Titre: Title:	Development of a bottom-up white-box building stock energy model for single-family dwellings
Auteurs: Authors:	Adam Neale, Michaël Kummert, & Michel Bernier
Date:	2022
Type:	Article de revue / Article
Référence: Citation:	Neale, A., Kummert, M., & Bernier, M. (2022). Development of a bottom-up white-box building stock energy model for single-family dwellings. Journal of Building Performance Simulation, 15(6), 735-756. https://doi.org/10.1080/19401493.2022.2082531

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/10473/
Version:	Version finale avant publication / Accepted version Révisé par les pairs / Refereed
Conditions d'utilisation: Terms of Use:	Tous droits réservés / All rights reserved

Titre de la revue: Journal Title:	Journal of Building Performance Simulation (vol. 15, no. 6)
Maison d'édition: Publisher:	Taylor & Francis
URL officiel: Official URL:	https://doi.org/10.1080/19401493.2022.2082531
Mention légale: Legal notice:	This is an Accepted Manuscript of an article published by Taylor & Francis in Journal of Building Performance Simulation (vol. 15, no. 6) in 2022, available online: https://doi.org/10.1080/19401493.2022.2082531

Development of a bottom-up white-box building stock energy model for single-family dwellings

Adam Neale*, Michaël Kummert and Michel Bernier

Département de Génie Mécanique, Polytechnique Montréal, Montreal, Canada;

Corresponding author: Adam Neale, adam.neale@polymtl.ca

Accepté pour publication dans Journal of Building Performance Simulation

– 19 mai 2022

Development of a bottom-up white-box building stock energy model for single-family dwellings

A new bottom-up white-box building stock energy modeling approach is presented in this paper. The model is a physics simulation-based stock model that can be used to compare the base case building stock with technological variations for comparative assessments. The process of characterizing the key parameters, such as the number of dwellings in each region and the heating, cooling and hot water system distributions, is described in detail. The model accuracy is compared to known stock data for a variety of categories, including end-use, energy source and building type. The model predicts the energy use of a building stock in a Canadian region with good agreement across all categories, with the total energy consumption of the model within 1.5% of the real stock energy use. The impact of the sample size of the modeled stock is evaluated, which demonstrates the importance of a sufficiently large sample to reduce the expected deviation for lesser-represented portions of the stock. A case study illustrates how the building stock model can be applied for a comparative assessment of different heating system distributions for the purpose of greenhouse gas emissions calculations, with an emphasis on the impact of measures on the peak electricity load of the building stock.

Keywords: residential; building stock energy model; bottom-up; white-box; greenhouse gas emissions

1 Introduction

Residential buildings account for approximately 17% of all energy use in Canada, compared to 9% for commercial buildings (NRCan 2014). Understanding the context of the energy use for residential buildings is essential for long term planning of energy efficiency measures and technology evaluation, which has become a priority for Canada (Ugursal 2017). There are a variety of applications for building stock modeling. For example, the provincial government of Québec has implemented a plan to reduce greenhouse gas emissions related to the heating of buildings by 50% by the year 2030 (Government of Québec 2020). There are nearly 2 million single-family dwellings in the province of Québec and non-electric energy usage represents between 22% and 42% of energy use for detached and attached homes, respectively, which is primarily for space heating and water heating (Figure 1). Space heating represents between 60 and 70% of the energy use of single-family homes in Québec, which has a predominantly cold climate. Reaching the Québec government's goal of 50% reduced emissions from space heating will require shifting fossil fuel energy sources to electricity and has significant implications on peak and overall electricity use.

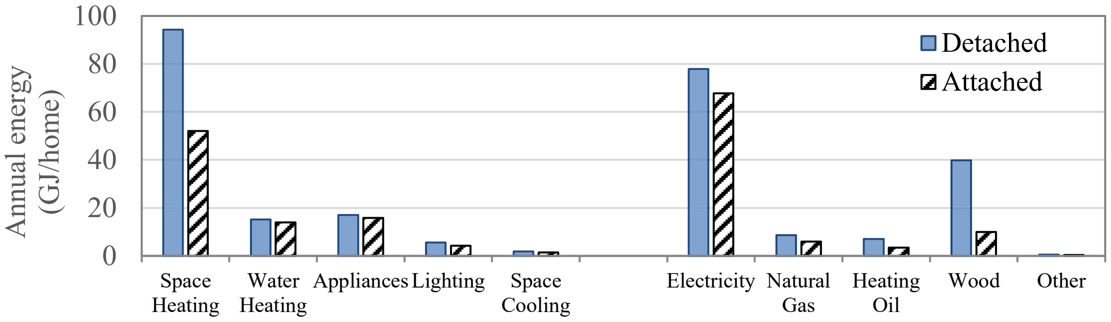


Figure 1. Annual energy consumption (GJ/home) by end-use and energy source for detached and attached residential dwellings in the province of Québec, Canada (NRCan 2017)

Accurately evaluating the impact of efficiency measures and technology changes on the annual and peak energy usage of homes at the municipal, regional or provincial level requires the use of some form of building stock energy model (BSEM).

1.1 BSEM approaches

Building stock energy modeling is the process of predicting the energy consumption of a large group of buildings, whether at the municipal, regional or national level. Comprehensive reviews of BSEM techniques have been performed previously by others (Swan and Ugursal 2009; Kavgic et al. 2010; Reinhart and Cerezo Davila 2016; Langevin et al. 2020). Methods have been typically divided in three broad categories: top-down, bottom-up (engineering), and bottom-up (statistics) (Swan and Ugursal 2009). More recently a new classification of BSEM has emerged that distinguishes stock modeling techniques among the following *"quadrants"* (Langevin et al. 2020):

- Quadrant 1: Top-down black-box
- Quadrant 2: Top-down white-box
- Quadrant 3: Bottom-up black-box
- Quadrant 4: Bottom-up white-box
- Multiple quadrants (hybrid)

The distinction *top-down* establishes a building stock's characteristics, such as energy consumption, in aggregate form and subsequently divides the stock according to various metrics, essentially starting at the so-called top level and working down to subdivided portions of the stock. *Bottom-up* models establish characteristics at the building level and subsequently aggregate the values to the desired level of accuracy, such as at the regional level. The *black-box* designation typically refers to models based on data, i.e. statistics-based models, while *white-box* refers to engineering-based cases that represent buildings and systems with physical models. A *multiple-quadrant* stock model combines several quadrants into a single approach (Langevin et al. 2020).

Langevin et al. (2020) also introduce a labelling system for building stock energy models, which categorizes models based on the general purpose of the model, the quadrant, the modeling technique used to produce the results, the availability of the model, and other factors. New stock models are encouraged to be expressed in terms of those descriptors, to facilitate the comparison with other works.

Booth et al. (2012) identified the following factors as common limitations of building stock modeling:

1) **Accuracy**: statistics-based models can more accurately represent energy consumption in homes because they are based on real energy use, whereas engineering-based models require many assumptions.

2) **Data collection**: all building stock models require significant information on the stock, which is often difficult to obtain. High-resolution energy data requires significant data storage space, especially when multiple energy sources must be considered.
3) **Computational time**: dynamic simulations require more time to perform than statistical models, and modeling an entire stock can be prohibitive in terms of total time and resources required.
4) **Decision-making**: it is difficult to evaluate the propagation of uncertainty in engineering-based models, and therefore difficult to establish the impact on decisions made using the stock model.
5) **Flexibility**: models based on statistical data are effective as long as the base assumptions of the stock do not significantly change over time. Statistics-based models are not able to easily adapt to leaps in technology or changes in occupancy behaviour, for example.

The factors identified by Booth et al. (2012) are further supported by Langevin et al. (2020), who also describe in detail the advantages and limitations of stock modeling methods. Top-down methods either have difficulty representing technological changes due to the aggregate approach of the model, or require extensive data and an expert user to properly disaggregate the stock data into a refined model. Black-box bottom-up models require a statistical approach towards establishing building energy consumption, relying on energy use data to establish trends. As with the top-down models, relying on existing energy consumption data results in a model that is inflexible given technological advancements and relies heavily on the availability of data. A bottom-up white-box model therefore presents an interesting opportunity that can evaluate the impact of various technologies provided the computational requirements can be mitigated.

1.2 Bottom-up white-box (BU/WB) building stock modeling

While stock models can be used for various applications, energy use (and subsequently greenhouse gas emissions) is the focus of the authors. A BU/WB building stock energy model (BSEM) has a number of advantages, including but not limited to:

1) Energy consumption at a high frequency can be determined for each household, allowing for more accurate stock peak energy use and greenhouse gas emissions based on time-of-use.
2) New technologies can be implemented gradually by changing probability distributions for equipment and evaluating the impact on the stock energy use.
3) Occupant behaviour can be modified, such as implementing more complex time-of-use energy incentives.
4) Regionally-targeted incentive measures can be evaluated and compared to the rest of the building stock.

A number of BU/WB building stock energy models exist in the literature. Several residential stock modeling examples are summarized in Table 1 and subsequently described in more detail.

Table 1. Examples of recent bottom-up white-box building stock energy models

Model name	**Stock**	**Stock size**[1]	**Building sample count (%)**	**Features**	**Market**	**Ref.**
CHREM	National (Canada)	~10 million households	16,952 (0.17%)	Hybrid approach between engineering and statistical techniques, national and provincial stock representation.	Residential	(Swan et al. 2012)
	Provincial (Québec)	~2 million households	3,670 (0.185%)		Residential	
UMI	Urban, user-defined	~30	30 (100%)	Flexibility, customized input of an urban building stock, multiple end-use applications	Commercial, residential	(Reinhart et al. 2013)
CityBES	Urban, user-defined	10 000	10 000 (100%)	Flexibility, customized input of an urban building stock, multiple end-use applications	Commercial, residential	(Hong et al. 2016)
ResStock	National (USA)	123 million households	350 000 (0.28%)	Visualization tools, national baseline. Energy sources and end-uses by state	Residential	(Wilson et al. 2017)
AutoBEM	Urban, user-defined	130 000	130 000 (100%)	Flexibility, uses a number of imaging techniques to build 3D maps of urban settings	Commercial, residential	(New et al. 2018)
Synthetic building stock tool	National (Switzer-land)	1.6 million households	10 000 (0.6%)	Auto generation of stock characteristic distributions	Residential	(Nägeli et al. 2018)
TREES	National (Japan)	53 million	16 000 (0.03%)	Detailed occupancy and house characteristics	Residential	(Taniguchi-Matsuoka et al. 2020)

[1] Stock size is based on the example use case provided by the authors of the tool.

The Canadian Hybrid Residential End-Use Energy and GHG Emissions Model (CHREM) combines a bottom-up engineering approach with neural networks to represent the end-use energy of the Canadian housing stock (Swan et al. 2012). The model is based on a database of 16,952 archetypal homes across Canada, which is used to model the energy use of individual dwellings. Considering the scope of the model, which represents approximately 10 million homes in Canada, the CHREM model provides a good approximation of the energy use at the provincial levels. As an example, for the province of Québec, Canada, the total energy use of the modeled homes is 15% higher than reference data, while electricity use is 33% higher.

The Urban Modeling Interface (UMI) allows a user to build up an urban building stock consisting of various buildings with commercial and/or residential end-uses (Reinhart et al. 2013). UMI is intended to allow a user to model any collection of urban buildings, which requires a user to input the details of the stock manually. Every building in the stock is modeled individually and therefore there is no sampling of the stock, which limits the model to smaller (urban) building stocks. It is unclear how accurate the energy prediction of UMI is, though it is based on EnergyPlus (US DOE 2013), which is a reliable energy simulation tool.

Similarly to UMI, the CityBES urban modeling software uses EnergyPlus as an engine to model a series of buildings in an urban setting (Hong et al. 2016). The user must input the details of the building stock to the model and specify the building parameters, which results in a flexible urban model requiring significant user manipulation. CityBES has a number of energy conservation measures already implemented in the tool, which allows a user to rapidly evaluate the impact of energy conservation measures on the stock energy consumption. Much like with UMI it is unclear how accurate the energy prediction of CityBES is on a building level, though it does include an auto-calibration feature if monthly building energy consumption is available. As stated by the creators of CityBES, the computational requirements for a large urban stock, such as one million buildings in New York City, become intractable and can require a more localized urban model.

The National Renewable Energy Laboratory (NREL) has developed a national residential building stock energy model called ResStock for visualization and energy prediction of dwellings across the United States of America (Wilson et al. 2017). ResStock uses conditional probability distributions to develop residential archetypes to represent dwellings across the country. Much like UMI and CityBES, ResStock relies on EnergyPlus to perform energy simulations. The National Baseline data viewer developed with ResStock (NREL 2021) is based on a subset of 350 000 buildings that represent approximately 123 million dwellings across the USA. The accuracy of the energy prediction across a number of housing types falls within ±20% in most cases, though some significant discrepancies are identified by the creators of the tool for certain housing combinations. Overall, the ResStock tool is a very detailed example of what is possible with a BU/WB model with sufficient data on building characteristics. However, due to the size of the housing stock in the USA, NREL must rely on a relatively small sample size (0.28%) for the aggregated energy consumption at the state and national levels.

The AutoBEM urban energy modeling tool developed by Oak Ridge National Laboratories (ORNL) again uses EnergyPlus as a basis for the prediction of urban energy consumption. The main differentiating factor for AutoBEM versus other urban tools, such as UMI and CityBES, is the use of multiple imaging techniques for the creation of 3D geometry of buildings for simulation purposes. It is difficult to evaluate the accuracy of AutoBEM as there is little detail provided by the authors on the validation of the model. An example of 130 000 mixed-use buildings in an urban setting is provided by ORNL in the form of a website (ORNL 2021).

Nägeli et al. (2018) propose a methodology and a tool based on a so-called *synthetic building stock model* using building stock characterisation and energy modeling. The energy consumption is calculated based on the monthly energy demand of each building. The tool is limited to a sample size of 10 000 buildings due to computational considerations. For the example given of the Swiss residential building stock, this results in a 0.6% sample of buildings simulated. The tool uses a detailed set of dwelling characteristics to represent each building, with a means to synthetically construct a set of dwellings and buildings based on input stock parameter distributions. While the tool has generally good agreement with building stock energy use, the tool authors acknowledge additional calibration could improve the results. In addition, given the monthly energy demand calculations it would be difficult to evaluate the impact of energy conservation measures on peak loads and other factors requiring high frequency energy consumption data.

The Total Residential End-use Energy Simulation (TREES) tool is a residential building stock energy model developed for Japan (Taniguchi-Matsuoka et al. 2020). Detailed occupancy and appliance characteristics are modeled to predict the space heating, space cooling, water heating and appliance energy consumption for randomly sampled dwellings. A simplistic thermal circuit network method is used to predict the heating and cooling loads. The sample size of the modeled stock compared to the total building stock is 0.03%, which results in significant discrepancies between reported stock energy consumption and the model prediction. Nevertheless, the structure of the TREES model illustrates how detailed dwelling characteristics can be implemented in a BU/WB stock tool.

The illustrated BU/WB stock models provide some relevant examples of energy prediction tools for dwellings at the urban and national scales. Accuracy of the models is a recurring issue, as the sample size is limited by the scope of the stock – urban cases can be modeled entirely, while national stocks require a very small sample due to computational resource limitations. It is difficult to achieve a high degree of accuracy when less than 0.1% of homes are simulated. Few details, if any, are provided on the impact of the sample size of BU/WB tools on overall accuracy. In addition, stock energy models appear to be mainly focused on energy consumption when peak loads are often an important consideration for stakeholders. Most tools are used for comparative studies, such as evaluating the impact of an energy conservation measure, in which case the absolute accuracy of the tool seems to be secondary to the tool authors.

2 Objectives

This paper describes a new bottom-up white-box building stock energy modeling

(BSEM) approach for single-family home building stocks. The proposed approach is a new method of modeling a large number of statistically-representative single-family dwellings, which are combined to construct an accurate residential stock model. While the general methodology presented in this paper is applicable to other building stocks, the authors have applied it to the single-family dwelling market in the province of Québec, Canada, a building stock representing 1.9 million detached, semi-detached, row and other single-attached homes. Multi-family buildings, such as apartment complexes, are not included in the proposed modeling approach. The resulting stock model is called the Québec Single-Family Building Stock Energy Model (QSFBSEM). The objective of the QSFBSEM is to provide a validated stock model that can evaluate different technology and building stock scenarios and study the impact on energy usage and peak loads for a variety of energy sources and end-uses. More specifically, this paper aims to:

1) Present a general methodology to develop a building stock model from a series of building energy simulations of individual houses.
2) Describe the characterisation process of the studied building stock, including a description of typical dwellings, population distribution, climate zones and common building systems.
3) Describe the characterization and implementation of the stock model in detail.
4) Validate the aggregate results of the proposed model with building stock energy consumption data.
5) Present an example application of the stock model.
6) Discuss how the proposed model improves upon previous works.

This work builds on a previous study by Neale et al. (2020) that presented a methodology to develop a virtual smart meter data set. Some of the building stock characterisation was presented previously, though the work by Neale et al. was for electricity use profiles only and did not consider other energy sources. In order to reduce repetition some details are summarised here and readers can refer to the paper by Neale et al. for further details. In many cases, data and probability distributions for characteristics have been updated, or additional details were added, and therefore are presented in this paper for clarity.

The main components of the modeling methodology are first described. The building stock characterisation process is then presented, with an emphasis on the distribution of houses across the studied region, the heating, cooling and domestic hot water systems, and other specific parameters of the stock. Some results are provided to illustrate the accuracy of the model based on energy source, building type and energy end-use. A case study is provided, followed by some discussion and conclusions.

3 Model description

The proposed building stock energy model uses a new approach that generates individual dwellings according to a detailed stock characterisation process described in Section 4 of the paper. Parameters are generated according to interdependent probability distributions combined using Bayes' Theorem, a process described in additional detail in Appendix 3. In brief, where applicable, conditional probability distributions are used to model houses consistent with the studied building stock.

Houses are generated independently according to the aforementioned probability distributions. The energy consumption for the subset of generated homes is scaled up to match the overall building stock size. As an example, typically 200 000 homes are generated for a building stock of approximately 1.9 million houses. The impact of the size of the building subset is addressed briefly in Section 5 of the present paper and in more detail in Appendix 4. The process of combining individually-modeled dwellings and aggregating them to the regional- and provincial-level classifies the model as a bottom-up white-box building stock energy model, as defined by Langevin et al. (2020).

The TRNSYS simulation program is used to model the energy use of each dwelling in the studied building stock (Klein et al. 2017). TRNSYS was selected as the building simulation tool because it had all of the required features to generate and simulate the individual houses and interface with another external program, which controls the batch simulation process. A user can execute a simulation for a variety of stock sample sizes and expect that the correct distribution of house parameters will be applied by the model, with some statistical variation. A parallel processing approach is adopted to distribute the task to multiple workers due to the number of houses required for a building stock simulation. A worker is defined here as a computational thread that a task can be assigned to for processing. Typically 20 workers are assigned to the task of modeling the building stock. As a point of reference, each set of 1000 houses modeled using the stock model requires approximately 30 minutes on a server with an Intel Core i9-7920X processor @2.9 GHz, 128 GB of RAM @2133 MHz and a SATA III solid-state hard drive. The modeling process is divided into a number of steps (Figure 2):

1) **Model initialization**: a new house simulation is initialized. Conditions for the stock simulation are registered at this stage. For example, limiting the stock simulation to a single region instead of all seven regions of the province.
2) **Input generation**: the building parameters (Figure 3) for each building are generated according to the probability distributions from the stock characterisation process and based on the scope of the simulation specified in 1).
3) **TRNSYS file preparation**: required TRNSYS simulation files are automatically prepared for the current building energy simulation according to the generated

building characteristics from step 2). The text file inputs have the building parameters, such as the dimensions of the house or the thermal properties of the wall materials, substituted for each house generated by the model. For the complete building stock model this represents 200 000 different sets of inputs.

4) **TRNSYS simulation**: an annual building energy simulation for each house is performed and energy consumption values are output at 15-minute intervals. Input characteristics for each house are also retained and matched with the energy consumption.
5) **Worker data**: when the desired number of house simulations is reached, the worker data is saved.
6) **Postprocessing**: worker data are compiled into a single data set. Data is categorized by end-use and energy source. Total stock electricity data is retained at 15-minute intervals while natural gas, heating oil and wood energy consumption is saved annually.

The Matlab software is used as a platform to distribute the building energy simulations to individual workers, generate the building input files and launch the TRNSYS building energy simulations (Mathworks Inc. 2018). The building parameters and their interdependencies are illustrated in Figure 3. The result is a flexible building stock energy model that can generate virtually any number of homes according to the input parameters provided to the model. Due to the probabilistic combination of building parameters, the generated homes are representative of real buildings in the stock and collectively represent the overall energy use of the building stock.

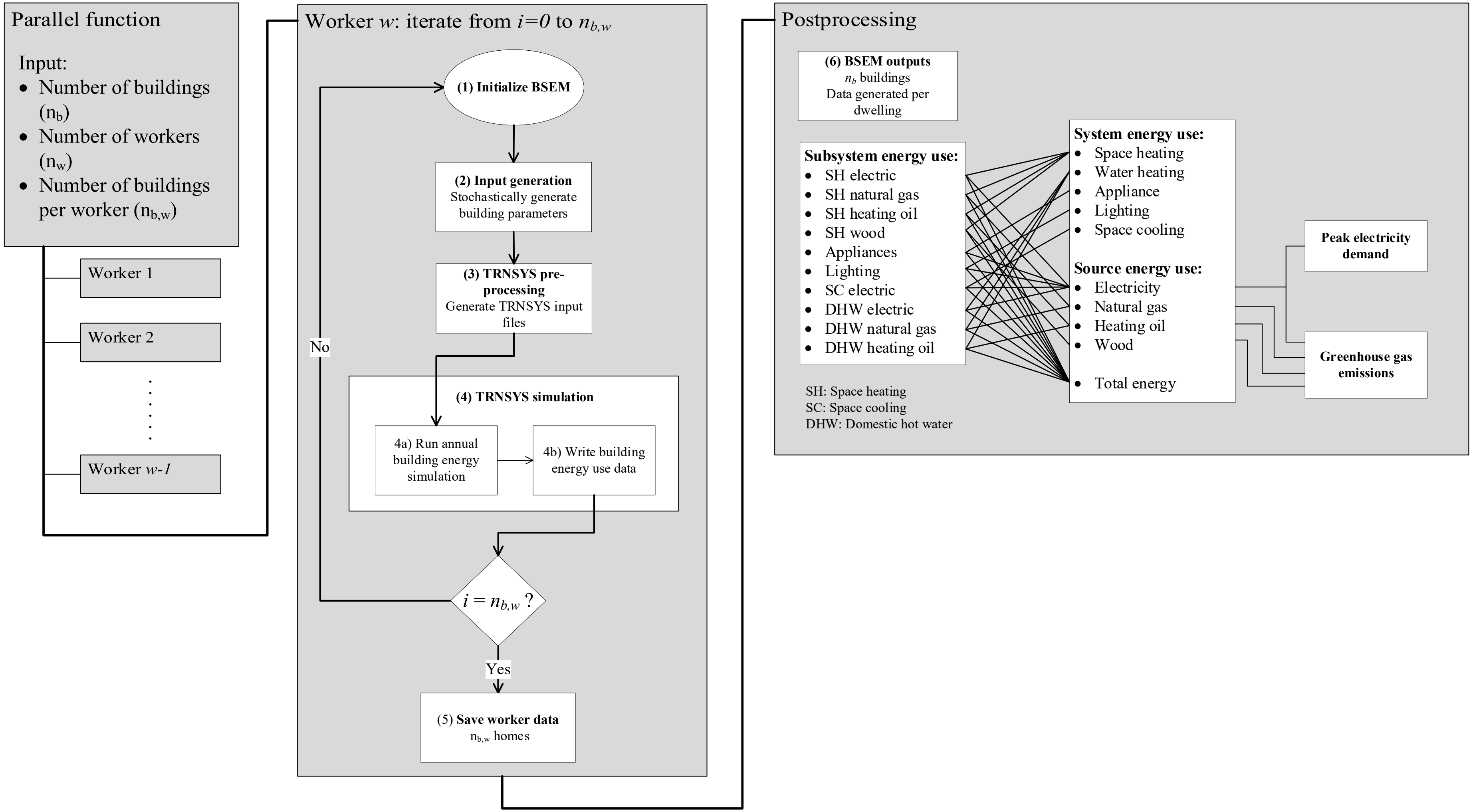


Figure 2. General building stock energy modeling approach and model outputs

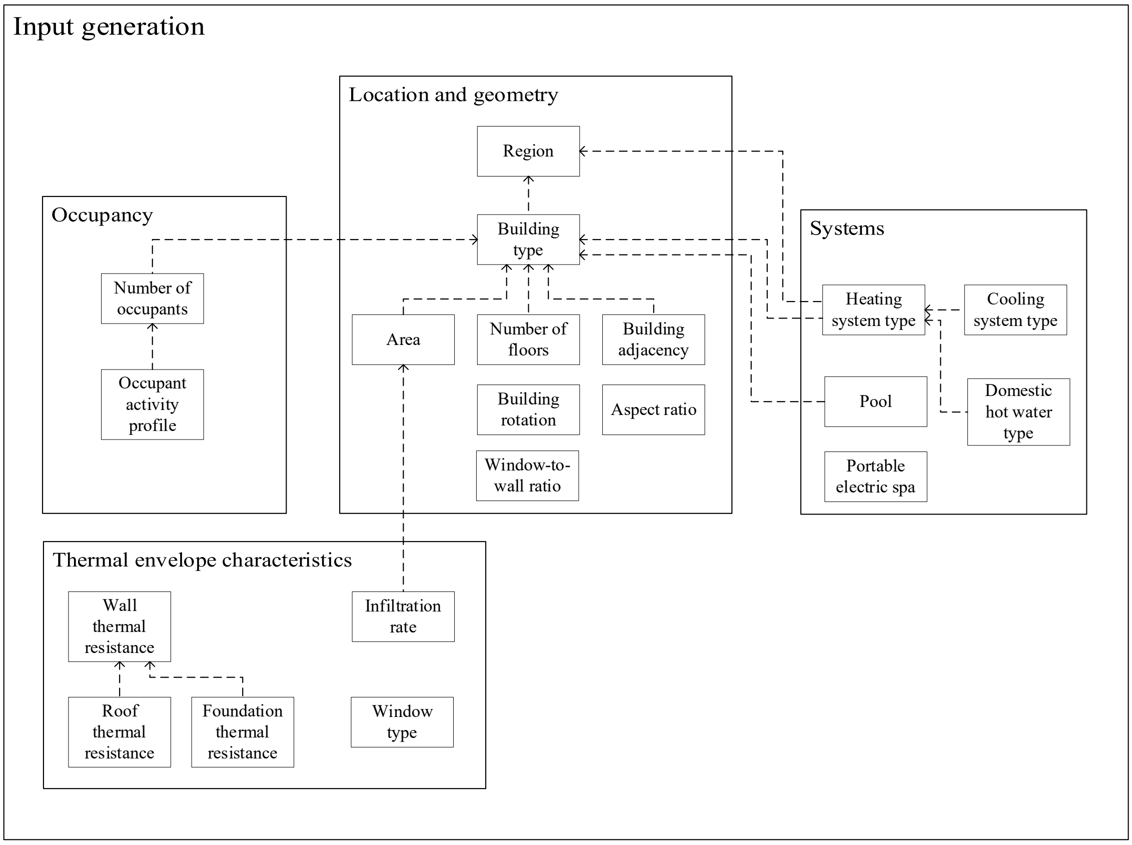


Figure 3. Building parameters generated for each house. Dependencies are illustrated with dashed arrows.

As mentioned previously, the modeling principles in TRNSYS and the approach for specific aspects of the building energy simulation, such as the internal loads, were described in another publication (Neale, Kummert, and Bernier 2020). Some features are described in brief detail in Table 2 for clarity, and interested readers can refer to the original publication for more detail. New or updated building characteristics are presented in Section 4 of the present paper.

Table 2. Overview of modeling choices for the building energy simulation of each house.

Model aspect	Detail	Reference (if applicable)
Geometry	Each house is modeled as a single-zone dwelling with a finished basement. Wall surfaces shared with adjacent dwellings, such as in row houses, are considered adiabatic. The window-to-wall ratio is applied on all aboveground wall surfaces not shared by another dwelling. House size is based on real house data.	
Building envelope characteristics	Wall, roof and foundation thermal resistance are applied separately to their corresponding surfaces. Infiltration rate is calculated using the Sherman-Grimsrud method. Window type is selected from a list of 24 single-, double-, and triple-glazed models.	Sherman and Grimsrud (1980)
Internal loads	Internal loads due to occupancy are produced using the CREST stochastic occupancy model, which includes the lighting and appliance loads. Loads are applied as fractional convective and radiative heat gains to the indoor environment.	McKenna and Thomson (2016)
Systems	Domestic hot water energy demand profiles are generated separately from the stock simulation using the TRNSYS software. Hot water use is determined using the CREST stochastic occupancy model as a function of the number of occupants currently active in the home.	McKenna and Thomson (2016)
	Heating systems are assumed to have infinite capacity unless otherwise specified in Table 7. Other characteristics of the heating systems are applied to address the building heating load as described in Table 7.	
	Cooling systems are assumed to have infinite capacity.	

4 Building stock characterisation

The segmentation (sometimes called classification) and characterisation processes are often used for building archetype development (Sokol et al. 2016), which is a bottom-up white-box technique commonly used for building stock modeling (Swan and Ugursal 2009; Langevin et al. 2020). Segmentation is the process of determining the parameters that differentiate different types of buildings, such as climate zones, house types, etc. Characterisation is used to identify the range of values of each parameter given the building stock composition, such as detached and semi-detached houses for the *house type* parameter.

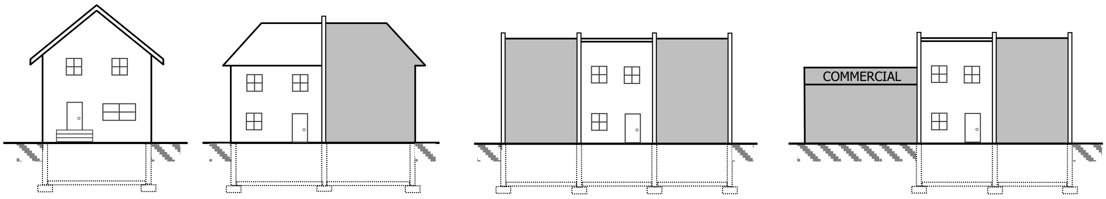
The characterisation process is applied to single-family dwellings in the province of Québec to generate accurate houses that fit the range of parameters found in the building stock. Characteristics of the housing stock, such as the types of dwellings, number of dwellings across the province, heating and cooling systems are described in detail. The data collected serves to establish probability distributions used to generate combinations of parameters that exist within the building stock.

4.1 Dwelling types

The province of Québec, Canada, is characterized by a residential building stock consisting of a mix of single-family dwellings (SFD) and multi-residential dwellings (MRD). From a modeling point of view there are distinct differences between SFD and MRD, such as the boundary conditions above and below a detached home are significantly different from an apartment in a multi-residential building. This study aims to develop a stock model for SFD only, and therefore targets the following types of homes:

- Single-detached homes (Det)
- Semi-detached homes (Semi)
- Row houses (Row)
- Other single-attached (OSA)

The four building types in the list above, depicted in Figure 4, are commonly used by national statistics and energy data publications in Canada to present dwelling distributions and energy consumption data. Occasionally reference data is presented for “Detached” and “Attached” homes, in which case Semi, Row and OSA homes are combined for the latter category as they all share external boundaries with adjacent buildings.



Detached Semi-detached Row Other single-attached

Figure 4. Single-family dwelling (SFD) types in the province of Québec, Canada.

Houses in the studied building stock are most commonly one or two-storeys plus a heated basement. The number of storeys is dependent on the type of dwelling, with Semi, Row and OSA more commonly having two-storeys (NRCan 2018). The heated surface area of the dwelling also depends on the building type and number of storeys.

4.2 Number of dwellings by region

There is a variety of sources for data related to the number of residential dwellings in the Province of Québec, Canada. The Canadian Census of Population Program (CCPP) from Statistics Canada is a reliable source for *Type of Dwelling* data since 2016, where census responders indicate relevant characteristics of the home that they live in (Statistics Canada 2016). The CCPP divides Canada into a number of Census Metropolitan Areas (CMA), which are population hubs of more than 100 000 people of one or more neighbouring municipalities where at least 50 000 individuals live in the urban core, and Census Agglomerations (CA), which have a core population of at least 10 000 (Statistics Canada 2016). The province of Québec has 6 CMA regions and 24 CA regions, which are illustrated on the population density map in Figure 5 (Statistics Canada 2019).

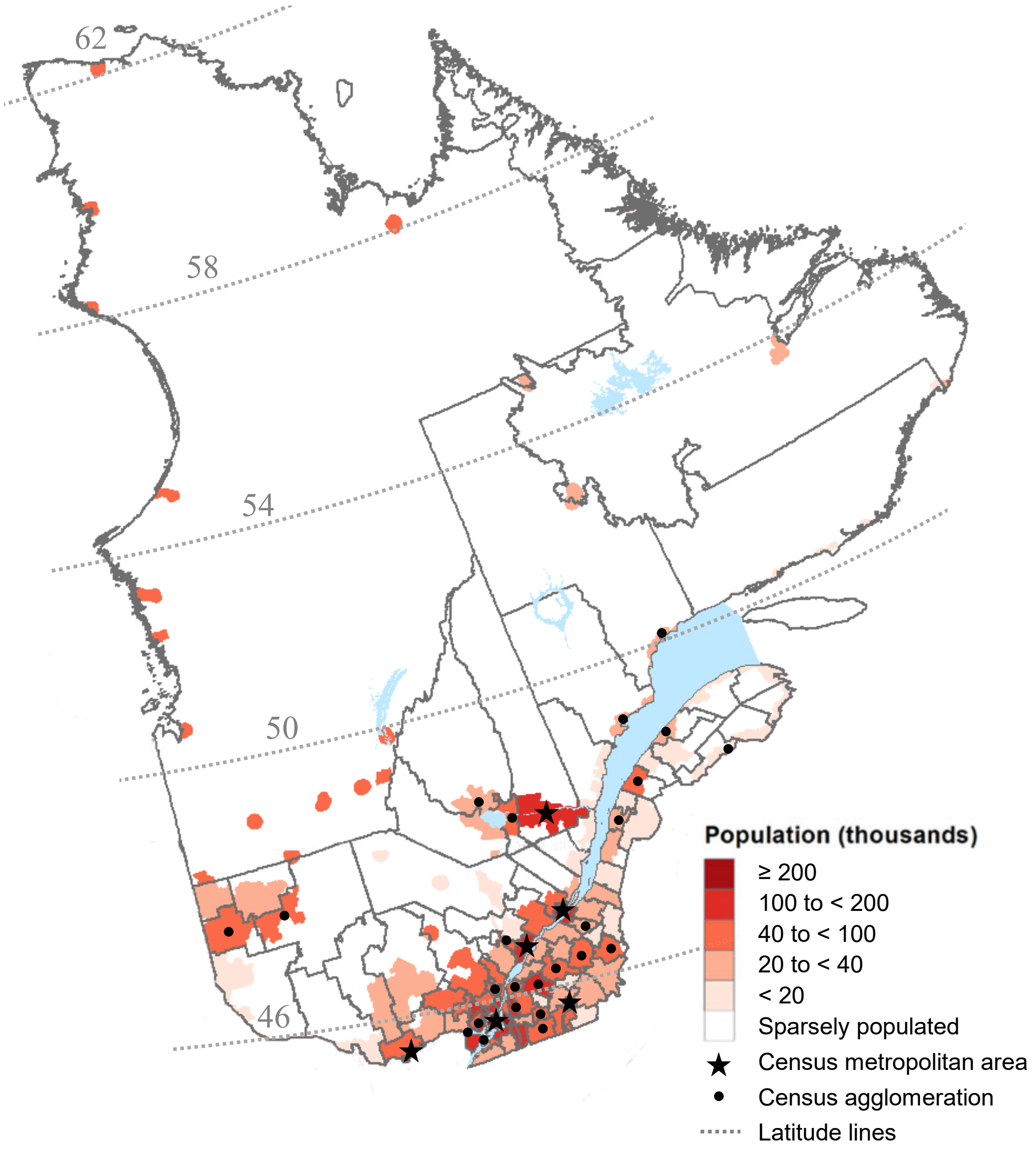


Figure 5. Population distribution of the province of Québec, Canada, with CMA and CA regions superimposed. Adapted from Statistics Canada (2019)*. Approximate latitude lines are indicated for reference.*

The six CMA regions represent significant concentrations of the population and, consequently, of dwellings. Energy use data for the province of Québec is often expressed in terms of these CMA regions, with a seventh region entitled "Québec non-CMA", which represents the remainder of the province (NRCan 2015). The distribution of single-family dwellings (SFD) across these seven regions is presented in Table 3 (Statistics Canada 2016).

Table 3. Distribution of occupied dwellings for Québec CMA areas. DF: dwelling fraction, SFD: single-family dwelling, Det: single-detached house, Row: row house, Semi: semi-detached house, OSA: other single-attached house (Statistics Canada 2016).

	Number of dwellings by region					**Dwelling fraction (DF) by region**				
Regions	**SFD**	**Det**	**Row**	**Semi**	**OSA**	DF_{SFD}	DF_{Det}	DF_{Row}	DF_{Semi}	DF_{OSA}
R1 Québec Non-CMA[1]	780,600	670,777	22,342	78,255	9,225	0.4108	0.8593	0.0286	0.1003	0.0118
R2 Saguenay	44,195	36,900	5,455	1,580	260	0.0233	0.8349	0.0358	0.1234	0.0059
R3 Québec City	180,380	148,965	21,135	8,935	1,345	0.0949	0.8258	0.0495	0.1172	0.0075
R4 Sherbrooke	49,780	42,630	4,320	2,455	375	0.0262	0.8564	0.0493	0.0868	0.0075
R5 Trois-Rivières	40,710	33,720	4,785	1,845	360	0.0214	0.8283	0.0453	0.1175	0.0088
R6 Montréal	713,710	564,230	86,460	56,770	6,250	0.3756	0.7906	0.0795	0.1211	0.0088
R7 Gatineau	90,845	64,710	17,890	7,955	290	0.0478	0.7123	0.0876	0.1969	0.0032
Total SFD	1,900,220	1,602,675	188,245	93,355	15,945					

[1] The dwelling fraction values for Québec non-CMA are extrapolated from the number of houses for the 24 census agglomerations regions, i.e. 230 000 dwellings are extrapolated to represent 780,600.

The seven identified regions represent 100% of the single-family dwellings in the province of Québec, Canada, which is approximately 1.9 million houses (2017 data). The value of DF_{SFD} represents the fraction of all single-family dwellings in each region. The following columns, DF_{Det} to DF_{OSA}, respresent the corresponding fraction of dwellings of that type for that region. As an example, 9.49% of single-family dwellings are located in R3 Québec City, and of those 82.6% are detached houses.

The 24 CA regions represent approximately 230 000 dwellings, which is 31% of the Québec non-CMA amount of 780 600 (Statistics Canada 2016). The remainder of the homes are located in less densely populated areas in relative proximity to the CA and CMA regions or in small settlements in the far north of the province. For the purpose of this study, the 24 CA regions identified in Figure 5 are considered to provide an accurate representation of the Québec non-CMA portion of the building stock, as they are the highest concentration of buildings outside of the metropolitan areas. The distribution of homes in the 24 CA regions is therefore extrapolated to be equal to the total number of homes in the remainder of the province. The description of the 24 sub-regions of R1 is provided in Appendix 1 in Table A-1. The results of the 24 sub-regions are aggregated and referred to as *Québec non-CMA* for the remainder of this study.

4.3 Weather data

Building energy models often use typical weather files to standardize the heating and cooling loads and to remove the annual fluctuation in conditions found in real weather. The Canadian Weather year for Energy Calculation (CWEC) files serve that purpose for Canadian regions (Government of Canada 2021). CWEC files contain 12 distinct typical meteorological months that are selected from a database of 30 years of weather data. Each combination of months is distinct for each region, which results in months selected from different years of data for each weather file. When modeling a building stock with multiple climate zones, non-coincidental peaks pose an issue as the peak energy use does not occur at the same moment for each region. As an example, the outdoor dry bulb temperature from CWEC files for the six CMA regions are illustrated in Figure 6 (left) for the month of December. Each region is identified with the corresponding year that the month of December is taken from, i.e. R2-2009 indicates the calendar year 2009 for region R2.

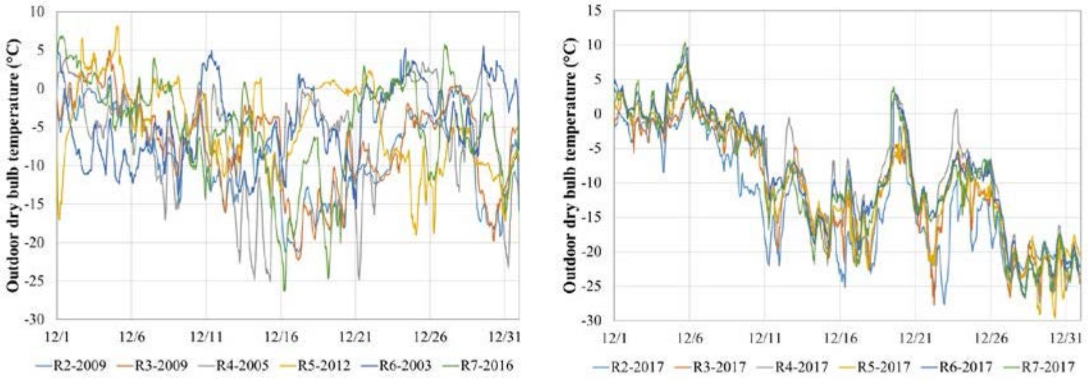


Figure 6. Outdoor dry bulb temperature for 6 CMA locations across the province of Québec for CWEC weather data (left) and 2017 CWEEDs data (right)

The temperature curves in Figure 6 (left) appear to be independent, except for R2 and R3, which both use the 2009 data for the month of December. For the purpose of comparison, the Canadian Weather Energy and Engineering Datasets (CWEEDs) data for 2017 is depicted in Figure 6 (right) for the same six CMA regions of the province (Government of Canada 2021). While there are regional differences, the general temperature trend is the same across the six illustrated regions for the CWEEDs data. For the purposes of modelling a building stock's energy consumption and peak load, a set of coincident weather data for the relevant regions is necessary for accurate prediction of the peak heating load. In addition, there is energy consumption data for validation purposes for specific years. After comparing multiple years of weather data, the calendar year 2017 is selected for the results presented in this paper, though the model can function with any year of weather data. For comparison purposes, the heating degree-days (HDD@18°C) for the 2017 CWEEDs data vary between approximately 4000 and 5400 degree-days across the studied building stock.

4.3.1 Weather for R1: Québec Non-CMA

Aside from the six census metropolitan areas, which represent a significant portion of the building stock, a seventh region is required to represent the remainder of the homes in the province. The 24 census agglomeration regions are selected to divide the remaining buildings into populated regions spread across Québec, as illustrated in Figure 5. The closest weather station to each CA region was identified geographically. A house in *Québec non-CMA* is therefore assigned to one of the 24 weather stations according to the probability distribution described in Table A-1 in Appendix 1.

4.3.2 CWEEDs missing cloud cover data

In the process of analysing the CWEEDs weather data for the 30 regions of this study it was discovered that Environment Canada no longer recorded Total Sky Cover or Opaque Sky Cover for most weather stations across the province of Québec from 2013 onwards. In some cases, such as at certain airports, Total Sky Cover was recorded but only at 3-hour intervals. The TRNSYS building energy simulation software used in this study requires the Opaque Sky Cover in order to determine the sky temperature for longwave radiation calculation (Klein et al. 2017). In order to correct this issue a methodology is applied to fill the Total and Opaque Sky Cover values in the weather files used for the building stock model, which is presented in Appendix 2.

4.4 Building envelope characteristics

The construction year of a home is not a strong indicator of the building envelope characteristics for the studied building stock. The building envelope is therefore better characterized directly by wall, roof, foundation thermal resistance, window type and leakage area, as opposed to construction vintage. Probability distributions for house characteristics related to the building envelope are determined from the Energuide Housing Database containing over 27 000 homes in the studied building stock (NRCan 2018).

4.5 System characterization

The single-family residential building stock studied in this paper relies on a variety of space heating, space cooling and water heating technologies. The characterization process for these systems is described in the following sections of the paper, which will be used for modeling purposes in the building energy simulation program.

4.5.1 Space heating

The data related to the prevalence of different heating technologies for attached and detached dwellings for the studied building stock is presented in Table 4, which originates from the Canadian Comprehensive Energy Use Database (CEUD) (NRCan

2017). Systems representing less than 1% of the building stock are excluded, whether for attached or detached dwellings. Heating systems are labelled from H1 to H8.

Table 4. Heating system distribution for the Province of Québec for single detached and single attached homes (NRCan 2017)

Primary heating system description		Detached		Attached	
		#units (thousands)	Probability	#units (thousands)	Probability
H1	Heating Oil – Medium Efficiency	134.5	0.078	28.0	0.085
H2	Natural Gas – High Efficiency	52.1	0.030	10.8	0.033
H3	Electric	741.3	0.432	227.4	0.689
H4	Heat Pump	212.3	0.124	16.6	0.050
H5	Wood/Electric	370.3	0.216	10.5	0.032
H6	Wood/Heating Oil	94.9	0.055	4.1	0.012
H7	Heating Oil/Electric	111.4	0.065	16.2	0.049
H8	Wood	0.0	0.000	16.5	0.050

The heating system data from (NRCan 2017) provides an average distribution across all regions for the studied building stock. There are significant differences between the detached and attached heating system distributions, most notably in the increased prevalence of electric baseboard heating (H3) in attached houses and the lack of wood-heated (H8) detached houses. In addition, the Survey of Household Energy Use of 2015 includes data on primary heating system energy type by region, illustrated in Table 5 (NRCan 2015).

Table 5. Fraction of homes by region based on primary heating energy (NRCan 2015)

	Electricity	Natural gas	Wood	Unknown[1]
R1 Québec Non-CMA	0.763	-	0.178	0.060
R2 Saguenay	0.920	-	-	0.080
R3 Québec City	0.870	-	-	0.130
R4 Sherbrooke	0.830	-	-	0.170
R5 Trois-Rivières	0.897	-	-	0.103
R6 Montréal	0.882	-	-	0.118
R7 Gatineau	0.289	0.623	-	0.088

[1] Data missing or unaccounted for

While there are significant gaps in the SHEU data set in Table 5, there are several important details that can be used to complement the data presented in Table 4. For example, there is a prevalence of natural gas heating in *R7 Gatineau* that is uncharacteristic of the remainder of the province, and wood heating is more common in the outlying regions of *R1 Québec Non-CMA*. Electricity-based systems are considered to be distributed according to the *electricity* fraction in Table 5. The *unknown* column represents missing data and is presumed to belong to heating systems not represented in the included data. As an example, for R2 Saguenay the 0.08 fraction of missing data is distributed among systems without electric primary heating, i.e. H1, H2, H6 and H8. The combination of the probability distributions in Table 4 and Table 5 result in regional probability tables for the heating systems of detached and attached buildings in Table 6.

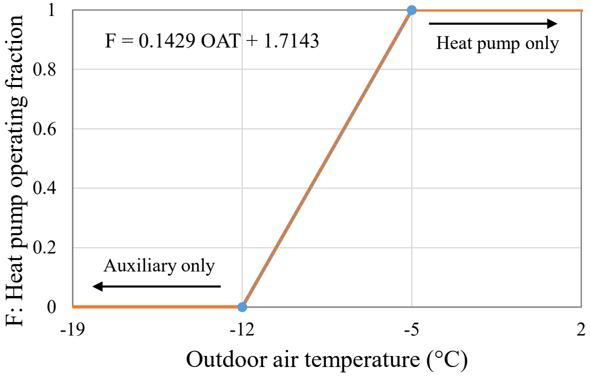
In order to illustrate how the values in Table 6 are determined, consider the case for R1 (*Québec Non-CMA*) for single-detached homes. The fraction of electric-heated homes for R1 is 0.763 from Table 5. Heating systems using electricity (H3, H4, H5 and H7) are distributed based on the weighted distribution in Table 4. Wood systems represent a fraction of 0.178, though only hybrid wood systems are found in detached homes for the studied building stock. Therefore the wood/heating oil systems are assumed to represent 0.178 of systems in R1, since wood/electric heating is considered primarily an electric system and there are no wood-only systems (H8) in R1. The remaining heating systems (H1 and H2) are therefore distributed according to the *unknown* fraction 0.060 from Table 5 according to their respective probabilities.

Table 6. Heating system probability by region. Highlighting by data source from Table 5: electric, natural gas, wood or unknown.

	Heating system fraction by region							
Region	H1 Heating oil	H2 Natural gas	H3 Electric	H4 Heat pump	H5 Wood/ Electric	H6 Wood/ Heating oil	H7 Heating oil/ electric	H8 Wood
Single detached								
R1	0.043	0.017	0.394	0.113	0.197	0.178	0.059	0.000
R2	0.038	0.015	0.475	0.136	0.237	0.027	0.071	0.000
R3	0.062	0.024	0.450	0.129	0.225	0.044	0.068	0.000
R4	0.081	0.032	0.429	0.123	0.214	0.057	0.064	0.000
R5	0.049	0.019	0.463	0.133	0.231	0.035	0.070	0.000
R6	0.056	0.022	0.456	0.130	0.228	0.040	0.068	0.000
R7	0.052	0.623	0.149	0.043	0.074	0.036	0.022	0.000
Single attached								
R1	0.043	0.017	0.641	0.047	0.030	0.035	0.046	0.142
R2	0.038	0.014	0.773	0.056	0.036	0.006	0.055	0.022
R3	0.061	0.024	0.731	0.053	0.034	0.009	0.052	0.036
R4	0.080	0.031	0.697	0.051	0.032	0.012	0.050	0.047
R5	0.049	0.019	0.754	0.055	0.035	0.007	0.054	0.029
R6	0.056	0.021	0.741	0.054	0.034	0.008	0.053	0.033
R7	0.051	0.623	0.242	0.018	0.011	0.007	0.017	0.030

The eight identified heating systems operate according to the descriptions provided in Table 7. Single energy source systems are assumed to cover the entire heating load of a dwelling. Hybrid systems are attributed specific fractional loads based on the prevalence of those types of systems in the studied province. For example, favourable electricity rates are provided to homeowners with hybrid electric and heating oil systems when the outdoor temperature is below -12°C (Hydro-Québec 2021), and therefore these types of systems are assumed to transition between systems at low temperatures.

Table 7. Detailed heating system descriptions. O: heating oil, NG: natural gas, E: electric, W: wood. OAT: Outdoor air temperature.

Heating system		**Description**	**System details**	**Secondary electricity use**	**System efficiency**
H1	Heating Oil – Medium Efficiency	Heating oil boiler or furnace	Heating load 100% covered by heating oil-fired system.	2.39%[1]	O: 0.78[2]
H2	Natural Gas – High Efficiency	Natural gas boiler or furnace	Heating load 100% covered by natural gas-fired system.	2.39%[1]	NG: 0.90[2]
H3	Electric	Electric baseboard or electric furnace	Heating load 100% covered by electric heating element system, i.e. baseboard heating or electric furnace.	N/A	E: 1.00[2]
H4	Heat Pump	Air-source heat pump with auxiliary electric element	OAT above -5 °C: Heating load 100% covered by heat pump. OAT below -12 °C: Heating load 100% covered by electric heating elements. OAT between -5 °C and -12 °C: linear fraction transitioning between the two systems as a function of temperature.  	N/A	E: 1.00[2] Heat pump[3]: COP = 0.0585 OAT + 3.115
H5	Wood/Electric	Electric baseboard or electric furnace with wood stove or fireplace	OAT below 0 °C: wood-fired system (i.e. fireplace or stove) contributes up to 40,800 kJ/h for detached houses and 23,800 kJ/h for attached houses to the heating load, with the remainder covered by electric heating elements.[1] OAT above 0 °C: heating load 100% covered by electric heating elements.	N/A	W: 0.50[2] E: 1.00[2]

Heating system		Description	System details	Secondary electricity use	System efficiency
H6	Wood/Heating Oil	Heating oil boiler or furnace with wood stove or fireplace	OAT below 0 °C: wood-fired system (i.e. fireplace or stove) contributes up to 40,800 kJ/h for detached houses and 23,800 kJ/h for attached houses to the heating load, with the remainder covered by heating oil system. OAT above 0 °C: heating load 100% covered by heating oil system.	2.39%[1]	W: 0.50[2] O: 0.78[2]
H7	Heating Oil/Electric	Heating oil boiler or furnace with electric baseboard or furnace	Considered a hybrid system following hybrid electricity rate operation[4], transitioning at low outdoor temperatures to non-electric systems. OAT below -12 °C: heating load 100% covered by heating oil system. OAT above -12 °C: heating load 100% covered by electric heating elements.	2.39%[1]	O: 0.78[2] E: 1.00[2]
H8	Wood	Wood stove, fireplace or furnace	Heating load 100% covered by wood-fired system.	N/A	W: 0.50[2]

[1] Based on 7 kWh/MMBtu average (NYSERDA 2013)
[2] NRCan (2017)
[3] Johnson (2013)
[4] Hydro-Québec (2021)

4.5.2 Space cooling

The presence of an air conditioning (AC) system varies by region according to the Survey of Household Energy Use of 2015 (NRCan 2015). The SHEU data set has data indicating the presence of a central air conditioner, window air conditioner or no air conditioner. However, since the data for homes with central versus window AC is unreliable or missing, the probability is expressed simply as whether AC is present or not for each region.

Table 8. Air conditioning prevalence in the studied building stock (NRCan 2015).

Region		**AC**	**No AC**
R1	Québec Non-CMA	0.376	0.624
R2	Saguenay, Québec	0.539	0.461
R3	Québec, Québec	0.366	0.634
R4	Sherbrooke, Québec	0.596	0.404
R5	Trois-Rivières, Québec	0.630	0.370
R6	Montreal, Québec	0.680	0.320
R7	Ottawa-Gatineau, Ontario/Québec	0.879	0.121

In addition, if a home has a heat pump heating system (H4), the house is presumed to have air conditioning as most residential heat pumps are reversible. Finally, the average Energy Efficiency Rating (EER) of cooling systems in the province of Québec is EER=14 (NRCan 2017), which is applied as a constant coefficient of performance (COP=4.1).

It should be noted that the data from SHEU for air conditioning is for all residential dwellings, including multi-residential apartment buildings. Other data sources related to air conditioning exist, but the SHEU data is retained because it differentiates the probability of a cooling system by region, which has a significant impact on the distribution of air conditioners in the province.

4.5.3 *Water heating*

Water heating for domestic purposes is predominantly electric for the studied building stock. The probability of a home's domestic hot water (DHW) energy source is presented in Table 9 (NRCan 2017). Water heating systems using the *Other* category are negligible for the studied building stock and are not represented in the model.

Table 9. Domestic hot water system distribution by energy source (NRCan 2017).

Water heater type	**P_{DHW}**
Electric	0.930
Natural gas	0.042
Oil	0.021
Other	0.007

The domestic hot water fuel type is dependent on the primary heating energy source. Homes with primary electric heating systems are 99.2% likely to have electric water heating (NRCan 2018). Dwellings with non-electric primary heating often have non-electric DHW systems, though this is not universally true. Homes with natural gas-based heating systems have a 66.1% probability of having natural gas water heating, and houses with heating oil-based heating systems have 22.5% probability of having heating oil domestic hot water heating (NRCan 2018).

4.6 Occupancy

The number of occupants for the studied building stock is based on the type of building, which in turn depends on the region where the house is located. Due to the limitations of the stochastic occupancy model used to represent the internal loads of each house, the maximum number of occupants is set to 5 (McKenna and Thomson 2016). Additional details on the modeling of internal loads are presented in Section 3 of this paper.

5 Stock model results

The energy consumption by end-use and energy source is available for detached and attached buildings in the studied building stock (NRCan 2017). The end-use values include space heating, water heating, appliances, lighting and space cooling, while energy sources include electricity, natural gas, heating oil and wood. The single-family dwelling building stock combines detached and attached dwellings to include approximately 1.9 million homes in total, which are represented by a sample of 200 000 simulated dwellings. The exact number of houses used to represent the stock has an impact on the total energy consumption, which is discussed in more detail in Appendix 4. With a subset of 200 000 dwellings, the expected deviation on the total stock energy consumption due to the random generation of the building characteristics is under 1%.

Considering the three dwelling categories and ten energy consumption values, this provides a total of 30 points of comparison for the building stock energy model, which are illustrated in Figure 7.

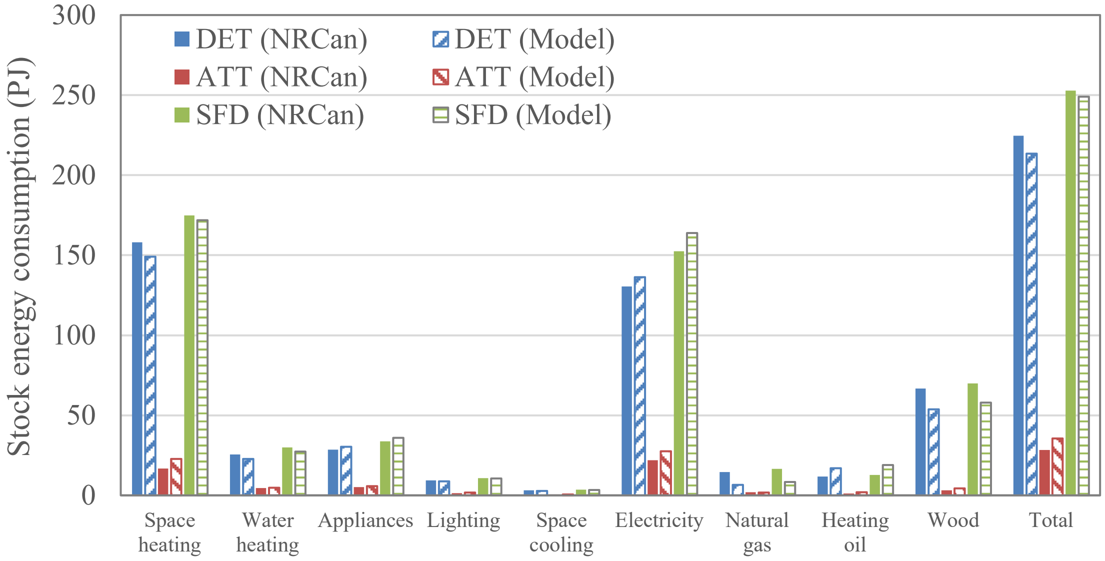


Figure 7. Model versus stock energy consumption for detached (DET), attached (ATT) and all single-family dwellings (SFD)

The total stock energy consumption values are illustrated in petajoules in Figure 7. There is good agreement for the end-use categories (space heating, water heating, appliances, lighting and space cooling). There are some differences in the natural gas, heating oil and wood categories, which predominantly contribute to space heating and water heating. Overall, the model provides total stock energy consumption within 1.58% of the stock data (NRCan value of 252.9 PJ compared to the model value of 248.9 PJ). In order to visualize these differences further, the modeled per-house energy consumption for all single-family dwellings is illustrated as a box and whisker plot in Figure 8.

The box in Figure 8 illustrates the interquartile range for the energy consumption by category for the 200 000 homes of the building stock energy model sample. Outliers are indicated as red data points. Most houses with natural gas and heating oil are considered outliers due to the fact that only a small percentage of homes are heated with those systems, and therefore most houses in the studied stock have zero natural gas and heating oil energy consumption. As a further example, the *Space heating* box plot illustrates the mean space heating energy consumption of the SFD stock as a black X symbol, the mean space heating energy according to known stock data (NRCan 2017) as a semi-transparent blue circle.

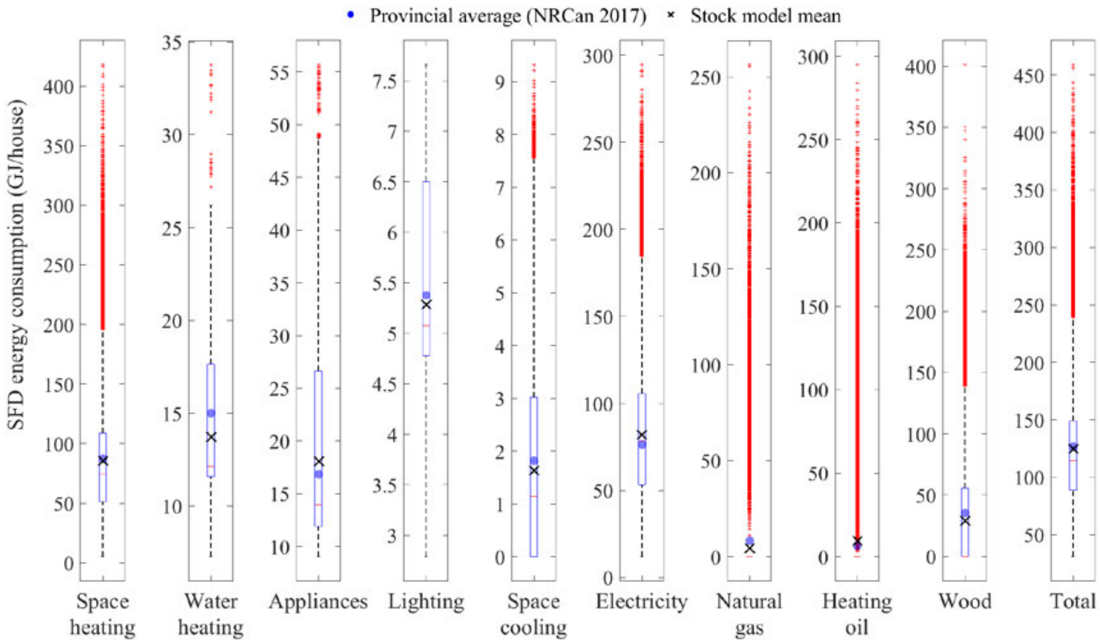


Figure 8. Box and whisker plot of the modeled single-family dwelling (SFD) energy consumption by end-use and energy source

Overall, the proposed building stock energy model has good agreement across the 30 studied end-use and source energy consumption categories. Further improvements could be made in the future to adjust the heating system distributions and/or refine the validation energy consumption values as additional information becomes available. The authors are satisfied that the proposed building stock energy model is a close representation of single-family dwellings in the province of Québec, Canada.

6 Case study

The government of Québec has the objective of reducing greenhouse gas emissions (GHG) related to space heating in buildings by 50% for 2030 (Government of Québec 2020). While the proposed building stock model is not a long-term energy projection model, a comparative assessment between two or more stock configurations can be performed. In order to study the effect of energy consumption changes on GHG emissions, the emissions factors for the province of Québec are first presented. The proposed scenarios are then compared, illustrating the impact of changes to the single-family dwelling space heating market on the energy consumption, GHG emissions and peak electricity use.

6.1 Greenhouse gas emission factors

As described previously, energy use in single-family dwellings in the province of Québec, Canada, consists primarily of electricity, natural gas, heating oil and wood. These energy sources each have distinct CO_2 equivalent emission factors, which are provided in Table 10 for electricity, natural gas, heating oil and wood.

Table 10. Greenhouse gas emission factors for energy sources in the province of Québec

Energy source	$gCO_2eq•kWh^{-1}$	Ref.
Electricity	2.0	TÉQ (2019)
Natural gas	178.3	TÉQ (2019)
Heating oil	254.9	TÉQ (2019)
Wood	84.5	NRCan (2017)

The emissions factors in Table 10 are considered constant for the entire year, with the exception of electricity. Electricity production in the province of Québec is predominantly hydroelectric and is estimated to generate CO_2 equivalent emissions at the rate of 2.0 $gCO_2eq•kWh^{-1}$ (Transition Énergétique Québec 2019). During peak electricity usage hours, a non-negligible portion of the electricity in the province is imported from neighbouring provinces and states, which is called *short-term imported electricity* and is illustrated in red in Figure 9 for December 2017 (Régie de l'énergie 2017).

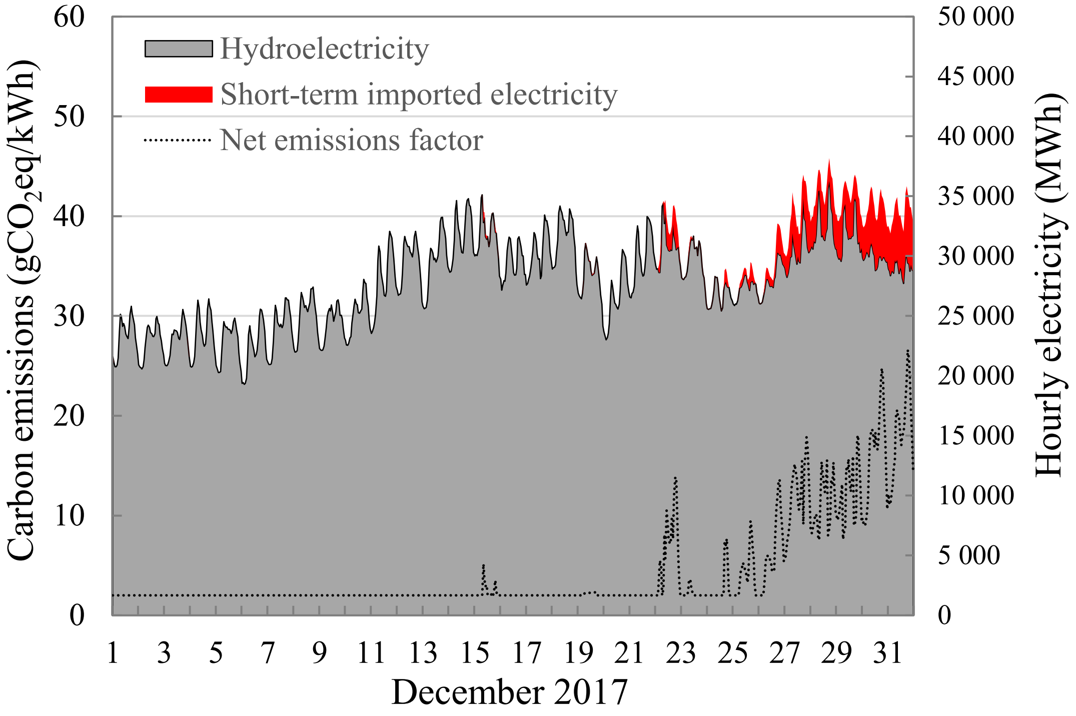


Figure 9. Calculated CO_2 equivalent emission rates in December 2017 for electricity in the province of Québec and electricity usage by source for all sectors.

According to the local electricity distributor, 97.5% of short term electricity purchases are from Ontario (Hydro-Québec 2017). To simplify the GHG case study analysis, the short term electricity imports for Québec are calculated using marginal seasonal emission factors for the province of Ontario, which vary hourly between 90 and 150 $gCO_2eq•kWh^{-1}$ (The Atmospheric Fund 2019). Combining the base electricity emissions factor of 2.0 $gCO_2eq•kWh^{-1}$ for local hydroelectricity with the short term imported electricity from Ontario results in a net emissions factor that varies approximately between 2.0 and 26.0 $gCO_2eq•kWh^{-1}$, depending on the day and time of year (Figure 9).

While a large increase to the peak electricity usage would theoretically result in more short-term imports, in reality sweeping changes to the building stock would not occur in a short time frame. It is likely that the local electricity distributor would adjust the hydroelectric production to account for any large increases to the stock electricity use. Therefore, emissions rates for additional electricity use are calculated at the rates as the current stock, as described above.

6.2 Case study: 50% reduction in GHG emissions for space heating

In order to study the impact of reducing GHG emissions due to space heating, three scenarios are compared:

1) Base case: the status quo single-family building stock for the province of Québec, Canada.
2) Scenario 1: 50% of heating systems with non-electric energy sources are converted to baseboard electric heating (heating system H3, Table 4), a common cheap alternative widely used in the province of Québec.
3) Scenario 2: 50% of heating systems with non-electric energy sources are converted to cold climate heat pump heating systems.

Cases are compared based on the annual space heating energy consumption, the GHG emissions for each energy source and the maximum peak load. Heating system distributions are updated by modifying the probabilities presented in Table 6, shifting heating systems with non-electric energy sources to electric systems as described in the scenario descriptions. This is accomplished by modeling different sets of 200 000 houses, each generated with specific probability distributions. In the case of Scenario 2, the coefficient of performance (COP) is implemented as a function of outdoor air temperature based on measured data (Johnson 2013). The total building stock energy for each hour of the year is then compared for the studied cases.

The total stock emissions and space heating energy consumption are presented for the base case and two scenarios in Figure 10. Total emissions and energy consumption are similar to the reference data by NRCan (2017). Most of the emissions in the province originate from the non-electric energy sources, and therefore reducing the heating systems using those energy types by 50% has the desired effect of reducing overall emissions by 48%. In terms of the energy consumption for Scenarios 1 and 2, reductions in total space heating energy use for single-family dwellings reach 10% and 21%, respectively. These decreases in space heating energy are largely due to the improvements in heating system performance when comparing electric systems to their nonelectric counterparts.

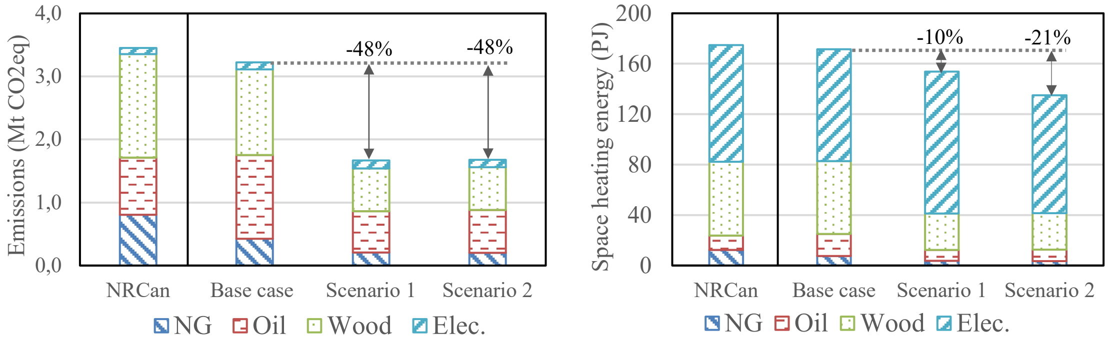


Figure 10. Annual CO_2 equivalent emissions and total space heating energy consumption for the studied cases. Stock reference data from Natural Resources Canada is also provided for comparison (NRCan 2017).

The difference in peak loads between the base case and Scenarios 1 and 2 are compared in Figure 11, which are illustrated for the month of December 2017. While emissions and annual space heating energy are reduced for both scenarios, the peak electricity use increases significantly. For Scenario 1, shifting nonelectric heating to baseboard heaters has the result of increasing the peak load over the normal SFD stock value by approximately 35%, or the equivalent of 4000 MW. For Scenario 2, the improved efficiency of the heat pumps mitigate the impact on the peak load, but still increase it by up to 20%, or approximately 2500 MW. During milder periods, the electricity load can occasionally decrease below the normal level of the base case scenario due to the higher efficiency of the heat pumps at milder temperatures. However, decreasing the load at other moments of the year does not aid the electricity distributor, as the maximum electricity production capacity of the stock is sized for the peak usage of the province, which would increase under both scenarios.

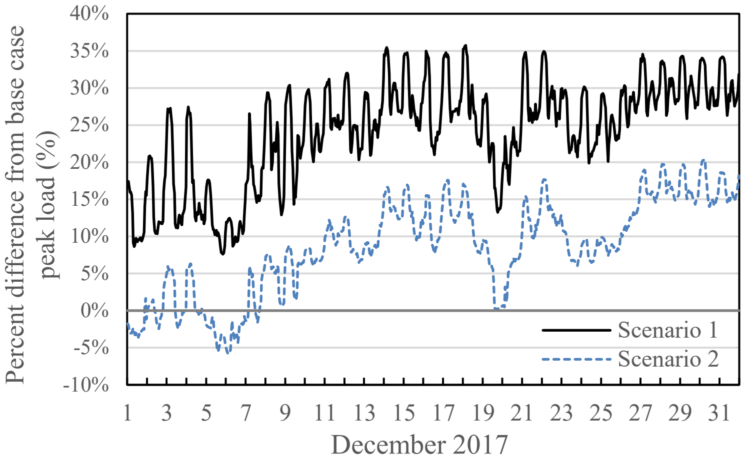


Figure 11. Peak load percent difference for Scenarios 1 and 2 with respect to the base case scenario.

The case study illustrates how the QSFBSEM can be used to compare greenhouse gas scenarios for different heating system distributions across the province. Similar studies could be performed at the regional level or considering other aspects of the building stock, such as retrofitting the building envelope or installing cooling systems. The case study demonstrates that the studied building stock can achieve the desired greenhouse gas reductions by shifting heating systems to electric alternatives, but at the cost of a significant increase to the peak electricity load. If not addressed properly by the local electricity distributor, the marginal electricity use for such scenarios could result in additional short-term electricity purchases from neighbouring sources with much higher electricity emission rates.

7 Discussion

Building stock energy models are an essential tool required for accurate estimation of energy efficiency measure evaluation, peak load assessments, comparative assessments of technology upgrades, and many other possible applications. The proposed model is designed in a way that produces a set of dwellings representative of the studied building stock. As shown in the case study section, the probability distributions used to generate the building parameters can be modified to compare between different configurations of building stocks, including the evaluation of a subset of buildings by region, by building type, by heating technology or other characteristic. The flexibility of the proposed stock model allows for many possible applications, for example:

- Evaluating the impact of a provincial energy efficiency upgrade measure targeting the replacement of double-glazed windows with triple-glazed windows, which could be accomplished by comparing the base case building stock with a new set of modeled homes with different window probability distributions.
- Determining the impact of a new community of 2000 homes on the electrical grid, which could be accomplished by modeling 2000 homes with representative

probability distributions and/or parameters fixed to correspond to the design of the new homes (i.e. wall thermal resistance values, etc.).

- Evaluating different occupancy patterns on the peak electrical demand, by modeling a new set of homes with specific occupancy patterns and comparing to the existing building stock.
- Etc.

The case study presented in the paper illustrates how a provincial greenhouse gas reduction target can potentially be achieved, but at the cost of a significant increase in peak electricity load. The proposed building stock energy model allows for an hourly comparison of electricity usage of the building stock, which allows for additional time-dependent analysis of greenhouse gas emissions and evaluation of the best measures based on marginal emission rates. Future studies can leverage the flexibility of the proposed model to evaluate a wide variety of building stock configurations.

8 Conclusion

The bottom-up white-box building stock energy model developed by the authors represents the single-family dwelling market for the province of Québec, Canada. The characterization process applied to the provincial stock data ensures that each region reflects the real distribution of systems and building characteristics according to the best available information. By implementing region-specific probability distributions, the stock model can then be applied to different areas of the province, such as the city of Montreal, rural areas, or to the entire province. The model is described using the stock energy model labelling system proposed by Langevin et al. (2020):

Country: Canada (province of Québec)
Model name: QSFBSEM
Model use: A static bottom-up white-box stock model for comparative assessment of residential stock energy use. Energy can be categorized by end-

use and by source, and peak electricity demand can be assessed. Appropriate for technological evaluation and greenhouse gas emissions studies.
Model classification quadrant: Q4 (physics simulation)
Additional details: N/A.

The QSFBSEM improves upon previous iterations of bottom-up white-box building stock energy models by accurately representing the sub-regional stock characteristics and by producing combinations of building parameters representative of the building stock. The compromise made in the development of the model is a high requirement in the number of buildings modeled to represent the stock, which demands a significant time investment to complete a simulation. The user can opt to model fewer houses however at the cost of some accuracy, as demonstrated in Appendix 4. Smaller portions of the stock, such as a community in one of the regions of the building stock, do not have this issue and can be modeled relatively rapidly with the proposed model. As a result, comparative studies for rural community building technologies can become an interesting application for the QSFBSEM. The capability to introduce technological changes to the building stock and evaluate them addresses one of the issues raised by Booth et al. (2012), described as the *flexibility* of building stock models. One of the key features of white-box models is the capability to evaluate technological changes immediately with simulation, rather than waiting for data to develop a statistics-based model. For the proposed model, the probability distributions can be altered to perform a direct comparison between different configurations of building stock characteristics, or a new technology can be introduced to a portion of the stock.

The total stock energy consumption of the model is within 1.6% of the provincial single-family dwelling stock data for 2017. As a point of comparison, the CHREM model (Swan et al. 2012) predicts total provincial energy consumption within 15% of the reference data. Electricity consumption is over-predicted by the CHREM model by 33%, compared to 7% with the proposed model. However, it is important to note that CHREM is a national model that covers the entire country of Canada and uses a different validation approach than the QSFBSEM, and therefore some differences are to be expected. The proposed model and methodology nevertheless represents a significant improvement over other existing models for the studied building stock.

References

Booth, A.T., R. Choudhary, and D.J. Spiegelhalter. 2012. “Handling Uncertainty in Housing Stock Models.” *Building and Environment* 48 (February). Pergamon: 35–47. doi:10.1016/J.BUILDENV.2011.08.016.

Government of Canada. 2021. “Engineering Climate Datasets.” *Engineering Climate Services*. https://climate.weather.gc.ca/prods_servs/engineering_e.html.

Government of Québec. 2020. *2030 Plan for a Green Economy*. Québec, Canada.

Hong, Tianzhen, Yixing Chen, Sang Hoon Lee, and Mary Ann Piette. 2016. “CityBES: A Web-Based Platform to Support City-Scale Building Energy Efficiency.” doi:10.1145/12345.67890.

Hydro-Québec. 2017. *Sustainable Development Plan 2020-2024.*

Hydro-Québec. 2021. “Rate DT – Dual Energy | Hydro-Québec.” https://www.hydroquebec.com/residential/customer-space/rates/rate-dt.html.

Johnson, R K. 2013. *Measured Performance of a Low Temperature Air Source Heat Pump Consortium for Advanced Residential Buildings*. Oak Ridge, TN, USA.

Kasten, Fritz, and Gerhard Czeplak. 1980. “Solar and Terrestrial Radiation Dependent on the Amount and Type of Cloud.” *Solar Energy* 24 (2). Pergamon: 177–189. doi:10.1016/0038-092X(80)90391-6.

Kavgic, M., A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic. 2010. “A Review of Bottom-up Building Stock Models for Energy Consumption in the Residential Sector.” *Building and Environment* 45 (7): 1683–1697. doi:10.1016/j.buildenv.2010.01.021.

Klein, S.A., W.A. Beckman, J.W. Mitchell, J.A. Duffie, N.A. Duffie, T.L. Freeman, J.C. Mitchell, et al. 2017. "TRNSYS 18: A Transient System Simulation Program, Solar Energy Laboratory, University of Wisconsin, Madison, USA." http://sel.me.wisc.edu/trnsys.

Langevin, J., J. L. Reyna, S. Ebrahimigharehbaghi, N. Sandberg, P. Fennell, C. Nägeli, J. Laverge, et al. 2020. "Developing a Common Approach for Classifying Building Stock Energy Models." *Renewable and Sustainable Energy Reviews* 133 (November). Pergamon: 110276. doi:10.1016/J.RSER.2020.110276.

Mathworks Inc. 2018. "Matlab R2018b."

McKenna, Eoghan, and Murray Thomson. 2016. "High-Resolution Stochastic Integrated Thermal–Electrical Domestic Demand Model." *Applied Energy* 165: 445–461.

Nägeli, Claudio, Clara Camarasa, Martin Jakob, Giacomo Catenazzi, and York Ostermeyer. 2018. "Synthetic Building Stocks as a Way to Assess the Energy Demand and Greenhouse Gas Emissions of National Building Stocks." *Energy and Buildings* 173 (August). Elsevier: 443–460. doi:10.1016/J.ENBUILD.2018.05.055.

National Renewable Energy Laboratory. 2021. "National Baseline Data Viewer - ResStock." https://resstock.nrel.gov/dataviewer/efs_v2_base.

Neale, Adam, Michaël Kummert, and Michel Bernier. 2020. "Development of a Stochastic Virtual Smart Meter Data Set for a Residential Building Stock – Methodology and Sample Data." *Journal of Building Performance Simulation* 13 (5): 583–605. doi:10.1080/19401493.2020.1800096.

New, Joshua Ryan, Mark B. Adams, Piljae Im, Hsiuhan Lexie Yang, Joshua C. Hambrick, William E. Copeland, Lilian B. Bruce, and James A. Ingraham. 2018. “Automatic Building Energy Model Creation (AutoBEM) for Urban-Scale Energy Modeling and Assessment of Value Propositions for Electric Utilities.” In *International Conference on Energy Engineering and Smart Grids*, 5. Oak Ridge, TN, USA: USDOE Office of Energy Efficiency and Renewable Energy, USDOE Office of Electricity.

NRCan (Natural Resources Canada). 2014. “Canada’s Secondary Energy Use by Sector, End Use and Subsector.” *Energy Use Data Handbook Tables*.

NRCan (Natural Resources Canada). 2015. “2015 Survey of Household Energy Use Data Tables by Census Metropolitan Area.” *National Energy Use Database*. https://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/menus/sheu-cma/2015/tables.cfm.

NRCan (Natural Resources Canada). 2017. “National Energy Use Database.” *Natural Resources Canada*.

NRCan (Natural Resources Canada). 2018. “Energuide Housing Database.”

NYSERDA. 2013. *Multifamily Performance Program Gas Furnace Electricity Usage Background*.

Oak Ridge National Laboratories. 2021. “Virtual EPB.” https://evenstar.ornl.gov/autobem/virtual_epb/.

Régie de l’énergie. 2017. “Relevés Des Livraisons d’énergie En Vertu de l’entente Globale Cadre Pour La Période Du 1er Janvier Au 31 Décembre 2017 - Version

Amendée .” *Audiences et Décisions, D-2016-143*. http://www.regie-energie.qc.ca/audiences/Suivis/Suivi_HQD_D-2016-143.html.

Reinhart, Christoph F., and Carlos Cerezo Davila. 2016. “Urban Building Energy Modeling - A Review of a Nascent Field.” *Building and Environment* 97 (February). Elsevier Ltd: 196–202. doi:10.1016/j.buildenv.2015.12.001.

Reinhart, Christoph F, Timur Dogan, Alstan Jakubiec, Tarek Rakha, and Andrew Sang. 2013. “UMI - an Urban Simulation Environment for Building Energy Use, Daylighting and Walkability.” In *13th Conference of International Building Performance Association*, 476–483. Chambéry, France.

Sherman, M.H., and D.T. Grimsrud. 1980. “Infiltration-Pressurization Correlation: Simplied Physical Modeling.” *ASHRAE Transactions* 86 (2): 778.

Sokol, Julia, Carlos Cerezo Davila, Christoph F. Reinhart, C. Cerezo, and Christoph F. Reinhart. 2016. “Validation of a Bayesian-Based Method for Defining Residential Archetypes in Urban Building Energy Models.” *Energy and Buildings* 134. Elsevier B.V.: 11–24.

Statistics Canada. 2016. “2016 Census.” *Statistics Canada*. http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm.

Statistics Canada. 2019. “Annual Demographic Estimates: Subprovincial Areas.” https://www150.statcan.gc.ca/n1/pub/91-214-x/2020001/section01-eng.htm.

Swan, Lukas G., and V. Ismet Ugursal. 2009. “Modeling of End-Use Energy Consumption in the Residential Sector: A Review of Modeling Techniques.” *Renewable and Sustainable Energy Reviews*. doi:10.1016/j.rser.2008.09.033.

Swan, Lukas G., V. Ismet Ugursal, and Ian Beausoleil-Morrison. 2012. “Hybrid Residential End-Use Energy and Greenhouse Gas Emissions Model - Development and Verification for Canada.” *Journal of Building Performance Simulation* 6 (1): 1–23. doi:10.1080/19401493.2011.594906.

Taniguchi-Matsuoka, Ayako, Yoshiyuki Shimoda, Minami Sugiyama, Yusuke Kurokawa, Haruka Matoba, Tomoya Yamasaki, Taro Morikuni, and Yohei Yamaguchi. 2020. “Evaluating Japan’s National Greenhouse Gas Reduction Policy Using a Bottom-up Residential End-Use Energy Simulation Model.” *Applied Energy* 279 (December). Elsevier: 115792. doi:10.1016/J.APENERGY.2020.115792.

The Atmospheric Fund. 2019. *A Clearer View on Ontario’s Emissions Electricity Emissions Factors and Guidelines*.

Transition Énergétique Québec. 2019. “Emission Factors and Conversion.” https://transitionenergetique.gouv.qc.ca/fileadmin/medias/pdf/FacteursEmission.pdf.

Ugursal, V I. 2017. *Canadian Building Stock Data: Notes from the Stock Modeling Workshop*. Ottawa, Canada.

US DOE. 2013. *EnergyPlus Engineering Reference - The Reference to EnergyPlus Calculations (v 8.1)*. Washington, DC, USA: US Department of Energy.

Wilson, Eric, Craig Christensen, Scott Horowitz, Joseph Robertson, and Jeff Maguire. 2017. *Energy Efficiency Potential in the U.S. Single-Family Housing Stock*. Denver, CO, USA.

Appendix 1: Weather stations for each region of the studied building stock

Table A-1. Weather stations for each region of the studied building stock. CMA: census metropolitan area, CA: census agglomeration.

Number	Region type	Fraction	Location name	Nearest weather station
R1-01	CA	0.0156	Alma	CAN_QC_JONQUIERE_7063370
R1-02	CA	0.0144	Baie-Comeau	CAN_QC_BAIE-COMEAU_704S001
R1-03	CA	0.0019	Campbellton	CAN_NB_CHARLO-AUTO_8100885
R1-04	CA	0.0053	Cowansville	CAN_QC_FRELIGHSBURG_7022579
R1-05	CA	0.0082	Dolbeau-Mistassini	CAN_QC_NORMANDIN_7065639
R1-06	CA	0.0440	Drummondville	CAN_QC_NICOLET_7025442
R1-07	CA	0.0362	Granby	CAN_QC_FRELIGHSBURG_7022579
R1-08	CA	0.0057	Hawkesbury	CAN_QC_MONTREAL-MIRABEL-INTL-A_7034900
R1-09	CA	0.0207	Joliette	CAN_QC_L'ASSOMPTION_7014160
R1-10	CA	0.0058	Lachute	CAN_QC_MONTREAL-MIRABEL-INTL-A_7034900
R1-11	CA	0.0104	Matane	CAN_QC_AMQUI_7050145
R1-12	CA	0.0288	Rimouski	CAN_QC_POINTE-AUX-PERE-(INRS)_7056068
R1-13	CA	0.0142	Rivière-du-Loup	CAN_QC_RIVIERE-DU-LOUP_7056616
R1-14	CA	0.0199	Rouyn-Noranda	CAN_QC_ROUYN-NORANDA-A_7086719
R1-15	CA	0.0175	Saint-Georges	CAN_QC_BEAUCEVILLE_7028754
R1-16	CA	0.0217	Saint-Hyacinthe	CAN_QC_MONTREAL-ST-HUBERT_7027329
R1-17	CA	0.0073	Sainte-Marie	CAN_QC_BEAUCEVILLE_7028754
R1-18	CA	0.0181	Salaberry-de-Valleyfield	CAN_QC_ST-ANICET-1_702FQLF
R1-19	CA	0.0125	Sept-Îles	CAN_QC_SEPT-ILES-A_7047911
R1-20	CA	0.0261	Shawinigan	CAN_QC_SHAWINIGAN_7018001
R1-21	CA	0.0220	Sorel-Tracy	CAN_QC_LAC-SAINT-PIERRE_701LP0N
R1-22	CA	0.0160	Thetford Mines	CAN_QC_BEAUCEVILLE_7028754
R1-23	CA	0.0151	Val-d'Or	CAN_QC_VAL-D'OR_7098603
R1-24	CA	0.0235	Victoriaville	CAN_QC_LEMIEUX_701Q009
R1	Québec non-CMA	0.4108		
R2	CMA	0.0233	Saguenay	CAN_QC_JONQUIERE_7063370
R3	CMA	0.0949	Québec	CAN_QC_QUEBEC-INTL-A_7016293
R4	CMA	0.0262	Sherbrooke	CAN_QC_LENNOXVILLE_7024280
R5	CMA	0.0214	Trois-Rivières	CAN_QC_NICOLET_7025442
R6	CMA	0.3756	Montréal	CAN_QC_MONTREAL-INTL-A_7025251
R7	CMA	0.0478	Ottawa - Gatineau	CAN_ON_OTTAWA-INTL-A_6106001

Appendix 2: Total and opaque cloud cover data correction

The cloud fraction can be calculated with Equation (A-1) using the global and diffuse horizontal radiation values that are available in the CWEEDs weather data (Kasten and Czeplak 1980).

$$f_{cloud} = 10 * \left(1.4286 \frac{E_{dif}}{E_{glob,h}} - 0.3\right)^{0.5} \qquad \text{(A-1)}$$

where E_{dif} and $E_{glob,h}$ are the diffuse and global horizontal radiation (Wh/m^2), respectively, and f_{cloud} is the cloud fraction in tenths. The Total and Opaque Sky Cover values in the CWEEDs files are assumed equal to the cloud fraction and are rounded to the neared whole integer, which is common practice (Government of Canada 2021). Since the global horizontal radiation in CWEEDs data is zero at night, the cloud cover data during night time is linearly interpolated between the last available value for cloud cover and the first available data point the next morning, as illustrated in Table A-2.

Table A-2. Example cloud cover data generated and filled for a 24 hour period. Night time values highlighted in grey are filled using linear interpolation between the two data points in bold text.

Hour of the day	Global horizontal radiation (Wh/m^2)	Diffuse horizontal radiation (Wh/m^2)	Cloud cover without interpolation (tenths)	Cloud cover with interpolation (tenths)
13	827	276	4	4
14	889	254	3	3
15	249	220	10	10
16	654	189	3	3
17	468	93	0	0
18	305	59	0	0
19	135	43	4	4
20	21	14	**8**	**8**
21	0	0	0	8
22	0	0	0	7
23	0	0	0	7
24	0	0	0	7
1	0	0	0	6
2	0	0	0	6
3	0	0	0	6
4	0	0	0	5
5	0	0	0	5
6	142	51	**5**	**5**
7	310	67	1	1
8	487	83	0	0
9	647	90	0	0
10	780	212	3	3
11	225	216	10	10
12	229	227	10	10

In summary, the general approach for completing the cloud cover data is as follows:

1) Verify whether any monitored data for Total Sky Cover was available, usually at 3 hour intervals. If yes, set the values for Total Sky Cover for hours 2 and 3 equal to the first hour and repeat for the whole year.
2) If no measured data is available, use Equation (A-1) to determine the cloud fraction for the hours where the global and diffuse horizontal radiation is available, and use linear interpolation to complete the night time values.

In all cases, the Opaque Sky Cover is assumed equal to the Total Sky Cover.

Appendix 3: Bayes' Theorem

The following section is an excerpt adapted from Neale et al. (2020) that describes how Bayes' Theorem is applied to create conditional probability distributions for dependent building parameters. Interested readers can refer to the original publication for additional detail.

Bayes' theorem, described in Equation (1), is applied for each of the connections in the network in [Figure 3], which allows for the determination of the conditional probability of a parameter given prior evidence.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

where A and B are dependent parameters, $P(A|B)$ is the conditional probability of A given B occurring, $P(A)$ is the prior probability distribution of A, $P(B)$ is the prior probability distribution of B, and $P(B|A)$ is the prior conditional probability of B given A occurring. $P(B|A)$ is typically based on prior knowledge, i.e. based on data obtained in the literature. Prior probability distributions such as $P(A)$ can be established based on evidence found in the literature and expressed using Equation (2).

$$P(A) = \frac{n_{Ai}}{n} \tag{2}$$

where n_{Ai} is the number of cases for class i of parameter A and n is the number of samples.

Appendix 4: Building stock sample size

Much like previous works, this study takes the approach of selecting a smaller subset of homes and scaling up the energy consumption of that subset. In such cases, an important aspect of building stock energy modeling is determining the size of the sample set of homes that accurately represents the overall stock. The impact of the sample size is expressed in terms of the normalized root mean square deviation (NRMSD) with respect to the mean of a sample. More precisely, the following procedure is followed:

1) The stock sample size b is selected, e.g. 5000 buildings, which will represent the building stock of size s, i.e. 1.9 million buildings for single-family dwellings.
2) The stock fraction SF is determined, where $SF = b/s$.
3) A building stock sample of size b is generated with the proposed model and the total energy consumption is determined for a variety of end-use and energy source categories. This process is repeated n times.
4) The size of the building stock sample is increased and steps (1) to (3) are repeated.
5) The $NRMSD$ for each tested stock sample size is determined, as described in Equation (3).

$$NRMSD = \frac{\sqrt{\frac{\sum_{i=1}^{n}(E_i - \bar{E})^2}{n}}}{\bar{E}} \tag{3}$$

where:

$NRMSD$ is the normalized root mean square deviation for a given sample size

E_i is the calculated stock energy for sample of buildings i (PJ)

$\bar{E}$ is the mean stock energy across all samples of size b (PJ)

n is the number of total sets of stock samples

The NRMSD values are illustrated in Figure A-1 as a function of total energy consumption and building type (attached, detached and single-family dwellings) for a variety of end-use and energy source categories. The NRMSD metric demonstrates the expected residual on the total energy consumption for a given stock fraction.

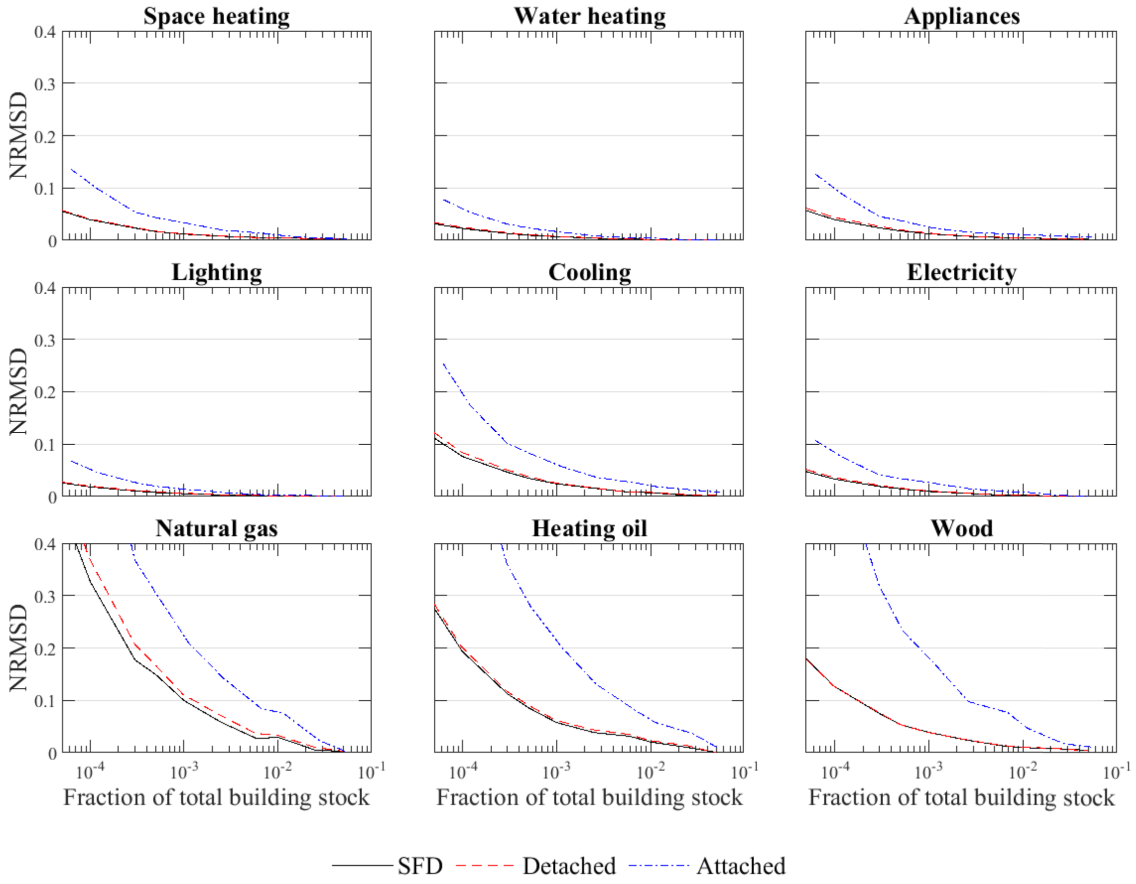


Figure A-1. NRMSD for energy consumption by end-use and energy source by fraction of the total building stock modeled.

The NRMSD of the energy consumption varies considerably depending on the stock fraction, building type and energy category. Categories strongly tied to electricity consumption have lower residuals than nonelectric categories, which is due to the prevalence of electricity in the studied building stock. Lesser-represented categories, such as natural gas, heating oil and wood, are more disposed to variations in the total energy use for smaller stock samples. Attached houses experience larger NRMSD values than detached houses for the same reason. As the fraction of the building stock approaches 10%, the NRMSD is well below 0.01 for all energy categories.

For a more specific comparison, the literature shows that stock model samples of 0.03% (Taniguchi-Matsuoka et al. 2020) and 0.6% (Nägeli et al. 2018) have been applied in previous works. For the building stock studied here, the NRMSD is illustrated for the same stock sample sizes used by Taniguchi-Matsuoka et al. and Nägeli et al. in Figure A-2. The same 10 energy use categories presently previously are are depicted.

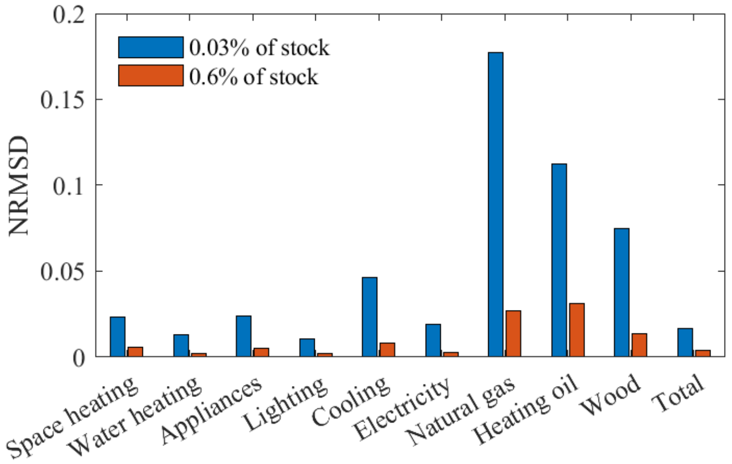


Figure A-2. NRMSD for SFD energy consumption categories for cases representing 0.03% and 0.6% of the modeled building stock.

The NRMSD is generally under 0.03 for the 0.6% stock sample, but can reach 0.18 for a sample of 0.03% of stock. In simpler terms, a deviation of 18% on the natural gas energy prediction is expected for a stock sample of 0.03% for the studied building stock. The *Total* energy use NRMSD is low in both cases, but if specific categories are required for a particular analysis then the resulting deviation can be quite large. For example, if the stock model is used to evaluate cooling energy savings due to a particular incentive measure, the NRMSD can be significantly improved with a large enough building stock sample. To apply this reasoning more generally, building stock energy modelers should be aware of the potential deviation due to stock size for lesser-represented portions of the building stock, particularly if they are relevant to the analysis they are pursuing.