Machine Learning-Powered Prediction Framework for Household Heating Demand

[1]Meng Zhang, [2]Zhihui Li, [3*]Zhibin Yu
[1]*Mechanical and Aerospace Engineering, School of Engineering, University of Liverpool, Liverpool, L69 7ZX, UK*
[2]*Mechanical and Aerospace Engineering, School of Engineering, University of Liverpool, Liverpool, L69 7ZX, UK*
[3*] *Mechanical and Aerospace Engineering, School of Engineering, University of Liverpool, Liverpool, L69 7ZX, UK*

Keywords: Energy demand in buildings, Heating demand prediction, Machine learning

Abstract. This study presents a framework for predicting building heating demand by integrating building simulation with two types of time-series machine learning models, i.e., Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). *DesignBuilder* is used to simulate energy performance, generating a dataset under various operational schedules and weather conditions. Eight heating schedules are analysed to represent diverse occupant behaviours, comfort levels, and operational settings. These simulation data are then collected to train the LSTM and GRU models. Both models indicate strong predictive performance in capturing temporal patterns of building heating demand. LSTM achieved R^2 values of 95.90% on testing data, with RMSE of 3.92%. GRU delivered competitive results, with R^2 values of 93.75% on testing data, with RMSE of 4.83%. While LSTM demonstrated superior accuracy and generalization, GRU required shorter training times, offering a trade-off between prediction performance and computational efficiency. This study highlights the potential of combining simulation and machine learning techniques to accurately predict building heating demand under varying scenarios.

1. Introduction

Climate problems caused by global warming are becoming increasingly critical, prompting countries to implement measures to limit greenhouse gas emissions [1]. Building operations contribute to 36% of global final energy consumption and 39% of energy-related greenhouse gas emissions [2]. The emissions resulting from inefficient buildings represent a significant issue that cannot be overlooked. Consequently, substantial research efforts have focused on improving energy efficiency in buildings, with a particular emphasis on accurately forecasting heating and cooling energy demand. Numerous factors influence a building's energy demand, including its structure, materials, HVAC systems, and occupant behaviour. Traditional forecasting methods consider all these factors and calculate a building's heating or cooling demand using complex heat balance principles, which is physical model. Another widely used approach is the data-driven model, which mainly relies on historical data rather than physical parameters.

Currently, physical models for building energy consumption calculations have become a well-established approach, leading to the development of various simulation tools. These tools are built on fixed physical principles, such as heat transfer, airflow, and light simulation, and are widely used for energy analysis. Since building energy calculations involve highly complex systems, such as envelope heat transfer, indoor airflow, and HVAC performance, and because these physical principles are already mature and well-defined, researchers often choose to build models using existing simulation tools rather than creating entirely new ones [3], [4]. Different tools focus on different aspects, which leads to variations in their solving algorithms. For example, *EnergyPlus* is more focused on the overall thermal balance of buildings, while *TRNSYS* emphasizes the flexibility of simulating individual energy module [5]. As a result, using energy consumption predictions to refine and validate these simulation tools remains an important area of research [6], [7]. Another research focus examines the impact of input parameters on the predicted energy demand of physical models. The sensitivity and uncertainty of energy consumption predictions to factors such as climate conditions, material properties, and usage patterns are evaluated by researchers through adjustments to these inputs [8]. In addition, physical models built on these simulation tools are widely applied for various purposes. Some are used for early-stage energy estimation during design [9], while others are focused on predicting outcomes for residential energy optimisation [10]. Overall, physical models are known for their high resolution and accuracy, making them a traditional yet reliable approach. However, in practical applications, the

complexity of physical characteristics often makes these models time-consuming. This challenge is reflected both in the collection of building environmental data and in the computational time required for simulations.
Unlike physical models, which usually take hours or days, data-driven models provide prompt extrapolation predictions based on training data. These models learn the relationship between inputs and outputs from historical data and adjust until the error is reduced to an acceptable level. Common models for building energy prediction include linear regression, support vector machines (SVM), and multi-layer perceptron (MLP) [11], [12], [13]. Linear regression assumes a linear relationship between input and output variables, is computationally simple but has limited capability in fitting nonlinear relationships and works best with a small number of features [14]. Regression models have low computational costs and are suitable for simple linear and weakly nonlinear building energy predictions. SVM handle nonlinear relationships through kernel functions and are commonly used for small-sample predictions but have imitations in handling highly complex feature interactions [15]. MLP possesses strong nonlinear modelling capabilities and optimises through multiple layers of neurons and weight adjustments. Gradients are backpropagated to automatically learn complex nonlinear relationships between input features and outputs, making MLP suitable for high-precision predictions on large-scale complex datasets. The basic type is the backpropagation neural network (BPNN) [16], which is not well-suited for handling time series data that are prevalent in heating demand prediction. Long Short-Term Memory (LSTM) networks are specifically designed for time series data, enabling long and short-term storage and utilisation of historical information, making them ideal for long-period dependencies in building energy prediction [17]. Gated Recurrent Unit (GRU), with a simpler structure and faster training speed, performs well in short-term time series modelling and is suitable for scenarios with limited computational resources [18].
Although data-driven models do not have the computational burden of physics-based models, they are highly dependent on the quality of historical data. When historical data is insufficient, achieving high-resolution predictions becomes challenging. To handle these challenges and combine the above complementary advantages, this study proposes a hybrid method to integrate physics-based and data-driven models for fast, accurate and high-resolution prediction. First, a physics-based model is developed by using building simulation software *DesignBuilder*. The heat demand generated by the physics-based model is then used to train the data-driven model. Finally, the trained model is applied to predict the building's heating demand, and the prediction results are discussed at the end.

2. Modelling

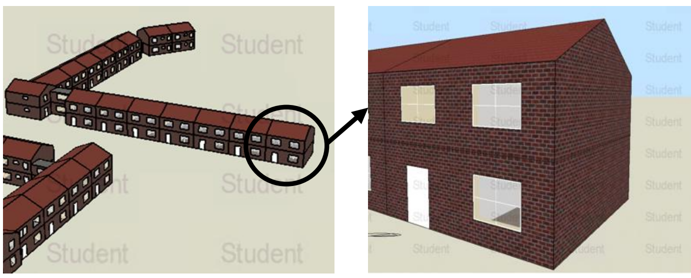

Figure 1 Physical model in *DesignBuilder*

In this study, an end-terraced house in a residential area of Manchester is used as a case study. The model constructed in *DesignBuilder*, based on fundamental building information, is shown in figure-1. The house was built between 1955 and 1964, with a total area of 94 square metres, divided into two floors. In the model, the insulation level is estimated based on the age of the building and it determines the U-values of different part of building. They are set as follows: 1.5 W/m²K for the walls, 0.75 W/m²K for the roof, 3.3 W/m²K for the glazing, and 0.84 W/m²K for the ground floor.

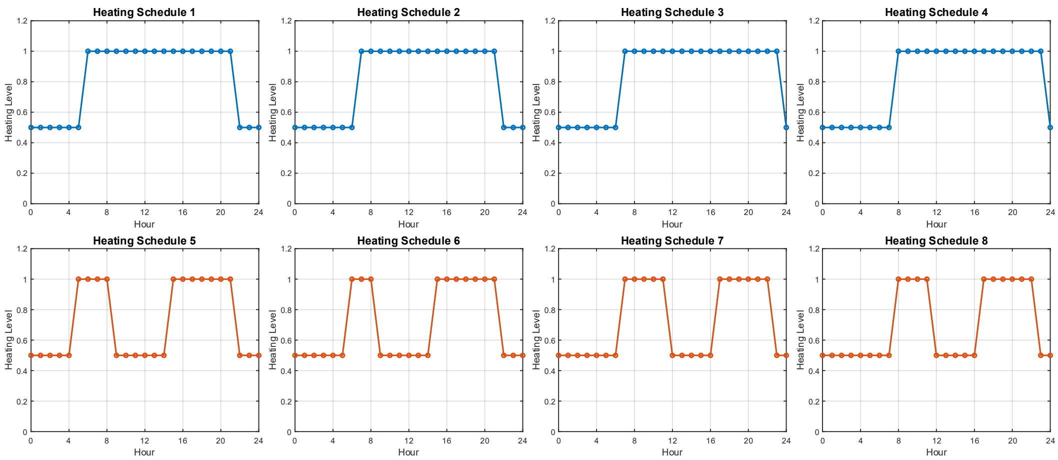

Figure 2 Eight different heating schedule for house

In addition to building information, the model requires local meteorological data as input. Historical weather data from 2020 is used in this study. A simplified HVAC system is implemented, where the heating setpoint temperature is set to 18°C and the setback temperature to 12°C. The model incorporates eight different heating schedules, as is shown in figure 2. When the indoor temperature falls below 18°C, a value of 1 indicates that the heating system maintains the room temperature at 18°C, while a value of 0.5 signifies that the temperature is maintained at 12°C during that period.

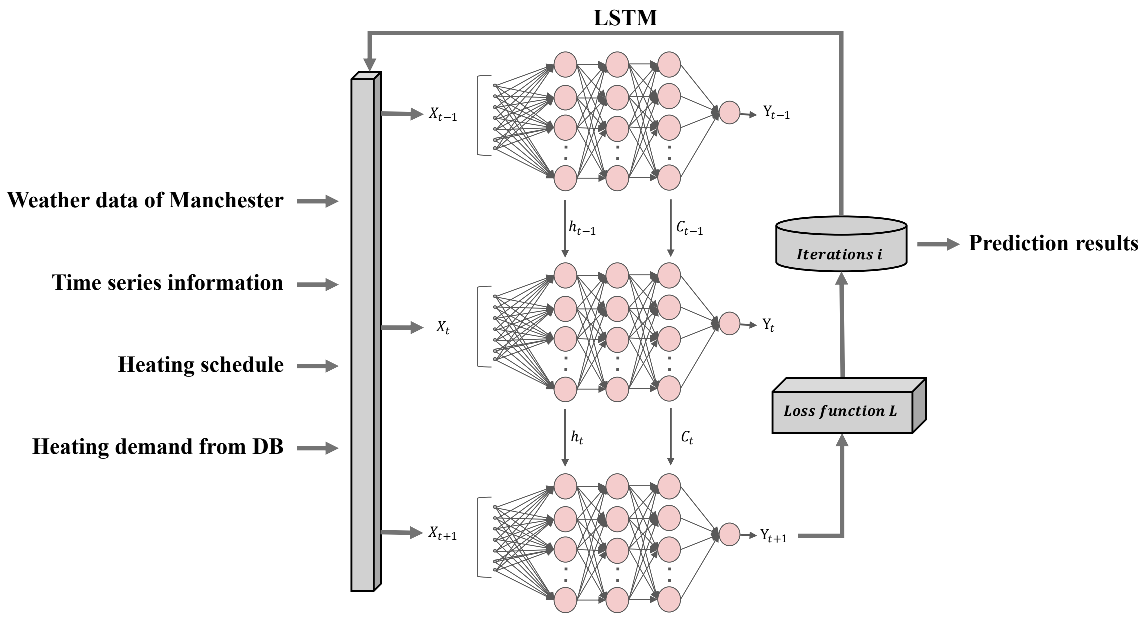

Figure 3 Block diagram of LSTM in the paper

Using LSTM as an example, figure 3 illustrates the process of making predictions with LSTM model. In this framework, the LSTM network takes three types of inputs: (1) weather data in Manchester, (2) time series information, and (3) heating schedule. The weather data includes dry bulb temperature, wet bulb temperature, humidity, global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI). The time series information consists of hour, week, and month, which help capture temporal patterns. These inputs, along with the heating schedule, are sequentially fed into the network to model temporal dependencies. The output of the model is the hourly heating demand of the house. The dataset is split based on time, with the first 80% of the time series used for training and the last 20% for testing. At each time step t, given the input x_t and the previous hidden state h_{t-1} and memory cell C_{t-1}, the LSTM network performs forward propagation to update the memory cell C_t and generate a new hidden state h_t. Here, x_t serves as the input, C_t retains long-term dependencies, and h_t is the output that represents the short-term dependency and is passed to the next layer or used for final predictions. Based on these states, the network produces the corresponding prediction Y_t. Subsequently, the loss function L calculates the error between the predicted and actual values. This error is backpropagated through the network to adjust the weights and bias parameters. During this process, the optimisation algorithm *Adam* is employed to update the gradients iteratively, gradually reducing the prediction error and enhancing model performance.

3. Results and analysis

The scatter plot in figure 4 presents the relationship between true values and predicted values for both GRU (blue) and LSTM (orange) models in the testing set. The $y = x$ reference line represents the ideal prediction scenario where predicted values perfectly match the true values. From the distribution of data points, it is evident that both models generally follow the reference line, indicating good predictive performance. However, GRU exhibits a slightly wider spread of points, suggesting higher variability in predictions. This observation aligns with the RMSE values, where LSTM achieves a lower RMSE of 3.92%, while GRU has a slightly higher RMSE of 4.83%. Despite this, GRU demonstrates a higher R^2 value of 0.9342, indicating a stronger overall correlation between predicted and true values compared to LSTM, which has an R^2 of 0.8965.

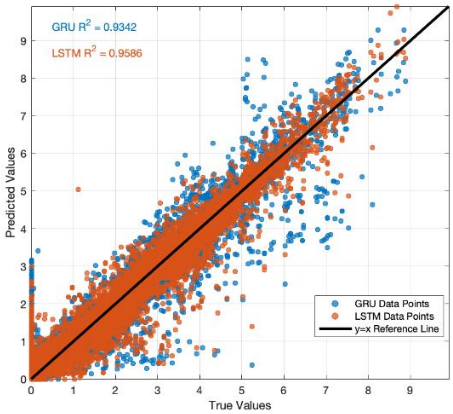

Figure 4 Regression results of the testing set for both LSTM and GRU

Figure 5 provides a detailed comparison of the prediction results of GRU and LSTM, illustrating the daily total heating demand alongside the true values. The top plot represents the GRU predictions, while the bottom plot shows the LSTM predictions. GRU's predictions closely follow the true values, particularly during peak demand periods. It slightly underestimates some of the highest peaks and exhibits more fluctuations during lower-demand periods. Conversely, LSTM also tracks the true demand well but produces smoother predictions. It sometimes overestimates certain peaks but maintains a more stable trend throughout the time series. Overall, both models demonstrate strong predictive capabilities.

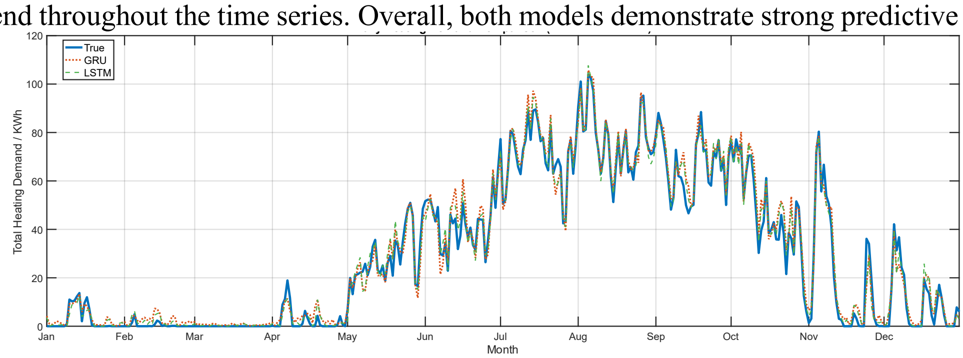

Figure 5 Total daily demand comparisons of test set for both LSTM and GRU

To further analyse the models' ability, Fig. 6 compares the hourly average heating and cooling demand in the training and test sets. The x-axis represents the hour of the day, while the y-axis indicates the average demand across all days. The top plot displays the hourly heating demand in the training set, while the bottom plot presents the hourly cooling demand in the test set. Each curve represents the true values (blue), GRU predictions (orange), and LSTM predictions (green). One key observation is that despite differences in demand profile between the training and test sets, both models effectively capture the overall trend. GRU and LSTM successfully align with the general demand patterns, indicating that they have effectively learned the underlying structure rather than simply memorising the training data. Although minor deviations occur, particularly near demand peaks, both models demonstrate strong predictive capabilities, closely following the true values. This suggests that the models are capable of adapting to variations in the test data while maintaining accurate demand predictions.

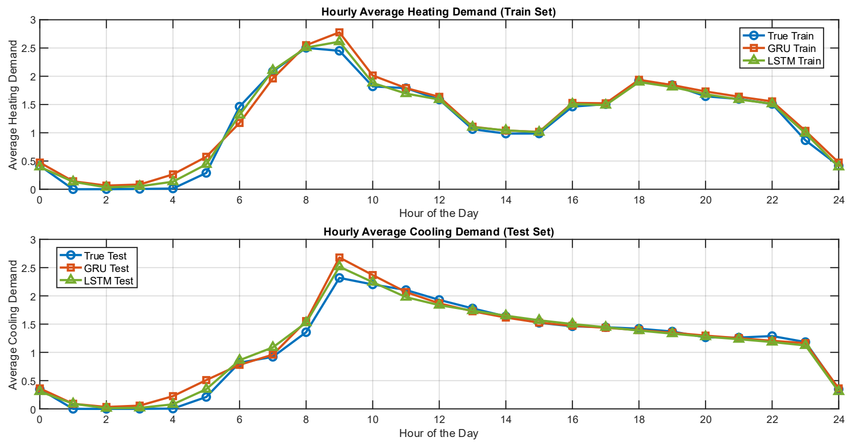

Figure 6 Average hourly heating demand comparisons between for both training and testing

Sensitivity analysis is crucial in machine learning and modelling because it helps to identify which input variables have the most significant impact on the model's predictions. Fig. 7 presents the Sobol Sensitivity Analysis results. Among all features, heating schedule has the highest sensitivity index (1.001), making it the most influential factor in the model. Hour also shows a high sensitivity index (0.239), indicating its significant role in shaping demand profile. It highlights the strong influence of time-dependent factors on the model's predictions. Both factors are directly determined by human decisions and are closely linked to user habits and preferences, as heating schedules are typically set based on occupancy patterns and daily routines. Meanwhile, wet bulb temperature (0.270) and direct normal irradiance (DNI) (0.218) are also higher the average sensitivity threshold, indicating the substantial contribution of meteorological factors. Wet bulb temperature, which reflects both temperature and humidity, strongly influences heating demand, while DNI, representing direct solar radiation, affects heat gains and losses in buildings. These findings confirm that both meteorological conditions and time-related factors significantly impact the predictions, emphasising the interplay between external climate variables and temporal heating demand patterns.

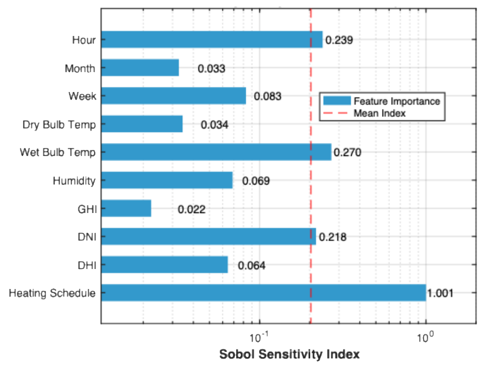

Figure 7 Sobol sensitivity analysis for prediction model

4. Conclusion

This study develops a framework that integrates a physical model with deep neural networks to predict building heating demand. Using a single building as a case study, different heating schedules are input to generate various heating demand profiles. The hourly data generated is then fed into LSTM and GRU models for training.

The results show that both models are able to accurately capture daily heating demand trends, closely following the true values. Even when the test data curve differs from the training data, both models successfully approximate the true hourly demand distribution on average. Additionally, sensitivity analysis highlights the critical impact of occupant behaviours and external weather conditions on heating demand. These findings confirm that both LSTM and GRU can effectively adapt to variations in heating demand. In future research, greater attention should be given to the impact of highly sensitive parameters on model performance to further enhance predictive accuracy.

Acknowledgement

The authors like to acknowledge EPSRC in the UK for financial support through the grants EP/X025322/1, EP/T022701/1, EP/V042033/2, EP/V030515/1, and EP/W027593/2, and the Royal Society for their financial support (Ref: IF\R1\231053). This project has also received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 101007976.

Reference

[1] A. Raihan, “A review of the global climate change impacts, adaptation strategies, and mitigation options in the socio-economic and environmental sectors,” Journal of Environmental Science and Economics, vol. 2, no. 3, pp. 36–58, 2023.

[2] M. Santamouris and K. Vasilakopoulou, “Present and future energy consumption of buildings: Challenges and opportunities towards decarbonisation,” e-Prime - Advances in Electrical Engineering, Electronics and Energy, vol. 1, p. 100002, 2021, doi: https://doi.org/10.1016/j.prime.2021.100002.

[3] Y. Chen, M. Guo, Z. Chen, Z. Chen, and Y. Ji, “Physical energy and data-driven models in building energy prediction: A review,” Energy Reports, vol. 8, pp. 2656–2671, 2022, doi: https://doi.org/10.1016/j.egyr.2022.01.162.

[4] G. M. Stavrakakis, D. Al. Katsaprakakis, and M. Damasiotis, “Basic Principles, Most Common Computational Tools, and Capabilities for Building Energy and Urban Microclimate Simulations,” Energies (Basel), vol. 14, no. 20, 2021, doi: 10.3390/en14206707.

[5] J. Yu, W.-S. Chang, and Y. Dong, “Building Energy Prediction Models and Related Uncertainties: A Review,” Buildings, vol. 12, no. 8, 2022, doi: 10.3390/buildings12081284.

[6] W. Li et al., “Modeling urban building energy use: A review of modeling approaches and procedures,” Energy, vol. 141, pp. 2445–2457, 2017, doi: https://doi.org/10.1016/j.energy.2017.11.071.

[7] A. Chong, Y. Gu, and H. Jia, “Calibrating building energy simulation models: A review of the basics to guide future work,” Energy Build, vol. 253, p. 111533, 2021, doi: https://doi.org/10.1016/j.enbuild.2021.111533.

[8] E. Prataviera, J. Vivian, G. Lombardo, and A. Zarrella, “Evaluation of the impact of input uncertainty on urban building energy simulations using uncertainty and sensitivity analysis,” Appl Energy, vol. 311, p. 118691, 2022, doi: https://doi.org/10.1016/j.apenergy.2022.118691.

[9] F. Del Ama Gonzalo, B. Moreno Santamaría, and M. J. Montero Burgos, “Assessment of Building Energy Simulation Tools to Predict Heating and Cooling Energy Consumption at Early Design Stages,” Sustainability, vol. 15, no. 3, 2023, doi: 10.3390/su15031920.

[10] J. Meiers and G. Frey, “Interfacing TRNSYS with MATLAB for Building Energy System Optimization,” Energies (Basel), vol. 18, no. 2, 2025, doi: 10.3390/en18020255.

[11] Y. Sun, F. Haghighat, and B. C. M. Fung, “A review of the-state-of-the-art in data-driven approaches for building energy prediction,” Energy Build, vol. 221, p. 110022, 2020, doi: https://doi.org/10.1016/j.enbuild.2020.110022.

[12] S. R. Mohandes, X. Zhang, and A. Mahdiyar, “A comprehensive review on the application of artificial neural networks in building energy analysis,” Neurocomputing, vol. 340, pp. 55–75, 2019, doi: https://doi.org/10.1016/j.neucom.2019.02.040.

[13] A. S. Ahmad et al., “A review on applications of ANN and SVM for building electrical energy consumption forecasting,” Renewable and Sustainable Energy Reviews, vol. 33, pp. 102–109, 2014, doi: https://doi.org/10.1016/j.rser.2014.01.069.

[14] G. Ciulla and A. D’Amico, “Building energy performance forecasting: A multiple linear regression approach,” Appl Energy, vol. 253, p. 113500, 2019, doi: https://doi.org/10.1016/j.apenergy.2019.113500.

[15] X. Li and R. Yao, “Modelling heating and cooling energy demand for building stock using a hybrid approach,” Energy Build, vol. 235, p. 110740, 2021, doi: https://doi.org/10.1016/j.enbuild.2021.110740.

[16] H. Sun, “Prediction of building energy consumption based on bp neural network,” Wirel Commun Mob Comput, vol. 2022, no. 1, p. 7876013, 2022.

[17] X. J. Luo and L. O. Oyedele, “Forecasting building energy consumption: Adaptive long-short term memory neural networks driven by genetic algorithm,” Advanced Engineering Informatics, vol. 50, p. 101357, 2021, doi: https://doi.org/10.1016/j.aei.2021.101357.

[18] C. Li, G. Li, K. Wang, and B. Han, “A multi-energy load forecasting method based on parallel architecture CNN-GRU and transfer learning for data deficient integrated energy systems,” Energy, vol. 259, p. 124967, 2022, doi: https://doi.org/10.1016/j.energy.2022.124967.