MCIT
Modeling, control & information technologies

Framework for Dataset Development in Building Energy Balance Simulations

https://doi.org/10.31713/MCIT.2024.065

Svitlana Alyokhina
Department Industrial Engineering
University of Applied Sciences Technikum Wien
Vienna, Austria
svitlana.alyokhina@technikum-wien.at

Natalia Orlova
Department of Modeling and Identification of Thermal Processes in Energy Technology Equipment
Anatolii Pidhornyi Institute of Power Machines and Systems of the National academy of sciences of Ukraine
Kharkiv, Ukraine

***Abstract* — The article presents a structured approach to forming initial datasets necessary for calculating the energy balance of buildings, tailored to the district's level of digitalization. The methodology outlined addresses the often-neglected step of data preparation, which critically influences the outcomes of energy balance simulations. Two distinct process flows for dataset formation are introduced, reflecting high and medium levels of district digitalization. The paper's novelty lies in formalizing the dataset creation process, enabling more accurate energy simulations, particularly in the early planning stages of positive energy districts.**

***Keywords*— energy balance, building, process flow, digitalization, simulation**

I. Introduction

Building energy balance calculations and simulations are crucial for understanding and optimizing energy usage in living districts and cities [1], [2], etc. By identifying areas of inefficiency, these calculations enable targeted improvements, leading to reduced energy consumption and lower utility costs. This promotes sustainability, as it allows for the integration of renewable energy sources and green building practices, contributing to environmental conservation and a smaller carbon footprint. Efficient energy use not only lowers operational costs but also enhances inhabitants' comfort, productivity, and well-being. Buildings designed or renovated with a focus on energy balance tend to retain their value over time, making them more attractive in the real estate market. In essence, building energy balance calculations are fundamental for achieving energy efficiency, sustainability, cost savings, occupant comfort, and long-term asset value in the built environment.

With the growing emphasis on sustainable development and energy efficiency, accurately modeling and simulating the energy balance of buildings is critical. Energy balance simulation allows optimization of building designs to achieve higher energy efficiency and reduce carbon footprints. With increasing complexity in building design and environmental factors, the need for sophisticated and dynamic dataset preparation methods is more pressing than ever.

II. Actuality

Accurate dataset preparation is essential for constructing energy balance simulations, which are used to optimize energy consumption in buildings and urban environments. A series of studies have been conducted on this topic, and a brief analysis of some of them is presented below.

In the study [3] developed a 3D Building Energy Model (BEM) dataset generation framework that leverages human-AI synergies in early-stage building design. This approach addresses the challenge of manually creating 3D models by incorporating machine learning and artificial intelligence (AI) into the dataset preparation process, which accelerates the development of high-performance buildings. One of the primary advantages of this method is its ability to handle complex datasets that combine manual and automated processes, reducing human error and time. However, this method is resource-intensive and requires expertise in AI, making it less accessible to smaller firms or individuals without computational resources.

EnergyPlus is a widely used energy simulation program that supports dataset preparation through a comprehensive collection of environmental data, including climate data like TMY2 (Typical Meteorological Year). Developed by Crawley et al. [4], it provides detailed models for heating, cooling, and ventilation systems within buildings. The strength of EnergyPlus lies in its accuracy and the large dataset library available for simulation, making it ideal for robust energy balance models. However, the complexity of the tool and its steep learning curve present significant barriers to widespread adoption, particularly for non-specialists.

Oraiopoulos and Howard [5] examined the accuracy of Urban Building Energy Modelling (UBEM). The study highlights the complexity of scaling energy balance simulations from individual buildings to urban environments, where the interactions between multiple buildings influence the overall energy use. While UBEM is valuable for studying large-scale urban environments and their collective energy needs, the model often suffers from accuracy issues related to scaling errors. These arise from the complexity of urban form and the interactions between buildings, climate, and energy systems.

Vartholomaios and Chatzidimitriou [6] applied Monte Carlo dynamic energy simulations to explore the thermal loads of urban buildings. By simulating different parameters and variations in building design, their method offers insights into how different building forms impact energy efficiency. The Monte Carlo method is beneficial for its robustness in handling uncertainty and variability in input data. However, this approach requires substantial computational power and complex dataset handling, which may be a disadvantage in real-world applications where resources are limited.

Labiadh [7] proposed the use of in building energy simulations, which dynamically adjust parameters in response to changes in the environment and building use. These models offer a more flexible and responsive way to simulate energy balance, as they account for real-time variations in energy demand. However, the limitation of this approach is the difficulty in scaling to larger datasets or complex urban environments, as the adaptability of the model may not function effectively across diverse building types.

Zhao [8] explored the use of artificial intelligence and machine learning models for energy consumption analysis in large buildings. AI-driven models improve the efficiency and accuracy of energy balance simulations by recognizing complex patterns in the dataset that traditional methods may overlook. While these models significantly enhance prediction capabilities, they require extensive datasets and computational resources, limiting their accessibility for smaller projects.

The studies reviewed demonstrate that dataset preparation is crucial for accurate building energy balance simulations, particularly as energy efficiency becomes a global priority. Advances in artificial intelligence and adaptive modeling have significantly improved the ability to handle large, complex datasets. However, computational intensity and the requirement for specialized knowledge remain common drawbacks across many methods. Tools like EnergyPlus offer robust datasets but may be difficult to use without expert training, while AI-driven approaches bring adaptability but also introduce new complexities in terms of data requirements and accessibility.

III. Problem formulation

The current framework for dataset development in building energy balance simulations must address several key challenges to remain relevant and effective:

- Diverse Building Types: Modern urban environments feature a broad array of building designs, materials, and technologies. Simulations require datasets that capture this diversity to produce accurate energy predictions across different contexts.
- Dynamic and Real-Time Data: Energy consumption patterns are not static; they vary with occupancy, external weather conditions, and operational changes. The framework must include adaptive datasets that can dynamically adjust to these real-time variables for more precise simulations.
- Climate Sensitivity: Buildings in different geographic regions face unique climate challenges. The framework should incorporate climate-specific data (e.g., temperature, humidity, solar radiation) to accurately model energy interactions, particularly in heating and cooling systems.
- AI and Machine Learning Integration: Traditional datasets often fall short in capturing complex, multi-variable interactions between building systems and their environments. By incorporating AI and machine learning, the framework can process large volumes of data, identify patterns, and optimize the energy balance simulation process.
- Scalability from Single to Urban Models: As the focus on urban sustainability grows, the framework should allow scalability, enabling simulation not only for individual buildings but also for entire urban clusters, considering interactions between multiple buildings and the surrounding environment.

The purpose of this study is to formalize the main stages and highlight features of dataset forming for building energy balance simulation.

IV. Methods

The paper utilizes a structured methodology to create flow diagrams and define roles in the initial dataset formation for energy balance calculations of buildings, with variations based on the region's level of digitalization. Flow diagrams are constructed using Business Process Model and Notation (BPMN) [9], a widely used system that visually represents sequences in the process flow, ensuring clarity in understanding the stages and tasks involved.

V. Results

The process of calculating the energy balance is a complex concept that combines closely interrelated subsets of methods, tools, and data (Fig. 1). The specificity of the data subset influences the formation of the methods subset that will be used to calculate the energy balance. The methods in turn determine the tools for calculating the energy balance. The tools influence the formation of an initial data subset, which will subsequently be used by them through the prism of methods.

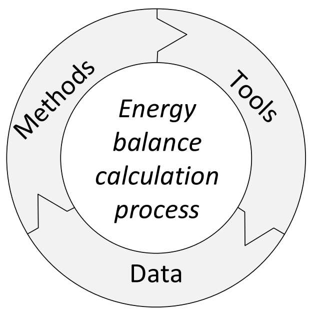

Figure 1. Subsets of energy balance calculation process

In general, the data required to analyze the energy balance of buildings can be divided into several groups (Fig. 2) [10].

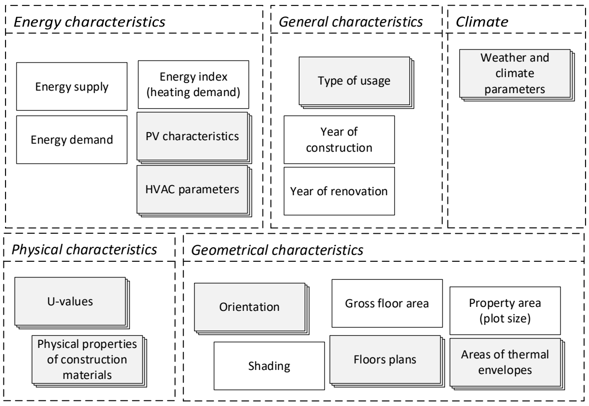

Figure 2. Initial dataset for buildings energy balance calculation

The data set “general characteristics” consists of data, that helps classify initial datasets and make adequate assumptions e.g. on base of construction years etc. if they are required. The “Energy characteristics” dataset describes parameters of heat, ventilation, and air conditioning systems, and PV systems if they are used. The “Geometrical” dataset includes dimensions of the analyzed district, building orientations, and shading, which in some ways correlate with energy consumption and demand. “Physical” parameters mostly describe buildings' construction materials and are closely related to their energy losses. The "Climate" dataset contains all required data about weather and insulation in the region, which is especially necessary for transient energy balance analysis. All these subsets of data form the so-called “initial” dataset for energy balance calculation.

From the roles point of view, the initial dataset (see Fig. 2) is forming as shown in Fig. 3 and includes two key players: district representatives and experts.

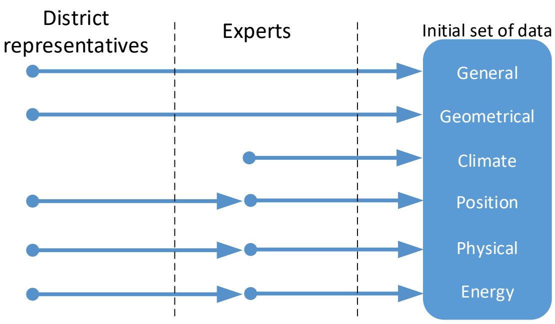

Figure 3. Roles in initial dataset forming

The process of the initial dataset forming always involves district representatives and experts. Some sub-sets could be directly formed by experts (e.g. climate and weather data), while others start on the district representatives’ level and go through experts where additional interpretations are made. It should be noted that for highly digitalized regions involvement of the district representatives is small and increasing with the digitalization level decreasing.

In general, the whole process of building energy balance calculation could be divided into three main stages:

- Focus district characterization.
- Dataset forming.
- Energy balance assessment.

In the first stage "district characterization" the boundaries of the district are defined. It means that energy flows, geographic boundaries, district development goals, key performance indicators, etc. should be defined [11]. This process is quite well described in [12]. It should be noted, that the focus district characterization weakly depends on computer technologies and the whole digitalization level as it requires mostly analytical work from experts with stakeholders’ support.

The second stage “Dataset forming” is the least formalized and practically not described in the open literature. However, it is the quality of preparation of the initial data set that will directly affect the quality of the results obtained in the third stage (e.g. [13]).

Thus, there is a need to formalize the process of preparing initial data for calculating the energy balance, and, as a result, improve the quality and efficiency of the calculation process.

Using the above-described methods and initial dataset there was identified algorithm (process flow) of the dataset formed for buildings energy balance calculation, for regions with high levels of digitalization process flow is shown in Fig. 4.

The process of dataset forming starts after focus district characterization and requires a description of district boundaries [12]. This information is analyzed and requires information to be taken from outer databases (climate/weather (e.g. [14]), energy monitoring/statistical (e.g. [15]), building documentation databases (e.g. [16]) and GIS systems (e.g. [17])). District representatives are involved remotely only if required data cannot be approached directly by experts. In case data cannot be obtained assumptions are made. Before forming the final dataset, all collected information is interpreted to the required format to be transferred to the next task of energy balance calculation.

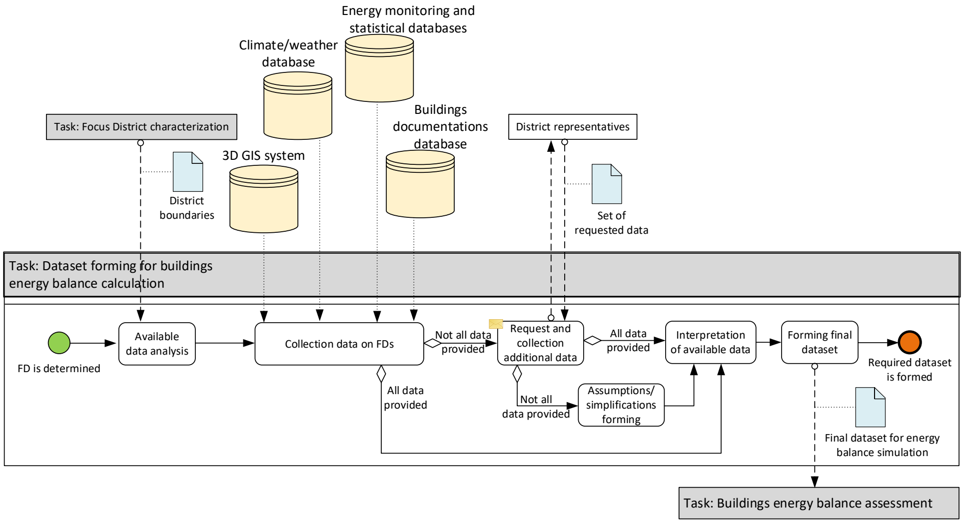

Figure 4. Roles in initial dataset forming

VI. CONCLUSIONS

In conclusion, this paper presents a comprehensive framework for the development of initial datasets required for building energy balance simulations, which could be adapted to different levels of district digitalization. The structured process flow and role identification provide a formalized approach that bridges the often-overlooked gap in the dataset preparation phase. The framework ensures adaptability to various district contexts, improving the accuracy and efficiency of energy balance simulations. The collaboration between district representatives and experts is a critical element, with the involvement of district representatives increasing as digitalization levels decrease. This study highlights the importance of well-prepared datasets in optimizing energy balance calculations, which are essential for the planning and implementation of positive energy districts. The formalization of this process enhances clarity and contributes significantly to the growing body of research focused on energy efficiency and sustainability in urban environments. The proposed framework offers a reliable foundation for future research and practical applications in urban energy management.

REFERENCES

[1] A. J. Marszal, P. Heiselberg, J. S. Bourrelle, E. Musall, K. Voss, I. Sartori, A. Napolitano, “Zero Energy Building – A review of definitions and calculation methodologies”, Energ. Buildings, 43, 2011, pp. 971–979. doi.org/10.1016/j.enbuild.2010.12.022.

[2] L. Vandenbogaerde, S. Verbeke, A. Audenaert, “Optimizing building energy consumption in office buildings: A review of building automation and control systems and factors influencing energy savings”, Journal of Building Engineering, 76, 2023, 107233. doi.org/10.1016/j.jobe.2023.107233.

[3] V. Vaidhyanathan, “Synthetic 3D Building Energy Model (BEM) Dataset Generation for Human+ AI Synergies in Early-Phase High Performance Building Design”, Doctoral dissertation, Carnegie Mellon University, 2021.

[4] D. B. Crawley, L. K. Lawrie, C. O. Pedersen, F. C. Winkelmann, “Energy plus: energy simulation program”, ASHRAE journal, 42(4), 2000, pp. 49-56.

[5] A. Oraiopoulos, B. Howard, “On the accuracy of urban building energy modelling”, Renewable and Sustainable Energy Reviews, 158, 2022.

[6] A. Vartholomaios, A. Chatzidimitriou, K. Ioannidis, “Estimating the influence of building and urban form on the thermal loads of urban dwellings in the Mediterranean climate using machine learning”, Energy Sources, Part B: Economics, Planning, and Policy, 16(8), 2021, pp. 687-706.

[7] M. Labiadh, “Methodology for construction of adaptive models for the simulation of energy consumption in buildings”, Modeling and Simulation. Université de Lyon, 2021.

[8] H. Zhao, “Artificial intelligence models for large scale buildings energy consumption analysis”, Ecole Centrale Paris, 2011.

[9] R. Dijkman, J. Hofstetter, J. Koehler, “Business Process Model and Notation”, Vol. 89, 2011. Berlin, Germany: Springer.

[10] P. Tuominen, R. Holopainen, L. Eskola, J. Jokisalo, M. Airaksinen, “Calculation method and tool for assessing energy consumption in the building stock”, Build. Environ., 75, 2014, pp. 153–160. doi.org/10.1016/j.buildenv.2014.02.001.

[11] X. Zhang, S. R. Penaka, S. Giriraj, M. N. Sánchez, P. Civiero, H. Vandevyvere, “Characterizing Positive Energy District (PED) through a Preliminary Review of 60 Existing Projects in Europe”, Buildings, 11 (8), 2021. doi.org/10.3390/buildings11080318.

[12] S. Schneider, T. Zelger, D. Sengl, J. Baptista, “A Quantitative Positive Energy District Definition with Contextual Targets”, Buildings, 13, 2023. doi.org/10.3390/buildings13051210.

[13] W. Tian “A review of sensitivity analysis methods in building energy analysis”, Renew. Sust. Energ. Rev., 20, 2013, pp. 411–419. doi.org/10.1016/j.rser.2012.12.014.

[14] A. Amin, M. Mourshed “Weather and climate data for energy applications”, Renewable and Sustainable Energy Reviews, 192, 2024. doi.org/10.1016/j.rser.2023.114247.

[15] Z. Tian, X. Zhang, S. Wei, S. Du, X. Shi, “A review of data-driven building performance analysis and design on big on-site building performance data”, Journal of Building Engineering, 41, 2021. doi.org/10.1016/j.jobe.2021.102706.

[16] T. Loga, B. Stein, N. Diefenbach, “TABULA building typologies in 20 European countries—Making energy-related features of residential building stocks comparable”, Energ. Buildings, 132, 2016, pp. 4–12. doi.org/10.1016/j.enbuild.2016.06.094.

[17] X. Liu, X. Wang, G. Wright, J. Cheng, X. Li, R. Liu, “A State-of-the-Art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS)”, ISPRS Int. Geo.-Inf., 6, 2017. doi.org/10.3390/ijgi6020053.