Preprints.org

Article

The Urban Building Energy Retrofitting Tool: An Open-Source Framework to Help Foster Building Retrofitting Using a Life Cycle Costing Perspective. First Results for Montréal

Oriol Gavaldà Torrellas [*] , Pilar Monsalvete , Ursula Eicker

Article

The Urban Building Energy Retrofitting Tool: An Open-Source Framework to Help Foster Building Retrofitting Using a Life Cycle Costing Perspective. First Results for Montréal

Oriol Gavaldà Torrellas [1,†] **, Pilar Monsalvete** [2] **and Ursula Eicker** [1,*]

1   Concordia University
2   Universidad de Cádiz
*   Correspondence: oriol.gavalda@concordia.ca
†   Current address: 2155 Rue Guy, ER-1431.35

Abstract: Building decarbonization is a major challenge for cities. Deciding what, when, and how to retrofit a building is difficult, given the complex interaction between energy costs and investment requirements. Several tools have been developed in the last years to help public and private stakeholders with these decisions, but none cover aspects the authors think are fundamental. For this reason, an urban building retrofit tool was developed and is presented in this article. This new tool is based on a bottom-up approach, with all buildings simulated individually, considering aspects such as shading or adjacencies. As a second step, three scenarios with different levels of ambition have been implemented in the tool, and the energy demand and emissions resulting from these scenarios have been calculated. As a third step, the retrofitting scenarios' initial investment and operational costs have been implemented via a detailed Life Cycle Costing (LCC) approach. The authors describe the robust and scalable structure that has been developed and the application of this structure to calculate LCC of different retrofitting scenarios in Montréal.

Keywords: urban building energy models; retrofitting; life cycle costs

---

1. Introduction

One of the most demanding segments of urban-scale decarbonization is building retrofits. The building sector accounts for more than 35% of global energy consumption and almost 38% of CO_2 emissions if we consider embodied energy [1].

The decision of what we do to existing buildings can significantly impact future CO_2 emissions. In the scientific community, there are significant doubts that demolishing and rebuilding net zero is a better option than retrofitting because of the reduced CO_2 embodied emissions of the latter option [2], [3], [4], but a strong effort must be made to ensure that it can be attractive to all stakeholders.

According to [5], even after 90 years of lifetime, retrofitting seems a better option than demolishing. In most cases, retrofitting seems the best option from a multi-factor social, environmental, and economic aspect. According to recent references, the stock of existing buildings will account for 75% of Green-House Gases (GHG) emissions, whilst new buildings will account for 25% (2021 World Green Building Council. WorldGBC Net Zero Carbon Buildings Commitment). According to recent references, the stock of existing buildings will account for 75% of GHG emissions whilst new buildings will account for 25% [6].

In Canada, as an example, the replacement rate of old buildings by new buildings is only 1-3% per year [7]. However, although the goal seems clear for most policymakers in different countries, massive and large-scale retrofitting is not happening. Very interesting examples of successful case studies have been happening worldwide if life cycle costs are taken into account[8], but large-scale implementation is not successful for several reasons. The most significant reasons [9] are the knowledge gap, the expertise gap, the lack of political commitment, inadequate policies, the speed to scale up, and the long return on investment. Some of those, especially the ones related to policymaking and the knowledge gap,

linked to the lack of detailed life cycle information on the impact of these measures [10], could be overcome using tools such as Urban Building Energy Modelling (UBEM).

Therefore, an Urban Building Energy Retrofit Tool based on UBEMs could be a step forward thanks to its scientific, modular, and scalable approach and ease of application to any city, given the reduced datasets necessary to implement it in a first rough approach.

2. Literature Review on UBEM and Retrofit Tools

2.1. Existing Urban Modelling Tools

Many tools have been developed in the last years to model energy flows in cities, trying to simplify the use and access to urban energy models but covering only partially the complexity of the energy system [11]. The so-called UBEM aims to provide understanding in all aspects of a city's energy flows, from building energy "consumption" to on-site renewable "production", including traffic or waste management. Within this broad field, many have focused on the building stock, which is the aspect the authors will focus on in this article. UBEM analyses the building demands in their urban context. [12] concluded in their review on urban models and tools that the building-to-building thermal and solar interactions, Urban Heat Island (UHI) or vegetation might actively modify the results of urban energy system simulations.

This building-urban interaction was considered for the first time in 2005 in the Sustainable Urban Neighbourhood modeling tool, `SUNtool` [13]. A thermal model of the building, coupled with Heating, Ventilation, and Air Conditioning (HVAC) and Environmental Control System (ECS) models, was integrated into the tool, also including an urban-sensitive radiation model, the Simplified Radiosity Algorithm (SRA) [14].

Heavily inspired by `SUNtool`, but further refined with more rigorous thermal and behavioral models was the comprehensive software `CitySim`. Citysim was a step forward with respect to previous UBEM tools, and it used a Reduced Order Models (ROM) developed at EPFL in Switzerland. The tool also allowed for the calculations with a simplified Dynamic Simulation Tool and integrations with other tools, like the transportation tool MATSIM.

One of the followers of Citysim is Simstadt. SimStadt is an urban workflow engine written in Java and developed by the University of Applied Sciences Stuttgart [15] in 2013. SimStadt can use both simplified geometry (extrusion of footprint) and, when available, it can use City Geography Markup Language (CityGML) format for detailed geometry (Level of Detail (LoD) 2) while considering the roof shapes. A study done by [16] modeled about 1000 buildings in Rotterdam, Netherlands, using the geometry format of CityGML. SimStadt benefits from the direct integration of , which contains thermophysical building attributes in a machine-readable format as an extension of CityGML [17]. Like other UBEM tools, SimStadt requires several preprocessing steps of input data to be suitable for energy analysis and has monthly output resolution. Finally, it does not provide transportation and mobility calculations and their connection with the UBEM. However, it provides life-cycle analysis using embodied carbon for building construction in a seamless tool [18].

At the same time, the Urban Modeling Interface (UMI) was developed at MIT in 2013, using EnergyPlus as the energy simulation engine. The input geometry used in the UMI is provided by Rhinoceros software.

CityBES was also developed running EnergyPlus as the simulation engine, but in this case, using the City Geography Markup Language (CityGML) format for 3D urban building modeling in LoD1 and LoD2 (considering the roof shapes). It can also be linked to EnergyADE for using the thermophysical building attributes.

Two other UBEM tools that use Resistance Capacitance (RC) models as a base engine are OpenIDEAS (an open-source platform) and CEA. The former uses an interesting approach to structure the entering data sources and the workflows to allow good input preprocessing, output postprocessing, and workflows to happen outside the main structure of the tool. CEA is one tool that integrates more aspects,

including several input sources, incorporating cost-benefit analysis, transportation use cases, and greenhouse gas emissions. Shadow analysis is also incorporated. Urban Renewable Building and Neighborhood Optimization (URBANOPT) is one of the latest tools developed by the National Renewable Energy Laboratory (NREL). This tool uses EnergyPlus and is open-source [19], [20]. It also contains lighting simulation but does not include transportation/mobility, life cycle, or cost-benefit analysis.

The Tool for Energy Analysis and Simulation for Efficient Retrofit (TEASER), developed at Aachen University (RWTH), is based on RC models, and it looks at several carbon emission accounting strategies[19] .

A summary of the different UBEM tools and their main aspects has been elaborated and is shown in tables 1, 2, 3, and 4.

Table 1. UBEM tools and fields of work

Tools	Buildings operational	Buildings Life Cycle	Transport	Microclimate	Waste	Other aspects
SUNTOOL	Yes	No	No	Yes	Yes	No
CITYSIM	Yes	No	Yes	No	No	No
SIMStadt	Yes	Yes	No	No	No	No
UMI	Yes	Yes	Yes	Yes	Yes	Food production, daylighting
CityBES	Yes	No	Yes	No	No	No
OpenIDEAS	Yes	No	No	No	No	No
CEA	Yes	Yes	Yes	Yes	Yes	No
UrbanOPT	Yes	No	No	No	No	Daylighting
TEASER	Yes	No	No	No	No	No

Table 2. UBEM tools and input formats

Tool	Geospatial Input formats accepted	Constructive inputs	Usage inputs defined	Shading analysis	Microclimate possibilities	Other aspects
SUNTOOL	Bespoke interface	To be defined by users	Static, deterministic and probabilistic	Yes, SRA	Yes	No
CITYSIM	CityGML, EnergyADE	To be defined by user, based on archetypes	Static, deterministic	Yes, SRA	No	No
SIMStadt	CityGML, EnergyADE	To be defined by user, based on archetypes	Static, deterministic	Yes, SRA	No	No
UMI	Rhino3D	To be defined by user, based on archetypes	Static, deterministic	Yes, based on Eplus	Yes (UWG)	Food production, daylighting
CityBES	CityGML, EnergyADE	To be defined by user, based on individual buildings	Static, deterministic	Yes, based on Eplus	No	No
OpenIDEAS	CityGML, EnergyADE	To be defined by user, based on archetypes	Static, deterministic and probabilistic	Not well explained and detailed (missing?)	No	No
CEA	CityGML, osm	To be defined by user, based on archetypes	Static, deterministic and probabilistic	Yes	Yes, EnviMET can be coupled to the models	No
UrbanOPT	CityGML, GeoJSON, osm, IDF	To be defined by user, based on archetypes	Static, deterministic	Yes, based on Eplus	No	Daylighting
TEASER (part of IDEAS)	CityGML, EnergyADE	To be defined by user, based on archetypes	Static, deterministic and probabilistic	Not well explained and detailed (missing?)	No	No

Table 3. USEM tools and software engine characteristics

Tool	Engine for building simulation	Engine for energy systems calculations	Engine for mobility calculations	Programming language	Output processing	Open-source architecture	Frequency data
SUNTOOL	R.O.M (grey-box model)	Yes	No	Java	No	No	Hourly
CITYSIM	R.O.M (RC model)	Own engine	MATSIM-c	C++ and java	Yes	Yes	Hourly
SIMStadt	Simplified energy balance methodology	Own engine	No	Java	Yes	Yes	Monthly
UMI	Energyplus	Energyplus	Own engine (python), and walkability based	Rhino Grasshopper (python)	Yes	Partly dependent on Rhino.	Hourly and sub-hourly
CityBES	Energyplus	Energyplus	No		Yes	Yes	Hourly and sub-hourly
OpenIDEAS	R.O.M (RC model)	Modelica	No	Python	Yes	Yes	Hourly
CEA	R.O.M (RC model)		No	Python	Yes	Yes	Hourly
UrbanOPT	Energyplus	Energyplus	No	Ruby, C++, Python	Yes	Yes	Hourly and sub-hourly
TEASER	R.O.M (RC model)	Modelica	No	Python	Yes	Yes	Hourly

Table 4. USEM tools and scenario creation

Tools	Incorporation of cost/benefit analysis	Incorporation of scenarios	UI tool	Level of scaling up of the UBEM tool
SUNTOOL	No	No	Yes	Low scaling-up
CITYSIM	No	No	Yes	Low scaling-up
SIMStadt	No	Yes	Yes	High scaling-up
UMI	No	Yes	Yes	High scaling-up
CityBES	Yes		Yes	Low scaling-up
OpenIDEAS	No		No	Low scaling-up
CEA	Yes	Yes	Yes	High scaling-up
UrbanOPT	No	Yes	Yes	High scaling-up
TEASER	No		Yes	Low scaling-up

2.2. Gaps to Be Overcome

The different UBEMs described in the previous chapters have established the seed for deploying relevant future urban system energy tools.

However, several research questions arise from the literature review and the tables mentioned. The capacity of these tools to answer real cost-benefit questions from the building decarbonization sector is often under doubt for several reasons. When dealing with such complex problems, the usual question is whether the UBEMs are designed with a goal in mind and whether they will be useful for more than this goal.

The capacity, therefore, for the aforementioned UBEMs for adaptability and modularity is small. Although they try to be multi-faceted tools, the design of a tool, if not considered modular from the beginning, tends to be adapted to the use case for which it was first developed. Each UBEM was developed with a back-end calculation methodology in mind, which is applicable in the field for which this methodology is useful.

Another important aspect is **the lack of scalability to a large number of buildings and different locations**. For example, most of the aforementioned tools have been applied to one or a few (under 5) use cases and, in most cases, linked to a small set of buildings. The use, in some cases, of simplified Reduced Order Models (Suntool, CitySIM, OpenIDEAS, CEA, TEASER), the quality of the results for the use of very simplified models makes it difficult to extend to . Conversely, using detailed dynamic simulation tools (UMI, CityBES, UrbanOPT) makes it hard to develop large-scale scenarios.

These tools **lack of capacity to deal with several types of input datasets** (as seen in table 2) make it necessary to clean and adapt the input data sources every time. This is a highly cumbersome process that sometimes takes months and limits the scaling up of the methodology. The typically incoherent input datasets from cities also make it necessary to compare several sources of information and choose the cleanest, which, in the cases of most current UBEMs, is a process that has to happen before using the tools.

The fourth is the **lack of data input automation**. Many authors [12],[21],[11],[22] [23] have delved deeply into the relevance of data input automation, which reduces the chance of input errors and also human costs as the data processing is, in most cases, one of the highest time consumers [24]. **Lack of open source developments** is also an issue, at least in some aspects of the different UBEM. UBEMs deal with very different sectors of GHG emitters (buildings, transportation, water), and the will to break the silos and integrate attractive models from academia in all those fields should be the goal. The development of UBEMs using proprietary software can partially block researchers from accessing these tools.

As a last point, **clear communication of results has long been an issue** in using urban simulation tools [12]. In the urban simulation tools field, GIS thematic mapping becomes a very intuitive tool for final users to access and understand results [25],[26],[27].

2.3. Reasons for the Development of a New Model

The research question that arises from the literature review and the gap detection is whether the current UBEM tools are designed to address the challenges required by different use cases. Is the structure behind the development of each UBEM conditioned by the use case for which each UBEMs was developed? After the review of the previous chapter, the authors think it is.

According to the authors, the answer to all those gaps is creating an open-source, flexible, and structured framework that allows multiple input files of city characteristics and tools to be used and integrated.

This framework should allow for the development of different use cases, all related to improving the sustainability of cities. The non-adherence to any concrete software, but the decision to create a framework to integrate them all is essential for this purpose. Some tools will be preferable to others when analyzing several use cases, and the platform will deal with them.

The authors will describe in detail a framework, the TOOLS4Cities hub that has been developed in the last three years and the application of this framework to the concrete case of an Urban Scale Retrofit Tool.

3. Materials and Methods

3.1. General Development of a UBEM Framework to Reach Retrofit

The basis for solidly developing the retrofitting use case depends on developing a scalable, modular, non-proprietary, and usable tool. Details of the UBEM developed by the authors of this paper can be found in [28]. The current paper will explain some aspects of the software architecture. The goal is to show how such a complex use case as a building retrofit using different input data sources and processing tools is represented using the UBEM framework developed at Concordia University's Canada Excellence Research Chair (from the CCERC), named the Tools4Cities HUB.

3.2. Structure of UBEM

3.2.1. UBEM Software Perspective

For good scalability, a comprehensive software structure has to be developed to allow for the evolution of the datasets and the adaptation to different tools while maintaining a robust ontology. The UBEM structure from the CERC uses five fundamental components: data models, models, factories, workflows, and data. Models are the representation of the properties and behaviors of a specific system. Workflows run all the models that are needed for a particular simulation. Data refers to the input data sources that enrich the UBEM and the results from the workflows developed in the UBEM.
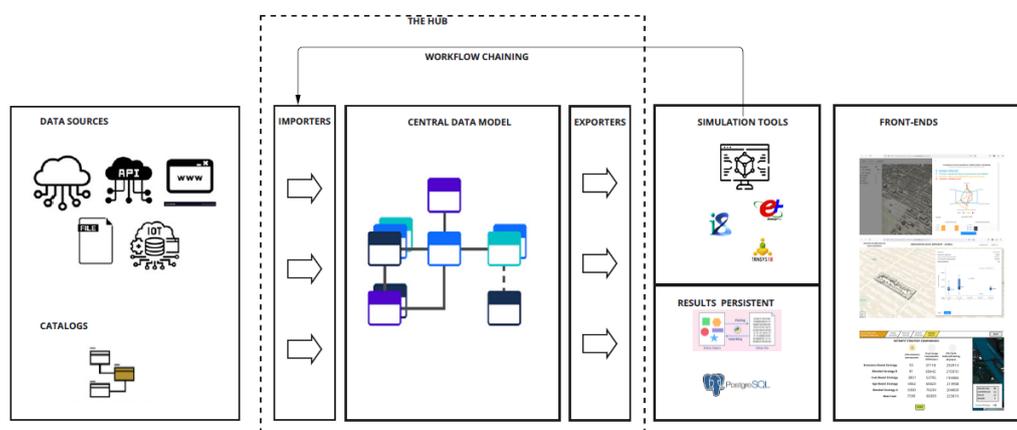

Figure 1. General view of the platform

The Appendix A develops, in detail, the software structure that we have developed to implement the new UBEM.

3.3. Retrofit Workflow

The retrofit workflow is a clear application of all the previously defined principles, and all the necessary steps towards the obtention of the results are defined. The retrofit workflow's essential objective is to analyze the energetic, environmental, and economic impact of retrofitting any building in a city. The idea would, therefore, be to pick any building in a city and get the investment costs, operational costs, and environmental emissions of leaving the building as-is or implementing some pre-defined retrofitting strategies. In the first approximation (or Minimum Value Product (MVP)) to the retrofit workflow, the retrofitting strategies were already predefined and homogeneous for all the buildings in the city. However, the tool's structure would allow for the implementation of more

detailed large-scale retrofitting strategies, and using heuristics to apply different strategies depending on different typologies of buildings.

3.3.1. Pipeline of Work

The main decisions regarding the pipeline of work were based on energy calculations, retrofitting scenario decisions, and economic analyses. As for energy simulation, following the structure defined in appendix A, once a city dataset is available, a **geometry factory** is called, capturing all the different shapes of a city's buildings, including the adjacencies between the buildings. The inclusion of several typologies of handlers, that can integrate and homogeneise different typologies of geospatial data, such as CityGML, geojson, 3ds and obj, facilitates the adaptability of the tool to the datasets available for every city. The archetypes of each of the different functions and vintages of the buildings are assigned in this first phase. Archetypes have been used for a long time in building simulation tools and are normally used to calculate the specific energy consumption of each building in the city, scaled to its surface [29]. However, in our case, using real geometries is prioritized as long as the quality of input geometrical datasets allows it. Nevertheless, in case some data is unknown, such as the Window to Wall Ratio, archetypes can fill the gaps. Once the geometry is captured, a **construction factory** is called to fill, with the aforementioned archetypes, the construction characteristics of each building. The construction factory uses, as well different handlers for different countries. At the moment, Us archetypes and Canadian archetypes have been developed. An endeavour to implement the archetypes defined in Tabula [23] is being implemented at the moment this article is being published. The **usage factory** comes afterward, filling the city instance with load profiles, occupancy gains, and setpoints for each building. Two usage factory handlers are available, the first based on the Canadian Energy Code [30], and the second on the Comnet profiles used by the DoE of the United States[31]. The **energy systems factory** is the last step of the workflow and allows for the incorporation of concrete energy systems for each of the city's buildings. Given the usual lack of building based information for energy systems, a method assigning randomly energy systems based on the building typology and the city-scale fuel consumption data has been developed. Once the whole city is created, deciding which simulation to use is crucial. Given the size of the endeavor (city-scale, with a first test made for Montréal), the two tools that could be used were a static demand tool (Monthly Energy Balance) and a dynamic simulation tool (Energyplus). For a first proof of concept, and since the final goal of the tool was understanding energy balances and long-term impacts instead of hourly calculations, the Monthly Energy Balance was chosen, coupled with a peak load calculation tool (based on the ASHRAE Heat Balance Method [32]). Once the structure for the MEB has been developed, the orchestration of the use of dynamic building simulation results will be developed.

3.3.2. Monthly Energy Balance Workflow

The MEB workflow involves several steps. The first one, the creation and first enrichment of the digital twin of the city, is performed using the four importers previously described: Geometry, Construction, Usage, and Systems. Once these steps are finished, the first simulation tool calculates the radiation on external walls. Simplified Radiosity Algorithm (SRA) [13] makes use of the geometry imported and the physical characteristics of the materials at the external layers of the walls, together with the weather data, to calculate the short wave radiation exchange between external walls and roofs. This calculation is performed in patches, as calculating together all the buildings belonging to the city in one run is computationally impossible. These patches overlap, allowing the selection of the surfaces with the most realistic outputs. Once this is done, all inputs required for the static calculation of the building demands are in place. So, the simulation tool (INSEL [33]) is called to perform the energy calculations. All buildings, one by one, are called, and their demands are calculated using the static model implemented in INSEL. In the third step, DHW and energy consumption are also calculated, considering the energy supply systems installed in each building. All results are saved and left available for use in the following steps.

3.3.3. Costs Workflow

In the case of the Life Cycle Costing (LCC) process, we have dealt with them in a separate Classes structure. The different nature of economic hypotheses, which are much more volatile than energy calculations, has pushed the team to develop a parallel structure coupled with the hub structure. This parallel structure will allow for further analysis, from the perspectives of economic scientists, like implementing dynamic cash flow modeling with sensitivity and risk analysis, a topic that has been detected as essential [34], or monetizing the externalities. The methodological approach for the economic analysis of the retrofit scenarios comprises the general LCC framework approach and the evaluation of externalities included in the general cost modeling methodology. LCC analysis, also known as total cost accounting, compares full costs during their useful life while defining alternative investment options in buildings or operating activities. For example, a baseline case can be compared with a potentially attractive one using the LCC of a retrofit scenario. LCC can showcase the different cost allocations of both scenarios, given that one of them already exists and does not imply extra investment costs. The new proposed scenario implies an extra investment but, on the other hand, minimizes yearly costs and, potentially, other externalities. LCC analysis design and practice is standardized in ISO 15686-5 [35], establishing the necessary criteria for analyzing Life Cycle Costing in the built environment. The global formulation is shown in equation (1).

$$
\begin{aligned}
LCC = CAPEX &- Subsidies + \\
\sum_{i=nyears} & \frac{\sum_{k=chaptinvest} CAPEX_{repositionk}(1+cpi)^i}{(1+d)^i} + \\
\sum_{i=nyears} & \frac{\sum_{j=nfuels} Operationalcostsj(1+epi)^i}{(1+d)^i} + \\
\sum_{i=nyears} & \frac{\sum_{l=nconcepts} Maintenancecostsl(1+cpi)^i}{(1+d)^i} + \\
\sum_{i=nyears} & \frac{\sum_{m=nincomes} Operationalincomesm(1+epi)^i}{(1+d)^i} + \\
\sum_{i=nyears} & \frac{\sum_{n=nextern} Externalitiescostsn(1+extpi)^i}{(1+d)^i} + \\
& \frac{Endoflifecosts(1+cpi)^i}{(1+d)^i}
\end{aligned}
$$

$$(1)$$

, where

$CAPEX$ = initial investment (product acquisition + installation and commissioning);

$CAPEX_{reposition}$ = investment cost for the reposition of the equipment after the end of life of the equipment;

$Operationalcosts$ = refers to the variable costs of operation;

$Maintenancecosts$ = refers to the variable costs of maintenance;

$Operationalincomes$ = refers to the variable incomes of operation, i.e. solar PV feed-in tariff;

$Externalitiescosts$ = refers to the variable costs of externalities, like CO2 or other emissions;

cpi = consumer price index increase. series of yearly values

epi = energy price index increase, which depends on the fuel type. series of yearly values

$extpi$ = externalities price index increase, which depends on the future externalities market, such as CO2 markets

d = discount rate for the retrofitting

CAPEX Calculation In our tool, we are using a data catalog structure to calculate the CAPEX costs. In the case of CAPEX, UNIFORMATII has been chosen as the format to classify building elements and

related site work [36]. The level of detail used in the first validation of the tool is significantly low, as the solutions are looked at from a global perspective. Within this LoD1, and for the sake of the first tests, only additional costs of the retrofitting have been considered, with average values per surface of the different elements, not distinguishing different solutions. The first structure of the data catalog for CAPEX costs is shown in table 5.

Table 5. Breakdown of UNIFORMATII costs used in the platform

Id	Concept
B	Shell
B10	Superstructure. Ground refurbishment
B20	Envelope
B2010	Opaque walls
B2020	Transparent walls
B30	Roofing
B3010	Roofing opaque
B3020	Roofing transparent
D	Services
D30	HVAC
D3010	Energy supply (PV systems)
D3020	Heat generating systems
D3030	Cooling generating systems
D3040	Distribution systems
D3060	Control and instrumentation
D3080	Other HVAC systems. AHU
D50	Electrical
D5020	Lighting and branch wiring
Z	Allowances
Z10	Design allowance
Z20	Overhead profit

As explained, the current structure of the central catalog model allows the connection of different datasets. In the first MVP, we connected this central catalog model to our first XML database, a simplified version of costs based on archetypes, localization of costs, and currency. A more detailed model is being prepared to include the communication of this central catalog model to external cost databases, like RSMEANS, following the structure the NRCAN BTAP team has been developing in the last years [37]. The different instances of the city and building classes, including opaque walls, transparent walls, roofs, and grounds, are used together with the values of the imported specific economic costs to evaluate the costs for each retrofitting scenario. However, it is necessary to outline that all the elements to be costed and evaluated will come from the different construction archetypes applied to each building instance (based on function or functions of the building and vintage of this building). All the endeavors dealing with the costing strategies(and even embodied CO2 strategies) in the future will depend on the first good set of archetypes developed for each concrete location. The CAPEX chapters consider the number of replacements of each of the different equipments for each of the cases during the lifetime of the analysis, given the lifetime of each of these types of equipment.

Subsidies Calculation A chapter dedicated to subsidies has been included in the calculation, including subsidies such as investment cost subsidies (divided between construction, energy systems, and renewable energy systems) or tax reduction subsidies.

OPEX Calculation In financing analysis, OPEX is normally created by adding operational and maintenance costs. The composition of operational costs of costs was developed using the same procedure. Depending on the chosen archetype, a central file with the details of energy costs (hourly, monthly, yearly) has been populated to account for the conventional specific costs. Those costs are coupled with the Monthly Energy Balance (MEB) results and the peak load calculations described in the previous sections. As for maintenance costs, the structure is the same, but the costs are assigned depending on

the equipment size, which is done during the MEB+peak loads workflow.

Other costs and incomes calculation Costs related to externalities (or incomes related to the same aspect) have been included in the workflow, as well as incomes related to energy exportation to the grid.

End of life costs calculation End-of-life costs have been calculated based on [38], and a first approximation based on the different archetypes has been implemented.

3.3.4. Decision on the Retrofitting Scenarios

A set of three retrofit scenarios for the first retrofit workflow have been established as a first test case to validate the full methodology. These scenarios are the following:

- Scenario 0, Business as usual: current status of the building, with its skin, but replacing the energy systems once its lifetime is finished (with the same performance as the original).
- Scenario 1, Skin retrofit: the building's envelope has been updated to current legislation.
- Scenario 2, HVAC systems optimization + PV: an update on HVAC systems to best available technologies is done, coupled with using the available roof space for Solar Photovoltaics, in case the shaded area is less than 30% of the total area.
- Scenario 3, Deep retrofit: a scenario including both previous scenarios

3.3.5. Structure of Results and Persistence Databases

During a workflow, the city object will instantiate the values of all the results in the adequate classes (building, thermal zone). However, these data have to be stored before the end of the workflow so as not to let it disappear. In the workflows that have, as a goal, to develop individual building simulations, the results can be analyzed via the output tool of the same Building Simulator, or an output file can be produced. However, the data storage structure is essential when dealing with massive amounts of information and when precalculating results, which is a desirable output. In the case of the TOOLS4Cities hub, a PostgreSQL database has been prepared with the outputs of some of the key KPIs of all the simulations in the city. The key KPIs have been stored in a PostgreSQL database, but since for research purposes, the input datasets are also crucial, all the data used for the simulations (construction layers, usage layers, geometry) have also been stored. Pickles have been used to reduce the space requirements for these amounts of data [39]. The pickles are meant to store a massive amount of data that won't initially be made accessible, but that can be recovered later to understand the hypotheses behind any simulation.. Another critical step in obtaining large-scale results is the orchestrator. In Urban Building Energy Models, small chunks of the city have to be simulated to optimize the processes. The overlap between these "small packages" is vital because shading and adjacencies won't be considered in the boundaries of those packages, and, therefore, a clear design of the overlap calculation methodologies has been put in place.

3.4. Other Features Implemented in the Tool

Other features implemented in the TOOLS4Cities hub that can allow for more in-depth analysis of retrofitting actions and methodologies in the future (but have not been included in the present article) are:

- Energyplus workflow: apart from the Monthly Energy Balance, the platform allows for the dynamic simulation of buildings using Energyplus. Tests to up to 3000 buildings have been done up to the moment, with excellent performance.
- Integration with TRNSYS. A workflow generating the template files (dck) and the building file (b18) has been implemented in the hub. Small sets of buildings are modeled (up to ten), but the capability will be extended in the following months.
- Energy systems detailed modeling tool. The capacity of the TOOLS4Cities Hub to chain two workflows has been put in practice with this use case. Once a first workflow (Energyplus

workflow) has been used, the next workflow has been developed to analyze the feasibility of changing energy systems to heat pump/storage simulations.

- District Heating workflow. A double-chained workflow (based on the EP workflow) showing the ideal district heating system connection between the different buildings, following the geojson of the roads and sizing pipes, taking into account simultaneity and several building priorities, is in place (being improved at the moment).

4. Proof of Concept. Results

4.1. Implementation of retrofit scenario for Montréal

4.1.1. Geospatial Data Treatment

The first proof of concept for developing a retrofit calculation for an entire city was implemented on the island of Montréal. The geospatial data cleaning process has been extremely time-consuming, especially because of the deficient quality of the input datasets and the necessary data cleaning. The dataset was created by combining several open-source datasets provided by NRCan and the Montreal city government (Montreal property assessment units [40], Montreal 3D buildings (LoD2 model with textures), LiDAR aerial dataset of Montreal [41], Building Footprints NRCAN [42] and Layer 2 cadastral data for disaggregated building use type [43].

However, even if the initial datasets were quite complete, the differences in quality and completeness have forced us to follow specific necessary steps, like (i) splitting the satellite-derived footprints based on information from the city's cadastral parcel data, (ii) correcting building heights using LiDAR data, improving the accuracy of the analysis and, especially, (iii) incorporating an algorithm to detect and clean the overlap between the buildings. The final integrated result was a geojson database, with corrected building heights and an essential number of characteristics of each building (vintage, function, out of a list of more than 250 functions, number of households, number of stories). This geojson database was finally refined with a methodology to recognize neighboring walls in 2D and 3D.

4.1.2. Development of a Set of Archetypes

The approach of the current platform uses accurate geometries (if they are available) to build 3d models of the buildings, with their relationships with the environment (shading and adjacencies). However, some reference archetypes are needed to populate these models with constructive and usage data. The development of new archetypes has been an essential task.

4.1.3. Geometry Factory

The first step in our platform is to capture the geojson datasets (the platform can also capture CityGML, 3ds, obj datasets). Once they are captured, an extruded 3D model of the city, with its buildings and neighboring walls, is created. This is the instance city that will be used for the simulations. Even though, given the poor quality of the input data, in this initial use case, a geojson file with extruded buildings and flat roofs was used, in the cases where acceptable LoD2 and LoD3 CityGMLs are available, the platform allows for the simulation of roofs of different shapes.

4.1.4. Construction Factory

Québec, with its rich architectural heritage and diverse population, boasts many buildings. Built at different times and reflecting various architectural influences, these structures form an essential part of the city's urban fabric. As Québec continues its journey towards sustainability and energy efficiency, it becomes imperative to understand the construction of these buildings. Building sets, or the compilation of materials and methods used to construct a building, are the fundamental elements of any architectural structure. In the case of Québec's buildings, these building sets are remarkably varied, given its long history and the successive waves of architectural trends it has witnessed. From the masonry-clad row houses of the late 19th century to modern high-rises where glazing dominates

the façade, construction techniques and materials have evolved considerably. The layering of these building assemblies, or how different materials are arranged and integrated, has a direct impact on a building's performance, longevity, and interaction with its environment. For example, the order in which insulation, structural materials, and vapor barriers are placed can significantly influence a building's thermal performance and resistance to harsh Quebec winters. However, defining these building assemblies and their layering comes with its own set of challenges. One of the main limitations is the lack of documentation or detailed records of the construction processes of older buildings. Over time, many original plans, material lists, and construction notes have been lost or are inaccessible. This lack of data makes it difficult to accurately define building layers, particularly for structures built in earlier eras. The first archetypes developed for Montréal were based on the developments of NRCAN for the BTAP tool, which inherits a significant amount of work done by NREL. Those archetypes, however, have a limited amount of vintage buckets regarding constructive solutions. For example, all the buildings before 1980 (which, in Montréal, account for more than 65 percent of the overall surface) are considered to have the same construction type. For that, several sources have been consulted, and, based on an initial endeavour using the datasets from BTAP [37] as a basis, we developed new constructive types based on references of academic studies [44] [45] and analysis from consultancies [46]. The compilation of the best cases of the different references can be found in Tables 6 and 7.

Table 6. Composition of the walls. Archetypes for residential buildings

Area	Opaque walls	Transparent walls
Pre-1950	Wall with 10 cm brick, 10 cm LW concrete, 10 cm air gap, 1.2 cm plasterboard. U=1.498 W/m2K	Window with a glazing conductivity value of U=3.10 W/m2K, a marc conductivity of U=4.20 W/m2K, and an SHGC of 0.66.
	Roof membrane, insulation to achieve U=0.823 W/m2K, metal surface	
	Floor with insulation to U=0.678 W/m2K, 4-inch concrete, carpeting	
1950-1980	Wall with 10 cm brick, 10 cm LW concrete, 5 cm insulation, 10 cm air gap, 1.2 cm plasterboard.	Window with a glazing conductivity value of U=3.10 W/m2K, a marc conductivity of U=4.20 W/m2K and an SHGC of 0.66.
	Roof membrane, insulation reaching U=0.823 W/m2K, metal surface	
	Floor with insulation to U=0.678 W/m2K, 4-inch concrete, carpeting	
1980-2010	Wall with 25 mm stucco, 5/8" plaster, virtual insulation to achieve U=0.426 W/m2K	Window with a glazing conductivity value of U=2.8 W/m2K, a marc conductivity of U=4.20 W/m2K and an SHGC of 0.66.
	Roof membrane, insulation reaching U=0.276 W/m2K, metal surface	
	Floor with insulation to achieve U=0.459 W/m2K, 4-inch concrete, carpet	
2011-2020	Wall with 25 mm stucco, 5/8" plaster, virtual insulation to achieve U=0.247 W/m2K	Window with a glazing conductivity value of U= 2.2 W/m2K, a frame conductivity of U=3.1 W/m2K and an SHGC of 0.39.
	Roof membrane, insulation reaching U=0.183 W/m2K, metal surface	
	Floor with insulation to reach U=0.183 W/m2K, 4-inch concrete, carpet	
>2020	Wall with 25 mm stucco, 5/8" plaster, virtual insulation to achieve U=0.247 W/m2K	Window with a glazing conductivity value of U=1.9 W/m2K, a frame conductivity of U=2.20 W/m2K and an SHGC of 0.39.
	Roof membrane, insulation reaching U=0.138 W/m2K, metal surface	
	Floor with insulation to reach U=0.156 W/m2K, 4-inch concrete, carpet	

Table 7. Composition of the walls. Archetypes for non-residential buildings

Area	Opaque walls	Transparent walls
Pre-1950	Brick/ Stone/ Terracotta/Concrete with an overall U value of U=0.9 W/m2K	Window with a glazing conductivity value of U=5 W/m2K, a frame conductivity of U=4.20 W/m2K, and an SHGC of 0.8.
	Roof membrane, insulation to achieve U=0.823 W/m2K, metal surface	
	Floor with insulation to U=0.678 W/m2K, 4-inch concrete, carpeting	
1950-1980	Steel structure/Curtain wall, Brick/Stone Cladding: 0.1, Concrete: 0.1, Gypsum Plastering: 0.013.	Window with a glazing conductivity value of U=3.10 W/m2K, a frame conductivity of U=4.20 W/m2K and an SHGC of 0.66.
	Roof membrane: 0.002 m, Asphalt cover board: 0.01 m, Rigid insulation (e.g., MW Glass Wool): 0.10 m - 0.12 m, Steel trusses, joists, concrete decks, parallel-chord trusses and joists	
	Floor with insulation to U=0.678 W/m2K, 4-inch concrete, carpeting	
1980-2010	Curtain wall with veneer or precast cladding: 0.1 m, Insulation (e.g., Rigid board insulation): 0.10 m - 0.12 m, Interior gypsum plastering: 0.013 m. / Metallic Cladding: 0.006 m, Gypsum Plastering: 0.013 m, Insulation (e.g., Rigid board insulation): 0.10 m - 0.12 m. U=0.426 W/m2K	Window with a glazing conductivity value of U=2.8 W/m2K, a frame conductivity of U=4.20 W/m2K and an SHGC of 0.66.
	Roof membrane: 0.002 m, Asphalt cover board: 0.01 m, Rigid insulation (e.g., MW Glass Wool): 0.10 m - 0.12 m, Steel trusses, joists, concrete decks, parallel-chord trusses and joists U=0.276 W/m2K	
	Floor with insulation to achieve U=0.459 W/m2K, 4-inch concrete, carpet	
2011-2020	Brick veneer, with air space and insulation on steel or wood framing. U=0.247 W/m2K	Window with a glazing conductivity value of U= 2.2 W/m2K, a frame conductivity of U=3.1 W/m2K and an SHGC of 0.39.
	Roof membrane, insulation reaching U=0.183 W/m2K, metal surface	
	Floor with insulation to reach U=0.183 W/m2K, 4-inch concrete, carpet	
>2020	Brick veneer, with air space and insulation on steel or wood framing. U=0.247 W/m2K U=0.247 W/m2K	Window with a glazing conductivity value of U=1.9 W/m2K, a frame conductivity of U=2.20 W/m2K and an SHGC of 0.39.
	Roof membrane, insulation reaching U=0.138 W/m2K, metal surface	
	Floor with insulation to reach U=0.156 W/m2K, 4-inch concrete, carpet	

Infiltration losses in a building are defined as heat losses produced by the involuntary infiltration of outside air through the building envelope, producing an untreated air renewal effect in the zone that significantly affects building heat loss. These losses can, in some cases (especially in Montréal), represent significant percentages of the total heat demand of the building. However, it is not usual to carry out a rigorous treatment of these losses when simulating buildings since constant reference values are usually used or, in the best case, variables estimated according to generic scenarios. Airtightness

tests are a widespread practice in the United States and Northern Europe to determine the level of infiltration in single-family homes. However, this practice is not as widespread in large commercial buildings (including offices) due to various factors probably related to the complexity of the analysis or the underestimation of the derived effects, in addition to the fact that it is not mandatory in terms of regulations. Several methods have been used in recent years to simulate infiltrations in buildings, using values from previous results and tests. The three most widespread methods to simulate them in energy plus are the design flow, the ELA, and the flow coefficient model. The model used in the simulations is based on the NIST model, developed by [47]. This model relies on the floor values per outdoor wall. We analyzed the Energuide datasets for all residential buildings (buildings under 4 stories) (defined as IRLMs by the same Energuide). A set of 120.000 buildings was found, and a boxplot analysis (2) was developed to see the values of ACH50 (the variable shown in the Energuide dataset), before any renovations had been carried out (EVALTYPE=D). In the case of MURBs and other high-rise buildings, several academic and industry-based documents have been used [44] and [48]. These values have been implemented per decade of construction in the construction libraries. All these values have been converted into infiltration per outdoor surface to incorporate them into the NIST model. The capacity of the UBEM to understand the walls that are exposed and those that aren't (adjacent to other buildings) will be used with the outdoor flow per area method.

Table 8. Infiltration values used in the archetypes

Vintage	Infiltration (l/sm2 at 75 Pa)
Pre-1950	6
1950-1980	5
1980-2010	3
2010-2020	1.5
>2020	0.5
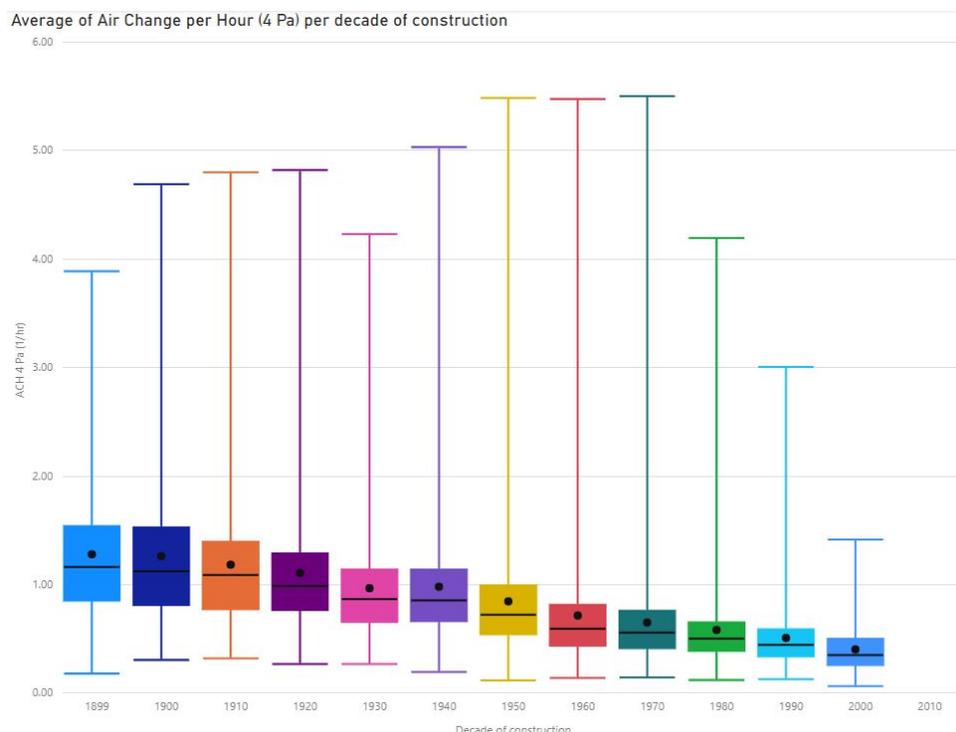
Figure 2. Selected zones to analyze single family/duplex buildings

Thermal bridges significantly affect building heat loss, particularly in MURBs with many balconies, which may constitute a potential condensation risk. Although the aim of the archetypes is not

to simulate in detail these situations, which are very specific for each project, we have carried out an initial analysis of the impact on MURBs to be able to incorporate its effect into the overall U-value of the various building construction archetypes, particularly in facades where thermal bridges will occur. The result has pushed us to increase the U value of the walls in MURBs by a certain percentage, depending on the vintage. When LoD3 buildings are incorporated, a more detailed analysis will be implemented using the Thermal Bridge Derating (TBD) methodology [49].

4.1.5. Usage factory

The usage factory uses the National Energy Code for Buildings of Canada, which uses fixed profiles for each building. We have two steps, one for the gains and another for the profiles. Those separate steps currently use a direct link to the repository of [37], where the NRCAN team has been updating the base data sources for the tool BTAP. As detailed in section 3, using public online data is a decision from the authors to reduce the maintenance costs of the Tools4Cities hub and allow for continuous updating.

Moreover, the factory structure already allows the integration of external stochastic generation profiles derived from [50], although it was decided not to use it for the retrofit use case.

4.1.6. Energy systems factory

For air renovation in buildings constructed after 2011, the air renovations will come directly from the NECB profiles, with heat recovery in the buildings from 2011 onwards.

Due to the lack of information available at the individual building scale, a methodology to randomly populate the energy systems factory has been implemented in the case of the HVAC factory. This methodology has established a separation between residential buildings, commercial buildings, and institutional buildings.

For the former, Figure 3 shows the primary sources and systems used to meet heating demand in Quebec residential buildings without distinguishing between single-family or multi-residential buildings. As the figure shows, electricity is the primary energy source for heating in Quebec. Not surprisingly, electric baseboard heaters are the province's most widely used heating appliance. In addition, natural gas is another popular fuel used in the province to run furnaces and boilers, mainly in low- and high-rise apartment buildings.
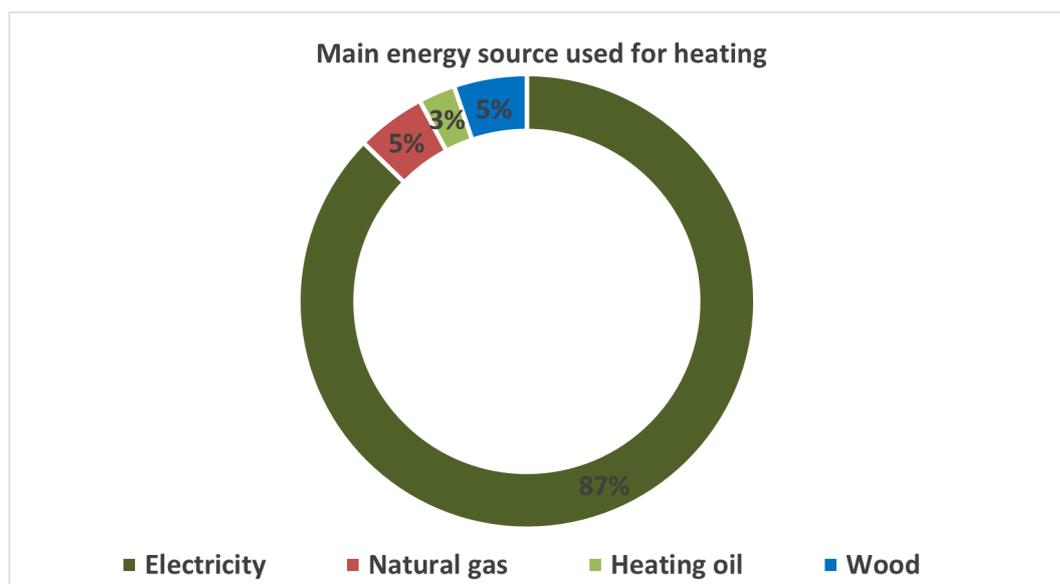
Figure 3. Splitting between different data sources

Figure 4. Assignation of energy systems per each type of building NECB

For the latter, the yearly balance in each sector was used to divide the buildings by gas and electricity consumption. Regarding the systems typologies, the NECB equivalence[51] was used (figure 4). These typologies show that a concrete HVAC system typology is suggested for every type of building. The systems, from System 1 to System 6 are:

- System 1: Unitary air conditioner with baseboard heater
- System 2: Four-pipe fan coil
- System 3: Single zone packaged rooftop unit with baseboard heaters
- System 4: Single zone make-up air unit with baseboard heating
- System 5: Two pipe fan-coil
- System 6: Multi-zone built-up system with baseboard heater

4.1.7. Cost Datasets and Hypotheses

The cost datasets used for the different subchapters for the Montréal case have been based on [52] and adapted to the local currency and increases in the price of materials since the edition of the book. We are conscious of the limitations of using worldwide references for local costs. Still, as a first approximation, the values look coherent to recent studies for Canada, as will be shown in section 4.2.4.

4.2. First Energy Results and Comparison with Real Datasets

Following the results of the first simulations, we analyzed in detail the origin of the differences between simulated and actual consumption. The sources of error were analyzed from two points of view: (i) input data and (ii) simulation assumptions. In the first stage, the analysis focused on discrepancies in area calculated by our tool, based on the city's various geospatial data sources (CityGML, geojson, and land valuation data) and consumption per square meter.

The authors followed two steps: the first was the consumption data analysis for specific typologies of buildings for which we had available metadata. We analyzed the data about two typologies of buildings: single-family homes/duplexes and public buildings.

The overall values for the full city were analyzed in the second step. The values for the simulation of the entire city have been compared with real datasets obtained from Montréal. In table 9, we can see the overall comparison of the results for the consumption of the different fuels yearly compared to the results of the first simulation of the base case scenario. From a global perspective, the results seem quite positive for a first global test. At this moment, detailed analysis procedures and calibration processes are being made to refine the archetypes and the geospatial inputs.

4.2.1. Single Family Houses/Duplex/Triplex

We first assessed the validity of the process in two significant areas of single-family / duplex / triplex buildings in Montreal: the first is in Pointe-Saint-Charles (between Wellington, Ash, Leber and Congrégation streets), with a majority of buildings built before 1946, and the second is in Saint Leonard (between Langelier boulevard, Robert boulevard, de Guyenne street and Louis Sicard street), with buildings from the 1960s-1990s. The ColouringMontreal tool (https://nextgenerations-cities. encs.concordia.ca/colouringmontreal/), developed by the CERC team as part of the ColouringCities initiative, also helps us detect the best neighborhoods to analyze. This gave rise to a total of 500 buildings.

Figure 5. Air renovations per hour 4 Pa

The results concerning the calculated building area show that the simulation input data gives higher values than the actual values obtained through property valuation data. The fact that the methodology captures the 3d exterior volume incorporates errors such as overestimating the surface area of second floors or incorporating partially covered outer volumes into the analysis. Figure 6 shows that the average floor area calculated with the models is 60% higher than the values extracted from the property data. The lack of good data is at the root of this error. Some novel methodologies are being applied to detect cleaner methodologies via AI and ML and will be incorporated into future platform developments [53].
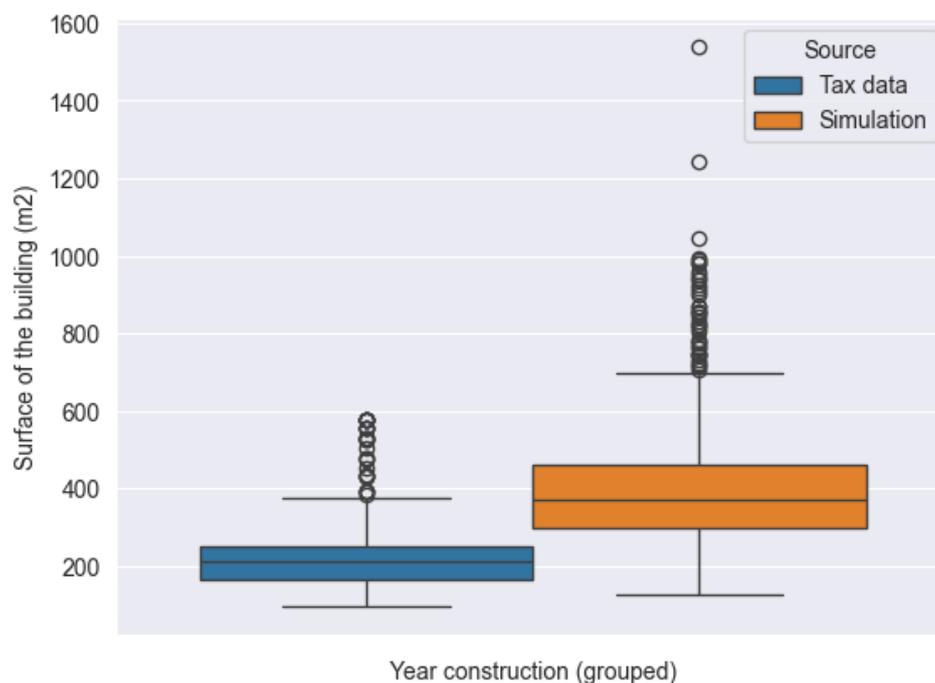
Figure 6. Comparison surface between tax office datasets and geojson extrapolated surface

Given the fact that the methodology developed by the CERC uses real geometry, including shadings and adjacencies, and scaling down these geometries without losing the geometrical links, the CERC team has decided to incorporate a downsizing factor based on the value of the real surface from tax evaluation data and apply the factor between the surface tax data and the simulated surface data to the final energy results.

Regarding energy consumption, since access to the actual data for individual building consumption for these buildings is not possible, the authors compared them to the average values for single-family homes in Quebec. There is a particular bias due to Quebec's different climates. Still, this analysis allows us to see the distance between simulated and real buildings, in this case, per square meter. We have grouped the buildings by the same groups in the National Energy Consumption Database document and extracted the heat and domestic hot water consumption data from our simulations to compare them. The data for buildings later than 1977 is not too high (less than 5% of buildings), but the trend shows us that the values obtained for the tool, concerning specific consumption per square meter, are very much in line with statistical consumption.

We have also added a comparison with 19.111 buildings from the Energuide dataset [54], filtered only for Montréal and for single-family/duplex/triplex houses. In this case, we can compare total final energy consumption with this dataset.

In both cases, the authors observe that the UBEM tool underestimates the consumption of old buildings and overestimates the consumption of new buildings. Given that the infiltrations for these types of buildings come from the Energuide dataset, the parameter to be adjusted for these buildings will be the constructive layers for both old and new buildings, especially from 1977 onwards.
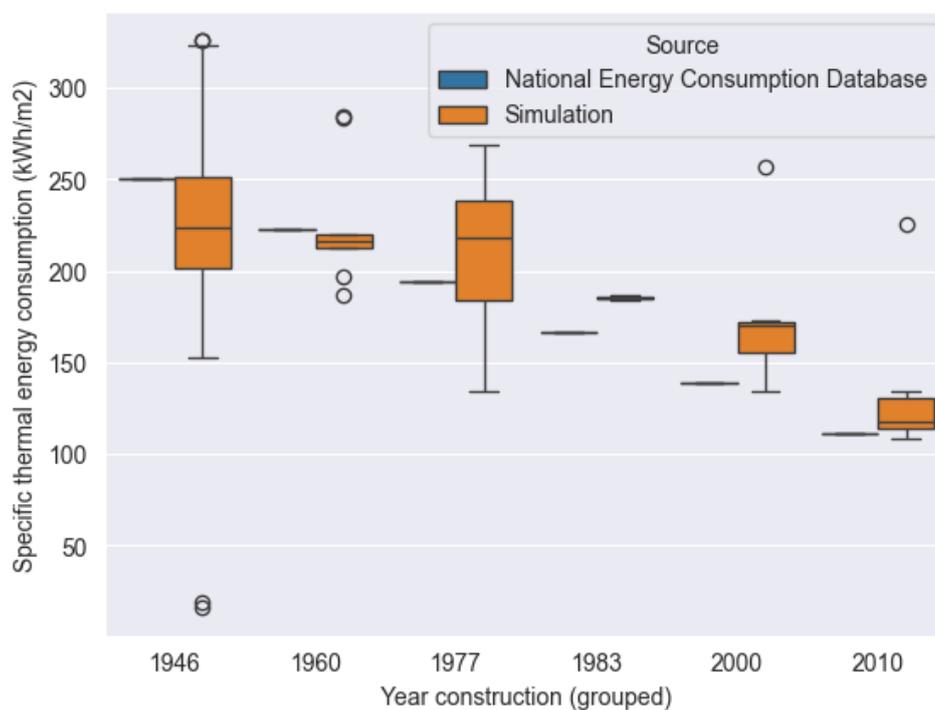
Figure 7. Final thermal specific energy in single-family/duplex/triplex buildings compared to National Energy Consumption data
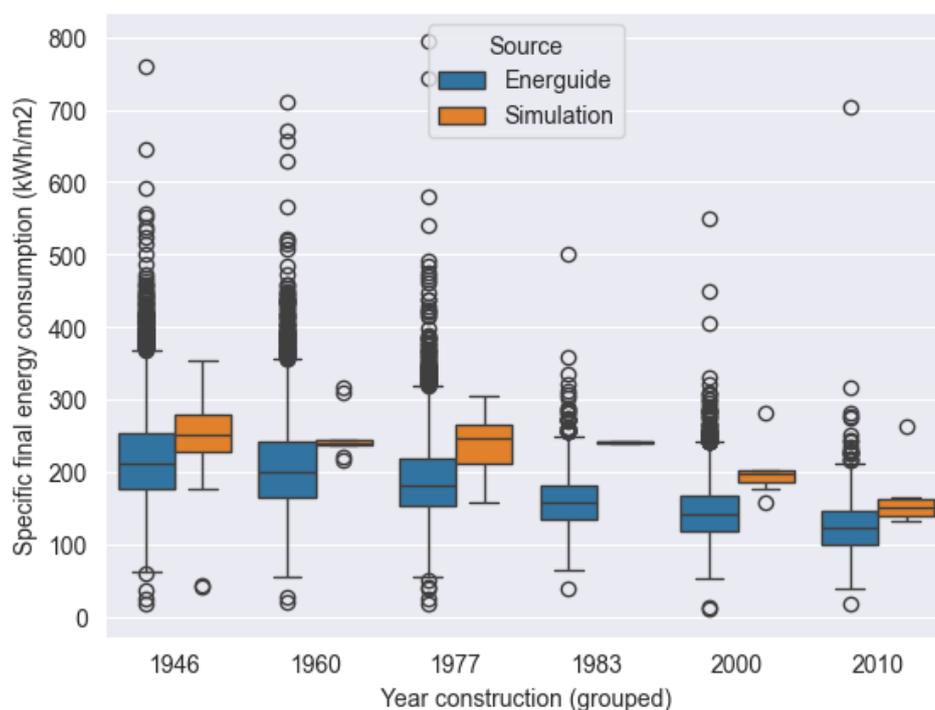

Figure 8. Final specific energy in single-family/duplex/triplex buildings from Energuide dataset

4.2.2. Institutional Buildings

For tertiary sector buildings, we analyzed the list of buildings published by the City of Montreal and their energy consumption. We evaluated 11 of these buildings in terms of surface area and consumption. In terms of surface area, the results are uneven. The differences are significant, and in

general, we observe that using the 3d models extracted from the city's data gives us 30% more surface area than the data published by the City.
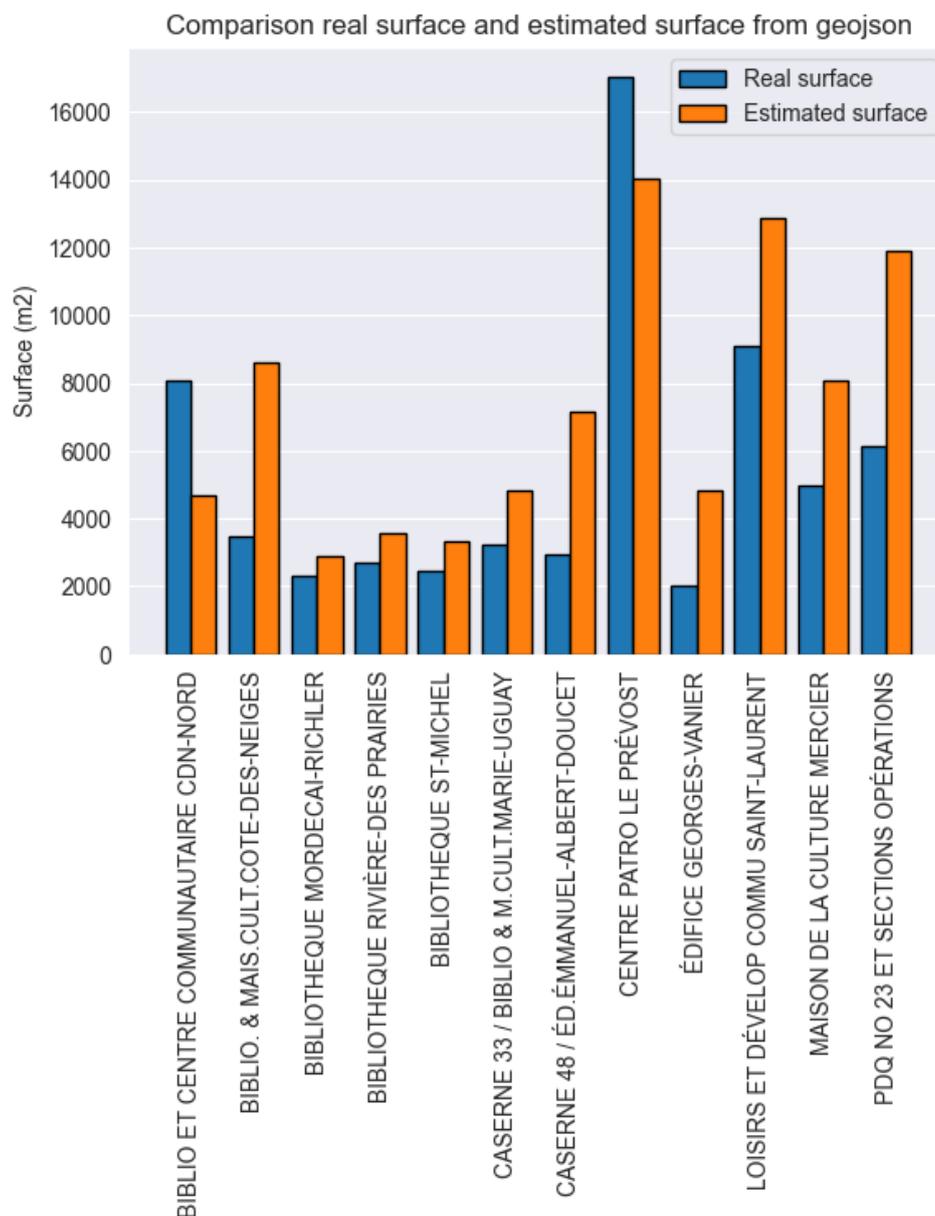

Figure 9. Comparison surface institutional buildings between city source and extracted data

Concerning energy consumption (final energy), we compare the results with the municipality's list. We note that the difference is significant in some cases, but we're in line in order of magnitude. The impact of meteorology has been analyzed (HQ data for the MacTavish weather station in Montreal compared with the epw file used), and, at day-degree levels, the impact is less than 5%.
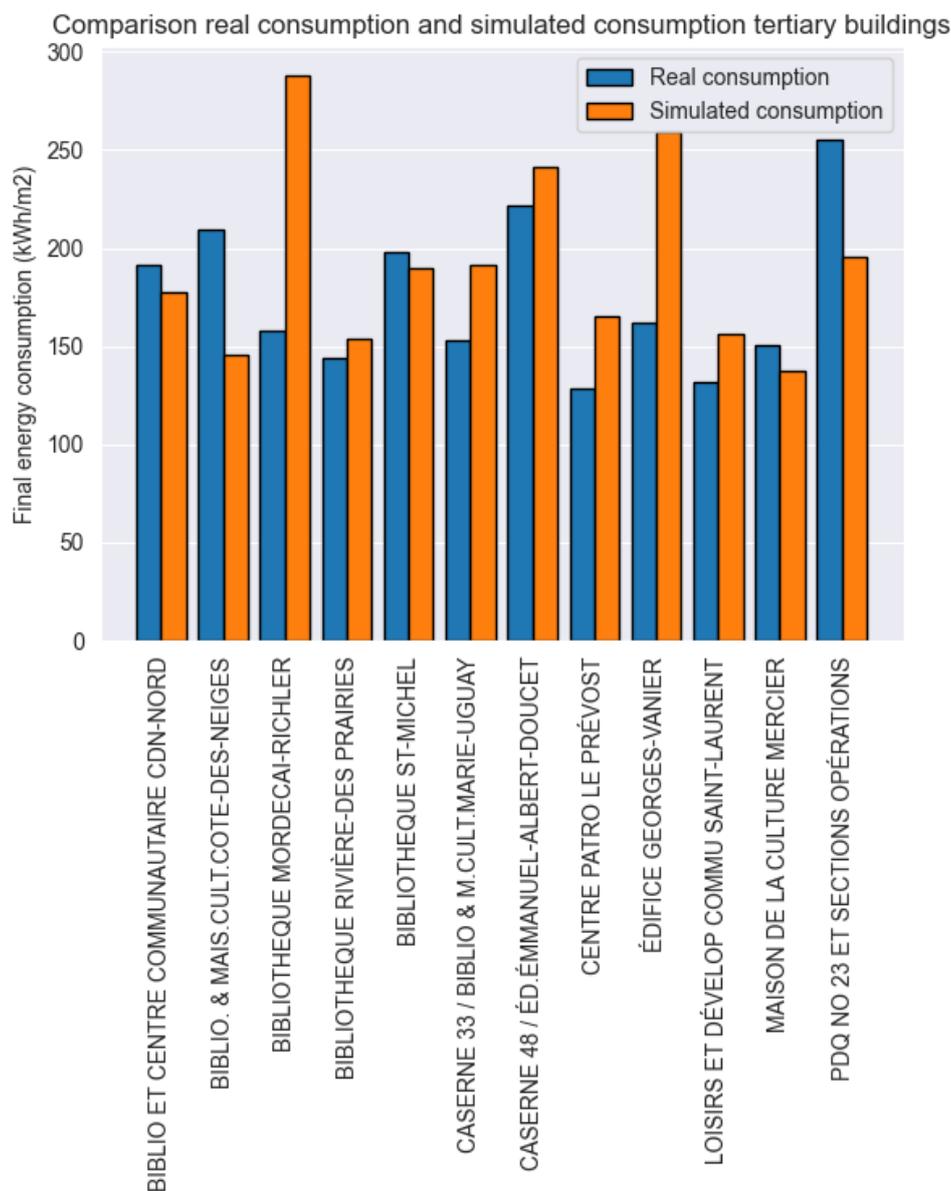
Figure 10. Specific final energy consumption institutional buildings

The results from the validation for individual buildings show that the tool can also be used to find inconsistencies between different datasets. The fact of having a tool that can help fill the gaps and detect anomalies in consumptions and geometrical data seems an exciting capability, complementary to the tasks that are being done using Machine Learning (ML) techniques in other areas of geospatial data treatment [55].

4.2.3. General Results for the Full Island

The availability of the aforementioned data from Hydro-Québec and Energir has allowed us to get the yearly data for the full island, except Westmount, where a municipal electrical company exists (Hydro-Westmount). The simulation for the full city's final energy consumption has been compared to the data coming from Hydro Québec and Energir. Since some of the city's final energy consumption comes from sources other than electricity or natural gas, there is a gap between those datasets. Although we understand it is a pessimistic hypothesis, we have used the data from 2015 [56] and considered no new development has been built using these fuels.

The results for the overall city in table 9 show an 11% difference between the data we have simulated and the real data from the city. This shows a good alignment between the simulation results and the city's global consumption. However, it should not hide the differences between real and simulated surfaces we have seen in single-family buildings. The non-validated sectors (commercial buildings) should be analysed in depth to confirm the UBEMs accuracy in evaluating their demands.

Table 9. Comparison of real consumption and simulated values

	Total final energy consumption	
Consumption from Hydro-Québec and Energir for 2022	42,311.52	GWh
Estimated diesel, kerosene and fuel-oil consumption	3,897.39	GWh
Subtotal	46,208.91	GWh
Simulation using tmy Montréal	40,923.96	GWh
Difference between simulation and reality	-11%	

4.2.4. Economic Results

Once the energy consumption for each building in Montréal has been calculated, the retrofit costs workflow has been implemented with the following economic parameters:

- number of years=31,
- percentage credit=0,
- interest rate=0.04,
- consumer price index=0.04,
- electricity peak index=0.05,
- CO2 index=0.06,
- CO2 price=30 $/ton ,
- electricity price index=0.05,
- gas price index=0.05,
- discount rate=0.03,
- retrofitting year construction=2020,

To ensure the adequate robustness and speed of the full process, a test case was done to calculate the costs for a set of buildings. Although this part of the process is not cross-validated with any actual project, references such as [57] and [58] have helped us see if we were aligned with the values of use cases. A first set of economic calculations for 108 residential buildings in the second zone selected in 4.2.1 has been implemented with the previously established hypotheses. The results are plotted in the following figures.

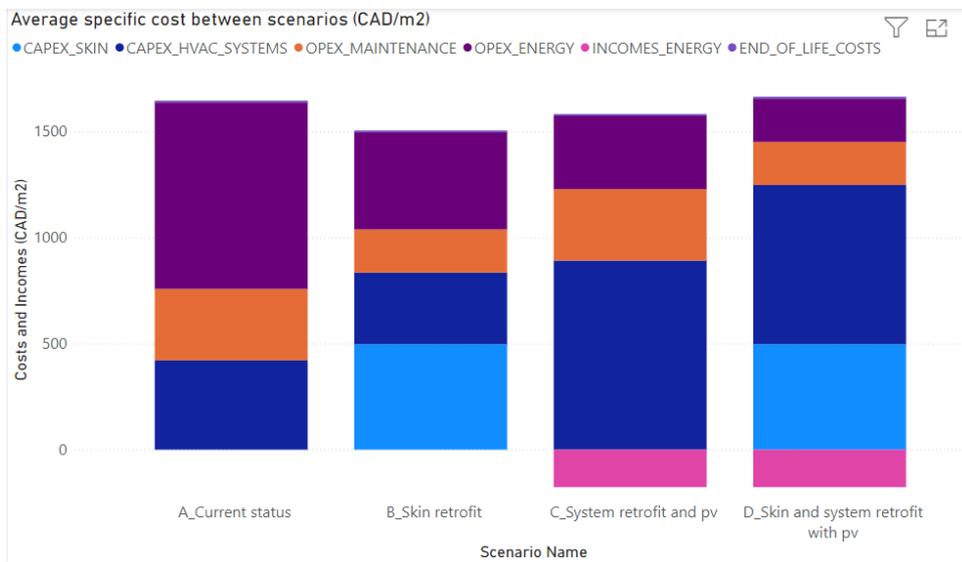**Figure 11.** Boxplot costs for the different scenarios

Figure 12. Average costs and incomes per categories different scenarios
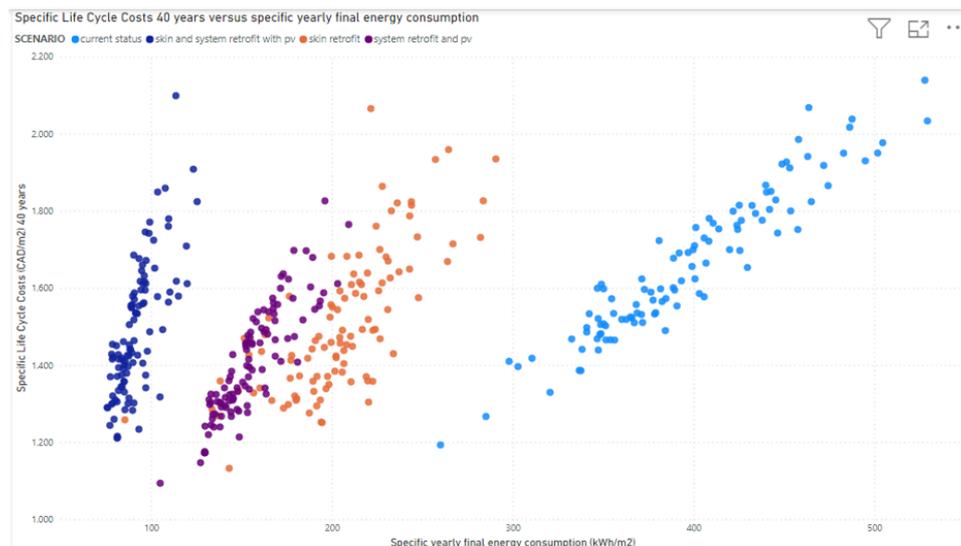
Figure 13. Pareto graph for the analyzed datasets, with three retrofit levels

The results show that the values mentioned above in the literature are coherent with CAPEX and OPEX of similar building types (residential)[59].

4.2.5. Visualization of the Results

A significant effort has been made in the presentation and accessibility to the results for the general public. As mentioned in 2.2, this is one of the big gaps we detected in other tools.

A web page showing all the buildings in the city is used to facilitate access to the city-simulated data and let the user interact with the data. The results are shown using two different services in the Citylayers test web page (https://nextgenerations-cities.encs.concordia.ca/citylayers/).

The first is the single building retrofit visualizer, which shows the individual results of the base case simulation and the three retrofit options once a building in the city is selected. Although the detailed results are stored in a PostgreSQL database, an effort was made to simplify the shown data. A qualitative radar shows the best option for the single building case according to different parameters and criteria. It is plotted with KPIs for maximum peak loads, 40-year life cycle costs, and yearly energy consumption.
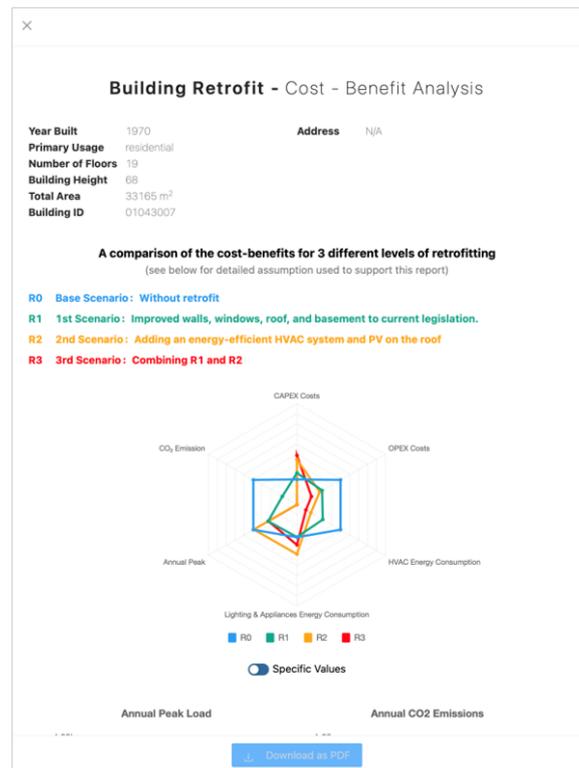
Figure 14. Single building retrofit front end

The second is the multiple-building retrofit tool, which can be used to select several city areas. A scatterplot comparing CO2 emissions, LifeCycle Costs, and energy consumption with the vintage of the buildings is shown to help the user decide on the best retrofit strategies for each building segment. The buildings can be selected in the scatterplot, and the different strategies can be applied to each, obtaining a final value for each retrofitting strategy (a combination of the previously defined Scenarios 0, 1, 2, and 3).
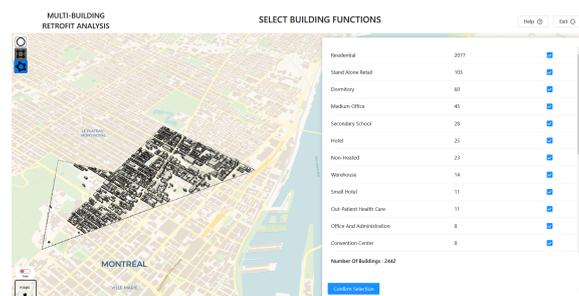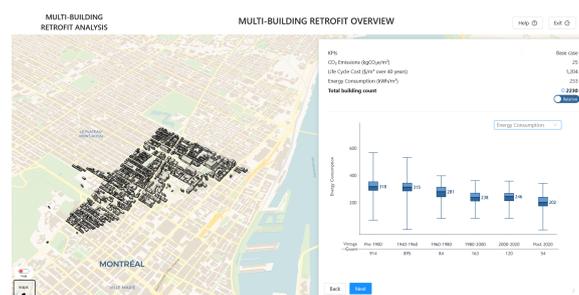

Figure 15. Multi-building retrofit selection

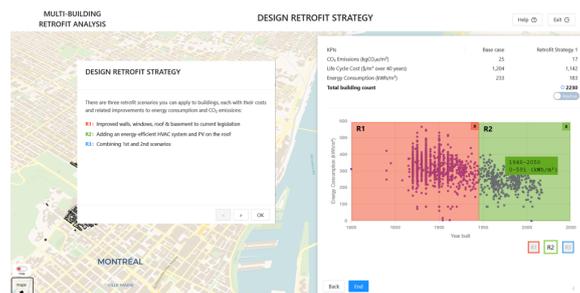**Figure 16.** Multi-building retrofit current scenario analysis

Figure 17. Multi-building retrofitting strategy selection

5. Conclusions

We presented a tool that can almost fully automatedly simulate large city areas (even in a big city such as Montréal). Unfortunately, part of the data processing must be manual, and a thorough processing effort was needed because of the errors and inconsistencies of all geospatial and constructive data sources. Nevertheless, the pipeline of the process is thought to detect those inconsistencies and, in most cases, can fill the gaps with sensible hypotheses. To fine-tune the models, we need access to much more detailed data at the temporary building level, especially in the tertiary sector. In the coming months, an initiative to capture data from the city of Montreal and evaluate it with Concordia's tools will be set up in collaboration with the city, Hydro Québec, Energir, the Ministère de l'Environnement, de la Lutte contre les Changements Climatiques, de la Faune et des Parcs, and Concordia. In conclusion, NGCI's new set of urban building simulation tools enables rapid analysis of the costs and carbon emissions of every building in a city. Based on archetypes for different building uses, the models make it possible to compare different renovation options. In this way, we hope to support and accelerate actions to decarbonize the municipal built environment.

6. Future Steps

The goal for the UBEM that we have developed, based on open-source criteria, modularity, flexibility, and integrability, is that it never stops evolving in the next years. The development goal is to make the framework available to the full research community and allow for integrating new workflows. Although any researcher can use the repository from the CERC, clone it, and use it mainly, we encourage using the manuals and the hub methodology to increase the tool's capabilities.

Author Contributions: Conceptualization, O.G. ; Literature search: P.A.M. and O.G. ; Methodology, O.G. and P.A.M.; Formal analysis, O.G.; Investigation, O.G.; Resources, O.G. and C.G.; Data curation, O.G.; Writing—original draft and review, O.G.; Writing—review & editing, P.A.M. and U.E. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: All data used in this project are available in the CKAN dataset from the Next Generation Cities Institute. Please contact the authors to have access to the necessary folders and the open-source tools from the NGCI. The gitea with all the data is in https://nextgenerations-cities.encs.concordia.ca/gitea/. The CKAN is in https://nextgenerations-cities.encs.concordia.ca/ckan/

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AI Artificial Intelligence. 19

BTAP Building Technology Assesment Platform. 10, 13, 17

CAPEX Capital Expenditures. 9, 10
CEA City Energy Analyst. 2
CERC Canada Excellence Research Chair for Sustainable Cities and Communities. 19, 20
CityGML City Geography Markup Language. 2

DHW Domestic Hot Water. 8

ECS Environmental Control System. 2
EnergyADE Energy Application Domain Extension. 2

GHG GreenHouse Gases. 1, 6
GIS Geographic Information System. 6

HVAC Heating, Ventilation, and Air Conditioning. 2, 11, 17, 18, 31

IoT Internet of the Things. 31

LCC Life Cycle Costing. 1, 9
LiDAR Laser Imaging Detection and Ranging. 12
LoD Level of Detail. 2, 10, 12, 17

MEB Monthly Energy Balance. 10, 11
ML Machine Learning. 19, 23
MURB Multi-Unit Residential Buildings. 16, 17
MVP Minimum Value Product. 7, 10

NECB National Energy Code for Buildings Canada. 18
NRCAN Natural Resources Canada. 10, 12, 13, 17
NREL National Renewable Energy Laboratory. 3, 13

OPEX Operational Expenses. 10

PV Solar Photovoltaics. 11

RC Resistance Capacitance. 2, 3
ROM Reduced Order Models. 2

SRA Simplified Radiosity Algorithm. 2, 8

TBD Thermal Bridge Derating. 17
TEASER Tool for Energy Analysis and Simulation for Efficient Retrofit. 3

UBEM Urban Building Energy Modelling. 2, 3, 5, 6, 7, 20, 24, 28, 29, 30
UHI Urban Heat Island. 2
UMI Urban Modeling Interface. 2
URBANOPT Urban Renewable Building and Neighborhood Optimization. 3

Appendix A

Appendix A.1

Appendix B Software perspective

Delving more into the details of each software component will allow a more detailed understanding of the full architecture of the development used to create the data structure that lies at the core of the UBEM.

Models

A model is specific to a particular domain (e.g., Transport or Waste or Building Physics, etc.) and focuses on representing a single system (or a small number of closely related systems) within that domain. A model is an algorithmic representation of an urban system capable of making predictions of the system's behavior or properties based on a definable set of inputs.

The TOOLS4Cities Platform architecture is designed to integrate with any model, provided it can run on a computer. This way, the Platform has no theoretical restriction on the scope of urban phenomena it can simulate, and it is also infinitely extendable over time as more models are added. Therefore, the platform uses 'Model' solely as defined above and will use the full names of these other models to differentiate them. Those models do not necessarily need to be developed with the same structure as the rest of the code developed by the CERC team, but they can be adapted to match the inputs and outputs required by the platform. However, they must be 'free-to-use' to be incorporated seamlessly into the platform. As such, they do not form part of the coding environment of the Platform; instead, they are separate programs that must be accessible to the Platform via their API (Application Programming Interface). The requirements established by the CERC platform, as with any model used by the platform, all third Party Models must support the following:

- Use a definable and program-accessible set of Input Data
- Algorithms or calculations that use the Input Data.
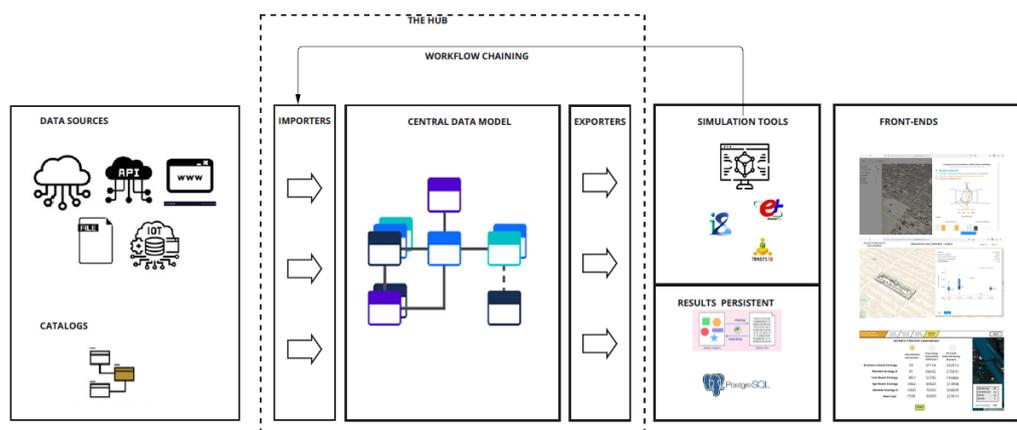- Product a definable and program-accessible set of Output Data.

Figure A1. General view of the platform

Central Data Model

A data model is a formalized structure that shows how data is to be represented within a system. Importantly, the data model is separate from the data itself and says nothing about where the data is physically stored. A data model is an abstract structure that allows us to understand the relationships between all pieces of data in the system.

Data Models

At the heart of the platform are the Data Models. Given all the aforementioned gaps in the existing UBEM models, it appears that it is a clear necessity to structure all the data. This data structure will be used to create clear interactions between the different sectors of urban decarbonization use cases and, at the same time, be able to use UBEM tools as multi-faceted tools and adapt the final tool at any moment.

For this reason, the authors developed a clear structure of how a city should be developed in Python

classes. Each class (see UML diagram) has clear relationships with others and can propose association, multiplicity, aggregation, composition, inheritance, or realization [60]. The objective is to create a metamodel of the full city, with clear relationships of the essential parameters of each element that compose it and the relationships between its agents.

Developing a generic metamodel of building data (especially energy data) has been an objective in the last years from several initiatives, especially from energy utilities and the HVAC IoT sector. IFC, Brick, Haystack, and Bigg [61], [62] are examples of metamodels and ontologies that have tried to structure and organize the metadata from the buildings to be able to understand their behavior and, in most cases, act over it.

The building-level perspective, though, is too detailed, and although some of the members of the same team have been developing a metamodel [61], the goal of the platform is to capture less data but is still compatible with the previously developed model.

The central data model consists, essentially, of a class called city that contains several components that will be stored in a structured manner. Buildings, roads, networks, public spaces, and waste separation plants are some of the objects that compose the city. Since the present article focuses on developing an urban retrofit model, a more detailed description of the building class will be developed.

The building class contains all the necessary information to develop global decarbonization analyses on buildings. The structure of internal zones, thermal zones, and thermal boundaries (surfaces, thermal openings, layers,..) contains all the input metadata that will allow for export to tools that can develop analyses such as energy consumption calculation, life cycle analysis, energy retrofitting analyses, demand response strategies (at urban level) or others. The flexible construction of the classes, with an orientation more similar to an ontology than a fixed metamodel, with links and relationships that can be changed, allows for future code re-structuring. Using classes in Python instead of fixed XML structures allows for the fast use of links and the flexibility of connections.

A data model can have classes with duplicated properties, but the Central Data Model uses the concept of inheritance to remove this duplication. The benefit of this is massively reducing the coding effort, making the code more reusable and easier to read. It also explicitly defines a relatedness between entities that helps the data model reflect their real interactions.

It is essential to underline that classes are only templates that guide how to represent the data of different entities. Once the classes are called, the values that fill the templates are objects that are instantiated, reserving RAM memory and initializing the property values. These objects disappear once the process is finished. Therefore, the goal of the central data model is not to produce anything but to give adequate structure for the later usage of the captured values.

Data

In the case of the Platform, the data used by the different workflows and models is divided between catalogs, datasets prepared by the CERC group, and online available data. The strategy of the CERC is to access, as much as possible, open data sources and, using the factories, link them to the central data model. Whenever a source of data requires deeper development or the origins of the data sources are unclear, a catalog is developed. The XML format is encouraged because of its clear structure and access facility. However, when the sources are well documented and maintained, external sources are directly accessed by the platform code. This strategy ensures that the amount of catalogs and data to be maintained is minimized and relies on maintaining these data sources by the external entities that are the data stewards. For example, the GitHub repository of the BTAP tool from NRCAN is used [63].

Central Data Model
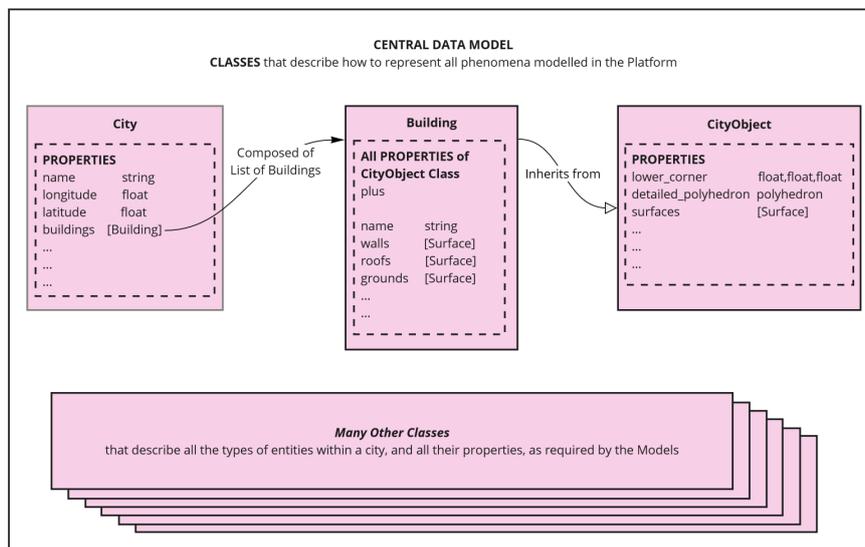
Figure A2. Characteristics of the central data model

Factories

As described in figure A4, there are two blocks with coding components called 'Import Factories' and 'Export Factories'. Factories are essential, and most of the authors' efforts to achieve homogenization of different datasets come from those factories. The Factories are specialized functions that manage exchanging data between various external (non-platform) data representations and the Platform's internal Central Data Model.

Figure A3. Detail on the platform workflow regarding input datasets

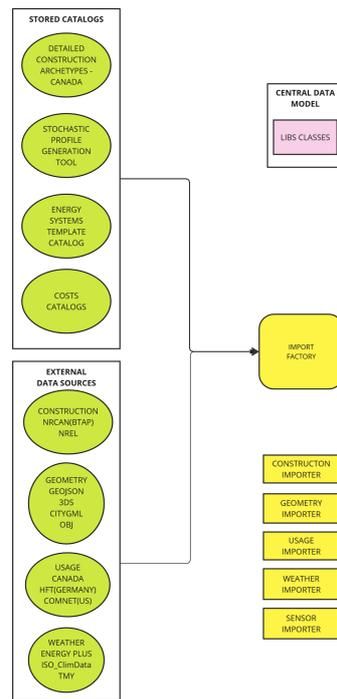When data is converted into the Central Data Model representation, it is said to be 'imported'. When data is taken out of the Central Data Model into an external representation, it is said to be 'exported'. All factories need a handler to know which type of format is being imported, as several ones are supported by each factory.
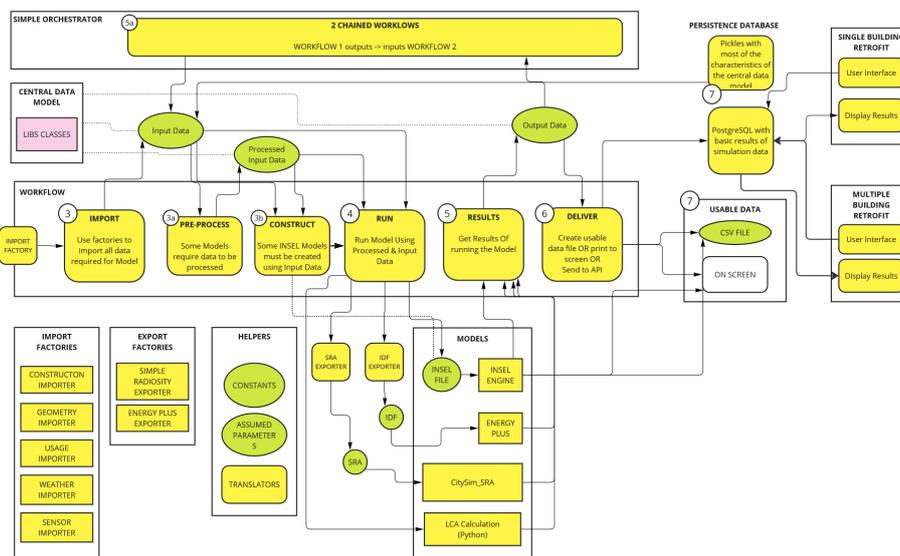
Figure A4. Steps of a complete workflow using the platform

Import factories   At the time of writing, there are Import Factories to handle the following categories of data: Construction, Geometry, Weather, Building Usage and Sensors. The Import Factory for each data category can then handle various data formats. For example, the Geometry Importer can import data in both CityGML format, Geospatial JavaScript Object (GeoJSON) format, STL format,

3DS (Rhino) and also OBJ format. The same platform includes the capacity to convert one object type into another. In this case, the handler of the **Geometry factory** supports at least cityGML, OBJ, rhino, and geojson. If a geojson file will be imported, we must provide the handler **geojson**. Weather importers incorporate how to enrich a City Object with weather data (from an EPW or a TMY file). The construction factory imports constructive characteristics for each of the buildings of a city class, which are structured in constructive archetypes. The constructive handlers available at the moment are NREL, using the data from the DOE archetypes, and NRCAN, using BTAP standards, The usage factory, at the moment, enriches each of the buildings with the usage associated with each of the building functions/archetypes and has three handlers: one using the COMNET values and profiles, another using the data from the National Energy Code of Canada, and the third using a stochastic profile generator developed by one of the researchers of the team.

Export factories   These do the same thing in reverse, taking some category of the data represented within a City Object and converting it into an externally recognized format. This externally recognized format will be used in the different models to run use cases. At the moment, several handlers have been developed in the hub and some more are being developed. The stable handlers, at the moment, are the Energyplus handler, the Monthly Energy Balance handler, the SRA handler. The handlers being built are an RC model handler based on an R1C1 model (CityBEM) and the TRNSYS handler.

Helpers   Helpers are considered part of the factories and are the tools used to perform several functions in support of the homogenization of the central data model parameters. Helpers are used to define assumptions for models, declare constants, and perform data translations.

Appendix B.0.1 Development of workflows

Once the pieces of the structure have been developed, the composition between its parts is to be defined. So far, we have considered a generalized Workflow and noted how its primary role is to manage the flow of data from external sources to one or more Models, which it runs and then collects their resulting output data.

Looking at more details, workflows need to do several specialized things to carry out this role. The schema above includes more workflow components, explaining the different workflow relationships. The schema also shows a series of numbered steps. These are the fundamental steps that every Workflow needs to take, regardless of which Model it is managing. The different steps followed by workflows are:

Step 1 – IMPORT

Every Workflow must first retrieve all the required data as input for the Model(s) it manages. Data sources may be in proprietary formats (for example, building geometry may be in CityGML or OBJ format, while weather information may be in EPW, TMY2 or ISO CLIM format). Whenever a new Model is integrated into the Platform, we also develop new Factory Importers if data formats are not already supported. The Import step is completed when the imported data is converted into a format that all other Workflows in the Platform can understand. This format is known as the Central Data Model and has been explained more fully previously. Suffice it to say, once the IMPORT step is complete, all the data from external sources is now fully represented internally in the Platform,

Step 1a – PRE-PROCESS (Optional)

Some Models require that the input data retrieved from external data sources be altered in some way, before it is ready to be used by the Model. There are many examples of this: 'Data Cleaning' (detecting and removing erroneous values; 'Data Calibrating'(modifying values based on auxiliary reference data); 'Data Pruning' (removing unnecessary data branches); 'Data Compression' (making

the size of data more manageable). We call these steps 'Pre-Processing' and the Workflow is responsible for carrying these out across all of the Input Data, as required.

Step 1b – CONSTRUCT (Optional)

For some models, the Input Data itself can change the Model. For example, an INSEL Model for calculating building energy demand may depend on the building geometry. The connectedness of thermal zones and surfaces within a building directly impacts the logic connections in a building simulation Model. As a result, the Input Data must be analyzed and used to create the required building simulation Model (INSEL, Energyplus, or TRNSYS). This is done programmatically by the CONSTRUCT step of the Workflow.

Step 2 – RUN

Once the Workflow has the Input Data ready it is time to RUN the Model(s). This could be as simple as making one call to the Model with one data set. Or it could involve calling the Model multiple times with subsets of the data. For example, the Input Data may include 300 buildings in a city district, and the model can calculate them all, either using a multi-zone approach to calculate different buildings or running all the buildings in a model one by one. If the Model is a Custom Model, written in Python, the Input Data can be sent using the format described by the Central Data Model. However, if the Model is an INSEL Model, the Input Data needs to be sent using an INSEL file template. And if the Model is a 3rd Party Model (e.g. Energy Plus), the Input Data must be sent in the proprietary format(s) understood by that Model. (e.g. IDF files for Energy Plus). As with importing different data formats, the Platform helps Workflows here, providing the Factories that can easily export data to the format required. Again, whenever a new model is integrated into the platform, we also develop new factory exporters if there are any data formats that are not already supported. The RUN step is complete once the Workflow has called the Model across all the necessary Input Data.

Step 3 – RESULTS

All Models generate some result or Output Data. The workflow must collect this data so that it can be used. INSEL Models can output their data directly to the screen. They can also write their data to files (CSV and other formats). Energy Plus outputs text files. Custom Models may do any of the above or return their results programmatically. The Workflow must retrieve the Output Data from any of these sources, using Factory Importers again where necessary, and convert the Output Data once more into the format described by the Central Data Model. This allows the data to be accessed by other Models later in the same Workflow or even other Workflows after this one completes. When the Workflow calls a Model multiple times with different subsets of Input Data, the Workflow must also retrieve the results after each call, building up a full set of Output Data.

Step 4 – DELIVER

With the Output Data now retrieved and stored, the final responsibility of the Workflow is to make the data available to whoever needs it. There are multiple options here. The Workflow may write the data to output files, display the data on the screen, and ensure the data is available for new Workflows; it may ensure the data is available for 3rd party Apps. In future versions of the Platform, this part of the Workflow is expected to be more formalized. Figure A4 shows the relationship between the Central Data Model and the Factories. The next two sections explain how these both support the Workflow

References

1. Programme, U.N.E.; .; Construction, G.A.f.B. 2020 Global Status Report for Buildings and Construction: Towards a Zero-emissions, Efficient and Resilient Buildings and Construction Sector - Executive Summary **2020**.

2.   Ding, G. Demolish or refurbish – Environmental benefits of housing conservation. *Australasian Journal of Construction Economics and Building* **2013**, *13*, 18–34. doi:10.5130/ajceb.v13i2.3322.

3.   Power, A. Does demolition or refurbishment of old and inefficient homes help to increase our environmental, social and economic viability? *Energy Policy* **2008**, *36*, 4487–4501. Foresight Sustainable Energy Management and the Built Environment Project, doi:https://doi.org/10.1016/j.enpol.2008.09.022.

4.   Schwartz, Y.; Raslan, R.; Mumovic, D. Refurbish or replace? The Life Cycle Carbon Footprint and Life Cycle Cost of Refurbished and New Residential Archetype Buildings in London. *Energy* **2022**, *248*, 123585. doi:https://doi.org/10.1016/j.energy.2022.123585.

5.   Crawford, K.; Johnson, C.; Davies, F.; Joo, S.; Bell, S. Demolition or Refurbishment of Social Housing?. Technical report, UCL Urban Lab and Engineering Exchange for Just Space and the London Tenants Federation, 2014.

6.   Tetteh, M.O.; Boateng, E.B.; Darko, A.; Chan, A.P.C. What Are the General Public's Needs, Concerns and Views about Energy Efficiency Retrofitting of Existing Building Stock? A Sentiment Analysis of Social Media Data. *Energy and Buildings* **2023**, p. 113721. doi:10.1016/j.enbuild.2023.113721.

7.   Zhang, H.; Hewage, K.; Prabatha, T.; Sadiq, R. Life Cycle Thinking-Based Energy Retrofits Evaluation Framework for Canadian Residences: A Pareto Optimization Approach. *Building and Environment* **2021**, *204*, 108115. doi:10.1016/j.buildenv.2021.108115.

8.   Lohse, R.; Mørck, O.C.; Zhivov, A., Eds. *Deep Energy Retrofit—Case Studies: Business and Technical Concepts for Deep Energy Retrofit of Public Buildings; Energy in Buildings and Communities Programme; Annex 61, Subtask A*; Springer International Publishing: Cham, 2023. doi:10.1007/978-3-031-17517-6.

9.   Volans. Pattern and APPG on Fair Banking. (2021), 2021.

10. Prabatha, T.; Hewage, K.; Karunathilake, H.; Sadiq, R. To Retrofit or Not? Making Energy Retrofit Decisions through Life Cycle Thinking for Canadian Residences. *Energy and Buildings* **2020**, *226*, 110393. doi:10.1016/j.enbuild.2020.110393.

11. Keirstead, J.; Jennings, M.; Sivakumar, A. A review of urban energy system models: Approaches, challenges and opportunities. *Renewable and Sustainable Energy Reviews* **2012**, *16*, 3847–3866. doi:10.1016/j.rser.2012.02.047.

12. Allegrini, J.; Orehounig, K.; Mavromatidis, G.; Ruesch, F.; Dorer, V.; Evins, R. A review of modelling approaches and tools for the simulation of district-scale energy systems. *Renewable and Sustainable Energy Reviews* **2015**, *52*, 1391 – 1404. doi:http://dx.doi.org/10.1016/j.rser.2015.07.123.

13. Robinson, D.; Campbell, N.; Gaiser, W.; Kabel, K.; Le-Mouel, A.; Morel, N.; Page, J.; Stankovic, S.; Stone, A. SUNtool–A new modelling paradigm for simulating and optimising urban sustainability. *Solar Energy* **2007**, *81*, 1196–1211. CISBAT 2005, doi:https://doi.org/10.1016/j.solener.2007.06.002.

14. Robinson, D.; Stone, A. Holistic radiation modelling with a fast simplified radiosity algorithm. Proc. Ninth Int. IBPSA Conf., Building Simulation. Citeseer, 2005.

15. Scartezzini, J.L.; Nouvel, R.; BRASSEL, K.H.; BRUSE, M.; Duminil, E.; Coors, V.; Eicker, U.; Robinson, D. SimStadt, a new workflow-driven urban energy simulation platform for CityGML city models **2015**.

16. Monien, D.; Strzalka, A.; Koukofikis, A.; Coors, V.; Eicker, U. Comparison of building modelling assumptions and methods for urban scale heat demand forecasting. *Future Cities and Environment* **2017**. doi:10.1186/s40984-017-0025-7.

17. Agugiaro, G.; Joachim, B.; Cipriano, P.; Nouvel, R. The Energy Application Domain Extension for CityGML: enhancing interoperability for urban energy simulations. *Open Geospatial Data Software and Standards* **2018**, *3*. doi:10.1186/s40965-018-0042-y.

18. Weiler, V.; Harter, H.; Eicker, U. Life cycle assessment of buildings and city quarters comparing demolition and reconstruction with refurbishment. *Energy and Buildings* **2017**, *134*, 319–328. doi:https://doi.org/10.1016/j.enbuild.2016.11.00

19. Ferrando, M.; Causone, F.; Hong, T.; Chen, Y. Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches. *Sustainable Cities and Society* **2020**, *62*, 102408. doi:https://doi.org/10.1016/j.scs.2020.102408.

20. El Kontar, R.; Polly, B.; Charan, T.; Fleming, K.; Moore, N.; Long, N.; Goldwasser, D. URBANopt: An Open-Source Software Development Kit for Community and Urban District Energy Modeling. 2020.

21. Kavgic, M.; Mavrogianni, A.; Mumovic, D.; Summerfield, A.; Stevanovic, Z.; Djurovic-Petrovic, M. A review of bottom-up building stock models for energy consumption in the residential sector. *Building and Environment* **2010**, *45*, 1683–1697. doi:10.1016/j.buildenv.2010.01.021.

22. Shaikh, P.H.; Nor, N.B.M.; Nallagownden, P.; Elamvazuthi, I.; Ibrahim, T. A review on optimized control systems for building energy and comfort management of smart sustainable buildings. *Renewable and Sustainable Energy Reviews* **2014**, *34*, 409–429. doi:10.1016/j.rser.2014.03.027.

23. Monsalvete Alvarez de Uribarri, P. A multiscale framework for predicting distributed renewable thermal energy integration, 2020. Publisher: University of Nottingham.

24. Huber, J.; Nytsch-Geusen, C. Development of modeling and simulation strategies for large-scale urban districts. Proceedings of Building Simulation, 2011, pp. 1753–1760.

25. Bollinger, L.A.; Evins, R. Facilitating model reuse and integration in an urban energy simulation platform. *Procedia Computer Science* **2015**, *51*, 2127–2136. International Conference On Computational Science, ICCS 2015, doi:https://doi.org/10.1016/j.procs.2015.05.484.

26. Nouvel, R.; Mastrucci, A.; Leopold, U.; Baume, O.; Coors, V.; Eicker, U. Combining GIS-based statistical and engineering urban heat consumption models: Towards a new framework for multi-scale policy support. *Energy and Buildings* **2015**, *107*, 204 – 212. doi:http://dx.doi.org/10.1016/j.enbuild.2015.08.021.

27. Xu, Z.; Coors, V. Application of system dynamics, GIS and 3D visualization in a study of residential sustainability. International Conference on Computational Science and Its Applications. Springer, 2011, pp. 300–314.

28. Dabirian, S.; Panchabikesan, K.; Eicker, U. Occupant-centric urban building energy modeling: Approaches, inputs, and data sources - A review. *Energy and Buildings* **2022**, *257*, 111809. doi:https://doi.org/10.1016/j.enbuild.2021.111809.

29. Palmer Real, J.; Møller, J.K.; Li, R.; Madsen, H. A data-driven framework for characterising building archetypes: A mixed effects modelling approach. *Energy* **2022**, *254*, 124278. doi:10.1016/j.energy.2022.124278.

30. Canada, N.R.C. National Energy Code of Canada for Buildings 2020. https://nrc.canada.ca/en/certifications-evaluations-standards/codes-canada/codes-canada-publications/national-energy-code-canada-buildings-2020, 2022.

31. Guidelines and quality standards for building energy modeling.

32. Spitler, J.D. *Load Calculation Applications Manual*; American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2010. Google-Books-ID: hiQ7SQAACAAJ.

33. Schumacher, J. INSEL Documentation. http://insel.eu/index.php?id=301.

34. Bleyl, J.; Bareit, M.; Casas, M.; Chatterjee, S.; Coolen, J.; Hulshoff, A.; Lohse, R.; Mitchell, S.; Robertson, M.; Ürge Vorsatz, D. Office building deep energy retrofit: life cycle cost benefit analyses using cash flow analysis and multiple benefits on project level. *Energy Efficiency* **2019**, *12*, 1–19. doi:10.1007/s12053-018-9707-8.

35. ISO. ISO 15686-5:2017. https://www.iso.org/standard/61148.html, 2017.

36. Charette, R.P.; Marshall, H.E. UNIFORMAT II Elemental Classification for Building Specifications, Cost Estimating, and Cost Analysis. Technical Report NIST IR 6389, National Institute of Standards and Technology, Gaithersburg, MD, 1999. doi:10.6028/NIST.IR.6389.

37. NRCAN. OpenStudio-Standards. National Renewable Energy Laboratory, 2023.

38. Vázquez-López, E.; Garzia, F.; Pernetti, R.; Solís-Guzmán, J.; Marrero, M. Assessment Model of End-of-Life Costs and Waste Quantification in Selective Demolitions: Case Studies of Nearly Zero-Energy Buildings. *Sustainability* **2020**, *12*. doi:10.3390/su12156255.

39. TheConnoisseur. Understanding Pickling in Python: A Versatile Data Serialization Technique, 2023.

40. Unités d'évaluation foncière - Site web des données ouvertes de la Ville de Montréal.

41. LiDAR aérien 2015 - Site web des données ouvertes de la Ville de Montréal.

42. Canada, N.R. Automatically Extracted Buildings - Prepackaged Shapefiles (download directory) - Open Government Portal.

43. Laval, U. Geoindex - Public.

44. Charbonneau, M. Simulating the Effects of Enclosure Retrofits on Post-War High-Rise Apartment Buildings in Cold Climates. PhD thesis, University of Waterloo, 2011.

45. Jadaa, D.A.; McArthur, J.J. Development of Granular Archetypes for Canadian Existing Commercial Office and Multi-Unit Residential Buildings in Climate Zone 5 **2024**.

46. Canada, G.B.C. Decarbonizing Canada's Large Buildings. https://www.cagbc.org/news-resources/research-and-reports/decarbonizing-canadas-large-buildings/.

47. Ng, L.C.; Ojeda Quiles, N.; Dols, W.S.; Emmerich, S.J. Weather correlations to calculate infiltration rates for U. S. commercial building energy models. *Building and Environment* **2018**, *127*, 47–57. doi:10.1016/j.buildenv.2017.10.029.

48. Study of Part 3 Building Airtightness - Resources.

49. TBD.

50. Dabirian, S.; Saad, M.M.; Hussain, S.; Peyman, S.; Rahimi, N.; Monsalvete Alvarez U, P.; Yefi, P.; Eicker, U. Structuring heterogeneous urban data: A framework to develop the data model for energy simulation of cities. *Energy and Buildings* **2023**, *296*, 113376. doi:10.1016/j.enbuild.2023.113376.

51. Canada, N.R.C. National Energy Code of Canada for Buildings 2020. https://nrc.canada.ca/en/certifications-evaluations-standards/codes-canada/codes-canada-publications/national-energy-code-canada-buildings-2020, 2022.

52. AECOM., Ed. *Spon's Architects' and Builders' Price Book 2018*; CRC Press: London, 2017. doi:10.1201/b22288.

53. Chowdhury, S. Rapid Building Feature Extraction and Geometry Formulation Using Machine Learning. Eleventh National Conference of IBPSA-USA Denver, Colorado, May 21 – 23, 2024, 2024.

54. Canada, N.R. EnerGuide energy efficiency home evaluations, 2018. Last Modified: 2023-11-29 Publisher: Natural Resources Canada.

55. Hecht, R.; Danke, T.; Herold, H.; Hudson, P.; Munke, M.; Rieche, T. Colouring Cities: A Citizen Science Platform for Knowledge Production on the Building Stock - Potentials for Urban and Architectural History. Research and Education in Urban History in the Age of Digital Libraries; Münster, S.; Pattee, A.; Kröber, C.; Niebling, F., Eds.; Springer Nature Switzerland: Cham, 2023; pp. 145–164. doi:10.1007/978-3-031-38871-2_9.

56. Montréal.

57. Echeverria, A.; Palacios, J.; Davila, C.; Zheng, S. Quantifying the financial value of building decarbonization technology under uncertainty: Integrating energy modeling and investment analysis. *Energy and Buildings* **2023**, *297*, 113260. doi:10.1016/j.enbuild.2023.113260.

58. Zhivov, A.; Lohse, R. *Deep Energy Retrofit—A Guide for Decision Makers*; 2021. doi:10.1007/978-3-030-66211-0.

59. Kennedy, M.; Frappé-Sénéclauze, T.P. Canada's renovation wave.A plan for jobs and climate. Technical report, The Pembina Institute, 2021.

60. Admin. What are the six types of relationships in UML class diagrams?, 2022.

61. Yefi, P.; Menon, R.; Eicker, U.; Guéhéneuc, Y.G. MetamEnTh: A Object-Oriented Metamodel for IoT Systems in Buildings. *IEEE Internet of Things Journal* **2024**, *PP*, 1–1. doi:10.1109/JIOT.2024.3373330.

62. project, B. biggproject/Ontology, 2024. original-date: 2022-09-07T08:50:03Z.

63. NREL/openstudio-standards, 2024. original-date: 2014-01-09T23:40:42Z.