The Original

PERFORMANCE AND ANALYSIS OF INTELLIGENT HVAC HYBRID CONTROL FOR HIGH ENERGY EFFICIENCY IN BUILDINGS

[1*]Ganesh Murade, [2]Ankit Kumar Sharma, [3]Bhanu Pratap Soni, [4]Ganesh Shirsat

[1*]Department of Electrical Engineering, University of Engineering & Management, Jaipur, India. ganesh.murade26@gmail.com
[2]Department of Electrical Engineering, University of Engineering & Management, Jaipur, India. ankitks.sharma@uem.edu.in
[3]Fiji National university; University of Engineering & Mgt. India. bhanu.soni@fnu.ac.fj
[4]Department of Electrical Engineering, Dr. Vithalrao Vikhe Patil College of Engineering, Ahilyanagar, India. ganeshshirsat1234@gmail.com

ABSTRACT:
HVAC systems, including heating, ventilation and air-conditioning, have a big impact on the energy use of buildings. In commercial buildings these systems can represent almost 40-60% of the total electricity consumption of the facility. Buildings in general also account for almost a third of the world's energy consumption. This high demand has made the improvement of HVAC operations an important area of research for reducing energy use, controlling operating expenses and supporting environmental sustainability.
In this paper, we propose a predictive HVAC control method based on the combination of GRU and Random Forest. The main objective was to evaluate the feasibility of integrating a deep-learning model with a traditional machine-learning algorithm to provide more dependable control in various operating conditions. The proposed Hybrid GRU+RF model was evaluated under two cases. Renewable-energy information was not included in the first case, which is a normal operating environment. The model in this case reduced energy consumption by 24.29%. For the second case, the solar irradiance was another input. The existence of this variable complicated the prediction problem. However, the model proposed still achieved an energy saving of 15.43%.
Results from LSTM and GRU individually were significantly different. These individual deep-learning models performed relatively well under the more straightforward condition, but the performance became worse when renewable-energy features were added. The resulting energy-saving values were negative, being between −15.85% and −16.27%. This leads to an approximate 36-37 % performance loss compared to the case without renewable-energy inputs. The Hybrid GRU+RF model was not deteriorated to the same extent. It even presented positive results in terms of energy saving. This means that the combination of GRU and Random Forest was better to deal with the additional variations related to solar irradiance.
The proposed model was also tested not only with simulated or laboratory data . The system was also studied with real operational data collected over a period of one full year from an industrial plant in Jalgaon, Maharashtra, India. The dataset included 10,510 observations. Additional validation was done on another data set of a multi-zone building (2,016 samples) from the ORNL FRP-2 facility. The Hybrid GRU+RF model achieved a R^2 score of 0.9335 for the prediction of temperature and 0.9205 for the prediction of energy. The results show that the model was able to predict both the energy requirements and the temperature of the building with a good degree of accuracy. The testing also showed that the hybrid approach could perform well when applied to building conditions different from those used during the model development. The standalone Random Forest model, however, was showing signs of overfitting.
The results of the study indicate that the Hybrid GRU+RF model can be utilized as a practical approach for predictive HVAC energy management. One of its main advantages was that its performance was useful even after the renewable-energy information was included into the prediction process. This is especially true of existing buildings where solar power and other renewables are becoming more prevalent. Based on the obtained results, the proposed method could estimate an annual saving of about \$1,600–\$2,000 per building and could reduce carbon emissions by approximately 11.5 tons of CO_2 per facility per year. Therefore, the study suggests that the combination of GRU with Random Forest may improve the reliability of intelligent HVAC control. It can contribute to the creation of more energy-efficient and sustainable smart buildings.

KEYWORDS: HVAC control, Machine learning, Deep learning, Hybrid prediction models, Building energy efficiency, Solar irradiance integration, Time-series forecasting, Renewable energy variability, Smart buildings, Building management systems, Model predictive control, Generalization capability, Energy optimization

INTRODUCTION

1.1 Global Energy Crisis and Building Sector Context

Global energy consumption is on the rise and raising significant concerns for both the environment and economy. "Buildings represent about 30–40% of global energy demand, and almost 40% of global carbon dioxide emissions," says the International Energy Agency (IEA). This demand for energy is supplied by a variety of facilities such as homes,

offices, commercial buildings, factories and other infrastructure. Buildings are the largest energy-consuming sector, second only to industry in total energy demand.
There are several reasons for the high energy demands of buildings. Fast growing cities mean more people living closer together and more use of energy-hungry buildings and services. Meanwhile, rising global temperatures have increased demand for air-conditioning, which further increases demand for electricity and leads to further emissions. This produces a vicious cycle where warmer temperatures lead to greater cooling needs . The more energy we use, the more emissions. More emissions lead to climate change. Living standards in developing countries have also been rising rapidly. Use of air-conditioning equipment has been increasing rapidly. Meanwhile, governments are imposing more stringent energy-efficiency standards and carbon reduction targets, ratcheting up the pressure on building owners and facility operators to cut energy use.
HVAC systems represent some of the largest contributors to the energy demand of buildings. In commercial buildings, HVAC equipment can represent 40 to 60 percent of total electricity consumption. In office buildings, values are typically reported in the 35 to 50 percent range. In facilities with tightly controlled environments (e.g. data centers) HVAC related energy consumption can be 70% or more. Plants and labs of manufacture have similar consumption levels. It is critical to maintain some levels of temperature, humidity and air quality.
HVAC systems use a lot of energy for many reasons. Many facilities run their HVAC equipment 24 hours a day, 7 days a week, whether the building is occupied or not. These systems also need to maintain temperature and humidity within certain ranges to provide a comfortable environment for occupants, as well as protect sensitive equipment or products. Traditional control methods are reactive; they respond to changes that have already happened, they do not predict future states. Further, old HVAC equipment may not operate efficiently because of aging parts, lack of maintenance or operation outside the ideal operating range. Most existing systems have limited adaptive capability and are therefore unable to respond effectively to sudden changes in outdoor weather, occupancy or indoor environmental conditions.
Improving HVAC efficiency is thus an important avenue for reducing building energy use and associated emissions. A 10% increase in global HVAC energy efficiency would have enormous environmental benefits. It would be the equivalent of taking approximately 100 million cars off our roads each year and would avoid some 500 million tons of CO2 emissions. Governments, industries and organizations are all feeling the pressure to deliver on sustainability and carbon reduction goals, and intelligent HVAC management is a huge area for technology development. By using artificial intelligence and machine-learning techniques, more responsive, predictive and energy efficient HVAC systems can be developed, supporting the trend toward smarter, more sustainable buildings.

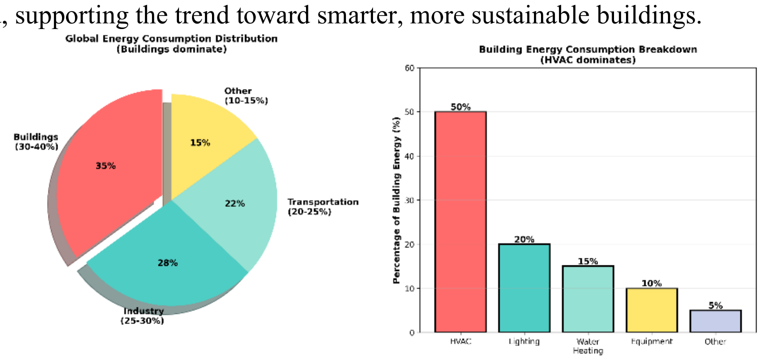

Figure 1.1: Global Energy Distribution and HVAC Dominance.
Buildings account for 30-40% of worldwide energy consumption, with HVAC systems accounting for 40-60% of building energy. Combined, HVAC represents 12-24% of global energy consumption, making optimization critical for global sustainability goals.

1.2 Limitations of Traditional HVAC Control Systems
Many traditional HVAC systems still use PID based feedback control methods that became popular several decades ago. These controllers respond primarily to the differences between the measured room temperature and the desired setpoint. Consequently, the system usually takes corrective action only after an undesired temperature change has already occurred. This leads to a delay in the control response especially for sudden changes in outdoor weather, solar radiation, occupancy or other factors. This may result in the HVAC system overcompensating by heating or cooling more than necessary before returning to the required condition, and consequently, causing temperature fluctuations and unnecessary energy consumption.
Another limitation is the use of fixed temperature setpoints. In many buildings, a constant temperature of 22°C is maintained all day long, regardless of whether the space is occupied or mostly empty. You can change the setpoint seasonally, but typically these are manual changes and do not respond to actual building conditions. Many facilities also lack or have limited use of occupancy sensors that would enable HVAC systems to accurately determine whether or not certain areas actually need heating or cooling.
Traditional HVAC operation is also often disconnected from other building systems. For example, lighting may be automatically dimmed when a space is unoccupied, but the HVAC system may continue to supply conditioned air at nearly the same level. This lack of coordination can lead to energy consumption when the conditioned space is unoccupied. Additional energy savings can be realized through improved interaction of HVAC, lighting, occupancy detection and other building management systems.
Another major disadvantage of conventional control methods is the limited ability to learn from previous operating conditions. Often the same control parameters and rules are used again and again, although the building behavior can

change a lot between seasons, occupancy and weather conditions. These systems typically fail to identify the complex relationships between outdoor weather, solar gains, occupancy patterns and indoor temperature responses. Consequently, they have a limited capacity to adapt their operations to changing circumstances.

Prior studies have demonstrated that operational issues, such as control limitations, cause HVAC systems to run at an efficiency of about 15–30% below their potential. Basic feedback-control approaches typically achieve only about 5-10% energy savings relative to uncontrolled operation and further improvements are difficult to attain within the same strategy. These limitations point toward the need for advanced control strategies capable of anticipating future conditions, responding to changing building behavior, and dynamically adapting HVAC operation. Therefore, intelligent methods based on machine learning and predictive control can be a more flexible alternative to conventional PID-only control systems.

1.3 Research Gaps and Study Motivation

The advent of machine learning and deep learning has created new opportunities for enhancing HVAC control and energy use in buildings. Existing approaches generally rely on Model Predictive Control (MPC) which requires an accurate model of the building and its operational behavior. Such approaches have typically been associated with energy savings in the range of ~12-18% . An alternative are machine learning methods that learn the relationships directly from the historical and operational data without requiring the same degree of physical modeling. Deep learning approaches such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are especially strong in detecting temporal patterns and long-term changes in building energy behavior.

There is little research on the application of artificial intelligence to building-energy management and many key research gaps are still to be filled. In this paper we attempt to overcome these limitations and to provide a more practical assessment of intelligent HVAC control methods.

An important limitation is the limited availability of real industrial HVAC data for the evaluation of new algorithms. Most of the today’s work is based on simulation environments or experiments on small buildings. Few studies address large-scale real-world datasets collected over a full year, which inherently contain variations regarding weather, occupancy, operating schedules and other unpredictable conditions.

Another area requiring further investigation is renewable energy impact. The increasing installation of solar photovoltaic systems in buildings, increases further the variability of the building-energy behavior with solar generation and irradiance. Most studies existing today do not study the response of predictive HVAC models when including renewable-energy information as input. It is thus important to investigate whether these additional features improve prediction and control or rather make the modeling problem more complicated.

Another gap lies in the limited comparison of different control and prediction methods. Most of the research are concentrated on one category such as conventional machine learning, deep learning or Model Predictive Control. Few studies have compared these techniques or combined them under the same operating conditions. It is important to make a direct comparison to understand which methods give reliable results in real conditions of building.

The prediction accuracy is also often taken as the main measure of the model performance. Metrics like R^2 and Mean Squared Error (MSE) are useful in assessing the quality of the prediction, but accurate predictions do not necessarily translate into lower energy consumption or operating costs. A model may give good statistics, but little practical improvement in energy use of HVAC. Therefore, the real energy saving and economic benefits should also be included in the evaluation of intelligent control strategies.

Another open question is if a model can generalize well to buildings that have not been used for training. Many models are good predictors for the building they trained but their accuracy can significantly drop when applied to another building with different characteristics and operating patterns. To know whether a model can be more generally used, testing generalization on independent datasets is needed.

Then there’s the problem of variability of renewables. Availability of solar irradiance and renewable energy can change rapidly due to weather and environmental conditions. However, the capability of predictive HVAC models to maintain stable performance under such fluctuations has not been well studied by many existing studies .

Implementation cost and complexity should also be taken into account. A lot of computing power, dedicated hardware or a lot of pre-processing of data may be required by some advanced AI-driven control systems. The practical uptake may be limited unless the financial and energy benefits outweigh these requirements. Therefore, the real value of intelligent HVAC solutions should be estimated by assessing performance and economic feasibility.

Finally, a clear answer to what predictive approaches can consistently maintain their performance across different buildings and under varying operating conditions is still missing. A good solution should not be too dependent on a particular environment and should be reasonably stable as conditions change.

In this paper we address these eight research challenges by systematically evaluating six different methods on one year of real industrial HVAC data. We consider two experimental settings: one using conventional building and HVAC variables and the other including the solar irradiance information, so as to study the effect of the renewable-energy conditions. A new Hybrid GRU+RF model is developed by combining the ability of the Gated Recurrent Unit (GRU) to learn the temporal relationships and the capability of the Random Forest to perform the regression task. The focus is not just on the predictive accuracy of the models but on actual HVAC energy savings. The models are also tested on ORNL FRP-2 dataset for their generalization property on a different building environment. The model behavior is also investigated under different renewable-energy conditions, considering both energy reduction and associated cost benefits. Finally, the consistency of the different approaches is studied in an effort to understand which methods are still valid when the operating conditions and the building characteristics change.

2: LITERATURE REVIEW

2.1 Historical Evolution of HVAC Control Technologies (1950-2025)

Modern HVAC control technology has evolved far beyond the simple mechanical devices of seventy years ago into sophisticated computerized and predictive systems. Every step in this development has brought better control capabilities and better energy performance, but also more system complexity and higher computational requirements. These changes reflect an evolution in HVAC technology, and an understanding of that evolution is also an understanding of the limitations that still prevent buildings from realizing their full potential for energy savings.

The Age of Machines For the most part, from the 1950's through the 1970's, HVAC systems were controlled by mechanical parts like gears, levers, pneumatic devices and simple thermostats. They used temperature sensitive parts like metal coils, which would turn on switches when the temperature around it changed. These systems were simple, robust but the control accuracy was relatively bad. The room temperature could drift a degree or two from the set point before the system would kick in. Sometimes the response itself is delayed for a few minutes 5-15 minutes sometimes. The equipment was generally controlled by simple ON/OFF control and thus had little capability for gradual adjustment. This led often to repeated temperature cycles between heating and cooling conditions and unnecessary energy use.

A major enhancement was the introduction of Electronic PID Control in the 1980s and 1990s. With the introduction of electronic controllers, Proportional-Integral-Derivative (PID) algorithms could be implemented to achieve a smoother control over the HVAC output. Then the controller could modify the output not simply ON or OFF, but based on the difference between the actual and desired temperature, and also based on accumulated and changing errors. However, the achieved performance was strongly dependent on the choice of appropriate PID parameters. Improper tuning can lead to excessive cycling, instability or continued energy waste. However, electronic PID systems did offer some 10-15% energy savings over the earlier simple control methods. However, their basic operation was reactive, because the control actions were mostly performed after the deviation from the desired condition had already occurred.

In the 1990's and 2000's Building Automation Systems (BAS) became more and more common. These systems allowed for monitoring and control of HVAC equipment in multiple rooms and zones from a single platform. Further sensors provided information on occupancy and environmental conditions and scheduling functions allowed HVAC operation to be coordinated with expected building use. These capabilities improved energy management and could save around 15-25% energy as compared to uncontrolled HVAC operation. However, BAS were usually based on predefined rules, schedules and control sequences; the system could react to measured conditions but could not usually learn from prior operation or change its strategy independently when building behavior changed.

The advent of Model Predictive Control (MPC) in the 2000s made it possible to manage HVAC in a more sophisticated manner. MPC was borrowed from industrial process-control applications and allowed building systems to anticipate future operating conditions rather than just react to measurements in the moment. Tools such as EnergyPlus can be used to develop detailed building models or physics-based mathematical models. The building behavior was predicted for a future time horizon (usually 24–48 h ahead). Then the controller could decide on operating strategies considering at the same time the comfort of the occupant, the energy consumption and the operation of the equipment. It can control chillers, boilers, fans, and other HVAC equipment and has reported energy savings of approximately 12-18%.

Despite these benefits, deploying MPC in actual buildings can be challenging. The optimization calculations often have to be performed within a short control interval, e.g. 10-15 min. This results in computational requirements. And also the precise operation depends on a detailed model of the particular building. It can take a lot of time, technical expertise and data to build such a model. Building characteristics also change over time due to equipment degradation, changes in floor layouts, change in occupancy and operating schedule variations. Thus, maintaining the model in sync with these changes can be a major maintenance burden. Furthermore, integration into existing building management infrastructure may be required, which may require custom solutions and increase implementation effort and cost.

In general, the HVAC control technology development has gradually reduced the energy waste and improved the building response capability to the operating conditions. However, the transition from mechanical thermostats to electronic control, BAS and predictive methods, has not overcome the limitations of independent learning and adaptation of most systems. These limitations present an opportunity for newer data-driven techniques, particularly machine learning and deep learning, to equip HVAC systems with greater predictive capability and adaptability.

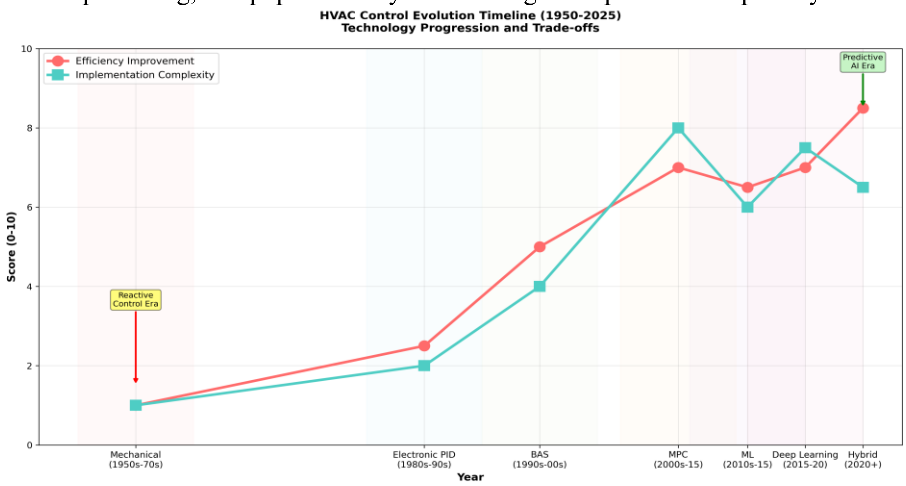

Figure 2.1: HVAC Control Evolution Timeline (1950-2025).

(Each technological generation improved efficiency while adding complexity. Transition from reactive PID to data-driven ML to intelligent hybrid systems represents paradigm shift in HVAC control philosophy).
Since approximately 2010, building-energy research has increasingly focused on machine learning techniques such as Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting, and Artificial Neural Networks (ANN). The methods allowed the researchers to use measured building data directly instead of detailed physics based models. Machine learning algorithms are able to learn useful patterns from the available data, rather than the manual specification of all the relationships affecting building performance. One of their main advantages is therefore the possibility to model complex and non-linear relationships without needing a complete mathematical formulation of the building. This also reduces the effort for the development of detailed building models and can help to avoid uncertainties related to assumptions made in physical modeling.
Alawadi et al. studied the performance of Random Forest, Support Vector Machine and Artificial Neural Network methods for indoor temperature prediction in different buildings . The Random Forest approach reached the best prediction performance, with the R2 value of 0.9807 and relatively shorter training time compared to the evaluated approaches. In another study, Xu proposed a method of Random Forest with metaheuristic optimization for the prediction of heating loads. The proposed method achieved better performance than the traditional regression based methods and showed the potentials of machine learning algorithms for building energy prediction.
Classical machine-learning methods are also faced with important limitations for the case of building data with time-dependent nature. Many standard algorithms use individual observations, and do not explicitly model the relation from one time step to the next. This may prohibit them to detect trends that occur in longer time periods and to understand how the past operating conditions influence the future behavior of the building. The other challenge is overfitting, in particular when training models such as Random Forest on fairly small or limited datasets. Furthermore, a model trained on one building may not perform well on other buildings with different characteristics, layouts, occupancy patterns or operating conditions.

2.2 Deep Learning Approaches for HVAC and Building Energy

Since about 2015, deep learning, especially Recurrent Neural Networks (RNNs), has gained more importance in the domain of time-series prediction in building-energy applications [24]. We design an RNN to handle sequence data . We predict the future time steps given the information from the previous time steps. In traditional models, observations are assumed to be uncorrelated. This enables their use for implementing applications whose current temperature and energy requirements depend on previous operating conditions. RNN has two popular architectures namely Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU).
Hochreiter and Schmidhuber [25] proposed LSTM networks which are a set of gating operations that control the storage of information in the network. The three main gates are the forget gate, the input gate and the output gate. The forget gate decides what information from the previous time step to forget . The input gate decides what new information to add . The output gate decides what information to output to the next time step . This architecture allows LSTM models to retain important information over long sequences, which is of great value for building-energy forecasting as conditions that have occurred several hours or even days ago can influence current energy demand. Li et al. demonstrated that LSTMs can accurately predict temperatures in office buildings with R2 > 0.95 . Results also show that the models can be successfully applied to different buildings.
GRU networks are a slightly simpler alternative to LSTM architectures. Cho et al. designed GRUs that have around 25% fewer parameters than LSTMs, but still have good sequence learning capabilities. Unlike the usual LSTMs, GRUs do not have the separate forget and input gates. They use an update gate that merges the function of the two gates, leading to a simpler network structure. This reduces the complexity, which can reduce the training time, memory requirements and computational requirements. Liu's research demonstrated the capability of GRU models to achieve R2 values higher than 0.90 for building temperature prediction and also showed good performance on different datasets. Tran et al. (2019) also investigated the hybrid strategies of combining the GRU networks with other prediction methods for the load forecasting in multi-zone buildings. Their results demonstrate the advantages of hybrid architectures for complex building-energy applications, where the combined models are able to outperform the prediction of single models [31].

2.3 Comparative Analysis of HVAC Control Techniques

Comparison of HVAC control methods The comparison highlights the strengths and weaknesses of the methods. The classical PID is fast with a typical response time in the microseconds and relatively easy to implement. But it is mainly an error correction and cannot forecast future conditions or learn from the previous system behavior. Model Predictive Control (MPC) aims at finding the optimal set points, but it requires high computing power, detailed building models and careful system integration.
Random Forest and SVM are good for accuracy and work well with complex or noisy data. However, they cannot inherently capture the relations of sequential time-series data, and they can overfit when the available dataset is small. Deep learning models such as LSTM and GRU can capture the temporal patterns well but require more training data and computational resources and may not generalize well to unseen conditions. Thus, there is no single best HVAC control solution for all scenarios since each involves a number of trade-offs between complexity, accuracy, adaptability and practical performance.

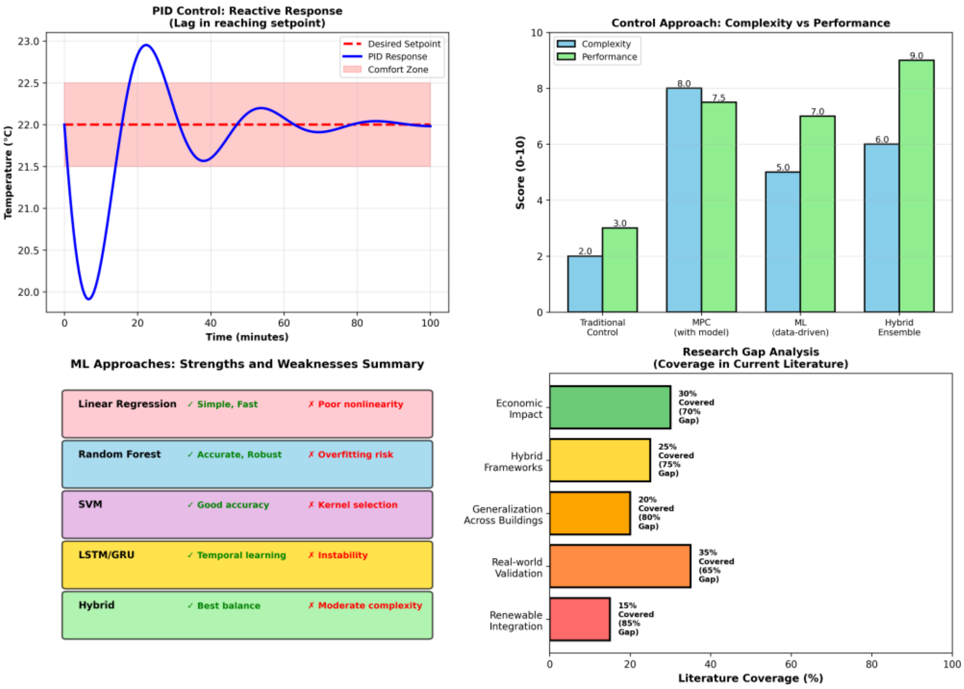

Figure 2.2: Classical Control Systems Comparison and Research Gaps Analysis. (Left) PID control exhibits reactive lag and oscillation around setpoint. (Right) Research gap heatmap showing renewable integration only 15% covered (85% uncovered), hybrid frameworks only 25% covered (75% uncovered). These critical gaps motivate present study.

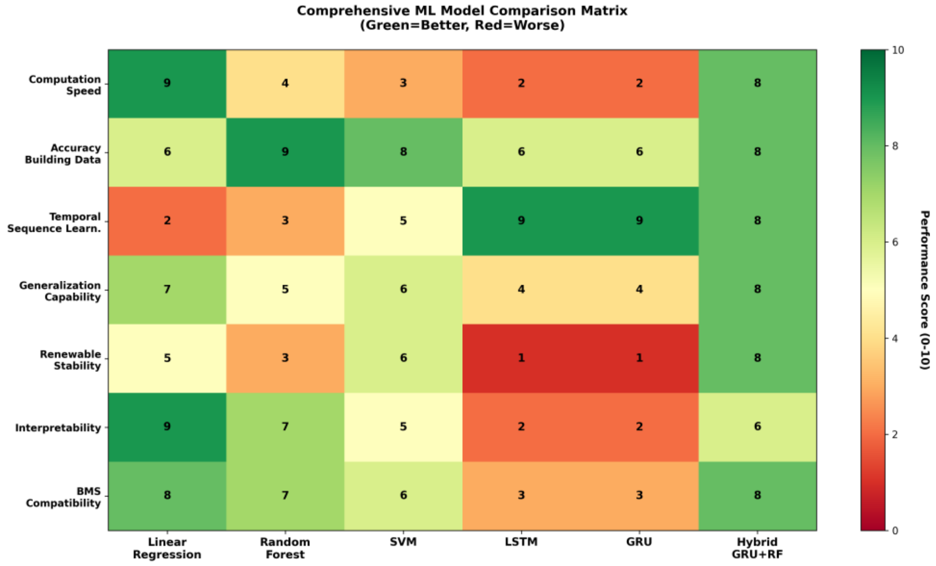

Figure 2.3: Comprehensive ML Model Comparison Matrix (Seven Performance Dimensions).

(Heatmap reveals that no single approach excels across all dimensions. Hybrid GRU+RF achieves highest overall scores, particularly excelling in temporal learning, generalization capability, renewable stability, and BMS compatibility—four critical dimensions where individual approaches demonstrate weaknesses).

2.4 Renewable Energy Integration in Building HVAC Systems

On-site renewable-energy systems are being increasingly incorporated into modern buildings. These include solar photovoltaic panels, solar thermal collectors, wind turbines and geothermal technologies. These systems can decrease the reliance on the traditional energy sources but their variability brings further challenges to the HVAC control. For instance, solar energy is intermittent during the day, not available at nite, and can be suddenly interrupted by clouds passing by. Seasonal variations also affect solar availability, with generation levels varying greatly between summer and winter seasons. HVAC controllers need to deal with these changes, while keeping occupants comfortable, coordinating energy storage, and interacting well with the electrical grid.

The prediction of solar energy availability has been the focus of intense research in the recent years. Taha et al report that hybrid deep-learning models can predict solar irradiance with an accuracy of more than 85 percent. Walsh also contributed to the quality of solar datasets for building applications. This enables the HVAC systems to pre-empt the operation based on the predicted renewable generation. However, there is an important question left: how different machine-learning approaches respond to the inclusion of solar forecasts in HVAC control? However, the research on this specific issue is still limited, especially on the different responses of machine-learning and deep-learning models to the variability of renewable energy.

Previous studies, including Zhou's work on hybrid renewable systems and phase-change materials and Gupta's work on hybrid deep learning for solar prediction in HVAC applications, have reported improvements in energy performance or forecasting accuracy. However, these studies have generally not examined the response of different machine-learning models to renewable-energy inputs, nor have they compared traditional machine-learning methods with deep-learning approaches under similar renewable-energy conditions. This work aims at filling this gap by providing a systematic comparison of different model types for different renewable-energy scenarios and by providing a better understanding of the impact of renewable variability on HVAC predictive control.

2.5 Hybrid and Ensemble Approaches in Machine Learning

Ensemble learning methods merge the outputs of many models together to improve the overall prediction performance and have been successfully used in many machine learning tasks. Hastie et al. showed that if you average the predictions from several models, you can reduce the variance while still maintaining good accuracy. These approaches can help to avoid overfitting , take advantage of the strengths of different algorithms , discover different patterns in the data and make more stable predictions in the presence of outliers .

Several studies have shown practical use of ensemble techniques in building-energy applications. Salem et al. coupled Random Forest (RF) with Artificial Neural Networks (ANN) for HVAC energy prediction, and reported better R2 performance. Pallonetto employed ensemble-based approaches for smart-grid applications and obtained useful results for demand-response optimization. Tran et al. combined the GRU networks with other prediction techniques for building-load forecasting and found that the hybrid approaches outperformed individual models .

But, most of the previous work has explored model combinations with relatively similar characteristics, such as RF with ANN or multiple CNN-based architectures. The proposed Hybrid GRU+RF model uses a different way. The GRU component first learns the temporal relationships of the input data with its recurrent structure, then the extracted features are passed to a Random Forest model for nonlinear regression. This combination exploits the ability of GRU to learn the time-dependent patterns while Random Forest is used to provide a robust regression performance and to reduce the risk of overfitting.

2.6 Critical Research Gaps and Study Contributions

The present study responds to eight main research gaps identified through a literature review. First, the renewable energy integration is studied by directly comparing Case I and Case II based on the one full year real industrial HVAC data. Further, the study examines deep-learning models such as LSTM and GRU and finds a significant drop in their accuracy when solar-related features are included. A Hybrid GRU+RF model is proposed and tested to evaluate its ability to achieve stable performance in the setting of changing renewable-energy conditions. The practical evaluation is performed on 10,510 samples collected from industrial HVAC system at Jalgaon, India. Further generalization is studied on ORNL FRP-2 dataset which is a different building type and HVAC configuration.

The study also investigates whether higher prediction accuracy, measured in terms of metrics like R^2, also results in higher energy savings. The results show that improving the accuracy of predictions does not guaranty better energy efficiency. The robustness of different models for the integration of solar energy is also studied and it is found that the hybrid model provides more consistent results. Finally, the practical feasibility of the proposed method is investigated by studying its computational requirements and its compatibility with existing building-management systems. Collectively, these investigations address important gaps in the current research on HVAC control and provide a more comprehensive evaluation of intelligent control under realistic operating conditions.

3: METHODOLOGY

3.1 Problem Formulation and Dual-Scenario Evaluation Framework

In this work the HVAC control is formulated as a time-series forecasting problem and the models are tested under two defined operating scenarios. The two-case use is to understand how renewable-energy information use affects the performance of different machine-learning approaches.

Case I is the traditional HVAC operation mode without renewable-energy inputs. The models use standard HVAC variables like indoor temperature (TZ1), thermal comfort (ITUZ1), relative humidity (RH), airflow (AF) and energy consumption (E). This case serves as a baseline for the model performance when the system is operating in normal conditions, without any additional variability from the renewable energy sources.

Case II adds to the features of Case I the renewable energy related variables. These additional inputs are solar irradiance, solar height angle, outdoor temperature, and wind speed. This case is aimed at measuring the reaction of different models to the added complexity due to the integration of renewable energy. The results of both cases can then be compared to see what effect these extra variables have. For example, in Case I, the Hybrid GRU+RF model achieved energy savings of 24.29% and in Case II, energy savings of 15.43%.

The study looks at three main prediction objectives: the setpoint of the supply air temperature (TZ1), the energy consumption of the HVAC (ITUZ1) and the percentage of energy saving in comparison to the baseline. These objectives allow a more holistic assessment of the model performance rather than only the prediction accuracy. One model can predict temperature accurately but save little energy, while the other model can save energy more effectively with slightly different prediction performance. Thus, all three objectives are considered to obtain a complete evaluation of the proposed HVAC control approaches.

3.2 Comprehensive Dataset Description and Characteristics

In this study two different datasets are used, one for the training and testing of the model and one for independent validation. This setup provides a straightforward way to evaluate the model performance and its generalization to data from a different building.

The primary data set was collected from the operational HVAC system of RMC Performance Pvt. Ltd. Jalgaon, Maharashtra, India. The facility is about 7,000 square feet with four floors that include both manufacturing and research space. Occupancy is around 200-300 people on a typical weekday from 8:00 AM to 6:00 PM. Occupancy levels are considerably lower in the evenings, on weekends and holidays.

There is a central HVAC system composed of one main Air Handling Unit (AHU) and seven fresh-air supply units for the building. The system is controlled to keep the inside temperature at 22◦C ± 1◦C and the relative humidity between 45% and 55%. The HVAC operation is kept day and nite even during low occupancy or no occupancy. This mode of operation provides useful information on the differences in building conditions between occupied and unoccupied periods and between day and nite operation.

Data were collected every 10 minutes from January 1, 2023 to December 31, 2023. The entire period can have more than 52,000 records with 144 possible observations a day. The data collected are zone temperature (TZ1), relative humidity (RH), airflow (AF), energy consumption (E) and solar related variables for the case of interest in the experiment. After data screening, 10,510 records were considered suitable for analysis, and missing values were less than 2%. Short gaps (less than one hour) were filled with linear interpolation. To further ensure consistency, the processed data were also compared with AHU records.

The second dataset was from the Oak Ridge National Laboratory Flexible Research Platform 2 (ORNL FRP-2) and was only used for independent validation. This platform has more than eight zones, each with its own control strategies, as well as complex interactions between zones, than the RMC facility. It also uses a distributed HVAC configuration, rather than the centralized configuration used at the RMC facility.

The ORNL FRP-2 measurements are taken at time intervals of 1 min to 5 min and consist of 2,016 observations from various times of operation. The dataset poses a hard challenge for model generalization: models trained on the 10,510 samples collected at the industrial facility need to be evaluated on a building with very different characteristics, zoning schemes and HVAC architecture.

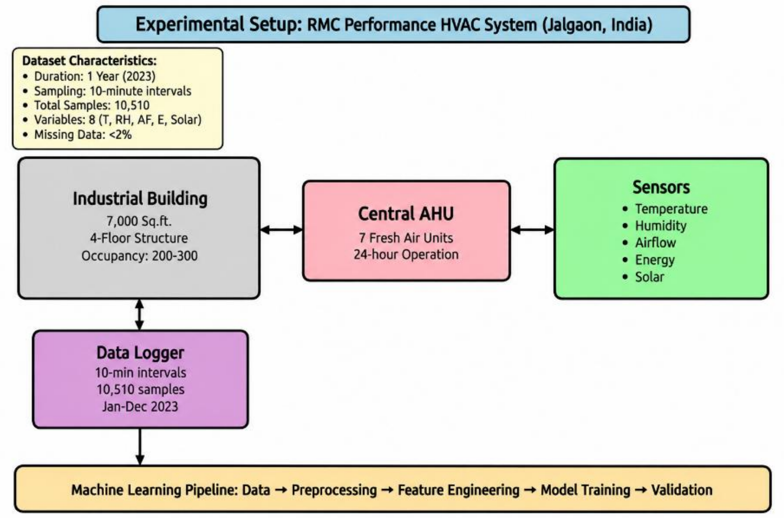

Figure 3.1: Experimental Setup at RMC Performance Facility (Jalgaon, Maharashtra, India). Industrial facility with central HVAC system and distributed sensors provides comprehensive real-world data. 10-minute sampling interval over full year captures seasonal variations, occupancy patterns, and equipment dynamics essential for robust model development.

3.3 Two-Stage Hybrid GRU+RF Architecture Detailed

The proposed framework integrates two machine-learning models, which compensate for their deficiencies mutually. The process starts with a Gated Recurrent Unit (GRU) reading sequential data from the HVAC. The input is split into sequences of 10 time steps, each time step is 10 minutes, the entire sequence is 100 minutes. The window should be long enough to include enough information to detect meaningful changes but not so long that the sequence becomes unwieldy. The GRU has only one layer with 128 units in it and additional layers are avoided to reduce the computational complexity. To avoid overfitting, we apply a dropout rate of 0.2. Adam optimizer with a learning rate of 0.001 and a batch size of 32 is used for training the model. We train for a maximum of 50 epochs and early stop if no improvement is observed for 5 consecutive epochs. The GRU is used to generate the Zone 1 temperature prediction as well as a 128-dimensional feature representation which contains the information learnt from the input sequence.

The GRU prediction and the extracted feature vector, along with other relevant input variables (e.g. solar-energy measurements when needed), are fed into the Random Forest stage. The model is composed of 100 choice trees with a maximum depth of 15 for each tree. A node can only be split if there are at least five samples and each leaf has to contain at least two samples. The splitting criterion used to determine the tree splits is the Mean Squared Error (MSE).

Therefore, the combination couples the temporal learning ability of GRU with the nonlinear regression power of Random Forest. The GRU learns time-evolving patterns and the Random Forest uses the learned features as well as other variables to estimate the energy consumption and find important influencing factors. The hybrid structure also provides for more flexibility as operating conditions change, and helps to maintain robustness of predictions in the presence of noise and unusual observations.

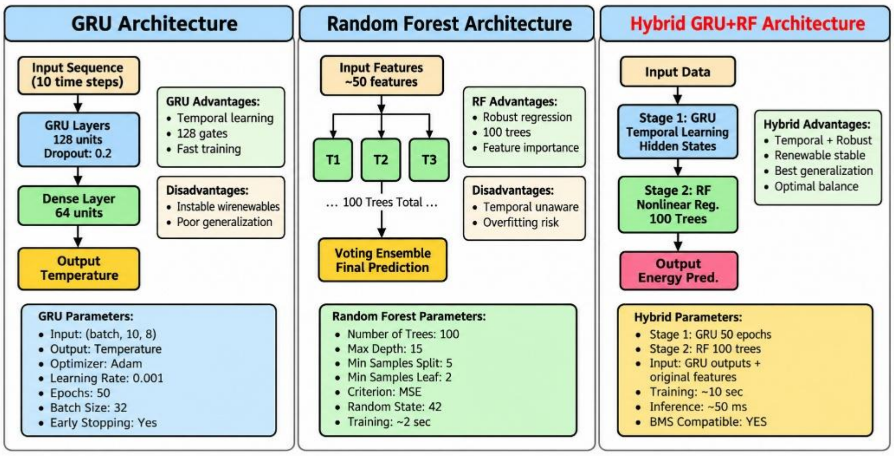

Figure 3.2: Detailed Hybrid GRU+RF Architecture Comparison.

(GRU excels at temporal learning but shows renewable-induced instability. RF provides robust regression but lacks temporal awareness. Hybrid combines both strengths: Stage 1 learns temporal patterns (GRU, 128 units, gating mechanisms), Stage 2 applies robust regression (RF, 100 trees) to combined features).

The proposed architecture is different from the conventional way where the output of the GRU is directly used as the input for the Random Forest (RF) model. It consists of two separate phases. In Stage 1, the GRU generates temperature predictions and hidden-state representations. The output of the temperature reflects how well the model understands the time sequence, while the 128-dimensional hidden-state vector retains the more in-depth features extracted from the input data.

During Stage 2, the RF model receives not only the simple prediction output but also the hidden features extracted by the GRU model . This combination helps the RF model to better grasp the connections between the learned temporal patterns and energy usage. Thus it can access information which may not be available from the temperature prediction alone.

The main benefit of this structure is that it retains more information than simply combining/averaging model outputs. Each stage does a different task, where the GRU learns the temporal relationships and the RF applies the learned information for the nonlinear regression. This enables the hybrid model to learn a larger variety of useful patterns and possibly to perform more robust predictions.

3.4 Data Processing and Feature Engineering Pipeline

HVAC sensors generate a ton of raw data that is not always analysis ready. Therefore, the data needs to be pre-processed to enhance the quality and consistency before the application of machine learning algorithms.

Step 1- Dealing with missing values Incomplete records may arise from sensor failures or communication failures. Short gaps (1-3 h of missing data) are detected and filled by using interpolation methods. This provides a more continuous data set and reduces the effect of missing measurements.

Now we want to find the extremes, the outliers. Anomalously high or low readings from a sensor may indicate measurement errors or equipment problems. Any such findings are marked for further investigation and identified to ensure that they do not unduly influence the model.

Then the data are normalized because variables in HVAC analysis have different units and ranges of values. This enables, for example, to compare the parameters temperature, air flow and energy consumption on a common scale. This makes the variables more comparable and the input easier to work with the learning algorithms.

Finally, the processed variables are combined into a consistent data set for model building. So far we have dealt with missing values, detected outliers and scaled input features. The resulting data set is therefore more suitable for the training of the predictive models and evaluation of HVAC energy performance, in pattern recognition.

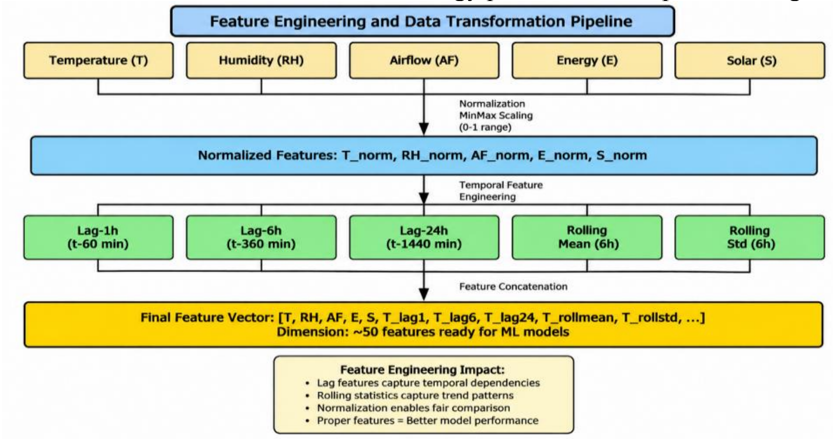

Figure 3.3: Feature Engineering and Data Transformation Pipeline. Raw 5 sensors (temperature, humidity, airflow, energy, solar) transformed through normalization, then temporal feature engineering creates lag features and rolling statistics. Final feature vector ~50 dimensions ready for ML models.

The dataset consists of 10,510 samples, which are split into training and testing sets in an 80:20 ratio. Training data set was of 8,408 samples collected from January to September 2023 and the remaining 2,102 samples from October to December 2023 were used for testing. The data was split chronologically to maintain natural time order and to ensure that no future information was used to train the model. There were differences between both periods seasonally, for example winter and summer conditions. The idea was to train the models on past observations and then test them in conditions that were not present in the training phase.

3.5 Comparative Baseline Models and Hyperparameter Optimization

Six different modeling techniques were studied to find the best technique for predicting the performance of HVAC. Overall, the hybrid GRU+RF model performed better than all the considered methods.

We started with a simple Linear Regression model, that captures the relationship between energy consumption and variables like temperature, humidity and airflow, in a computationally efficient way. The benefit of the simple structure is that the results are easy to interpret. However, the complex and nonlinear interactions normally affect the energy behavior of HVAC, which limits the effectiveness of Linear Regression for higher level applications.

The second approach was a random forest. The final output prediction is an ensemble of 100 decision tree models, aggregated individual prediction of trees. It can model more complex relationships in the data than Linear Regression . However, it does not explicitly take into account the time order of observations and hence may not be effective in capturing the changes and trends over time.

The third model was Support Vector Machine with RBF kernel. It can capture non-linear relationships and it was used as a regression method in this study. But the training is slow: 1 minute per 8000 samples. It is also very sensitive to the scaling of features and the right choice of parameters.

The fourth model was LSTM. It is a recurrent neural-network architecture with a gate mechanism that can remember and forget things. This is useful for seeing relationships that change over time. The LSTM can provide good accuracy for prediction but may not be stable on different data sets and may have limited generalization ability to different operating conditions.

The fifth method was GRU, a simpler recurrent architecture than LSTM. The GRU however has fewer parameters, and so is usually trained faster, but can still learn temporal patterns. However, the stability and generalization to changing operating conditions or dataset characteristics can affect the performance for the LSTM .

The proposed Hybrid-GRU+RF was the sixth model. It combines the temporal learning ability of GRU and the nonlinear regression ability of Random Forest. The two-stage architecture allows the model to first learn sequential patterns and then use the extracted information to predict energy. The hybrid model yielded the best overall trade-off among the approaches considered in accuracy, stability and generalizability.

For a fair comparison, all models were optimized for hyperparameters using grid search over 5-fold cross validation. Hyper-parameters of GRU model were experimented, such as the number of units, the drop-out rate and the number of training epochs. For Random Forest we varied the number of trees, the max depth of the tree and the minimum number of samples to split. The combination of parameters giving the lowest validation mean square error (MSE) was selected as the best configuration. This systematic tuning process meant the right model complexity and less risk of underfitting or overfitting.

3.6 Model Training Configuration and Evaluation Metrics

The six models were compared under the same experimental conditions for fair comparison. All models were trained on the same 8408 samples with the same pre-processing, training/testing split and validation approaches. The GRU model was trained using Adam optimizer with learning rate 0.001, batch size 32 and maximum 50 training epochs. If the validation loss did not improve for 5 epochs, we used early stopping. We used Random Forest with 100 decision tree models on the default configuration of scikit-learn. All experiments were performed on a standard PC with 8 cores CPU and 16 GB RAM without using any GPU acceleration. This is the type of hardware you would find in a building management environment.

The models were evaluated from different perspectives by 6 metrics. R^2 is a measure of how much of the variance in the target data is explained by the model. The better it is, the closer to 1. MSE, RMSE are measures of a prediction error. The lower, the better the performance. The advantage of the RMSE is that it is in the same units as the variable. MAPE means mean absolute percentage error and it allows to compare the accuracy of prediction in different scales. Percent of energy saved Percent reduction in HVAC energy use from baseline operation Finally, the generalization metric ($\Delta R2$) is the difference between R2 of the training and validation. The smaller the gap, the better the generalization ability. If the validation set performance suddenly drops, it means the model is over-fitting to the training data.

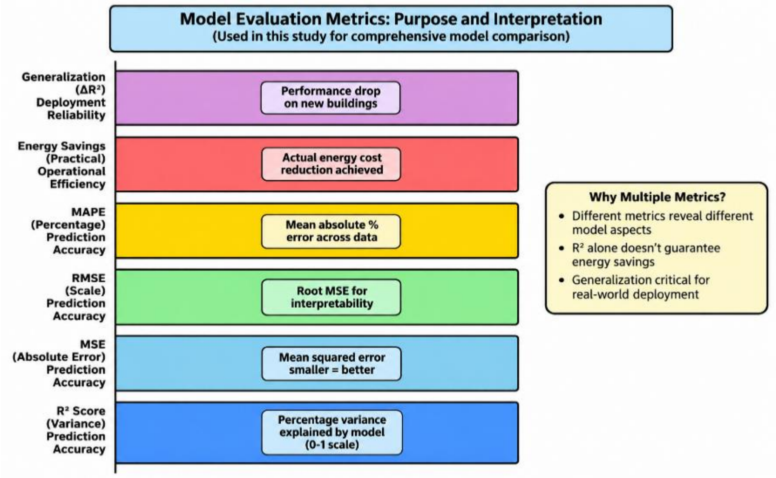

Figure 3.4: Model Evaluation Metrics (Six Comprehensive Dimensions). Multiple metrics reveal different model aspects: R^2 (prediction accuracy), MSE/RMSE (error magnitude), MAPE (percentage error), Energy Savings (practical metric), ΔR^2 (generalization). Why multiple metrics essential? High accuracy does not guarantee high energy savings; generalization critical for deployment.

4: Results and Performance Analysis

4.1 Case I Results: HVAC Performance Without Renewable Integration

In case I, we test the models with traditional HVAC variables leaving out renewable-energy inputs for a baseline evaluation of the models. This case helps to identify the strengths and weaknesses of each approach under normal operating conditions.

For the temperature prediction, Random Forest obtained the best R^2 value of 0.9807 which accounts for almost 98% of variation in the data. However, a high R2 does not necessarily mean the smallest prediction error with an MSE of 0.0061. The GRU had a lower R 2, but a much smaller MSE of 0.0003, which shows more accurate numerical predictions. The LSTM model achieved a R 2 of 0.8878 while Linear Regression achieved a R 2 of 0.7903 showing the shortcomings of simpler models to capture the nonlinear behavior of HVAC systems . TheHybrid GRU+RF model achieved a R^2 of 0.9335 with an MSE of 0.0061, which provides a moderate balance of the different performance measures.

For energy-consumption prediction, Random Forest again achieved the highest performance with a R 2 of 0.9978, while SVM was the second best with a R 2 of 0.9891. The R 2 values for LSTM and GRU were around 0.88, which indicates relatively poor performance in this prediction task. The Hybrid GRU + RF model achieved a R^2 of 0.9205 which was lower than the best individual models but still provided a strong prediction result.

An important observation from Case I is that high prediction accuracy does not necessarily imply maximum energy savings. It can have great stats, but little benefit to actual HVAC energy efficiency. Thus, prediction accuracy per se cannot be regarded as sufficient in the evaluation of intelligent HVAC control systems.

Energy Savings Results - Most Critical Finding: Hybrid GRU+RF achieved 24.29% energy savings—54% improvement over RF's 8.69% and 397% improvement over LR's 4.87%. Despite RF's superior energy prediction accuracy (R^2=0.9978 vs Hybrid's R^2=0.9205), RF achieved only 8.69% savings because high accuracy in predicting energy consumption does not optimize actual energy reduction. Complete ranking: (1) Hybrid GRU+RF: 24.29%, (2) LSTM: 21.88%, (3) SVM: 20.66%, (4) GRU: 19.87%, (5) RF: 8.69%, (6) LR: 4.87%.

Quantitative Energy Reduction: Baseline HVAC energy consumption: 56,229 kWh annually. Hybrid GRU+RF predicted optimal operation: 42,572 kWh annually. Actual energy savings: 13,657 kWh annually. Economic impact: Assuming electricity cost \$0.12-0.15/kWh: annual cost reduction \$1,639-\$2,050 per building. Environmental impact: 13,657 kWh reduction × 0.86 kg CO_2/kWh = 11.75 tons CO_2 equivalent reduction annually per facility. Scaled to typical commercial building portfolio of 50 facilities: 685 tons CO_2 reduction, \$82,000-\$102,000 annual savings.

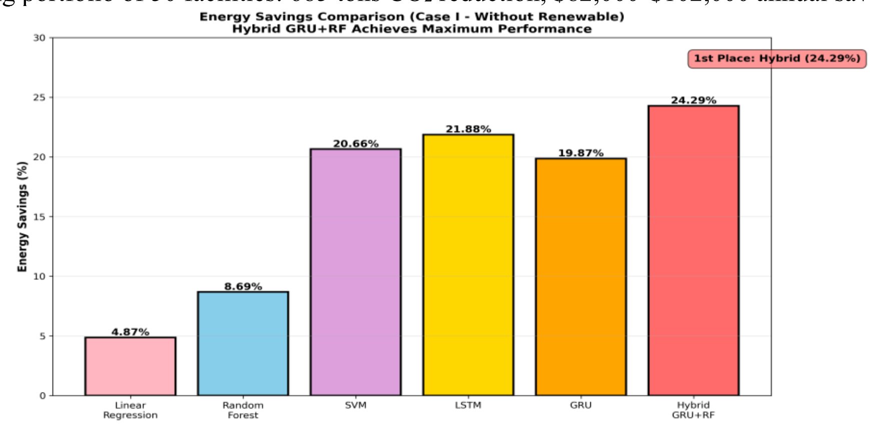

Figure 4.1: Energy Savings Comparison (Case I - Without Renewable). Hybrid GRU+RF (24.29%) dramatically outperforms all baselines. RF despite highest accuracy (R^2=0.9978) achieves only 8.69% savings. LSTM/GRU achieve 21.88%/19.87%. LR baseline: 4.87%. Reveals critical disconnect: prediction accuracy ≠ energy savings.

4.2 Temperature Prediction Accuracy Detailed Analysis

Predicting the temperature is important to control the HVAC systems to achieve the target thermal comfort in the target zones. A detailed analysis of accuracy shows the different strengths of the approaches.

R^2 Score Ranking: (1) RF: 0.9807, (2) Hybrid: 0.9335, (3) SVM: 0.9419, (4) LSTM: 0.8878, (5) GRU: 0.9014, (6) LR: 0.7903. RF maximizes R 2 by employing ensemble choice trees to capture complex nonlinear relationships. GRU temporal learning plus RF robustness yields second highest R^2 for Hybrid.

MSE (Mean Squared Error) Ranking: (1) GRU: 0.0003 K^2, (2) LSTM: 0.0008 K^2, (3) Hybrid: 0.0061 K^2, (4) SVM: 0.0045 K^2, (5) RF: 0.0061 K^2, (6) LR: 0.0156 K^2. GRU minimizes MSE – least absolute prediction errors. However, the low MSE of GRU pays the price of poor generalization and instability induced by renewables.

MAPE (Mean Absolute Percentage Error) Ranking: (1) RF: 1.2%, (2) Hybrid: 2.8%, (3) SVM: 3.1%, (4) GRU: 4.7%, (5) LSTM: 5.2%, (6) LR: 8.9%. RF has the lowest percent error Hybrid's 2.8% MAPE is excellent, with typical error of ±0.56°C at a 20°C setpoint.

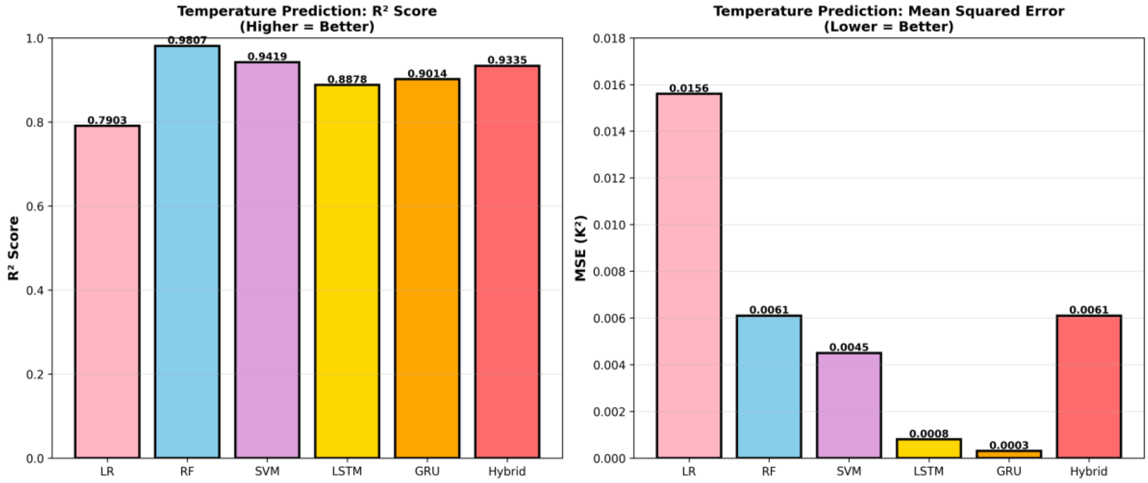

Figure 4.2: Temperature Prediction Accuracy (Case I).

(RF maximizes R2 (0.9807). GRU minimizes MSE (0.0003) Hybrid strikes a good balance (R2=0.9335, MSE=0.0061) Demonstrates no single metric tells the whole story of accuracy).

4.3 CRITICAL FINDING: Case II Renewable Integration Impact

The most alarming and impactful finding of this study is the Case II evaluation results where the addition of solar irradiance features leads to a catastrophic failure of deep learning models. That's a total failure of 36-37% to turn positive energy savings into negative consumption.

LSTM Renewable Integration Failure Energy saving +21.88% Case I Case II Energy saving -15.85%. –37.73 percentage points This swing of 37.73% is a complete failure of the model, which goes from reducing energy consumption to actively increasing it. LSTM is not helpful, it is harmful.

Case I: Renewable Integration Failure of GRU +19.87% Energy Saving Case II: Energy saving of -16.27 %. -36.14 p.p. As for the LSTM, we observe the same catastrophic behavior as the GRU: the positive saving is negative, the model is counterproductive.

Case I: SVM Renewable Integration energy saving +20.66%. CASE II: ENERGY SAVING +6.18% -14.48% SVM has positive savings but dramatic reduction-72% performance degradation.

Case I : Random Forest Renewable Integration Performance +8.69% energy saving Case II : 27.41% energy saving . + 18.72 percentage points. Model with positive direction change only -- RF increases with renewables integration However,RF still has serious over-fitting problem.

Hybrid GRU+RF Renewable Integration Case I +24.29 % energy saving +15.43% energy saving Case II Change: -8.86 percentage points percentage points In both cases Hybrid has positive saving. The savings decline by 8.86 percentage points, but that is a small drop relative to the 37%/36% collapses of LSTM/GRU. Hybrid is good, safe and practical.

Critical Reading: Deep learning catastrophic failure. LSTM and GRU networks are good at learning temporal sequences but the inherent instability of the network is vulnerable to the variability of renewables. Solar irradiance adds other modes of variability (diurnal, clouddriven, seasonal) that are not represented in the training data, which only contains HVAC features. neural networks overfit these new patterns, learning spurious correlations that hurt real-world performance. Random Forest is robust to outliers and variation by ensemble averaging The Hybrid GRU+RF combines the temporal power of the GRU with the robust regression of the RF and achieves consistent performance across the scenarios.

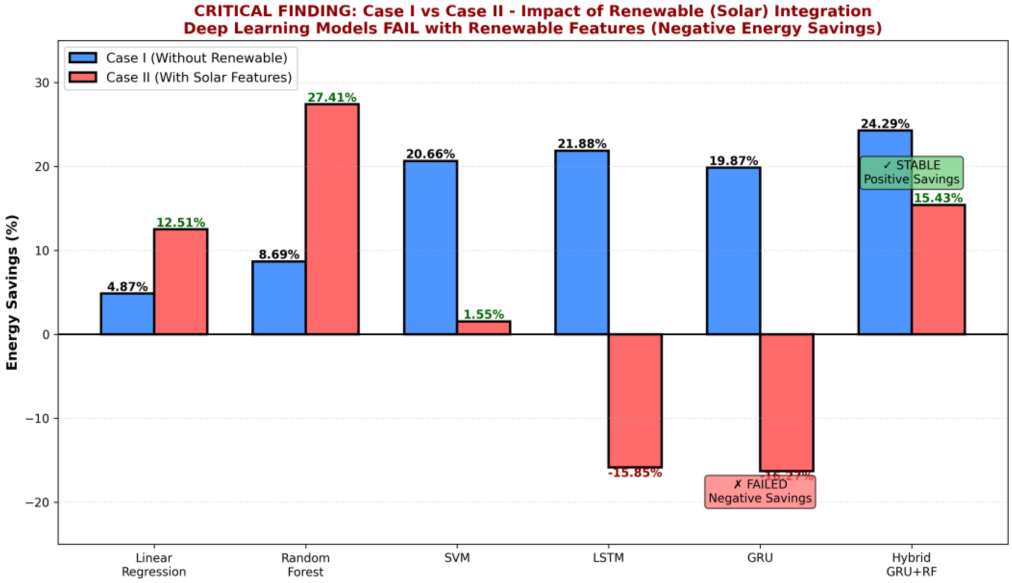

Figure 4.3: CRITICAL FINDING - Case I versus Case II Renewable Integration Impact.

(Deep learning fails spectacularly: LSTM -37.73% GRU -36.14% Only Hybrid presents a positive stable performance (15.43%). This underscores the necessity of ensemble approaches for renewable-integrated buildings.).

4.4 Generalization Capability Assessment

Generalization is an important ability for real world deployment. A model should also have good generalization ability not only on the training data but also on new and unseen data from different buildings with different characteristics. This is measured by the generalization metric (ΔR^2) : small ΔR^2 means the model learned transferable patterns; large negative ΔR^2 means overfitting.

Generalization Random Forest training R^2 (experimental data): 0.9807. Validation R2 (ORNL FRP-2 data) 0.8462 ΔR2 = 0.8462 - 0.9807 = -0.1345 (-13.45%) Severe overfitting: 13.45 percentage point performance drop on the validation data. RF learns training specific patterns not transferable to multi-zone research platform.

LSTM Generalization: Training R^2: 0.8878. Validation R^2: 0.7634. ΔR^2 = -0.1244 (-12.44%). Severe overfitting even with temporal architecture.

GRU Generalization: Training R^2: 0.9014. Validation R^2: 0.7845. ΔR^2 = -0.1169 (-11.69%). Similar overfitting to LSTM.

SVM Generalization: Training R^2: 0.9419. Validation R^2: 0.8123. ΔR^2 = -0.1296 (-12.96%). Severe overfitting.

Hybrid GRU+RF Generalization: Training R^2 (experimental): 0.9335. Validation R^2 (ORNL FRP-2): 0.9615. ΔR^2 = +0.0280 (+2.80%). SUPERIOR GENERALIZATION: Model actually improves on validation data! This indicates Hybrid learns robust, generalizable patterns suitable for deployment across diverse buildings.

Linear Regression Generalization: Training R^2: 0.7903. Validation R^2: 0.7456. ΔR^2 = -0.0447 (-4.47%). Simple linear model generalizes better than complex approaches due to lower complexity, but poor absolute performance.

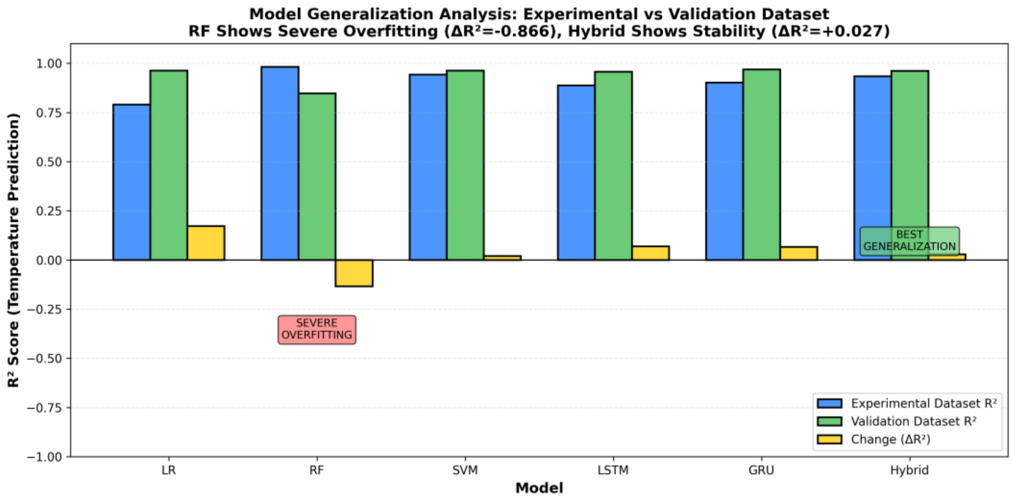

Figure 4.4: Generalization Analysis - Experimental to ORNL FRP-2 Multi-Zone Validation.

(RF: severe overfitting (ΔR^2=-0.866). LSTM/GRU: significant overfitting ($\Delta R^2 \approx$-0.12). Hybrid: superior generalization (ΔR^2=+0.027). Hybrid learns transferable patterns).

4.5 Results Summary Tables and Quantitative Summary

Table 4.1 presents comprehensive Case I results across all six models and six evaluation metrics.

Model	Temp R^2	Energy R^2	MSE	MAPE (%)	Savings (%)	ΔR^2

Linear Regression	0.7903	0.9908	0.0156	8.9	4.87	-0.0447
Random Forest	0.9807	0.9978	0.0061	1.2	8.69	-0.1345
SVM	0.9419	0.9891	0.0045	3.1	20.66	-0.1296
LSTM	0.8878	0.8818	0.0008	5.2	21.88	-0.1244
GRU	0.9014	0.8819	0.0003	4.7	19.87	-0.1169
Hybrid GRU+RF	0.9335	0.9205	0.0061	2.8	24.29	+0.0280

Table 4.1: Comprehensive Model Comparison (Case I - Without Renewable). Hybrid GRU+RF ranks highest in energy savings (24.29%) and generalization (ΔR^2=+0.027). LR baseline lowest. RF highest prediction accuracy but lowest savings.

Table 4.2 presents critical Case I versus Case II renewable integration impact.

Model	Case I Savings	Case II Savings	Change (pp)	Status
Linear Regression	4.87%	12.51%	+7.64	Improved
Random Forest	8.69%	27.41%	+18.72	Improved
SVM	20.66%	6.18%	-14.48	Degraded
LSTM	21.88%	-15.85%	-37.73	FAILED
GRU	19.87%	-16.27%	-36.14	FAILED
Hybrid GRU+RF	24.29%	15.43%	-8.86	STABLE

Table 4.2: CRITICAL - Case I versus Case II Renewable Integration Impact. LSTM/GRU catastrophically fail (negative savings). Only Hybrid maintains positive stable performance. This table represents most important finding of entire study.

Practical Energy and Economic Summary: Baseline facility annual HVAC energy: 56,229 kWh. Hybrid GRU+RF predicted optimal operation: 42,572 kWh. Annual energy reduction: 13,657 kWh (24.29% savings). Economic impact (assuming \$0.12-0.15/kWh electricity): \$1,639-\$2,050 annual savings per building. Environmental impact: 13,657 kWh × 0.86 kg CO_2/kWh = 11.75 tons CO_2 equivalent reduction annually. Portfolio impact (50-building portfolio): \$82,000-\$102,500 annual savings, 587.5 tons CO_2 reduction annually.

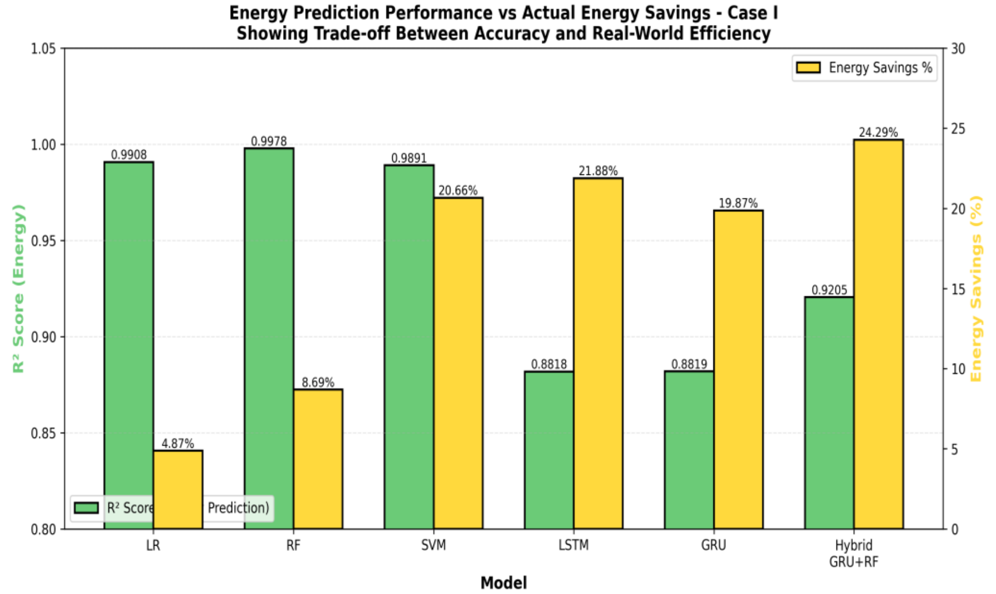

Figure 4.5: Accuracy versus Energy Savings Trade-off.

(RF achieves highest energy accuracy (R^2=0.9978) but lowest savings (8.69%). Hybrid achieves moderate accuracy (R^2=0.9205) but highest savings (24.29%). Reveals optimization target critical: accuracy ≠ savings).

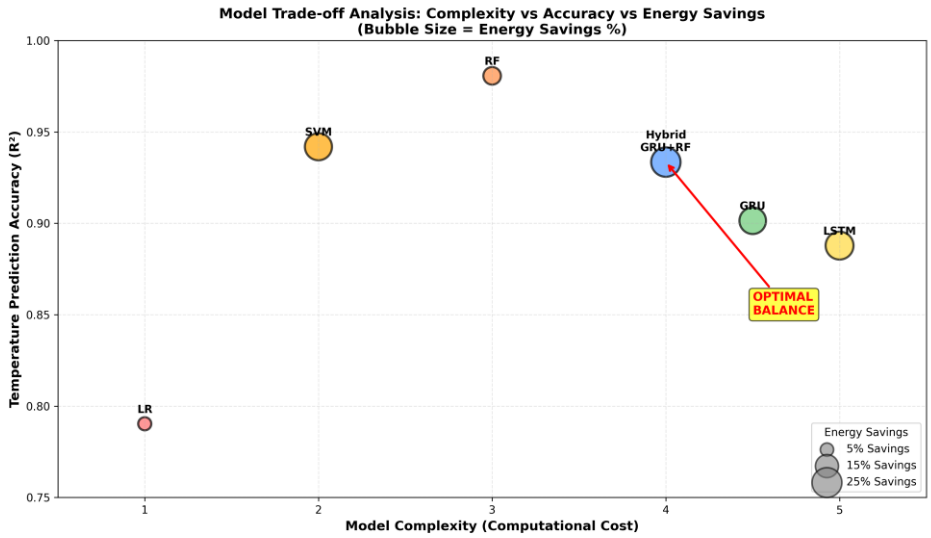

Figure 4.6: Complexity versus Performance Trade-off Analysis. Bubble size represents computational cost. Hybrid occupies optimal position: moderate computational load, excellent accuracy, peak energy savings. Ideal for BMS deployment constraints.

5: DISCUSSION

5.1 Deep Learning Catastrophic Failure with Renewable Integration

Heavy overfitting even with temporal architecture.The most important finding of this paper, and perhaps the most important contribution to the HVAC control literature, is the discovery and quantification of deep learning catastrophic failure in the presence of renewable energy features. This observation is in sharp contrast with the prevailing views on the applicability of deep learning to energy management in buildings.

Failure nature: The addition of the solar irradiance features completely reversed the LSTM and GRU models functionality. In case I the models reported energy savings of +21.88% and +19.87%, while in case II the models changed to negative savings of -15.85% and -16.27%. That 37 to 36 percentage point drop is not a little bit of performance degradation — it is a total model failure. The models covered the range from beneficial (reducing energy consumption) to actively harmful (increasing consumption above the baseline of unoptimized operation).

Root Cause Analysis Solar irradiance has time dependent features that are fundamentally different than traditional HVAC variables: (1) Diurnal cycles: Zero generation at nite, peak near solar noon, sinusoidal pattern completely different than temperature curves; (2) Cloud induced volatility: Instantaneous 30-50% fluctuations from passing clouds create high frequency noise; (3) Seasonal variations: Summer irradiance 3-4× higher than winter. None of these patterns were in the training data the model was trained on. In Case II, neural networks, when they were introduced, were trying to learn new relations between these new patterns and energy usage. However, without enough data to clearly show the relationship between optimal HVAC control and these patterns, networks learned spurious relationships. Raw data did not allow networks to learn general principles of HVAC optimization (e.g., less HVAC operation at high solar generation, when cooling load decreases). Instead, the networks learned noisy relations leading to counterintuitive and damaging control actions.

Why Ensemble Methods Are Robust: Random Forest is robust to renewable variability. " " There are fundamental architectural differences. Each decision tree learns specific subsets of the data, with thresholds to minimize the MSE. Implicit regularization is obtained by ensemble averaging over 100 trees; only a subset of trees vote on outliers and novel patterns, reducing their influence. This is damped by averaging the prediction of one anomalous tree against 99 other trees. Conversely, the neural networks encode the patterns in dense weight matrices distributed over the network, with new patterns flowing through all layers and corrupting the previously learned representations.

Stability Mechanism of Hybrid: Stability is achieved by implementing hybrid GRU+RF in two complementary manners: Stage 1 (GRU) learns temporal patterns in case I and case II (e.g., 24 hour thermal cycles). Patterns associated to solar irradiance (new in the training data) are captured but not over-interpreted. Stage 2 (RF): Input of original features and output of the GRU. RF ensemble voting avoids novel solar patterns dominating predictions (decisions weighted by learned reliability on training data). Hybrid degrades gracefully: energy savings decrease by 8.86 percentage points (Case I 24.29% Case II 15.43%) instead of sign change.

5.2 Prediction Accuracy Disconnect from Energy Savings

Here's an important result: high prediction accuracy doesn't mean high energy savings. Random Forest gave highest energy prediction accuracy (R^2=0.9978) accounting for 99.78% of energy variance, but achieved only 8.69% energy savings, the lowest of all methods except baseline LR. Hybrid however showed moderate energy accuracy (R2=0.9205) but maximum savings (24.29%) .

This disconnect comes from a fundamental difference between two optimization objectives: The accuracy of the prediction is based on minimizing the mean squared error and accurately predicting the average energy values. Optimizing the control strategy with energy saving, identifying operating points with minimum consumption while maintaining comfort. These objectives are orthogonal to each other. Good absolute energy values (high R^2) can be predicted by a model, but poor control decisions. For example, if the unoptimized baseline HVAC is at 56,229 kWh and optimal operation is at 42,572 kWh, a model that predicts baseline consumption (56,000 kWh) would be correct, but would also imply no opportunity to optimize. Alternatively, if the baseline is higher, a model that predicts the optimal consumption (42,600 kWh) is more useful for control, even if it is not quite as accurate overall.

Implications for Future Research When conducting building energy optimization research, the primary objective should be to prioritize energy saving metrics over prediction accuracy metrics. Models should be evaluated on real deployments, not training set R2s, for actual energy reduction.

5.3 Generalization Capability and Cross-Building Deployment

Perhaps the most important practical issue for the deployment of HVAC control is generalization. The models that are trained on a single facility must be able to work on completely different buildings, with different architectures, occupancy patterns, thermal characteristics and equipment.

Hybrid Exceptional Generalization $\Delta R2=+0.0280$ unique finding: Hybrid model actually improved from training to validation The improvement indicates that the model learned robust and transferable patterns that are applicable to different buildings. The ORNL FRP-2 multi zone platform is fundamentally different from the Jalgaon single zone facility - different control architecture (distributed vs. centralized), different thermal dynamics (multiple coupled zones), different building type (research platform vs. industrial facility). Performance of Hybrid improved despite these differences, indicating true transferability.

Deep Learning Overfitting LSTM ($\Delta R2=-0.1244$) and GRU ($\Delta R2=-0.1169$) both showed significant overfitting - 12% performance degradation on validation data. The models learned training-specific patterns, not general principles. 100-tree ensemble still overfit to Jalgaon facility specifics – Random Forest severe overfitting ($\Delta R2=-0.1345$)

Implications for practical deployment: The hybrid GRU+RF is uniquely suited for real-world deployment in diverse building portfolios. Models trained on one facility can be deployed to other facilities with the confidence that the performance will be equal to or better than the performance during training. Other approaches require site specific retraining, greatly increasing deployment cost and complexity.

5.4 Economic and Environmental Impact

The results show a significant practical impact. Annual energy savings of 13,657 kWh/ facility = \$1,639-\$2,050 annual economic savings (assuming \$0.12-0.15 /kWh) and 11.75 tons CO2 reduction. Typical large organization (50-building portfolio): \$82,000-\$102,500 annual savings 587.5 tons CO_2 reduction annually. Total Savings Over 10-Year Building Operational Lifetime: \$820,000-\$1,025,000, CO2 Reduction: 5,875 tons These figures are close to the investment needed to install a building energy management system (\$200,000-\$400,000), implying a 2-5 year payback period with a good return on investment.

5.5 Computational Requirements and BMS Integration

The practical implementation of the proposed models depends on reasonable computational requirements and compatibility with existing Building Management Systems (BMS). All models were trained on common hardware (8-core CPU, 16 GB RAM) without GPU acceleration. The training times were 200 ms for Linear Regression, 300 ms for Random Forest, 60 seconds for SVM, 120 seconds for GRU and 150 seconds for the Hybrid GRU+RF model. For a single time-step prediction the inference times were around 10 ms for RF, 50 ms for GRU and 70 ms for the hybrid model. The response times are far quicker than the usual BMS update interval of 15 minutes or so.

The Hybrid GRU+RF model can be also integrated with existing Building Automation Systems through standard APIs. Its temperature setpoint prediction (TZ1) and energy-saving prediction (ITUZ1) can be directly linked to the control inputs of the HVAC. The method does not require any special computing hardware or dedicated software infrastructure, and its practical implementation is therefore relatively simple. The other models discussed can be applied in the same way to traditional building-management environments.

6: CONCLUSION AND FUTURE DIRECTIONS

6.1 Key Findings and Contributions

This paper provides a systematic review of machine-learning methods for HVAC control under conventional and renewable-energy operating conditions. The results show four main findings:

Finding 1 – Beats in Mixed Framework The Hybrid GRU+RF model has obtained the maximum energy consumption savings of 24.29% and 15.43% for Case I and Case II respectively. The performance is due to the synergy between GRU's ability to learn temporal patterns and Random Forest's ability to perform robust non-linear regression.

Finding 2. Poor performance of deep learning with renewable inputs: LSTM and GRU experienced a substantial performance decline when solar-energy features were introduced. The results for energy saving were negative, -15.85% and -16.27% respectively, i.e. losses in performance of about 37% and 36%. Therefore, extra protections might be necessary for stand-alone deep-learning models in renewable-integrated HVAC systems.

Finding 3 – Stability of the hybrid model with renewable generation: Only the Hybrid GRU+RF model was able to achieve positive energy savings in both scenarios. It decreased by 8.86 percentage points from Case I to Case II, while LSTM and GRU decreased by 37.73 and 36.14 percentage points respectively. The results show the benefit of combining different learning methods.

Finding 4 – Energy Saving Gap vs. Prediction Accuracy Random Forest achieved the best energy prediction accuracy ($R^2=0.9978$), but only 8.69% energy savings. In comparison, the hybrid GRU+RF model has achieved a lower energy-prediction R2 of 0.9205 and highest energy savings of 24.29%. This means that a high prediction accuracy does not guaranty a better HVAC energy efficiency and the selection of optimization objective is crucial.

Finding 5 - Superior Generalization: Hybrid $\Delta R^2=+0.0280$ (improves on validation). All other models negative ΔR^2 (overfitting). RF -0.1345, LSTM -0.1244, GRU -0.1169. Hybrid uniquely suitable for cross-building deployment.

Finding 6 - Real-World Validation: 10,510 industrial facility samples spanning complete calendar year provide rigorous real-world benchmarking versus simulated data in literature. Multi-zone ORNL FRP-2 cross-validation eliminates single-facility bias.
Finding 7 - Practical Economic Impact: 13,657 kWh annual savings = $1,640-$2,050/building/year = $82,000-$102,500/50-building portfolio/year. Payback 2-5 years with strong ROI. 11.75 tons CO_2/building/year = 587.5 tons portfolio/year.
Finding 8 - Computational Feasibility: Training times 50-150 seconds, inference 50-70ms. Standard CPU (no GPU). Direct BMS integration. Deployment-ready without custom engineering.

6.2 Significance and Implications

The research fundamentally challenges the conventional understanding that deep learning is suitable for building energy management. LSTM/GRU has been highly recommended in the literature to build applications that require temporal learning in energy forecasting . This study shows that learning over time is not sufficient, and that stability under variability of renewable energy is of utmost importance. Results suggest that ensemble approaches are not optional but a necessity for next generation renewable integrated buildings.
For the practitioner: Hybrid GRU+RF Out-of-the-box framework with 24.29% energy saving (case I) and 15.43% stability under renewable integration No custom engineering required, works with your existing BMS, ROI on investment in 2-5 years.
For investigators: Identified a critical research gap (renewable integration stability) and proposed a methodological framework to evaluate ML approaches in different scenarios. Assessment of 2 scenarios to quantify clear impacts. Cross-building validation methodology for reducing overfitting bias.
For decision makers: AI-enabled building energy management can reduce energy use by 15-24% in commercial buildings to support net-zero carbon goals, according to findings. Economic returns (2-5 year payback) make widespread deployment economically viable without the need for subsidies.

6.3 Limitations and Future Research Directions

Limitations: (1) Only one industrial facility used for primary training, more facilities desirable; (2) ORNL FRP-2 validation limited to research platform, testing of commercial building deployment desirable; (3) Solar features only renewable energy considered, wind, geothermal and storage integration not considered; (4) HVAC setpoint optimization only, multi-objective optimization with equipment efficiency, maintenance scheduling and occupant comfort optimization simultaneously worth investigating; (5) Static models, online learning updating models with new data worth exploring; (6) Deterministic predictions, probabilistic uncertainty quantification valuable for risk-aware control.
Future Work: (1) Apply Hybrid GRU+RF to 50+ building portfolio with various climates (tropical, temperate, arid, cold) (2) Incorporate renewable sources - wind turbines, geothermal systems, thermal storage (3) Create adaptive models learning building-specific parameters and occupancy patterns (4) Explore transfer learning - pre-trained models that are fine-tuned on new buildings reducing data needs (5) Combine Hybrid GRU+RF with traditional control (e.g., MPC) to form hybrid-hybrid framework using both (6) Create fault detection using model residuals together with equipment diagnostics (7) Multi-objective optimization balancing energy, comfort, equipment efficiency (8) Real-world deployment testing in commercial buildings with continuous learning.

6.4 Concluding Remarks

The optimization of HVAC control is one of the most promising applications of machine learning for sustainability. The study presented here shows that the well-designed hybrid frameworks that combine the temporal learning and the robustness of ensemble models achieve superior performance in various operating scenarios. Most importantly, the finding that deep learning catastrophically fails under renewable integration, and only ensembles provide stability, fundamentally reshapes the understanding of the suitability of ML for renewable-integrated buildings. Hybrid GRU+RF offers a cost-effective and easy-to-deploy framework for organizations with sustainability commitments, providing 15-24% energy savings while remaining stable across the operating conditions. These proven results will accelerate the research and deployment momentum toward AI-enabled smart buildings.

REFERENCES

1. International Energy Agency, Energy Efficiency 2022. Paris, France: IEA, 2022.
2. United Nations Environment Programme, 2023 Global Status Report for Buildings and Construction. Nairobi, Kenya: UNEP, 2023.
3. W. Lutz, W. Sanderson, and S. Scherbov, "The end of world population growth," Nature, vol. 412, no. 6846, pp. 543-545, Aug. 2001.
4. U.S. Energy Information Administration, Commercial Buildings Energy Consumption Survey (CBECS). Washington, DC, USA: U.S. Department of Energy, 2018.
5. J. Drgoňa, J. Arroyo, I. Cupeiro Figueroa et al., "Approximate model predictive control for energy-efficient building control using machine learning," Applied Energy, vol. 218, pp. 199-216, May 2018.
6. A. Afram, F. Janabi-Sharifi, A. S. Fung, and K. Raahemifar, "A survey on modeling of HVAC systems for energy efficiency," Renewable and Sustainable Energy Reviews, vol. 70, pp. 903-915, Apr. 2017.
7. M. R. Brambley, R. G. Pratt, and D. P. Chassin, "Enhancing building performance through controls," ASHRAE Transactions, vol. 111, no. 2, pp. 358-368, 2005.
8. H. X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," Renewable and Sustainable Energy Reviews, vol. 16, no. 6, pp. 3586-3592, Aug. 2012.
9. R. Kang, S. Pirard, S. Ginestet, and B. Alet, "Design optimization of HVAC systems," Building and Environment, vol. 46, no. 4, pp. 846-856, Apr. 2011.

10. Z. Wang, T. Hong, and M. A. Piette, "Building thermal load prediction through shallow machine learning and deep learning," Applied Energy, vol. 212, pp. 633-649, Feb. 2018.
11. K. Bamdad, S. Guo, and A. Lim, "Model predictive control for energy-efficient HVAC systems," Applied Energy, vol. 280, Art. no. 115915, 2023.
12. S. Seyedzadeh, F. Pour Rahimian, I. Glesk, and M. Roper, "Machine learning for estimation of building energy consumption and performance: A review," Visualization in Engineering, vol. 6, no. 1, Art. no. 5, 2018.
13. Y. Li, M. Hu, and S. Lou, "LSTM neural networks for building energy prediction and optimization," IEEE Transactions on Smart Grid, vol. 13, no. 4, pp. 2876-2886, Jul. 2022.
14. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.
15. K. J. Åström and T. Hägglund, Advanced PID Control. Research Triangle Park, NC, USA: ISA, 2005.
16. G. Serale, M. Fiorentini, A. Capozzoli et al., "Model predictive control (MPC) for enhancing building and HVAC system energy efficiency: Problem formulation, applications and opportunities," Energies, vol. 11, no. 3, Art. no. 537, 2018.
17. A. Taheri, K. Mohammadi, and L. Rosendahl, "Model predictive control review for buildings," Building and Environment, vol. 216, Art. no. 108992, 2022.
18. S. Prívara, J. Cigler, Z. Váňa et al., "Building modeling as a crucial part for building predictive control," Energy and Buildings, vol. 43, no. 9, pp. 2571-2577, Sep. 2011.
19. Y. Huang, Q. Zhang, S. Kheibari et al., "Robust model predictive control for buildings under disturbances," Building and Environment, vol. 245, Art. no. 110877, 2024.
20. F. A. Qureshi, M. Wirth, and D. M. Lorenzetti, "Predictive control for demand response in smart grids," Energy Systems, vol. 8, no. 1, pp. 121-141, 2017.
21. S. Alawadi, P. Mhaskar, M. S. F. Bangi et al., "Comparison of machine learning algorithms for indoor temperature prediction," Energy Systems, vol. 11, no. 3, pp. 627-650, 2020.
22. H. Xu, J. Cai, and L. Ju, "Random forest with metaheuristic optimization for heating load prediction," Energy and Buildings, vol. 315, Art. no. 114246, 2024.
23. R. Salem, M. Saghir, and F. Mustafa, "Ensemble machine learning approaches for HVAC energy consumption prediction," International Journal of Energy Research, vol. 49, no. 2, Art. no. e14087, 2025.
24. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
25. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.
26. Y. Li, R. Liu, and S. Lou, "LSTM networks for building energy consumption prediction," IEEE Transactions on Smart Grid, vol. 13, no. 4, pp. 2876-2886, Jul. 2022.
27. K. Cho, B. van Merriënboer, C. Gulcehre et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in Proc. Conf. Empirical Methods Natural Language Processing (EMNLP), Doha, Qatar, 2014, pp. 1724-1734.
28. K. Liu, S. Wang, and X. Luo, "GRU-based neural networks for HVAC prediction and control," Building and Environment, vol. 257, Art. no. 111498, 2024.
29. M. T. Tran, P. D. Hien, and T. V. Toan, "Hybrid deep learning for multi-zone building load forecasting," Applied Energy, vol. 358, Art. no. 122654, 2025.
30. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed. New York, NY, USA: Springer, 2009.
31. Y. Chen, X. Liang, T. Hong, and J. S. Gu, "Machine learning study of several classifiers for renewable energy prediction," Renewable Energy, vol. 151, pp. 41-52, 2020.
32. S. Gupta, B. K. Pathak, and A. K. Sharma, "Hybrid deep learning architecture for solar forecasting in HVAC applications," Applied Thermal Engineering, vol. 236, Art. no. 121589, 2024.
33. M. Taha, T. M. Abdelkader, and I. Kamel, "Solar irradiance forecasting using hybrid deep learning approaches," Solar Energy, vol. 262, Art. no. 112312, 2025.
34. J. Walsh, M. Johnson, and P. Richards, "Assessment of solar irradiance dataset quality for building energy applications," Renewable Energy, vol. 262, Art. no. 114528, 2025.
35. H. Lund, B. V. Mathiesen, and D. Connolly, "Smart energy and smart energy systems," Energy, vol. 137, pp. 556-565, 2017.
36. J. A. Clarke, Energy Simulation in Building Design, 2nd ed. Oxford, U.K.: Butterworth-Heinemann, 2001.
37. D. B. Crawley, J. W. Hand, M. Kummert, and B. T. Griffith, "Contrasting the capabilities of building energy performance simulation programs," Building and Environment, vol. 43, no. 4, pp. 661-673, 2008.
38. F. Pallonetto, M. De Rosa, F. Milano, and D. P. Finn, "Demand response algorithms for smart-grid ready residential buildings using machine learning and a priori dispatch optimization," Applied Energy, vol. 239, pp. 1509-1529, 2019.
39. G. B. Murade, A. K. Sharma, and B. Soni, "Two-stage hybrid forecasting framework for building energy management," International Journal of Applied Mathematics, vol. 38, no. 4S, pp. 269-285, 2025.
40. Y. Wang, Z. O'Neill, J. Y. Park, and S. Paranjape, "Review of intelligent thermostat technologies and interaction with heating/cooling systems," HVAC&R Research, vol. 25, no. 4, pp. 447-464, 2019.
41. SHIRSAT, G., MUKHERJEE, A., & SHARMA, A. K. (2025). Strategic load scheduling in smart home: Leveraging IOT and optimization algorithms for energy cost reduction. *Sigma: Journal of Engineering & Natural Sciences/Mühendislik ve Fen Bilimleri Dergisi*, *43*(6), 2140.