A Multi-stage Approach to Understand GIS Model Enrichment Used for Decision-Making Support When Developing Energy Retrofit Strategies on a Neighborhood Level

Christian-Dominik Thiele, **Puyan A. Zadeh**, **Najme Hashempour**, **Sheryl Staub-French**, and **Uwe Rüppel**

Abstract In the AEC sector, energy performance targets of buildings continuously increase for contributing to reduce carbon dioxide. This is usually done on building level, but the focus continuously shifts to larger scales such as neighborhoods, e.g. for identifying buildings with the most retrofitting potential. For this, low detailed GIS models can serve as a basis for energy simulations and are broadly available. However, neighborhood energy simulations hold many challenges, such as the lack of accurate and sufficient data to perform reliable simulations. Information such as window positions or thermal parameters of the building elements can thereby help to increase the quality of the energy simulation results. Therefore, in this paper, challenges of data collection are presented and discussed. To enable users to find a trade-off between accuracy and reliability of a neighborhood simulation and the effort to provide this data, the authors developed the concept of the Neighborhood Model States (NMS). Furthermore, occurring challenges in enriching the GIS model for each NMS are discussed on the example of buildings from the UBC campus.

Keywords Energy Simulation · Neighborhood Level · Neighborhood Model States · Level of Detail (LOD) · Level of Granularity · Geometric information · Non-geometric information · BIM · GIS · Retrofitting

1 Introduction

Climate change is certainly one of the most pressing problems in this day and age. As current reports show, there is a huge potential to save energy in the AEC sec-

C.-D. Thiele (✉) · U. Rüppel
Institute of Numerical Methods and Informatics in Civil Engineering, Technical University of Darmstadt, Darmstadt, Germany
e-mail: thiele@iib.tu-darmstadt.de

P. A. Zadeh · N. Hashempour · S. Staub-French
Project Management and Construction Group, Department of Civil Engineering, University of British Columbia, Vancouver, Canada

tor [1]. This can be reached e.g. through developing suitable retrofit strategies by replacing fossil fuel-based heating systems with electric heat pumps and improving building insulation. As a result, many governmental agencies and control bodies in different countries have developed strategies and encourage energy retrofits through implementing certain laws and offering financial incentives to building owners. To reach Global Climate targets as set in the Paris Agreement [2, 3], it is important to focus on larger scales and conduct energy analysis at neighborhood levels instead of individual buildings. For this aim, it is necessary to develop and use tools that can estimate the energy consumption of neighborhoods [4]. As the authors in [5, 6] highlight, the estimation of energy consumption at the neighborhood level enables the opportunity to combat climate change by creating livable and energy efficient neighborhoods, as well as to support energy efficiency, sustainability, and management of cities. More particularly, automated energy simulations on a neighborhood level can be used for evaluating the impacts of potential retrofitting measures, or for identifying specific critical buildings with a priority to be energy retrofitted [5, 7]. According to [7], changing the retrofit analysis perspective from individual buildings to neighborhoods has a great potential for bigger savings and lower investments. In this way, the decision makers can quickly identify the most critical buildings that need retrofitting and can then perform more precise and intensive analysis on those specific buildings. This can be especially beneficial to large owners of buildings with close proximity to each other, such as university and hospital campuses, etc. However, most currently available energy simulation tools are designed to analyze only individual buildings, which require a large amount of data with high level of precision. This is why it is a fundamental research challenge to change the perspective from individual buildings to neighborhood levels [4].

To perform energy simulations on any level, several parameters must be known. These include, on the one hand, geometric information such as the height and footprint of the buildings. On the other hand, non-geometric information like the thermal parameters of the exterior wall layers and the type of the installed HVAC system. However, on a large scale, such as neighborhoods, there are significant challenges that surface when it comes to availability of the data, data inconsistencies, and data privacy issues [8]. Although these challenges can also apply to individual buildings, they become even more challenging on a larger scale. While data for individual buildings can be collected on site from Building Information Models (BIMs), building drawings, and other project documents, this level of data collection is no longer feasible for larger scales, such as neighborhoods, due to the availability and consistency of the required data. This is why, when analyzing a collection of buildings, such as in a neighborhood, practitioners are often limited to only relying on the data from Geographic Information Systems (GIS) with low levels of detail. Since GIS data that is used for performing neighborhood energy simulations is often inhomogeneous regarding the provided amount of information as well as their reliability [9], practitioners need to find ways to enrich the GIS data to be able to conduct useful and reliable simulations that can be used for developing retrofit strategies. However, enriching GIS data is complicated, cumbersome and expensive [10].

Therefore, the main research objective is to address this gap between the required data for energy simulations and the available data. This is done by investigating and understanding the challenges and resulting efforts when gathering additional geometric and non-geometric information (Sect. 4).

For understanding the GIS model enrichment, an innovative Neighborhood Model State (NMS) concept is developed. This concept consists of four model states with their respective geometric and non-geometric information components that can be added to the original GIS data. The information content for each NMS increases progressively, and each state is respectively analyzed for the challenges in data acquisition and their potential impacts on the quality of the energy simulation outcomes. To demonstrate the concept of NMS, two representative use case buildings, the Engineering Student Center (ESC) and the Centre for Interactive Research on Sustainability (CIRS) at the University of British Columbia (UBC) Vancouver Campus in Canada are used in this work. A GIS model for the entire UBC Vancouver campus is available, where the selected buildings could be extracted for further analysis.

In the following Section, an overview about the role of the data in neighborhood energy simulations, as well as energy simulations on large scales in general, is given. Then, in Sect. 3, the research methodology is presented. In Sect. 4 the concept of the NMS are introduced and examples for challenges in data acquisition for different neighborhood model states are presented and discussed. Finally, in Sect. 5 the results of the research are concluded.

2 Background

As described, for existing buildings, usually only limited digital data is available, and what is available, is mostly unstructured. Therefore, extensive and time-consuming on-site explorations and measurements can be the consequence to acquire needed information like wall layers, materials, geometric dimensions etc. On a small scale, the collection of these geometric and non-geometric information is manageable, but for multiple buildings e.g. on a neighborhood level this procedure is too time-consuming and in consequence not applicable. To give an insight about how energy advisors and engineers gather their data for energy simulations on neighborhood levels, a detailed literature review was conducted.

2.1 The Role of Data for Energy Simulations

As is well known, BIM's on the scale of individual buildings can serve as the basis for different stakeholders during the planning and operational phase. It thereby can be a collection of different specialist models and can include necessary information for energy simulations. Energy related specialist models are commonly referred to as BEM (Building Energy Modeling). By increasing the scope to a neighborhood level, a comparable method exists, namely UBEM (Urban Energy Modeling) or USEM

Table 1 Needed data to perform energy simulations (based on [9])

	Minimum information	Additional information
Geometric	Footprint	Glazing (WWR)
	Envelope composition (layers)	Basement
		Roof type and shape
		Interaction with adjacent
		buildings (shading)
Non-geometric	Weather data	Year of construction
	Heating system	Thermal zones/stories
		HVAC information
		Number of Occupants
		Occupant behavior

(Urban-Scale Energy Modeling), for which researchers see a strong potential [8, 11] but also are challenging due to the complexity of urban energy systems [12]. Energy simulation on large scales can basically be classified into two approaches: top-down and bottom-up. The top-down approach is usually data-driven and can be based on statistical energy use and historical data, among others [8]. On the other hand, there is the bottom-up approach, which is physics-based and engineering models and simulations are used. Therefore, usually more individual data is necessary which can have significant uncertainties in building energy estimates at an urban scale, as the authors in [8] are stating. Additionally, when buildings are modeled individually, it can require a higher computational power regarding the provided information [13].

In this context, the bottom-up physics-based approach is chosen, since it is very suitable for in-depth urban scale analyses [8]. This results in a need for extensive data that has to be gathered as it builds the foundation for the simulation. Therefore, an overview about the needed data is given in Table 1. These parameters are in the literature often categorized as geometric and non-geometric information, which is taken up in this publication. Furthermore, the authors distinguish the data in between minimum mandatory data that is needed to get at least a result from the simulation, and additional data that can lead to more reliable results if provided.

To grade the complexity and amount of included geometric and non-geometric information of the respective data, the concept of Level of Detail (LOD) is common and well known on the building level from the BIM methodology. For GIS, this concept has been adopted, although different GIS have different definitions.

One of the most well-known GIS is CityGML on which the concept of LOD will be elaborated. CityGML provides geometrical data of neighborhoods or even municipalities by using an XML-based format and is defined in the OGC CityGML Encoding Standard [14]. This geospatial data is largely available, e.g. for most European countries [15]. CityGML is currently being revised, the upcoming version of CityGML will be version 3. In the previous version, a LOD was describing the whole building (e.g. LOD4 means very detailed facade including furniture) but is now getting harmonized in version 3 with the definition of BIM. There, the LOD does not obligatorily mean the whole building, but can describe specific components individually. For

example, it is now possible to have a very low LOD of the outer shell combined with a highly detailed inner interior model. So, all the building parts can have their own LOD, as known from BIM [16]. Additionally, the LOD of CityGML models varies significantly depending on regions [17]. Furthermore, Biljecki et al. [18] showed, that the CityGML 2.0 LOD concept as currently defined in [14] is inconclusive, since each LOD can be interpreted in multiple ways [18]. However, in the upcoming CityGML version 3 the usability and inconsistencies are improved [19].

When it comes to geometric representations, most of the existing tools are using GIS models (namely CityGML). Even though low detailed CityGML models offer only a specific amount of information as well as a limited accuracy regarding geometric representations, this data source is commonly used as a basis for energy simulations on neighborhood levels. Next to CityGML, more possibilities to exchange geospatial data are GeoJSON and Shapefile. Unfortunately, these data types do not provide schemas to further define building properties [20]. It is also worth mentioning KML (Keyhole Markup Language) as an alternative file format and data source for retrieving geometric and geospatial data. KML is used, for example, by Google to display geographic data in Google Earth and is based on the XML standard [21]. While the acquisition of geometrical data on neighborhood levels is well documented, existing reviews with profound discussions on non-geometric acquisition for UBEMs are lacking [10].

The quality of the simulation can be continuously improved by adding more known details as expected and by increasing its LOD. The influence of the different input parameters have been dissected through several sensitivity analyses [4, 9, 17, 18, 22]. As an example, the authors in [4] showed the influence of different parameters on the quality of energy simulations by performing a sensitivity analysis through varying individual geometrical and physical factors. Therefore, they used six case study buildings of different building types. They found out that for the geometry, the deviation of between LOD1 and LOD2 models is smaller or equal to 10%, the comparison between LOD2 and LOD3 is again smaller than 12%. Furthermore, they varied parameters such as the windows-to-wall ratio, U-values of the walls among others. They concluded that depending on the data availability and assumptions being made, errors up to 80% could occur. For the user behavior parameters, errors up to 40% have been encountered. These numbers underline the necessity of having reliable underlying data for energy simulations.

The authors from another study that is concerned with a sensitivity analysis is [9]. Here, the authors used a comprehensive neighborhood data set of approx. 8,600 buildings from a city in Germany. They limited their analysis to the comparison of LOD1 and LOD2 representations, even though they looked at a considerably large dataset. The authors looked into the effect of varying parameters like the error of using LOD2 instead of LOD1 models, the role of basements, attics, window-to-wall ratio, air change behavior, internal gains, among others. Eventually, they come to a similar conclusion as the authors from [4]. They noted that the available LOD affects other parameters with sizable roles regarding the result of the energy simulation. Due to a lack of LOD3 data, they couldn't assume their influence on the energy simulation result. Finally, they provided a ranking about must-have parameters (besides a LOD1

city model), which are absolutely essential for doing neighborhood simulations, such as the building year of construction, the building function, refurbishment information, and residence type. Further information has a less impactful consequence when missing than these (error over 30%).

However, since this assessment in both case studies have been done by the simulation platform SimStadt, which is inferring information from data sets and benchmarking data libraries [9], it may not be transferable to other simulation tools.

2.2   Energy Simulations on Neighborhood Level

The Building Energy Simulation Tools web directory (BEST-D) lists over 170 different tools to perform energy simulations. More than 50 can be used for a whole-building energy simulation [23]. However, most building energy evaluation tools are for the analyses of individual buildings and require a high amount of data [4]. There are some tools that originated mostly from research projects, to overcome this problem. Popular ones, SimStadt, CitySim, UrbanSim and CityBES will be presented in the following.

The urban simulation tool SimStadt and SimStadt 2.0 respectively was developed in research projects finished in 2015 and 2020. The platform analyzes districts or even regions regarding their heating requirements, photovoltaic studies, and renewable energy supply scenarios [24]. For that, it uses a bottom-up physic-based approach. Its focus was less on getting as accurate models as possible, but to provide a reliable simulation tool using existing data points for basic decision-making. It is validated through three case studies [25].

CitySim is another tool that tries to support urban energy planners to reduce energy consumption and emission of greenhouse gasses. It thereby provides the possibility to enrich geometrical buildings with thermophysical properties. The calculation is based on statistical values for occupants' presence and behavior and offers typical HVAC systems. As a simulation engine, CitySim Solver was developed and comes with its own proprietary XML file format. The engine was validated in field studies [26].

City Buildings, Energy, and Sustainability (CityBES) is a web-based data and computing platform sponsored by Lawrence Berkeley National Lab. It uses CityGML as an open standard and is based on EnergyPlus as a simulation engine. The tool allows adding additional data like weather data, information about the building stock through CityGML and GeoJSON and further standards and codes. It supports scenarios like energy benchmarking and energy retrofit analysis [27].

As the literature review shows, many research projects are focusing on the development of tools that can support city planners when it comes to energy related questions. However, it remains an open problem that for bottom-up approaches a significant amount of data is required. The research presented in this paper seeks to address this gap by identifying openly available data and enriching base models used for analysis, such as CityGML files.

3 Methodology

The main objective of this research is to investigate the relevance of the granularity in geometric and non-geometric building information for energy simulations on the neighborhood level, and to highlight the challenges in performing such simulations. For this aim, the research team conducted a thorough review of the related literature to understand the necessity of different impact factors when performing energy simulations for neighborhoods (Sect. 2).

For understanding the information required to perform energy simulations, several energy simulation tools were reviewed, and the respective data gathering effort were identified. Based on this, a better understanding of a needed trade-off between the effort to gather the data and the accuracy of the energy simulation was gained. The results of this analysis are presented in Sect. 4.1.

For the development of the neighborhood model state (NMS) concept, a CityGML model of the UBC campus was used as an example. In addition, models of the City of Vancouver were examined. These models were then decomposed into individual buildings and evaluated for their suitability for energy simulations and their LOD. BIMs of selected campus buildings were also used and compared to the CityGML models for accuracy. This ultimately allowed the identification of challenges and difficulties in neighborhood-level data collection. Based on the literature review and the conducted research, difficulties in data acquisition were crystallized and the concept of the neighborhood model states was developed which is presented in Sect. 4.2.

The developed NMS concept was then applied to two representative example buildings of the UBC campus. The specific challenges resulting from those examples are discussed and presented in Sect. 4.3.

4 Neighborhood Model States (NMS)

4.1 Data Acquisition and Enrichment for Energy Simulation on Neighborhood Level

As the authors in [7] state, energy simulations for neighborhoods can support decision makers to better prioritize effective retrofitting measures. Furthermore, the authors in [17] highlight that decision-makers need access to reliable quantification of the energy demand of all (or most of) buildings within a neighborhood to be able to have a correct assessment of the feasibility of district energy systems.

When it comes to the shifting from single buildings to a higher scale, such as a neighborhood, there is a trade-off between the accuracy of the simulation outcomes and the ease of data acquisition and simplicity of compiling energy simulations. Although reaching the highest level of accuracy is desirable, the main purpose of such analyses is often identifying the low performing buildings in a neighborhood
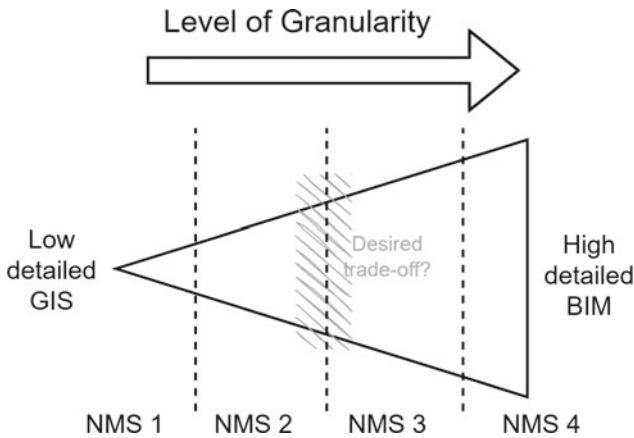
Fig. 1 Discrepancy between NMSs regarding the granularity of geometric and non-geometric information

and choosing suitable investment strategies to address their performance deficiencies. In other words, a highly detailed energy simulation wouldn't be necessarily required for this stage and such detailed analysis can be made once the specific low performing buildings are selected for the retrofitting process. At that time, comprehensive data collection from those specific buildings and a detailed energy analysis can be conducted.

Therefore, while the granularity of building data can vastly vary from NMS1 (rough GIS models) to NMS4 (detailed BIMs), when performing energy simulations for neighborhoods, the desired level of granularity in the data can settle somewhere between a rough geospatial representation and a highly detailed individual 3D visualization as shown in Fig. 1.

Since the highly granular models such as BIMs are currently not expected to be available for most neighborhoods, the question emerges as to what extent low LOD GIS models need to be enriched to achieve sufficient energy simulation outcomes for neighborhoods. On this basis, low performing buildings and their performance discrepancy to the other neighborhood buildings can be identified while considering the accessibility to the needed information.

To answer this question, it is necessary to investigate whether changes, i.e. increasing the accuracy in geometric data would lead to noticeable better results. The same investigation must be conducted for the non-geometric data as well. These investigations are not mutually exclusive, but the amount of efforts for theextra data acquisition

and enrichment for each data point must be considered and evaluated to better understand the trade-off between the reliability of the simulation outcomes and the ease of data acquisition. For this aim, the new classification of the neighborhood model states can be used to provide a better understanding of such trade-offs.

4.2 Neighborhood Model States for Energy Simulations

It is often the case that high granularity building data, including BIMs, are not available for the entire buildings of a neighborhood. On the other hand, there are numerous databases available that provide GIS models for many cities and their neighborhoods [17], which can be used for different neighborhood-based analyses. However, as for the energy simulations, the available GIS models are mostly limited to simple geometric representations, and lack in non-geometric data, which potentially can lead to imprecise energy simulation results for neighborhoods [9]. This is especially critical when using energy simulators designed for working with detailed building models for the neighborhood analysis purposes. Therefore, GIS models need to be adjusted and enriched by adding more detailed information to be more suitable for energy simulation purposes.

As discussed in Sect. 2, there are different interpretations of LOD when dealing with GIS models and BIMs. Furthermore, the LOD for GIS models can vary significantly depending on regions, and also each GIS LOD can be interpreted in multiple ways [18]. To avoid misinterpretations and inconsistencies, this research proposes a new classification for the different neighborhood model states based on the required geometric and non-geometric data for each state, in accordance with the difficulty of acquiring this data.

This new classification demonstrates a gradual increase in the granularity of the geometric and non-geometric data, which ultimately can be used for high resolution energy modelings. In defining these model states, we paid extra attention to the readiness and accessibility of each data point. In this classification, the NMS1 represents the lowest granularity of the data with the highest accessibility potential, while the NMS4 represents the highest granularity of the data and is comparable to the building data available in high detailed BIMs. The higher the level of NMS is, the more challenging the data acquisition gets so that NMS4 data can often be obtained only by having detailed building plans or conducting extensive in-situ explorations using LiDAR and other technologies. The details of the new proposed classification for neighborhood model states are shown in Table 2, where there is a distinction made between geometric and non-geometric data for each neighborhood model state.

Table 2 Classification of the neighborhood model states (NMS) for energy simulations on neighborhood level.

NMS1	NMS2	NMS3	NMS4
<u>Geometric</u>	<u>Geometric</u>	<u>Geometric</u>	<u>Geometric</u>
1) Rough geometry ($\leqslant$ $LOD1$[a])	1) Rough geometry ($\leqslant LOD1$)	1) Rough geometry ($\leqslant$ $LOD1$)[c]	1) ~~Rough geometry ($\leqslant LOD1$)~~
2) GIS position and orientation	2) GIS position and orientation	2) GIS position and orientation	2) GIS position and orientation
3) Rough zone determination	3) Rough zone determination	3) ~~Rough zone determination~~	3) ~~Rough zone determination~~
	4) Partially fenes-tration ratio[b]	4) ~~Partially fenestration ratio~~[b]	4) ~~Partially fenestration ratio~~[b]
	5) Roof shape	5) Roof shape	5) ~~Roof shape~~
		6) Full fenes tration ratio[a]	6) ~~Full fenestration ratio~~[a]
			7) Basement
			8) Exact window locations
			9) Inner layout & exact zones
			10) More accurate geometry ($\geqslant LOD3$)
<u>Non-geometric</u>	<u>Non-geometric</u>	<u>Non-geometric</u>	<u>Non-geometric</u>
a) Weather data	a) Weather data	a) Weather data	a) Weather data
	b) Building usage type	b) Building usage type	b) Building usage type
	c) Year of construction	c) Year of construction	c) Year of construction
	d) Occupancy numbers	d) Occupancy numbers	d) Occupancy numbers
		e) Heating system	e) Heating system
		f) Wall layers	f) Wall layers
			g) Occupancy behavior
			h) Historical operational data (bills)
			i) Exact data of HVAC system

[a] In accordance with the CityGML 3.0 definition.
b2 WWR: Window-to-Wall Ratio.
c The dimensions of the geometry should be validated.
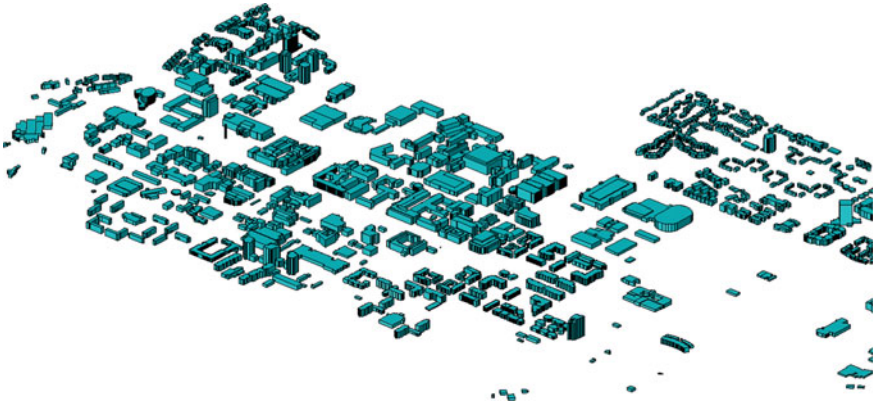Crossed: obsolete, replaced by a more accurate feature.

Fig. 2 CityGML Model (LOD1, provided by UBC) of UBC Campus, Vancouver

4.3 Challenges of Data Acquisition on the Example of UBC Buildings

To discuss the data acquisition and its challenges for each NMS, an exemplary neighborhood as a use case is selected. The chosen neighborhood is the campus of the University of British Columbia (UBC) in Vancouver, Canada. For this neighborhood, a low detailed CityGML GIS model is available (Fig. 2). Considering UBC as the exemplary neighborhood also has the convenience of having access to operational data when it comes to the validation of the simulations.

As example buildings for the UBC neighborhood, the Engineering Student Center (ESC) as well as the Centre for Interactive Research on Sustainability (CIRS) were selected. The architecture of the ESC is simple and squares well regarding the complexity. It was opened in 2015 as a break and study room for undergrad engineering students. The CIRS building however represents a complex geometry with a versatile facade, which was created as a case study building for an energy efficient building. The CIRS building was opened in 2011.

The selected buildings will be used as examples in the following subsection, where the process of collecting geometric and non-geometric data and their respective challenges for the data acquisition for each NMS will be discussed.

Neighborhood Model State 1 (NMS1). The geometry data for NMS1 is very basic and can be obtained from various sources, e.g. from municipalities, databases or operators of large facilities (such as universities, hospitals etc.). This can be GIS models such as CityGML models. These data are usually widely-used and easy to obtain, but they are mostly not more detailed than LOD1 (floor area and average height of the building). Therefore, the quality of the geometry may vary. As for the UBC sample buildings, the geometry of the CIRS building is very accurate. In contrast, the geometry of the ESC building in the GIS model is too small by a

factor of 2 in each dimension, resulting in a significantly underestimated volume. The fact, that the building is surrounded by large buildings could be a contributor to this outcome. If the number of stories is determined by the average height of the building, these geometric inaccuracies can have an impact. Assuming an average story height of 3.00 m, the ESC building would have only one story, since it's CityGML model height is 4.35 m. This is not the case, since the ESC building actually consists of 2 stories.

The originally GIS models usually do not come with many non-geometric information and therefore an adjustment should be conducted. If the simulation should be run based on the available information of NMS1, many additional assumptions need to made, such as the layers of the walls, the building type, setpoint temperatures and the heating system.

Neighborhood Model State 2 (NMS2). The NMS2 is based on NMS1, which is basically the rough CityGML geometry. In NMS1, important features like the Window-To-Wall-Ratio (WWR) and information about the roof shape are usually missing. These should be added in NMS2. The fenestration of a building is a major feature for energy simulations, since it can have a significant influence on the U-Value of the building envelope. This is especially because windows usually have a higher heat transmission than the surrounding wall.

Some non-geometric parameters such as building usage type, year of construction, and information about the number of occupants, are further features that can contribute to a more reliable simulation. These parameters are often easy to access and therefore should be considered as model enrichment measures to reach NMS2. While the building type may be found in some NMS1 models, the year of construction and the number of occupants are usually not available. The building type can give hints for the schedule and density of the building occupants. The year of construction can be important, since the building envelope of a building could be estimated based on typical related archetypes. The number of occupants in accordance with the building schedule can have a significant impact on the simulation.

Neighborhood Model State 3 (NMS3). If a rough WWR is known, influences like the orientation of the windows can be additionally considered for an even more accurate simulation. The orientation of the windows can thereby have a significant impact on the heat transmission through the building envelope [28]. However, obtaining the orientation of the windows is even more challenging than the guessing of a rough WWR. Therefore, advanced algorithms might be necessary. Regarding the possible errors in the geometric representation, a check of the underlying model should be performed to ensure the basic geometry is correct.

At this stage, further non-geometric parameters like the heating system and the wall layers should be considered more precisely. Thereby, it needs to be differentiated between basically three common heating systems: district heating, fossil fuel based heating (gas, oil), and electric heat pumps. However, complex neighborhoods like UBC are also likely to have complex heating systems. In the example of the ESC building, the heating systems consist of a heat pump which is supported by district heating. The heating system of the CIRS building is even more complex. Here

multiple electrical heat pumps working together, and the system is even overarching multiple buildings that exchange heat. Identifying the exact layer composition of walls is a comparable challenge. Even for individual buildings, energy advisors need to guess the layers in case it is not even known to the homeowner. Due to the risk of releasing harmful particles like asbestos, energy advisors are urged to not drill or open the walls in any way for investigating the wall layers, which makes it nearly impossible to reliably determine the construction of the walls on a large scale.

Neighborhood Model State 4 (NMS4). The last NMS is the most detailed one. It includes all the features that are necessary for a reliable energy simulation. That level of detail required for this NMS is comparable to the design BIM version of a building with an LOD of 300 and above. However, the data acquisition for this NMS is the most challenging one regarding the availability of the data and on a district level usually not feasible. Here, different additional data parameters can be considered, such as the information about a basement, accurate zone information, the inner layout, the exact heating system and the exact wall layers. This information is usually available in BIMs, however even here can be considerable differences between design, construction and as-build models. Also, non-geometric information like historical operational data and the accurate used heating system should be added to reach this enrichment. This information can be obtained from owners, but is usually not accessible on larger scales. Probably the most challenging part of the NMS4 is the acquisition of the occupancy behavior (such as ventilation habits or the actual desired temperature).

5   Discussion

As the conducted literature review showed, energy simulations on a neighborhood level can contribute to reducing the emission of greenhouse gasses and are a promising instrument to reach the self-imposed objectives in the fight against climate change. However, the literature also mentions when performing energy simulations on a large scale, it comes with problems that impact the reliability of the simulation. Therefore, the authors proposed a multi-stage approach to understand GIS model enrichment on a neighborhood level, the Neighborhood Model States (NMS) concept. This consists of four levels with increasing amounts of information, based on an initial (and usually low detailed) GIS model. The NMS levels are thereby oriented on the feasibility of gathering the additional information to enrich the basic GIS model.

In the course of developing the NMS concept, many problems arose that underlined the challenges of enriching low level GIS models with reliable and accessible data. To demonstrate the concept, two representative buildings of the UBC campus were chosen, the ESC and the CIRS building. Even though both buildings came from the same CityGML model, one showed a significant underestimated volume. This can lead to not reliable energy simulations. However, these models usually serve as basis for neighborhood energy simulations and are used unverified, which is due to a

lack of verification possibilities. This shows, that there is a strong need of algorithms which uses additional databases to further enhance GIS models.

Since data availability is one of the most challenging problems when it comes to simulation on large scales, the authors also want to encourage further deployment of comprehensive and accessible databases. For neighborhoods, cities can serve as an example, where much more relevant data is being made available through open data platforms.

To be able to do all these kinds of data enrichment, much manual work is required, and therefore it is not feasible to do it for multiple buildings on a large scale by hand. Based on this work, the authors are currently working on an automatization approach for enriching the individual buildings in a neighborhood. For this, they are working on AI algorithms to automatically extract window ratios and positions from openly available satellite images as well as on an automated dimension validation method. This will increase the reliability of the simulations significantly. Also, the authors aim to operationalize the NMS concept by scaling it to the extent of the UBC campus. Furthermore, the authors are interested in the investigation of the challenges in the data exchange between GIS models, UBEMs and energy simulation software.

References

1. United Nations Environment Programme (2021) 2021 global status report for buildings and construction: towards a zero-emission, efficient and resilient buildings and construction sector. Technical report, Nairobi
2. UNFCCC: The Paris agreement. https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement
3. United Nations (2015) Paris agreement
4. Strzalka A, Monien D, Koukofikis A, Eicker U (2015) Sensitivity analysis for minimization of input data for urban scale heat demand forecasting. In: Sustainable energy, p 79
5. Mutani G, Todeschi V (2020) Building energy modeling at neighborhood scale. Energy Effi 13:1353–1386
6. Reinhart CF, Davila CC (2016) Urban building energy modeling - a review of a nascent field. Build Environ 97:196–202
7. Stanica D-I, Karasu A, Brandt D, Kriegel M, Brandt S, Steffan C (2021) A methodology to support the decision-making process for energy retrofitting at district scale. Energy Build 238:110842
8. Ali U, Shamsi MH, Hoare C, Mangina E, O'Donnell J (2021) Review of urban building energy modeling (UBEM) approaches, methods and tools using qualitative and quantitative analysis. Energy Build 246:111073
9. Nouvel R, Zirak M, Coors V, Eicker U (2017) The influence of data quality on urban heating demand modeling using 3D city models. Comput Environ Urban Syst 64:68–80
10. Wang C, Ferrando M, Causone F, Jin X, Zhou X, Shi X (2022) Data acquisition for urban building energy modeling: a review. Build Environ 217:109056
11. Hong T, Chen Y, Luo X, Luo N, Lee SH (2020) Ten questions on urban building energy modeling. Build Environ 168:106508
12. Abbasabadi N, Ashayeri M (2019) Urban energy use modeling methods and tools: a review and an outlook. Build Environ 161:106270

13. Ascione F, Bianco N, Mauro GM, Napolitano DF (2021) Knowledge and energy retrofitting of neighborhoods and districts. A comprehensive approach coupling geographical information systems, building simulations and optimization engines. Energy Convers Manag 230:113786
14. Gröger G, Kolbe TH, Nagel C, Häfele K-H (2012) OGC city geography markup language (CityGML) encoding standard
15. About INSPIRE: INSPIRE. https://inspire.ec.europa.eu/about-inspire/563
16. Kolbe TH, Kutzner T, Smyth CS, Nagel C, Roensdorf C, Heazel C (2021) OGC city geography markup language (CityGML) part 1: conceptual model standard
17. De Jaeger I, Reynders G, Ma Y, Saelens D (2018) Impact of building geometry description within district energy simulations. Energy 158:1060–1069
18. Biljecki F, Ledoux H, Stoter J, Vosselman G (2016) The variants of an LOD of a 3D building model and their influence on spatial analyses. ISPRS J Photogramm Remote Sens 116:42–54
19. Kutzner T, Chaturvedi K, Kolbe TH (2020) CityGML 3.0: new functions open up new applications. PFG J Photogramm Remote Sens Geoinf Sci 88:43–61
20. Chen Y, Hong T, Luo X, Hooper B (2019) Development of city buildings dataset for urban building energy modeling. Energy Build 183:252–265
21. KML tutorial: keyhole markup language: Google developers. https://developers.google.com/kml/documentation/kml_tut
22. Malhotra A, Shamovich M, Frisch J, Treeck C (2019) Parametric study of the different level of detail of CityGML and energy-ADE information for energy performance simulations, pp 3429–3436
23. Irmler R (2019) Datensynthesemethoden für die Generierung dynamisch-thermischer Simulationsmodelle im Wohngebäudebestand am Beispiel von EnergyPlus. Ph.D. thesis, Shaker Verlag, Düren
24. SimStadt documentation. https://simstadt.hft-stuttgart.de/
25. Coors V, Weiler V, Klöber A, Brüggemann T, Goll L, Balbach B, Spath F, Gärtner B (2001) EnEff:Stadt – SimStadt 2.0. Schlussbericht, Stuttgart, Dresden, Leimen, Mainz
26. CitySim software. https://www.epfl.ch/labs/leso/transfer/software/citysim/
27. CityBES. https://citybes.lbl.gov/
28. Kim S, Zadeh PA, Staub-French S, Froese T, Cavka BT (2016) Assessment of the impact of window size, position and orientation on building energy load using BIM. Procedia Eng 145:1424–1431