Optimizing Comfort and Energy Efficiency: Smart HVAC Control with Reinforcement Learning and Time Series Forecasting

By

Abdullah

Fall 2021-MSEE (AI&AS)-00000362516

Department of Electrical Engineering

School of Electrical Engineering & Computer Science (SEECS)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan.

(February 2024)

Optimizing Comfort and Energy Efficiency: Smart HVAC Control with Reinforcement Learning and Time Series Forecasting

By

Abdullah

Fall 2021-MSEE (AI&AS)-00000362516

Supervisor

Dr. Muhammad Shahzad Younis

Department of Electrical Engineering

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Electrical Engineering (MS EE)

In

School of Electrical Engineering & Computer Science (SEECS) ,

National University of Sciences and Technology (NUST),

Islamabad, Pakistan.

(February 2024)

Approval

It is certified that the contents and form of the thesis entitled "Optimizing Comfort and Energy Efficiency: Smart HVAC Control with Reinforcement Learning and Time Series Forecasting" submitted by Abdullah . have been found satisfactory for the requirement of the degree

Advisor : Dr. Muhammad Shahzad Younis

Signature: ____________

Date: 06-Feb-2024

Committee Member 1:Dr. Wajid Mumtaz

Signature:

06-Feb-2024

Committee Member 2:Dr. Muhammad Yousif

Signature: ____________

Date: 06-Feb-2024

Signature: ____________

Date: ____________

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis entitled "Optimizing Comfort and Energy Efficiency: Smart HVAC Control with Reinforcement Learning and Time Series Forecasting" written by Abdullah ., (Registration No 00000362516), of SEECS has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/M Phil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: ______________________

Name of Advisor: Dr. Muhammad Shahzad Younis

Date: 06-Feb-2024

HoD/Associate Dean: ______________________

Date: 12/02/2024

Signature (Dean/Principal): ______________________

Dr. Muhammad Ajmal
Principal
NUST School of Electrical Engg & Computer Science
H-12, Islamabad

Date: 12 Feb, 2024

295

FORM TH-4

National University of Sciences & Technology

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: (Student Name & Reg. #) Abdullah . [00000362516]

Titled: Optimizing Comfort and Energy Efficiency: Smart HVAC Control with Reinforcement Learning and Time Series Forecasting

be accepted in partial fulfillment of the requirements for the award of Master of Science (Electrical Engineering) degree.

Examination Committee Members

1. Name: Wajid Mumtaz — Signature: ______ 28-Feb-2024 10:42 AM

2. Name: Muhammad Yousif — Signature: ______ 28-Feb-2024 10:42 AM

Supervisor's name: Muhammad Shahzad Younis — Signature: ______ 28-Feb-2024 1:31 PM

Associate Dean EE
SEECS - NUST
Sector H-12, Islamabad

HoD/Associate Dean

29-Feb-2024
Date

COUNTERSINGED

29Feb, 2024
Date

Dr. Muhammad Ajmal
Principal
NUST School of Electrical Engineering & Computer Science
H-12 Islamabad

Dean/Principal

FMs
74096

Dedication

This thesis is dedicated to the unwavering support and love of my family, whose encouragement has been my anchor throughout this academic journey. To my parents, who instilled in me the value of education and resilience, and to my brother, who cheered me on through every triumph and setback.

I extend heartfelt gratitude to Dr. Shahzad Younis, my mentor and supervisor, for his unwavering guidance, insightful feedback, and continuous support. His expertise has been instrumental in shaping the trajectory of this research.

Lastly, to anyone who finds themselves flipping through these pages, may this thesis contribute in some small way to the collective pursuit of knowledge and understanding.

Certificate of Originality

I hereby declare that this submission titled "Optimizing Comfort and Energy Efficiency: Smart HVAC Control with Reinforcement Learning and Time Series Forecasting" is my own work. To the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics, which has been acknowledged. I also verified the originality of contents through plagiarism software.

Student Name:Abdullah .

Student Signature: ________________

Acknowledgments

In the name of Allah, the Most Gracious, the Most Merciful. I extend my deepest gratitude to Allah (S.W.T) for His continuous guidance and blessings that have illuminated my path throughout my research at NUST. Special thanks to my parents, whose unwavering support and sacrifices have been my foundation. I am grateful for the encouragement and understanding from my brother, whose presence has been a source of strength.

I would also like to express my appreciation to my supervisor Dr. Shahzad Younis for their guidance, mentorship, and valuable insights, which significantly contributed to the success of this research. Furthermore, I express my gratitude to all my friends for their unwavering support and encouragement.

Abdullah

Abstract

The Deep Reinforcement Learning (DRL) Framework offers a sophisticated solution for optimizing HVAC control strategies in diverse building environments, utilizing deep learning algorithms and reinforcement learning. Adaptable to various building types and HVAC systems, the framework enhances energy efficiency and occupant comfort. Successful implementation requires consideration of system complexity, data availability, and computational resources. Integrating time series forecasting models, specifically 1D CNN and LSTM, with reinforcement learning proves effective in predicting system parameters. These predictions guide the reinforcement learning agent in making sequential decisions for HVAC control actions, resulting in improved total rewards and validation loss during training. This holistic approach, supported by experiments, demonstrates tangible benefits in achieving optimal HVAC system management, highlighting increased energy efficiency and cost-effectiveness through the synergy of predictive analytics and adaptive control.

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

Abbreviations

HVAC	Heating, Ventilation, and Air Conditioning
DRL	Deep Reinforcement Learning
DQN	Deep Q Network
LSTM	long short-term memory networks
1D CNN	1D Convolutional Neural Network

CHAPTER 1

Introduction

1.1 Background

In recent years, the field of building automation has witnessed a transformative shift towards intelligent and adaptive control systems. Among these, Heating, Ventilation, and Air Conditioning Heating, Ventilation, and Air Conditioning (HVAC) systems play a pivotal role in maintaining indoor environmental comfort and energy efficiency. As buildings become more sophisticated and dynamic in their energy demands, there is a growing need for advanced control strategies that can adapt to changing conditions and optimize HVAC system performance.

Buildings use 36% of the energy consumed worldwide, according to the United Nations Environment Program. Interestingly, a large amount of this energy consumption is accounted for by integrated HVAC (heating, ventilation, and air conditioning) systems [1]. These technologies are essential for finding a balance between maintaining occupant comfort and energy economy. However, it is more important than ever to maximize energy use without sacrificing indoor comfort due to rising energy costs and growing environmental concerns [2, 3].Figure 1.1 explains how various things like sunlight, building design, location, and human activity all affect how much heating and cooling a building needs. The effective management of HVAC systems may be the most important component in guaranteeing appropriate energy saving, cost reduction, environmental sustainability, and occupant comfort in building structures in order to meet such a wide range of goals. In the world of commercial buildings specially offices, comfort comes at a hefty price: 75% of their energy goes towards keeping occupants comfortable with HVAC and lighting. With HVAC alone chomping down approximately 40%. As shown in Figure 1.2, these two factors dominate the pie chart representing an office building's typical energy consumption breakdown.

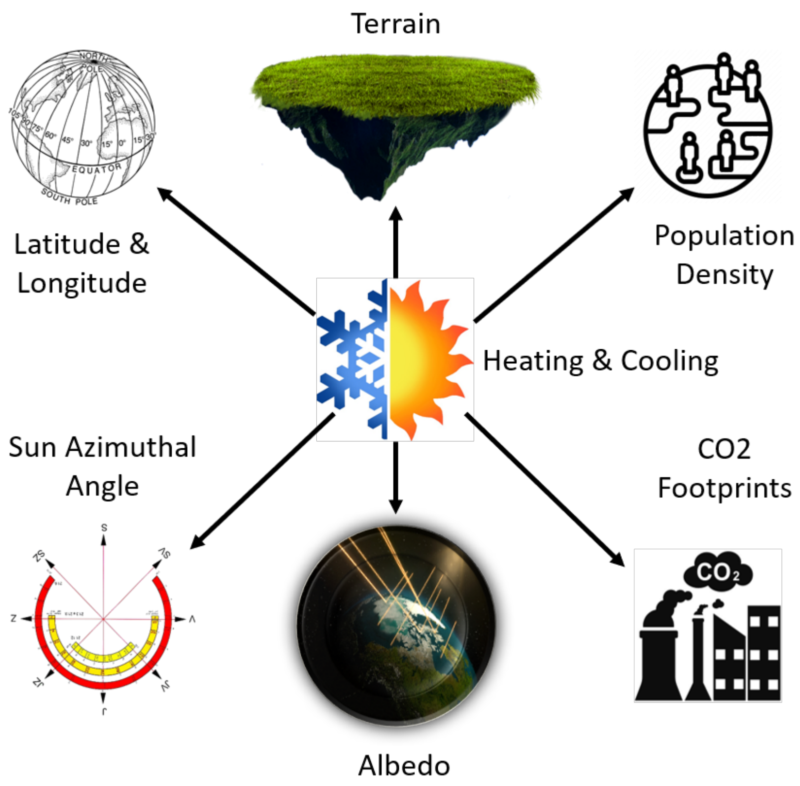


Figure 1.1. Exploring the Potential Influencing Factors of Heating and Cooling Degree: A Comprehensive Analysis.

In order to control HVAC systems, residents and building users have long relied on basic manual techniques, such as twiddling thermostats, opening windows, or turning fans on and off in response to sudden comfort demands [4]. Though these manual methods offer a concrete and immediate means of affecting the indoor atmosphere, they frequently fall short in terms of efficient energy use and long-term maintenance of constant comfort. Because of this, it is now evident that a more complex control strategy is required.

Which initially led to advancements like simple timers and programmable thermostats (see figure 1.3). These offered a step towards automation, allowing for preset schedules and temperature control, enhancing basic comfort levels. However, their rigidity limited adaptability to individual preferences and real-time environmental changes, often resulting in energy loss during idle times or unforeseen weather fluctuations. This inherent inflexibility necessitates the exploration of more sophisticated control themes [2, 3, 5].

Rule-based control, or RBC, is another popular technique for controlling HVAC systems in addition to conventional automation. The rules, which define conditions and related actions like

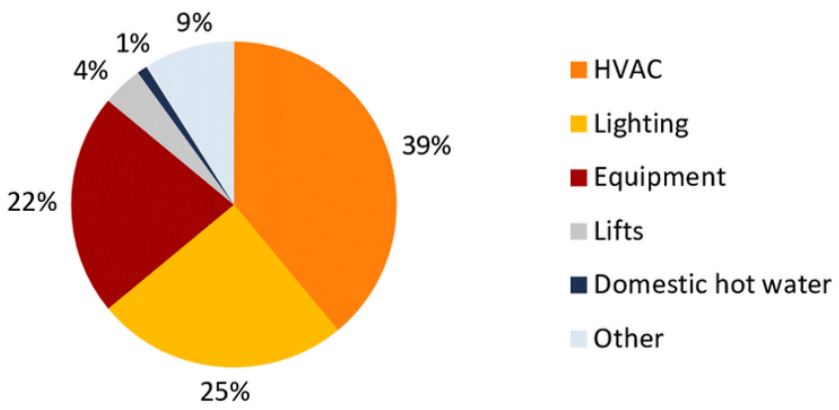


Figure 1.2. Analysis of energy usage in a standard office facility.

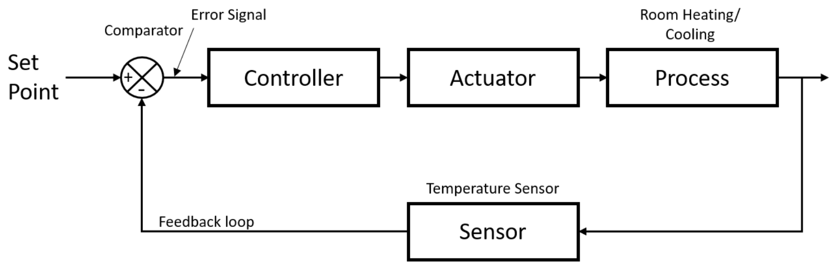


Figure 1.3. Block diagram of a room heating and cooling system with a proportional-integral (PI) controller.

reducing temperatures during off-peak hours or regulating ventilation based on occupancy levels, are frequently shaped by empirical observations and expert knowledge. Since HVAC systems' operation is subject to a variety of influences, including outside weather, building occupancy, and equipment variability, which are frequently unpredictable or difficult to predict by a set of predefined rules, such control schemes are inevitably unable to provide real-time optimal control for HVAC systems [6]. Moreover, the RBC technique typically ignores cost-effectiveness or energy efficiency in favor of preserving comfort levels. In addition, complex interactions between various HVAC components are not integrated, which might result in unsatisfactory performance of the system overall [7, 8].

In order to effectively handle the associated complexity and non linearity of the thermal dynamics in buildings, algorithm-inspired control strategies, such as intelligent model-based or model-free methodologies, have been developed in response to this uncertainty and complexity, which has presented significant challenges to the efficient operation of HVAC systems [9, 10]. Because model-based control systems can handle restrictions and optimize over a prediction, like the well-known model predictive control (MPC) approach, they are becoming more and more common in

HVAC control. Model-based techniques appear to be significantly more sophisticated than RBC, but putting them into practice has their own set of difficulties. Initially, these methods just rely on mathematical representations of the system to forecast future actions and enhance control measures according to these forecasts.

Because mathematical models are complex and nonlinear, developing correct models has substantial drawbacks. Furthermore, these models fall short in representing the dynamic nature of several important variables that affect HVAC efficiency, such as occupancy and weather, which are unpredictable external circumstances [11, 12]. Second, because they involve solving an optimization issue at each control step, techniques like maximum power consumption (MPC) may be computationally demanding, particularly for large multi functional systems. This result highlights a major difficulty in putting such methods into practice for real-time control, especially in systems with fast dynamics where frequent control decisions are required [8].

Approaches for model-free control have become attractive substitutes in view of these difficulties. For the control of complex frameworks like traffic management systems (TMSs), electric vehicle charging stations (EVCSs), building energy management systems (BEMSs), IoT ecosystems, and even autonomous robotic navigation, model-free control strategies provide a number of benefits [13–15]. In light of this, model-independent control approaches offer a successful control strategy for complex autonomous HVAC control. Firstly, such solutions are not strictly dependent on explicit models of the system, removing the need for the time-consuming and often erroneous process of system modeling [16]. This is very useful for HVAC control because, especially in large-scale building constructions, system dynamics are extremely nonlinear, complicated, and impacted by a multitude of external influences. Furthermore, model-free control methods are sufficient for directly learning from data, adjusting to changes, and managing intricate, nonlinear system dynamics. Examples of these algorithms are the widely recognized reinforcement learning and artificial neural network systems. These kinds of data, which might be historical or current, allow control rules to be improved over time. Because of their flexibility, they can perform better than more conventional control techniques like RBC when the dynamics of the system are unpredictable or changing [17].

A possible solution to these issues is model-free reinforcement learning, in which an RL agent interacts directly with the building environment to learn the building dynamics through firsthand experience. This agent may autonomously and rationally make decisions on behalf of a building owner to accomplish pre-established goals by having its reward function stated. This work contributes to the optimization of energy use for commercial buildings, a project in which ASP Lab is creating effective smart energy meters for buildings.

This thesis explores a cutting-edge approach to HVAC control through the application of Deep Reinforcement Learning Deep Reinforcement Learning (DRL) frameworks. DRL has emerged as a revolutionary paradigm in artificial intelligence, exhibiting unprecedented capabilities in learning complex tasks through interaction with an environment. By integrating DRL techniques into HVAC control systems, this research seeks to revolutionize the way buildings manage their thermal comfort and energy consumption.

1.2 Problem Statement

Conventional Heating, Ventilation, and Air Conditioning (HVAC) control methods, often based on rule-based or model-predictive approaches, face limitations in handling the dynamic and non-linear nature of building systems. These methods typically rely on predefined rules or models, which may not capture the intricate inter dependencies and uncertainties inherent in real-world environments. As a consequence, these systems often struggle to adapt to changing conditions, leading to sub-optimal energy usage and comfort levels.

Moreover, the increasing complexity of modern building structures, coupled with the rising emphasis on sustainability, necessitates more sophisticated and adaptive control strategies. Traditional approaches may struggle to provide the level of flexibility and efficiency required to meet the evolving demands of today's smart buildings.

1.3 Why Deep Reinforcement Learning?

In contrast, Deep Reinforcement Learning leverages neural networks and reinforcement learning algorithms to enable systems to learn and adapt autonomously. By allowing the HVAC control system to learn from its interactions with the building environment, DRL holds the promise of optimizing energy consumption, enhancing occupant comfort, and ultimately contributing to a more sustainable and intelligent built environment.

This thesis aims to investigate the potential of a DRL-based HVAC control framework to overcome the limitations of conventional methods and pave the way for a new era of adaptive, self-learning building automation systems. Through an in-depth exploration of DRL principles and their application to HVAC control, this research seeks to contribute valuable insights to the ongoing discourse on the future of smart building technologies.

1.4 Aims & Objectives

The objective is to employ Deep Reinforcement Learning (DRL) techniques to create an HVAC (Heating, Ventilation, and Air Conditioning) control system capable of optimizing zone mean air temperature and minimizing consumption costs. The primary contributions of this study include:

Data Generation and Pre-processing: This section focuses on the generation and pre-processing of data essential for training and evaluating the HVAC control system. It outlines the sources of data, the steps involved in data collection, and the pre-processing techniques applied to ensure the quality and relevance of the dataset.

Machine Learning Prediction: The Machine Learning Prediction section delves into the application of traditional machine learning techniques to predict HVAC system behavior. It explores how regression or classification models are employed to forecast key parameters, providing a baseline for comparison with deep learning approaches.

Deep Learning Models: Here, various deep learning models are discussed, including neural networks and other advanced architectures. The focus is on their application in understanding complex patterns within HVAC systems, with an emphasis on feature learning and representation.

Deep Reinforcement Learning Control: This section details the integration of deep reinforcement learning into HVAC control strategies. It covers the formulation of the reinforcement learning problem, the design of state and action spaces, and the implementation of the Deep Q-Network (DQN) architecture for learning optimal control policies.

Analysis of Environmental Factors: The analysis of environmental factors section explores the key external elements influencing HVAC system performance. It includes a detailed examination of outdoor temperature variations, solar radiation effects, wind speed impact, and occupancy patterns, providing insights into their roles in building thermal dynamics.

Performance Evaluation: In this section, the performance of the proposed HVAC control system is thoroughly evaluated. Key metrics, such as zone mean air temperature deviation, consumption cost, discomfort score, and adaptability to environmental changes, are analyzed. Comparative assessments against baseline scenarios and real-world applicability considerations are also discussed.

1.5 Thesis Organization

This thesis is organized as follows: Chapter 2 covers existing literature, the recent works involving HVAC Control using deep learning. Chapter 3 covers the Overview of Reinforcement Learning. Details of datasets available in literature and simulated in this thesis, proposed methodology and implementation . In Chapter 4, results obtained using the proposed method are described and explained. Chapter 5 includes a comprehensive discussion on the results of time series forecasting models and their integration with reinforcement learning. Lastly in Chapter 6, the conclusion and potential future directions are discussed.

CHAPTER 2

Literature Review

2.1 Building HVAC System Control Approaches

In the research literature, many approaches have been proposed to make heating, ventilation, and air conditioning (HVAC) systems in buildings more energy-efficient [18–21]. These methods typically involve using simplified models of the thermal dynamics of buildings to control energy consumption in real-time. For example, the authors of the study by [20] developed a nonlinear model that covers the entire cooling system, incorporating thermal storage banks, chillers, and cooling towers. They also introduced a model predictive control (MPC) scheme aimed at minimizing energy usage.

In another study by [18], the proposed system model is described as bilinear concerning inputs, states, and weather parameters. The optimization for control is framed as a sequential linear programming (SLP) problem in this case.

2.2 Energy Landscape in Commercial Buildings

The building sector, which includes both residential and commercial end users, is responsible for a significant 20.1% of the world's energy consumption. Keeping an eye on the business sector, which encompasses places of employment, malls, medical facilities, and so forth, air conditioning and lighting systems have a significant impact on the energy environment. Together, these systems account for more than 70% of energy consumption in the commercial building sector, of which about 45% comes from air conditioning and 25% from lighting. This is especially important in areas like the subcontinent where cooling is needed more than elsewhere, especially in places like workplaces, marts, shopping centers, and hospitals where people's health is at stake. Energy usage in these settings frequently exceeds 56%, highlighting the need for focused energy-efficient solutions and environmentally friendly procedures in commercial infrastructure [22].

2.3 Time-Series Forecasting in HVAC System Performance

Thermal comfort and system energy consumption are the two primary goals of time-series forecasting for HVAC systems. Convolutional neural networks (CNN), attention mechanism (AM), and bidirectional long short-term memory (BiLSTM) are the three primary LSTM-based extension strategies. By including a loop unit to enable forward and backward movement for the detection of the influence of future information, the bidirectional operations improve prediction performance. Studies have shown its contribution to the recursive prediction accuracy of HVAC energy consumption [23]. In the literature, attention mechanisms combined with LSTM/BiLSTM are more common. By allocating weights to various priority levels of information, the attention mechanism facilitates the more efficient identification of valid information. Additionally, the attention mechanism can be added to the front [24] or back [25] of LSTM layers, as well as the temporal [26, 27] or feature [24] dimension.

2.4 Evaluation of LSTM-Based Models and Recursive Prediction

Anjun Zhao employed a dual attention mechanism superimposed on top of the long short-term memory (LSTM), with the temporal attention layer placed after the LSTM layers and the feature attention layer placed before of them [28]. In the case of indoor temperature prediction, the improvement of long prediction period stability has been shown in [29]. It has also been proven that LSTM [26] and BiLSTM [24, 30] improve accuracy in energy consumption prediction and that this accuracy is resistant to overfitting [27]. Although they aid in the processing of sequential data, convolutional alternatives are frequently employed in the field of computer vision. For data pre-processing, a 1D convolution layer is frequently positioned before an RNN [31]. It has been demonstrated that using CNNs makes it easier to filter many features, reduce noise, and enhance prediction [32]. With significantly less computation time, such a technique can yield competitive results [23]. The resilience of LSTM-based models under recursive prediction scenarios, which is essential for the integration of predictive models with RL agents, has not been rigorously assessed despite the proposal of several different HVAC performance prediction models. Further exploration is necessary, since several studies have indicated that LSTM-based algorithms may be extended towards recursive prediction [23, 31].

2.5 Thermal Comfort Evaluation in HVAC Demand Response Control

Assessment of thermal comfort in HVAC demand response management The difficulty of creating building temperature control rules that are both energy-efficient and comfortable for occupants is covered in the document. The authors suggest a framework for thermal comfort control and energy optimization in smart buildings that is based on deep reinforcement learning. They frame the issue as one of cost reduction that takes into account occupants' thermal comfort as well as the energy usage of HVAC systems. The authors employ Deep Deterministic strategy Gradients (DDPG) to develop the thermal control strategy after initially using a deep neural network to forecast the thermal comfort of the occupants. The proposed approach's performance is assessed using a building thermal control simulation system, which demonstrates increased accuracy in predicting thermal comfort and decreased HVAC system energy usage.

2.6 Novel Metrics for Thermal Comfort in HVAC Demand Response Control

In this study, thermal comfort is evaluated for HVAC demand response control [33]. In order to evaluate demand responsiveness and thermal comfort in a university building, the authors suggest an entirely novel tool termed the Daily Discomfort Score (DDS). Operational temperature is used as the primary input parameter by the DDS, which also includes a punishment mechanism for temperature deviations over comfortable bounds and consecutive hours of pain. Baseline and preconditioning scenarios are used to show how successful the DDS is. The importance of interior thermal comfort in demand response techniques and the potential for HVAC control to save expenses and consumption of energy are also emphasized in the paper. The process for assessing the DDS entails creating a model for thermal dynamic simulation, putting an energy cost model into practice, and rating HVAC demand response control according to thermal comfort and cost parameters.

2.7 Building Thermal Behavior Modeling and Control

In [19], a tracking linear-quadratic regulator (LQR) for HVAC control is proposed, and building thermal behavior is modeled as RC networks. Using a building model akin to [19], the work in [21] creates an MPC-based algorithm for co-scheduling HVAC control with other demands and supplies. The precision of the building thermal dynamics model, which must also be effectively solved using mathematical tools for realistic runtime control, is crucial to the performance and dependability of these approaches [34]. The surrounding environment, which includes the temperature, humidity, and intensity of solar radiation, as well as interior heat gains from people, lighting systems, and other equipment, all have an impact on the building's temperature. As a result, when a modeling is insufficient, the building temperature frequently displays erratic behaviors. All things considered, creating a building dynamics model precise and efficient enough for runtime HVAC control is frequently unachievable. The recently developed deep reinforcement learning (DRL) method appears as a potent data-driven approach to tackling complicated control issues. It has been demonstrated to be successful in playing games such as Go and Atari [35, 36]. By constructing a deep neural network to relate the value estimations and related state-action pairs, the DRL technique can handle huge state spaces, which is a limitation of conventional reinforcement learning.

2.8 Deep Reinforcement Learning (DRL) in Building Control

In order to optimize energy usage, the [37] article suggests using deep reinforcement learning (DRL) to control the indoor air quality and temperature in a classroom and laboratory. It is crucial to adjust air conditioning and ventilation systems in accordance with occupant needs and external factors in order to strike a balance between thermal comfort and energy use. Energy consumption from air conditioners and ventilation fans, indoor air quality (CO_2 levels), and thermal comfort are all balanced by the DRL algorithm, notably the double Q-learning technique. After training on simulated historical data, the algorithm is put to the test in actual settings. The findings show that, in comparison to traditional management techniques, the suggested agent effectively regulates and balances the thermal environment, indoor air quality, and energy usage. With reduced energy use, the agent can minimize CO_2 levels and keep the indoor climate at a comfortable level.

The literature review explores various approaches to enhance the energy efficiency of Heating, Ventilation, and Air Conditioning (HVAC) systems in buildings. Different control strategies, including Model Predictive Control (MPC), Bilinear modeling, and various extensions of Long Short-Term Memory (LSTM) networks, are discussed for real-time energy consumption management. The commercial building sector's significant contribution to global energy consumption is highlighted, with a focus on HVAC and lighting systems. Time-series forecasting methods, such as Convolutional Neural Networks (CNN), Attention Mechanism (AM), and Bidirectional LSTM, are examined for predicting HVAC system performance.

The evaluation of LSTM-based models in recursive prediction scenarios is discussed, emphasizing their accuracy and resistance to overfitting. Additionally, the review explores the integration of Deep Reinforcement Learning (DRL) to optimize HVAC control, balancing energy usage and thermal comfort. Novel metrics like the Daily Discomfort Score (DDS) are introduced for assessing thermal comfort in HVAC demand response control. The significance of precise building thermal behavior modeling for effective HVAC control is acknowledged, with attention to the challenges associated with conventional modeling methods.

The subsequent sections present an organized overview of key research contributions, methods, and focuses in the field. The included studies cover a range of topics, including data-driven predictive control, model-free HVAC control, reinforcement learning applications, and optimization of energy usage, thermal comfort, and indoor air quality in smart buildings.

Table 2.1. Literature Review of HVAC Control Methods

Authors	Year	Title	Method	Focus	Contribution
Zhuang et al.	2023	Data-driven predictive control for smart HVAC system in IoT-integrated buildings with time-series forecasting and reinforcement learning	Data-driven predictive control, time-series forecasting, reinforcement learning	Smart HVAC system in IoT-integrated buildings	Integrated data-driven predictive control with time-series forecasting and reinforcement learning
Wang et al.	2023	Comparison of reinforcement learning and model predictive control for building energy system optimization	Reinforcement learning, model predictive control	Building energy system optimization	Compared reinforcement learning and model predictive control for optimization
Sivamayil et al.	2023	A systematic study on reinforcement learning based applications	Reinforcement learning	Reinforcement learning applications	Conducted a systematic study on reinforcement learning applications
Michailidis et al.	2023	Model-Free HVAC Control in Buildings: A Review	Model-free reinforcement learning	Model-free HVAC control	Reviewed and compared different model-free reinforcement learning approaches
					Continued on the next page

Table 2.1 – continued from previous page

Authors	Year	Title	Method	Focus	Contribution
Mao et al.	2023	Data analysis and interpretable machine learning for HVAC predictive control: A case-study based implementation	Data analysis and interpretable machine learning	HVAC predictive control	Combined data analysis and interpretable machine learning
Yu et al.	2021	A review of deep reinforcement learning for smart building energy management	Deep reinforcement learning	Smart building energy management	Provided a review of deep reinforcement learning in smart building energy management
Yu et al.	2021	Optimization of thermal comfort, indoor quality, and energy-saving in campus classroom through deep Q learning	Deep Q learning	Thermal comfort, indoor quality, and energy-saving	Optimized thermal comfort, indoor quality, and energy-saving
Brandi et al.	2020	Deep reinforcement learning to optimise indoor temperature control and heating energy consumption in buildings	Deep reinforcement learning	Indoor temperature control and heating energy consumption in buildings	Utilized deep reinforcement learning for optimization
Zou et al.	2020	Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network	Deep reinforcement learning, recurrent neural network	Optimal control of air handling units	Applied deep reinforcement learning and recurrent neural network for optimal control

Continued on the next page

Table 2.1 – continued from previous page

Authors	**Year**	**Title**	**Method**	**Focus**	**Contribution**
Azuatalam et al.	2020	Reinforcement learning for whole-building HVAC control and demand response	Reinforcement learning	Whole-building HVAC control and demand response	Developed and tested a reinforcement learning approach
Karevan & Suykens	2020	Transductive LSTM for time-series prediction: An application to weather forecasting	Transductive LSTM	Time-series prediction for weather forecasting	Introduced a Transductive LSTM model
Valladares et al.	2019	Energy optimization associated with thermal comfort and indoor air control via a deep reinforcement learning algorithm	Deep reinforcement learning	Energy optimization, thermal comfort, and indoor air control	Developed a deep reinforcement learning algorithm
Gao et al.	2019	Energy-efficient thermal comfort control in smart buildings via deep reinforcement learning	Deep reinforcement learning	Energy-efficient thermal comfort control	Proposed a deep reinforcement learning approach

CHAPTER 3

Methodology

The methodology is structured into two integral components, each crucial to the comprehensive analysis of the thermal dynamics and energy optimization within the National Science and Technology Park (NSTP) building. In the first part, denoted as Fig. 3.1 - "Data Generation," the process initiates with SketchUp modeling to intricately represent the building's physical structure. Subsequently, the model undergoes OpenStudio translation, followed by compilation and EnergyPlus simulation, ultimately culminating in the generation of extensive and detailed data sets. This phase lays the foundation for understanding the building's energy consumption and thermal behavior. In the second part, illustrated in Fig. 3.2 - "Reinforcement Learning Framework," the focus shifts to leveraging the generated data. This involves a sequence of steps, including Data Preprocessing, Feature Selection, Time series forecasting, Reinforcement Learning, and the application of advanced techniques for HVAC Control & Energy Optimization. Together, these two components form a cohesive methodology for a thorough exploration of the HVAC control system based on Deep Reinforcement Learning, enhancing the understanding and potential for energy-efficient practices within the NSTP building.

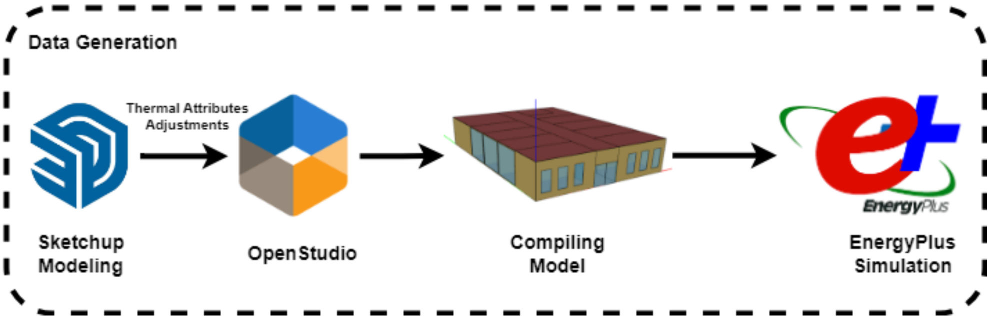


Figure 3.1. Data generation process for building energy simulation.

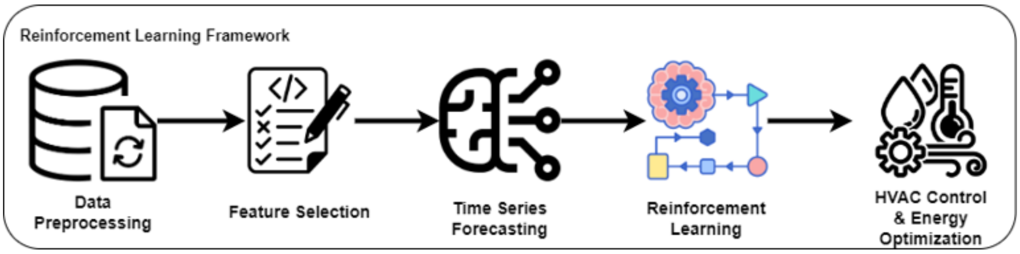


Figure 3.2. Reinforcement learning framework for the energy optimization and thermal comfort.

3.1 Building Description and Focus Area

In this research, the focus is on modeling and analyzing the 1st floor of Wing 04 within the National Science and Technology Park (NSTP) at the National University of Science & Technology (NUST), Islamabad, serving as a pertinent case study. The thermal model for this particular wing was constructed using SketchUp, and the subsequent parametrization was carried out utilizing OpenStudio. EnergyPlus was employed as the simulation engine to simulate the HVAC (Heating, Ventilation, and Air Conditioning) Framework.

The NSTP building, in its entirety, comprises three floors. However, for the purposes of this investigation, the attention is directed specifically towards the dimensions and characteristics of Wing 04 on the 1st floor. The dimensions of this wing are outlined as follows: length (73.00 meters), width (52.00 meters), and height (12.00 meters). The total area encompasses 2,032 square meters of conditioned spaces and an additional 1,135 square meters designated as non-conditioned spaces. Table 3.1 shows the materials used in the construction of this building.

Within this wing, there are various functional areas including 8 offices, stairwells, 2 washrooms, elevators, and other auxiliary spaces. To augment the analysis, features essential for the calculation of cost and DDS (presumably Demand Dispatch Services) were extracted and tailored from the dataset. It is worth noting that the utilization of OpenStudio and Energy-Plus in this context implies a robust approach to modeling and simulating the thermal dynamics, allowing for a comprehensive exploration of the HVAC control system based on Deep Reinforcement Learning.

The baseline schedule serves as a foundational reference for the heating, ventilation, and air conditioning (HVAC) system, establishing specific temperature parameters for different modes of operation. In heating mode, the set point is fixed at a comfortable 21°C, ensuring that the indoor environment remains warm and conducive to occupant well-being. Conversely, during cooling mode, the baseline schedule maintains a set point of 24°C, contributing to a cool and pleasant atmosphere within the designated spaces. Figure 3.4 shows the Heating and Cooling Schedule of

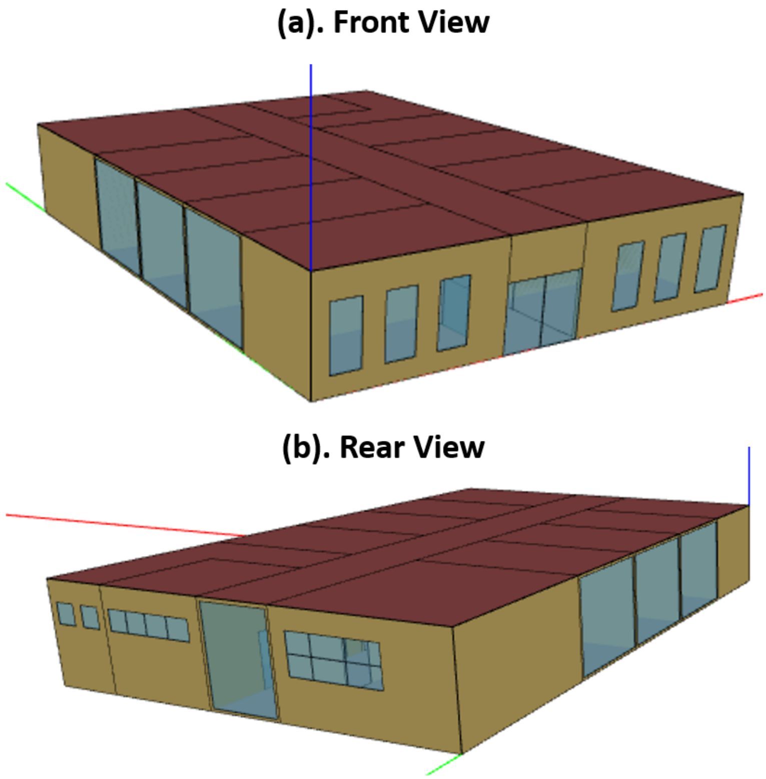


Figure 3.3. Front and rear view of the building.

the building.

These temperature standards are particularly applicable during typical working days and within specified hours from 08:00 to 20:00. This time frame aligns with regular business hours and reflects the anticipated occupancy and activity levels within the building. Notably, the baseline schedule does not encompass weekends, creating a differentiation between weekdays and non-working days. This intentional distinction aims to optimize energy efficiency by tailoring HVAC settings to the specific patterns of building usage.

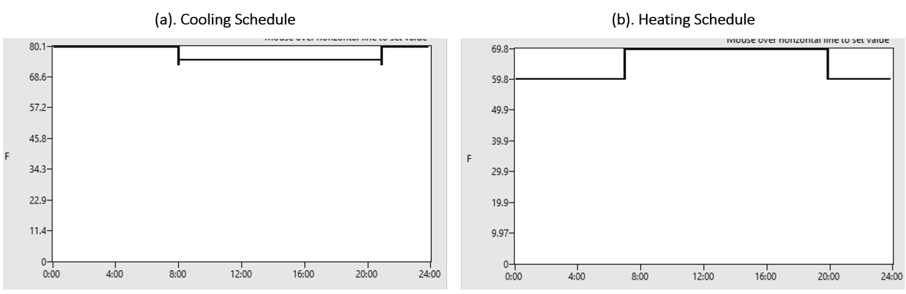


Figure 3.4. Heating and Cooling Schedule of the building.

Table 3.1. Wing 04, 1st Floor - Updated Structure Material Summary

Element	**Material**
Walls	Bricks with external plaster and internal paint (Varies in thickness depending on wall type)
Internal Walls	Gypsum board with and paint
Floors	Ceramic Tiles
Ceilings	None
Roof	RCC (Reinforced Concrete) slab with lantern and waterproofing membrane (Pitch varies depending on section of the roof)
Windows	Aluminum framed, double-glazed with low-emissivity coating and two-layer glass (Size and type may differ)
Doors	Glass doors with aluminium frame

3.2 Dataset Preparation

The dataset preparation process for this study involves a systematic and detailed approach, encompassing various stages and tools to ensure the generation of a high-quality dataset for the development and application of a Deep Reinforcement Learning (DRL) framework for HVAC control.

1. **SketchUp Modeling:** In the initial stage, the building's architectural features and spatial layout are meticulously modeled using SketchUp. This 3D modeling tool allows for a detailed representation of the building's physical structure, including walls, windows, doors, and other relevant elements. The accuracy of this model is crucial, as it serves as the baseline for subsequent energy simulations.

2. **OpenStudio Translation:** Following the SketchUp modeling, OpenStudio is employed to translate the detailed 3D model into a format suitable for energy simulations. OpenStudio facilitates the conversion of the architectural representation into an energy simulation-compatible format, preserving the intricacies of the building's design and ensuring that the

subsequent simulations are based on an accurate representation of the physical structure.

3. **EnergyPlus Simulation:** EnergyPlus, a state-of-the-art energy simulation software, is utilized for simulating the building's energy performance. The simulation process takes into account a myriad of factors, including building materials, insulation, HVAC system specifications, occupancy patterns, and external environmental conditions. By running simulations over a two-year period from January 2021 to December 2022, the software generates detailed energy data for multiple zones within the building, capturing variations in energy consumption across different seasons and environmental conditions.

3.2.1 Data Collection Period

The dataset spans a substantial two-year time frame, allowing for the comprehensive collection of data covering diverse environmental scenarios and seasonal changes. This extended period ensures that the dataset captures long-term trends and variations, providing a holistic understanding of the building's energy dynamics.

3.2.2 Dataset Characteristics

The dataset is characterized by its dimensions and includes various features essential for HVAC control model development. The key characteristics are:

- Total Number of Features: 17
- Total Number of Samples: 218,305
- Number of Zones: 12
- Samples per Zone: 18,193 (expected)

Notably, the dataset deviates slightly from the expected total sample count due to 11 missing values on the final day of the data collection period, which will be addressed during data preprocessing.

3.3 Dataset Preprocessing

In this project, the EnergyPlus simulation data was preprocessed thoroughly in order to extract useful data. The main dataset was first put into a Pandas DataFrame. It came from a simulation of a building's energy performance. The dataset was arranged into distinct column groupings, including lighting, thermostat settings, zone temperatures, occupant-related indicators, and environmental conditions. This arrangement provided the way for a methodical and perceptive examination of the energy dynamics of the structure.

3.3.1 Consumption Cost Calculation and Peak Hour Strategies

In the intricate landscape of HVAC control, the meticulous calculation of consumption costs is paramount for both operational efficiency and fiscal responsibility. The Consumption Cost (CC) is derived through a comprehensive formula that encapsulates various factors influencing energy expenditure:

$$CC = \text{Zone Lights Electricity Rate (Rupees)} \times \left(\frac{\text{Lights Power (kWh)}}{1000}\right) \times \text{Peak Hour Factor} \tag{3.3.1}$$

Breaking down this formula provides a deeper understanding:

- **Zone Lights Electricity Rate (Rupees):** This denotes the cost of electricity specifically attributed to lighting within each zone, expressed in Rupees.
- **Lights Power (kWh):** Reflecting the power consumption of lighting, measured in kilowatt-hours.
- **Peak Hour Factor:** An adjusting factor that takes into account whether the consumption transpires during designated peak hours.

To delve further into the dynamics of cost considerations, it is essential to explore the nuances of peak hours and their consequential impact on pricing structures. The per-unit price during peak hours stands at Rs 49.35, whereas outside of these peak periods, the rate is Rs 33.3 per unit.

Peak Hours Calculation

Understanding and strategically leveraging peak hours is instrumental in optimizing consumption costs. The delineation of peak hours is contingent upon the prevailing season:

- **Summer Months (June to September):** Peak hours extend from 5 PM to 11 PM, aligning with periods of heightened energy demand.

- **Winter Season (October to May):** Peak hours are scheduled from 6 PM to 10 PM, capturing the distinct energy consumption patterns during colder months.

To facilitate proactive planning and informed decision-making, a Peak Hour Notification link is provided, offering real-time insights into current peak periods.

In essence, a granular comprehension of consumption cost calculations and the strategic consideration of peak hours empower HVAC systems to operate judiciously. This approach not only contributes to efficient energy management but also aligns with cost-effective strategies for HVAC control. By navigating the intricacies of consumption costs and peak hour dynamics, organizations can concurrently prioritize sustainability and financial prudence in their energy utilization strategies.

3.3.2 Discomfort Score Calculation

Building upon the concept of the discomfort score (DS), our methodology draws inspiration from the Adaptive Comfort Standard (ACS) ASHRAE Std. [38]. This standard serves as an alternative to the Predicted Mean Vote (PMV)-based method and establishes a relationship between outdoor temperature (T_{out}) and indoor operative temperature (T_{in}). The optimal comfort indoor operative temperature (T_{op}) is modeled by the regression equation:

$$T_{\text{op}} = 17.8 + 0.31 \times T_{\text{out}} \tag{3.3.2}$$

Figure 3.5 illustrates this relationship, showcasing the 90% acceptability area between the green dashed lines, denoting the temperature range accepted by 90% of participants in the standard's experiment. This range extends up to $T_{\text{op}} \pm 2.5°C$. The red lines delineate the 80% acceptability area, encompassing temperatures within $\pm 3.5°C$ of the optimal line.

While the ASHRAE ACS standard was originally designed for naturally ventilated buildings, our methodology adapts it for mixed-mode buildings featuring both open windows and locally controlled HVAC systems. This adaptation assumes adherence to EN 15251 [39] compliant set points for indoor temperature control, ensuring prevention of overheating or overcooling. Additionally, our methodology leverages hourly outdoor air temperature values to extend the range of acceptable indoor operative temperatures, aligning with the ACS ASHRAE standard.

To calculate T_{op},[33] we determine the vertical difference $\Delta = T_{\text{in}} - T_{\text{comf}}$ (as illustrated in Figure 3.5) based on hourly indoor operative temperatures and concurrently measured outdoor temperatures. This approach enables a nuanced understanding of thermal comfort dynamics, essential for the successful implementation of our Deep Reinforcement Learning Framework for HVAC control.

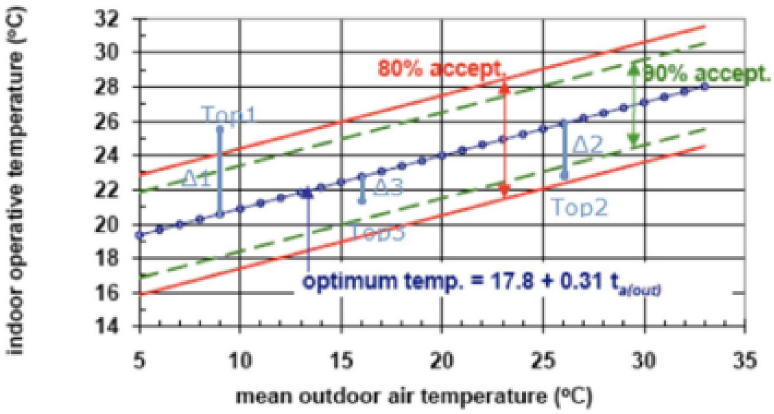


Figure 3.5. Correlation of indoor operative temperature and mean outdoor air temperature.

3.3.3 Feature Engineering

Feature engineering was crucial for improving the interpretability and analytical depth of the Energy Plus dataset. Creating dedicated columns for key factors like consumption cost, peak hour, and temperature delta enhanced granularity and insights. These features isolated critical parameters and streamlined subsequent analyses, allowing focused examination of the building's response to external conditions. Further, introducing zone-specific metrics like occupant count and lighting rate through thoughtful feature engineering enabled a nuanced understanding of individual thermal zones. This process not only organized the data for improved clarity but also paved the way for comprehensive insights into the intricate dynamics of energy consumption and thermal regulation within the simulated environment.

Zone Identification:

Description: Assigns a unique identifier to each specific area within the building.

Purpose: Enables the model to distinguish and understand the distinct characteristics and energy dynamics of different zones.

Environmental Conditions:

Features:

- Captures Environment Temperature
- Measures Environment Solar Radiation
- Monitors Environment Relative Humidity
- Records Environment Barometric Pressure
- Tracks Environment Wind Speed

Purpose: Captures external environmental factors influencing the building's thermal behavior, allowing the model to adapt to varying climate conditions.

Occupancy Information:

Feature: Records People Occupant Count

Purpose: Incorporates the number of occupants within a zone, providing insight into internal heat gains and aiding in predicting occupancy-related energy patterns.

Lighting-Related Data:

Feature: Monitors Zone Lights Electricity Rate [W]

Purpose: Accounts for the electricity consumption associated with lighting in each zone, a crucial factor in overall energy consumption.

Zone Temperature Characteristics:

Features:

- Measures Zone Mean Air Temperature
- Sets Zone Thermostat Heating Setpoint
- Sets Zone Thermostat Cooling Setpoint

Purpose: Describes the thermal conditions within each zone, including setpoints, which are vital for assessing and controlling the HVAC system.

Seasonal Information:

Feature: Identifies the current season (e.g., winter, spring)

Purpose: Indicates the current season, allowing the model to adapt control strategies based on seasonal variations in energy demand.

Temperature Delta:

Feature: Represents the difference from a reference temperature

Purpose: Provides insights into deviations from the desired thermal conditions.

Time-Related Indicators:

Features:

- Captures Time of Day
- Identifies Peak Hour
- Records Time (Timestamp)

Purpose: Incorporates temporal information to account for diurnal variations, peak energy demand periods, and precise timestamps for data recordings.

By extracting these features, the dataset becomes a rich and multidimensional representation of the building's energy dynamics. Each feature serves a specific purpose, collectively providing the DRL framework with the necessary information to learn and optimize HVAC control strategies based on complex interactions and dependencies within the building environment. This meticulous feature extraction ensures the dataset's effectiveness in training a robust and adaptive HVAC control model.

3.3.4 Feature Selection

Feature selection is a critical step in optimizing the performance and efficiency of machine learning models. In this project, the features for building energy dataset were carefully selected based on their relevance and impact on the target variables. The selection process involved utilizing a correlation matrix to identify the most influential features that contribute significantly to the understanding of energy dynamics within the simulated environment.

Correlation Matrix Analysis

The correlation matrix was employed to quantify the relationships between different features in the dataset. By examining the correlation coefficients, it was possible to gauge the strength and direction of the linear relationships between pairs of variables. This analysis facilitated the iden-

Table 3.2. Features Categorized for Building Energy Dataset.

Category	**Features**
Zone Identification	Zone
Environmental Conditions	Environment Temperature
	Environment Solar Radiation
	Environment Relative Humidity
	Environment Barometric Pressure
	Environment Wind Speed
Occupancy Information	People Occupant Count
Consumption Related Feature	Zone Lights Electricity Rate [W]
Zone Temperature Characteristics	Zone Mean Air Temperature
	Zone Thermostat Heating Setpoint
	Zone Thermostat Cooling Setpoint
Seasonal Information	Season
Temperature Delta	Temperature Delta
Engineered Features	Time of Day
	Peak Hour
	Consumption Cost

tification of features that exhibit high correlation with the target variables, making them valuable for predicting and optimizing HVAC control strategies.

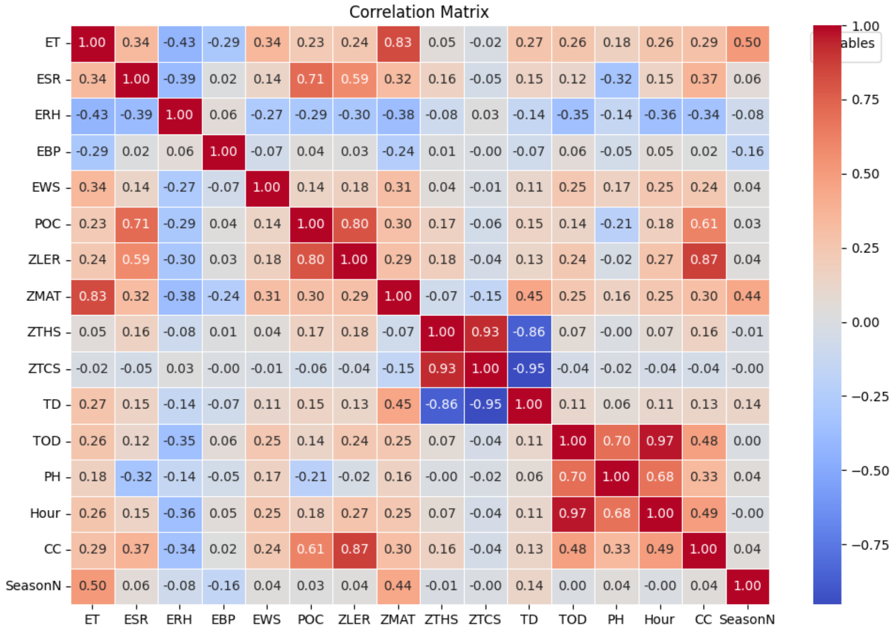


Figure 3.6. Correlation Matrix for feature selection.

3.3.5 Selected Features

The following features were selected for inclusion in the feature vector based on their correlation with the target variables:

Feature Vector for Zone Mean Air Temperature Prediction:

To predict Zone Mean Air Temperature, the model relies on a set of carefully selected features that capture various aspects of the environment and occupancy. The table below lists the specific features incorporated into the prediction model.

Feature Vector for Consumption Cost Prediction:

For predicting Consumption Cost, a distinct set of features is considered to capture the factors influencing energy consumption and associated costs. The table below provides an overview of

Table 3.3. Selected Features for Zone Mean Air Temperature Prediction

Features	Targeted Feature
Environment Temperature Environment Solar Radiation Environment Wind Speed People Occupant Count Zone Lights Electricity Rate Time of Day Temperature Delta Consumption Cost Season	**Zone Mean Air Temperature**

the selected features for the Consumption Cost prediction model.

Table 3.4. Selected Features for Consumption Cost Prediction

Features	Targeted Feature
Environment Temperature Environment Solar Radiation People Occupant Count Zone Lights Electricity Rate Zone Mean Air Temperature Time of Day Peak Hour Hour	**Consumption Cost**

These selected features encapsulate a diverse set of environmental, occupancy, and temporal factors that significantly influence both the zone temperature and consumption cost. By focusing on these key features, the model is expected to achieve a balance between predictive accuracy and computational efficiency, leading to an effective and optimized HVAC control strategy within the simulated building environment. The selected features provide the necessary information to train the machine learning model, enabling it to learn and adapt to the complex dynamics of the energy consumption patterns.

3.4 Time Series Forecasting Model Construction

After preprocessing the dataset, the next step involves employing various techniques on the dataset to forecast the targeted features, namely Zone Mean Air Temperature and Consumption Cost. In this context, five machine learning models, including random forest, decision tree, linear regression, support vector regressor, and gradient boosting, are utilized. Additionally, two deep learning models, specifically Long Short-Term Memory long short-term memory networks (LSTM) and one-dimensional Convolutional Neural Network (1D CNN), are incorporated for predictive analysis.

3.4.1 Forecasting Horizons

In our predictive modeling approach, we have defined three distinct forecasting horizons to cater to different planning and decision-making needs:

Short Term (Next 6 Hours)

This forecasting horizon focuses on predicting 'Consumption_Cost' and 'Zone_Mean_Air_Temperature' for the immediate future, specifically the next 6 hours.

Short-term predictions are valuable for real-time energy management and environmental control within the specified time frame.

Moderate Term (Next 24 Hours)

For a slightly extended planning horizon, the moderate-term forecasting targets the prediction of 'Consumption_Cost' and 'Zone_Mean_Air_Temperature' for the upcoming 24 hours.

This time frame is crucial for facilities and energy planners to optimize resource allocation and operational strategies.

Long Term (Next 7 Days - 1 Week)

The long-term forecasting horizon extends the prediction scope to cover a week, providing insights into energy consumption and air temperature patterns over an extended period.

Long-term forecasts are beneficial for strategic planning, allowing stakeholders to prepare for variations in energy demand and environmental conditions.

Each forecasting horizon requires specialized models and considerations due to the inherent differences in the temporal scales and factors influencing short-term, moderate-term, and long-term trends. The models developed for these horizons take into account the unique challenges and characteristics associated with each time frame.

3.4.2 Machine Learning

To develop regression models for predicting both 'Consumption_Cost' and 'Zone_Mean_Air_Temperature,' we initiated the process with data preprocessing. Using the Pandas library in Python, we extracted the necessary features, including 'Environment_Temperature,' 'Environment_Solar_Radiation,' 'Environment_Wind_Speed,' 'People_Occupant_Count,' 'Zone_Lights_Electricity_Rate,' 'Time_of_Day,' 'Temperature_Delta,' and 'SeasonN' (encoded season column) from the dataset.

The dataset was then split into training and testing sets, with a test size of 20% and a random state of 42. The training set was utilized for fitting the regression models, while the testing set served for evaluating their performance.

We employed a variety of regression models on the training data, each with its own architecture:

Linear Regression: Linear regression assumes a linear relationship between the input features and the target variable. The equation is given by:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \tag{3.4.1}$$

where y is the target variable, β_0 is the intercept, $\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients for each feature $x_1, x_2, \ldots, x_n$, and ε is the error term. See Figure 3.7 for its architecture.

Decision Tree Regression: Decision trees partition the feature space into disjoint regions based on predictor variables, capturing non-linear relationships. The decision tree doesn't have a direct mathematical representation, but it involves recursive splitting of the dataset based on features to minimize the sum of squared differences within each split. Refer to Figure 3.8 for the architecture.

Random Forest Regression: Random forests are ensembles of decision trees, reducing overfitting and improving overall performance. The architecture is a combination of individual decision trees. For a single tree, the equation is the same as in Linear Regression (Equation 3.4.1). Figure 3.9 illustrates its architecture.

Support Vector Regressor (SVR): SVR uses support vectors to find a hyperplane that best rep-

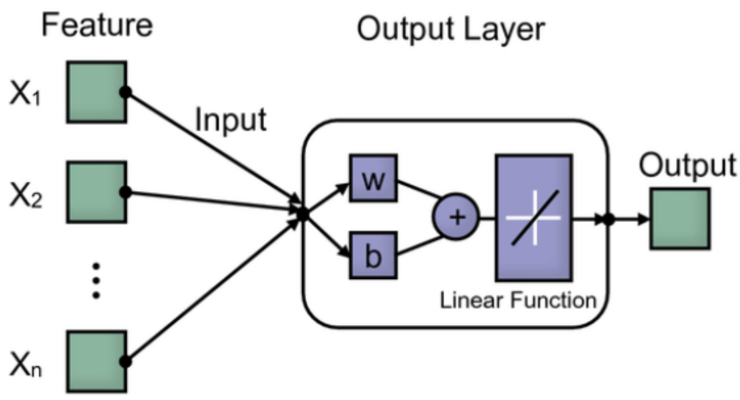


Figure 3.7. Architecture of Linear Regression.

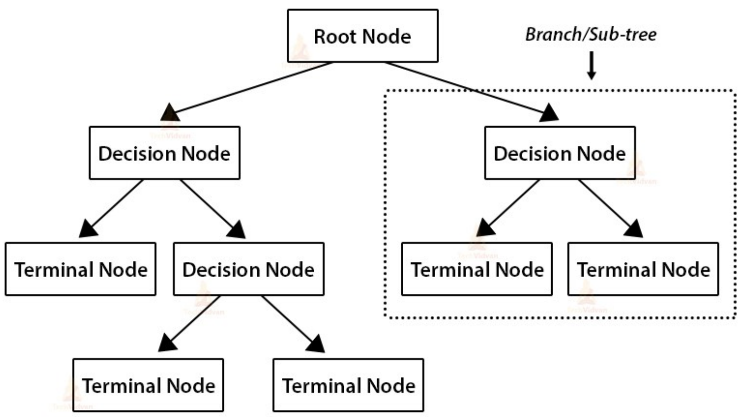


Figure 3.8. Architecture of Decision Tree Regression.

resents the relationship between features and target variables. The equation is given by:

$$y = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b \tag{3.4.2}$$

where y is the target variable, α_i are the support vector coefficients, x_i are the support vectors, K is the kernel function, and b is the bias term. The architecture is shown in Figure 3.10.

Gradient Boosting Regression: Gradient boosting builds an ensemble of weak learners (typically decision trees) to create a strong predictive model. The overall prediction is a sum of weak learners. The equation is given by:

$$F(x) = F_0(x) + \eta F_1(x) + \eta^2 F_2(x) + \ldots \tag{3.4.3}$$

where $F(x)$ is the predicted value, $F_i(x)$ represents the weak learner at iteration i, and η is the learning rate. Figure 3.11 displays its architecture.

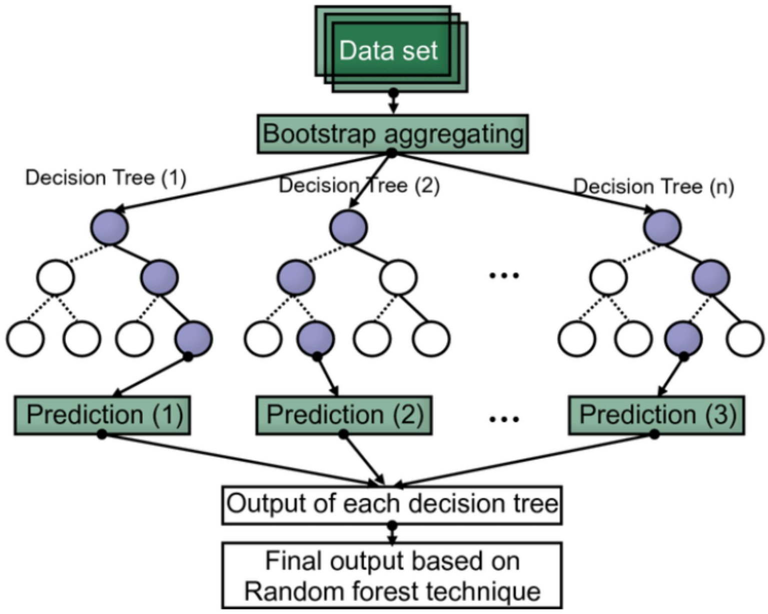


Figure 3.9. Architecture of Random Forest.

The Scikit-Learn library in Python facilitated model fitting, test data prediction, and the computation of Mean Squared Error (MSE) and R-squared (R^2) scores for each prediction target.

The Mean Squared Error (MSE) [40] is calculated as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3.4.4}$$

where n is the number of samples, y_i is the true value, and $\hat{y}_i$ is the predicted value.

The R-squared (R^2) score [41] is calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3.4.5}$$

where $\bar{y}$ is the mean of the true values.

After initially training uni-target models for each variable independently, we performed k-fold cross-validation for each prediction target, 'Consumption_Cost' and 'Zone_Mean_Air_Temperature,' to obtain more robust estimates of the models' performance. Using five folds, data shuffling, and a random state of 42, the Scikit-Learn library was utilized to fit the models to the training data, predict the test data, and calculate the MSE. The cross-validation MSE was then obtained by computing the mean of the MSE scores across all folds, providing a more reliable assessment of the models for both prediction targets.

It's important to note that all five algorithms were trained for multi-target variables, considering both 'Consumption_Cost' and 'Zone_Mean_Air_Temperature' as the target variables.

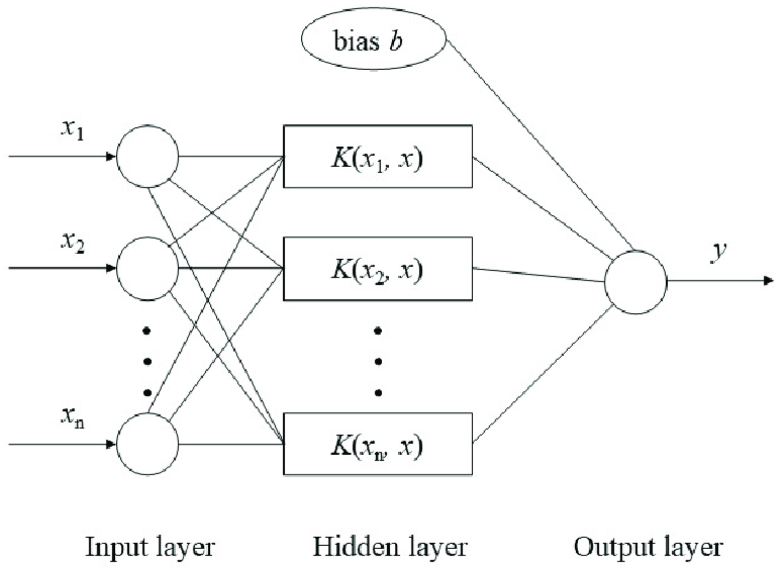


Figure 3.10. Architecture of Support Vector Regressor.

3.4.3 Deep Learning

The integration of deep learning techniques into HVAC systems involves the utilization of two sophisticated models for predicting 'Consumption_Cost' and 'Zone_Mean_Air_Temperature': the Long Short-Term Memory (LSTM) model and the 1D Convolutional Neural Network (1D CNN). Each of these models brings unique strengths to the table, addressing the temporal complexities inherent in energy consumption and air temperature fluctuations.

Long Short-Term Memory (LSTM)

The LSTM model, belonging to the family of recurrent neural networks (RNNs), stands out for its prowess in modeling sequential data, making it an ideal choice for handling time-series data prevalent in HVAC systems. The architecture of the LSTM regression model is illustrated in Figure-3.12.

Model Configuration: The LSTM regression model consists of a single LSTM layer with 50 units, followed by a dense output layer with one unit. The LSTM layer employs the Rectified Linear Unit (ReLU) activation function and has an input shape of (1, number of features). The output layer, with no activation function, comprises a single unit.

Compilation and Training: The model is compiled using the Adam optimizer and the mean squared error loss function. During training, the mean absolute error serves as a metric to evaluate

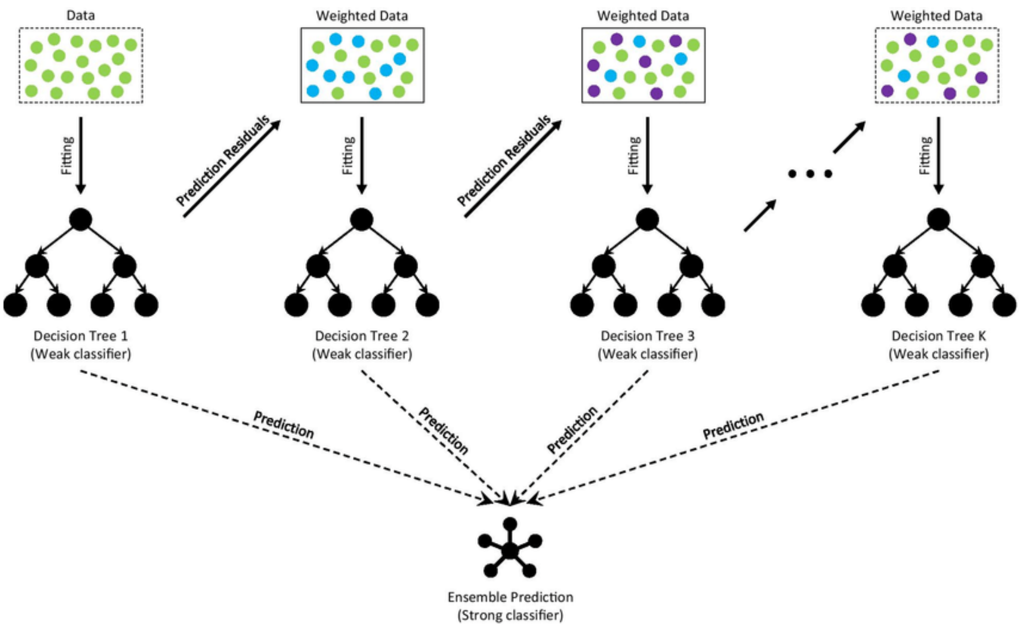


Figure 3.11. Architecture of Gradient Boosting Regression.

the model's performance. To prevent overfitting, the model incorporates dropout and recurrent dropout parameters, randomly dropping out some of the LSTM layer's output values.

The LSTM layer can be mathematically represented as follows:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \cdot \tanh(C_t) \end{aligned} \tag{3.4.6}$$

where:

- f_t, i_t, and o_t are the forget, input, and output gates, respectively.
- $\tilde{C}_t$ is the new candidate value for the cell state.
- C_t is the updated cell state.
- h_t is the output of the LSTM unit.
- $W_f, b_f, W_i, b_i, W_C, b_C, W_o, b_o$ are weight and bias parameters.

The training process spans 100 epochs with a batch size of 32, with periodic evaluations on a validation set. Post-training, the model's performance is assessed using metrics such as mean squared error and mean absolute error on the testing set.

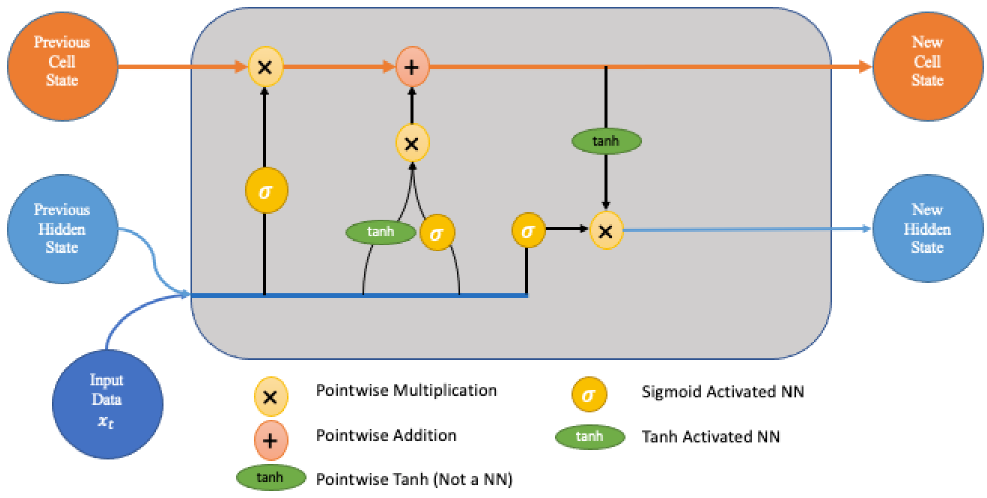


Figure 3.12. Architecture of Long Short-Term Memory (LSTM).

Multi-Target Long Short-Term Memory (LSTM)

Expanding the Long Short-Term Memory (LSTM) model into a multi-target framework involves predicting multiple output variables simultaneously, offering a nuanced perspective on HVAC system dynamics. Below is the architecture and training details for the Multi-Target LSTM:

Model Architecture: The LSTM model is built using the Keras Sequential API with L2 regularization. The model comprises a single LSTM layer with 50 units, taking into account the temporal dependencies within the input data. The input shape is adjusted to match the dimensions of the training data, ensuring compatibility with the time-series nature of HVAC-related features. The inclusion of L2 regularization with a specified strength (0.01 in this example) enhances the robustness of the model.

Compilation and Training: The model is compiled using the Adam optimizer and the mean squared error loss function. Early stopping with a patience of 10 epochs is employed to prevent overfitting, restoring the best weights when triggered. The training process spans 300 epochs with a batch size of 32, with validation data used to monitor performance and ensure generalizability.

1D Convolutional Neural Network (1D CNN)

The 1D Convolutional Neural Network (1D CNN) extends the principles of Convolutional Neural Networks to one-dimensional sequences, making it adept at capturing temporal dependencies—an essential feature in energy consumption prediction.

Model Architecture: The cornerstone of the 1D CNN is the Conv1D layer, equipped with a set of learnable filters. These filters operate on local regions of the input sequence, extracting distinct temporal features. The subsequent application of the ReLU activation function introduces non-linearities crucial for capturing complex relationships within the data.

Pooling layers, specifically MaxPooling1D, play a pivotal role in downsampling spatial dimensions, retaining essential information while reducing computational complexity. This step focuses on the most salient features within the temporal sequence.

The Flatten layer transforms the output of convolutional and pooling layers into a one-dimensional vector, preparing the data for processing by densely connected layers. The final dense layer, acting as the output layer, provides predictions for the target variables associated with energy consumption and thermal comfort. The model's architecture is depicted in Figure 3.13.

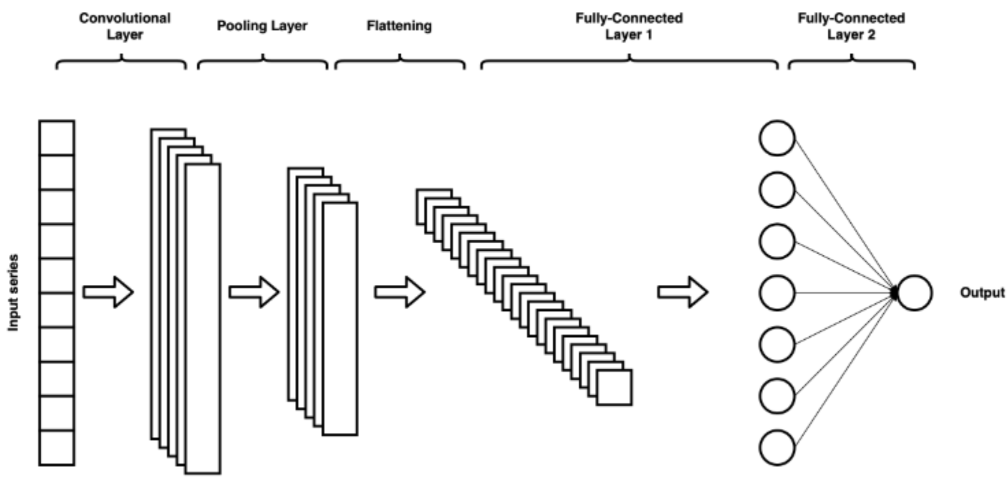


Figure 3.13. Architecture of 1D Convolutional Neural Network (1D CNN).

Model Compilation and Training: The compilation phase involves configuring the learning process, with optimization algorithms like Adam employed to minimize the Mean Squared Error (MSE) loss function. The training process spans iterative epochs, with the model adjusting its internal parameters to minimize the defined loss function.

The 1D CNN layer can be mathematically represented as follows:

$$
\begin{aligned}
&Y_t = f\left(\sum_{i=1}^{n} w_i \cdot X_{t+i-1} + b\right) \\
&\text{where:} \\
&\quad Y_t \text{ is the output at time } t, \\
&\quad X_t \text{ is the input sequence at time } t, \\
&\quad w_i \text{ are the learnable weights of the filters,} \\
&\quad b \text{ is the bias term,} \\
&\quad f \text{ is the ReLU activation function.}
\end{aligned}
\tag{3.4.7}
$$

Validation data plays a crucial role during training, providing insights into the model's performance on unseen data and aiding in preventing overfitting. The utilization of deep learning models in HVAC systems marks a significant advancement, enabling accurate predictions while considering the dynamic interplay between energy consumption and thermal comfort.

Multi-Target 1D Convolutional Neural Network (1D CNN)

Expanding the 1D Convolutional Neural Network (1D CNN) into a multi-target context involves predicting multiple output variables concurrently, providing a holistic approach to HVAC system forecasting. Below is the architecture and training details for the Multi-Target 1D CNN:

Model Architecture: The 1D CNN model is constructed using the Keras Sequential API, tailored to capture temporal dependencies within the input data. The architecture includes a Conv1D layer with 32 filters and a kernel size of 5, followed by a ReLU activation function. A MaxPooling1D layer with a pool size of 3 is incorporated for downsampling. The Flatten layer prepares the output for processing by densely connected layers, consisting of a hidden layer with 50 units and a ReLU activation function, as well as an output layer matching the number of target variables.

Compilation and Training: The model is compiled using the Adam optimizer and the mean squared error loss function. Training spans 300 epochs with a batch size of 32, and validation data is utilized to monitor performance. Additionally, predictions are generated on the test set for further evaluation.

3.5 Reinforcement Learning for HVAC Control: A Holistic Approach to Energy Efficiency and Comfort

In the pursuit of creating sustainable and comfortable indoor environments, the fusion of Reinforcement Learning (RL) with Heating, Ventilation, and Air Conditioning (HVAC) systems introduces a dynamic and adaptive control mechanism. This innovative approach leverages environmental states, controllable actions, and a refined reward system to strike an optimal balance between energy consumption and thermal comfort.

Environmental Dynamics: States and Actions

The RL algorithm operates within a state space (S) defined by real-time variables such as Environment Temperature, Consumption Cost, and Zone Mean Air Temperature. Controllable actions (A) involve discrete adjustments to the Zone Mean Air Temperature, providing a flexible and responsive means of HVAC system control. The selected set of discrete temperatures $\{22, 23, 24, 25, 26, 27\}$ offers a practical range for indoor climate management.

Enhanced Reward Function: Bridging Energy and Comfort

Central to the RL framework is a refined reward function that encapsulates the system's objectives. The reward (R) is computed as the negative minimum between Consumption Cost (CC) and a predetermined threshold (500). This structure emphasizes cost containment while introducing a critical element—thermal comfort. The integration of a weighted Discomfort Score (DS) enhances the reward function, facilitating a nuanced trade-off between energy efficiency and occupant satisfaction.

The reward function is given by:

$$R = -\min(\text{CC}, \text{Threshold\_Cost}) - \alpha \times \text{DS} \tag{3.5.1}$$

where α is the weighting factor, and DS is the Discomfort Score. To adjust the hourly discomfort score (DS) and determine the upper cap ($\Delta_{\text{Upper DS}}$) in the reward function, we need to integrate the discomfort score calculation into the given context. Now, the updated reward function is as follows:

$$R = -\min(\text{CC}, Threshold\_Cost) - \alpha \cdot \min(\text{DS}, \text{DS}_{\text{Upper}}) \tag{3.5.2}$$

Here, DS is the discomfort score, and we want to adjust it based on the Δ value. We can define DS as a function of Δ in the following way [33]:

$$\text{Discomfort Scale} = \begin{cases} 1, & \text{if } \Delta \geq 3.5 \\ 0.5, & \text{if } 2.5 \leq \Delta < 3.5 \\ 0, & \text{otherwise} \end{cases} \tag{3.5.3}$$

This ensures that the upper cap (DS_{upper}) is selected based on the same conditions used to calculate Δ, and the reward function takes the minimum of DS with the coefficient α.

Weighting Factor (α): Fine-tuning the Trade-off

A key feature of the RL framework is the introduction of a weighting factor (α), enabling system operators to tailor the algorithm's behavior. This factor dictates the balance between energy efficiency and thermal comfort. A higher α assigns greater importance to thermal comfort in the reward function, allowing for a personalized and adaptable trade-off based on specific operational goals.

In short, the integration of Reinforcement Learning in HVAC control transcends traditional paradigms, offering a holistic solution that optimizes indoor environments. This framework, adept at navigating the complex interplay between energy consumption and occupant well-being, paves the way for intelligent and responsive climate control systems.

Figure 3.14 illustrates the Reinforcement Learning Block Diagram.

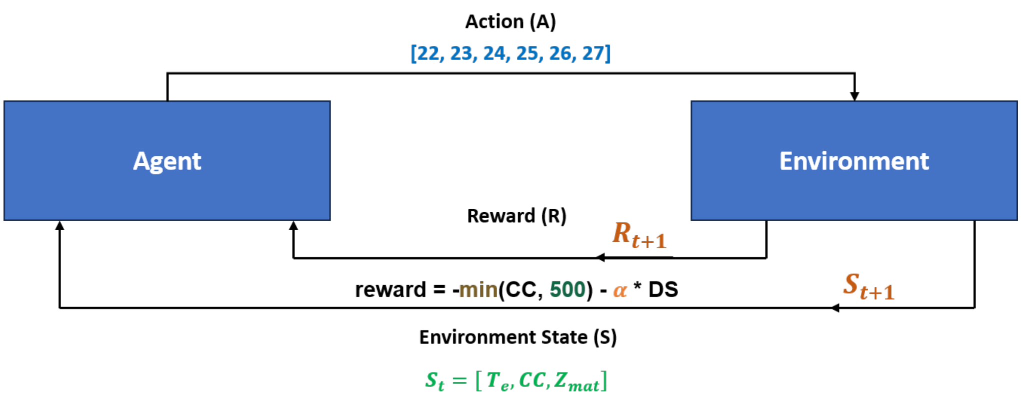


Figure 3.14. Reinforcement Learning Block Diagram

The transition from S_t to S_{t+1} involves selecting an action, adjusting zone mean air temperature, calculating discomfort, obtaining consumption cost, calculating the reward, and moving to the next time step. This repeats until the termination condition is met.

Model Architecture

The Q-network model, a fundamental component of the Deep Q-Learning agent, plays a crucial role in approximating the Q-values for different state-action pairs. Below is the architectural structure of the Q-network:

Deep Q-Network Architecture: The Deep Q-network Deep Q Network (DQN) is implemented as a neural network using the Keras library. It consists of several densely connected layers, with the final layer providing the Q-values for each possible action (see Figure 3.15).

- **Dense Layer 1:** 128 units with ReLU activation function.
- **Dense Layer 2:** 128 units with ReLU activation function.
- **Output Layer:** Number of units equal to the number of possible actions, with linear activation function.

The Q-network takes the environmental state as input and outputs Q-values for each action. This architecture enables the agent to learn and approximate the optimal action-value function during the training process.

Model Compilation: The model is compiled using the Adam optimizer with a specified learning rate. The mean squared error loss function is employed, aligning with the nature of Q-learning objectives.

Training and Validation: The training loop involves interacting with the HVAC environment, updating the Q-network based on observed rewards, and employing techniques like experience replay. Validation is performed to monitor the model's performance on unseen data.

Training Strategy and Hyperparameters

The successful training of the Deep Q-Learning agent involves careful selection of hyperparameters and the adoption of effective training strategies. Below are key details regarding the training process:

- **Number of Episodes:** The training process is conducted over a specified number of episodes,

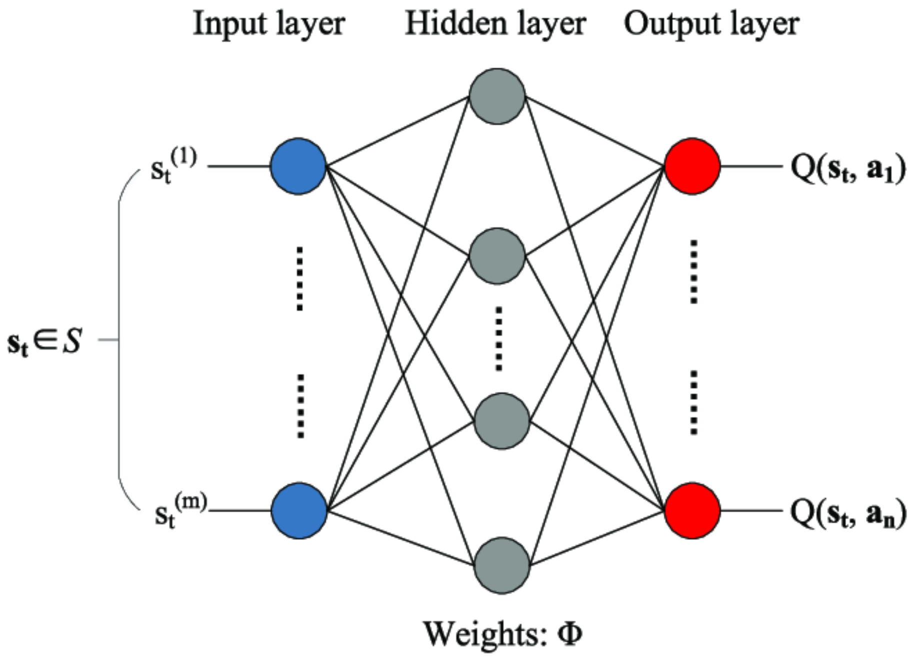


Figure 3.15. Architecture of the Q-network.

representing the total interactions between the agent and the environment. In this implementation, the number of episodes is set to 1000.

- **Exploration-Exploitation Strategy:** During each episode, the agent employs an exploration-exploitation strategy to balance between exploring new actions and exploiting the learned knowledge. The strategy involves using an epsilon-greedy policy, where the agent selects a random action with a certain probability (ε) and otherwise selects the action with the highest Q-value.

- **Learning Rate (α):** The learning rate determines the step size during the optimization process of updating the Q-network's weights. In this implementation, a learning rate of 0.001 is utilized.

- **Discount Factor (γ):** The discount factor (γ) controls the importance of future rewards in the agent's decision-making process. It is set to 0.99 in this implementation, indicating a high consideration of future rewards.

- **Exploration Rate Decay:** The exploration rate (ε) is decayed over time to gradually shift from exploration to exploitation. The decay factor is set to 0.995, ensuring a smooth transition.

- **Minimum Exploration Rate:** To ensure a minimum level of exploration, the exploration rate is constrained to not fall below 0.01.
- **Experience Replay:** Experience replay is employed to break the temporal correlation between consecutive experiences. The agent stores experiences in a replay buffer and samples batches during training.
- **Target Network Update Frequency:** The target Q-network is periodically updated to stabilize training. In this implementation, the update frequency is set to every 5 episodes.
- **Early Stopping:** To prevent overfitting, early stopping is applied with a patience of 20 episodes. If there is no improvement in the validation loss within the specified patience, training is halted.
- **Initial Exploration Phase:** In the initial episodes, the agent prioritizes exploration with a higher exploration rate to gather diverse experiences.

These hyper parameters and strategies contribute to the overall training process, enabling the agent to learn an effective policy for HVAC control over the specified number of episodes.

3.6 Integrating Time Series Forecasting Models and Reinforcement Learning

The methodology aims to achieve optimal HVAC system control and enhance energy efficiency through the seamless integration of time series forecasting models and reinforcement learning (RL). Utilizing advanced models such as 1D Convolutional Neural Network (1D CNN) and Long Short-Term Memory (LSTM), the key parameters of zone mean air temperature and consumption cost is accurately predicted. The time series models contribute by capturing temporal dependencies and providing essential predictions for the RL agent.

3.6.1 Reinforcement Learning Framework

The RL framework includes the predicted features from time series forecasting models—specifically, zone mean air temperatures and consumption cost into the state representation. The RL agent, with a focus on real-time decision-making, utilizes these projected features to optimize HVAC control strategies. The training process aims to fine-tune policies that maximize cumulative rewards, leading to enhanced total rewards and positive trends in validation loss.

3.6.2 Real-Time Decision Emphasis

While decisions are made with consideration for future states, the RL framework prioritizes real-time decision-making for the current state of the HVAC system. Immediate objectives, such as energy efficiency and comfort maintenance, are addressed through decisions made in the present.

3.6.3 Forecasted Features as Predictive Inputs

Forecasted features, obtained from time series forecasting models, serve as predictive inputs for the RL agent. These inputs provide insights into anticipated changes in temperatures and consumption costs, enhancing the agent's adaptability.

3.6.4 Proactive Decision-Making

With the help forecasted features, the RL agent engages in Proactive decision-making. It strategically considers both the current state and predicted future states to proactively adjust HVAC settings and adapt to upcoming changes.

3.6.5 Adaptive Control Strategies

The integration of forecasted features empowers the RL agent to dynamically adapt its control strategies based on the evolving conditions of the HVAC system. This adaptability enhances the system's responsiveness and overall efficiency.

CHAPTER 4

Results

This section presents the outcomes of two key components: the Reinforcement Learning (RL) approach for HVAC control and Time Series Forecasting. The RL agent's performance is evaluated based on its interactions with the HVAC environment, focusing on optimizing energy efficiency and thermal comfort. Additionally, the results of time series forecasting models, namely Long Short-Term Memory (LSTM) and 1D Convolutional Neural Network (1D CNN), are discussed in the context of predicting 'Consumption_Cost' and 'Zone_Mean_Air_Temperature.' The evaluation encompasses various metrics, visualizations, and analyses to provide a comprehensive understanding of both the RL-based HVAC control system and the accuracy of the forecasting models. The discussion includes observed trends, challenges, and potential areas for improvement in both approaches, offering valuable insights for further research and practical applications.

4.1 Time Series Forecasting Results

In this section, we present the results of time series forecasting using both 1D CNN and LSTM models for predicting 'Consumption Cost' and 'Zone Mean Air Temperature' based on different sequence lengths.

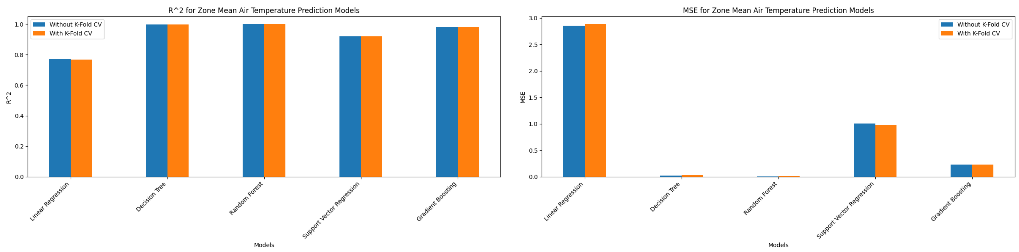
Machine Learning Models - Zone Mean Air Temperature Prediction

In this subsection, we explore the performance metrics for different machine learning algorithms used in predicting Zone Mean Air Temperature. The table below compares the models with and without K-Fold Cross Validation (5 Folds), utilizing Mean Squared Error (MSE) and R-squared (R^2) as evaluation metrics.

Table 4.1. Performance metrics for Zone Mean Air Temperature Prediction models with and without K-Fold Cross Validation.

	Algorithm	Without K-Fold CV		With K-Fold CV (5 Folds)	
		MSE	R^2	MSE	R^2
1	Linear Regression	2.8519	0.7688	2.8838	0.7662
2	Decision Tree	0.0221	0.9982	0.0321	0.9974
3	Random Forest	0.0100	0.9992	0.0131	0.9989
4	Support Vector Regression	1.0050	0.9185	0.9750	0.9209
5	Gradient Boosting	0.2298	0.9814	0.2350	0.9809

The results indicate that the Decision Tree, Random Forest, and Gradient Boosting models perform well in predicting Zone Mean Air Temperature, with high R^2 values and relatively low MSE. The use of K-Fold Cross Validation helps in assessing the models' generalization performance.



(a) R^2 for Zone Mean Air Temperature Prediction.

(b) MSE for Zone Mean Air Temperature Prediction.

Figure 4.1. Comparison of performance metrics for Zone Mean Air Temperature Prediction models.

Figure 4.1a shows the R^2 values for Zone Mean Air Temperature Prediction, and Figure 4.1b shows the MSE values.

Machine Learning Models - Consumption Cost Prediction

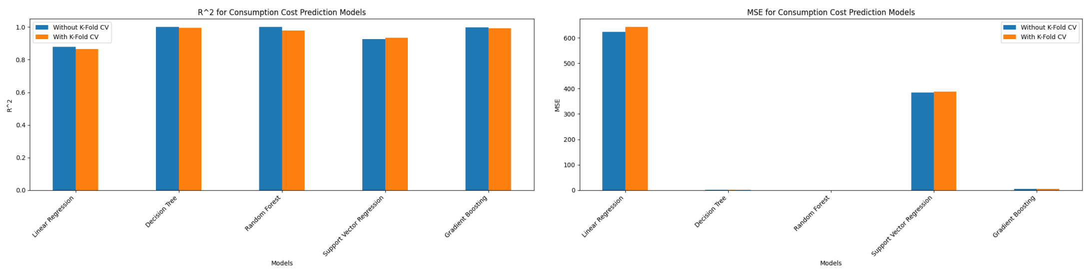
In this subsection, we present the outcomes of machine learning models applied to predict Consumption Cost. The table below displays the results, showcasing the performance of various algorithms without K-Fold Cross Validation. Metrics such as MSE and R^2 are reported for evaluation.

The results suggest that the Random Forest and Gradient Boosting models excel in predicting

Table 4.2. Performance metrics for Consumption Cost Prediction models without K-Fold Cross Validation.

Algorithm	Without K-Fold CV	MSE	R^2
Linear Regression		623.83	0.879
Decision Tree		1.854	1.0
Random Forest		0.0006	0.9999
Support Vector Regression		384.75	0.9256
Gradient Boosting		4.5953	0.9991

Consumption Cost, exhibiting very low MSE and high R^2 values. These models demonstrate strong predictive capabilities for the given task.



(a) R^2 for Consumption Cost Prediction. **(b)** MSE for Consumption Cost Prediction.

Figure 4.2. Comparison of performance metrics for Consumption Cost Prediction models.

Figure 4.2a shows the R^2 values for Consumption Cost Prediction, and Figure 4.2b shows the MSE values.

4.1.1 1D CNN Model - Consumption Cost Prediction

The forecasting model for 'Consumption Cost' was trained using 1D CNN with various sequence lengths. The model performance metrics and the corresponding loss plots are provided below:

Model Performance Metrics

The 1D CNN model for 'Consumption Cost' prediction was evaluated with varying sequence lengths, demonstrating its performance under different temporal contexts. The results suggest that the model performs well, achieving competitive Mean Squared Error (MSE), Normalized MSE (NMSE), Mean Absolute Error (MAE), and R^2 values across different sequence lengths.

Additionally, the table includes scenarios involving historical data from previous days and weeks, showcasing the model's adaptability to various input configurations.

Table 4.3. Model Performance Metrics for 1D CNN 'Consumption Cost' Prediction

Sequence Length	**MSE**	**NMSE**	**MAE**
6	1222.25	0.23487	16.5005
12	1063.15	0.20568	15.8219
24	1096.02	0.21142	15.7157
48	1098.46	0.21012	14.5309
60	1043.73	0.20194	14.1603
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week	89.495	0.017	4.81
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week + 6 hrs. in Last to Last Week	89.77	0.017	4.79
Previous Day 24 hrs. to Predict Next 6 hrs	131.59	0.024	5.21
Previous Last Day and Corresponding Day in Last Week	248.28	0.047	9.09
Previous Last Day and Corresponding Day in Last Week + Corresponding Day in Last to Last Week	233.06	0.044	9.27
Previous last week and corresponding week in last year's month	411.97	0.079	12.89

Loss Plots

The model trained to predict 'Consumption Cost' based on the consumption patterns of the previous week and the corresponding week from the previous year's month demonstrates a good fit. This is evident in the loss plot, where the convergence of the training and validation curves suggests effective learning without significant signs of overfitting, particularly in Figure 4.12.

When examining the loss plots for various sequence lengths and prediction intervals, as shown in Figures 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, and 4.11, we observe fluctuations, divergences, or plateauing, indicative of potential overfitting issues in other scenarios. These findings underscore the importance of carefully tuning hyperparameters and employing regularization techniques to strike a balance between model complexity and generalization.

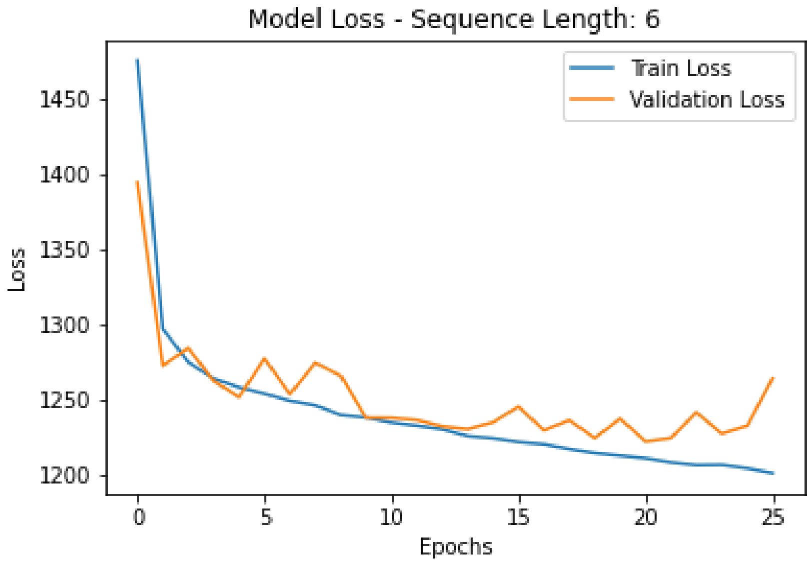


Figure 4.3. Loss Plot for 'Consumption Cost' Prediction - Sequence Length: 6

4.1.2 LSTM Model - Consumption Cost Prediction

The forecasting model for 'Consumption Cost' using LSTM was trained with various sequence lengths. The model performance metrics and loss plots are provided below:

Model Performance Metrics

The LSTM model for 'Consumption Cost' prediction was assessed with various sequence lengths, demonstrating its performance under different temporal contexts. The model exhibited competitive results in terms of Mean Squared Error (MSE), Normalized MSE (NMSE), Mean Absolute Error (MAE), and R^2 across different sequence lengths. The table also includes scenarios involving historical data from previous days and weeks, showcasing the model's adaptability to diverse input configurations.

Table 4.4. Model Performance Metrics for LSTM 'Consumption Cost' Prediction

Sequence Length	**MSE**	**NMSE**	**MAE**
6	1202.62	0.2311	15.82
12	1058.10	0.2047	14.52
24	1056.35	0.2038	14.83
48	1097.75	0.2100	15.25
60	1059.03	0.2049	15.01
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week	1142.3	0.2242	16.044
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week + 6 hrs. in Last to Last Week	1173.5	0.2264	15.933
Previous Day 24 hrs. to Predict Next 6 hrs	1563.4.5	0.2273	16.015
Previous Last Day and Corresponding Day in Last Week	1250.4	0.2417	17.002
Previous Last Day and Corresponding Day in Last Week + Corresponding Day in Last to Last Week	1497.3	0.2879	19.973
Previous last week and corresponding week in last year's month	765.92	0.147	17.21

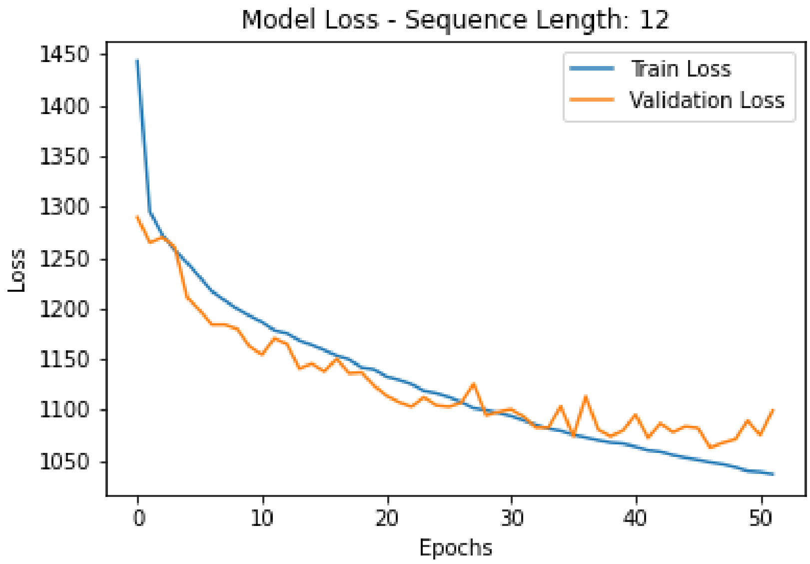


Figure 4.4. Loss Plot for 'Consumption Cost' Prediction - Sequence Length: 12

Loss Plots

The following loss plots provide insights into the training and validation performance of the LSTM model for 'Consumption Cost' prediction:

In Figure 4.18a, 4.18b, and 4.18c, a concise comparison of LSTM and 1DCNN models for Consumption Cost Prediction is presented. Each subfigure represents a specific metric: NMSE, MSE, and MAE, providing a quick visual assessment of the models' performance. The figures offer insights into the normalized error, squared error, and absolute error, respectively, aiding in the evaluation and selection of the most suitable model for Consumption Cost Prediction.

4.1.3 1D CNN Model - Zone Mean Air Temperature Prediction

Similarly, the forecasting model for 'Zone Mean Air Temperature' was trained using 1D CNN with various sequence lengths. The model performance metrics and the corresponding loss plots are presented below:

Model Performance Metrics

The 1D CNN model for 'Zone Mean Air Temperature' prediction was evaluated with varying sequence lengths, demonstrating its performance under different temporal contexts. The model displayed competitive results in terms of Mean Squared Error (MSE), Normalized MSE (NMSE),

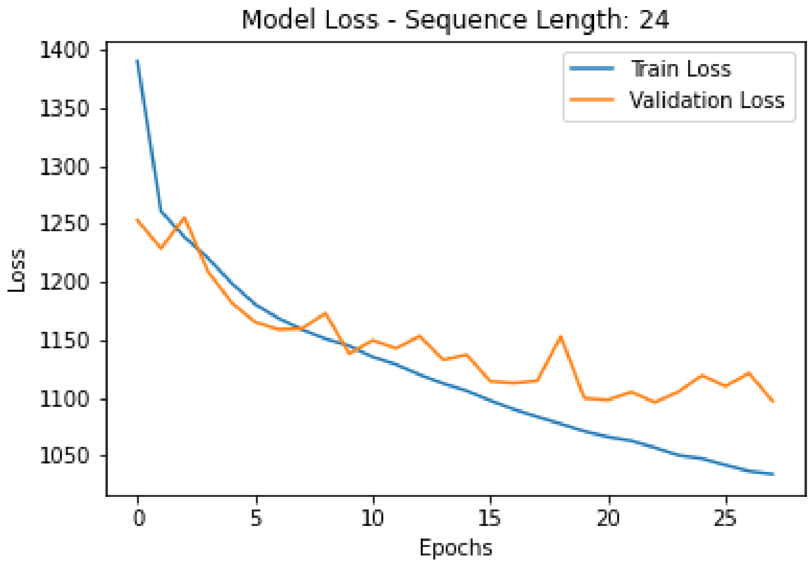


Figure 4.5. Loss Plot for 'Consumption Cost' Prediction - Sequence Length: 24

Mean Absolute Error (MAE), and R^2 across different sequence lengths. The table also includes scenarios involving historical data from previous days and weeks, highlighting the model's adaptability to diverse input configurations.

Loss Plots

The following loss plots provide insights into the training and validation performance of the 1D CNN model for 'Zone Mean Air Temperature' prediction:

4.1.4 LSTM Model - Zone Mean Air Temperature Prediction

The forecasting model for 'Zone Mean Air Temperature' using LSTM was trained with various sequence lengths. The model performance metrics and loss plots are provided below:

Model Performance Metrics

The LSTM model for 'Zone Mean Air Temperature' prediction exhibited strong performance across various sequence lengths, showcasing its ability to capture temporal dependencies. The metrics include Mean Squared Error (MSE), Normalized MSE (NMSE), Mean Absolute Error (MAE), and R^2, providing a comprehensive evaluation of the model's predictive capabilities. The table also explores scenarios involving historical data from previous days and weeks, highlighting

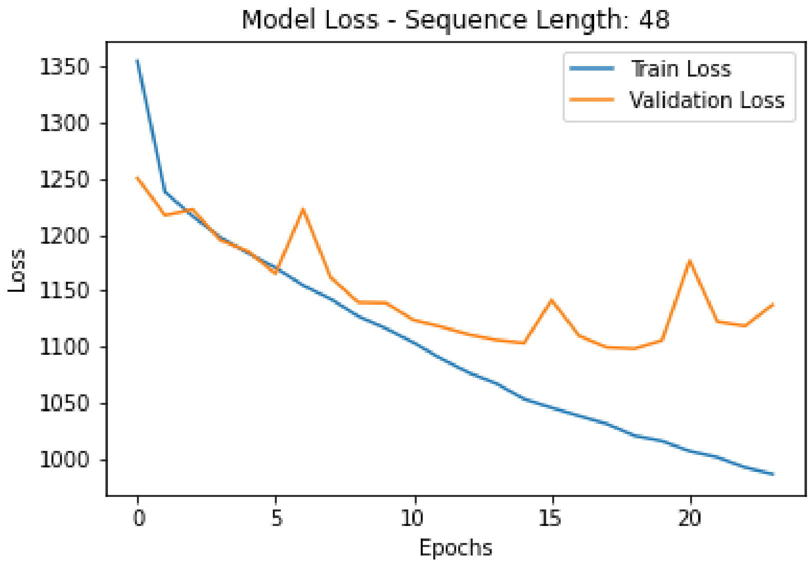


Figure 4.6. Loss Plot for 'Consumption Cost' Prediction - Sequence Length: 48

the model's adaptability to different input configurations.

Loss Plots

The following loss plots offer insights into the training and validation performance of the LSTM model for 'Zone Mean Air Temperature' prediction:

In Figure 4.33a, 4.33b, and 4.33c, a concise comparison of LSTM and 1DCNN models for zone mean air temperature Prediction is presented. Each subfigure represents a specific metric: NMSE, MSE, and MAE, providing a quick visual assessment of the models' performance. The figures offer insights into the normalized error, squared error, and absolute error, respectively, aiding in the evaluation and selection of the most suitable model for zone mean air temperature Prediction.

In summary, both 1D CNN and LSTM models were applied to time series forecasting for 'Consumption Cost' and 'Zone Mean Air Temperature'. The comparison of performance metrics and loss plots for different sequence lengths can guide the selection of the most suitable model for each target variable. The results provide insights into the models' accuracy and effectiveness in capturing the temporal patterns of energy-related parameters.

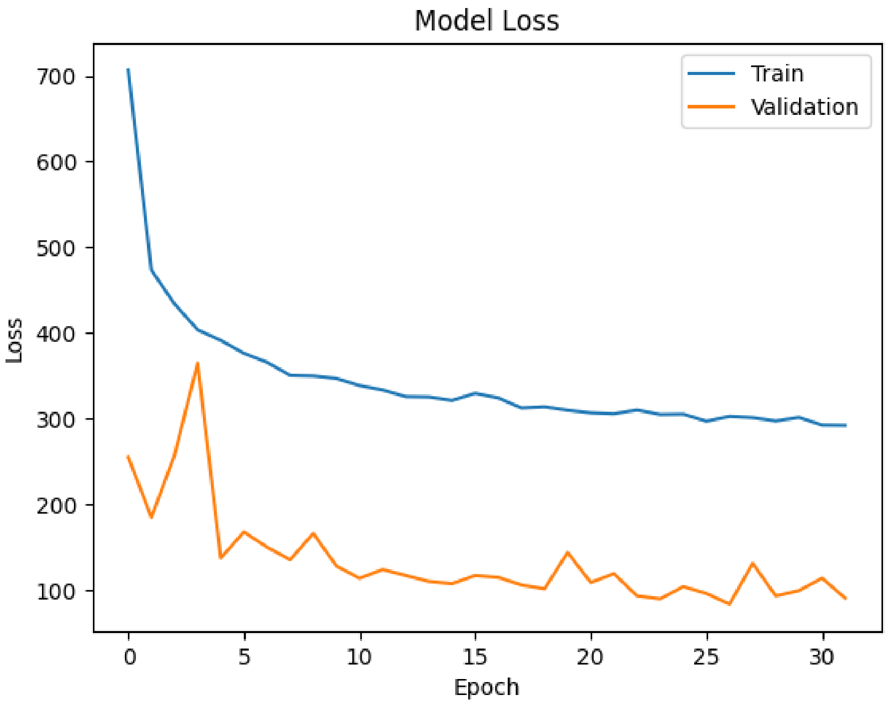


Figure 4.7. Loss Plot for 'Consumption Cost' Prediction - Previous 12 hrs to predict next 06 hrs.

4.1.5 Multi-Target Prediction

In this section, we present the results of multi-target time series prediction using LSTM and 1D CNN models. The target variables include 'Zone Mean Air Temperature' and 'Consumption Cost,' and various sequence lengths are explored.

LSTM Model

The LSTM model demonstrates robust performance in predicting both 'Zone Mean Air Temperature' and 'Consumption Cost.' The model was trained with various sequence lengths, and the following table presents the performance metrics:

The table provides a comprehensive overview of the model's performance under different input configurations. Notably, the LSTM model showcases strong predictive capabilities, as reflected in low MSE, MAE, and NMSE values, along with a high R^2 score.

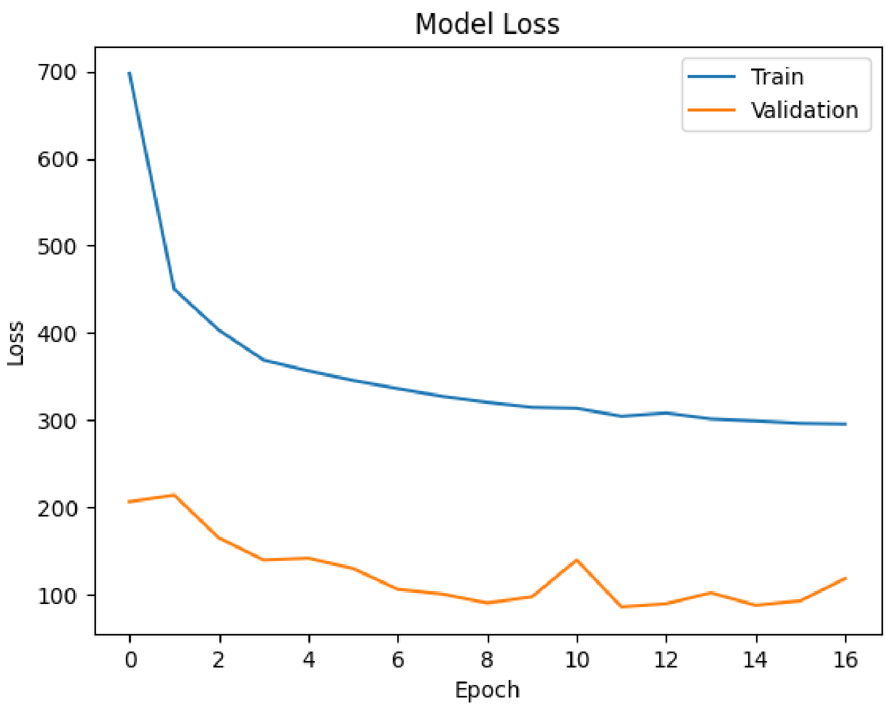


Figure 4.8. Loss Plot for 'Consumption Cost' Prediction - Previous 18 hrs to predict next 06 hrs.

1D CNN Model

The 1D CNN model, similarly, exhibits strong performance in predicting 'Zone Mean Air Temperature' and 'Consumption Cost.' The table below summarizes the model's performance metrics for different sequence lengths:

The performance metrics table showcases the 1D CNN model's accuracy and robustness across different scenarios and sequence lengths. Notably, the model demonstrates low MSE and MAE values, indicating precise predictions. The high R^2 score signifies the model's capability to capture the variance in the data.

Figures 4.36 and 4.37 present visualizations of the model predictions for 'Zone Mean Air Temperature' and 'Consumption Cost' in THERMAL ZONE 1, respectively.

4.1.6 Visualization of Predictions

In this section, the visual representation of the forecasted results obtained from our time series forecasting models are given below. Visualization plays a crucial role in enhancing the interpretability of predictions, allowing stakeholders to grasp patterns, trends, and potential anomalies within the predicted data.

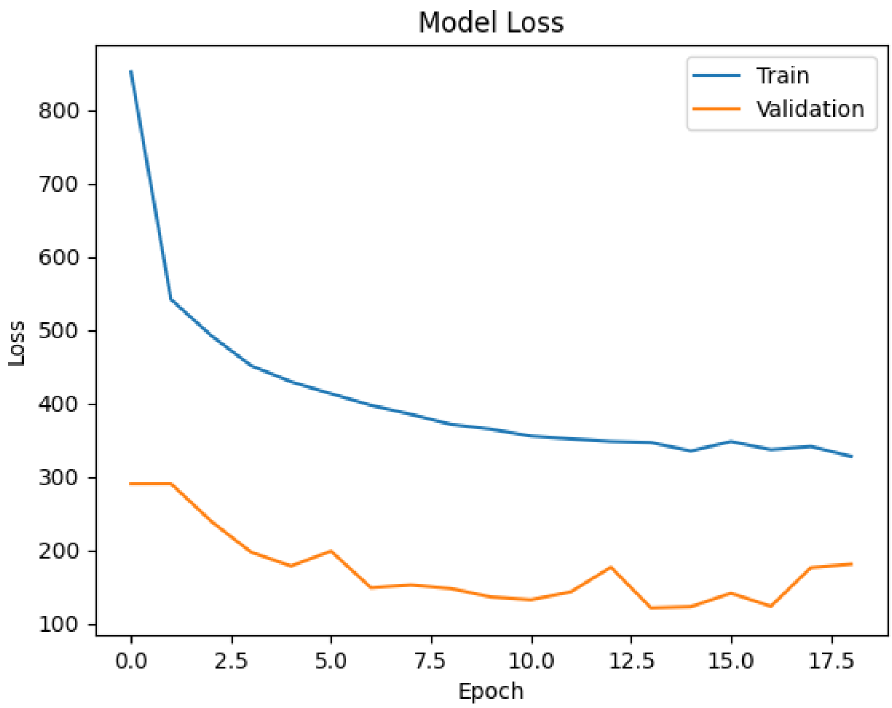


Figure 4.9. Loss Plot for 'Consumption Cost' Prediction - Previous 24 hrs to predict next 06 hrs.

The visualizations depict forecasted results from various time series forecasting models. Figures 4.38 and 4.39 showcase the 1DCNN model's monthly and weekly predictions for Consumption Cost (CC), revealing patterns and trends in monthly and weekly consumption costs, respectively. Figure 4.40 illustrates the 1DCNN model's ability to predict CC for the next day, aiding stakeholders in short-term decision-making. Similarly, Figures 4.41 and 4.42 display the monthly and weekly predictions for Zone Mean Air Temperature (ZMAT) using the same model. The 1DCNN ZMAT_next_day Prediction in Figure 4.43 indicates the model's capability to forecast daily temperature fluctuations. Additionally, the LSTM model predicts ZMAT and CC for the next day, as depicted in Figures 4.44 and 4.45, leveraging its capacity to capture sequential dependencies in the data. Furthermore, Figures 4.46 to 4.50 present predictions for ZMAT using various machine learning algorithms, including Support Vector Regression, Linear Regression, Gradient Boosting Regression, Decision Tree, and Random Forest, showcasing diverse modeling approaches for temperature forecasting.

4.2 Reinforcement Learning for HVAC Control Results

In this section, we present the results of applying reinforcement learning for HVAC control using the implemented Deep Q-Learning algorithm. The key metrics and aspects evaluated include total

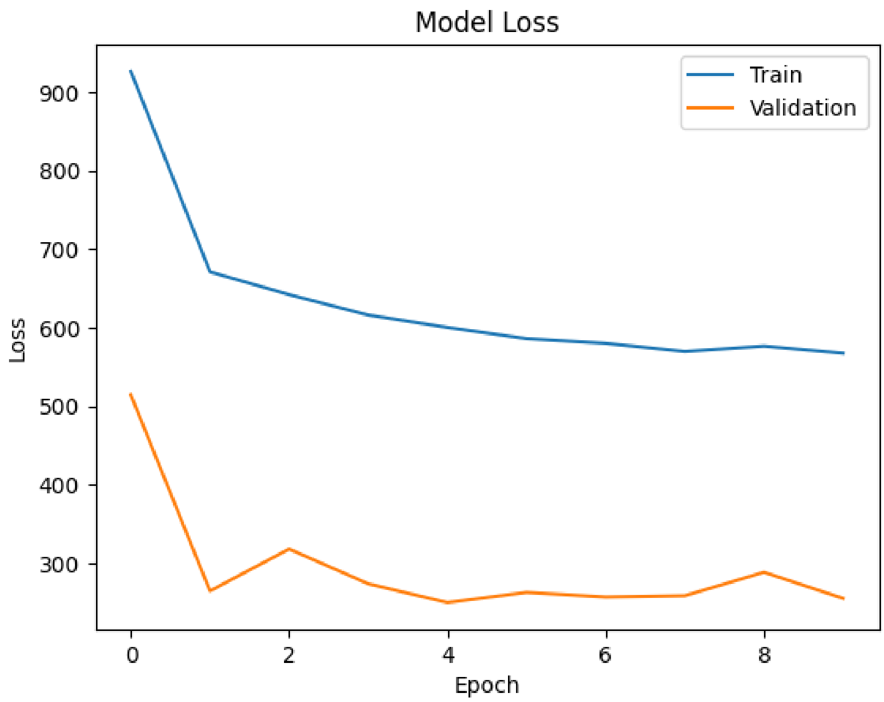


Figure 4.10. Loss Plot for 'Consumption Cost' Prediction - Previous 48 hrs to predict next 24 hrs.

rewards, validation loss, training and exploration strategies, model convergence, stability, and a comparison with a baseline.

The performance of the HVAC control system is assessed using various metrics. Table 4.9 summarizes the key evaluation metrics obtained during the reinforcement learning process.

4.2.1 Total Rewards

The total rewards obtained during the training episodes are a crucial indicator of the agent's performance. Figure 4.51 shows the plot of total rewards against the number of training episodes.

4.2.2 Validation Loss

Validation loss is an important metric to assess the generalization performance of the trained model on unseen data. The validation loss is calculated on a separate validation set. Figure 4.52 presents the validation loss over the training episodes.

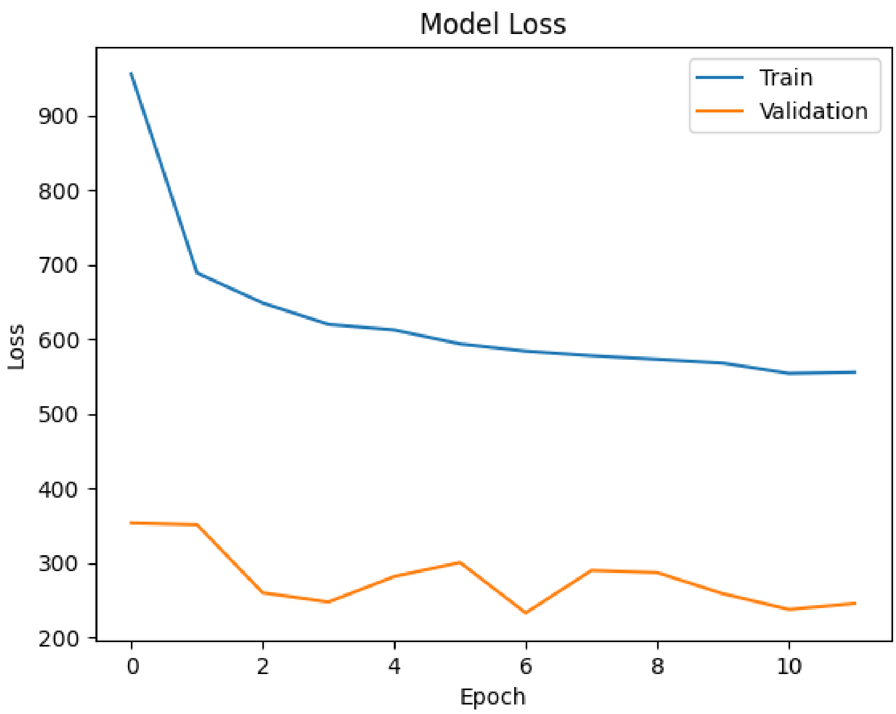


Figure 4.11. Loss Plot for 'Consumption Cost' Prediction - Previous 72 hrs to predict next 24 hrs.

4.2.3 Training and Exploration Strategies

The agent's training strategy involves updating the Q-network using the Deep Q-Learning algorithm. The exploration-exploitation strategy is controlled by the epsilon-greedy policy. Figure 4.53 illustrates the decay of epsilon values over the training episodes.

4.2.4 Model Convergence and Stability

Model convergence and stability are crucial for the successful application of reinforcement learning. The training loss during episodes indicates the convergence of the Q-network. Figure 4.54 depicts the training loss over the training steps.

In summary, the results demonstrate the effectiveness of the reinforcement learning approach for HVAC control, as evidenced by the increasing total rewards, decreasing validation loss, and stable training process. The exploration-exploitation strategy ensures a balance between exploring new actions and exploiting learned knowledge, contributing to the overall success of the trained agent.

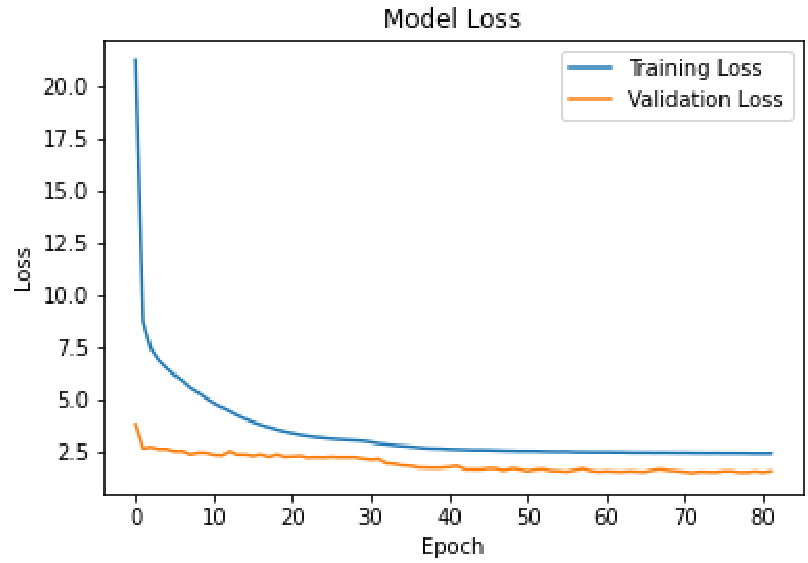


Figure 4.12. Loss Plot for 1D CNN 'Consumption Cost' Prediction - Previous last week and corresponding week in last year's month.

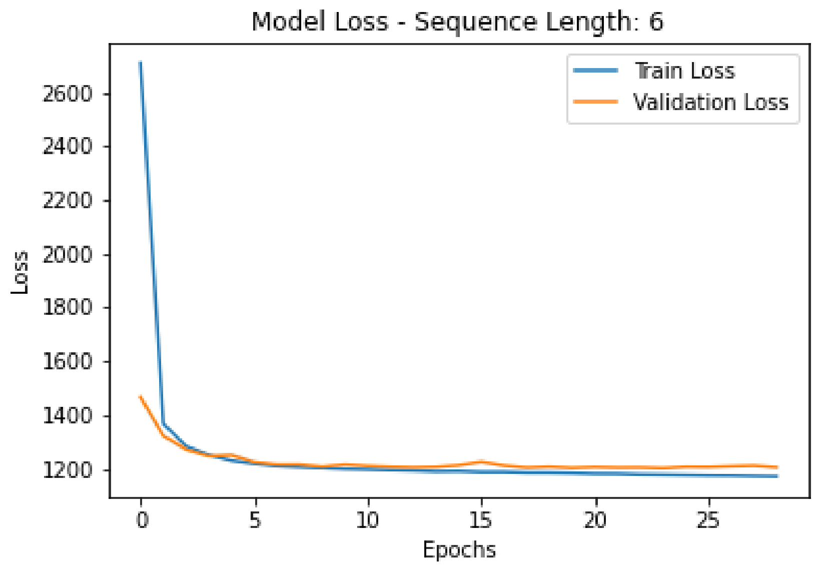


Figure 4.13. Loss Plot for LSTM 'Consumption Cost' Prediction - Sequence Length: 6

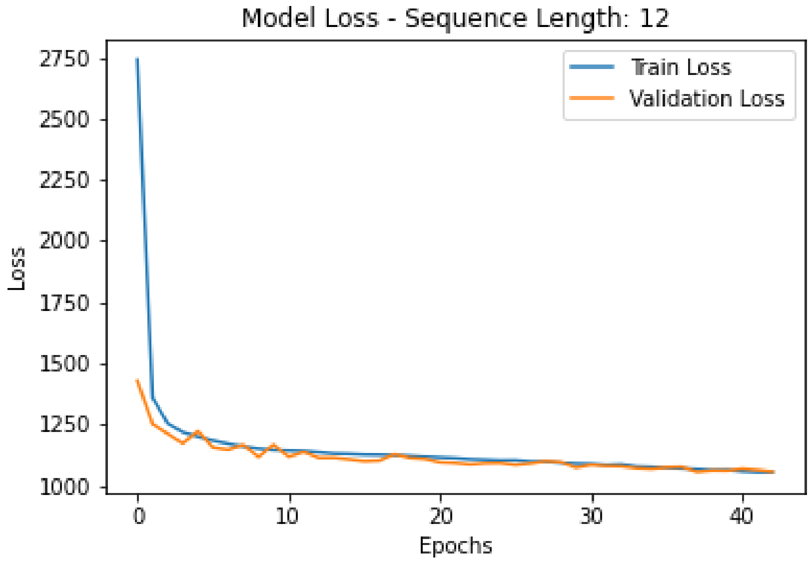


Figure 4.14. Loss Plot for LSTM 'Consumption Cost' Prediction - Sequence Length: 12

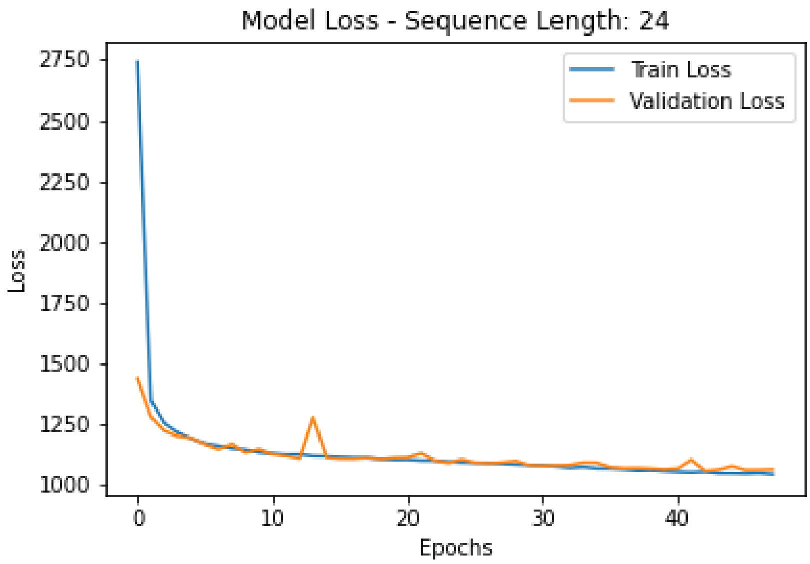


Figure 4.15. Loss Plot for LSTM 'Consumption Cost' Prediction - Sequence Length: 24

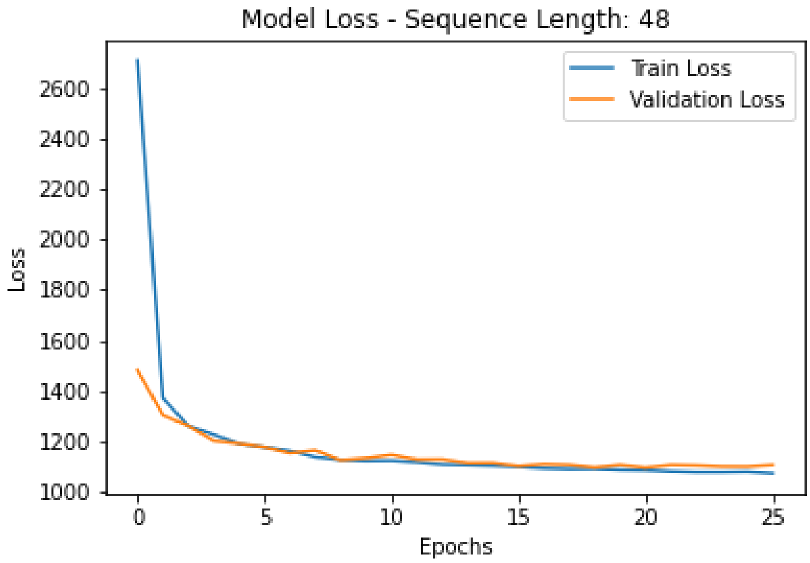


Figure 4.16. Loss Plot for LSTM 'Consumption Cost' Prediction - Sequence Length: 48

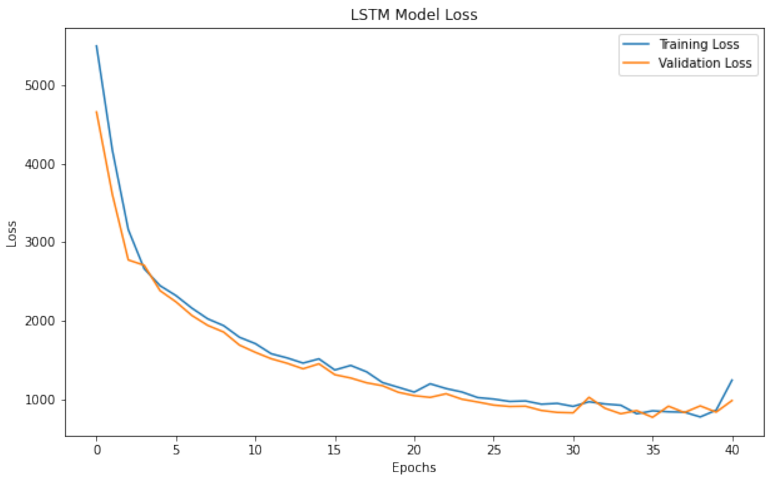
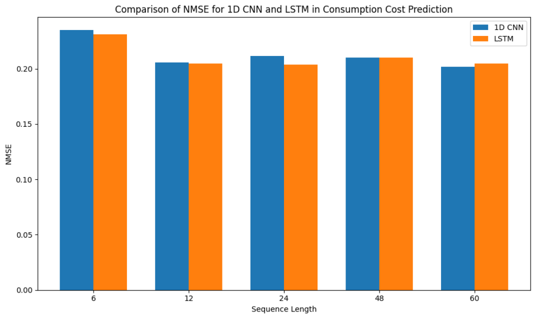
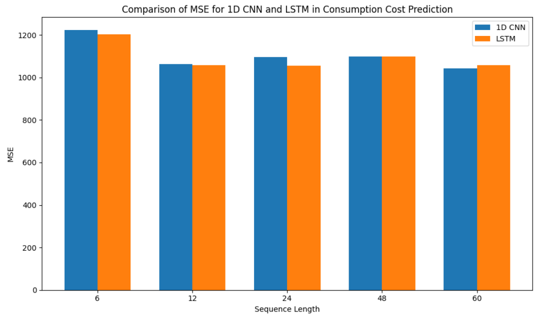


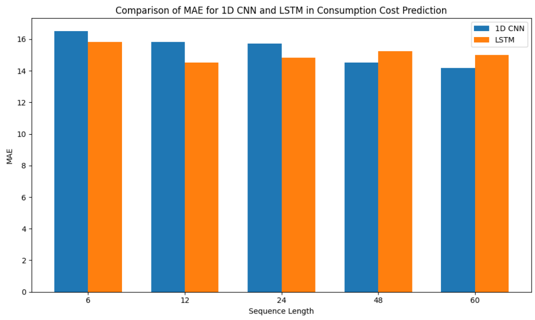
Figure 4.17. Loss Plot for LSTM 'Consumption Cost' Prediction - Previous last week and corresponding week in last year's month.





(a) NMSE for Consumption Cost Prediction in deep learning.

(b) MSE for Consumption Cost Prediction in deep learning.



(c) MAE for Consumption Cost Prediction in deep learning.

Figure 4.18. Comparison of performance metrics for Consumption Cost Prediction models in deep learning.

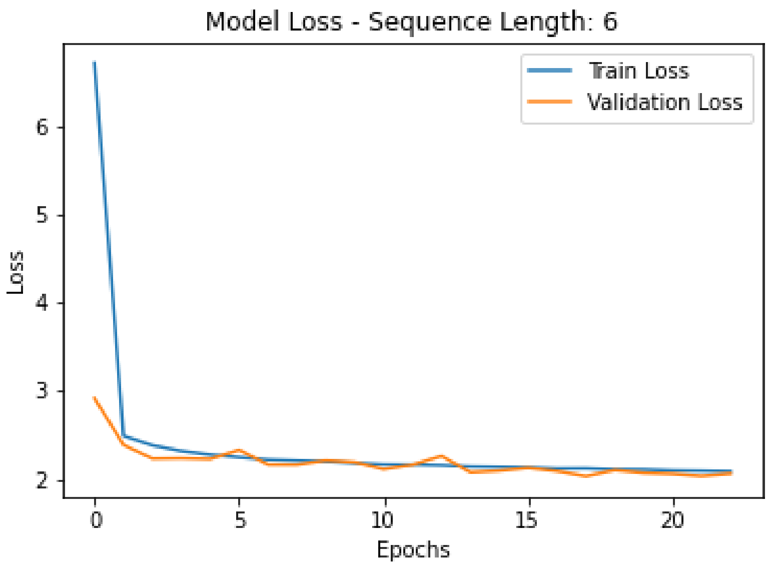


Figure 4.19. Loss Plot for 'Zone Mean Air Temperature' Prediction - Sequence Length: 6

Table 4.5. Model Performance Metrics for 1D CNN 'Zone Mean Air Temperature' Prediction

Sequence Length	**MSE**	**NMSE**	**MAE**
6	2.033	0.145	0.945
12	2.042	0.145	1.008
24	2.102	0.149	0.980
48	1.999	0.142	0.968
60	2.058	0.146	0.947
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week	0.639	0.045	0.533
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week + 6 hrs. in Last to Last Week	0.453	0.032	0.427
Previous Day 24 hrs. to Predict Next 6 hrs	0.877	0.062	0.626
Previous Last Day and Corresponding Day in Last Week	0.706	0.049	0.571
Previous Last Day and Corresponding Day in Last Week + Corresponding Day in Last to Last Week	0.763	0.054	0.606
Previous last week and corresponding week in last year's month	1.53	0.11	0.93

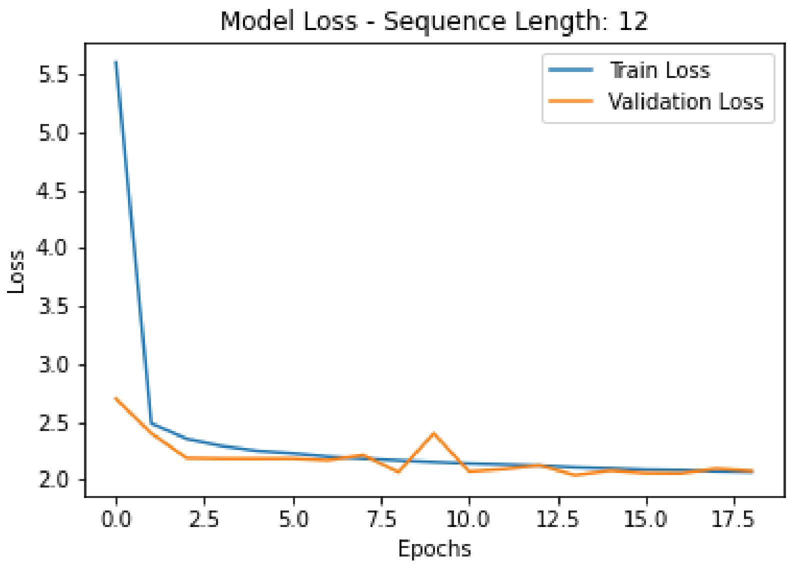


Figure 4.20. Loss Plot for 'Zone Mean Air Temperature' Prediction - Sequence Length: 12

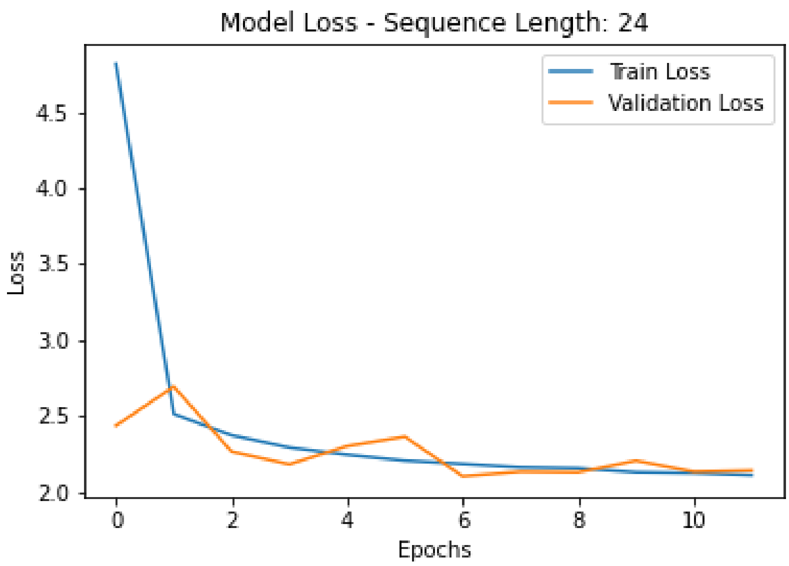


Figure 4.21. Loss Plot for 'Zone Mean Air Temperature' Prediction - Sequence Length: 24

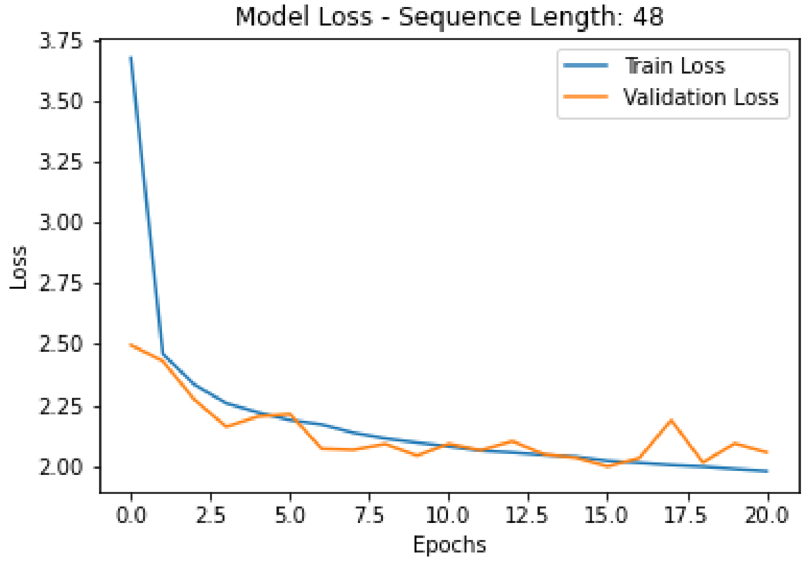


Figure 4.22. Loss Plot for 'Zone Mean Air Temperature' Prediction - Sequence Length: 48

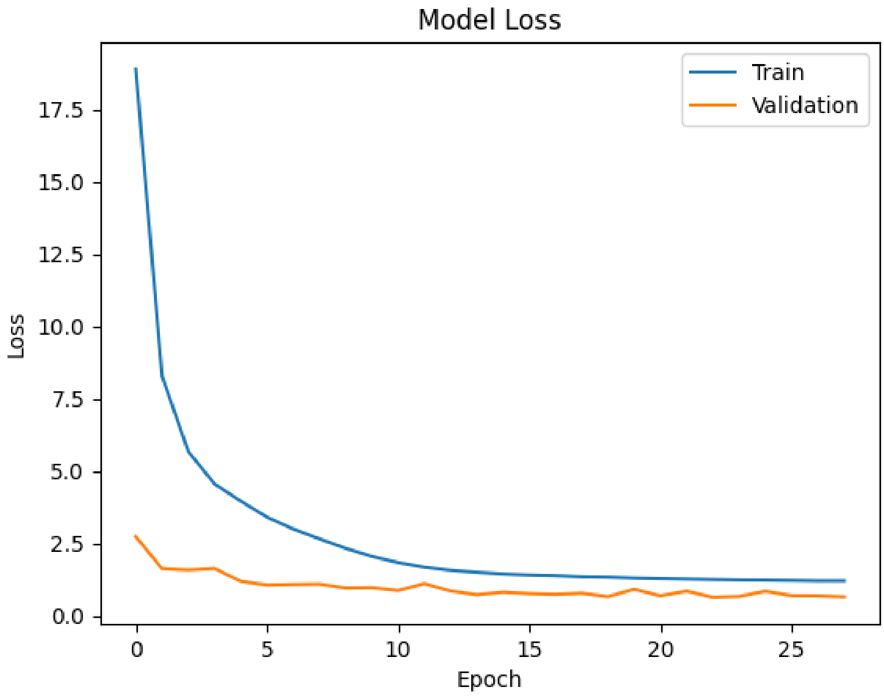


Figure 4.23. Loss Plot for 'Zone Mean Air Temperature' Prediction - Previous 12 hrs to predict next 06 hrs.

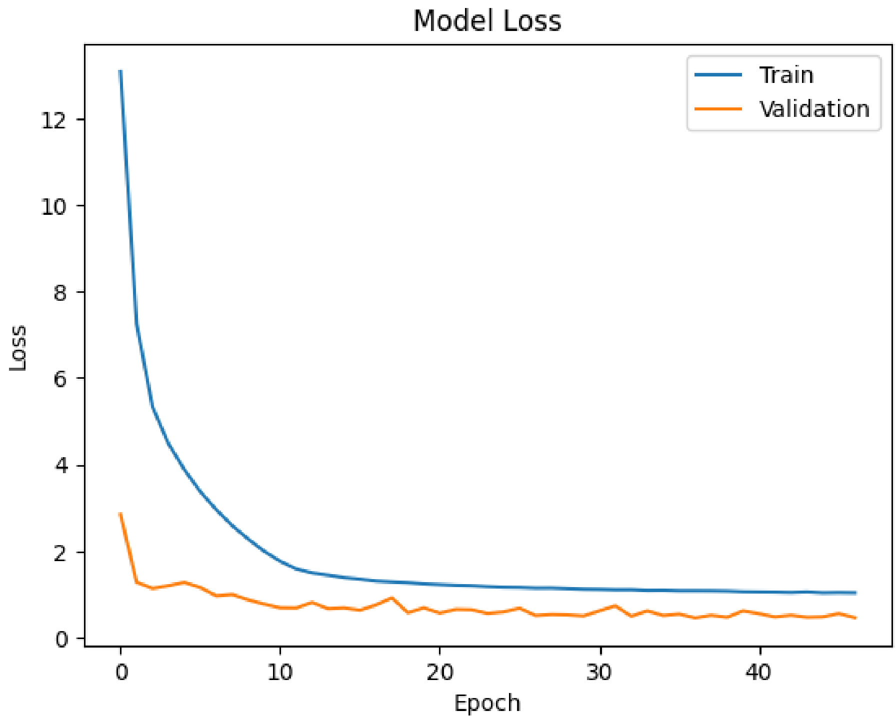


Figure 4.24. Loss Plot for 'Zone Mean Air Temperature' Prediction - Previous 18 hrs to predict next 06 hrs.

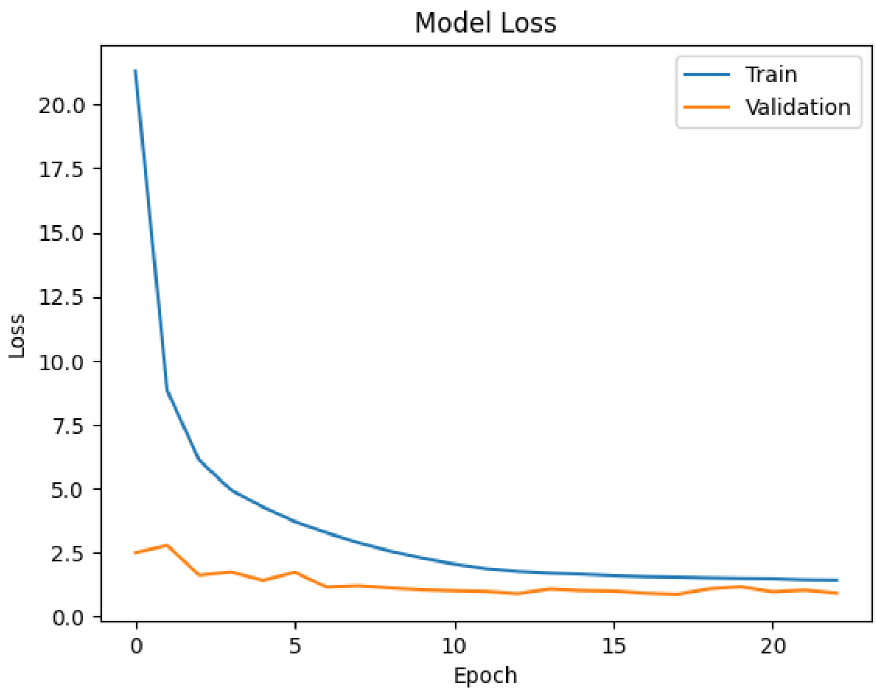


Figure 4.25. Loss Plot for 'Zone Mean Air Temperature' Prediction - Previous 24 hrs to predict next 06 hrs.

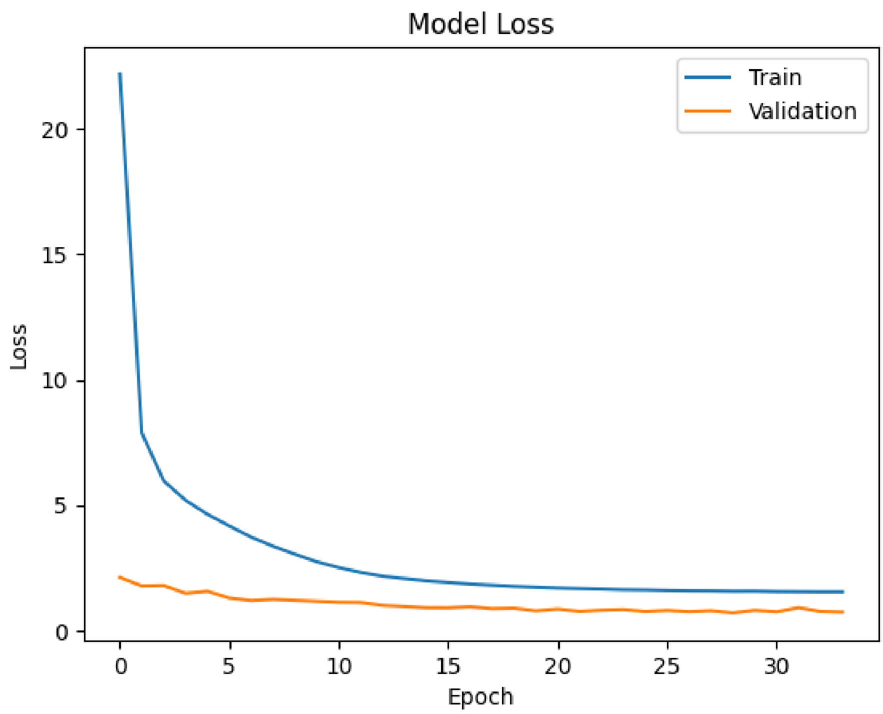


Figure 4.26. Loss Plot for 'Zone Mean Air Temperature' Prediction - Previous 48 hrs to predict next 24 hrs.

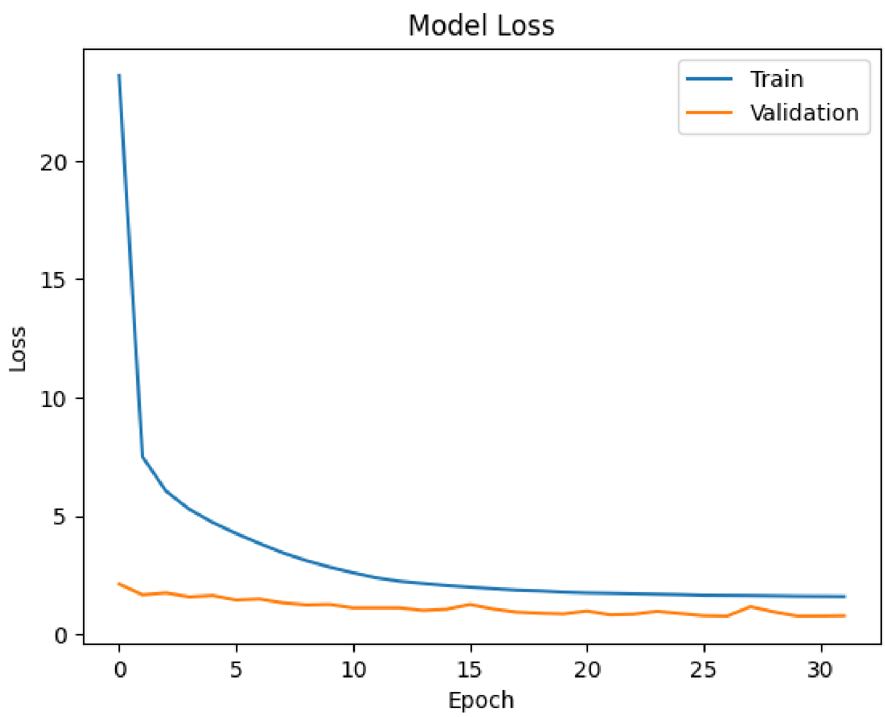


Figure 4.27. Loss Plot for 'Zone Mean Air Temperature' Prediction - Previous 72 hrs to predict next 24 hrs.

Table 4.6. Model Performance Metrics for LSTM 'Zone Mean Air Temperature' Prediction

Sequence Length	**MSE**	**NMSE**	**MAE**
6	1.9576	0.1395	0.9243
12	1.9045	0.1354	0.9236
24	1.8811	0.1333	0.9084
48	1.9067	0.1356	0.9251
60	1.8675	0.1325	0.9060
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week	1.928	0.1374	0.9293
Previous Day 6 hrs. and Corresponding Day 6 hrs. in Last Week + 6 hrs. in Last to Last Week	1.8966	0.1345	0.9160
Previous Day 24 hrs. to Predict Next 6 hrs	2.24	0.1460	0.9745
Previous Last Day and Corresponding Day in Last Week	1.97	0.1405	0.9565
Previous Last Day and Corresponding Day in Last Week + Corresponding Day in Last to Last Week	2.003	0.1425	0.9705
Previous last week and corresponding week in last year's month	1.16	0.084	0.79

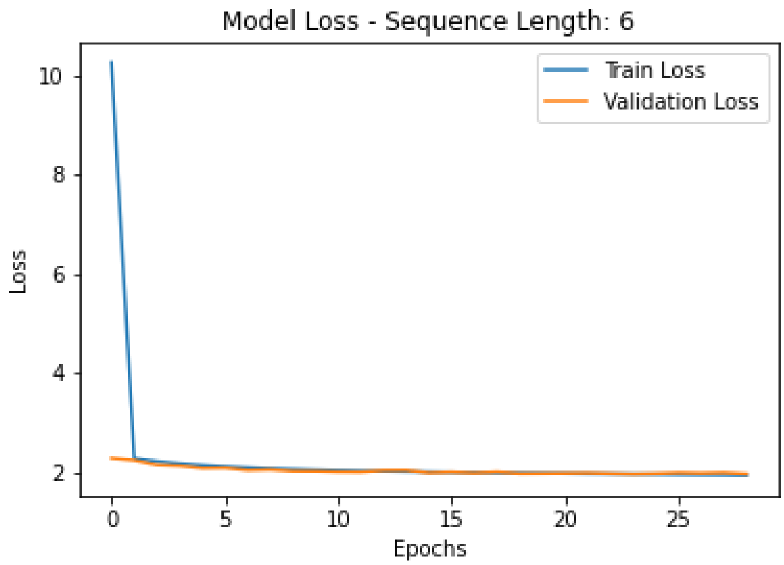


Figure 4.28. Loss Plot for LSTM 'Zone Mean Air Temperature' Prediction - Sequence Length: 6

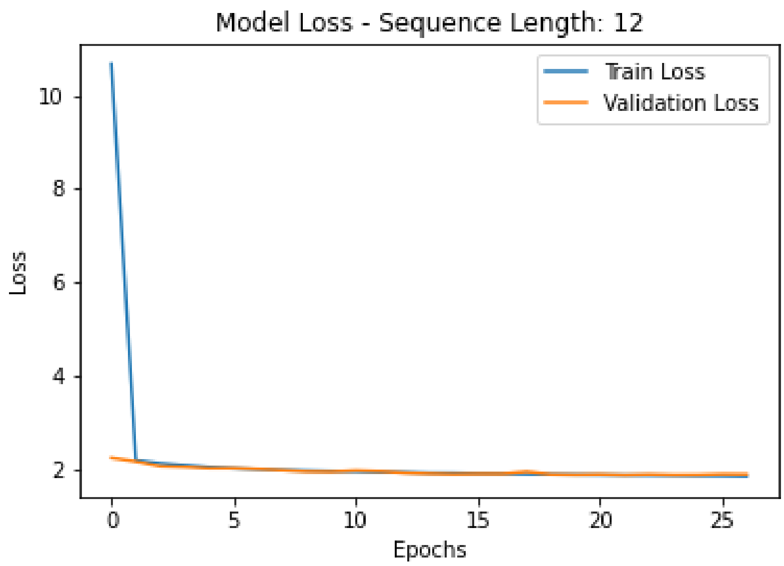


Figure 4.29. Loss Plot for LSTM 'Zone Mean Air Temperature' Prediction - Sequence Length: 12

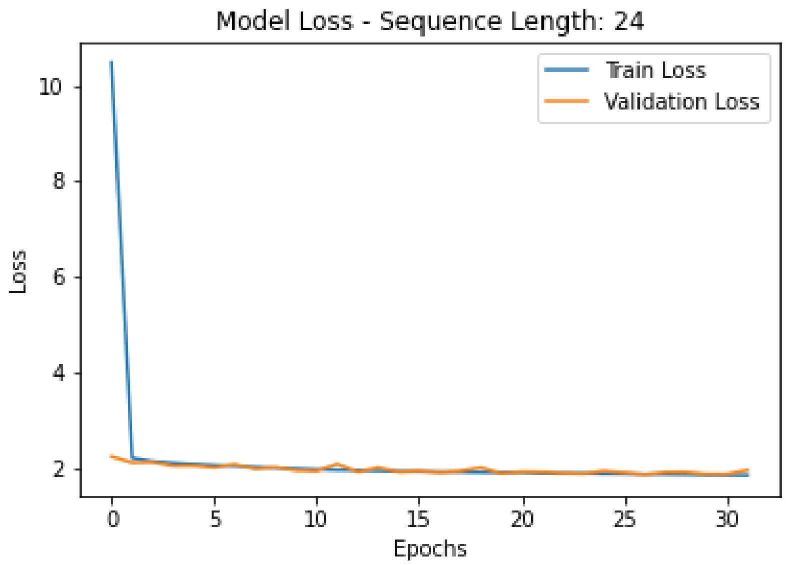


Figure 4.30. Loss Plot for LSTM 'Zone Mean Air Temperature' Prediction - Sequence Length: 24

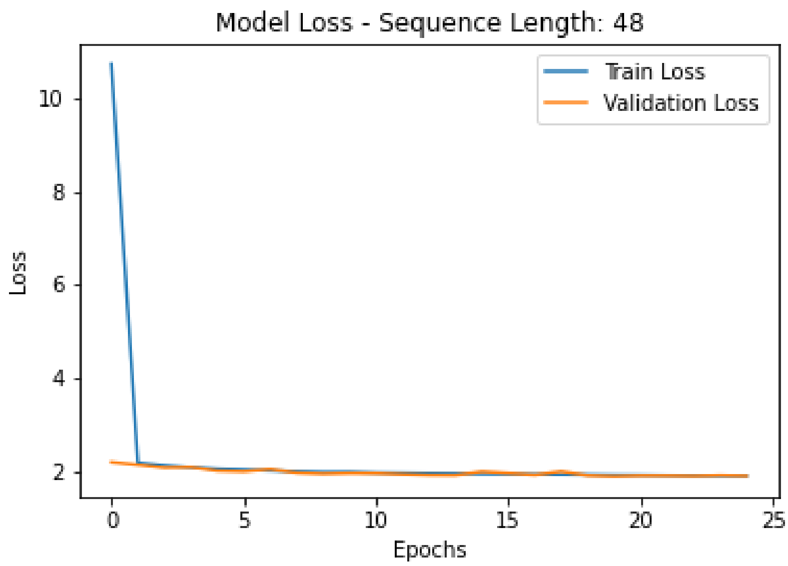


Figure 4.31. Loss Plot for LSTM 'Zone Mean Air Temperature' Prediction - Sequence Length: 48

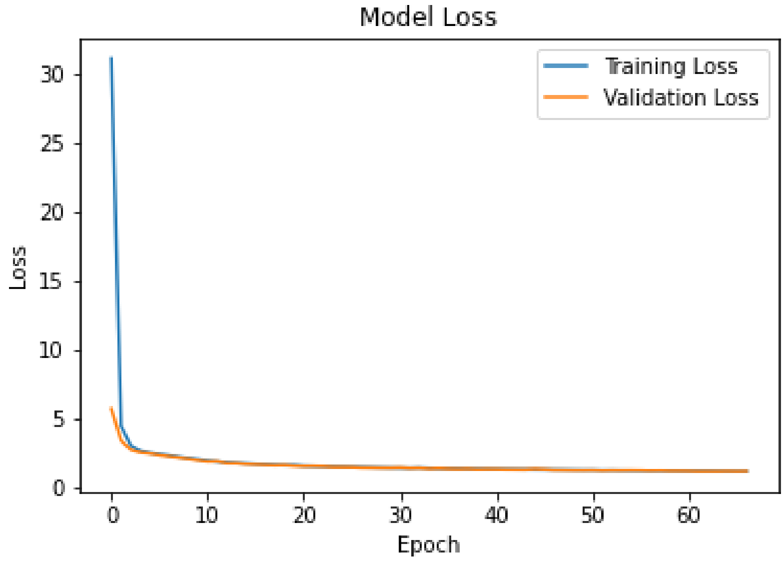
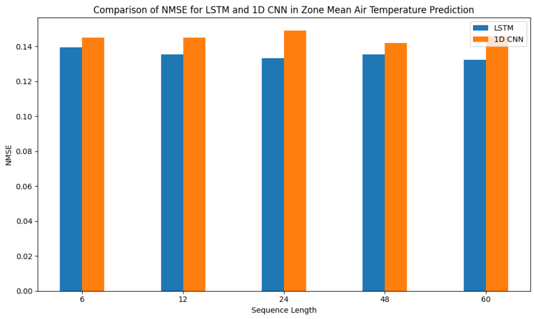
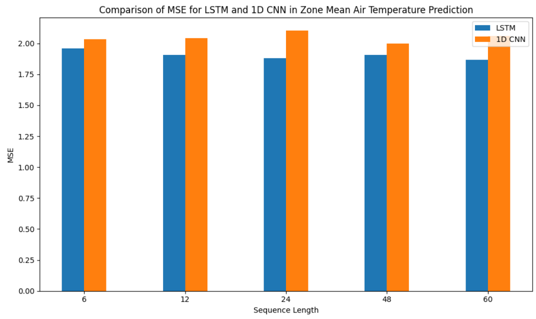


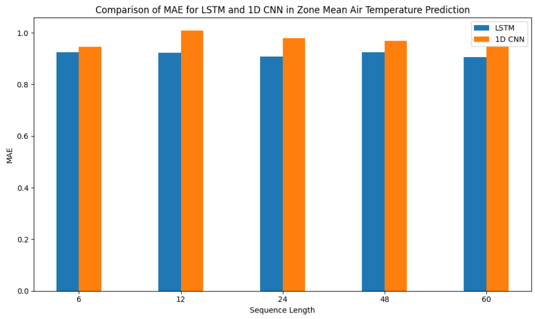
Figure 4.32. Loss Plot for LSTM 'Zone Mean Air Temperature' Prediction - Previous last week and corresponding week in last year's month.



(a) NMSE for Consumption Cost Prediction in deep learning.



(b) MSE for Consumption Cost Prediction in deep learning.



(c) MAE for Consumption Cost Prediction in deep learning.

Figure 4.33. Comparison of performance metrics for Zone Mean Air Temperature Prediction models in deep learning.

Table 4.7. Model Performance Metrics for LSTM 'Zone Mean Air Temperature' & 'Consumption Cost' Prediction

Sequence Length	MSE	MAE	NMSE
Previous day 6 hrs. and corresponding day 6hrs in last week	0.002	0.028	0.051
Previous day 6 hrs. and corresponding day 6hrs in last week + 6 hrs. in last to last week	0.002	0.028	0.047
Previous day 24 hrs. to predict next 6hrs	0.002	0.027	0.048
Previous last day and corresponding day in last week	0.002	0.028	0.051
Previous last day and corresponding day in last week and corresponding day in last to last week	0.002	0.028	0.050
Previous last week and corresponding week in last year's month	4.39	0.78	0.001

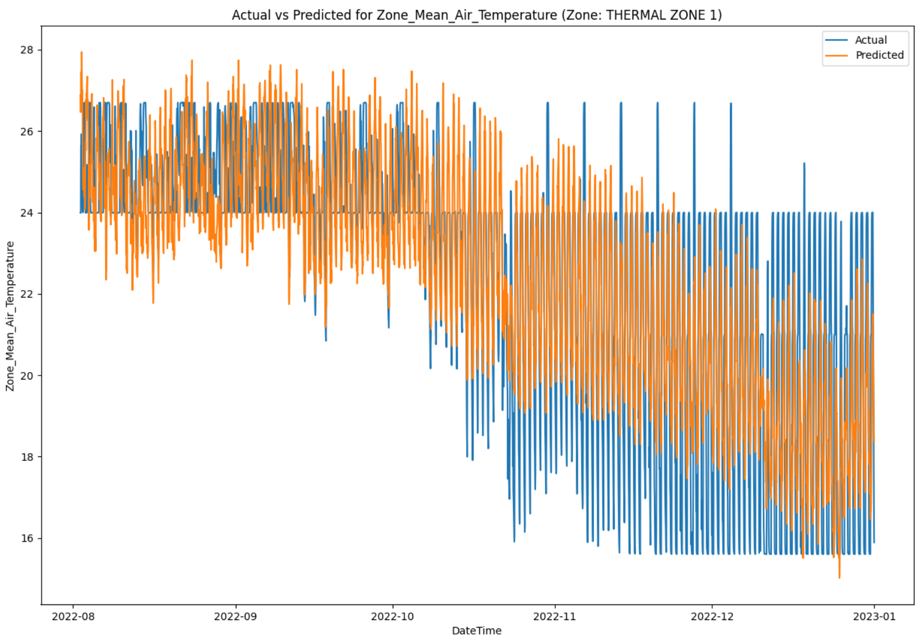


Figure 4.34. LSTM Multi Zone Mean Air Temperature Zone THERMAL ZONE 1

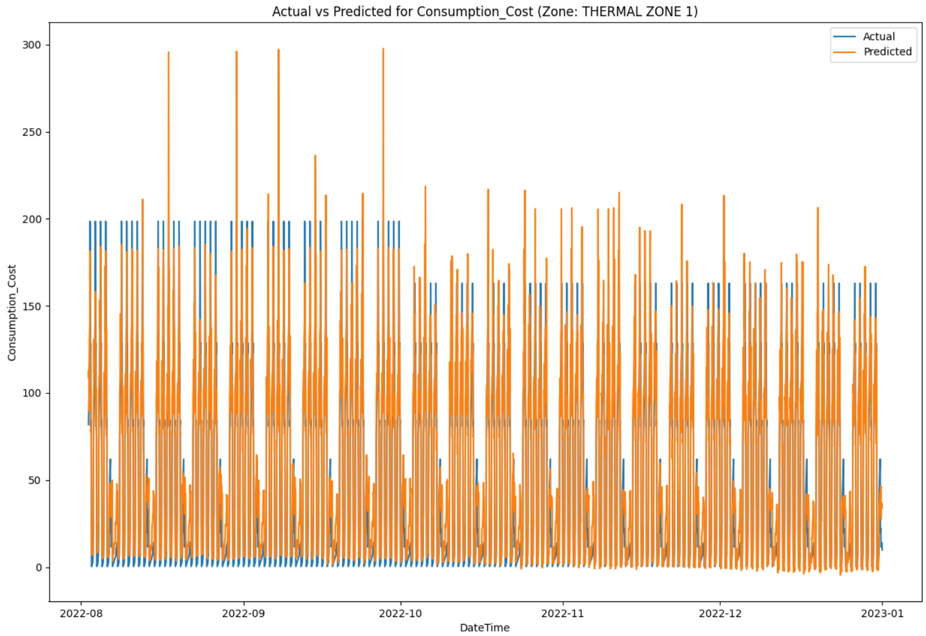


Figure 4.35. LSTM Multi Consumption Cost Zone THERMAL ZONE 1

Table 4.8. Model Performance Metrics for 1D CNN 'Zone Mean Air Temperature' & 'Consumption Cost' Prediction

Sequence Length	MSE	MAE	NMSE
Previous day 6 hrs. and corresponding day 6hrs in last week	0.0020	0.0280	0.0520
Previous day 6 hrs. and corresponding day 6hrs in last week + 6 hrs. in last to last week	0.0020	0.0280	0.0520
Previous day 24 hrs. to predict next 6hrs	0.0020	0.0290	0.0510
Previous last day and corresponding day in last week	0.0030	0.0310	0.0580
Previous last day and corresponding day in last week and corresponding day in last to last week	0.0030	0.0340	0.0690
Previous last week and corresponding week in last year's month	131.75	8.08	0.051

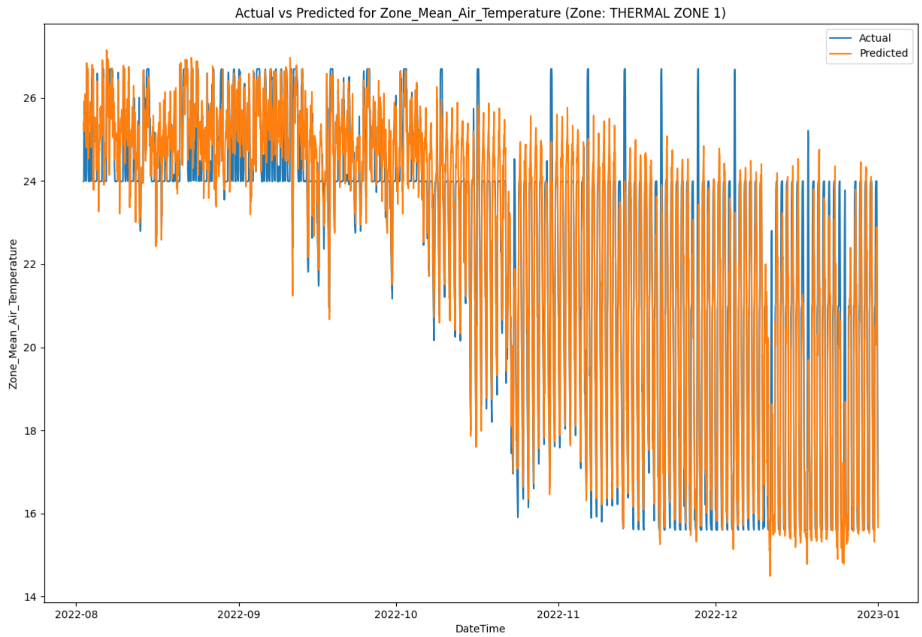


Figure 4.36. 1DCNN Multi Zone Mean Air Temperature Zone THERMAL ZONE 1

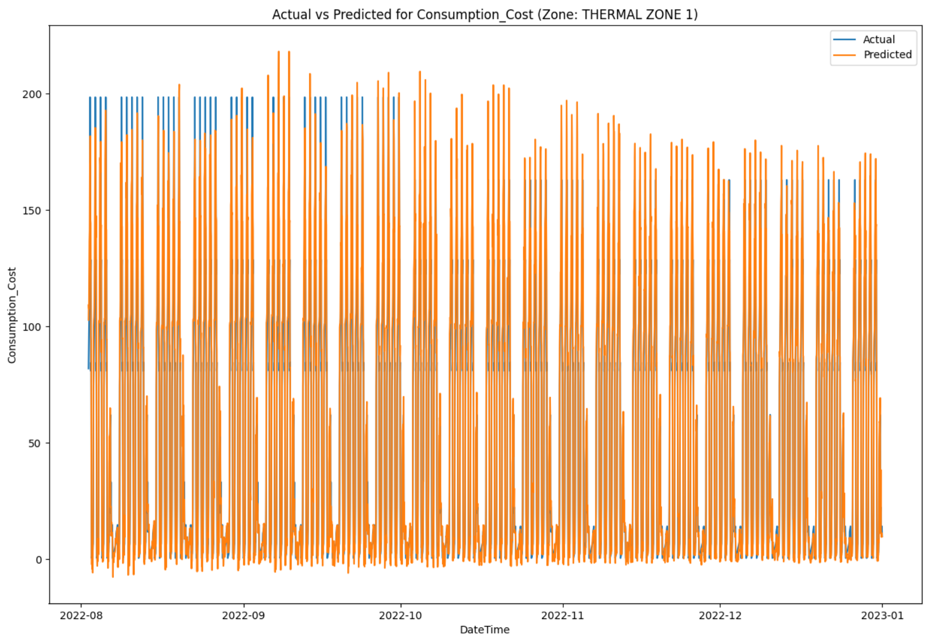


Figure 4.37. 1DCNN Multi Consumption Cost Zone THERMAL ZONE 1

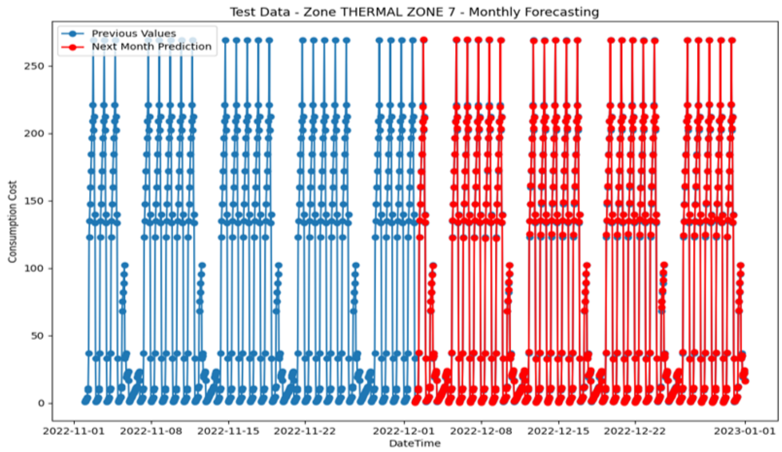


Figure 4.38. 1DCNN CC_Monthly Prediction

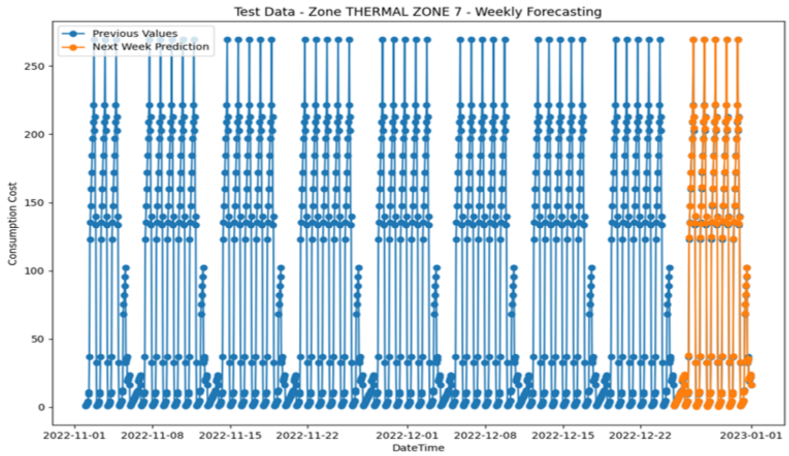


Figure 4.39. 1DCNN CC_Weekly Prediction

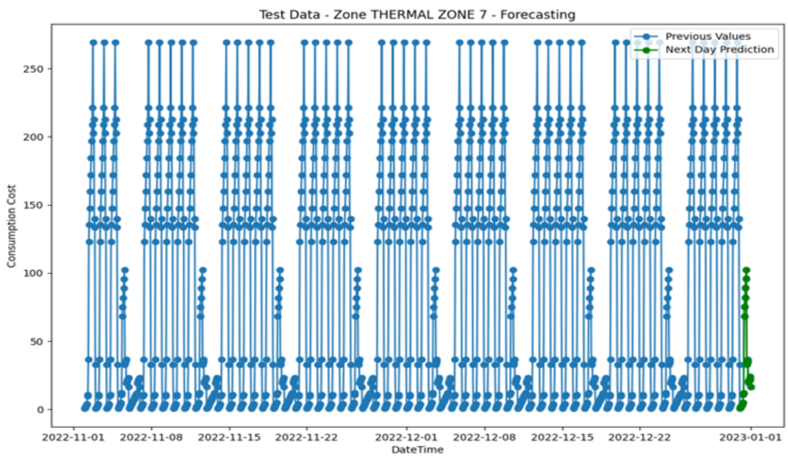


Figure 4.40. 1DCNN CC_next_day Prediction

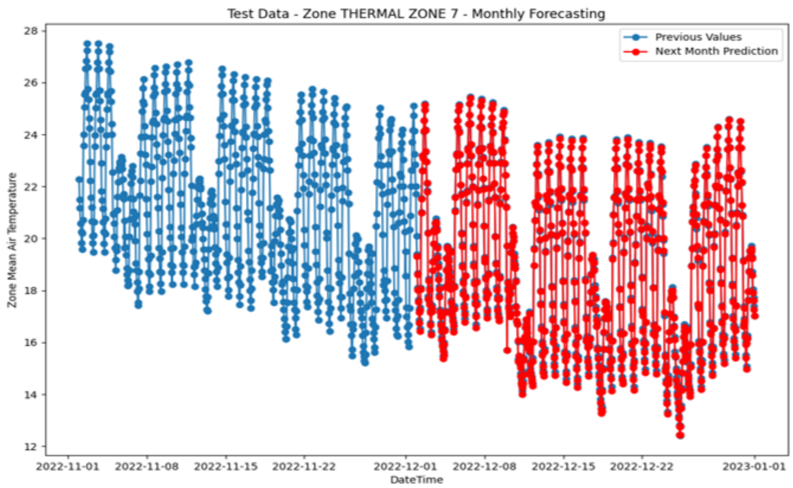


Figure 4.41. 1DCNN ZMAT_Monthly Prediction

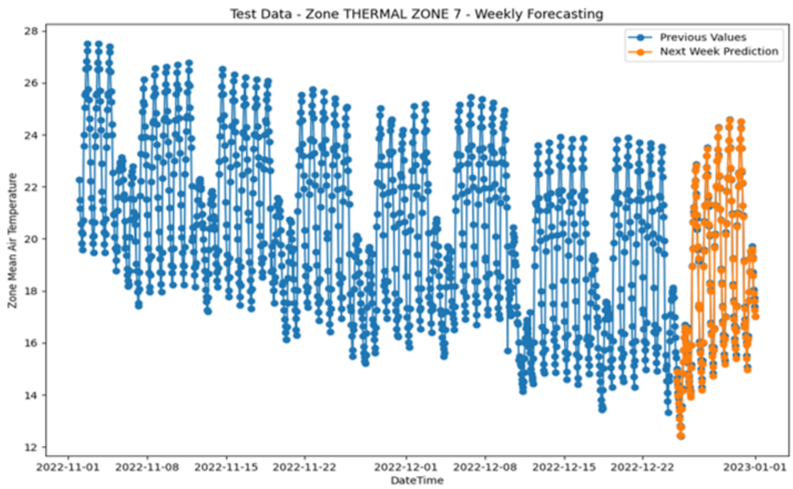


Figure 4.42. 1DCNN ZMAT_Weekly Prediction

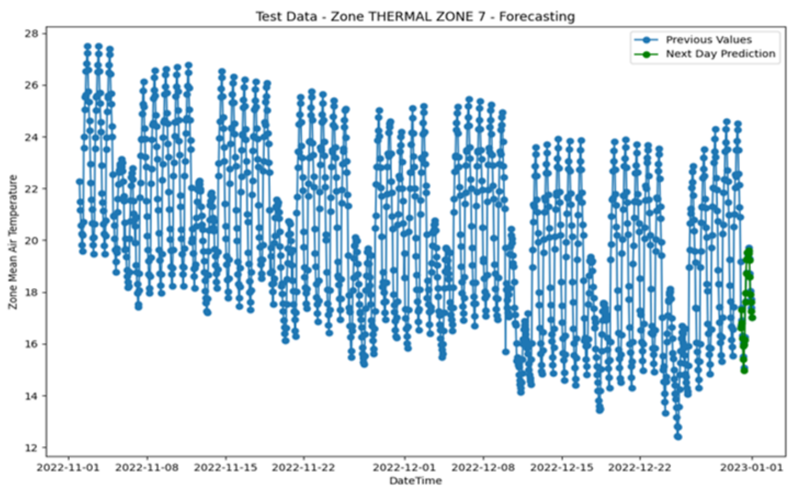


Figure 4.43. 1DCNN ZMAT_next_day Prediction

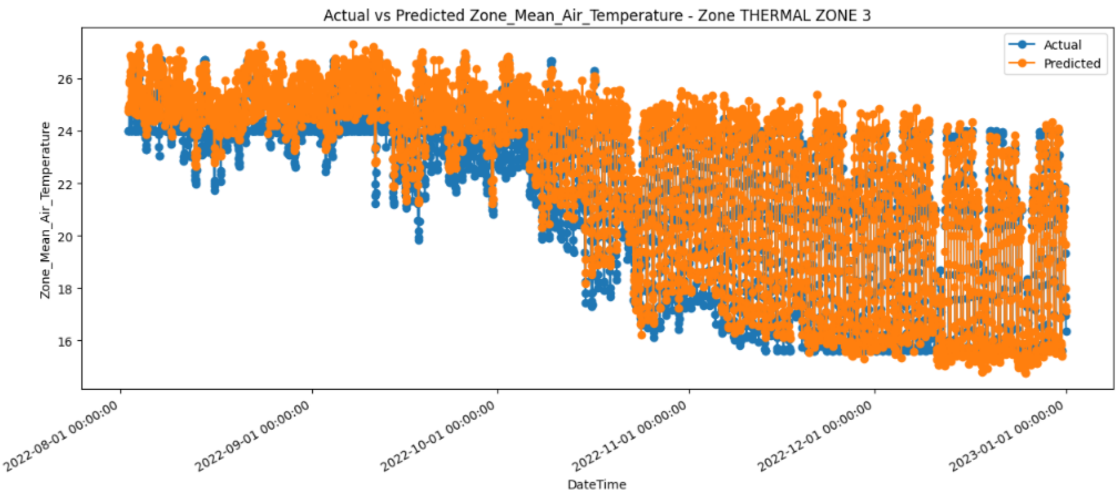


Figure 4.44. LSTM ZMAT_next_day Prediction

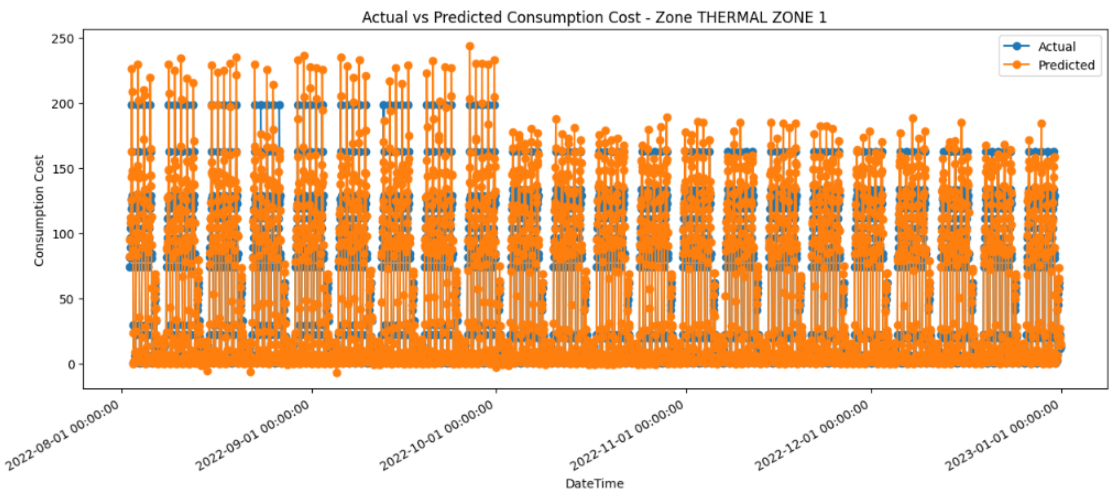


Figure 4.45. LSTM Consumption_Cost_next_day Prediction

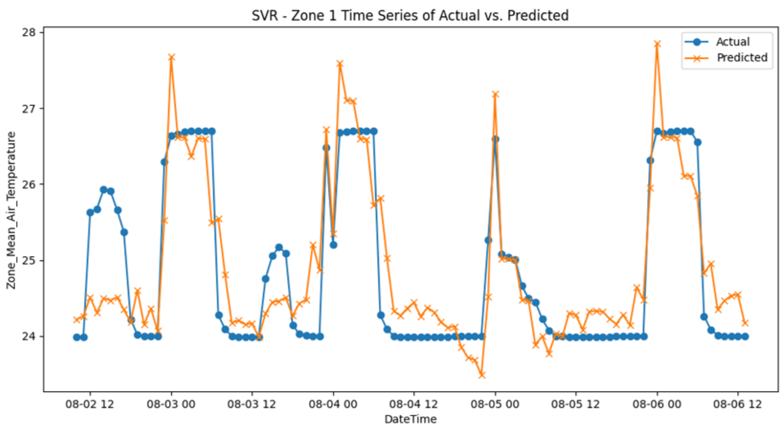


Figure 4.46. ZMAT SVR Prediction

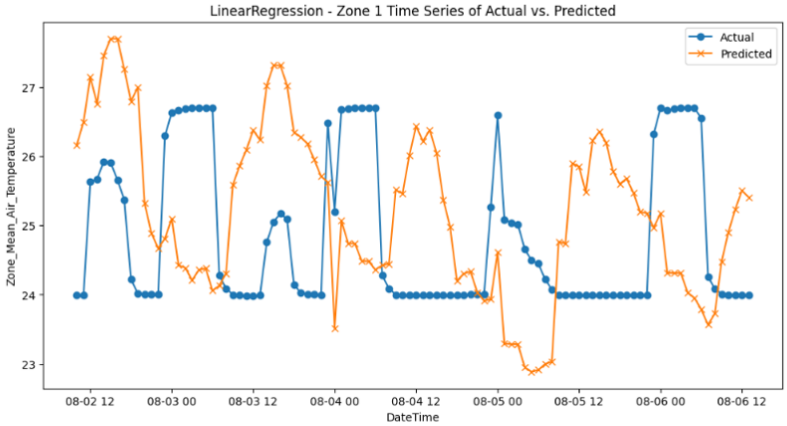


Figure 4.47. ZMAT Linear Regression Prediction

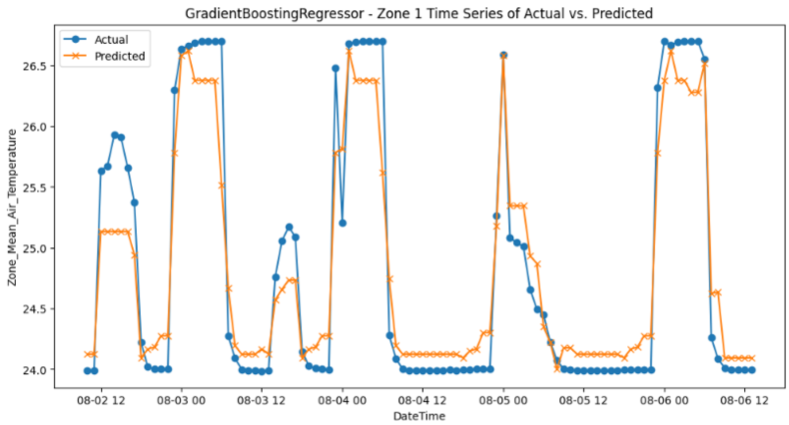


Figure 4.48. ZMAT Gradient Boosting Regression Prediction

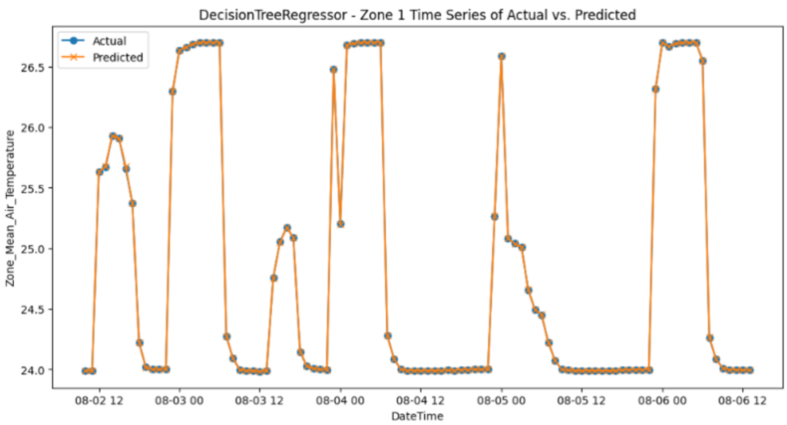


Figure 4.49. ZMAT Decision Tree Prediction

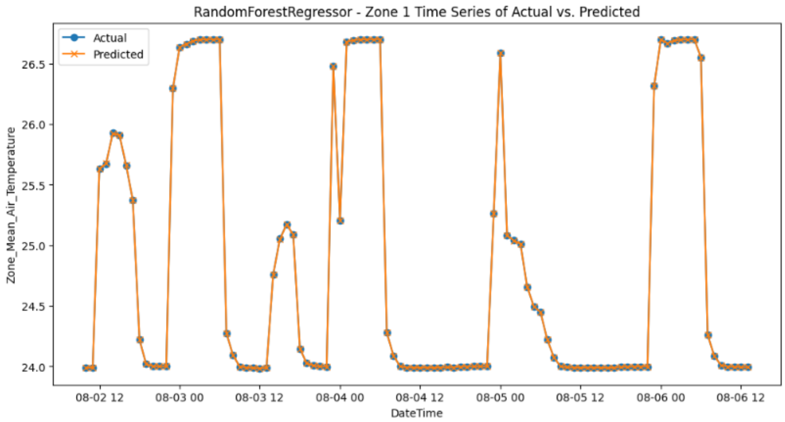


Figure 4.50. ZMAT Random Forest Prediction

Metric	Value
Mean Squared Error (MSE)	0.9924
Normalized Mean Squared Error (NMSE)	0.15725
Mean Absolute Error (MAE)	0.5657
R-squared (R2)	0.8427

Table 4.9. Evaluation Metrics for DRL for HVAC Control

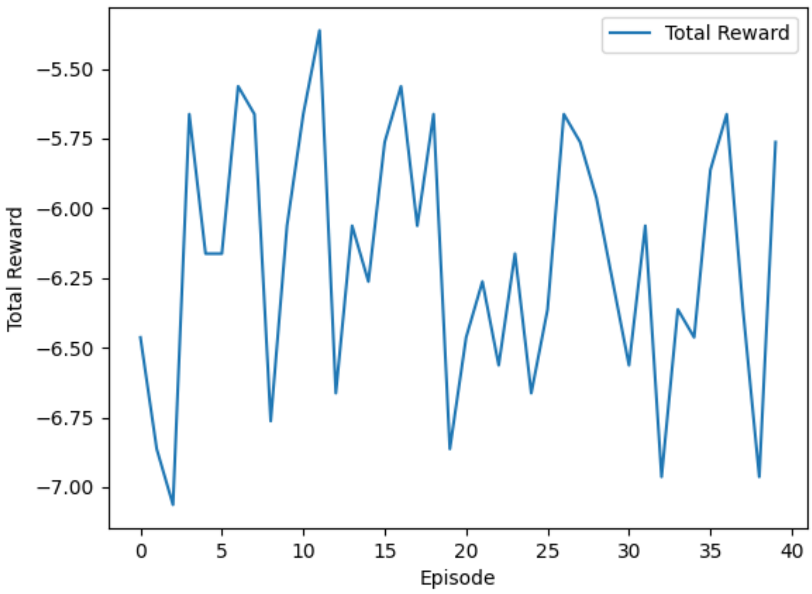


Figure 4.51. Visualization of Total Rewards Plot in Deep Q-Network (DeepQ) Training Process

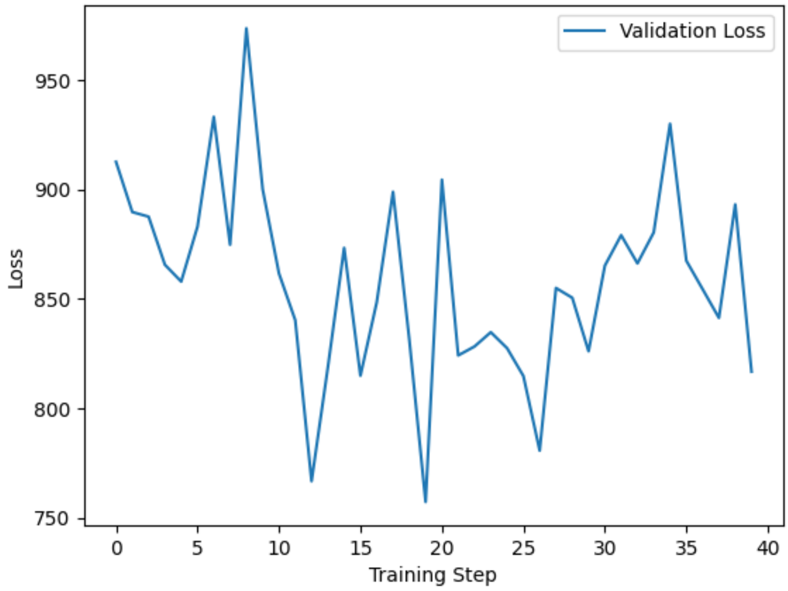


Figure 4.52. Visualization of Validation Loss Plot in Deep Q-Network (DeepQ) Training Process

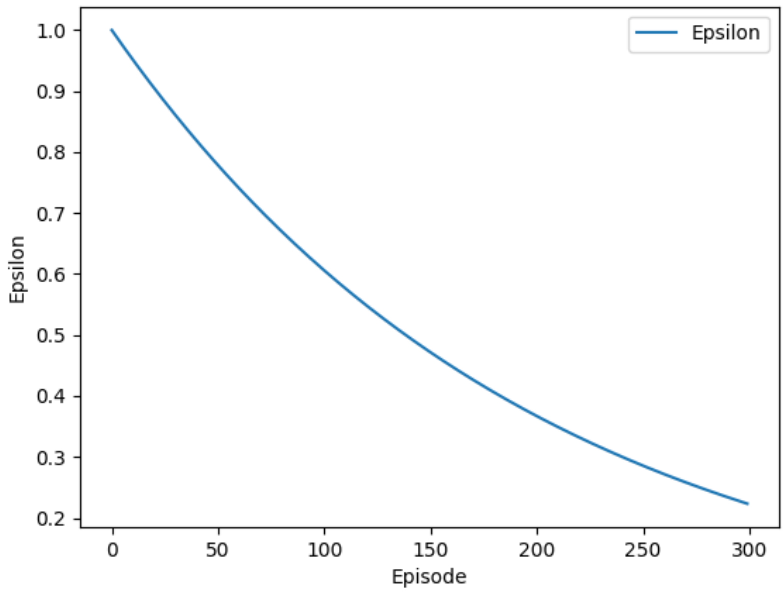


Figure 4.53. Visualization of Epsilon Decay Plot in Deep Q-Network (DeepQ) Training Process

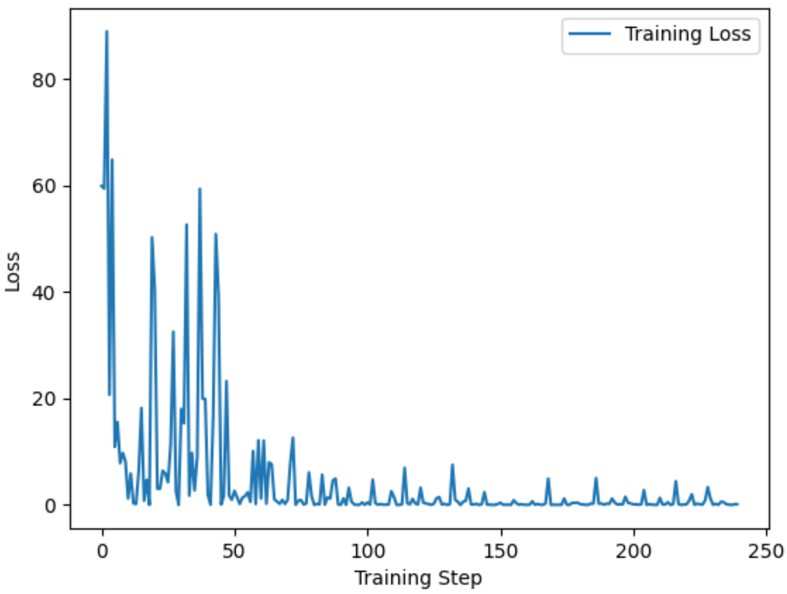


Figure 4.54. Visualization of Training Loss in Deep Q-Network (DeepQ) Training Process

CHAPTER 5

Discussion

Time Series Forecasting

1D CNN Model - Consumption Cost Prediction

The 1D CNN model consistently demonstrated high performance, quantified by decreasing Mean Squared Error ($\mathrm{MSE}_{\mathrm{CNN}}$(L)) and consistently high $R^2_{\mathrm{CNN}}(L)$ values. Specifically, let $\mathrm{MSE}_{\mathrm{CNN}}(L)$ represent the MSE for a given sequence length L. The $R^2_{\mathrm{CNN}}(L)$ values indicated a strong linear relationship between predicted and actual consumption costs.

Different forecasting strategies, incorporating data from the previous day and corresponding day in the last week, revealed promising results with low $\mathrm{MSE}_{\mathrm{CNN}}(L)$ and high $R^2_{\mathrm{CNN}}(L)$. Loss plots, displaying the training and validation losses, exhibited stable convergence, with the loss function $\mathscr{L}_{\mathrm{CNN}}(L)$ decreasing as sequence length increased.

LSTM Model - Consumption Cost Prediction

Similar to the 1D CNN model, the LSTM model's performance was assessed through $\mathrm{MSE}_{\mathrm{LSTM}}(L)$ and $R^2_{\mathrm{LSTM}}(L)$. These metrics illustrated a decreasing MSE and high $R^2_{\mathrm{LSTM}}(L)$ as sequence length increased, highlighting the model's efficacy. Loss plots, depicting $\mathscr{L}_{\mathrm{LSTM}}(L)$, demonstrated stability and convergence patterns comparable to the 1D CNN model.

1D CNN Model - Zone Mean Air Temperature Prediction

For predicting Zone Mean Air Temperature using the 1D CNN model, performance was evaluated using $\mathrm{MSE}_{\mathrm{CNN,\,Temp}}(L)$ and $R^2_{\mathrm{CNN,\,Temp}}(L)$. Results showed consistent improvements with increasing sequence length. Loss plots, indicating $\mathscr{L}_{\mathrm{CNN,\,Temp}}(L)$, displayed stable convergence patterns.

LSTM Model - Zone Mean Air Temperature Prediction

The LSTM model's performance in Zone Mean Air Temperature prediction was quantified by $\mathrm{MSE}_{\mathrm{LSTM,\,Temp}}(L)$ and $R^2_{\mathrm{LSTM,\,Temp}}(L)$. Similar to the 1D CNN model, varying performance was observed across forecasting strategies, with some achieving low $\mathrm{MSE}_{\mathrm{LSTM,\,Temp}}(L)$ and high $R^2_{\mathrm{LSTM,\,Temp}}(L)$. Loss plots for the LSTM model exhibited stability and convergence trends.

Multi-Target Prediction

In the joint prediction of 'Zone Mean Air Temperature' and 'Consumption Cost,' both the LSTM and 1D CNN models demonstrated their effectiveness. Performance metrics for multi-target prediction, denoted as $\mathrm{MSE}_{\mathrm{Multi}}$ and R^2_{Multi}, highlighted the models' ability to handle simultaneous predictions.

Reinforcement Learning for HVAC Control

The reinforcement learning approach for HVAC control exhibits positive outcomes. The increasing trend in total rewards indicates that the agent is learning and adapting its control strategy to optimize energy efficiency and thermal comfort. The validation loss shows a decreasing trend, suggesting that the learned policy generalizes well to unseen scenarios. The epsilon decay plot demonstrates a gradual reduction in exploration as the agent becomes more knowledgeable about the environment.

The training loss plot indicates model convergence, and the stability of the training process is crucial for the successful application of reinforcement learning. The results collectively affirm the effectiveness of the deep Q-learning algorithm in training an HVAC control agent. Model convergence, observed through decreasing training loss $\mathscr{L}_{\mathrm{RL}}$, indicated stability in the training process.

Common Trends

For both time series forecasting models, the improvement in performance with longer sequence lengths was expressed through mathematical metrics. Various forecasting strategies were quantified in terms of MSE and R^2, providing numerical insights into their effectiveness.

Comparison between 1D CNN and LSTM

In the context of time series forecasting for 'Consumption Cost' and 'Zone Mean Air Temperature,' both 1D CNN and LSTM models were evaluated. The comparison below highlights key aspects of their performance across different sequence lengths.

1D CNN vs. LSTM for 'Consumption Cost' Prediction

- **Model Performance Metrics:** Both 1D CNN and LSTM models were trained and evaluated for predicting 'Consumption Cost' at various sequence lengths. The performance metrics, including Mean Squared Error (MSE), Normalized Mean Squared Error (NMSE), Mean Absolute Error (MAE), and R^2 were measured and compared.

- **Loss Plots:** Loss plots for both models at different sequence lengths provide insights into the convergence and performance of each model. Figures 4.3, 4.4, 4.5, and 4.6 depict the loss evolution for 1D CNN, while Figures 4.13, 4.14, 4.15, and 4.16 show the corresponding plots for the LSTM model.

- **Observations:** The comparison of performance metrics and loss plots reveals that the choice between 1D CNN and LSTM depends on the specific sequence length and the trade-off between computational complexity and accuracy. In some cases, one model may outperform the other, highlighting the importance of tailoring the model to the specific characteristics of the data.

- **Best Practices:** Understanding the strengths and weaknesses of both models allows for informed decision-making. For example, if capturing long-term dependencies is crucial, LSTM may be preferred, while 1D CNN might excel in capturing local patterns. The choice could also depend on computational efficiency and training time.

1D CNN vs. LSTM for 'Zone Mean Air Temperature' Prediction

- **Model Performance Metrics:** Similar to the 'Consumption Cost' prediction, 1D CNN and LSTM models were compared for predicting 'Zone Mean Air Temperature' at various sequence lengths. Metrics such as MSE, NMSE, MAE, and R^2 were measured and compared.

- **Loss Plots:** Loss plots for both models at different sequence lengths provide insights into the convergence and performance of each model. Figures 4.19, 4.20, 4.21, and 4.22 depict the loss evolution for 1D CNN, while Figures 4.28, 4.29, 4.30, and 4.31 show the corresponding plots for the LSTM model.

- **Observations:** The comparison between 1D CNN and LSTM for 'Zone Mean Air Temperature' prediction follows a similar trend as in the 'Consumption Cost' prediction. The choice between models depends on factors such as the sequence length, the ability to capture temporal dependencies, and the specific characteristics of the data.

- **Best Practices:** Understanding the strengths and weaknesses of both models allows for informed decision-making. Consideration of factors such as the dataset's characteristics, computational efficiency, and the importance of short-term vs. long-term dependencies guides the selection of the most suitable model for the task.

In conclusion, the comparison between 1D CNN and LSTM models for time series forecasting emphasizes the importance of tailoring the choice of model to the specific requirements of the task at hand. Both models have their advantages, and the selection should be based on a thorough understanding of the data and the desired outcome.

Multi-Target Prediction

Mathematical metrics $\mathrm{MSE}_{\mathrm{Multi}}$ and R^2_{Multi} supported the discussion on the models' efficacy in multi-target prediction.

Reinforcement Learning for HVAC Control

Mathematical representations of total rewards, validation loss, and epsilon decay added precision to the assessment of reinforcement learning for HVAC control.

Practical Implications

Mathematical evaluations of forecasting models' accuracy contribute to their practical implementation for optimizing HVAC systems. Reinforcement learning's impact on energy efficiency and cost savings was emphasized through quantitative metrics.

Areas for Improvement

The need for further investigation into specific strategies and hyperparameter tuning was emphasized through quantitative assessments. Continuous monitoring and refinement of reinforcement learning algorithms were suggested to enhance long-term HVAC control.

Time Series Forecasting Summary Table

Table 5.1. Time Series Forecasting Performance Comparison with Trends

	'Consumption Cost' Prediction				**'Zone Mean Air Temperature' Prediction**			
Model	**6h**	**12h**	**24h**	**48h**	**6h**	**12h**	**24h**	**48h**
1D CNN MSE	1222.25	1063.15	1096.02	1098.46	2.033	2.042	2.102	1.999
LSTM MSE	1202.62	1058.10	1056.35	1097.75	1.9576	1.9045	1.8811	1.9067

In the table 5.1, red shades indicate decreasing trends, while green shades represent increasing trends. For the 'Consumption Cost' prediction task, the 1D CNN MSE model shows a decreasing trend in forecast accuracy from 6 hours to 12 hours, followed by an increasing trend from 12 hours to 48 hours. The LSTM MSE model, on the other hand, exhibits a decreasing trend in forecast accuracy from 6 hours to 24 hours but increases at the 48-hour mark.

For the 'Zone Mean Air Temperature' prediction, both models generally show an increasing trend in accuracy over time, with some fluctuations. The 1D CNN MSE model has a slight decrease in accuracy at 24 hours, while the LSTM MSE model demonstrates a decrease at 6 hours.

These trends provide insights into the performance of the models over different prediction horizons, aiding in the selection of the most suitable model for specific forecasting requirements.

5.1 Integration of Time Series Forecasting Models and Reinforcement Learning for Enhanced HVAC Control and Energy Efficiency

In the context of HVAC systems, integrating time series forecasting models with reinforcement learning (RL) proves to be a robust approach for achieving optimal control and energy efficiency. The models utilized, specifically the 1D CNN and LSTM, play a pivotal role in predicting key system parameters such as zone mean air temperatures and consumption costs.

These time series forecasting models contribute valuable insights by capturing temporal dependencies and understanding how the HVAC system responds to diverse external conditions. The accurate predictions generated by the models serve as crucial observations for the RL agent, enhancing its decision-making capabilities.

In the RL framework, the predicted features, including zone mean air temperatures and consumption costs, become integral components of the state representation. The RL agent leverages this information to make sequential decisions regarding HVAC control actions. The training process involves optimizing the agent's policies to maximize cumulative rewards, reflecting the effectiveness of learned control strategies.

The success of this integration is evident in improved total rewards and validation loss trends observed during the RL agent's training. The link between time series forecasting and RL establishes a synergy that enables the agent to adapt more effectively to dynamic environmental changes. This holistic approach, combining predictive analytics with adaptive control, contributes to enhanced HVAC system performance, ultimately resulting in increased energy efficiency and cost-effectiveness.

The integration is particularly effective across different time scales:

Short-term (6 hours): Up to a few hours, mainly for real-time temperature adjustments based on immediate needs and anticipated occupancy changes.

Medium-term (24 hours): A few hours to a day, useful for pre-conditioning the building before peak occupancy periods or adapting to expected weather changes.

Long-term (1 week): One day to a week, primarily for planning energy-efficient operation schedules and anticipating broader weather patterns.

This integrated approach is not just a theoretical concept; the results obtained from the experiments and evaluations conducted support the effectiveness of combining time series forecasting models with RL in the HVAC domain. The demonstrated improvements in total rewards and validation loss underscore the practical applicability of this synergy, providing tangible evidence of its benefits in achieving optimal HVAC system management.

CHAPTER 6

Conclusion

The time series forecasting models, multi-target prediction, and reinforcement learning for HVAC control all showcase promising results. The choice of sequence length is a critical factor, with longer sequences generally leading to improved accuracy. The inclusion of information from previous days and weeks consistently enhances the models' performance, suggesting the significance of capturing temporal dependencies and patterns.

The multi-target prediction results demonstrate the models' ability to handle diverse datasets and predict multiple variables simultaneously. This capability is valuable in real-world applications where multiple factors influence the system's behavior.

The reinforcement learning approach for HVAC control aligns with the growing interest in applying AI techniques to optimize energy consumption. The positive trends in total rewards, validation loss, and epsilon decay indicate the agent's ability to learn an effective control strategy.

In conclusion, the presented results offer valuable insights into the application of machine learning techniques for HVAC control and time series forecasting. The models demonstrate adaptability to different sequence lengths and the ability to capture complex temporal patterns. The visualizations enhance the interpretability of the results, providing a comprehensive understanding of the models' performance. The findings contribute to the broader field of building energy management and pave the way for further research and practical implementations.

6.1 Future Work

The present research lays the foundation for advancing the field of HVAC system optimization through the integration of sophisticated time series forecasting models and reinforcement learning techniques. While the current study has yielded valuable insights into the effectiveness of such an integrated approach, numerous avenues for further exploration and enhancement remain. Addressing the below mentioned future work areas is anticipated to contribute to the ongoing development and application of intelligent HVAC control systems, fostering energy efficiency, cost-effectiveness, and user comfort.

The proposed future work encompasses the following key aspects:

1. **Hybrid Models and Ensemble Learning:**
   Explore the potential benefits of combining 1D CNN and LSTM models in a hybrid architecture. Hybrid models or ensemble learning techniques could leverage the strengths of both models, potentially improving overall forecasting accuracy.

2. **Advanced Hyperparameter Tuning:**
   Conduct a more extensive hyperparameter search to identify optimal configurations for the forecasting models. Advanced tuning techniques, such as Bayesian optimization or genetic algorithms, could be employed to efficiently explore the hyperparameter space.

3. **Transfer Learning:**
   Investigate the applicability of transfer learning for time series forecasting. Pre-training models on similar datasets or tasks and fine-tuning them on the target task may enhance performance, especially in cases where labeled data is limited.

4. **Online Learning and Adaptability:**
   Implement online learning strategies to enable models to adapt in real-time to changing patterns and dynamics in the HVAC system. This could involve continuous model updates based on incoming data streams to ensure optimal performance over time.

5. **Interactions Between Subsystems:**
   Extend the scope of the HVAC control system to consider interactions between multiple subsystems. For instance, incorporate insights from lighting or occupancy sensors to improve overall building management and energy efficiency.

6. **Cost-Benefit Analysis:**
   Conduct a comprehensive cost-benefit analysis of the integrated approach. Evaluate the economic impact of implementing the proposed system in terms of energy savings, maintenance costs, and potential environmental benefits.

7. **Real-world Deployment and Validation:**
   Test and validate the integrated system in real-world settings, potentially in collaboration with building management systems or smart building initiatives. Evaluate its performance over an extended period to understand long-term benefits and challenges.

By addressing these future work areas, the integrated approach to HVAC control and energy efficiency can be further refined, leading to more robust and adaptive systems with tangible benefits in terms of cost savings, energy conservation, and user comfort.

Bibliography

[1] Yuanzheng Li et al. “A review of studies involving the effects of climate change on the energy consumption for building heating and cooling”. In: *International Journal of Environmental Research and Public Health* 18.1 (2021), p. 40.

[2] Farinaz Behrooz et al. “Review of control techniques for HVAC systems—Nonlinearity approaches based on Fuzzy cognitive maps”. In: *Energies* 11.3 (2018), p. 495.

[3] Gianluca Serale et al. “Model predictive control (MPC) for enhancing building and HVAC system energy efficiency: Problem formulation, applications and opportunities”. In: *Energies* 11.3 (2018), p. 631.

[4] Shaimaa Seyam. “Types of HVAC systems”. In: *HVAC System* (2018), pp. 49–66.

[5] Amir Rafati, Hamid Reza Shaker, and Saman Ghahghahzadeh. “Fault detection and efficiency assessment for hvac systems using non-intrusive load monitoring: A review”. In: *Energies* 15.1 (2022), p. 341.

[6] Panagiotis Michailidis et al. “Enabling optimal energy management with minimal IoT requirements: A legacy A/C case study”. In: *Energies* 14.23 (2021), p. 7910.

[7] Filip Belic, Zeljko Hocenski, and Drazen Sliskovic. “HVAC control methods-a review”. In: *2015 19th International Conference on System Theory, Control and Computing (ICSTCC)*. IEEE. 2015, pp. 679–686.

[8] Ye Yao and Divyanshu Kumar Shekhar. “State of the art review on model predictive control (MPC) in Heating Ventilation and Air-conditioning (HVAC) field”. In: *Building and Environment* 200 (2021), p. 107952.

[9] Md Washim Akram et al. “Global prospects, advance technologies and policies of energy-saving and sustainable building systems: A review”. In: *Sustainability* 14.3 (2022), p. 1316.

[10] Dongsu Kim et al. “Energy modeling and model predictive control for HVAC in buildings: a review of current research trends”. In: *Energies* 15.19 (2022), p. 7231.

[11] Iakovos T Michailidis et al. “Balancing energy efficiency with indoor comfort using smart control agents: a simulative case study”. In: *Energies* 13.23 (2020), p. 6228.

[12] S Ali et al. *A Review of DC Microgrid Energy Management Systems Dedicated to Residential Applications. Energies 2021, 14, 4308*. 2021.

[13] Christoforos Keroglou et al. “A Survey on Technical Challenges of Assistive Robotics for Elder People in Domestic Environments: The ASPiDA Concept”. In: *IEEE Transactions on Medical Robotics and Bionics* (2023).

[14] Georgios D Karatzinis et al. “Coordinating heterogeneous mobile sensing platforms for effectively monitoring a dispersed gas plume”. In: *Integrated Computer-Aided Engineering* Preprint (2022), pp. 1–19.

[15] Georgios Salavasidis et al. “Autonomous trajectory design system for mapping of unknown sea-floors using a team of AUVs”. In: *2018 European Control Conference (ECC)*. IEEE. 2018, pp. 1080–1087.

[16] Cheng Gao and Dan Wang. “Comparative study of model-based and model-free reinforcement learning control performance in HVAC systems”. In: *Journal of Building Engineering* 74 (2023), p. 106852.

[17] Dimitrios Vamvakas et al. “Review and Evaluation of Reinforcement Learning Frameworks on Smart Grid Applications”. In: *Energies* 16.14 (2023), p. 5326.

[18] Frauke Oldewurtel et al. “Energy efficient building climate control using stochastic model predictive control and weather predictions”. In: *Proceedings of the 2010 American control conference*. IEEE. 2010, pp. 5100–5105.

[19] Mehdi Maasoumy, Alessandro Pinto, and Alberto Sangiovanni-Vincentelli. “Model-based hierarchical optimal control design for HVAC systems”. In: *Dynamic systems and control conference*. Vol. 54754. 2011, pp. 271–278.

[20] Yudong Ma et al. “Model predictive control for the operation of building cooling systems”. In: *IEEE Transactions on control systems technology* 20.3 (2011), pp. 796–803.

[21] Tianshu Wei, Qi Zhu, and Mehdi Maasoumy. “Co-scheduling of HVAC control, EV charging and battery usage for building energy efficiency”. In: *2014 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE. 2014, pp. 191–196.

[22] Adrian R Katili, Rabah Boukhanouf, and Robin Wilson. “Space cooling in buildings in hot and humid climates—a review of the effect of humidity on the applicability of existing cooling techniques”. In: *14th International Conference on Sustainable Energy Technologies â€“SET*. 2015.

[23] Cheng Fan et al. “Assessment of deep recurrent neural network-based strategies for short-term building energy predictions”. In: *Applied energy* 236 (2019), pp. 700–710.

[24] Yeming Dai et al. “Improving the Bi-LSTM model with XGBoost and attention mechanism: A combined approach for short-term power load prediction”. In: *Applied Soft Computing* 130 (2022), p. 109632.

[25] Won Hee Chung, Yeong Hyeon Gu, and Seong Joon Yoo. “District heater load forecasting based on machine learning and parallel CNN-LSTM attention”. In: *Energy* 246 (2022), p. 123350.

[26] Yue Li et al. “A data-driven interval forecasting model for building energy prediction using attention-based LSTM and fuzzy information granulation”. In: *Sustainable Cities and Society* 76 (2022), p. 103481.

[27] Zahra Fazlipour, Elaheh Mashhour, and Mahmood Joorabian. “A deep model for short-term load forecasting applying a stacked autoencoder based on LSTM supported by a multi-stage attention mechanism”. In: *Applied Energy* 327 (2022), p. 120063.

[28] Anjun Zhao et al. “Prediction of functional zones cooling load for shopping mall using dual attention based LSTM: A case study”. In: *International Journal of Refrigeration* 144 (2022), pp. 211–221.

[29] Ben Jiang et al. “Attention-LSTM architecture combined with Bayesian hyperparameter optimization for indoor temperature prediction”. In: *Building and Environment* 224 (2022), p. 109536.

[30] Dandan Li et al. “Short-mid term electricity consumption prediction using non-intrusive attention-augmented deep learning model”. In: *Energy Reports* 8 (2022), pp. 10570–10581.

[31] Furkan Elmaz et al. “CNN-LSTM architecture for predictive indoor temperature modeling”. In: *Building and Environment* 206 (2021), p. 108327.

[32] Tae-Young Kim and Sung-Bae Cho. “Predicting residential energy consumption using CNN-LSTM neural networks”. In: *Energy* 182 (2019), pp. 72–81.

[33] Nikos Kampelis et al. "Thermal comfort evaluation in HVAC Demand Response control". In: *Energy Procedia* 134 (2017), pp. 675–682.

[34] Lei Yang et al. "Reinforcement learning for optimal control of low exergy buildings". In: *Applied Energy* 156 (2015), pp. 577–586.

[35] Volodymyr Mnih et al. "Human-level control through deep reinforcement learning". In: *nature* 518.7540 (2015), pp. 529–533.

[36] David Silver et al. "Mastering the game of Go with deep neural networks and tree search". In: *nature* 529.7587 (2016), pp. 484–489.

[37] William Valladares et al. "Energy optimization associated with thermal comfort and indoor air control via a deep reinforcement learning algorithm". In: *Building and Environment* 155 (2019), pp. 105–117.

[38] Gail S Brager and Richard De Dear. "Climate, comfort, & natural ventilation: a new adaptive comfort standard for ASHRAE standard 55". In: (2001).

[39] CEN Comite'Europe'en de Normalisation. "Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics". In: *EN 15251* (2007).

[40] Zhengbo Zou, Xinran Yu, and Semiha Ergan. "Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network". In: *Building and Environment* 168 (2020), p. 106535.

[41] Tao Liu et al. "A novel deep reinforcement learning based methodology for short-term HVAC system energy consumption prediction". In: *International Journal of Refrigeration* 107 (2019), pp. 39–51.