, for training. Error is averaged over all rounds. [131]

– **Leave-One-Out** The process of calculating error is the same as k-fold. However, in this method, $k = n$ (the number of data points in the entire sample). [132]

Degrees of Freedom (Residual) A measure of certainty that the sample population is representative of the entire population. Degrees of freedom are a function of sample size and the number of independent variables. Each addition parameter included in the model reduces the degrees of freedom by 1. Generally, the more degrees of freedom a model has, the more accurately sampled the population is. [133, 134]

Freedman Paradox/model selection bias Bias is introduced into the model when variables with a weak relationship to the response variable are selected as significant. Their selection is often a result of small effects being magnified within a particular dataset. A bias will always exists when selecting the "best" model from a large selection of models employing many predictor variables. The Freedman Paradox is the extreme case of this bias where predictor variables having no relationship to the response are included in the model thus spuriously inflating the R^2. [83, 135]

Interquartile Range A descriptive statistic that measures statistical dispersion in a dataset. Quartiles each represent 25% of the data; the range is the distance between the first and third quartile. [136]

Log Likelihood A statistic used to measure how likely a particular population is to produce an observed sample. When used to compare models, the value can be used to validate the plausibility of one model to that of another. [137]

Nominal Cost of Electricity The value of the good or service at the time it was produced.

Principle of Parsimony The ideal that a model should neither be under- nor over-fitted. The model should be as simple as possible with respect to the parameters, model structure, and number of variables. In practice, parsimony demands a tradeoff between bias/variance and the number of estimated variables in the model. [73]

Real Cost of Electricity The nominal value adjusted for inflation. A positive inflation over time will result in a decrease in the real value if the nominal price remains the same.

Root Mean Square Error A measure of the difference between predicted values from a model and actual observed values (residuals). Calculated by taking the square root of the mean of the squares of the deviation, the RMSE is a good measure of accuracy. [138]

- **Coefficient of Variation of the Root Mean Square Error** A normalized RMSE to the mean of the observed value – often reported as a percentage.

Standard Deviation A statistic measuring how far each value in the dataset is from the mean. The measure is expressed in the same units as the mean. [139, 140]

Variance A statistic measuring how far each value in the dataset is from the mean. The measure is expressed in square units. [139, 140]