MODELS FOR PREDICTING BUILDING ENERGY PERFORMANCE: REVIEW

Makhbuba Shermatova

Turin Polytechnic University in Tashkent

Abstract. This paper provides a comprehensive overview of the principles, methods, and challenges associated with modeling and predicting building energy performance. It examines various modeling approaches—including physics-based (white-box), statistical (black-box), and hybrid models—and evaluates their strengths and limitations in forecasting energy consumption across different building types and urban contexts. The study highlights the growing role of GIS-integrated urban building energy models (UBEMs) for city-scale analyses and discusses their application in energy retrofitting, load-shifting, and policy planning. Furthermore, the article identifies persistent challenges such as data scarcity, occupant behavior variability, and model generalization issues, emphasizing the need for advanced AI and machine-learning techniques combined with interpretable and physics-informed modeling. The findings underscore the importance of accurate energy prediction models as a foundation for sustainable design, energy policy, and urban decarbonization strategies.

Keywords: Building energy performance, energy Modeling, Urban Building Energy Models (UBEM), energy efficiency, Physics-Based Simulation, Data-Driven Models, GIS Integration, sustainable urban development, energy forecasting.

Аннотация. В данной статье представлен всесторонний обзор принципов, методов и проблем, связанных с моделированием и прогнозированием энергоэффективности зданий. В ней рассматриваются различные подходы к моделированию, включая физические модели (белый ящик), статистические модели (черный ящик) и гибридные модели, а также оцениваются их сильные и слабые стороны при прогнозировании энергопотребления для различных типов зданий и условий городской среды. В исследовании подчеркивается растущая роль интегрированных в ГИС моделей энергопотребления городских зданий (UBEM) для анализа в масштабах города и обсуждается их применение в модернизации энергосистем, перераспределении нагрузки и планировании политики. Кроме того, в статье рассматриваются такие сохраняющиеся проблемы, как дефицит данных, изменчивость поведения жильцов и проблемы обобщения моделей, что подчеркивает необходимость использования передовых методов искусственного интеллекта и машинного обучения в сочетании с интерпретируемым и физически обоснованным моделированием. Результаты подчеркивают важность точных моделей прогнозирования энергопотребления как основы для устойчивого

ISSN 3030-3907

проектирования, энергетической политики и стратегий декарбонизации городов.

Ключевые слова: Энергоэффективность зданий, энергетическое моделирование, энергетические модели городских зданий (UBEM), энергоэффективность, физическое моделирование, модели на основе данных, интеграция ГИС, устойчивое городское развитие, энергетическое прогнозирование.

Annotatsiya. Ushbu maqolada binolarning energiya samaradorligini modellashtirish va hisoblash bilan bog'liq bo'lgan qoidalar, usullar va muammolar bayon etilgan. Ushbu maqolada turli xil modellashtirish yondashuvlari, jumladan, fizikaga asoslangan (oq quti), statistik (qora quti) va gibrid modellar ko'rib chiqiladi va turli xil binolar va shahar miqyosida energiya iste'molini hisob qilishda ularning imkoniyat va cheklovlarini baholaydi. Tadqiqotda shahar miqyosidagi tahlillar uchun GIS bilan integratsiyalangan shahar qurilishi energiya modellarining (UBEMs) ortib borayotgan rolini ko’rsatadi va ularni energiyani rekonstruksiya, yuklamalar o’zgarishi va siyosatni rejalashtirishda qo'llash muhokama qilingan. Bundan tashqari, maqolada ma'lumotlar yetishmovchiligi, bino ekspluatatsiya qilishda ichidagi odamlar xatti-harakatlarining o'zgaruvchanligi va modelni umumlashtirish muammolari kabi doimiy muammolar haqida aytib o’tilgan va ilg'or sun’iy intellekt va Machine learning usullarini izohlash mumkin bo'lgan va fizikadan ma'lumotga ega modellashtirish zarurligini ta'kidlaydi. Tahlillar shuni ko’rsatadiki, barqaror loyihalash, energiya samaradorlik siyosati va shaharlarda CO_2 ta’sirini kamaytirish strategiyalari uchun asos sifatida aniq energiya iste’molini hisoblash modellarining muhimligini ta'kidlaydi.

Kalit so'zlar: Binoning energiya samaradorligi, energiya modellashtirish, shahar qurilishi energiya modellari (UBEM), energiya samaradorligi, fizikaga asoslangan simulyatsiya, ma'lumotlarga asoslangan modellar, GAT(GIS) integratsiyasi, barqaror shahar rivojlanishi, energiya iste’molini hisoblash.

Introduction

The accurate prediction of building energy performance has become a pivotal aspect of sustainable urban development and energy policy formulation. With global energy consumption in the building sector accounting for a significant share of total demand, optimizing performance is vital for reducing greenhouse gas emissions, improving operational efficiency, and ensuring thermal comfort. Building energy performance modeling involves complex interactions among architectural design, material properties, mechanical systems, occupant behavior, and climatic conditions. Consequently, researchers have developed diverse modeling approaches—ranging from simplified statistical regression models to

ISSN 3030-3907

advanced physics-based simulations and hybrid machine-learning frameworks—to quantify and predict energy use patterns at both building and urban scales.

Recent advancements in computational modeling and data analytics have enhanced the precision of such predictions, facilitating the assessment of retrofit measures, load-shifting strategies, and renewable integration. Furthermore, the integration of Geographic Information Systems (GIS) with urban building energy models (UBEMs) allows for spatially explicit analyses of energy flows, enabling planners and policymakers to evaluate the effectiveness of interventions and establish decarbonization strategies across entire city districts. Despite these developments, several challenges persist—particularly the heterogeneity of building stocks, limited availability of high-resolution data, and the "black-box" nature of many artificial intelligence (AI) models—which hinder the transferability and interpretability of results. Addressing these challenges requires unified evaluation frameworks, improved data governance, and interdisciplinary approaches that combine building physics, data science, and urban informatics.
Thus, understanding and improving building energy performance through accurate prediction models is not only a technical endeavor but also a strategic pathway toward achieving resilient, low-carbon, and equitable urban environments.

Background on Building Energy Performance

Accurate prediction of building energy performance is a multifaceted challenge, involving complex interactions between building physics, occupant behavior, and environmental factors [1]. Therefore, sophisticated modeling approaches are essential to accurately forecast energy consumption and optimize building design and operation for improved efficiency. These models range from simplified statistical methods to complex physics-based simulations, each offering distinct advantages depending on the scope and required precision of the analysis. For example, urban building energy models leverage geospatial vector datasets to process geometric data for hundreds to thousands of buildings, allowing for city-scale energy assessments crucial for efficiency, cost, and carbon reduction [2]. This comprehensive approach allows for the identification of optimal retrofitting strategies and the quantification of their potential impacts on urban energy landscapes. Furthermore, understanding building performance is crucial for implementing energy efficiency strategies like retrofits, advanced conversion technologies, and energy load-shifting, especially when considering the spatiotemporal characteristics of building demands in conjunction with demographic data and renewable resources [3].

Moreover, the integration of these models with advanced Geographic Information Systems provides a robust framework for visualizing energy flows and

ISSN 3030-3907

understanding the spatial dimensions of energy demand across urban environments [4]. This allows for a nuanced assessment of energy use patterns, facilitating targeted interventions for energy conservation and sustainable urban development.

A critical aspect of these advanced models is their capacity to integrate diverse data, encompassing building characteristics like age, type, and occupancy, which subsequently inform the implementation of load-shifting strategies within the urban building energy model itself [5]. This allows for the development of targeted urban decarbonization strategies by integrating multi-scale energy performance evaluation within the design development process [6]. This feedback loop between Urban Building Energy Models and geographical information systems further supports iterative testing of load-shifting scenarios, enabling spatiotemporal visualization of UBEM outputs. The application of these models extends to assessing energy demand at the city district level, utilizing statistical models to project energy use from individual buildings across an entire urban area [7]. These models are particularly valuable for urban planners and energy policy makers, enabling them to evaluate the impact of various interventions on reducing the overall energy footprint of a city [8]. The reliability of such urban-scale models in predicting energy demand is often proven through metrics like CVRMSE and MAPE, allowing for the effective testing of various scenarios and the identification of optimal retrofitting strategies [9]. Beyond retrofits, these predictive models also enable the evaluation of urban design policies by comparing various proposed strategies and understanding their community-level implications, thereby linking municipal goals with specific interventions at the district or city scale. Moreover, these urban building energy models can inform individual citizens about estimated energy consumption behaviors, encouraging reductions in electricity bills and contributing to broader community-level energy efficiency targets.

Urban building energy modeling represents a practical approach for stakeholders in the energy industry to predict energy use within the building sector under diverse design and retrofit scenarios, particularly for large-scale applications [10]. These models play a crucial role in understanding and mitigating the energy impact of urban environments by offering insights into consumption patterns and the efficacy of various energy-saving strategies [11] [12].

Importance of Energy Performance Prediction

Accurate energy performance prediction is fundamental for developing effective energy policies and optimizing resource allocation in urban planning, enabling decision-makers to anticipate future energy demands and proactively implement sustainable solutions [13]. This predictive capability allows for more informed investments in renewable energy infrastructure, grid modernization, and

ISSN 3030-3907

energy efficiency programs, ultimately contributing to a more resilient and sustainable urban energy system [14]. Moreover, these models provide a critical foundation for analyzing granular-level building energy performance, which assists local authorities, city planners, and energy policymakers in formulating improved sustainable energy policy decisions [15]. Such detailed analysis is crucial for identifying disparities in energy consumption and informing targeted interventions to promote energy equity across diverse neighborhoods [16].

Furthermore, this foresight facilitates the implementation of advanced energy management strategies, including intelligent control systems and demand-side management, which depend on precise estimations of energy consumption to optimize energy use and reduce waste [17]. The ability to accurately predict building energy consumption is also vital for the successful integration of renewable energy sources into the urban grid, allowing for better balancing of supply and demand.

Challenges in Energy Performance Modeling

Despite the advancements, challenges persist in developing robust and universally applicable models, primarily due to the inherent complexities of building physics and occupant behavior. These complexities are further compounded by the heterogeneity of urban building stocks, varying climatic conditions, and the dynamic nature of energy markets, all of which introduce significant uncertainties into prediction models. Moreover, the lack of standardized, high-resolution datasets across different urban contexts hinders the development of generalizable models, often necessitating problem and data-specific algorithmic approaches [18]. The integration of diverse data sources, such as real-time occupancy statistics derived from mobile phone data, can address these limitations by offering a more precise understanding of building energy use patterns [19]. This multi-layered approach enables a more accurate calibration of urban models, particularly when actual energy consumption data is scarce or unavailable, thereby allowing for detailed energy efficiency analyses and scenario planning [20].

Additionally, addressing data scarcity for bottom-up building energy models through innovative methods, such as disaggregating census tract data to localized buildings, can enhance the realism and precision of urban energy demand profiles. Furthermore, advanced machine learning techniques, including neural networks, offer enhanced predictive capabilities, especially when uniform building data is available, outperforming more traditional methods like KNN for certain data distributions. These advanced models are particularly effective in characterizing urban energy models more accurately, capturing feature variability across diverse neighborhoods, and ensuring equity in energy analysis and the implementation of

ISSN 3030-3907

new technologies [16]. However, the inherent complexities of decentralized energy systems, coupled with intermittent renewable energy generation and dynamic demand-side management, necessitate continuous advancements in modeling approaches to accurately capture system interactions and future scenarios [21]. This continuous evolution of modeling techniques is essential to overcome limitations such as data scarcity and the inherent variability of building characteristics and occupant behaviors, which often lead to discrepancies between simulated and actual energy consumption [22] [18].

Moreover, the increasing amount of macro-level energy data presents opportunities for developing secure management platforms by utilities and authorities, which necessitates robust anonymization techniques to handle privacy concerns and facilitate data distribution effectively [23].

Traditional Energy Performance Models

These models typically involve physics-based simulations or statistical approaches, each with distinct advantages and limitations regarding computational intensity and data requirements. Physics-based models, often referred to as white-box models, rely on detailed building characteristics and environmental parameters to simulate energy flows, providing high fidelity but demanding extensive input data and computational resources. Conversely, statistical models, or black-box models, leverage historical energy consumption data to identify correlations and predict future performance, requiring less detailed physical input but often lacking the explanatory power of physics-based approaches [24].

Hybrid models aim to combine the strengths of both approaches, integrating physical principles with data-driven techniques to enhance accuracy and reduce computational burden. Nevertheless, even these sophisticated hybrid models often struggle to capture the significant variability in energy demand across different buildings and neighborhoods, which can lead to misrepresentation of energy peaks and suboptimal planning. Therefore, a crucial direction for future research involves evaluating the efficacy of these sophisticated statistical models in forecasting actual building energy performance [25]. This assessment necessitates a careful consideration of the models' ability to generalize across diverse building types and operational conditions, especially when data quality and availability are limited [26]. This evaluation is particularly challenging given the disparate nature of existing datasets, which often vary in granularity and explanatory variables, thereby complicating direct comparisons of model performance [27]. Further research is needed to develop unified metrics and benchmarks that can reliably assess the predictive accuracy and robustness of various modeling approaches under diverse real-world conditions.

Physics-Based Simulation Models

ISSN 3030-3907

These models, also known as white-box models, are constructed upon fundamental thermodynamic principles and heat transfer equations to simulate the dynamic energy performance of a building. They require comprehensive inputs, including detailed building geometry, material properties, HVAC system specifications, and occupancy schedules, to generate highly granular predictions of energy consumption and indoor environmental conditions [28]. The accuracy of these models is heavily dependent on the precision and completeness of the input data, often necessitating extensive calibration against measured performance to reduce discrepancies between simulated and actual energy use. Despite their high fidelity, the computational intensity and extensive data requirements often limit their application to individual buildings or small clusters, posing challenges for widespread urban-scale energy analysis [29]. Conversely, reduced-order models offer a more pragmatic alternative, simplifying the physics-based simulations by utilizing normative parameters to quickly assess building energy performance with reduced data demands [30]. However, these simplified models may sacrifice some predictive accuracy, especially in complex building configurations or when specific operational nuances are critical to energy performance. Therefore, the selection of an appropriate modeling strategy, whether detailed physics-based or simplified reduced-order, must balance the need for accuracy with the available computational resources and data accessibility. Furthermore, the accuracy of physics-based models is highly contingent on accurate parameter settings related to building thermal dynamics, which can be difficult to obtain and can lead to significant variations in simulated outcomes [31]. These models typically require expert knowledge and significant time for development and validation, further complicated by the need for individual models for each unique building [32].

Statistical Regression Models

These models, often categorized as black-box models, diverge significantly from white-box approaches by utilizing historical energy consumption data to predict future performance without explicit consideration of physical building parameters [33]. Instead, they identify statistical relationships between energy use and various independent variables, such as weather data, occupancy patterns, and building characteristics [34]. Multiple regression, artificial neural networks, and decision trees represent common statistical techniques employed for predicting electrical energy consumption. These models excel at rapid predictions and can achieve high accuracy when trained on sufficiently diverse and representative datasets, but their validity is often constrained to the specific context of their training data, making generalization to different building types or operational conditions challenging [35]. A significant amount of data is typically required for these statistical techniques, with hourly or daily resolutions generally preferred

ISSN 3030-3907

over monthly aggregates for better predictive power. Despite their ability to quickly generate predictive energy consumption data, these models often lack transparency in their decision-making processes [36], making it difficult to interpret the underlying relationships between variables. This characteristic, often termed the "black-box" nature, can hinder efforts to identify specific energy-saving opportunities, as the models prioritize predictive accuracy over interpretable insights [37].

Data Requirements and Preprocessing

The effectiveness of any predictive model, regardless of its architectural complexity, is profoundly influenced by the quality, quantity, and relevance of the input data. Inadequate or noisy data can lead to biased models with poor generalization capabilities, irrespective of the sophistication of the algorithms employed. Therefore, rigorous data preprocessing, including cleaning, normalization, and feature engineering, is paramount to ensure the integrity and utility of the dataset for training robust energy performance prediction models. This process often involves addressing missing values through imputation techniques, handling outliers to prevent skewed model training, and scaling features to ensure uniform influence during the learning process [38].

Furthermore, feature engineering, which involves creating new variables from existing ones to better represent underlying relationships, can significantly enhance model performance, particularly in capturing complex temporal and contextual patterns in building energy consumption data. The selection of appropriate features and the rigor of the preprocessing pipeline are critical, as they directly impact a model's ability to discern meaningful patterns from raw data, especially when dealing with the high dimensionality and temporal dependencies inherent in building operational time-series data [39]. The nascent field of research into how various data types influence prediction methods indicates that current efforts are largely concentrated on implementing AI methods, which have been likened to "black box" models due to the complexities associated with deciphering their internal algorithmic processes.

Case Studies and Applications

This section delves into real-world implementations and empirical findings of energy prediction models, highlighting their practical utility and performance across diverse building types and operational contexts. These case studies often demonstrate the advantages and limitations of different modeling approaches, providing empirical evidence for their applicability in informing energy management strategies and design decisions [40].

For instance, several studies have successfully applied machine learning models to predict energy consumption in commercial buildings, demonstrating

ISSN 3030-3907

superior accuracy compared to traditional statistical methods. This improved accuracy often stems from the ability of advanced supervised learning techniques like artificial neural networks and ensemble methods to capture the non-linear relationships inherent in building energy consumption patterns [41]. However, these models frequently necessitate extensive, well-curated datasets for training and validation, underscoring the critical role of robust data preprocessing to maximize their predictive efficacy.

For residential buildings, long-term forecasts of energy usage, particularly concerning lighting patterns, require more focused attention in data-driven models. Furthermore, the duration and type of data used for model training significantly impact performance, with studies predominantly utilizing real-world datasets collected over periods ranging from one month to one year, although some leverage simulated or benchmark data.

Moreover, commercial and educational buildings are frequently targeted for predictive modeling due to the abundant historical energy consumption data available, which simplifies the application of artificial intelligence, statistical, and hybrid methods. The predictive accuracy of these models is significantly influenced by various external factors, including meteorological conditions, occupant behavior, and building operational schedules [42]. The integration of diverse data streams, such as real-time occupancy sensing and granular weather forecasts, further enhances the robustness and predictive power of these models, moving beyond reliance on historical consumption alone [43].

Limitations and Future Directions

Despite the advancements in predictive modeling for building energy performance, several challenges persist, including the "black box" nature of complex AI models, the intensive data requirements, and the difficulty in generalizing models across diverse building typologies and climates [44] [45]. These limitations highlight the ongoing need for more interpretable models, robust data collection methodologies, and transfer learning techniques that can adapt pre-trained models to new environments with minimal additional data. A significant gap also exists in the diversity of data collection case studies, with mixed-use buildings and older structures being notably underrepresented, despite their prevalence. Furthermore, the development of models capable of accurately predicting energy consumption in post-retrofit buildings is significantly hampered by the scarcity of available performance data following such interventions, necessitating novel approaches to data generation and model training [46].

Conclusion

In conclusion, the field of building energy performance prediction has seen substantial progress through the application of advanced computational methods,

ISSN 3030-3907

particularly artificial intelligence and statistical models. These models offer significant potential for optimizing energy efficiency, reducing operational costs, and supporting sustainable building practices across various sectors. However, challenges related to data availability, model interpretability, and generalization across diverse building types remain, necessitating continued research into hybrid approaches that combine physics-based simulations with AI techniques [47].

REFERENCES

1. J. García-López, M. Hernández-Valencia, J. Roa-Fernández, E. J. Mascort-Albea, and R. Herrera-Limones, "Balancing construction and operational carbon emissions: Evaluating neighbourhood renovation strategies," *Journal of Building Engineering* , vol. 94, p. 109993, Jun. 2024, doi: 10.1016/j.jobe.2024.109993.

2. T. Santhanavanich *et al.* , "ENABLING INTEROPERABILITY OF URBAN BUILDING ENERGY DATA BASED ON OGC API STANDARDS AND CITYGML 3D CITY MODELS," *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences* , p. 97, Dec. 2023, doi: 10.5194/isprs-annals-x-1-w1-2023-97-2023.

3. E. Naboni *et al.* , "Parametric Workflow to Conceive Facades as Indoor and Outdoor Climate Givers," *Research Portal Denmark* , p. 11, Jan. 2019, Accessed: Jul. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=arkc&id=arkc-7851f7a0-5cf1-4777-9c00-789161a4fbe5&ti=Parametric%20Workflow%20to%20Conceive%20Facades%20as%20Indoor%20and%20Outdoor%20Climate%20Givers

4. A. Nowacka and F. Remondino, "GEOSPATIAL DATA FOR ENERGY EFFICIENCY AND LOW CARBON CITIES – OVERVIEW, EXPERIENCES AND NEW PERSPECTIVES –," *The international archives of the photogrammetry, remote sensing and spatial information sciences/International archives of the photogrammetry, remote sensing and spatial information sciences* , p. 467, Sep. 2018, doi: 10.5194/isprs-archives-xlii-4-467-2018.

5. A. Kamari *et al.* , "Unleashing the diversity of conceptual building renovation design: Integrating high-fidelity simulation with rapid constraint-based scenario generation," *Research Portal Denmark* , vol. 51, p. 29, Jan. 2019, Accessed: Jul. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=au&id=au-1e73a96c-9f1d-4fe5-923b-49c632d380d2&ti=Unleashing%20the%20diversity%20of%20conceptual%20building%20renovation%20design%3A%20Integrating%20high-fidelity%20simulation%20with%20rapid%20constraint-based%20scenario%20generation

6. A. G. di Stefano, M. F. Ruta, and G. Masera, "Advanced Digital Tools for Data-Informed and Performance-Driven Design: A Review of Building Energy

ISSN 3030-3907

Consumption Forecasting Models Based on Machine Learning," *Applied Sciences* , vol. 13, no. 24. Multidisciplinary Digital Publishing Institute, p. 12981, Dec. 05, 2023. doi: 10.3390/app132412981.

7. M. Hosseingholizadeh, V. Coors, H. Ostadabbas, and F. Friesecke, "DEVELOPMENT OF QGIS PLUGIN FOR URBAN ENERGY SIMULATION USING 3D CITY MODEL AT THE CITY DISTRICT LEVEL," *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences* , p. 81, Dec. 2023, doi: 10.5194/isprs-annals-x-1-w1-2023-81-2023.

8. S. T. Moghadam, J. Toniolo, G. Mutani, and P. Lombardi, "A GIS-statistical approach for assessing built environment energy use at urban scale," *Sustainable Cities and Society* , vol. 37, p. 70, Nov. 2017, doi: 10.1016/j.scs.2017.10.002.

9. M. Gholami, D. Torreggiani, P. Tassinari, and A. Barbaresi, "Narrowing uncertainties in forecasting urban building energy demand through an optimal archetyping method," *Renewable and Sustainable Energy Reviews* , vol. 148, p. 111312, Jun. 2021, doi: 10.1016/j.rser.2021.111312.

10. E. Kamel, "A Systematic Literature Review of Physics-Based Urban Building Energy Modeling (UBEM) Tools, Data Sources, and Challenges for Energy Conservation," *Energies* , vol. 15, no. 22, p. 8649, Nov. 2022, doi: 10.3390/en15228649.

11. I. Dochev *et al.* , "Calculating urban heat demands: An analysis of two modelling approaches and remote sensing for input data and validation," *Energy and Buildings* , vol. 226, p. 110378, Aug. 2020, doi: 10.1016/j.enbuild.2020.110378.

12. S. Choi and S. Yoon, "AI Agent-Based Intelligent Urban Digital Twin (I-UDT): Concept, Methodology, and Case Studies," *Smart Cities* , vol. 8, no. 1, p. 28, Feb. 2025, doi: 10.3390/smartcities8010028.

13. J. Rodríguez-Álvarez, "Urban Building Energy Modelling for the Renovation Wave: A Bespoke Approach Based on EPC Databases," *Buildings* , vol. 13, no. 7, p. 1636, Jun. 2023, doi: 10.3390/buildings13071636.

14. F. Yang and Q. Mao, "Auto-Evaluation Model for the Prediction of Building Energy Consumption That Combines Modified Kalman Filtering and Long Short-Term Memory," *Sustainability* , vol. 15, no. 22, p. 15749, Nov. 2023, doi: 10.3390/su152215749.

15. U. Ali, M. H. Shamsi, C. Hoare, E. Mangina, and J. O'Donnell, "A data-driven approach for multi-scale building archetypes development," *Energy and Buildings* , vol. 202, p. 109364, Aug. 2019, doi: 10.1016/j.enbuild.2019.109364.

16. R. E. Kontar *et al.* , "An Open-Source Framework for Characterizing Urban Energy Models: Integrating Top-Down and Bottom-Up Methods to Predict Residential Buildings Characteristics: Preprint," *OSTI OAI (U.S. Department of*

ISSN 3030-3907

Energy Office of Scientific and Technical Information) , Aug. 2024, Accessed: Aug. 2025. [Online]. Available: https://www.osti.gov/biblio/2438558

17. K. Chen and W. Zhu, "The Utilization of a Multi-Layer Perceptron Model for Estimation of the Heating Load," *International Journal of Advanced Computer Science and Applications* , vol. 15, no. 6, Jan. 2024, doi: 10.14569/ijacsa.2024.0150614.

18. W. Khan, S. Walker, and W. Zeiler, "A bottom-up framework for analysing city-scale energy data using high dimension reduction techniques," *Sustainable Cities and Society* , vol. 89, p. 104323, Nov. 2022, doi: 10.1016/j.scs.2022.104323.

19. E. Barbour, C. C. Davila, S. Gupta, C. Reinhart, J. Kaur, and M. C. González, "Planning for sustainable cities by estimating building occupancy with mobile phones," *Nature Communications* , vol. 10, no. 1, Aug. 2019, doi: 10.1038/s41467-019-11685-w.

20. V. Costanzo, R. Yao, X. Li, M. Liu, and B. Li, "A multi-layer approach for estimating the energy use intensity on an urban scale," *Cities* , vol. 95, p. 102467, Nov. 2019, doi: 10.1016/j.cities.2019.102467.

21. M. Yazdanie and K. Orehounig, "Advancing urban energy system planning and modeling approaches: Gaps and solutions in perspective," *Renewable and Sustainable Energy Reviews* , vol. 137, p. 110607, Dec. 2020, doi: 10.1016/j.rser.2020.110607.

22. G. Tardioli, A. Narayan, R. Kerrigan, M. Oates, J. O'Donnell, and D. Finn, "A methodology for calibration of building energy models at district scale using clustering and surrogate techniques," *Energy and Buildings* , vol. 226, p. 110309, Jul. 2020, doi: 10.1016/j.enbuild.2020.110309.

23. P. Gianniou, "Energy demand models for buildings in a smart cities context," *Research Portal Denmark* , no. 390, p. 216, Jan. 2018, Accessed: Jul. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=dtu&id=dtu-49fffdbd-563b-4986-841c-0d6bacf2668e&ti=Energy%20demand%20models%20for%20buildings%20in%20a%20smart%20cities%20context

24. C. Koulamas, A. Καλογεράς, R. Pacheco-Torres, J. Casillas, and L. Ferrarini, "Suitability analysis of modeling and assessment approaches in energy efficiency in buildings," *Energy and Buildings* , vol. 158, p. 1662, Dec. 2017, doi: 10.1016/j.enbuild.2017.12.002.

25. L. Yue, X. Liu, and S. Chang, "Appraising the Building Cooling Load via Hybrid Framework of Machine Learning Techniques," *International Journal of Advanced Computer Science and Applications* , vol. 15, no. 6, Jan. 2024, doi: 10.14569/ijacsa.2024.01506137.

ISSN 3030-3907

26. V. Sánchez and T. Gómez-Acebo, “Building energy performance metamodels for district energy management optimisation platforms,” *Energy Conversion and Management X* , vol. 21, p. 100512, Dec. 2023, doi: 10.1016/j.ecmx.2023.100512.
27. B. Grillone, S. Danov, A. Sumper, J. Cipriano, and G. Mor, “A review of deterministic and data-driven methods to quantify energy efficiency savings and to predict retrofitting scenarios in buildings,” *Renewable and Sustainable Energy Reviews* , vol. 131. Elsevier BV, p. 110027, Jul. 10, 2020. doi: 10.1016/j.rser.2020.110027.
28. M. Jradi and J. Bjørnskov, “A Proposed Methodology Towards Adaptable and Effective Building Digital Twins : Energy Models and Data Requirements, DanRETwin Project Report, October 2024.,” Technical University of Denmark, Jan. 2024. Accessed: Jul. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=sdu&id=sdu-1fe80316-5929-4b65-823d-03595836fb18&ti=A%20Proposed%20Methodology%20Towards%20Adaptable%20and%20Effective%20Building%20Digital%20Twins%20%3A%20Energy%20Models%20and%20Data%20Requirements%2C%20DanRETwin%20Project%20Report%2C%20October%202024 .
29. L. V. Krannichfeldt, K. Orehounig, and O. Fink, “Combining physics-based and data-driven modeling for building energy systems,” *Applied Energy* , vol. 391, p. 125853, Apr. 2025, doi: 10.1016/j.apenergy.2025.125853.
30. C. Zhu, X. Ye, J. Du, Z. Hu, Y. Shen, and D. Retchless, “Simulating urban energy use under climate change scenarios and retrofit plans in coastal Texas,” *Urban Informatics* , vol. 3, no. 1, Jun. 2024, doi: 10.1007/s44212-024-00046-8.
31. R. Yan, T. Zhao, Y. Rezgui, S. Kubicki, and L. Yu, “Transferability and robustness of a data-driven model built on a large number of buildings,” *Journal of Building Engineering* , vol. 80, p. 108127, Nov. 2023, doi: 10.1016/j.jobe.2023.108127.
32. A. Prabowo *et al.* , “BTS: Building Timeseries Dataset: Empowering Large-Scale Building Analytics,” *arXiv (Cornell University)* , Jun. 2024, doi: 10.48550/arxiv.2406.08990.
33. G. Chaudhary, H. Johra, L. Georges, and B. Austbø, “Synconn_build: A python based synthetic dataset generator for testing and validating control-oriented neural networks for building dynamics prediction,” *MethodsX* , vol. 11, p. 102464, Oct. 2023, doi: 10.1016/j.mex.2023.102464.
34. F. Pallonetto, M. Д. Poca, F. D’Ettorre, and D. Finn, “On the assessment and control optimisation of demand response programs in residential buildings,”

Renewable and Sustainable Energy Reviews , vol. 127, p. 109861, Apr. 2020, doi: 10.1016/j.rser.2020.109861.

35. I. Staffell, S. Pfenninger, and N. Johnson, "A global model of hourly space heating and cooling demand at multiple spatial scales," *Nature Energy* , vol. 8, no. 12, p. 1328, Sep. 2023, doi: 10.1038/s41560-023-01341-5.

36. Q. Qiao, A. Yunusa-Kaltungo, and R. E. Edwards, "Towards developing a systematic knowledge trend for building energy consumption prediction," *Journal of Building Engineering* , vol. 35, p. 101967, Nov. 2020, doi: 10.1016/j.jobe.2020.101967.

37. C. Zhang, L. Ma, Z. Luo, H. Xing, and T. Zhao, "Forecasting building plug load electricity consumption employing occupant-building interaction input features and bidirectional LSTM with improved swarm intelligent algorithms," *Energy* , vol. 288, p. 129651, Nov. 2023, doi: 10.1016/j.energy.2023.129651.

38. E. Khaoula, A. Baïna, M. Bellafkih, and A. F. Deifalla, "Machine Learning Algorithms for Predicting Energy Consumption in Educational Buildings," *International Journal of Energy Research* , vol. 2024, p. 1, May 2024, doi: 10.1155/2024/6812425.

39. Y. Meng, Y. Sun, S. Rodríguez, and B. Xue, "Transforming Building Energy Management: Sparse, Interpretable, and Transparent Hybrid Machine Learning for Probabilistic Classification and Predictive Energy Modelling," *Architecture* , vol. 5, no. 2, p. 24, Mar. 2025, doi: 10.3390/architecture5020024.

40. H. Kiavarz, M. Jadidi, and P. Esmaili, "A graph-based explanatory model for room-based energy efficiency analysis based on BIM data," *Frontiers in Built Environment* , vol. 9, Sep. 2023, doi: 10.3389/fbuil.2023.1256921.

41. C. Fan, Y. Sun, Y. Zhao, M. Song, and J. Wang, "Deep learning-based feature engineering methods for improved building energy prediction," *Applied Energy* , vol. 240, p. 35, Feb. 2019, doi: 10.1016/j.apenergy.2019.02.052.

42. M. U. Danish, M. SureshKumar, T. Fonseka, U. Uthayakumar, and V. Galwaduge, "An Investigation into Seasonal Variations in Energy Forecasting for Student Residences," *Research Square (Research Square)* , Feb. 2025, doi: 10.21203/rs.3.rs-5632243/v2.

43. A. Mishra, H. R. Lone, and A. Mishra, "DECODE: Data-driven energy consumption prediction leveraging historical data and environmental factors in buildings," *Energy and Buildings* , vol. 307, p. 113950, Jan. 2024, doi: 10.1016/j.enbuild.2024.113950.

44. B. A. Salami, S. I. Abba, A. A. Adewumi, U. A. Dodo, G. K. Otukogbe, and L. O. Oyedele, "Building energy loads prediction using bayesian-based metaheuristic optimized-explainable tree-based model," *Case Studies in Construction Materials* , vol. 19, Nov. 2023, doi: 10.1016/j.cscm.2023.e02676.

ISSN 3030-3907

45. R. Tariq, A. Mohammed, A. Alshibani, and M. Soledad, “Complex artificial intelligence models for energy sustainability in educational buildings,” *Scientific Reports* , vol. 14, no. 1, Jul. 2024, doi: 10.1038/s41598-024-65727-5.
46. [46] X. Yang *et al.* , “Energy-saving potential prediction models for large-scale building: A state-of-the-art review,” *Renewable and Sustainable Energy Reviews* , vol. 156. Elsevier BV, p. 111992, Dec. 17, 2021. doi: 10.1016/j.rser.2021.111992.
47. H. Wefki, R. Khallaf, and A. M. Ebid, “Estimating the energy consumption for residential buildings in semiarid and arid desert climate using artificial intelligence,” *Scientific Reports* , vol. 14, no. 1, Jun. 2024, doi: 10.1038/s41598-024-63843-w.

ISSN 3030-3907