energies

Article

Toward Improved Urban Building Energy Modeling Using a Place-Based Approach

Guglielmina Mutani [1,*], Pamela Vocale [2] and Kavan Javanroodi [3]

1 DENERG-Department of Energy, Politecnico di Torino, Corso Duca Degli Abruzzi 24, 10129 Torino, Italy
2 Department of Engineering and Architecture, University of Parma, Parco Area Delle Scienze 181/A, 43124 Parma, Italy; pamela.vocale@unipr.it
3 Division of Building Physics, Department of Building and Environmental Technology, Faculty of Engineering, LTH Lund University, 22363 Lund, Sweden; kavan.javanroodi@byggtek.lth.se
* Correspondence: guglielmina.mutani@polito.it; Tel.: +39-0110904528

Abstract: Urban building energy models present a valuable tool for promoting energy efficiency in building design and control, as well as for managing urban energy systems. However, the current models often overlook the importance of site-specific characteristics, as well as the spatial attributes and variations within a specific area of a city. This methodological paper moves beyond state-of-the-art urban building energy modeling and urban-scale energy models by incorporating an improved place-based approach to address this research gap. This approach allows for a more in-depth understanding of the interactions behind spatial patterns and an increase in the number and quality of energy-related variables. The paper outlines a detailed description of the steps required to create urban energy models and presents sample application results for each model. The pre-modeling phase is highlighted as a critical step in which the geo-database used to create the models is collected, corrected, and integrated. We also discuss the use of spatial auto-correlation within the geo-database, which introduces new spatial-temporal relationships that describe the territorial clusters of complex urban environment systems. This study identifies and redefines three primary types of urban energy modeling, including process-driven, data-driven, and hybrid models, in the context of place-based approaches. The challenges associated with each type are highlighted, with emphasis on data requirements and availability concerns. The study concludes that a place-based approach is crucial to achieving energy self-sufficiency in districts or cities in urban-scale building energy-modeling studies.

Keywords: Urban Building Energy Modeling UBEM; Urban Scale Energy Models USEM; place-based approach; geo-database; geographic information system; QGIS; sustainable cities and communities

Citation: Mutani, G.; Vocale, P.; Javanroodi, K. Toward Improved Urban Building Energy Modeling Using a Place-Based Approach. *Energies* **2023**, *16*, 3944. https://doi.org/10.3390/en16093944

Academic Editor: Fabio Bisegna

Received: 5 April 2023
Revised: 29 April 2023
Accepted: 5 May 2023
Published: 7 May 2023

1. Introduction

In European urban contexts, where 74.5% of the EU-27 population live [1], buildings account for 36% of greenhouse gas (GHG) emissions [2] and the residential sector for 28% of final energy use [3]. The current issues of climate change, pandemics, wars, and energy price crises highlight the centrality of the energy sector at a territorial level. One of the crucial future challenges that European policies face is achieving energy self-sufficiency in high-density urban contexts, meeting lower energy demand by producing energy by boosting the available renewable energy sources (RESs) [4]. High-energy intensity, combined with the numerous constraints, limited availability, and the associated uncertainties of RESs, shifts the spatial scale of the analysis from a local scale (i.e., building or block of building scale) to an urban or regional scale [5]. In order to ensure a clean energy transition and carbon neutrality by 2050 [6], integrated models and methods that incorporate fine spatial and temporal resolution data from different sources at an urban scale are crucial [7]. Urban building energy modeling (UBEM) is one such method that can be used to simulate and analyze the energy performance of buildings at the neighborhood scale. However, the potential of UBEM to facilitate sustainable development in terms of built environments is limited by its dependence on conventional physics-based inputs [8]. This issue becomes even more critical for district- or city-scale studies, in which data availability is also a concern. To this end, this study investigates a set of methodologies and models to improve the effectiveness of UBEM to enhance the sustainability of built environments using a place-based approach. In particular, this study presents a methodological paper that describes a new energy modeling technique with a place-based approach to provide an effective means of simulating and analyzing energy use in urban built environments, considering fine spatiotemporal resolution data. The energy modeling presented also explains the importance of managing large databases, identifying other variables or surrogate variables, and representing the results. The pre-modeling part is of fundamental importance because it influences the choice and accuracy of the model, and the representation of the data allows for the dissemination of the results.

This paper is organized as follows. The 'Literature review' section reports on the current state of the art and is followed by the 'Knowledge gap and objectives' of this study; 'Place-based Urban Building Energy Modeling' describes the proposed energy modeling; the last section, 'Conclusion and Remarks', presents the main strengths and opportunities of USBEMs using a place-based approach via typology, with a description of future developments.

2. Literature Review

There are several examples of building energy modeling (BEM) discussed in the literature [9]. These models are implemented at the individual building level, and their outcomes are extrapolated to the urban level [10], mostly based on predefined archetypes [11]. In order to consider the geometrical interplay between buildings and their surroundings, urban building energy modeling (UBEM) and urban-scale energy model (USEM) gradually took hold to access large-scale static and dynamic simulations of different building types and urban morphologies [12,13].

By extending the UBEM and USEM to the district or territorial scale, a new place-based energy model can be introduced, namely urban-scale building energy modeling (USBEM) [14]. Similar to UBEM, USBEM is based on two main approaches: top-down models, in which aggregated historical energy data are provided (e.g., municipal data), and bottom-up models, in which energy consumption is provided for many buildings and for several years (e.g., building data) [15]. Several scientific reviews (e.g., [16,17]) agree upon classifying energy scale modeling into three main groups: data-driven (black box), which relies on statistical models [18] and AI models (e.g., [19,20]); process-driven (white-box), which is based on process-driven models [21]; hybrid (grey-box) models, which use a combination of data-driven and process-driven models [22,23] and are mainly used by environmental and urban planners, public entities, and policymakers.

Literature reviews provide an overview of the existing tools and platforms, considering the type of energy model [24–26]:

- Data-driven tools: rely on statistical models [18] for energy analysis (e.g., EnergyProforma and CRECM) and AI models (e.g., [19,20]) for energy analysis and benchmarking (e.g., DUE-S and DUE-B);
- Process-driven tools: based on process-driven models [21] with EnergyPlus as a simulation engine (e.g., CityBES, UMI, and UrbanOPT) or using other dynamic simulation engines (e.g., CitySim and SEMANCO);
- Hybrid tools use a combination of data-driven and process-driven models [22,23] and are mainly used by environmental and urban planners and policymakers (e.g., SimStadt, TEASER, and CEA);
- Integrated platforms optimize energy consumption models with RES simulation tools [27] (e.g., SynCity, Epic-hub, EnerGIS, and LEAP).

Currently, USBEM is a promising field of research, but no integrated modeling platform that can encompass various urban environments, climates, types of populations, and data availability can be found in the literature. In addition to the USBEM tools and methods mentioned above, it is important to note that conducting geospatial analysis poses a significant challenge in urban building energy modeling, as highlighted in [17]. The use of a place-based approach enables the description of spatial interrelationships among various urban components and variables across different layers [28]. This approach also allows for the analysis of spatial-autocorrelations and the dynamic behavior of physical phenomena that impact energy consumption.

A place-based approach is a promising approach for better understanding the natural, built, and anthropogenic urban environment by analyzing the spatial interrelationships between various elements. This approach is crucial in addressing the complexities of urban energy systems and can serve as a powerful tool for various energy applications [29,30]. Particularly, the three significant advantages of the implementation of the different typologies of USBEM using a place-based approach are as follows: the increase in the number, typology, and accuracy of the energy-related variables, the possibility of expanding the range of energy models to choose from, and, lastly, the opportunity to identify the most effective energy model for each specific application. Additionally, by considering technical, environmental, and socio-economic variables that can affect actual and future energy scenarios, it is possible to identify proper site-specific energy policies and plans [31,32]. Place-based USBEM can help redistribute sustainability targets across the territories, adapting them to the peculiarities of each context, exploiting opportunities, and compensating the measures for local constraints. The place-based approach is functional in reducing and translating global strategies into effective local actions [33].

3. Knowledge Gap and Objectives

Research into building physics and geomatics lacks an integrated approach, which hinders the creation of place-based urban-scale building energy modeling (USBEM). Building physics focuses on the physical phenomena that regulate building energy consumption at a building or block scale. Geomatics deals with databases at the urban and territorial scale and uses open-source geographic information systems such as QGIS to extract and calculate geo-referenced information (e.g., raster or vector data), which can be used for energy modeling at the urban scale. Combining these two fields can create a more comprehensive and effective approach to USBEM.

In order to accurately calculate the energy consumption of a building or block, extensive information on the building envelope, technological systems, and surrounding context must be incorporated into the energy balance equation. However, this level of detailed information is not always available for entire cities. In these cases, QGIS can be utilized to calculate the missing variables and incorporate them into the balance equations. This process ensures a more comprehensive and accurate calculation of energy consumption at the urban scale. This work provides insight into place-based USBEM by explaining the different types of USBEM models, from pre-modeling to energy modeling, result representation, and application fields. The pre-modeling stage is enriched by considering spatial autocorrelation, which identifies clusters or areas with similar characteristics that influence energy consumption. This is a novel improvement, as spatial autocorrelation has not been investigated in energy modeling, even in place-based USBEM. Incorporating this improvement into urban-scale models can be significant, as some information may not be available, and it enables accounting for different energy-use types.

According to the literature, USBEM using a place-based approach offers numerous advantages, such as performing site-specific analyses and considering each building's characteristics that influence its energy consumption. These models can work on various layers and scales, use all the available data, and understand the spatial variations within different areas of a city. They can easily shift from the building to the urban scale and compare different scenarios and analyze them from various perspectives with an interactive interpretation of the results. Place-based USBEM can evaluate the impact of localized interventions on a larger scale. However, it has limitations, including limited data availability on a large scale, language barriers, and privacy concerns. Nevertheless, with expertise in buildings physics and geomatics and experience in locating and managing vast datasets, it is possible to process data or surrogate data to create models for predicting building energy consumption at the urban scale.

The aim of this paper is to assess the effectiveness of the place-based approach in urban building energy modeling (UBEM). Specifically, this study will explore the challenges associated with the scarcity of data on a large scale, language barriers, and privacy concerns and propose methods to overcome these limitations. The goal is to provide insights into how place-based UBEM can be improved to better predict the spatial distribution of energy consumption and compare different scenarios at the urban scale.

4. Place-Based Urban Building Energy Modeling

In this section, the implementation of the place-based approach in USBEM is described in detail. In order to treat all of the energy models found in the literature, the place-based approach is implemented into data-driven, process-driven, and hybrid energy models.

The methodology presented here is illustrated in Figure 1 and described in the steps below:

1. Pre-modeling with:
    - Data collection: the collection of input data/geo-databases and geo-localization of urban environment data.
    - Pre-processing phase: correction, integration, and spatialization of databases and evaluation of spatial correlations and local climate conditions.
    - Geo-database creation: the creation of a complete and accurate geo-database for energy modeling;
2. Energy modeling with: USBEM using a place-based approach: application of the place-based approach to data-driven, process-driven, and hybrid modeling;
3. Calibration: error evaluation and adjustments to input data to minimize errors between the data measured and calculated by the model, making the model more robust.

The geographic information system QGIS is used in all phases of the place-based process because it allows the user to upload geo-referenced data, yet it also has various plugins that manage the databases, calculate other data, operate with information at different scales, and helps create data-driven models through specific tools and algorithms.

After collecting the available data and processing the missing data, the geo-database is created, and it provides the information to build USBEM. Then, the geo-database splits into two datasets to train and then test the model. The calibration allows for the minimization of errors and to improve the robustness of the model.

For USBEM using a place-based approach, some of the useful QGIS plugins and tools are "Smart maps", "R" and "Orfeo Toolbox provider" for statistical analyses and machine learning algorithms.

Figure 1 is explained in more detail in the following paragraphs with some examples of the results to explain the different types of outcomes, such as the numerical values, tables, graphs, and maps.

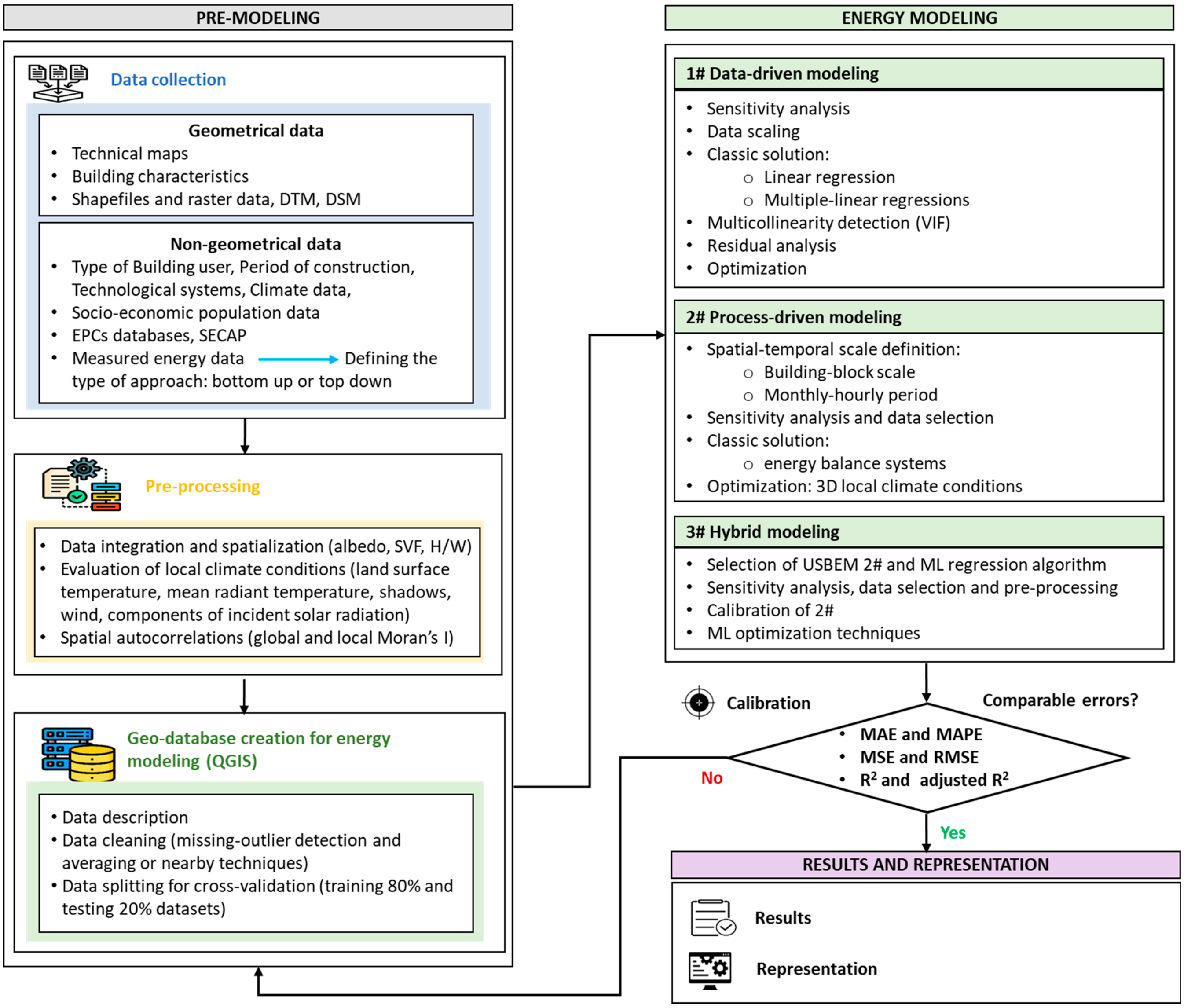

Figure 1. Methodology of USBEM using a place-based approach.

4.1. Pre-modeling

This is one of the most important phases of urban buildings energy modeling because it will influence the type of energy modeling and its accuracy.

4.1.1. Data Collection

Having a comprehensive and precise database with all information about a territory is a crucial phase in urban energy modeling, yet it remains a challenge [34]. Ensuring data availability is a crucial factor to consider when developing effective urban energy models. However, accessing data, especially disaggregated data, can be challenging due to privacy concerns [35]. Despite these challenges, there is a current political trend towards sharing databases and geo-databases containing comprehensive information about territories to facilitate knowledge dissemination, analysis, and problem-solving by various stakeholders such as citizens, companies, government, and public bodies. These datasets cover various fields of applications, including the orography of the territory, the characteristics and morphology of the built environment, and the population distribution by type.

Each geo-database allows for the integration of data with their corresponding geographic positions. These data comprise geometric and non-geometric datasets. Geometric data are mainly acquired from cartographic databases, websites, and remote sensors, such as technical maps (e.g., shapefiles), terrain orography using digital terrain models DTMs, digital surface models (DSMs) (e.g., Lidar flight), and land use, among others. Additionally, non-geometric data are also included, such as climate data, the typological features of buildings, the registry of thermal plants, building archetypes, energy performance certificates (EPCs), socioeconomic data of the population (from statistical census databases), and measured energy consumption data (annual, monthly, daily, and hourly data).

In order to apply the place-based approach to energy performance evaluations, the free and open-source Geographic Information System, QGIS 3.28 (Quantum GIS), can be used. QGIS is among the most extensively used software that enables geographical observations of the real world within digital systems, making it possible to formalize places with their mathematical and physical features using GIS MODELING [36,37]. The collection of input data and the categorization of existing databases are carried out in relation to the spatial and temporal scale of the information, the level of ownership, and integrations with other data.

Some examples of typical Italian databases useful for USBEM are

- National Geoportal [38];
- ISPRA (Institute for Environmental Protection and Research, Department for the Geological Service of Italy) [39];
- National Territorial Data: [40].

In general, all countries have these geo-databases, but the language of the country itself is used.

For Europe, the most known databases for USBEM (in English) are

- The official portal for European data [41];
- INfrastructure for SPatial Information (INSPIRE) Geoportal [42];
- European Centre for Disease Prevention and Control [43].

In the world, some data for USBEM (in English) can be found via

- International Energy Agency (IEA) [44];
- National Aeronautics and Space Administration (NASA) POWER Project [45];
- Group on Earth Observation (GEOS) [46].

4.1.2. Preprocessing Phase

QGIS enables the visualization and management of large geo-databases while providing a wide range of tools to add new input data. The preprocessing phase is critical in drafting the most accurate and comprehensive input geo-dataset, as the outcome of the model depends on the quality of the initial information. The preprocessing phase encompasses the following steps:

- acquisition of additional information by utilizing the QGIS tool and plugins; for input data, the geo-localization consents to enrich the set of information even if it uses different scales and accuracies. Moreover, the spatial representation of the data consents to visualizing the superimposition of the data, investigating more aspects;
- allowing a qualitative assessment of the spatial relationships between nearby geometries and features through the calculation of the statistics and indicators that describe spatial autocorrelations (e.g., global and local Moran's I index). The use of an adjacency matrix and spatial-temporal weights can help calibrate the model by adjusting the energy-related variables, leading to a better explanation of the spatialized results.

This work introduces spatial autocorrelation, which allows for the association of the numerical values of a variable with their spatial position. This technique has been used, above all, for research on flora and fauna [47] at the territorial scale, but also for assessments on the risk of flooding [48] or COVID-19 contagion [49]. As far as energy models are concerned, it would be very important to introduce this when some of the data related to energy use are not known, yet the spatial correlations can be identified.

4.1.3. Creation of Geo-Database for Energy Modeling

In order to prepare for energy modeling, it is necessary to create a final geo-database that combines and validates the previous two phases based on the specific investigation and case study application. In QGIS, the spatial analyst tools consent to the combination of databases while considering the location of the variables. The validation of the databases on a territorial scale is fundamental to identifying and eliminating missing data and avoiding anomalous data [18]. Statistical tools can be used to check the validity of the databases by detecting null values, completing databases (if possible), and identifying and correcting/removing anomalous data, such as outliers or missing values.

The process of creating the final geo-database information takes place through

- Data description: an exploratory study of the dataset using basic statistical calculations (e.g., count, null values, mean, standard deviation, etc.) to describe data distribution (e.g., normal or gamma distribution) and data type (e.g., integer, categorical, etc.);
- Data cleaning: detecting and handling missing and outlier values using different methods based on the nature of the dataset (i.e., averaging or nearby techniques);
- Data splitting for cross-validation: the database is split into two datasets, namely training and testing data, to train and test the models. This allows for the generalization and strengthening of the energy modeling [50].

The dimension of a geo-database depends on the quality of the selected energy-related variables, which are closely linked to the available measured energy consumption data. Depending on the aggregation level at which the energy consumption data is provided, either a top-down or bottom-up approach [51] will be used in the subsequent phase of energy modeling for different buildings. The literature suggests that the bottom-up model is more accurate, especially with a place-based approach [15,52]; thus, this work focuses mainly on bottom-up models.

4.2. Energy Modeling

In this section, three types of USBEMs that are associated with the place-based approach are described: data-driven, process-driven, and hybrid models.

4.2.1. Data-Driven Models

Data-driven modeling comprises statistical and AI approaches, with the latter primarily relying on machine learning (ML) techniques. In order to achieve accurate outputs, both approaches require a large amount of data on buildings and urban environments. They imply the use of regression algorithms of an increasing level of complexity to find a more accurate association between the influential parameters and energy consumption.

Statistical approaches are widely used, but fast processing and user-friendly methods collide with the inability to include complex energy patterns and detailed descriptions of real urban peculiarities [37].

ML algorithms have proven their capability in handling complex and non-linear datasets, especially those that are multivariate and prone to noise, which can often occur when dealing with urban energy models. The computational power of ML techniques enables the consideration of stochastic variables, the identification of new patterns, and the creation of clusters, leading to improved accuracy for energy predictions.

However, when it comes to black-box models, the intricate nature of urban systems can sometimes make it challenging to interpret the outputs and establish connections with the inputs [53]. Thus, the outcome model can be strictly site-specific, requiring detailed and hard-to-access spatiotemporal data, leading to crucial challenges in dealing with their scalability and generalizability in urban-scale studies [54]. It is worth noting that the interpretation of model outcomes is crucial for decision-making frameworks and policymakers [55]. Despite the challenges posed by black-box models, significant efforts have been made to improve their interpretability, and data-driven models remain at the forefront of urban studies. ML algorithms have restricted application fields when implemented at the urban scale because sometimes they become too complicated and non-interpretable.

Data-driven models (in Figure 2) are concerned with the following:

- Sensitivity analysis: this helps to identify the most influential variables on energy consumption. Univariate and multivariate analyses are the most common techniques to investigate the relationship between one or more variables with the outcome; they can be performed by using Pearson's coefficient, correlation matrices, or heat maps (principal component analysis). For large geo-databases, principal component analysis (PCA) is widely used [18];
- Data scaling: this includes normalization and standardization techniques. The normalization process measures the similarity of two datasets (e.g., Kernel function) and consists of scaling individual samples from 0 to 1; the standardization of datasets is required mainly by ML estimators and when variables have a normal distribution;
- Classic solutions: these include linear regression (LR), multiple linear regression (MLR), or logarithmic regression models. More accurate models can be implemented: polynomial regression (PR), support vector machine (SVM), random forest (RF), decision tree (DT), artificial neural network (ANN), Gaussian process (GP), and gradient-boosted regression trees (GBRT). For the general regression problems in the energy sector, RF, GBRT, ANN, SVM, and GP are the most used models [56,57];
- Multicollinearity detection: this enables the user to test the independence between energy-related variables; the most common technique is VIF (variable inflation factors);
- Residual analysis or homoscedasticity: this concerns the homogeneous variance of the residuals; the variance in the errors should not depend on the variables (e.g., White test).

Figure 2 presents an example of the results of a typical statistical bottom-up energy model applied in the city of Turin (Italy) for residential and non-residential building archetypes [37]. The mean energy consumption of some archetypes have been applied to the buildings of the city of Turin, considering the main energy-related variables, such as volume, building use, and period of construction. The spatial distribution of the results provides a clearer and more immediate picture of the energy consumption of the city, thus making information accessible to non-technical personnel as well. These results were used to evaluate the spatial distribution of natural gas consumption and the emissions due to building use and to highlight critical neighborhoods.

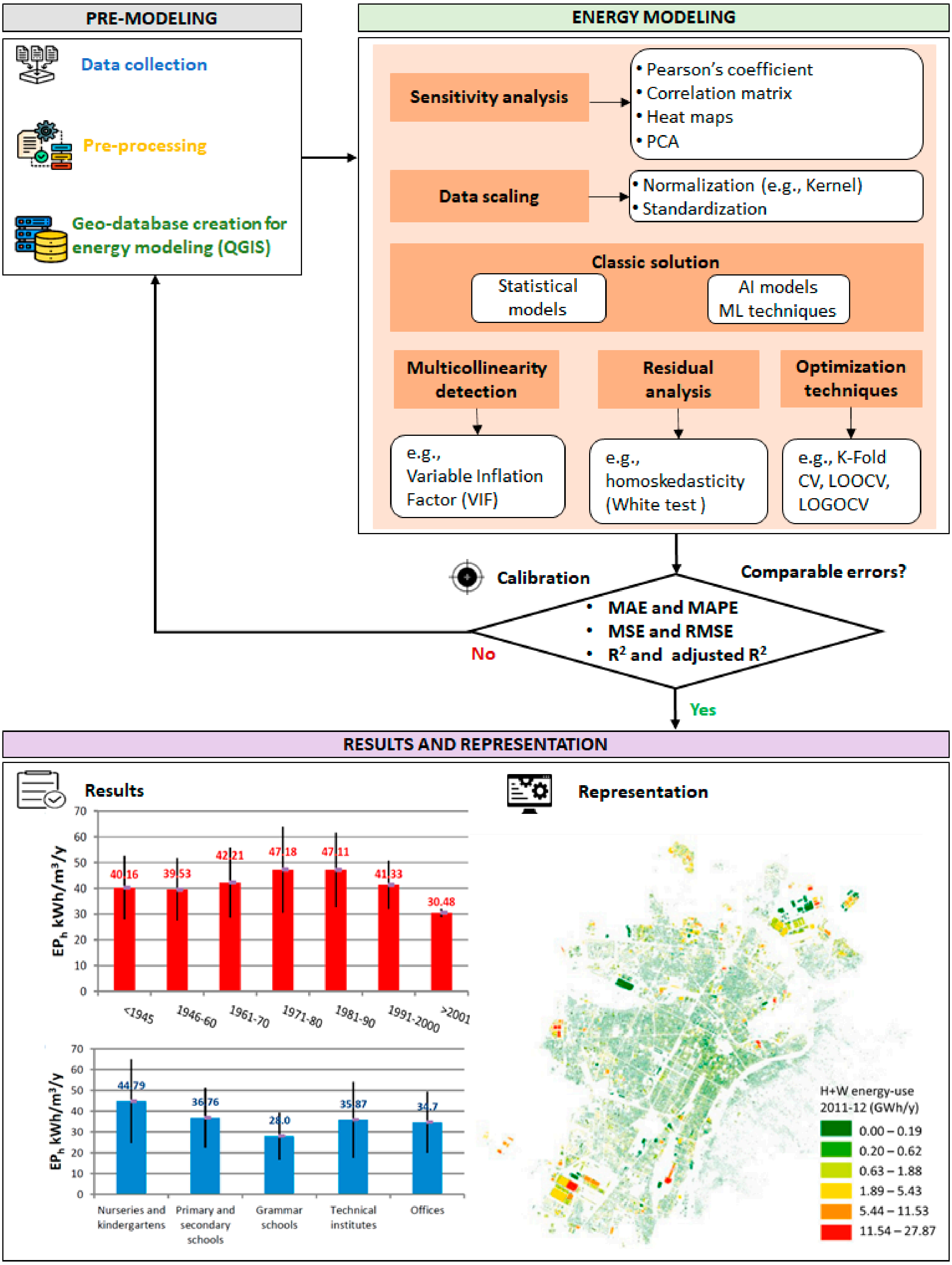

Figure 2. Data-driven energy modeling.

4.2.2. Process-Driven Models

Process-driven models rely on physical laws to calculate the energy performance of buildings. They require a complete set of data for a detailed description of the physical phenomena. Usually, these models are based on energy balance equations between buildings and outdoor environments, considering the geometry and characteristics of the buildings, human behavior, the urban environment, and climate data. Simulation engines are used to describe the energy and heat transfer mechanisms of buildings, providing great flexibility regarding the application fields; once the model is validated, the energy balance equations can be adapted to any context and spatial-temporal scales [52]. Limitations occur when considering uncertainties and stochastic variables that do not follow physical phenomena [20].

Process-driven models (Figure 3) are concerned with the following:

- Spatial-temporal scale definition: this enables the definition of an energy balance system. It is necessary to describe the thermodynamic system and the spatial-temporal scale on which the energy balance equations will be applied (e.g., spatial boundary and temporal period); the spatial boundary of the thermodynamic system could occur at various scales; generally, it is applied at the building scale, and then the results are aggregated at the district and urban scales;
- Sensitivity analysis and data selection: these include the choice of variables and typical data useful for describing the energy balances between heated/cooled built spaces and the outside environment. These data describe the whole characteristics of urban environments considering the operational indoor conditions according to thermal comfort, air quality, and lighting requirements.
- Three-dimensional local climate conditions evaluation: this concerns the definition of a detailed climate database considering the measured data survey. For models using a place-based approach, the evaluation of local climate conditions is required for a three-dimensional environment for an accurate description of the main climate-driven variables in the energy balance equations; research is in progress to develop QGIS plugin tools [56];
- Classic solution: this is represented by an energy balance system that has the aim of evaluating the energy consumption of buildings for different services [52,58]. The energy balance system takes into account various equations for each energy service and their interplay, which is in line with the prevailing standards for assessing building energy performance. Usually, to calculate the energy demand for space heating and cooling, three equations based on an iterative procedure between three thermodynamic systems (TSs) are used (in Figure 3): the opaque envelope, the glazing components, and the indoor building spaces, which include internal partitions, horizontal structures, air, the occupants, and the furniture [56]. Typically, the energy demand for hot water production and electrical use is also incorporated into the energy balance system.

Figure 3 presents some results of a typical process-driven energy model for residential buildings [52,58]. The hourly energy demand of a building for space heating was calculated by mainly considering the construction period, according to the materials and technologies used, and the solar exposition and the geometry characteristics calculated with QGIS. For a winter day, the graph of the results shows the heat flows that are considered in the balance equations, with the power supplied by the heating system in red. The pick power in the early morning was used to evaluate the built volumes that can be connected to the district heating network [36], taking into account all the existing technical, economic, environmental, and social constraints. On the right, the different components of the heat fluxes are presented for the average monthly day to describe consumption variation during the year.

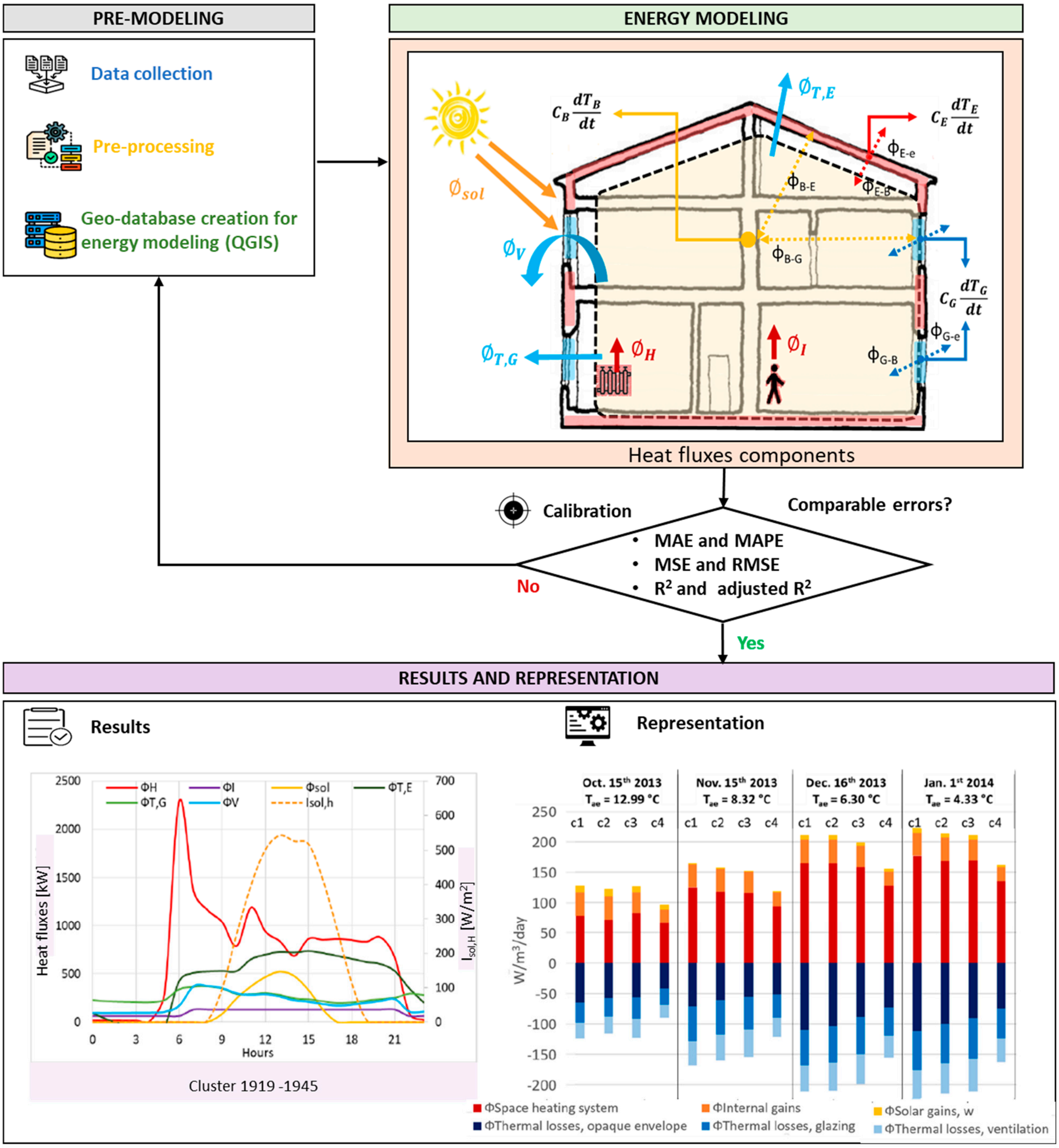

Figure 3. Process-driven energy modeling.

4.2.3. Hybrid Models

Hybrid modeling is an approach that combines the strengths of both data-driven and process-driven modeling to achieve the best possible outcome. By leveraging the computational efficiency of the former and the ability to elucidate the physical relationships between the variables of the latter, hybrid modeling represents a powerful synthesis of these two approaches [57]. Recent works demonstrate that hybrid models have better performance, handling complex situations and unexpected trends with high accuracy [59]. For USBEM, the evaluation of geometrical data and the identification of typological features can sometimes be imprecise (e.g., user profiles cannot be estimated with average data). However, by using a process-driven model in conjunction with a supportive ML technique, it is possible to select a range of variability for certain energy-related variables, thus allowing for better model calibration [56]. Alternatively, a data-driven model can be employed with physical-based correlations introduced. For USBEM using a place-based approach, physical-based modeling has more advantages in describing energy consumption due to its spatial morphology dependence and its site-specific applications.

The proposed Hybrid model (Figure 4) is concerned with the following:

- Selection of process-driven model;
- The use of an ML algorithm can improve the accuracy of the results with the optimized use of the energy-related variables and constant data in the energy balance. Some of the most used ML algorithms are RF, GBRT, ANN, SVM, and GP [56,57]. The identification of their hyperparameters entails an overall good balance between modeling performance and accuracy and simulation time.

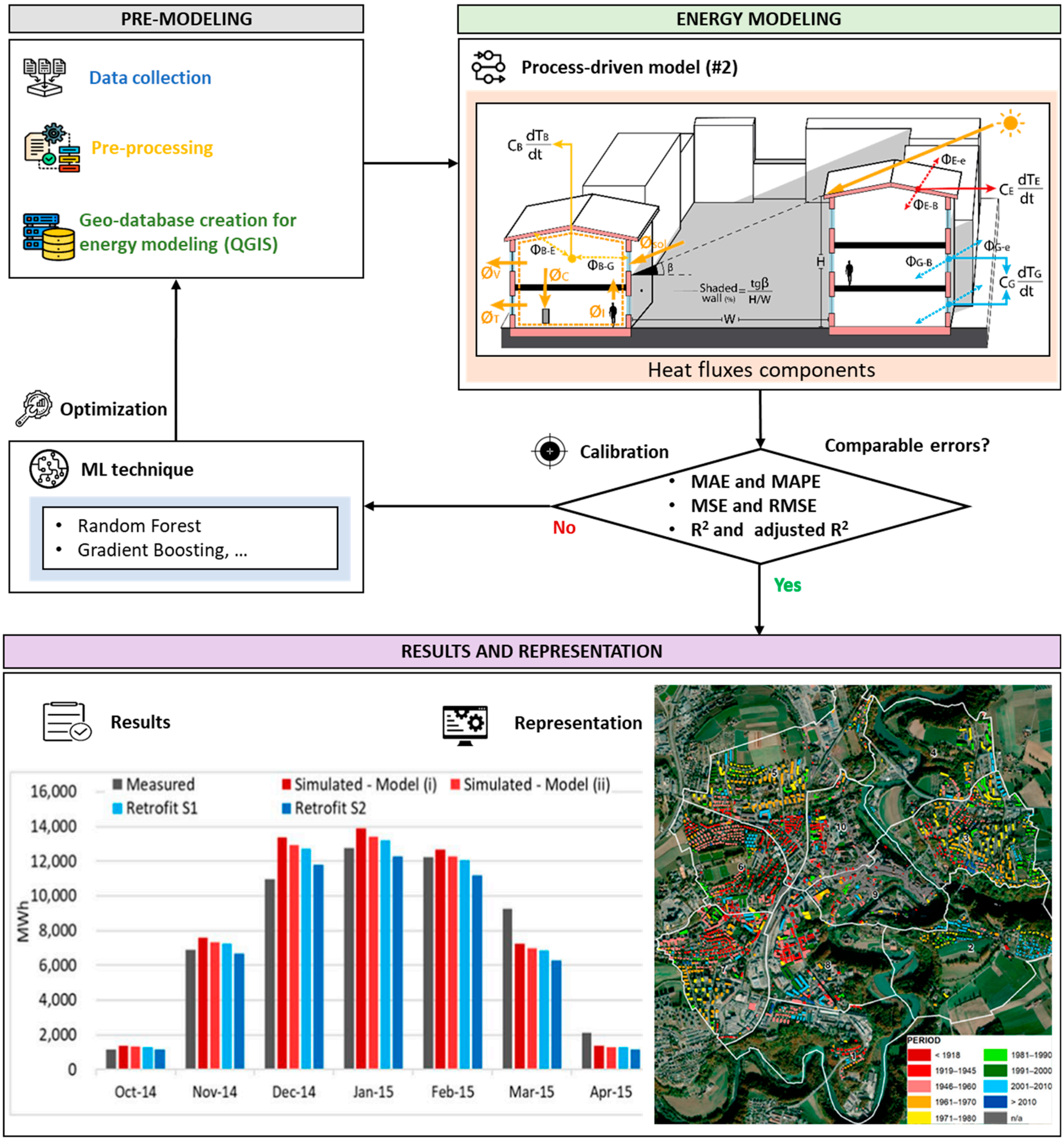

Figure 4. Hybrid energy modeling.

Figure 4 describes some of the results of hybrid modeling based on a process-driven model and with a machine learning optimization algorithm that improves the accuracy of the modeling [36,59,60]. The process-driven model is an hourly balance that considers all thermal fluxes between the buildings and the external environment. The random forest machine learning algorithm was used to train the model using an initial dataset of the more energy-related variables and constants that could be improved to reduce the error of the modeling [58]. The resulting graph compares the monthly energy consumption with two different hypotheses using the datasets and considering two retrofit scenarios. The representation map describes the characteristics of the heritage of the building in the city of Fribourg, with its spatial distribution used for many purposes, such as identifying the energy efficiency level of the building and selecting the more convenient retrofit interventions or energy-saving policies in each specific city.

4.2.4. Calibration

In order to ensure that the model accurately predicts energy consumption, it undergoes a calibration process aimed at minimizing the differences between the predictions and the measured data. This process can be manual or automatic, with the objective being to adjust or modify the input data to reduce errors [59]. Manual calibration is very time-consuming but can be driven by knowing the physical process, allowing the user to search, create, and add other variables to the model. Automatic calibration uses an algorithm that is more or less complex, using an iterative process that selects the more energy-related data, identifies the range of variation of each variable, and mixes the order of the iterative process to obtain the best results.

In order to evaluate the accuracy of the model, various error metrics are considered, which may include but are not limited to the following:

- Mean absolute error (MAE) and mean absolute percentage error (MAPE): these are calculated by the mean value of absolute error or the mean percentage of the absolute error;
- Mean square error (MSE) and root mean square error (RMSE): mean square error is the ratio between the sum of the square error and the number of data (the deviation between the predicted result and the actual value, i.e., the variance). Its square root allows for the evaluation of the standard deviation;
- R square and adjusted R square: R square is calculated via the ratio between the mean square error and the total mean square error (it varies between 0 and 1, a bigger value indicates a better fit); adjusted R square considers the number of independent variables used.

ML techniques are also used to refine and improve the accuracy of energy models through the adjustment of the input data in regard to the continuous output. It also allows for the management of key energy-related physical and stochastic variables, assigning confidence intervals and testing ranges to address the uncertainties of the model [61]. Some ML methods can also check the training–testing databases: k-fold cross-validation (k-Fold CV), leave-one-out cross-validation (LOOCV), leave-one-group-out cross-validation (LOGOCV), or nested cross-validation [62].

5. Conclusions and Remarks

Urban building energy modeling is strongly influenced by the availability of data. A large amount of data is needed, and the completeness of any database is fundamental for the accuracy of the model. At the territorial level, this means working on thousands of buildings, with each of them having many data on different scales and levels of precision. The data collection and preprocessing phases are often long, difficult, onerous, and sometimes impossible. As an example, data on energy consumption are often not open data due to privacy problems and, therefore, are provided in an aggregated way and cannot be used with a bottom-up model.

The place-based approach can describe and represent energy-related parameters considering spatial specificities (e.g., climate and socio-economic conditions, urban forms, and solar exposition) with high flexibility, adaptability, and scalability. The use of a Geographic Information System (GIS) consents to identify the real use of energy due to the technical, economic, environmental, and social aspects of a territory. The spatial component is fundamental in evaluating the energy use, energy-saving potential, and energy production of a specific neighborhood to reach energy self-sufficiency. The flexibility of place-based modeling can be used to evaluate how retrofit interventions and low-carbon technological systems can reduce local or global benefits by using an energy-economic-environmental-social point of view.

There are various energy modeling approaches, each with its advantages and disadvantages that can affect the selection process. Process-driven models are dependable, need less data, and have short simulation times. However, as they rely on heat balance equations, they necessitate a complete dataset to describe the physical phenomenon, and they may not take into account stochastic aspects, such as human behavior or climate/technological anomalies.

Data-driven statistical models are widely used because they are user-friendly and have fast processing times. Nonetheless, they are often limited to specific applications, usually consider a limited number of building archetypes, and fail to capture complex patterns. Machine Learning models can learn and identify the aspects that affect energy consumption during the training phase, resulting in precise and fast predictions. However, they may lack the ability to generalize and have limited adaptability and replicability.

Data-driven energy modeling has a point of weakness because it works by using a provided geo-database and does not drive the researcher to look for other data, as is the case when using process-driven models, where all the components of the energy balance have to be accessed. Besides, they can explain stochastic phenomena better.

Hybrid models, which combine the advantages of process-driven and data-driven models, are the most interesting ones and are widely used. These models utilize the robust energy balance system of process-driven models and the data selection of variables and constants through a machine learning optimization algorithm.

The results of these models, together with a place-based approach, make it possible to create spatial representations and maps that can better explain how energy is used through the superimposition of variables. The place-based approach consents to also consider technical, environmental, and socio-economic constraints that can limit retrofit interventions, the installation of clean technologies, and the adoption of policies in a territory; still, policies can force sustainable development in some areas with fewer criticalities to achieve an overall sustainability target in the territory.

The research on place-based USBEMs is still under development. The models have already been applied to small, large, and medium-sized cities. These models have potential applications of great interest, especially in the current energy, economic, and environmental crisis. The place-based approach used in conjunction with the models discussed in this paper enables researchers to create spatial representations and maps that consider site-specific building characteristics, spatial variations within a specific area of a city, and different scales. By utilizing all the available data, researchers can easily shift between scales and compare different scenarios, enabling them to evaluate the impact of localized interventions on a larger scale.

Author Contributions: Conceptualization, G.M. and P.V.; methodology, G.M. and P.V.; writing—original draft preparation, G.M., P.V. and K.J.; writing—review and editing, G.M., P.V. and K.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. United Nations. Available online: https://population.un.org/wup/Download/ (accessed on 22 February 2023).
2. European Commission. Available online: https://transport.ec.europa.eu/media-corner/publications/statistical-pocketbook-2022_en (accessed on 22 February 2023).
3. European Commission. Available online: https://ec.europa.eu/eurostat/web/main/data/database (accessed on 22 February 2023).
4. Todeschi, V.; Marocco, P.; Mutani, G.; Lanzini, A.; Santarelli, M. Towards energy self-consumption and self-sufficiency in urban energy communities. *Int. J. Heat Technol.* **2021**, *39*, 1–11. [CrossRef]
5. Perera, A.T.D.; Javanroodi, K.; Wang, Y.; Hong, T. Urban cells: Extending the energy hub concept to facilitate sector and spatial coupling. *Adv. Appl. Energy* **2021**, *3*, 100046. [CrossRef]
6. European Commission. Available online: https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en (accessed on 22 February 2023).
7. Perera, A.T.D.; Javanroodi, K.; Mauree, D.; Nik, V.M.; Florio, P.; Hong, T.; Chen, D. Challenges resulting from urban density and climate change for the EU energy transition. *Nat. Energy* **2023**, *8*, 397–412. [CrossRef]
8. Heidelberger, E.; Rakhahttps, T. Inclusive urban building energy modeling through socioeconomic data: A persona-based case study for an underrepresented community. *Build. Environ.* **2022**, *222*, 15. [CrossRef]
9. Harish, V.S.K.V.; Kumar, A. A review on modeling and simulation of building energy systems. *Renew. Sust. Energy Rev.* **2016**, *56*, 1272–1292. [CrossRef]
10. Pagliarini, G.; Rainieri, S.; Vocale, P. Energy efficiency of existing buildings: Optimization of building cooling, heating and power (BCHP) system. *Energy Environ.* **2014**, *25*, 1423–1438. [CrossRef]
11. Mancini, F.; Romano, S.; Basso, G.L.; Cimaglia, J.; De Santoli, L. How the Italian residential sector could contribute to load flexibility in demand response activities: A methodology for residential clustering and developing a flexibility strategy. *Energies* **2020**, *13*, 3359. [CrossRef]
12. Ang, Y.Q.; Berzolla, Z.M.; Reinhart, C.F. From concept to application: A review of use cases in urban building energy modeling. *Appl. Energy* **2020**, *279*, 115738. [CrossRef]
13. Basu, S.; Bale, C.S.E.; Wehnert, T.; Topp, K. A complexity approach to defining urban energy systems. *Cities* **2019**, *95*, 102358. [CrossRef]
14. Abolhassani, S.S.; Amayri, M.; Bouguila, N.; Eicker, U. A new workflow for detailed urban scale building energy modeling using spatial joining of attributes for archetype selection. *J. Build. Eng.* **2022**, *46*, 103661. [CrossRef]
15. 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. *Sustain. Cities Soc.* **2020**, *62*, 102408. [CrossRef]
16. Swan, L.G.; Ugursal, V.I. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renew. Sust. Energy Rev.* **2009**, *13*, 1819–1835. [CrossRef]
17. Ali, U.; Shamsi, M.H.; Hoare, C.; Mangina, E.; O'Donnell, J. Review of urban building energy modeling (UBEM) approaches, methods and tools using qualitative and quantitative analysis. *Energy Build.* **2021**, *246*, 111073. [CrossRef]
18. Malhotra, A.; Bischof, J.; Nichersu, A.; Häfele, K.-H.; Exenberger, J.; Sood, D.; Allan, J.; Frisch, J.; van Treeck, C.; O'Donnell, J.; et al. Information modelling for urban building energy simulation—A taxonomic review. *Build. Environ.* **2022**, *208*, 108552. [CrossRef]
19. Sunm, Y.; Haghighatm, F.; Fungm, B.C.M. A review of the-state-of-the-art in data-driven approaches for building energy prediction. *Energy Build.* **2020**, *221*, 110022. [CrossRef]
20. Yang, X.; Liu, S.; Zou, Y.; Ji, W.; Zhang, Q.; Ahmed, A.; Han, X.; Shen, Y.; Zhang, S. Energy-saving potential prediction models for large-scale building: A state-of-the-art review. *Renew. Sust. Energy Rev.* **2022**, *156*, 111992. [CrossRef]
21. Gassar, A.A.A.; Cha, S.H. Energy prediction techniques for large-scale buildings towards a sustainable built environment: A review. *Energy Build.* **2020**, *224*, 110238. [CrossRef]
22. Sun, J.; Gong, M.; Zhao, Y.; Han, C.; Jing, L.; Yang, P. A hybrid deep reinforcement learning ensemble optimization model for heat load energy-saving prediction. *J. Build. Eng.* **2022**, *58*, 105031. [CrossRef]
23. Wang, L.; Lee, E.W.M.; Yuen, R.K.K. From concept to application: Novel dynamic forecasting model for building cooling loads combining an artificial neural network and an ensemble approach. *Appl. Energy* **2018**, *228*, 1740–1753. [CrossRef]
24. Wong, C.H.H.; Cai, M.; Ren, C.; Huang, Y.; Liao, C.; Yin, S. Modelling building energy use at urban scale: A review on their account for the urban environment. *Build. Environ.* **2021**, *205*, 108235. [CrossRef]
25. 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. *Renew. Sust. Energy Rev.* **2015**, *52*, 1391–1404. [CrossRef]
26. Doma, A.; Ouf, M. Modelling occupant behaviour for urban scale simulation: Review of available approaches and tools. *Build. Simul.* **2023**, *16*, 169–184. [CrossRef]
27. Bouw, K.; Noorman, K.J.; Wiekens, C.J.; Faaij, A. Local energy planning in the built environment: An analysis of model characteristics. *Renew. Sust. Energy Rev.* **2021**, *144*, 111030. [CrossRef]
28. Mutani, G.; Todeschi, V. Urban Building Energy Modeling: An hourly energy balance model of residential buildings at a district scale. *J. Phys. Conf. Ser.* **2020**, *1599*, 012035. [CrossRef]
29. Wink, R.; Kirchner, L.; Koch, F.; Speda, D. There are Many Roads to Reindustrialization and Resilience: Place-based Approaches in Three German Urban Regions. *Eur. Plan. Stud.* **2016**, *24*, 463–488. [CrossRef]

30. Devine-Wright, P. Decarbonisation of industrial clusters: A place-based research agenda. *Energy Res. Soc. Sci.* **2022**, *91*, 102725. [CrossRef]
31. Johari, F.; Peronato, G.; Sadeghian, P.; Zhao, X.; Widén, J. Urban building energy modeling: State of the art and future prospects. *Renew. Sust. Energy Rev.* **2020**, *128*, 109902. [CrossRef]
32. Mutani, G.; Todeschi, V. Optimization of Costs and Self-Sufficiency for Roof Integrated Photovoltaic Technologies on Residential Buildings. *Energies* **2021**, *14*, 4018. [CrossRef]
33. Mutani, G.; Santantonio, S.; Brunetta, G.; Caldarice, O.; Demichela, M. An Energy Community for Territorial Resilience. The Measurement of the Risk of Energy Supply Blackout. *Energy Build.* **2021**, *240*, 110906. [CrossRef]
34. Mutani, G.; Todeschi, V.M. An Urban Energy Atlas and Engineering Model for Resilient Cities. *Int. J. Heat Technol.* **2019**, *37*, 936–947. [CrossRef]
35. Abbasabadi, N.; Mehdi Ashayeri, J.K. Urban energy use modeling methods and tools: A review and an outlook. *Build. Environ.* **2019**, *161*, 106270. [CrossRef]
36. Mutani, G.; Todeschi, V. GIS-based urban energy modelling and energy efficiency scenarios using the energy performance certificate database. *Energy Effic.* **2021**, *14*, 47. [CrossRef]
37. Mutani, G.; Todeschi, V. Space heating models at urban scale for buildings in the city of Turin (Italy). *Energy Procedia* **2017**, *122*, 841–846. [CrossRef]
38. Italian National Geoportal. Available online: http://www.pcn.minambiente.it/mattm/en/ (accessed on 5 January 2023).
39. ISPRA (Italian Institute for Environmental Protection and Research, Department for the Geological Service of Italy). Available online: http://portalesgi.isprambiente.it/en (accessed on 5 January 2023).
40. Italian National Territorial Data. Available online: https://geodati.gov.it/geoportalRNDTPA/rest/find/document?f=html&searchText=apiso.Language%3Aeng (accessed on 5 January 2023).
41. The Official Portal for European Data. Available online: https://data.europa.eu/data/datasets?locale=en (accessed on 5 January 2023).
42. INfrastructure for SPatial Information (INSPIRE) Geoportal. Available online: https://inspire-geoportal.ec.europa.eu/ (accessed on 5 January 2023).
43. European Centre for Disease Prevention and Control. Available online: https://www.ecdc.europa.eu/en/publications-data/ecdc-geoportal (accessed on 5 January 2023).
44. International Energy Agency (IEA). Available online: https://www.iea.org/data-and-statistics/data-product/energy-and-emissions-per-value-added-database (accessed on 5 January 2023).
45. National Aeronautics and Space Administration (NASA) POWER Project. Available online: https://power.larc.nasa.gov/ (accessed on 5 January 2023).
46. Group on Earth Observation (GEOS). Available online: https://www.geoportal.org/?m:activeLayerTileId=osm&f:dataSource=dab (accessed on 5 January 2023).
47. Browning, E.; Freeman, R.; Boughey, K.L.; Isaac, N.J.B.; Jones, K.E. Accounting for spatial autocorrelation and environment are important to derive robust bat population trends from citizen science data. *Ecol. Indic.* **2022**, *136*, 108719. [CrossRef]
48. Li, H.; Zhang, C.; Chen, M.; Shen, D.; Niu, Y. Data-driven surrogate modeling: Introducing spatial lag to consider spatial autocorrelation of flooding within urban drainage systems. *Environ. Model. Softw.* **2023**, *161*, 105623. [CrossRef]
49. Freitas, W.W.L.; de Souza, R.M.C.R.; Amaral, G.J.A.; De Bastiani, F. Exploratory spatial analysis for interval data: A new autocorrelation index with COVID-19 and rent price applications. *Expert Syst. Appl.* **2022**, *195*, 116561. [CrossRef]
50. Shams Amiri, S.; Mueller, M.; Hoque, S. Investigating the application of a commercial and residential energy consumption prediction model for urban Planning scenarios with Machine Learning and Shapley Additive explanation methods. *Energy Build.* **2023**, *287*, 112965. [CrossRef]
51. Hong, T.; Chen, Y.; Luo, X.; Luo, N.; Lee, S.H. Ten questions on urban building energy modeling. *Build. Environ.* **2020**, *168*, 106508. [CrossRef]
52. Mutani, G.; Todeschi, V.M. Building energy modeling at neighborhood scale. *Energy Effic.* **2020**, *13*, 1353–1386. [CrossRef]
53. Manandhar, P.; Rafiq, H.; Rodriguez-Ubinas, E. Current status, challenges, and prospects of data-driven urban energy modeling: A review of machine learning methods. *Energy Rep.* **2023**, *9*, 2757–2776. [CrossRef]
54. Adilkhanova, I.; Ngarambe, J.; Yun, G.Y. Recent advances in black box and white-box models for urban heat island prediction: Implications of fusing the two methods. *Renew. Sust. Energy Rev.* **2022**, *165*, 112520. [CrossRef]
55. Chen, Z.; Xiao, F.; Guo, F.; Yan, J. Interpretable machine learning for building energy management: A state-of-the-art review. *Adv. Appl. Energy* **2023**, *9*, 100123. [CrossRef]
56. Todeschi VBoghetti, R.; Kämpf, J.H.; Mutani, G. Evaluation of Urban-Scale Building Energy-Use Models and Tools—Application for the City of Fribourg, Switzerland. *Sustainability* **2021**, *13*, 1595. [CrossRef]
57. Boghetti, R.; Fantozzi, F.; Kampf, J.; Mutani, G.; Salvadori, G.; Todeschi, V. Building Energy Models with Morphological Urban-Scale Parameters: A Case Study in Turin. In Proceedings of the 4th IBPSA-Italy Conference on Building Simulation Applications, BSA 2019, Bolzano, Italy, 19–21 June 2020. [CrossRef]
58. Mutani, G.; Todeschi, V.; Beltramino, S. Energy Consumption Models at Urban Scale to Measure Energy Resilience. *Sustainability* **2020**, *12*, 5678. [CrossRef]

59. Todeschi, V.; Javanroodi, K.; Castello, R.; Mohajeri, N.; Mutani, G.; Scartezzini, J.-L. Impact of the COVID-19 pandemic on the energy performance of residential neighborhoods and their occupancy behavior. *Sustain. Cities Soc.* **2022**, *82*, 103896. [CrossRef] [PubMed]
60. Mutani, G.; Todeschi, V.; Santantonio, S. Urban-Scale energy models: The relationship between cooling energy demand and urban form. *J. Phys. Conf. Ser.* **2022**, *2177*, 012016. [CrossRef]
61. Østergård, T.; Jensen, R.L.; Maagaard, S.E. A comparison of six metamodeling techniques applied to building performance simulations. *Appl. Energy* **2018**, *211*, 89–103. [CrossRef]
62. Heydari, A.; Nezhad, M.M.; Garcia, D.A.; Keynia, F.; De Santoli, L. Air pollution forecasting application based on deep learning model and optimization algorithm. *Clean Technol. Environ. Policy* **2022**, *24*, 607–621. [CrossRef]