BATTERY THERMAL MANAGEMENT SYSTEM FOR ELECTRIC VEHICLES: DESIGN, OPTIMIZATION, AND CONTROL

by

Yuanzhi Liu

APPROVED BY SUPERVISORY COMMITTEE:

Jie Zhang, Chair

Babak Fahimi

Justin Koeln

Xianming Dai

To my family.

BATTERY THERMAL MANAGEMENT SYSTEM FOR ELECTRIC VEHICLES:

DESIGN, OPTIMIZATION, AND CONTROL

by

YUANZHI LIU, BS

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN

MECHANICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS

December 2021

ACKNOWLEDGMENTS

This work would not have been possible without the support, encouragement, and assistance of different people.

I would like to thank my supervising adviser, Dr. Jie Zhang, for guiding me throughout my PhD studies and funding this research. His passion and profession always provided me inspiration and motivation. I could not have imagined having a better advisor and mentor for my PhD study. He is definitely a role model in my career and life.

I would also extend my sincere appreciation to my supervising committee members, Dr. Babak Fahimi, Dr. Justin Koeln, and Dr. Xianming Dai, for their valuable and continuous guidance and suggestions in my PhD journey.

I would like to thank my colleagues in the Design and Optimization of Energy Systems (DOES) lab, including Dr. Mao Li, Dr. Mingjian Cui, Dr. Xiaobang Wang, Dr. Binghui Li, Dr. Cong Feng, Dr. Mucun Sun, Zhenke Wang, Li He, Xin Li, Samragni Dutta Roy, Jubeyer Rahman, and Roshni Anna Jacob, for their assistance and friendship. Moreover, I would like to thank Dr. Dani Fadda for his encouragement and support.

I would be remiss not to thank faculty and staff in the Mechanical Engineering department, the Erik Jonsson School of Engineering and Computer Science, and the university for their support and help in the last five years. At last, I dedicate my dissertation to my parents and family members for their unconditional love and support.

December 2021

BATTERY THERMAL MANAGEMENT SYSTEM FOR ELECTRIC VEHICLES:

DESIGN, OPTIMIZATION, AND CONTROL

Yuanzhi Liu, PhD
The University of Texas at Dallas, 2021

Supervising Professor: Jie Zhang, Chair

We are witnessing a fast-growing demand in vehicle electrification nowadays due to the widespread environmental consciousness, stringent emission regulations, and carbon neutrality implementation. As one of the most promising energy storage and electrification solutions, lithium-ion battery has been widely employed for electric vehicles (EVs) due to its excellent properties like high energy density, low maintenance, and long cycle life. However, there still exist multiple critical challenges in using lithium-ion battery at large scale as the major power source, such as reliability issues, safety concerns, and especially the range anxiety. Several promising solutions have been explored in the EV industry to mitigate the drawback of range anxiety, such as larger capacity with high energy density and ultra-fast charging. All these approaches challenge the temperature sensitive battery system as a side effect by bringing in extra overburdened waste heat. Given these concerns, battery thermal management system (BTMS) plays an indispensable role in maintaining the maximum temperature and temperature uniformity for EVs.

This dissertation proposes a novel J-type air-based cooling structure via re-designing conventional U- and Z- type structures. Aiming to further improve the thermal performance, a surrogate-based optimization framework with two-stage cluster-based resampling is developed for BTMS structural optimization. Compared with the U- and Z- type, the novel

J-type structure is proved with significant advancements. Based on the optimized J-type configuration, an operation mode switching module is designed to mitigate the temperature unbalance by controlling the opening degree of two outlet valves. Tested by an integrated driving cycle, results reveal that the J-type structure with its appropriate control strategy is a promising solution for light-duty EVs using an air cooling technology.

Improving the energy efficiency is another potential approach to mitigate range anxiety. In this dissertation, a model predictive control (MPC)-based energy management strategy is developed to simultaneously control the BTMS, the air conditioning system, and the regenerative power. A vehicle velocity forecasting framework is integrated with the MPC-based energy management to further improve the energy efficiency. Deep learning and image-based traffic light detection techniques have been leveraged for velocity forecasting. Results show that the proposed energy management method has significantly improved the overall EV energy efficiency.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

We are witnessing a fast-growing demand in vehicle electrification nowadays due to the widespread environmental consciousness, stringent emission regulations, and carbon neutrality implementation. During the past decade, the global passenger electric vehicle (EV) market penetration has increased from zero to nearly 4.5% with a number of more than 10 million by 2020. As one of the most promising energy storage and electrification solutions, lithium-ion battery has been widely employed for EVs due to its excellent properties like high energy density, low maintenance, and long cycle life. However, compared with internal combustion engines, there still exist multiple critical challenges in using lithium-ion battery at large scale as the major power source, such as reliability issues caused by temperature sensitivity and gradual aging, safety concerns like catching fire or penitential explosion, and especially the range anxiety resulted from low energy density and small capacity.

It can be concluded that the competitions between EVs and conventional internal combustion engine vehicle actually lie on the endurance range. Multiple promising solutions have been explored in the EV industry aiming to enhance the driving range and mitigate the drawback of range anxiety, such as increasing the energy density of battery cell by testing varying materials, building larger battery packages, employing ultra-fast charging technologies. All these approaches tend to challenge the battery system as an inevitable side effect by generating extra overburdened waste heat. Note that the appropriate operating temperature for lithium-ion battery ranges between 20°C and 45 °C. An extremely high temperature caused by heat accumulation will jeopardize the performance, health status, and safety of the battery pack. Given these concerns, battery thermal cooling/heating structural design, optimization, and control have emerged as one of the essential research fields in EV applications and innovations.

The maximum operating temperature and the temperature uniformity are two of the most important evaluation metrics for battery thermal management system (BTMS). A large amount of research has been performed during the past decade to examine and explore a wide range of heat transfer mediums integrated with its appropriate structure design and optimization for BTMS. State-of-the-art heat transfer mediums include air, fluid, phase change material, heat pipe, and an integration of them. All the mediums other than air and fluid are still under laboratory experimental stages due to their complexities and unstability. For air-based cooling technologies, optimizing the existing structure and further redesigning new structures are two mainstream research directions in both academic and industry. A large number of cooling structures have been developed for different types of battery under varying operation scenarios. Some of these air-based structures have also been successfully applied in EV industry. The air-based cooling technology was primarily utilized in hybrid EVs to satisfy the thermal constraints for a small-size battery pack, such as the early models of Nissan Leaf and Toyota Prius. Following the aforementioned advancements in the EV industry with larger battery packs, the fluid-based cooling technology has been proved with a higher thermal capacity to address the newly-rising challenges like fast charging. Although the fluid-based thermal control approach is the most popular technology in practice, the air-based cooling technology is still worth to be explored especially associated with its structure optimization and optimal control strategies, due to its excellent performance for compact-size or cost-competitive EVs.

Aiming to address the challenge of range anxiety, besides these improvements from the view of battery power source like increasing battery capacity and fast charging, another promising alternative solution is to improve the energy efficiency of varying on-board devices at device/system level. As discussed above, these technical advancements focus more on the hardware perspective, while the approach of improving the energy efficiency emphasizes more on the control and optimization algorithm for EVs. Furthermore, improving the power allocation will also reduce the heat generation and thermal burden for the BTMS as a return. Considering the energy regeneration from the braking system, the energy usage can be potentially enhanced mainly from two aspects: (i) to optimize the discharging sequence by scheduling the operations of different devices and subsystems to avoid overlapped high power output based on real-time driving conditions; (ii) to fully utilize the regenerative energy from the braking system rather than recharging back to the battery system. Apart from the driving motor and its assisted subsystems that mainly depend on actual traffic conditions, other primary systems, including the air conditioning system and BTMS, could be operated with a more flexible energy-efficient schedule that incorporates with the fixed driven power output jointly through load shifting.

To optimize the operation schedule, model predictive control is a promising approach by considering the power demands of the driven motor system several steps ahead. The sequential actions of energy allocation are obtained via solving the predictive optimization problem to guarantee optimal energy usage, while retaining the system constraints from electric, thermal, and fluid dynamic perspectives. As the forecasted inputs for predictive energy management algorithm, vehicle velocity forecasting is leveraged to estimate the power demand of the driven motor and its assisting systems. For repeated commuting routes, the piece-wise segmentation is a promising approach, in which the whole driving cycle is divided into segments according to the locations of intersections and stops. Based on the cycle segments, a forecasting pool that consists of a collection of base forecasting models is established to yield primary predictions. These basic forecasting results can be further improved via off-line and on-line model selection and ensemble approach. Besides the predictive algorithm, another feasible approach to improve the energy efficiency is to perform real-time energy distribution integrating with the detected traffic light signal and velocity acceleration signal. The traffic light detection-based real-time energy distribution is potentially an alternative to the predictive algorithm for urban routes.

The goal of this dissertation is to develop promising cooling structures and corresponding control algorithms for BTMS to mitigate the range anxiety barriers by leveraging surrogate-based optimization and data driven-based techniques. With the optimized BTMS cooling structure and enhanced control implementations, the dissertation also aims to optimize the energy efficiency for the whole EV via predictive control-based energy management with real-time velocity forecasting and real-time energy allocation with traffic light detection.

The remainder of this dissertation is organized as follows. Chapter 2 presents the literature review on different related topics, including battery cooling structure design and optimization, thermal control and energy management, individual velocity forecasting and its impacts on energy management. Chapter 3 presents a comparative study and surrogate-based optimization for a J-type air-based cooling structure. Chapter 4 develops a thermal control strategy for J-type air-based BTMS. A model predictive control-based algorithm is developed and implemented for energy management in Chapter 5, followed by individual vehicle velocity forecasting in Chapter 6. The impacts of velocity forecasting on vehicle energy management are investigated in Chapter 7. Conclusions and future work are discussed in Chapter 8.

CHAPTER 2

LITERATURE REVIEW[1234]

There are a variety of heat transfer mediums that have been employed to mitigate the thermal impacts from battery charging and discharging, e.g., air, fluid, phase change material, heat pipe, and a hybrid combination of the mediums. Compared with other thermal dissipation approaches, the air-based cooling technology has always been one of the research priorities due to its distinctive nature. A significant amount of research on air-based BTMS has been performed in the literature, including the BTMS design, optimization, and control methods. Moreover, the energy management strategy also plays an important role in improving the overall energy efficiency, especially integrated with vehicle velocity forecasting. In this chapter, the technology advancements in these aspects will be comprehensively reviewed.

2.1 Battery Cooling Structure Design and Optimization

Air-based cooling technologies have been widely applied in EV industrial applications. Most of the compact size EVs only employ passive air cooling as a trade-off between vehicle weight and cruising capacity, so does the plug-in hybrid electric vehicle (PHEV) powered by lithium iron phosphate (LFP) battery with more stable thermal characteristics. For example, Nissan Leaf and BYD Song have successfully updated for several generations, and the passive cooling

---

[1] Y. Liu and J. Zhang (2019), Design A J-type Air-based Battery Thermal Management System through Surrogate-based Optimization, *Applied Energy*, Vol.252, pp.113426. Reprinted with permission from Elsevier.

[2] Y. Liu and J. Zhang (2019), Self-adapting J-type Air-based Battery Thermal Management System via Model Predictive Control, *Applied Energy*, Vol.263, pp.114640. Reprinted with permission from Elsevier.

[3] ©2021 ASME. Reprinted, with permission, from Y. Liu and J. Zhang (2021), Electric Vehicle Battery Thermal and Cabin Climate Management Based on Model Predictive Control, *Journal of Mechanical Design*, Vol.143, Issue 3, 2021, pp.031705.

[4] ©2021 ASME. Reprinted, with permission, from Y. Liu and J. Zhang (2021), A Repeated Commuting Driving Cycle Dataset with Application to Short-term Vehicle Velocity Forecasting, *Journal of Autonomous Vehicles and Systems*, 1-44

systems have been proved to be reliable. PHEV with lithium nickel manganese cobalt oxide (NMC) battery like Toyota Prius normally employs a fan-driven active air cooling system, since NMC battery is more sensitive to temperature than the LFP battery in daily operations [96].

2.1.1 Air-based Cooling Structure Design

These industrial applications also reveal the inadequacy of air-based BTMS like non-uniform and limited heat dissipation capability, contamination from external cooling air, and potential noise or vibration. The majority of the existing literature on active pure air-based BTMS focuses on structure design improvement and flow optimization, attempting to uniformize the internal temperature profiles of the battery pack while retaining the maximum temperature simultaneously. Based on the conventional structures like U- and Z-type, several structural modifications of pure air-based BTMS have been developed in the literature. For example, a simulation-aided comparative study of convectional U-type and Z-type ventilation structures was conducted by Park [66], in which irregular flow shapes like tapered manifolds were proved to be effective. Based on the U-type configuration, a reciprocating structure with two flip door was developed by Mahamud et al. [55], where a 72% temperature uniformity improvement was yielded via changing the settings of the flow direction. A counterflow arrangement developed by Xun et al. [94] was validated by computational fluid dynamics simulations, suggesting that changing the flow direction periodically could improve the thermal performance. The effects of driven-fan location were investigated by Wang et al. [90] using both numerical simulation and experiments, which revealed that the top fan location coupled with cubic arrangement outperformed other settings. A series-parallel mixed cooling structure with aligned bank, staggered bank, and trapezoid configuration developed by Yang et al. [95] was assessed, showing better thermal characteristics than the benchmark case. Moreover, auxiliary devices or components that attempt to balance the flow rate were also testified with simulations or experiments, which implies that it is an effective and competitive approach to meet the thermal expectations regarding maximum temperature and temperature uniformity, such as the parallel ventilation structure with multiple vortex generators and jet inlet developed by Shahid et al. [77], the straight-forward cooling configuration with mist generator developed by Saw et al. [75], and other specific structures with assisting components that seek to allocate the flow rate uniformly [32]. There hasn't been any study available to comprehensively compare the pros and cons of all the proposed structures.

Pure air-based BTMS may encounter thermal challenges under intense or instantaneous working conditions like extreme fast charging and rapid acceleration, due to its limited thermal conductivity [45]. Based on the specific heat differences between air and other mediums, several studies have attempted to develop a hybrid cooling system by adding an extra intermediate medium as a thermal buffer in between the battery cells and the cooling air. For instance, a mixed phase changing material (PCM) system can be integrated with passive air cooling, in which additional material can be added to the main PCM (paraffin for most cases) to enhance the thermal conductivity, such as metal foams [69], graphene, expanded graphite [43], and graphite of nanocomposite structure [37]. It is also reported that a PCM system coupled with forced air cooling tends to yield more desired performances than pure PCM or pure forced convective air cooling, especially for cases after a long time operation [44]. Moreover, a hybrid heat pipe cooling system coupled with forced air ventilation is able to effectively control the temperature field within an acceptable difference at the battery pack level under abusive discharging conditions [12]. It is worth noting that though the mixed cooling technologies have desirable potentials in mitigating the thermal concerns for EVs, none of these studies have been reported to be successfully integrated into industrial EV applications due to their complexity and instability.

2.1.2 Air-based Cooling Structure Optimization

Based on these existing prototypes of pure or mixed air-based cooling structures, structural configuration optimization is another feasible alternative to enhance thermal performances. Several thermal or fluid dynamic characteristics like temperature limitation, uniformity, and flow efficiency are usually treated as the optimization objectives based on different purposes. The channel and manifold configuration optimization is one of the most straight-forward but effective approach for any U- or Z-based parallel cooling structures. For example, a channel size reconfiguration performed by Chen et al. [14] observed a 45% reduction of the maximum temperature difference, in which the Newton method and a network-based flow model were employed for the optimization. A similar channel size optimization for a U-type structure was also conducted by Li et al. [42], where a parametric investigation regarding the channel size was firstly performed before adaptive iterative optimization. Besides the traditional iterative optimization approach, data-driven surrogate-based optimization integrated with stochastic solvers has also been reported. For example, Wang et al. [91, 92] performed a multidisciplinary optimization via genetic algorithm (GA) to modify a U-type BTMS by taking into account both the thermal performance and the battery lifetime. The Pareto frontier as an output indicated that the final configuration was yield as a trade-off between the thermal performance and the cycle life during the decision-making process.

Given these concerns, this dissertation seeks to (i) design a novel air-based cooling structure (J-type) by integrating the basic U- and Z- structures, and (ii) optimize the setting of the J-type structure using a two-stage adaptive cluster-based resampling method and surrogate-based optimization.

2.2 Thermal Control and Energy Management

2.2.1 Vehicle Thermal Control Strategy

There exist a large number of published studies focusing only on the optimization of the geometry size or configuration for a single balanced steady state without any modification and control strategies. The thermal performance has rarely been tested or verified under dynamic operating conditions. As a result, when the battery condition changes, e.g., under a dynamic driving condition, a fast charging schedule, or seasonal and regional temperature variations, the thermal control system may possibly fail to work as effectively as its original optimal design or even lead to critical issues like severe temperature imbalance and potential safety concerns. Moreover, the aforementioned complex structures, as well as the side effects that brought upon themselves, e.g., noise and vibration, are very challenging to be generalized and applied in practice.

Developing an appropriate control strategy for the existing air-based BTMS structure is another effective alternative at a higher level to achieve the goal of maintaining the operating temperature and uniformity. It is also noticed in recent years that the research foci tend to shift from the conventional steady-state structure design and modification to a dynamic integrated control and optimization. However, since it generally involves with multidisciplinary analysis, the emerging trend has a very limited number of studies available in integrating appropriate control strategy with the current air-based structures at the system level. For instance, Gao et al. [21] developed a fuzzy logic control unit for a straight-forward air-based cooling system, whose results showed that the maximum temperature could be controlled within expectations. He et al. [87] exploited the air-based reciprocating cooling with a hysteresis control method to achieve optimal cooling effectiveness, which resulted in a desirable 84% reduction in terms of the parasitic power consumption. Vatanparvar et al. [85] developed a thermal and energy management methodology that optimized the utilization of battery and ultra-capacitor, and the control results showed significant improvements regarding both the thermal performance and energy consumption. It can be summarized that the control strategy for air-based BTMS can be implemented via straight-forward algorithms like proportional–integral–derivative (PID) and fuzzy control, in which the airflow rate is usually regarded as one of the major state variables, though the system inputs may differ. Through a frequency or voltage modulation control for the cooling fan or compressor, the airflow rate is able to be adaptively adjusted according to dynamic operating conditions. However, owing to the intrinsic limitations of conventional U-type, Z-type, or other through-type structures, controlling the airflow rate has a significant impact on temperature rise, but makes little difference regarding the temperature uniformity, especially under large flow rate conditions.

In addition to the conventional control approaches, model predictive control (MPC) and its family algorithms (e.g., constrained linear MPC, and nonlinear MPC) have been employed in several studies because of the unique characteristics, i.e., MPC forecasts steps ahead to determine optimal control solutions. It is worth noting that the MPC for thermal management usually involves with the power output and energy consumption of the battery system. For instance, Masoudi et al. [59] extended the study of parallel cooling using the model predictive control algorithm, in which a thermal improvement compared to Toyota Prius baseline performance was observed. Tao et al. [83] used MPC to regulate the refrigerant compressor and cooling air flow rate to keep an ideal cooling temperature for a hybrid cooling system, and found that both the temperature uniformity and energy efficiency could be improved significantly. Amini et al. [5] extended the existing work and presented a hierarchical two-layer MPC scheme to schedule optimal thermal trajectories for the cabin and battery cooling in hybrid electric vehicles. The aforementioned studies have revealed the effectiveness of MPC and highlighted further potential applications in battery systems. However, the majority of studies implemented the thermal control directly on pre-designed air-cooling structures without further optimization, and there has not any control co-design approach reported in the literature to design a battery cooling system. Additionally, lumped battery electric-thermal-fluid models were usually adopted for simplification, in which the battery temperature was investigated only at the cell level. The temperature distribution at the battery pack level has not yet been evaluated and validated with simulations or experiments. Overall, by implementing system control on a fixed cooling structure, the battery thermal performance has been significantly improved for dynamic cases but probably far from optimum.

2.2.2 Vehicle Energy Management Strategy

It is also noticed that the battery charging/discharging with an instant high power output may potentially shorten the cycle-life and affect the discharge power density [46]. From the perspective of thermal control, a high power density tends to generate a large amount of heat and bring about inevitable thermal impacts on BTMS. In consequence, it is reasonable to optimize the discharging sequence of EVs by scheduling the operations of different devices to avoid overlapped high power output based on real-time driving conditions. Moreover, judging from the energy perspective, investigations have also revealed that the energy used for thermal control itself ranks third among all the consumptions for EVs, after the driven power and the air conditioning system. Optimizing the discharging schedules of all the power consumption sources is an essential and promising approach to enhance energy efficiency. Apart from the driving motor and its assisted subsystems that mainly depend on actual traffic conditions, other primary systems including the cabin climate control system /air conditioning system (AC), as well as BTMS, can be operated with a flexible energy-efficient schedule that incorporates with the fixed power outputs jointly through load shifting. It is worth noting that there are similar energy management studies for plug-in hybrid electric vehicles, which focuses more on apportioning the driven power between the internal combustion engine and electric motor instead of the internal consumed energy distribution [39]. For EVs, this is a fast-emerging research field but with very few publications available [59, 83, 5].

Considering regeneration effects, discharge scheduling and optimization refer to two approaches: (i) optimize the operation sequences of different devices and subsystems to avoid a peak demand; (ii) utilize the regenerated power directly instead of recharging the battery system. There exist similar studies in the literature regarding discharging scheduling optimization for EVs [22]. However, the majority of previous studies were performed and evaluated merely in terms of energy-saving using the aforementioned schedule optimization approach [4]. Moreover, previous studies did not consider the side effects of the recharged energy, as well as its corresponding thermal impact on the whole battery pack under dynamic driving conditions [97].

To bridge these gaps, there are two directions that could be further improved: (i) develop an intelligent thermal control algorithm for the novel BTMS structure (J-type); (ii) develop an energy management strategy via MPC and taking into account the internal power distributions.

2.3 Vehicle Velocity Forecasting

Short-term traffic forecasting has been extensively investigated in the past decade as a potential feasible solution to mitigate the growing concern of traffic congestion, especially with the advent of connected and automated vehicles (CAV), big data, artificial intelligence, and internet of things [84, 86]. It has been shown that the traffic can be significantly improved by integrating the existing road network with smart traffic light control and intelligent transportation systems, based on the real-time traffic measuring and forecasting, such as the traffic flow volume, traffic density, average traffic velocity, and travel time [36].

2.3.1 Network Traffic Forecasting

The majority of these studies usually focus on a broader scope of traffic forecasting, from point-level, street-level, to network-level, serving to provide further insights for transportation management and policy making. At the very-beginning stage, several statistic and machine learning-based algorithms were exploited for varying traffic conditions. For instance, as a statistical analysis model, the auto-regressive integrated moving average (ARIMA) method has been broadly adopted for point-level or street-level traffic forecasting [23]. ARIMA and its families were usually adopted as a benchmark to be compared with other advanced approaches, such as machine learning-based methods (e.g., support vector regression (SVR) [89], back propagation artificial neural network (BPANN) [100], and radial basis function neural network (RBFNN)), probability-based methods (e.g., Kalman filter and hidden Markov chain (HMM) [27]), and deep learning-assisted methods (e.g., long short-term memory network (LSTM) [107, 40]). Among these reported methods, the probability-based Kalman filter and HMM methods produced stochastic forecasts, while others produced deterministic forecasts. Historical driving records are usually considered as the training dataset for forecasting. Though only the temporal relationship is taken into account, the aforementioned algorithms are capable of predicting the traffic situations in most cases with reasonable accuracy. It is also worth noting that there is no published literature that indicates the superiority of any of these algorithms, due to diverse traffic data sources.

However, some of these forecasting algorithms may not perform well for a large area with complicated transportation networks and uncertain environmental factors. To alleviate this arising challenge, advanced deep neural networks with topological feature embedding have been employed to characterize the spatial-temporal characteristics in forecasting. Massive efforts have been performed on convolutional neural network (CNN)-based algorithms [54], which generally falls into two categories: convolution-based LSTM that integrates CNN and LSTM [93, 102], and temporal graph convolutional networks that combine graph neural network with gated recurrent units/networks [106, 53, 60]. By comparing with the reported forecasting performance, these deep learning-based methods have shown overwhelming superiority over the aforementioned classic parametric or simple-structured machine learning-based approaches. This newly-emerging trend is very likely to continue in traffic forecasting [18].

2.3.2 Individual Vehicle Velocity Forecasting

It should be noted that for the aforementioned studies, all the data has been collected from a single or a series of fixed observation locations using sensors like inductive-loop detector, wireless magnetometer, microwave radar, and video image processor. These studies emphasize more on the networked vehicles rather than an individual passenger vehicle. However, velocity forecasting for individual vehicles has drawn significant attention in the past decade, especially along with the fast-growing demand for vehicle electrification. Velocity forecasting plays a critical role in improving the energy efficiency for electric or hybrid vehicles. Generally, velocity forecasting serves as the system input for a model predictive control-assisted or reinforced learning-based energy management system to optimize the charging/discharging schedule, the regenerative power harvest, and the operation of an on-board air-conditioning system, especially for repeated fixed routes inside or between cities [50, 5, 3, 108]. Additionally, forecasted velocity is also regarded as an indispensable prerequisite to generate varying scenarios and networked/individual vehicle ecosystem for multidisciplinary control Co-design [92, 65].

There are generally three major discrepancies that the individual vehicle velocity forecasting differs from the network traffic forecasting. First, individual vehicle velocity forecasting utilizes the floating velocity trajectory as the data source instead of the network traffic records. Second, individual vehicle velocity forecasting requires a significantly shorter prediction horizon at seconds, compared to the network traffic forecasting at minutes or hours timescales. Third, the networked traffic forecasting can be facilitated by local/cloud-based powerful computing tools with sophisticated deep learning-based structures; while the velocity forecasting for individual vehicles tends to directly utilize on-board computing devices, leaving no alternative but to implement computationally efficient forecasting algorithms.

The traffic forecasting algorithms discussed above are still applicable and practicable to individual vehicle velocity forecasting, which can generally be categorized into stochastic and deterministic approaches. For instance, as one of the most popularly used stochastic methods, HMM was modified by Jing et al. [29] with a fuzzy logistic model to predict individual vehicle speed 8 seconds ahead, and Zhou et al. [109] developed a self-learning multi-step Markov chain model based on simulated data. However, stochastic methods are usually eclipsed by deterministic approaches regarding prediction accuracy. For example, Sun et al. [82] revealed that RBFNN and ANN performed significantly better than HMM. Liu et al. [47] reported similar results that both LSTM and ARIMA outperformed HMM in 10 seconds ahead speed forecasting based on a real urban driving dataset. Moreover, among all the deterministic methods, it seems that LSTM possesses prevailing advantages for a same driving dataset. For instance, a comparative study between LSTM and other deep learning-based algorithms like CNN and CNN-LSTM conducted by Rabinowitz et al. [70] revealed that LSTM dominates other deep learning-based and machine learning-based forecasting technologies with a considerable higher accuracy. The feasibility of embedding LSTM on board has also been tested and verified by Gaikwad et al. [20] with an on-board processor. Besides, other deep learning networks such as deep belief network and stacked auto-encoder have also been investigated for speed forecasting for a highway speed dataset [41].

Acting as an indispensable element, individual vehicle velocity forecasting has been deeply integrated with the connected and automated system by leveraging vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) technologies. It is worth noting that the forecasting performance can be further improved by considering the surrounding traffic situations via vehicle-to-everything communications. For example, Moser et al. [61] and Zhang et al. [103] proved that individual vehicle velocity forecasting could yield more accurate predictions by knowing the states of traffic lights in advance. Upon perceiving the traffic conditions of the local area, velocity forecasting will become a dynamic process by choosing an optimal eco-routing [72]. There is also a growing trend to predict the vehicle behaviors at intersections including acceleration and deceleration, aiming to achieve optimal velocity trajectory for energy control [13].

The electric vehicle industry has witnessed huge advancements in embedding with advanced driver assistance systems (ADAS) and predictive optimal energy management strategies (POEMSs). However, one of the challenges is that the policy-making and construction of intelligent transportation system are falling far behind the electrification and intellectualization of passenger vehicles. It is still economically and technologically prohibitive to enable all passenger vehicles being connected to the intelligent transportation system and receive real-time traffic information. Based on the existing urban transportation infrastructure and vehicle installations, how to improve the velocity forecasting accuracy is still a stringent and challenging problem for vehicle energy management.

To further improve the performance of individual vehicle velocity forecasting, this dissertation will: (i) generate a repeated driving cycle dataset and develop a hybrid velocity forecasting algorithm using on-board GPS devices, and (ii) utilize image-based object detection techniques to identify the traffic light status to further improve the velocity forecasting accuracy.

CHAPTER 3

COMPARATIVE STUDY AND SURROGATE-BASED OPTIMIZATION FOR A J-TYPE AIR-BASED COOLING STRUCTURE[1]

Air-based battery thermal management system (BTMS) has been widely employed in the EV industry due to its remarkable advantages like lightweight, simple structure, and low cost, especially for plug-in hybrid EVs where the battery system works only in low-duty operation scenarios. The existing air cooling systems attempt to modify the configurations of the channel and manifold to uniformize air flow rate with a lower pressure drop so as to enhance the temperature uniformity and energy efficiency simultaneously. However, it is challenging to achieve these goals with the conventional U- or Z-type BTMS with a fixed structure under changing working conditions (e.g., discharging, charging, extreme fast charging, etc.), especially for large battery packs. To address this challenge, this chapter proposes a flexible air cooling structure, named J-type, emerging from the existing Z- and U-type. The proposed J-type BTMS has two outlets with control valves, which adds more cooling flexibility to the BTMS under varying battery working conditions. By controlling the opening degree of the two valves simultaneously, the J-type BTMS can be adaptively controlled in real time to modify the flow field to provide an optimal cooling strategy to the whole battery pack.

A comparative parametric study among U-, Z-, and J-type structures is performed to further explore the sensitivities and effects of key system parameters, e.g., channel size, manifold configuration, charging/discharging rate, temperature, and cooling air flow rate. Based on the parametric analyses, a suite of key design parameters and constraints are determined to perform optimization. The grouped-channel optimizations of the three structures are performed using surrogate-based optimization. The pros and cons of the novel J-type structure

---

[1] Y. Liu and J. Zhang (2019), Design A J-type Air-based Battery Thermal Management System through Surrogate-based Optimization, *Applied Energy*, Vol.252, pp.113426. Reprinted with permission from Elsevier.

are elaborated by comparing with the optimal U- and Z-type structures. A further J-type optimization regarding the manifold configuration is also conducted to show that the optimal settings of a BTMS vary with battery working conditions, and the J-type BTMS is capable of switching BTMS modes with controlling valves and air flow rate in real time to satisfy the cooling requirement. This study serves to develop a basic design concept of the J-type structure and to establish a pioneering foundation for further BTMS control or co-design framework.

3.1 Battery Electro-thermal Model

It is essential to establish a battery electro-thermal model before conducting the thermal control studies. A comprehensive electro-thermal model helps to better understand the heat generation mechanisms inside the battery, which serves to develop appropriate redesign and optimization algorithms. Extensive studies have revealed that the volumetric heat generation rate of LIB is strongly influenced by the charging/discharging current, operating temperature, state of charge (SoC) [2], and cycles.

3.1.1 Battery Thermal Model

A number of electrochemical/electrothermal models have been proposed in the literature to interpret the thermal mechanism. First proposed by Bernardi, then improved by Rao and Newman [73], a simplified LIB thermal model is expressed as:

$$\dot{Q} = I(V - V_{oc}) + IT\frac{\partial V_{oc}}{\partial T} \tag{3.1}$$

where $\dot{Q}$ represents the battery heat generation rate, V and V_{oc} denote the cell voltage and open circuit voltage, respectively. T is the battery cell operating average temperature, $\partial V_{oc}/\partial T$ is named as the entropic heat coefficient, and I is the battery current, which is defined as positive for charging and negative for discharging. The thermal model can be established by measuring the entropic heat coefficient. The first term is irreversible, mainly from the internal resistance ohmic losses, while the second term is reversible, known as the entropic heat from chemical reactions.

For simplification, the internal cell condition is assumed to be homogeneous, and the heat source derived above is assumed to be distributed uniformly inside the cell. Radiation heat transfer is neglected here since the temperature difference is limited. With these assumptions, the battery thermal behavior can be estimated using a lumped thermal model, as shown in Eq. 3.2.

$$mC_p \frac{\partial T}{\partial t} = \dot{Q} - hA(T_{cell} - T_{\infty}) \tag{3.2}$$

where m denotes the mass of battery cell, C_p is the average heat capacity, h represents the convective heat transfer coefficient, A is the effective surface area, and T_{∞} is the free stream temperature of the cooling media. This approach requires a precise measurement of the dynamic heat by using either the method of accelerating rate calorimeter (ARC) or isothermal heat conduction calorimeter (IHC).

3.1.2 Battery Equivalent Circuit Model

Compared to the electrochemical model, an equivalent circuit model is more straightforward to characterize the relationship between battery electrical characteristics and its thermal behaviors. Figure 3.1 shows the first order equivalent circuit, which consists of an ideal voltage source, an internal ohmic resistance, and a parallel RC circuit. The RC circuit is utilized here to interpret the dynamic response. All the parameters are contingent on the SoC, operating temperature, and battery cycle. The mathematical expression of the equivalent circuit is derived as follows:

$$v_t = V_{oc} - V_D - V_o + V_D e^{-\frac{t-t_0}{\tau}} \tag{3.3}$$

where $V_D = I \cdot R_D$ is the potential drop on the RC circuit, $V_o = I \cdot R_o$ is the potential drop on the internal resistance, and $\tau = R_D \cdot C_D$ denotes the time constant.

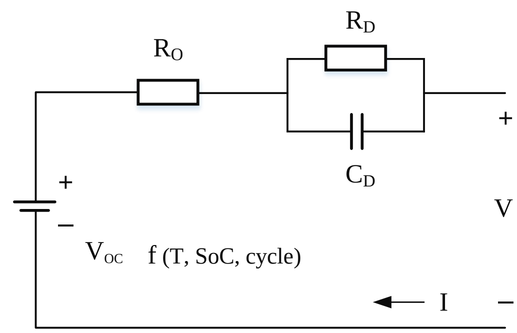


Figure 3.1: First order equivalent circuit for LIB

These parameters in Eq. 3.3 can be measured, extracted, and calculated with the method of hybrid pulse power characterization (HPPC) test [98, 25]. In this study, Graphite/$LiMn_2O_4$ pouch battery cells with a capacity of 1.6 Ah, and a nominal voltage of 3.75 V are used for experiments. A total of 200 groups of the characteristic data (R_o, R_D, and C_D) are extracted from various experimental settings. The heat generated by the resistances R_o and R_D is considered to be equal to the battery internal heat source. More details about the experiment setup, parameters analyses, and results validation can be found in Ref. [99]. The experimental data is then utilized to establish the battery electro-thermal model by using a surrogate model, which consists of three input variables: the current, SoC, and operating temperature.

3.1.3 Surrogate-based Electro-thermal Model

After k-fold cross-validation, a Kriging approximation with second order polynomial regression and exponential error estimation is utilized to create a surrogate model based on the experimental data. A deterministic response $\mathcal{G}(I, SoC, T)$ with three dimensional variables is formulated with the Kriging surrogate model, given as:

$$\mathcal{G}(I, SoC, T) = \mathcal{F}(\lambda, I, SoC, T) + \mathcal{R}(\omega, I, SoC, T) \tag{3.4}$$

where $\mathcal{F}$ is defined as the regression model with a second-order polynomial kernel, and $\mathcal{R}$ is the approximation error, given as

$$\mathcal{F}(\lambda, I, SoC, T) = f(I, SoC, T)\lambda \tag{3.5}$$

$$f(I, SoC, T) = [1, I_N, SoC_N, T_N, I_N^2, I_N SoC_N, I_N T_N, SoC_N^2, SoC_N T_N, T_N^2] \tag{3.6}$$

$$\mathcal{R}(\omega, I, SoC, T) = r(I, SoC, T)\omega \tag{3.7}$$

$$r(\gamma, I, SoC, T) = e^{(-\gamma_I|I-I_q|-\gamma_{SoC}|SoC-SoC_q|-\gamma_T|T-T_q|)} \tag{3.8}$$

where $f(I, SoC, T)$ denotes a vector of the normalized variables with orders 0, 1 and 2, in which the normalized variable is defined as $\alpha_N = (\alpha - \mu_\alpha)/\sigma_\alpha$, μ and σ are the mean and standard deviation, respectively. I_q, SoC_q, and T_q are training data. All the Kriging parameters λ, ω, and γ are calculated by the generalized least squares estimation method [52, 16].

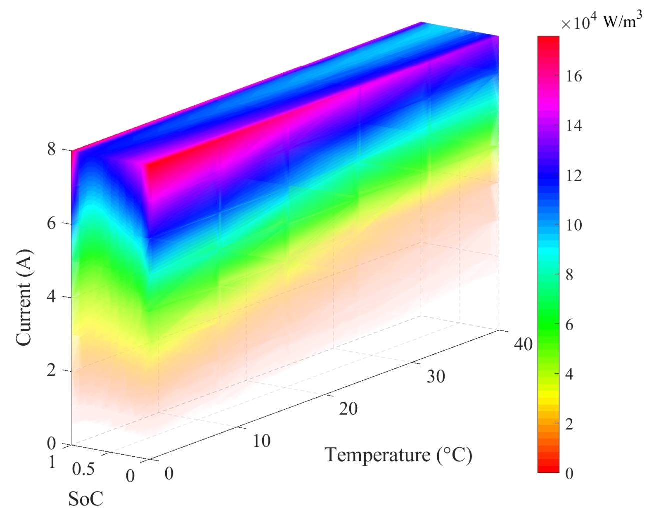


Figure 3.2: The distribution of battery equivalent volumetric heat generation rate

Figure 3.2 shows the equivalent volumetric heat generation rate distribution with respect to the operation current, SoC, and temperature. The electro-thermal model only covers the feasible operation ranges of current and temperature for LIB, and 10 A is suggested as the critical safety current. It indicates that decreasing the operating temperature tends to increase the internal resistance and thus induce a huge augment of heat generation rates. The calorific value is comparatively low around 60% SoC. Additionally, the overall battery thermal performance is more sensitive to the operating current than other parameters. This electro-thermal model can be applied to both charging and discharging conditions since the reversible heat is relatively smaller than the irreversible heat, especially under high current conditions.

3.2 Numerical Comparative Study of Air-based Cooling Structure

Compared with the hybrid air-based BTMS, pure air cooling has incomparable advantages in terms of stability, maintenance, and vehicle power-to-weight ratio. Existing air cooling systems attempt to modify the configurations of the channel and manifold to uniformize flow rate with a lower pressure drop so as to enhance the temperature uniformity and energy efficiency simultaneously. However, It is challenging to achieve these goals with the conventional U- or Z-type BTMS with a fixed structure under changing working conditions (e.g., discharging, charging, extreme fast charging, etc.), especially for large battery packs. To address this challenge, this task develops a flexible air cooling structure, named J-type, emerging from the existing Z- and U-type. The proposed J-type BTMS has two outlets with control valves, which adds more cooling flexibility to the BTMS under varying battery working conditions. By controlling the opening degree of the two valves simultaneously, the J-type BTMS can be adaptively controlled in real-time to modify the flow field to provide an optimal cooling strategy to the whole battery pack. In this task, we aim to provide promising insights for the further optimization.

3.2.1 Concept Design of a J-type Cooling Structure

By taking the advantages of both U-type and Z-type air-based BTMS, this paper proposes a novel BTMS structure named J-type. A conceptual design of the J-type BTMS is illustrated in Fig. 3.3. The J-type BTMS prototype consists of ten battery cells with geometry sizes of 151 mm in height, 65 mm in length, and 16 mm in width based on Ref. [66].

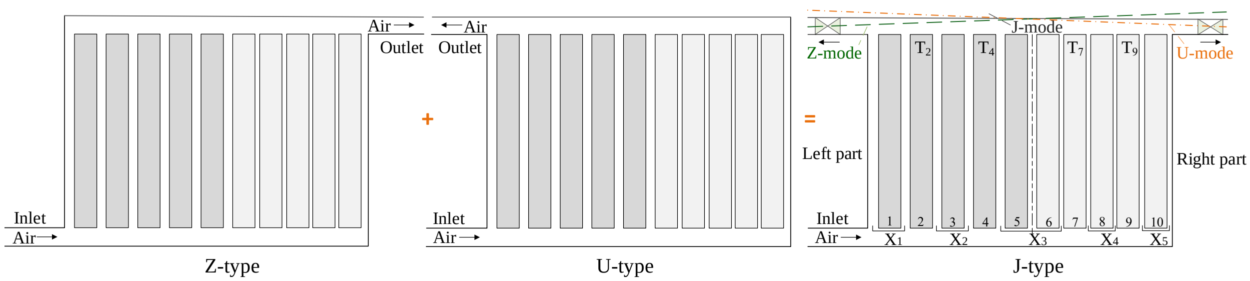


Figure 3.3: The conceptual design of a J-type air-based BTMS

As illustrated in Fig. 3.3, the size of channel and inlet manifold are first optimized under multiple working conditions, while the air flow rate and the openness of the two outlet valves will be adaptively controlled in real-time. By changing the air flow field via the openness of the two outlet valves, BTMS is able to cool down the hot area, thereby ensuring the temperature uniformity within a narrow range and improving the overall thermal performance. Moreover, the concept design is readily extended and applied to the battery pack level, since the concept module is arranged in a standard shape and can be extended from different directions.

The J-type BTMS can easily switch to U-type by completely closing the Valve Z and opening the Valve U, in which the front battery cells maintain a lower temperature than the rear ones. Similarly, the J-type BTMS can also switch to Z-type, by closing the Valve U and opening the Valve Z, which leads to better thermal performance for the rear battery cells. In between the two extreme conditions, the system can optimally adjust the opening degree of each valve based on the module temperature and battery working conditions.

To evaluate the effectiveness of the proposed J-type BTMS, a comparative study is performed among J-, Z-, and U-type BTMS. Three-dimensional (3D) computational fluid dynamics(CFD) simulation models of the three types are built up in ANSYS Fluent with the k-epsilon (k-ϵ) turbulence model. The total size of the models converges to around 1,700,000 elements after grid dependence analysis, as shown in Fig. 3.4. The mass flow rate inlet and pressure outlet are selected as the inlet and outlet settings, respectively. The radiative heat transfer is neglected here due to the very limited temperature difference. The battery cell in the CFD model is assumed to have a uniform heat source and isdefined by a user-defined function. A simulation takes approximately 25 minutes to converge on a six-core workstation.

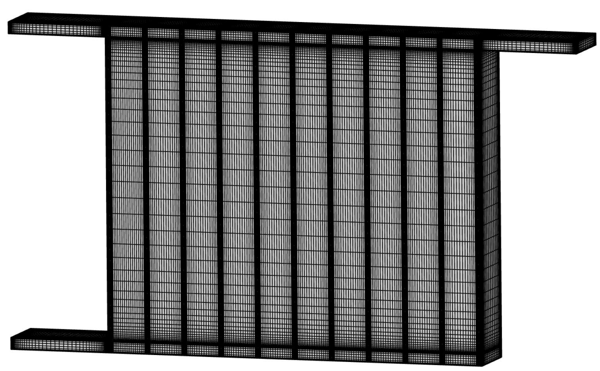


Figure 3.4: The CFD meshing of J-type air-based BTMS

A benchmark model with an even 3 mm channel size, an equal 6 mm inlet/outlet manifold size, and an initial inlet mass flow rate 7.1 g/s is set up for comparisons among the three types of BTMS. The initial temperature, discharging current, and SoC are set as 295 K, 3 C-rate, and 1, respectively, which correspond to a heat generation rate of 33,800 W/m^3. The J-type outlet structure could be simplified into a tapered manifold for simulation convenience, as indicated by the red dashed line in Fig. 3.3.

3.2.2 Comparative Study

A collection of parametric studies among U-, Z-, and J-type structures are performed based on CFD simulations, including the structural shape, controlled variables, and premodified structure.

Effects of Structure Size

The BTMS structure uses three main parameters to define geometry size, i.e., channel interspacing size, collecting manifold size, and distribution manifold size. The numerical and experimental parametric studies of the U-type BTMS with respect to the channel size, equivalent heat generation rate, and mass flow rate have been well conducted in the authors' previous studies [49, 42]. In this study, comparative parametric analyses among the three BTMS structures are performed to study and explore the sensitivity of BTMS to major design variables under certain conditions, so as to provide fundamental understandings for further BTMS optimization.

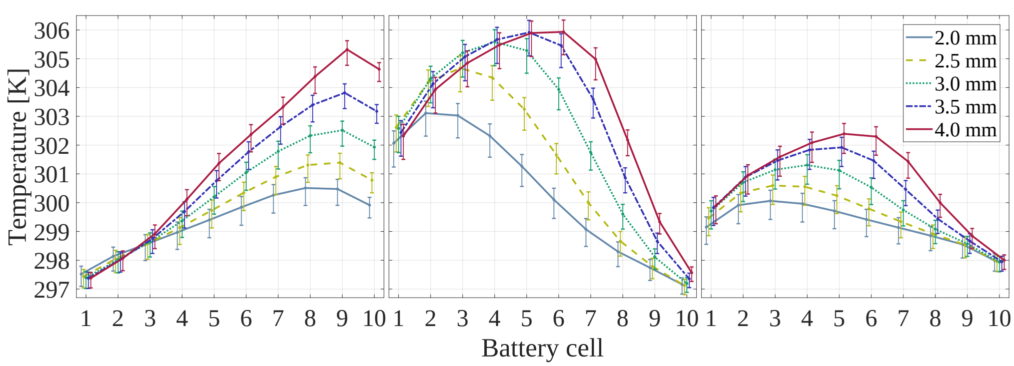


Figure 3.5: Parametric analyses of channel size (U-Z-J-type, respectively)

Figure 3.5 shows the battery temperature distribution of U-type, Z-type, and J-type with different channel sizes. The curves represent the face-weighted average temperature of the battery pack. The maximum and minimum temperature of the same battery cell are represented by the upper bar and lower bar that are connected with vertical line segments, respectively. It is seen that the temperature distributions of U-type and Z-type are similar to each other, except that the distributions are opposite in terms of the maximum temperature location, in which the highest temperature of U-type occurs at the rear side, while the highest temperature of Z-type is located at the front side. The J-type's thermal performance is always better than the U- and Z-type structures, since it comprises an extra ventilation outlet. It also suggests that a narrow channel tends to uniformize the temperature distribution and lower the temperature rise simultaneously. However, as a side effect, a narrow channel may also lead to the augment of pressure drop and result in a higher pumping energy consumption, as shown in Fig. 3.6a. This is one of the most critical design considerations regarding maintenance and energy efficiency. Note that the parametric investigation mainly emphasizes on the ultimate thermal performance of the battery pack, though the mass flow rate and pressure drop vary from channel to channel.

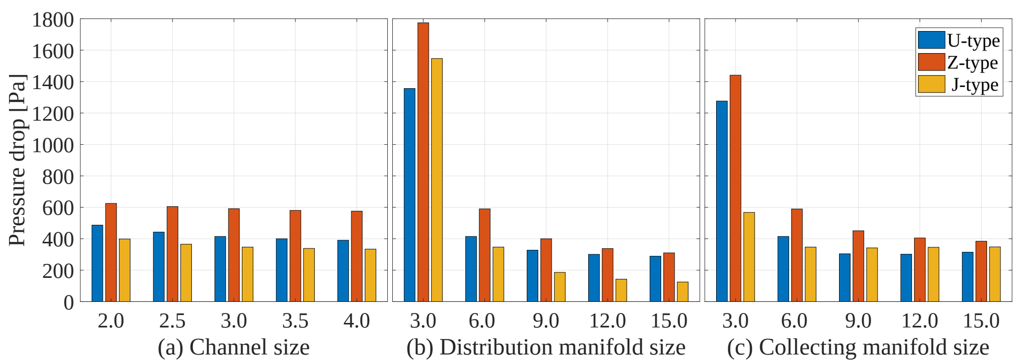


Figure 3.6: The pressure drop of different structures with respect to different key parameters

Figure 3.7 shows the relationship between the distribution manifold size and temperature rise. It is impressive that the 3 mm distribution manifold reduces and uniformizes the temperature distribution in both U- and Z-type. However, as shown in Fig. 3.6b, the 3 mm manifold almost triples in pressure drop compared to the benchmark cases, making it unsustainable in energy consumption. For J-type as shown in Fig. 3.7, increasing the size of the distribution manifold improves the thermal performance, and little difference is observed when the size is larger than 9 mm. Figure 3.6b also indicates that the distribution manifold size for all three BTMS structures should be larger than 6 mm due to the significant augment of pressure drop.

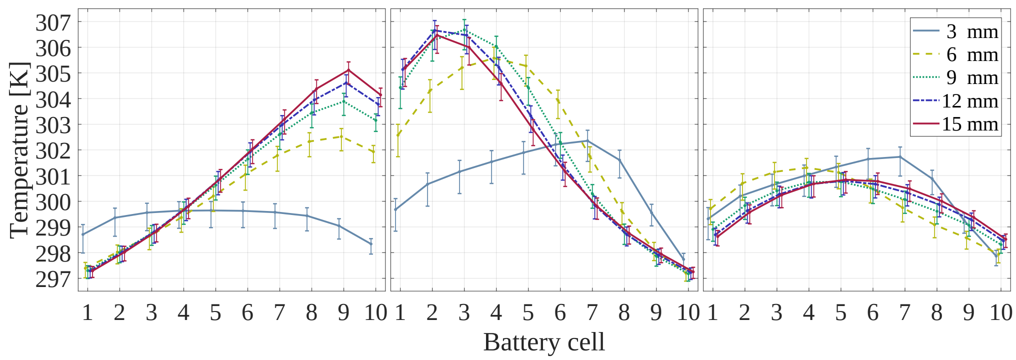


Figure 3.7: Parametric analyses of distribution manifold spacing size (U-Z-J-type, respectively)

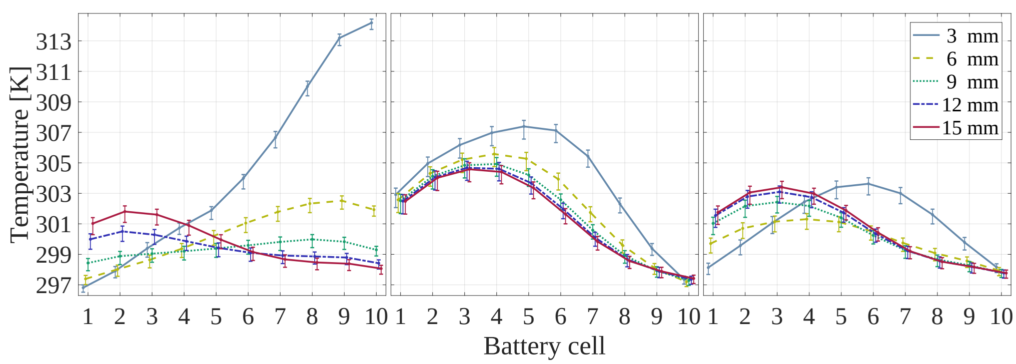


Figure 3.8: Parametric analyses of collecting manifold spacing size(U-Z-J-type, respectively)

In contrast to the distribution manifold, the collecting manifold performs differently, in particular with a small manifold size, as shown in Fig. 3.8. Under the benchmark flow rate condition, a 3 mm collecting manifold is too small for all the three structures, which results in unexpected high temperature (Fig. 3.8) and pressure drop (Fig. 3.6c). For the U-type BTMS, the 9 mm collecting manifold case performs the best regarding temperature rise and uniformity. For Z- and J-type, little discrepancy is observed when the collecting manifold is larger than 12 mm, which implies that enlarging the manifold size does not help to lower the temperature rise. Therefore, by analyzing Figs. 3.6−3.8, it is found that an appropriate manifold size ranges between 6 mm and 12 mm, which should also satisfy the constraints imposed by the entire battery pack volume and energy density.

Effects of Controlled Variables

Besides the structural parameters, controlled variables such as operating temperature, heat generation rate, and air mass flow rate may also have strong influences on BTMS thermal performance. The Graphite/$LiMn_2O_4$ battery is vulnerable to temperature. There is a closed-loop coupled relationship between the operating temperature and the battery heat generation rate, in which different operating temperature leads to varying heat generation rates, and thus impacts the battery pack temperature. For simplification, the SoC and battery current are fixed as 1 and 3 C-rate in simulations, respectively. Figure 3.9 shows the effects of operating temperature on BTMS performance for U-, Z-, and J-type. It is seen that the temperature rise is significantly less in high-temperature environment than that of low-temperature environment, so does the temperature difference of the battery pack, since the internal resistance tends to decline in high-temperature environments.

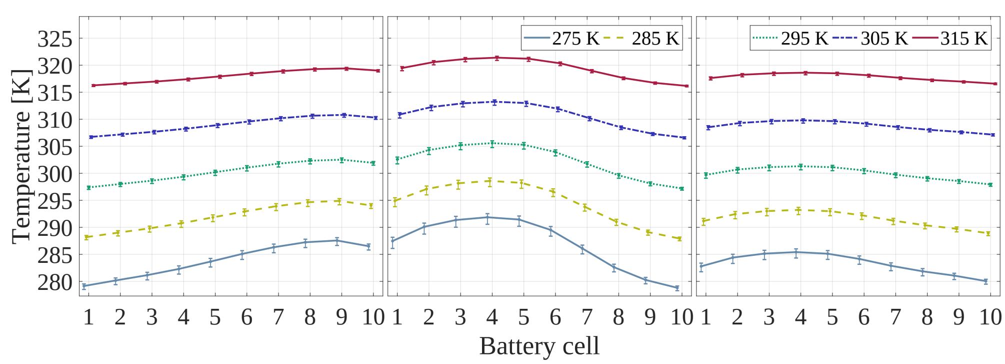


Figure 3.9: Parametric analyses of operating temperature (U-Z-J-type, respectively)

As discussed in the previous section, there is a quadratic relationship between the battery equivalent heat generation rate and the charging/discharging current. To study the effects of operation current, the SoC and operating temperature are set as 1 and 295 K, respectively. The current investigation range is constrained between 1 C-rate and 3 C-rate, which equal to 3,700 and 33,800 W/m^3 for heat generation rate, respectively. As seen from Fig. 3.10, the temperature rise presents a quasi-quadratic relationship with the C-rate current input. Both U- and Z-type deteriorate as the heat generation rate increases. The J-type has significant advantages over the other two types with high heat generation rates.

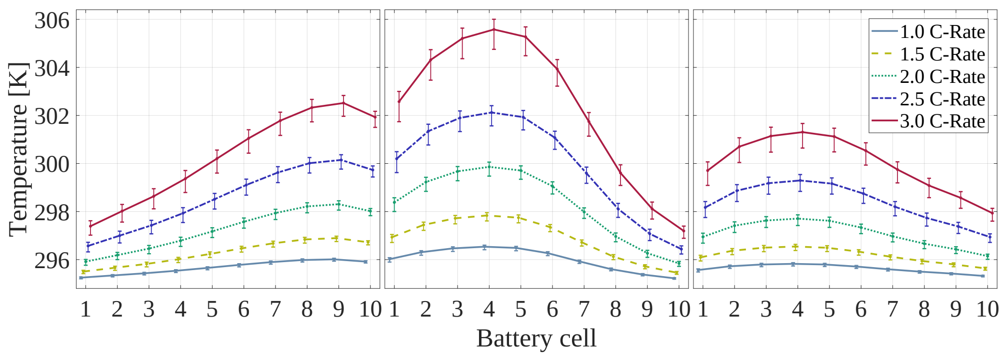


Figure 3.10: Parametric analyses of charging/discharging current (U-Z-J-type, respectively)

The impacts of air mass flow rate on the performances are presented in Fig. 3.11. It shows that the temperature rise is approximately linear to the mass flow rate for both the U- and J-type BTMS. The only exception occurs at the rear side of Z-type: the temperature dose not change significantly as the mass flow rate increases, due to that most of the cooling air flows through the rear side channels.

Effects of Modified Structures

Extensive studies have been conducted in the literature, seeking to uniformly distribute the flow rate [9, 88]. Specifically, the modified special structures with better fluid characteristics for BTMS can be summarized into three types: uneven channels, tapered distribution manifold, and tapered collecting manifold. Figure 3.12 shows the temperature distribution of priori grouped-channel cases, in which the benchmark cases (i.e., 3 mm even channels) are

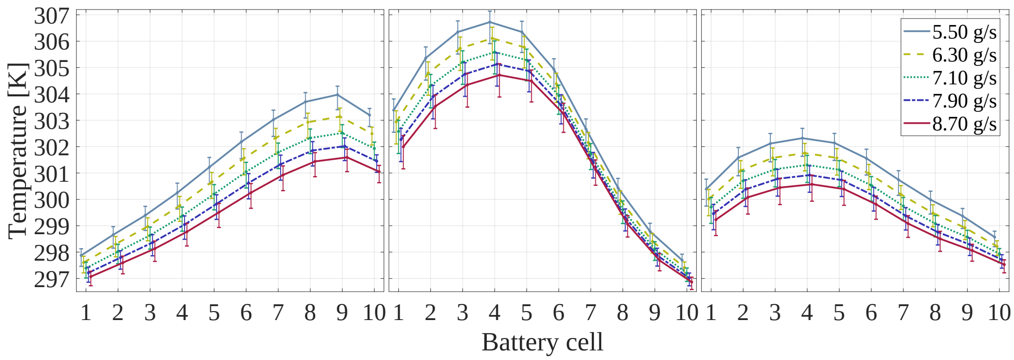


Figure 3.11: Parametric analyses of mass flow rate (U-Z-J-type, respectively)

presented for comparisons. The channels are divided into three groups with combinations of every 4, 3, and 4 channels in different interspacing sizes. Note that the numbers in the legend (e.g., U 2-3-4) represents the sizes of grouped channels in mm, and U 3-3-3 denotes the benchmark case with even channels for the U-type BTMS. It is seen that all grouped-channel cases perform significantly better than the benchmark cases, where the temperature rise is reduced and the uniformity is also improved. The pressure drop is slightly increased for grouped-channel cases as shown in Fig. 3.13a.

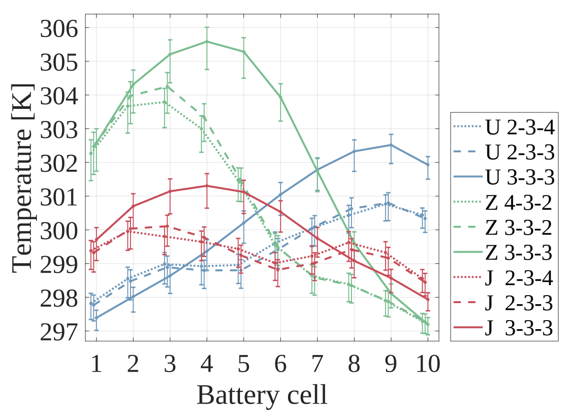


Figure 3.12: Parametric analyses of grouped-channel sizes

Figures 3.14 and 3.15 show the effects of tapered manifold modifications on the BTMS performance. The geometry of a tapered manifold is defined by the heights of two ends, as shown in the legend (e.g., 6-3 mm). The larger number represents the inlet size of the

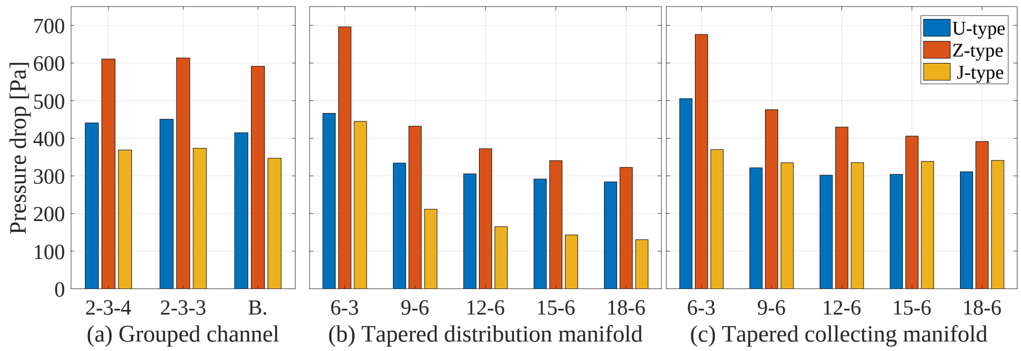


Figure 3.13: The pressure drop of different modified structures (B.:Benchmark cases)

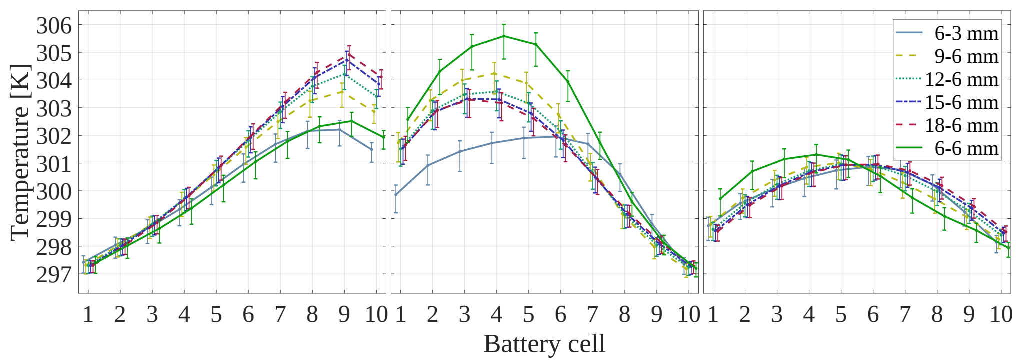


Figure 3.14: Parametric analyses of tapered distribution manifold size (U-Z-J-type, respectively)

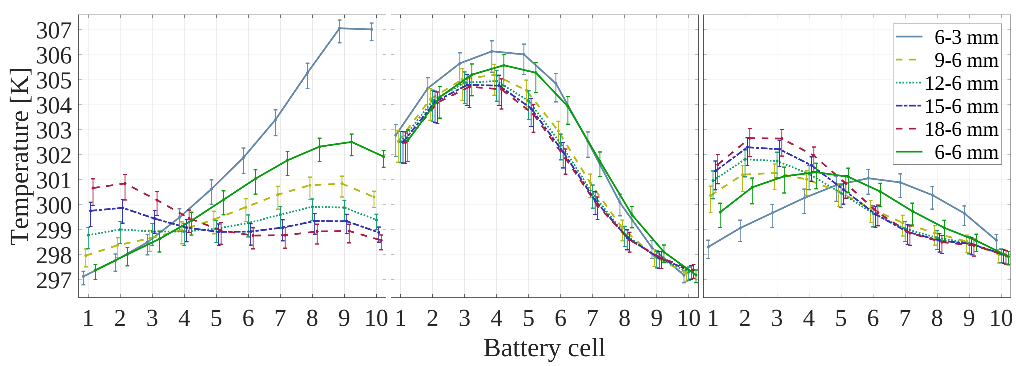


Figure 3.15: Parametric analyses of tapered collecting manifold size (U-Z-J-type, respectively)

distribution manifold or the outlet size of the collecting manifold. Note that it is infeasible to have a smaller inlet for the distribution manifold or a smaller outlet for the collecting manifold in BTMS structure design, since it may generate back flows or eddies with extra vibration and noise inside the manifold. It is observed that the tapered distribution manifold (in Fig. 3.14) does not enhance the BTMS performance for U-type, while the tapered collecting manifold (in Fig. 3.15) has improved the uniformity and lowered the temperature rise for most cases. For the J-type BTMS with tapered manifolds, the tapered distribution manifold has improved the uniformity, while the tapered collecting manifold does not show a clear improvement. For Z-type, both the distribution and collecting tapered manifolds have improved the uniformity and lowered the temperature rise. The pressure drop is strongly affected by the manifold size, as illustrated in Fig. 3.13, and a structure with larger flow channels generally leads to less pressure drop. The overall impacts of modified structures are summarized in Table 3.1.

Table 3.1: Summary of the modified structures

Improvement (Y/N)	U-type	Z-type	J-type
Grouped Channels	Y	Y	Y
Tapered Distribution Manifold	N	Y	Y
Tapered Collecting Manifold	Y	Y	N

Based on the analyses performed above in this section, it is recognized that the geometry of the flow field, the channel size and the manifold configuration in particular, has a significant impact on the thermal performance and flow efficiency. The manifold configuration is highly related to the pressure drop, and the appropriate size ranges between 6 mm and 12 mm. Not all of the priori optimized structures can surely enhance the performance. For example, the tapered distribution manifold does not improve the U-type performance, and so does the tapered collecting manifold for the J-type structure.

3.3 Surrogate-based Optimization of Air-based BTMS

Based on the parametric analyses above, it is recognized that the BTMS structure has remarkable influences on system performance. BTMS is a complex system, especially when considering modified structures and various working conditions, e.g., normal battery arrangement with time-dependent electric characteristics, uneven channels with tapered manifolds in steady stage, and J-type structure under real-time control conditions. To address the computational challenges in BTMS structure optimization with expensive CFD simulations, a surrogate-based optimization method is proposed in this study. The whole optimization is divided into two stages: Stage 1 optimizes the grouped-channel size for all three types (U, Z, and J), and Stage 2 further optimizes the distribution and collecting manifolds under different working conditions for the J-type BTMS.

3.3.1 Optimization Methodology

In this study, three groups of surrogate models are constructed based on a limited number of high fidelity CFD simulations of U-, Z-, and J-type BTMS. To set up CFD simulations, the 11 battery channels (as shown in Fig. 3.3) are divided into 4 groups with a layout of 3-3-2-3 channels from the front side to the rear side. The 4 grouped-channel sizes are considered as the design variables, ranging from 2 mm to 5 mm. In the process of design of experiments (DoE), the optimal Latin Hypercube method is employed to generate these experiment points and settings [80]. The paralleled manifolds and other controlled variables remain the same as the benchmark case. A total of 500 simulations are conducted for each type, 75% of which are utilized as the training data, and the rest are used for validation and testing.

The surrogate-based optimization framework is illustrated in Fig. 3.16. Based on the CFD simulation results, a large pool of surrogate models is first constructed, which consists of 5 major groups with 62 surrogate models regarding different kernel functions or hyper parameters, e.g., Artificial Neural Network (ANN), Kriging/Gaussian Process Regression (GPR), Support Vector Machine (SVM), Radial Basis Functions (RBF), and Polynomial Response Surface (PRS). During the model training and selecting processes, a weighted evaluating criterion of two metrics is adopted here to evaluate the accuracy using K-fold cross-validation, which include the normalized maximum absolute error (NMAE) and normalized root mean square error (NRMSE), as given by:

$$NMAE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{\hat{y}_k - y_k}{y_{max} - y_{min}}\right| \tag{3.9}$$

$$NRMSE = \frac{1}{y_{max} - y_{min}}\sqrt{\frac{\sum_{k=1}^{n}\left(\hat{y}_k - y_k\right)^2}{n}} \tag{3.10}$$

where $\hat{y}$, y, y_{max} and y_{min} denote the corresponding estimated value, actual value, maximum value, and minimum value, respectively. n is the number of test data used in evaluating the performance. The top-ranked models are selected for optimization.

Other than these base surrogate models, a hybrid ensemble model are established with multiple basic models after model evaluation, given as:

$$Y_{hybrid} = \sum_{i=1}^{n_s} w_i \hat{y}_i \tag{3.11}$$

where w_i and $\hat{y}_i$ are the weight factor and estimated value of ith surrogate, respectively. n_s denotes the number of ensemble surrogate members. Note that the weight factors are solved in the validation process. Then a local impact factor β, defined as the ratio between the actual value and the hybrid model estimated value, is added based on the methodology from the Extensive Adaptive Hybrid Function (E-AHF) in Ref. [104, 81] to further tune surrogate estimates, as given in Eq. 3.12. This tunning step is only suggested for low-dimensional cases, since the uncertainty brought by high-dimensional features may worsen the hybrid surrogate estimation.

$$Y_{final} = \beta Y_{hybrid} \tag{3.12}$$

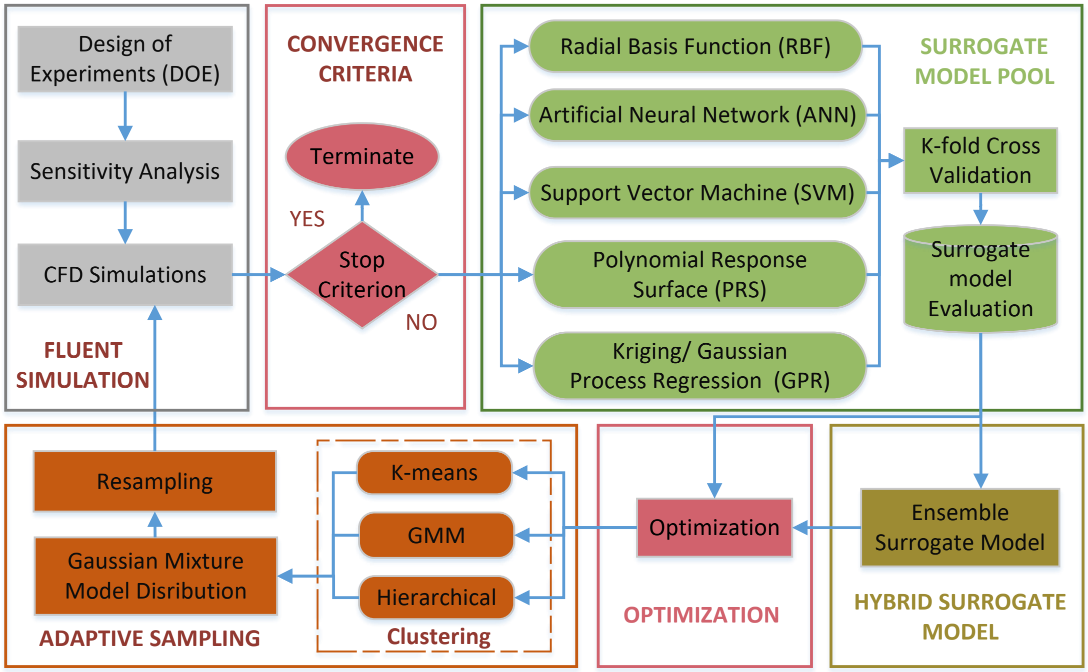


Figure 3.16: The overall framework of adaptive surrogated-based optimization for BTMS

Based on the selected base and hybrid models, the genetic algorithm (GA) is adopted here to solve the black-box optimization. It is worth noting that some of the optimization solutions are very likely to lie on a very small region since they follow the same algorithm but with hyperparameter variants. Given this concerns, during the adaptive resampling step, a two-stage sampling method will be employed to generate more experiment points, in which the first stage is to cluster the optimum candidate solutions, and the second stage is to generate adaptive samples based on the clustered groups using a Gaussian mixture model (GMM). The sampling probability is proportional to the size of clusters. Combining with the obtained optimal solutions, these resampling points are then validated via CFD simulations. By comparing with the previous resampling results, the convergence and stop criterion is defined as:

$$Bias = \left| \frac{Y_k^* - Y_{k-1}^*}{Y_{k-1}^*} \right| \leq 0.001 \tag{3.13}$$

where Y_k^* and Y_{k-1}^* are the best optimization result after the k-th resampling and the (k-1)-th resampling, respectively, and 0.001 is a predefined convergence tolerance. If the convergence criteria is met, the whole process terminates, and the best result among all the resampling trials will be regarded as the global optimum.

3.3.2 Stage I: Channel Size Optimization

In the first stage, the optimization objective is to minimize the maximum temperature $T_{max} = f(x_1, x_2, x_3, x_4)$, and the range of the grouped-channel size is considered as constraints. The optimization problem is formulated by:

$$\begin{aligned} \arg\min_{x} \quad & T_{max} = f(x_1, x_2, x_3, x_4) \\ \text{subject to} \quad & 2.0 \leq x_i \leq 5.0 \quad i = (1, 2, 3, 4) \end{aligned} \tag{3.14}$$

where x_1-x_4 represent the grouped-channel size from the front to the rear side. Table 3.2 summarizes a subset of surrogate models with high accuracy from k-fold cross-validation. The surrogate models for the three types of BTMS have a similar level of accuracy, since they are all established based on fluid dynamics. However, the model accuracy decreases slightly from U-, Z- to J-type as the complexities of physical models increase. The surrogate estimation and error of the ensemble model and selected member models are shown in Fig. 3.21. Note that the ensemble model does not necessarily perform the best at every local estimation due to the high nonlinearity of the problem. However, the ensemble model captures the overall trend of the problem and provides the best global accuracy.

The whole process terminates after two rounds of resampling, as shown in Fig. **??**, and the resampled results converge to a small design range. The optimal results of the first and second rounds are 304.989 K and 304.954 K, respectively, and the normalized bias decreases to 1.4E-10^4. The best sample among all the resampling data is treated as the optimal solution, as shown in Fig. 3.18. Compared with the case of even channels, it is seen that

Table 3.2: Evaluation results of a subset of surrogate models

	U-type		Z-type		J-type	
Model-Kernel	NMAE	NRMSE	NMAE	NRMSE	NMAE	NRMSE
RBF-TPS	**2.47**	4.39	4.92	7.99	4.37	5.94
GPR-Matern32	2.49	4.09	3.49	5.60	4.11	5.57
RBF-Cubic	2.50	4.14	4.32	6.64	4.14	5.40
GPR-Ardmatern32	2.58	4.29	**3.31**	**4.98**	**3.97**	**5.23**
SVR-Polynomial	2.71	**4.03**	4.54	6.56	5.31	7.59
PRS-Cubic	2.87	4.09	4.64	7.11	4.27	5.59
KRG-Poly2gauss	3.01	4.62	4.61	6.67	4.73	6.41
SVR-RBF	3.02	4.26	4.54	7.92	5.38	7.12
ANN-RBFN	3.57	6.03	3.35	6.50	5.04	7.24

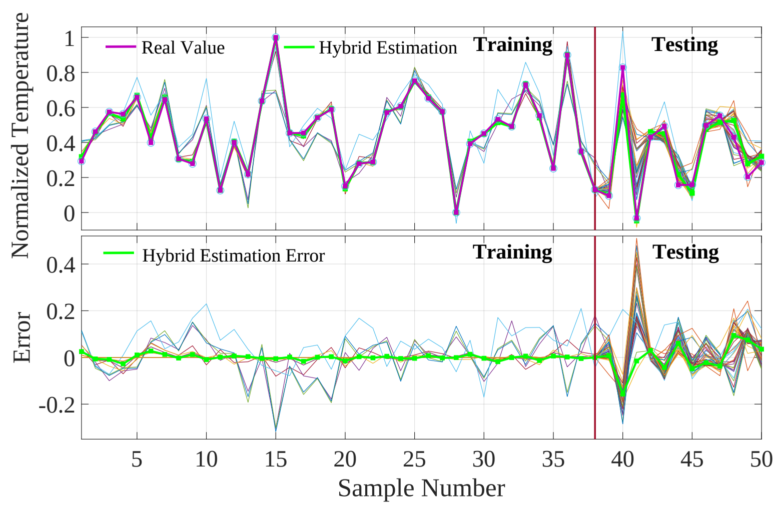


Figure 3.17: Surrogate model estimation and error from both ensemble and individual member models for J-type (only a portion of the scattered points are presented)

the maximum temperature has decreased from 307.18 K to 304.95 K, and the uniformity (represented by standard deviation) has significantly improved from 1.46 K to 0.42 K.

Our previous studies have also considered the total volume, temperature standard deviation, and pressure drop to perform a multiobjective optimization [91, 48]. However, according to the parametric studies above, it is found that both the volume and pressure

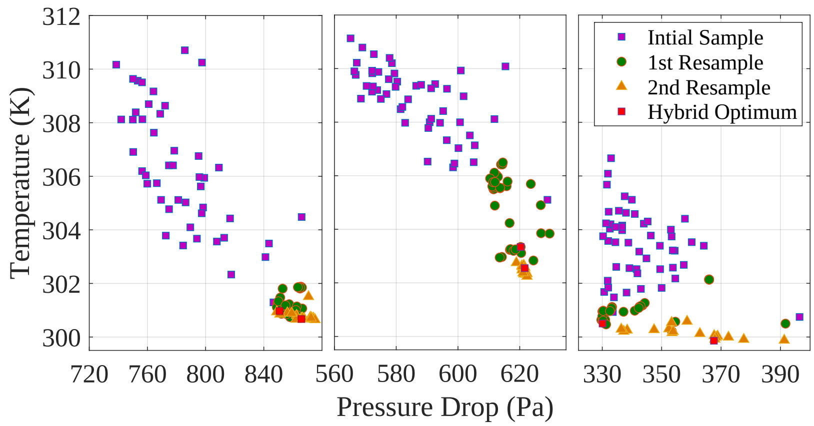


Figure 3.18: The optimization and resampling results (U-Z-J-type, from left to right)

drop are highly related to the geometry size of the flow field. Once the mass flow rate and other controlled variables are fixed, the temperature uniformity is indirectly reflected by the maximum temperature. Though the parameters of pressure drop and temperature uniformity are not directly modeled in the objective function, they are still regarded as reference metrics to evaluate the BTMS performance.

The optimal thermal management system of all three types are summarized in Table 3.3. Compared with the benchmark case, the optimized U-type has a 35.3% reduction in temperature rise, and a 63.4% improvement in temperature uniformity with a cost of 7.5% augment in pressure drop. Similarly for Z-type, the temperature rise and temperature standard deviation decrease by 46.6% and 69.1%, respectively, while the pressure drop increases by 5.0%. For J-type, the optimal arrangement has reduced the temperature by 31.18% and improved the uniformity by 67.8%, but increased the pressure drop by 12.7%.

Figure 3.19 shows the CFD simulations of the three optimal designs. By comparing U- and Z-type structures, it is seen that the optimal maximum temperature of U-type is 21% lower than that of Z-type, but the pressure drop of U-type is 28% higher. Overall, U-type is more competitive in temperature-sensitive cases, while Z-type is more competitive in energy efficiency-sensitive cases. J-type has shown advantages in terms of both temperature and

Table 3.3: The optimal designs of U-, Z-, and J-type BTMS

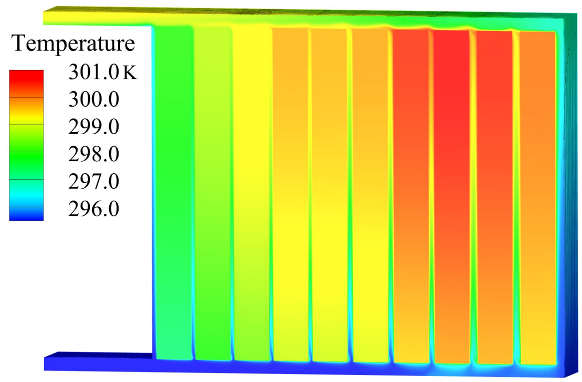
	Design Variable (mm)				Evaluation Criteria					
Type	x_1	x_2	x_3	x_4	$T_{max}(K)$	ΔT	$T_\sigma(K)$	ΔT_σ	$\Delta P(Pa)$	$\Delta\Delta P$
U-Benchmark	3	3	3	3	303.75	-	2.30	-	801.48	-
U-Optimum	2.01	2.14	2.81	3.86	300.66	-35.3%	0.84	-63.4%	862.20	7.50%
Z-Benchmark	3	3	3	3	308.41	-	4.11	-	592.70	-
Z-Optimum	4.96	2.51	2.14	2.01	302.16	-46.6%	1.27	-69.1%	622.34	5.0%
J-Benchmark	3	3	3	3	302.15	-	1.43	-	347.25	-
J-Optimum	2.74	2.66	2.13	2.02	299.98	-31.18%	0.46	-67.8%	391.28	12.7 %

pressure drop. Due to the space occupied by the extra outlet, J-type is more suitable in volume-insensitive applications like hybrid EVs.

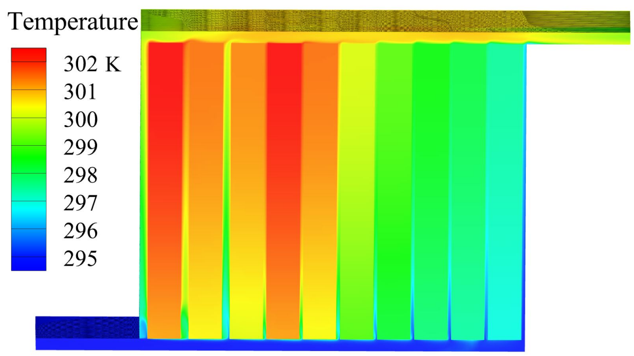
3.3.3 Stage II: J-type Structure Manifold Optimization

For the J-type BTMS as illustrated in Fig. 3.3, the system can optimally adjust the opening degree of the two valves based on the module temperature and the air mass flow rate. The openness degree of the two valves corresponds to the manifold size under a specific battery working condition. The goal of performing manifold optimization here is to validate the hypothesis that controlling the manifold size (via valves control) could further improve the thermal performance under varying working conditions. In this study, based on the optimal grouped-channel BTMS obtained from the previous optimization, a further step is performed to optimize the manifold under two working conditions.

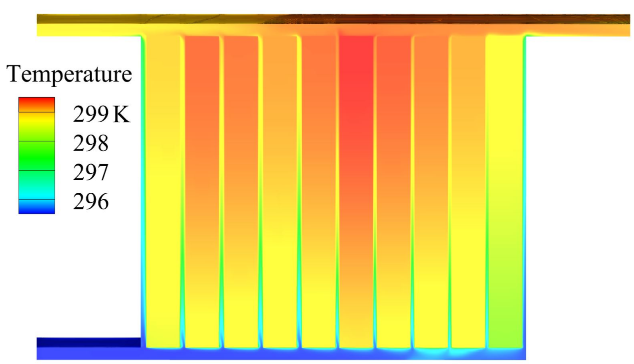
The four manifold sizes are considered as the design variables, as shown in Fig. 3.20. The inlet of the distribution manifold is defined as b_1+b_2 to prevent backflow. Similarly, the



(a) U-type



(b) Z-type



(c) J-type

Figure 3.19: CFD simulation results of the optimal designs of U-, Z-, and J-type BTMS

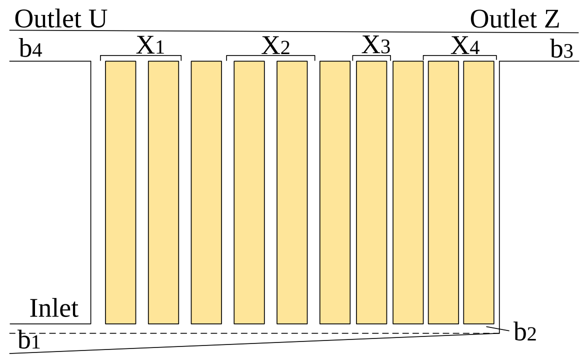


Figure 3.20: The second stage optimization of J-type BTMS

objective is to minimize the maximum temperature $T_{max} = f(b_1, b_2, b_3, b_4)$, given by:

$$\begin{aligned} \arg\min_{b} \quad & T_{max} = f(b_1, b_2, b_3, b_4) \\ \text{subject to} \quad & 0 \leq b_1 \leq 9 \\ & 3 \leq b_2 \leq 7 \\ & 0.5 \leq b_3 \leq 15 \\ & 0.5 \leq b_4 \leq 15 \end{aligned} \tag{3.15}$$

The optimizations are performed under two different working conditions: (1) the benchmark condition, and (2) the comparative condition that has double heat generation rate and double air flow rate compared to the benchmark condition. A number of surrogate models are also constructed here using the same surrogate modeling method discussed in the previous section, and the performances are shown in Fig. 3.21. The entire process terminates after two rounds of resampling and optimization. Figure 3.22 shows the optimization and resampling results under both the benchmark and comparative conditions. It is seen that the resampling under the comparative condition is not as concentrated as the benchmark condition due to accuracy difference.

Table 3.4 highlights the optimal configuration results under the two conditions. By comparison, the optimal manifold configuration varies between the benchmark and the comparative conditions. Thus, the optimal setting of BTMS changes with battery working

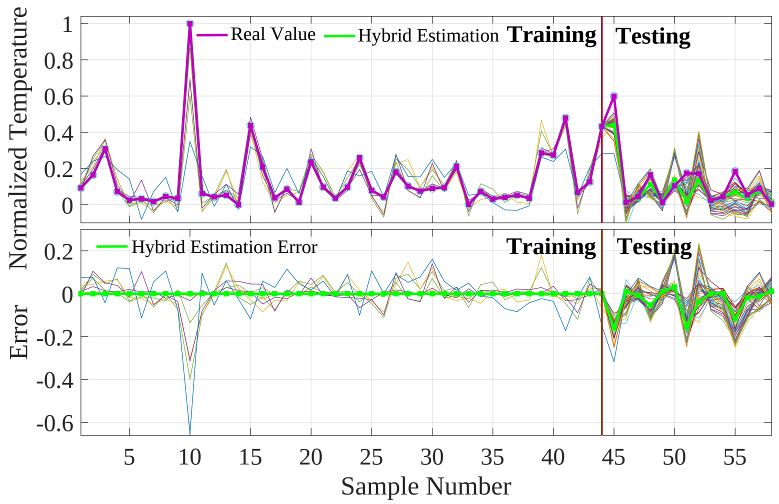


Figure 3.21: Surrogate model evaluations of J-type under the benchmark condition (only a portion of the scattered points are presented)

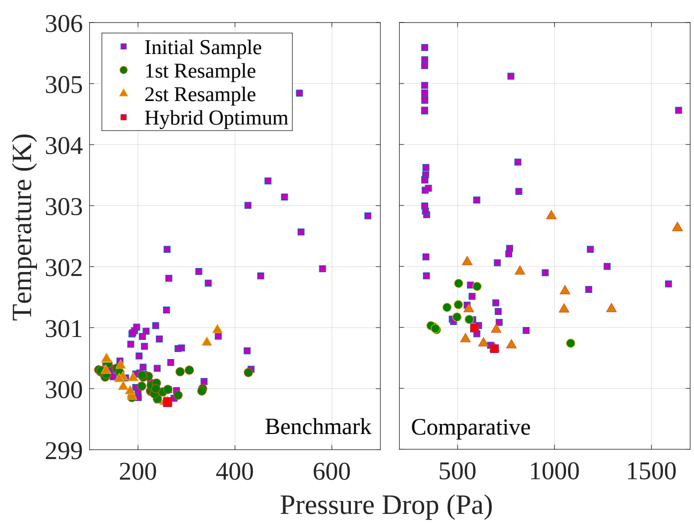


Figure 3.22: The J-type manifold optimization process for the benchmark and comparative conditions

condition, and an air-based BTMS with a fixed structure is unable to always work in its optimal settings under changing working conditions. The proposed J-type cooling system has introduced more flexibility with two controlling valves. Coupled with its optimal control strategy, the J-type system is capable to adapt itself to the changing conditions.

Table 3.4: The J-type optimized results under different conditions

	Design Variable (mm)				Benchmark condition			Comparative condition		
Types	b_1	b_2	b_3	b_4	$T_{max}(K)$	$T_\sigma(K)$	$\Delta P(Pa)$	$T_{max}(K)$	$T_\sigma(K)$	$\Delta P(Pa)$
J-stage1	0	6.0	6.0	6.0	299.98	0.46	391.28	301.9	0.86	1,233.3
J-stage2B	1.28	6.05	13.62	8.23	**299.78**	0.28	260.8	301.56	0.75	894.25
J-stage2C	2.5	6.55	11.31	10	300.1	0.37	261.9	**300.66**	0.34	690.4

J-stage1: optimal case in Stage 1 optimization
J-stage2B: optimal case in Stage 2 optimization under the benchmark condition
J-stage2C: optimal case in Stage 2 optimization under the comparative condition

To validate the optimized results, an experimental platform is established to study the thermal performance of different pure-air based BTMS, in which different types of BTMS can be shifted flexibly. As shown in Fig. 3.23, the platform consists of three parts, i.e., the air and power supply section, the battery model section, and the measurement section. The equivalent heat source of the battery cell is replaced with two heaters, which are controlled by the DC power source. The detailed settings of the platform can be found in Ref. [49]. Several groups of experiments are performed to mutually validate the benchmark and optimized cases.

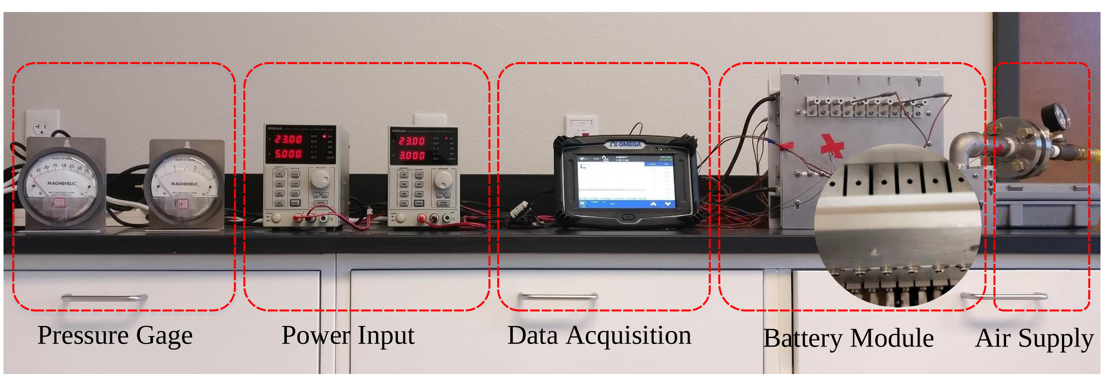


Figure 3.23: The air-based BTMS experimental platform

Figure 3.24 compares the simulation and experimental temperature distributions for both the benchmark and optimal cases. There are only small discrepancies observed between simulations and experiments, which is mainly attributed to the settings of temperature measuring position. The maximum temperature in experiments is measured by a K type thermocouple that is inserted deep inside the top part of an aluminum model cell, while the maximum temperature is directly extracted from the lateral surface in simulations. Other potential factors such as measurement accuracy, initial condition settings, the bias of grouped-channel size, and CFD turbulence model, may also contribute to the small differences between simulations and experiments. Overall, the experiment results agree with the simulation results, which further validates the parametric analysis and optimization results in this study.

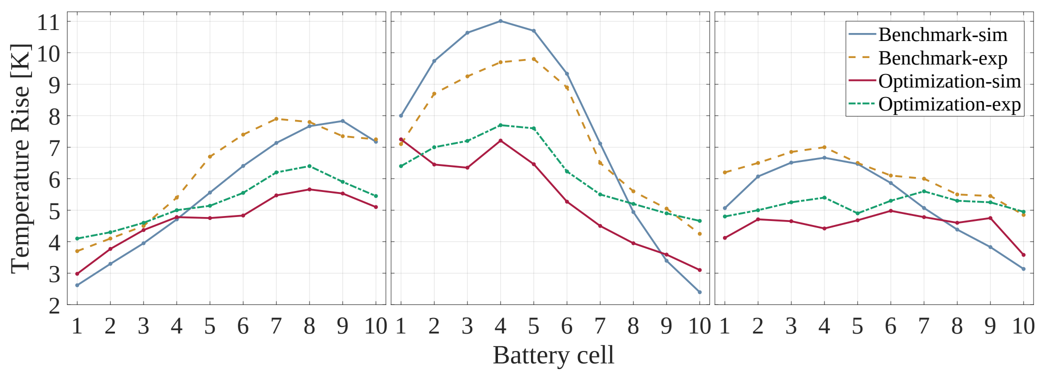


Figure 3.24: Comparing results between simulations and experiments (From left to right: U-, Z-, and J-type)

3.4 Summary

This chapter proposed a novel J-type air-based battery thermal management system by integrating the conventional U-type and Z-type structures. An electro-thermal model for Lithium-ion battery was developed, based on which a comparative parametric study of several structural and controlled variables was performed. It is recognized that the geometry of the flow field has a significant impact on thermal performance and flow efficiency. Several priori optimized structures like grouped-channel and tapered manifold were also investigated to develop fundamental understandings for further optimization. Results showed that the tapered distribution manifold brings no improvements to the U-type structure, and so does the tapered collecting manifold to the J-type structure.

A surrogate-based optimization framework was proposed, which consists of a surrogate pool and a two-stage clustering-based resampling approach. Structural channel size optimizations were first performed to improve the grouped-channel structures for U-, Z-, and J-type battery thermal management system, in which the Gaussian mixed model resampling method was adopted to improve the accuracy. The optimal results showed that there were 35.3%, 46.6%, and 31.18% reduction in the temperature rise for U-, Z-, and J-type, respectively. A J-type manifold configuration optimization was also performed to further explore the optimal settings of the system under different working conditions. It demonstrated that the optimal settings of J-type structure vary across the working conditions, and the J-type system is able to be adaptively controlled to the optimal settings by adjusting the two valves. The simulations and optimizations were further verified by experiments.

CHAPTER 4

THERMAL CONTROL STRATEGY FOR J-TYPE AIR-BASED BTMS[1]

To mitigate the gap between structural design and its corresponding control strategy, a novel control-friendly J-type air-based BTMS has been proposed in Chapter 3. By integrating the advantages of conventional U- and Z-type structures, the J-type structure is distinguished by threefold: (i) there are two outlets in the J-type structure compared to one outlet in the U- and Z-type structure; (ii) the flexibility of the controllability have been significantly enhanced by employing two control valves to adjust the opening degree of each outlet simultaneously according to varying driving conditions; (iii) three control modes are predefined for simplification and the temperature uniformity can be enhanced by switching between different working modes, namely, U-, Z-, and J-mode. This chapter extends the control strategy for the J-type BTMS conjointly and seeks to develop an adaptive control approach for the air-based cooling system via a model predictive control (MPC) strategy. It is worth mentioning that prior to any control implementations, the channels of the J-type structure need to be pre-optimized by using a surrogate-based optimization algorithm, for the reason that the J-type optimal configuration is able to uniformize the temperature distribution within a narrow range and lengthen the operation time in J-mode with a higher efficiency.

4.1 Thermal Control System Modeling

In this chapter, we aim to establish the plant model and the controller model for the air-based BTMS via neural network (NN), where the modeling data is obtained from transient CFD simulations with a time step of 5 seconds. This model consists of three parts: the plant, the controller, and a control mode switcher. It also employs a hybrid driving cycle as the system input.

[1]Y. Liu and J. Zhang (2019), Self-adapting J-type Air-based Battery Thermal Management System via Model Predictive Control, *Applied Energy*, Vol.263, pp.114640. Reprinted with permission from Elsevier.

4.1.1 J-type Operation Mode Design

For the J-type structure, based on the framework developed in Chapter 3, we rearrange the channels into 5 groups and performed a channel size optimization as a trade-off between the optimization efficiency and computational complexity. Another purpose is to design a J-type structure with an appropriate temperature distribution that enables flexible thermal control, e.g., the maximum temperatures occur at the battery cells near the two outlets.

Theoretically, the two outlet valves are expected to be continuously adjusted and controlled according to the predefined optimal structure settings and battery operating conditions. However, the implementation of continuous control may unnecessarily consume excessive energy and require a more powerful real-time computational capacity as well. Consequently, it is reasonable to restrain the action interval or the change rate of the valve opening level. According to the parametric studies in Ref. [51], only three structures and their corresponding control modes are predefined in this paper for the sake of simplification, namely, J-mode, Z-mode, and U-mode. The opening of the two valves under different working modes are represented by the configuration sizes of their corresponding outlets, as predefined in Table 4.1. The time intervals between switching actions among the three modes are not predefined, which highly depends on the real-time condition.

Table 4.1: Predefined control mode settings

Control mode	Bottom inlet (mm)	Left outlet (mm)	Right outlet (mm)
J-mode	6	6	6
U-mode	6	8	4
Z-mode	6	4	8

Based on the aforementioned predefined three operation modes, transient CFD simulations are performed to generate the raw data during dynamic responses. Figure 4.1 shows the framework of the transient flow simulation, where the model inputs I_t and SoC_t stand for the Euclidean norm of current and the arithmetic average of SoC_t within a single time step, respectively; and T_t denotes the real-time battery temperature. In consequence, the response output of the transient simulation is the dynamic temperature augments of the battery cells labeled as 2, 4, 7, and 9 within a time step δt, namely, $([T_{2t},\ T_{4t},\ T_{7t},\ T_{9t}])$. In this paper, the time step δt is set to be 5 s after comprehensive considerations of the computational cost, algorithm, and simulation accuracy. Each transient simulation takes approximately 45 min to converge on a six-core workstation. Note that the incoming cooling air temperature is only limited to 300 K, and scenarios with varying incoming temperatures are not yet considered in this study.

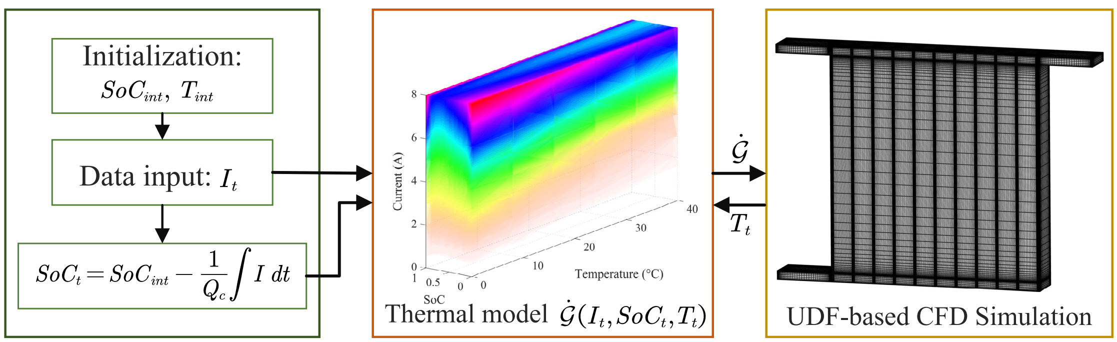


Figure 4.1: The UDF framework of the transient flow CFD model

4.1.2 Battery Temperature Prediction Model

The overall framework of the J-type BTMS thermal control system is presented in Fig. 4.2. In the primary stage concept design, the thermal control system prototype is implemented in the absence of an actual battery pack system since it is not readily available. A battery temperature prediction model is established here to represent the dynamic response of the battery system, which is also referred to as the plant model. Owing to the complexity and abusive assumptions of multi-channel parallel flow, it is impractical to estimate the dynamic temperature distribution using an analytical approach. In this study, a data-driven approach is adopted to implement the system identification, which reveals the multi-input multi-output

(MIMO) relationship between the plant inputs (e.g., mass flow rate, initial temperature, and equivalent battery heat source) and the plant outputs (e.g., the updated temperatures of next time step), as defined by:

$$[T_{2_{t+1}}, T_{4_{t+1}}, T_{7_{t+1}}, T_{9_{t+1}}] = f(T_{ave}, \dot{m}, \dot{\mathcal{G}}) + [T_{2_t}, T_{4_t}, T_{7_t}, T_{9_t}] \tag{4.1}$$

where T_{ave} is defined as the average of the initial battery temperature. Note that the equivalent heat source is calculated based on the initial electric characteristics and temperature in each time step.

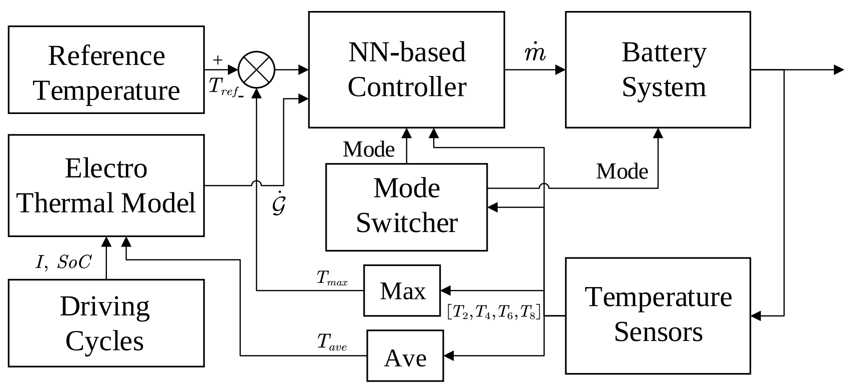


Figure 4.2: The thermal control framework of the J-type BTMS

According to the concept design, three control modes are predefined for different scenarios. There are approximately 650 cases for each mode to be conducted with the transient flow CFD model based on the optimized BTMS structure. Sixty percent of the cases are generated using the Latin hypercube design of experiments algorithm, while the rest samples emphasize on the feasible design areas near the lower and upper bounds to increase the accuracy of sensitive regions.

Several typical stochastic black-box algorithms have been developed for the data-driven discrete time system identification, i.e., the auto-regressive moving average model, auto-regressive moving average with exogenous input, sparse identification of nonlinear dynamics with control, and neural network (NN) model [76, 11]. Since all the raw samples are extracted from CFD simulations separately and randomly, the back propagation NN approach is employed here with the Levenberg-Marquardt training algorithm to establish the plant model. The hidden layers and neurons are determined by exhaustive searching after cross-validation, as tabulated in Table 4.2, where the multivariate normalized maximum absolute error (NMAE) and multivariate normalized root mean square error (NRMSE) are defined as:

$$NMAE_{mv} = \sum_{q=1}^{m}[\alpha_q(\frac{1}{n}\sum_{k=1}^{n}\left|\frac{\hat{y}_{qk} - y_{qk}}{y_{q_{max}} - y_{q_{min}}}\right|)] \tag{4.2}$$

$$NRMSE_{mv} = \sum_{q=1}^{m}[\beta_q \frac{1}{y_{q_{max}} - y_{q_{min}}}\sqrt{\frac{\sum_{k=1}^{n}\left(\hat{y}_{qk} - y_{qk}\right)^2}{n}}] \tag{4.3}$$

$$\sum_{q=1}^{m}\alpha_q = 1 \qquad \sum_{q=1}^{m}\beta_q = 1 \tag{4.4}$$

where m and n stand for the sample dimension and sample size, respectively. The coefficient α_q and β_q are assigned with an equal weight, α_q=β_q=0.25.

Table 4.2: Plant model accuracy evaluation

Mode	U-mode	Z-mode	J-mode
Layers&neurons	[5,4]	[7,5]	[4,4]
NMAE(%)	6.31	6.28	7.72
NRMSE(%)	5.95	5.86	7.37

The mapping results are not as accurate as those of the surrogate models established for structure optimization in Ref. [51], in which the training data is based on steady-state simulation. The main reason is that transient simulations have a large number of cases that have a relatively small mass flow rate, in which the channel flow rate distribution largely differs from the benchmark case. Note that by adding more samples, the accuracy of the plant model could be potentially further improved. Given the highly nonlinear nature of fluid dynamics, the accuracy of the plant model via back propagation NN is reasonable and acceptable to be integrated into the thermal control system.

4.1.3 Neural Network-based Controller

As regards the controller, conventional control strategies like proportional-integral-derivative (PID) control may be not adequate to satisfy the control performance requirement for a fluid system. Advanced strategies like fuzzy logic control, adaptive neuro–fuzzy inference system, artificial neural network (ANN), NN-based adaptive PID control, and PID NN, have been successfully implemented in the literature [1]. Neural network is adopted here to construct and tune the controller using the same training data from transient simulations, in which the mass flow rate input is reversed as the output of the NN-based controller. The control goal is to generate an appropriate mass flow rate based on the temperature bias and operating mode, so as to closely follow the predefined temperature trajectory. Note that the actuator (e.g., cooling fan or compressor) is not considered here, since the output of the controller is same as the actuator. The relationships of the controller are expressed as:

$$\begin{aligned} &\dot{m} = g(T_{bias}, T_{ave}, \dot{\mathcal{G}}) \\ &T_{bias} = T_{ref} - max(T_2, T_4, T_7, T_9) \\ &T_{ave} = (T_2 + T_4 + T_7 + T_9)/4 \end{aligned} \tag{4.5}$$

where the temperature bias is defined as the disparity between the reference temperature T_{ref} and the maximum temperature T_{max}, and the operating temperature is represented by the average temperature T_{ave}. The feasible range and change rate of the mass flow rate are constrained by the actuator's actual operating limitations. Here, the constraints of the NN-based controller are predefined as:

$$\begin{aligned} &0 \leq \dot{m} \leq 0.014 \\ &|\Delta \dot{m}| \leq 0.0035 \quad (kg/s) \end{aligned} \tag{4.6}$$

When the whole control system is put into practice, advanced adaptive self-tuning control algorithms, e.g., model identification adaptive control and model reference adaptive control, can be applied to adjust and improve both the plant model and the controller simultaneously by integrating the simulated results and real measured data.

4.1.4 Control Mode Switcher

As demonstrated in Fig. 3.3, the J-mode is designed for normal operation with higher efficiency, while the U-mode does bring in a stronger heat dissipation capacity to the left part, and the Z-mode acts similarly on the right part. By quasi-periodically switching among the three modes, the temperature difference is expected to be fully constrained within a reasonable range. The switch logic for U- and Z-mode is defined as:

$$|T_{left} - T_{right}| = |max(T_2, T_4) - max(T_7, T_9)| \geq T_{uz} \tag{4.7}$$

where T_{left} and T_{right} refer to the maximum temperature of the left and right parts, respectively. The critical temperature for switching to U- and Z-mode is set as 0.5 K. When the left part has a higher temperature, the working mode will be switched to U-mode, and vice versa.

After three switches between the U- and Z-mode, the next action is to return back to the J-mode, and a switching cycle completes, as given by:

$$|T_{left} - T_{right}| \leq T_j \tag{4.8}$$

where the critical temperature switching to the J-mode (T_j) is set to be 0.2 K. Each cycle traverses a J-, U-, Z-mode, and a half period of either U- or Z-mode in sequence, as illustrated in Algorithm 2. The duration of a switching cycle highly depends on the dynamic working conditions.

4.2 Case Study and Discussion

To test and validate the thermal control system, a hybrid driving cycle is employed as the dynamic inputs, while a predefined temperature trajectory serves as the target for thermal control. For a driving vehicle, the equivalent traction power can be estimated using the

Algorithm 1: Control mode switching logic

Criterion U : $T_{left} - T_{right} \geq T_{uz}$
Criterion Z : $T_{right} - T_{left} \geq T_{uz}$
Criterion J : $|T_{left} - T_{right}| \leq T_j$
Initialization: $Mode$=J, $Record$=[J]
while *control prcoess contiunues* **do**
    if *Criterion U* & $Mode \neq U$ **then**
        $Record \Leftarrow [Record, U]$
        $Mode \Leftarrow U$
    end
    if *Criterion Z* & $Mode \neq Z$ **then**
        $Record \Leftarrow [Record, Z]$
        $Mode \Leftarrow Z$
    end
    if $Record$==([J, U, Z, U]||[J, Z, U, Z]) & *Criterion J* **then**
        $Record \Leftarrow [J]$
        $Mode \Leftarrow J$
    end
end

power expression, given by:

$$P = \frac{V}{1000\eta}[mg\mu cos\alpha + mgsin\alpha + \frac{1}{2}\rho A_f C_d V^2 + m\frac{dV}{dT}] \tag{4.9}$$

Note that the regenerative energy from the braking system with an efficiency of 0.8 will be recharged back into the battery system instantly. Specific vehicle parameters and driving cycle conditions are tabulated in Table 4.3, where the vehicle is assumed to be running on a level road. For the electrical settings of the vehicle, the EV's battery system consists of 6 battery modules in parallel, and every module has 110 battery cells in series. For the specific open-circuit voltage (OCV)-SoC, polynomial regression model is adopted here to represent the relationship between OCV and SoC, given as:

$$V_{oc} = -40.2SoC^6 + 138.6SoC^5 - 186.2SoC^4 + 123.5SoC^3 - 42.4SoC^2 + 7.5SoC^1 + 3.32 \tag{4.10}$$

Table 4.3: EV specification and driving condition (TESLA Model 3)

Mass	m	1,875 kg	Road gradient	α	0
Windward area	A_f	2.22 m^2	Air friction Coeff	C_d	0.24
Standard gravity	g	9.8 m/s^2	Motion efficiency	η	0.98
Air density	ρ	1.16 kg/m^3	Rolling resistance	μ	0.01
Velocity	V	-	Regenerative Coeff	η_r	0.8

In a dynamic process, the SoC is generally defined and calculated using the Coulomb counting [17], as given by:

$$SoC = SoC_{int} - \frac{\int I dt}{Q_c} \tag{4.11}$$

where SoC_{int} and Q_c denote the SoC at the initial stage and the battery capacity, respectively. I is the battery current, which is positive for discharging and negative for charging.

To be consistent with the transient simulations, the time step is extended to 5 s (δt=5 s) for simulation and control convenience. The quadratic mean of current and the arithmetic mean of SoC are derived to represent the initial electric characteristics for simulations. Moreover, the reference temperature trajectory is supposed to be determined and adjusted adaptively according to the real-time driving conditions, e.g., strenuous driving requires an aggressive temperature augment trajectory, while a gradual rise temperature curve is sufficient enough for smooth driving. In this paper, considering the driving cycles illustrated above, the reference temperature trajectory consists of a climbing section from 303 K to 310 K in 50 min and a stable section of 310 K. The maximum temperature is expected to be fully restrained within 313.15 K with a target temperature of 310 K.

4.2.1 Neural Network-based Control without Mode Switching

Based on the driving cycles as the system inputs, two control studies are conducted depending on whether the operation mode switching module is activated or not, in which the basic

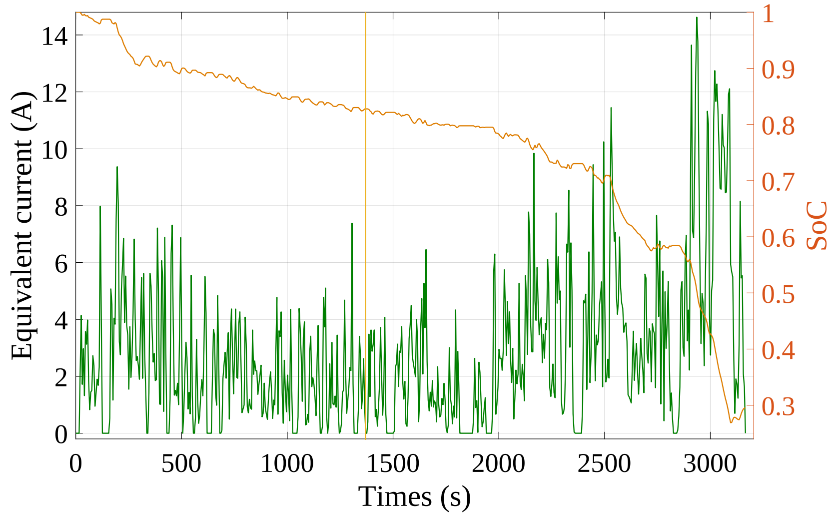


Figure 4.3: The profiles of equivalent charging/discharging current and SoC

control strategy without mode switching module is conducted as a benchmark. As defined in Fig. 4.2, the mode switching module is bypassed. As a result, the battery system only operates in J-mode. By applying the comprehensive driving test cycle, the dynamic temperature distributions and mass flow rate are obtained, as shown in Fig. 4.4. It is observed that the maximum temperature of the battery pack is able to follow the reference temperature by controlling and adjusting the mass flow rate. However, the trend of temperature deviation starts in approximately 2,500 s due to the accumulation of a large amount of heat, though the mass flow rate has already been controlled and increased to its upper bound, as presented in Fig. 4.5.

Another evaluation criterion, the temperature uniformity characterized by the temperature standard deviation, also deteriorates in the meantime, as shown in Fig. 4.5. It should be noted that the J-type structure is optimized under the setting of approximately 52% full mass flow rate. Under the full mass flow rate condition, the cooling effects are more significant on the right part of the battery pack, making the heat dissipation of the 7^{th} and 9^{th} cells much stronger as shown in Fig. 4.4, due to the distinct nature of the optimized J-type structure. Once the external heat source drops off, it is foreseeable that the maximum

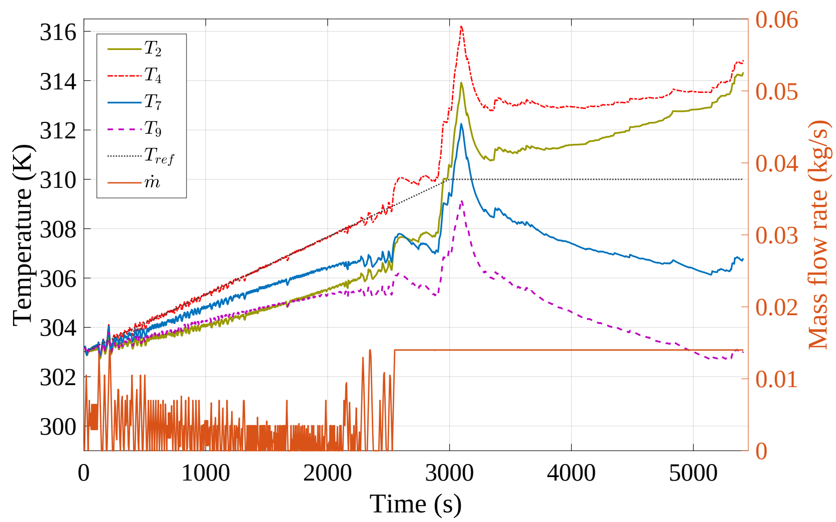


Figure 4.4: The battery temperatures and mass flow rate using the NN-based thermal control without mode switching

temperature will be restored back to the reference temperature, while the lower temperature will be kept at the current level.

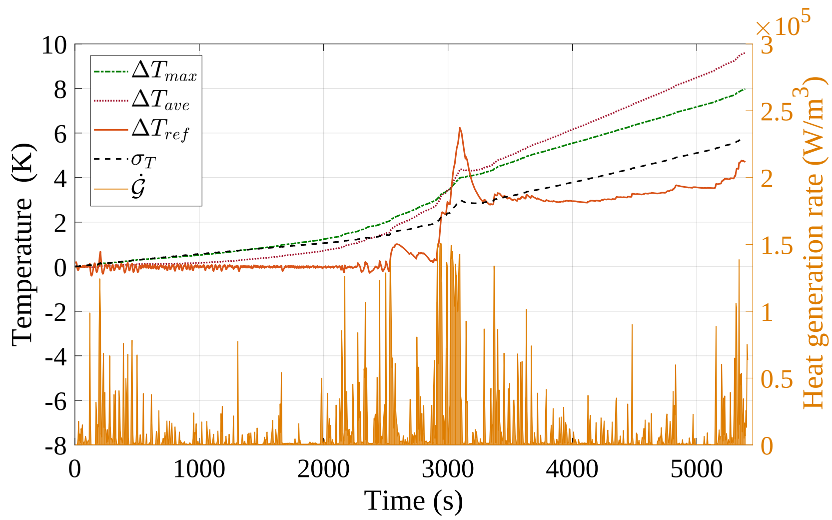


Figure 4.5: Temperature uniformity and equivalent heat source (ΔT_{max} is the temperature difference between the maximum temperature of the left part and the right part, which is also the control temperature for mode switching. ΔT_{ave} is the difference between the average temperature of the left part and the right part. ΔT_{ref} is the difference between the maximum temperature of the battery pack and the reference temperature. σ_T is the temperature standard deviation of the battery pack.)

It is seen from Fig. 4.4 that the maximum battery pack temperature difference under the NN-based control without mode switching is 11 K, due to the temperature deviations accumulated during the dynamic process. Our previous study has already shown the distinguished advantages of the J-type structure in terms of temperature uniformity compared to the traditional U- and Z-type structures [51]. Thus the dynamic thermal performance will be even worse for traditional structures, i.e., U-type and Z-type. There is scarcely any solution to address this issue but to increase the mass flow. Instead of increasing the capacity of mass flow rate, we will explore if the mode switching technique can further improve the performance of BTMS in this study.

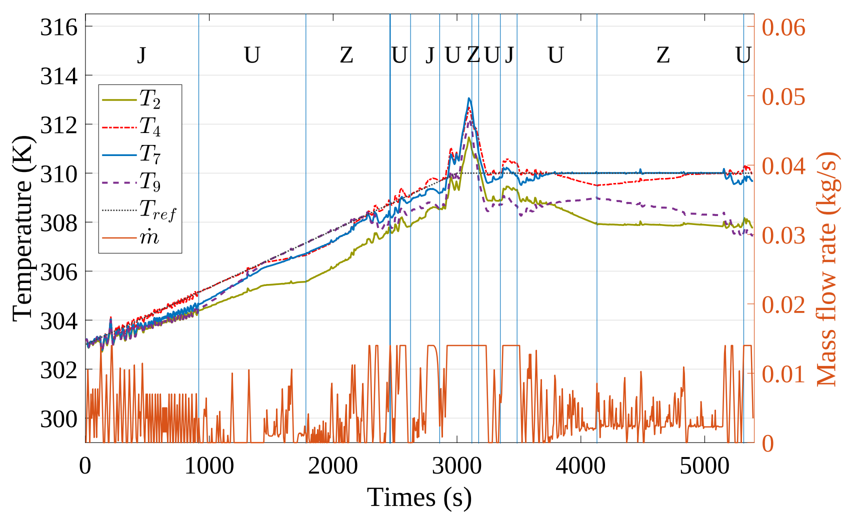


Figure 4.6: The battery temperatures and mass flow rate using the NN-based thermal control with mode switching

4.2.2 Neural Network-based Control with Mode Switching

Considering the mode switching module in the control framework as shown in Fig. 4.2, the corresponding responses regarding the temperature and mass flow rate are calculated, as shown in Fig. 7.8. Compared to the results without mode switching in Fig. 4.4, it is seen that the temperature difference between T_2 and T_9 exceeds the switching threshold 0.5 K at around 900 s, the temperature difference stops increasing by switching the operation mode from J-mode to U-mode, and followed by mode switching from U-mode to Z-mode, and continues. The battery temperatures closely follow the reference trajectory during the entire process, except under the intense driving conditions at around 3,000 s, where the battery temperature slightly exceeds the predefined ceiling temperature.

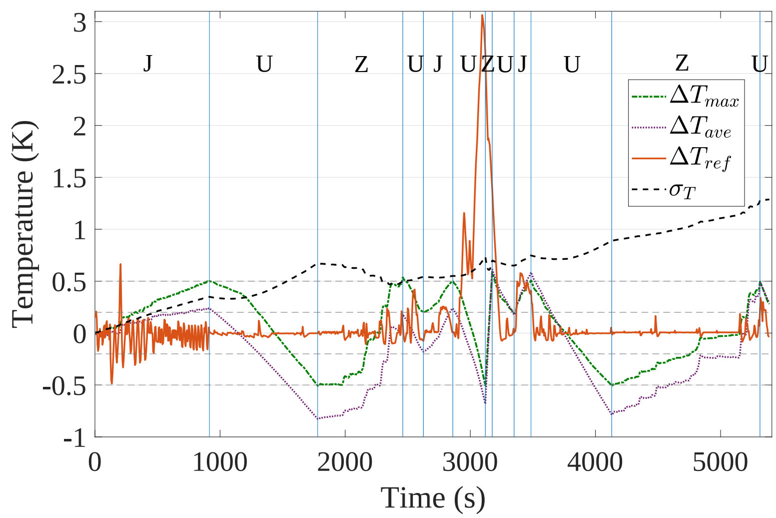


Figure 4.7: The mode switching details and battery temperature uniformity of the NN-based control

The temperature uniformity under the NN-based control with mode switching is also significantly better than that of the control without mode switching, as shown in Fig. 4.7. In addition, Fig. 4.7 also shows the tendency of the temperature difference between the left part and the right part during the entire transition process. Overall, we find that the NN-based control with mode switching is able to control the maximum temperature and balance the temperature uniformity simultaneously.

4.2.3 Neural Network-based Model Predictive Control

In this task, an MPC approach is employed to integrate with the NN-based control system. As shown in Fig. 7.8, there exist very rapid changes to the mass flow rate, while the battery temperatures are well controlled according to the reference temperature. This is mainly due to that the NN-based control strategy only considers the temperature bias of the last step and the external input at the current step, and does not foresee the possible approaching drastic changes in the external input. By foreseeing several steps ahead, the MPC approach is expected to further improve the control performances in terms of thermal requirements and energy efficiency.

The overall control framework is illustrated in Fig. 4.8, where the discrete time nonlinear MPC module is added to forecast and optimize the controller output. The main objective of MPC is to minimize the cost function as follows:

$$\begin{aligned}
\arg\min_{\dot{m}} \; J = & \sum_{k=n}^{n+N} (\alpha_k \sum_{b=2}^{[2,4,7,9]} (T_{ref_k} - T_{b_k})^2) + \sum_{k=n}^{n+N-1} \beta_k (\dot{m}_{k+1} - \dot{m}_k)^2 \\
\text{subject to} \quad & 0 \leq \dot{m} \leq 0.014 \\
& |\Delta \dot{m}| \leq 0.0035
\end{aligned} \tag{4.12}$$

where n is the current control step, and N is the control horizon, here, N=4. The parameter α_k is a weight factor that reflects the relative importance of the temperature difference, while the coefficient β_k penalizes the relatively big changes in the mass flow rate. By adjusting the two coefficients, different control strategies can be achieved for different purposes.

Owing to the high nonlinearity of the control problem, the genetic algorithm (GA) optimizer is adopted here to solve the optimization problem by considering several steps ahead. It should be noted that GA does not have high computational efficiency. For optimal control, to better balance the performance and computational efficiency, a local optimum (or a better solution) is acceptable here as long as the performance is improved. The optimization convergence criteria can be adjusted and tuned based on the time requirement and accuracy settings.

The battery temperatures and mass flow rate with the NN-based MPC strategy is shown in Figure 4.9. While the battery temperatures well follow the reference temperature at most

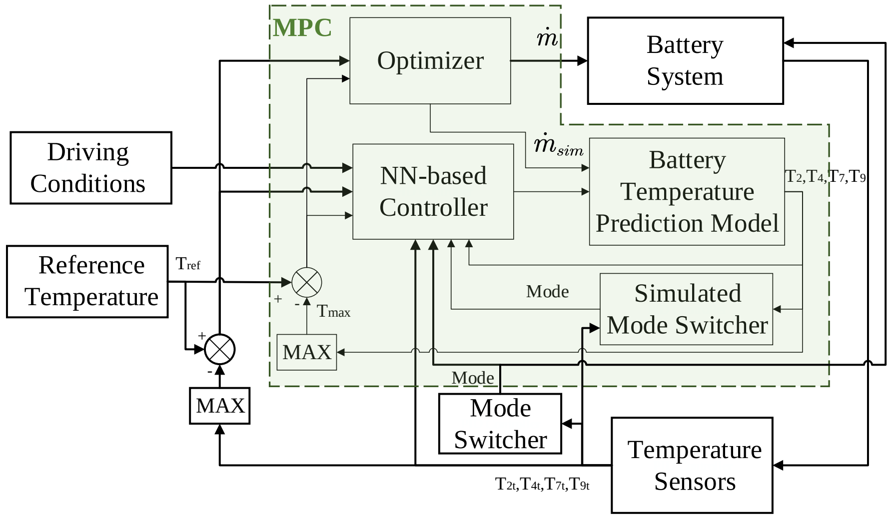


Figure 4.8: The overall framework of the NN-based MPC strategy

of the time, there still exist deviations at several time periods, e.g., 400-1,000 s, 2,000-2,400 s, and after 3,200 s. The main reason is that the penalization term (i.e., the second term in Eq. 5.8) dominates the cost function. Whenever the penalty cost of increasing the mass flow rate prevails over that of the temperature deviation, the MPC strategy tends to reduce the variability of the mass flow rate, so as to improve the smoothness of the thermal control system as well as the energy efficiency. Similarly, the temperature changes during the mode switching process and temperature uniformity are presented in Fig. 4.10. It is seen that the temperature uniformity is 7.4% better than that without MPC, though the temperatures deviate from the reference trajectory.

4.2.4 Discussion

To compare the energy performance, it is assumed that the cooling air is provided by an axial flow fan for simplification, and the energy consumption of BTMS can be calculated by:

$$N_p = \int_0^n \frac{\dot{m}\Delta P}{\rho\eta} dt \tag{4.13}$$

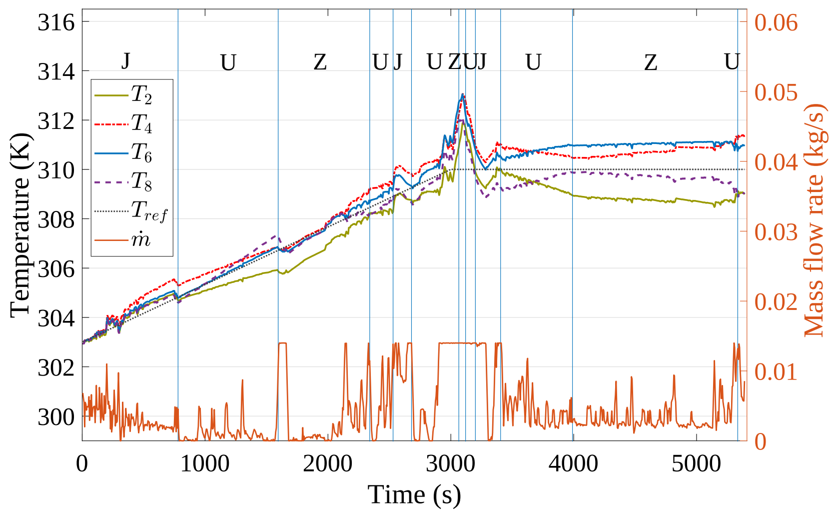


Figure 4.9: The battery temperatures and mass flow rate with the NN-based MPC strategy

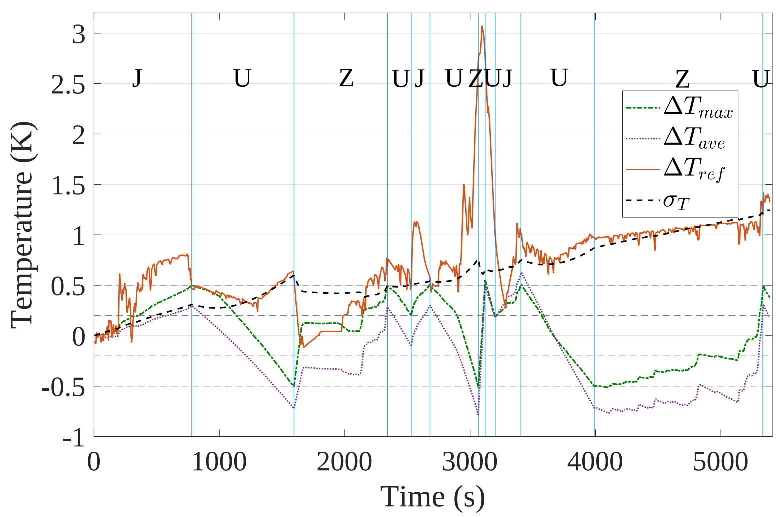


Figure 4.10: The mode switching details and battery temperature uniformity of the NN-based MPC strategy

where ΔP is the pressure augment of the fan, ρ is the air density, and η is the overall efficiency (η=0.75).

The pressure augment relates to the mass flow rate, which can also be obtained via transient flow simulations. In this study, a support vector regression model is established to represent the pressure augment as a function of the mass flow rate. The energy consumption, as well as the maximum temperature and the temperature uniformity, is tabulated in

Table 4.4. It is observed the NN-based control without mode switching fails to meet the thermal requirements in terms of the maximum temperature and temperature uniformity. In addition, by employing MPC, the energy efficiency has an approximately 15.8% improvement compared to NN-based with mode switching, with a cost of slight deviation from the reference temperature. Given the critical operation temperature is 313.15 K, it is reasonable to employ the MPC strategy with a reference temperature of 310 K.

Table 4.4: Summaries of the three control strategies

Control strategy	Maximum temperature (K)	Temperature uniformity (K)	Energy (J)
Control without mode switching	316.10	5.81	51,384
Control with mode switching	313.17	1.33	17,412
MPC with mode switching	313.12	1.25	14,678

4.3 Summary

This chapter developed a self-adaptive intelligent air-based J-type battery thermal management system via neural network-based model predictive control. Based on the optimized structure and the established electro-thermal-fluid model, a large number of transient fluid dynamics simulations were conducted for the three predefined operation modes (i.e., J-, Z-, and U-mode) to build control models using the neural network algorithm. By applying dynamic driving cycles, the impacts of mode switching was investigated in the case study. Results showed that the NN-based control without mode switching failed to control the thermal system in terms of both the maximum temperature and temperature uniformity, while the NN-based control with mode switching was able to regulate the maximum temperature within the critical safety temperature and balance the temperature uniformity within 1.5 K

under varying working conditions. It was also found that there was a 15.8% energy efficiency improvement by employing model predictive control.

It was verified that the developed self-adaptive BTMS control strategy was able to meet the thermal requirements for battery system. However, the model predictive control in this study emphasized more on the energy efficiency rather than the maximum temperature control. More studies need to be performed on the two coefficients in model predictive control to develop a balanced control strategy.

CHAPTER 5

MODEL PREDICTIVE CONTROL-BASED ENERGY MANAGEMENT STRATEGY FOR ELECTRIC VEHICLES[1]

Building on state-of-the-art thermal management strategies, this chapter seeks to investigate the discharging scheduling and load shifting from the thermal perspective, in which the process starts with a thermal control of the optimized battery pack and AC system, and ends up with evaluations of systematic energy efficiency. The research objective is to develop an MPC-based strategy to improve the overall energy efficiency and battery cycle-life while well retaining thermal constraints of the battery pack. The established neural network-based J-type air-based thermal control system is inherited from Chapter 4, in which both the plant model and controller are established with data-driven models. By controlling the operation mode and the mass flow rate simultaneously, the developed BTMS has been proved to be able to maintain both the maximum temperature and the uniformity within expected ranges simultaneously. Moreover, an air precooling module is added to the existing J-type BTMS with extra cooling capability to mitigate thermal impacts from severe working conditions.

5.1 Dynamic Modeling for EVs

5.1.1 Air Conditioning System Model

It is essential to establish a dynamic electro-thermal model of the vehicle-mounted air conditioning system for energy management. According to an auxiliary system impact report from the Idaho National Laboratory [34], the air conditioning system may consume up to 30% of the traction battery energy for cooling, depending on the air flow volume, and the temperature difference between the ambient environment and the cabin.

[1] from Y. Liu and J. Zhang (2021), Electric Vehicle Battery Thermal and Cabin Climate Management Based on Model Predictive Control, *Journal of Mechanical Design*, Vol.143, Issue 3, 2021, pp.031705.

Energy Model

For simplification, the heat balance model of the electrical air conditioning system is established separately with the air conditioner model and the thermal load model, and only the temperature feature is considered in this study. There are three major energy consuming components in the AC system, i.e., the compressor, the evaporator blower, and the condenser fan, as presented in Fig. 5.1. Regarding the air conditioner modeling, while a full range component level AC simulation model provides more detailed evaluations of various components with a higher accuracy, it is very challenging to manage the mathematical and computational complexity in terms of phase-changing flow, thermodynamics, and heat transfer [68]. Since this work emphasizes more on the energy perspective, the AC energy model developed in Ref. [31] is adopted. The overall power consumption of the AC system can be simplified and estimated using a ratio coefficient between the cooling capacity provided and its corresponding power consumed by the AC system, which is also referred to as the coefficient of performance (η_{cop}), given by:

$$\eta_{cop} = \frac{Q_{ac}}{P_{ac}} = \mathcal{F}(T_{in}, T_{ex}, P_{lr}) \tag{5.1}$$

where η_{cop} is a function of the uniform internal and ambient temperatures, and the partial load ratio P_{lr}. The model is implemented using Gaussian process regression based on the essential data from the study conducted by Pino et al. [67].

Thermal Load Model

For the thermal load model, several external and internal heat sources are generally identified and considered in the cabinet thermal model, as illustrated in Fig. 5.2 and tabulated in Table 5.1 after reasonable assumptions and simplifications. The cabin is assumed to be a trapezoidal box that has a roof panel and an interior base, surrounded by windows. Solar radiation as well as ambient air have significant impacts on internal climate via the roof

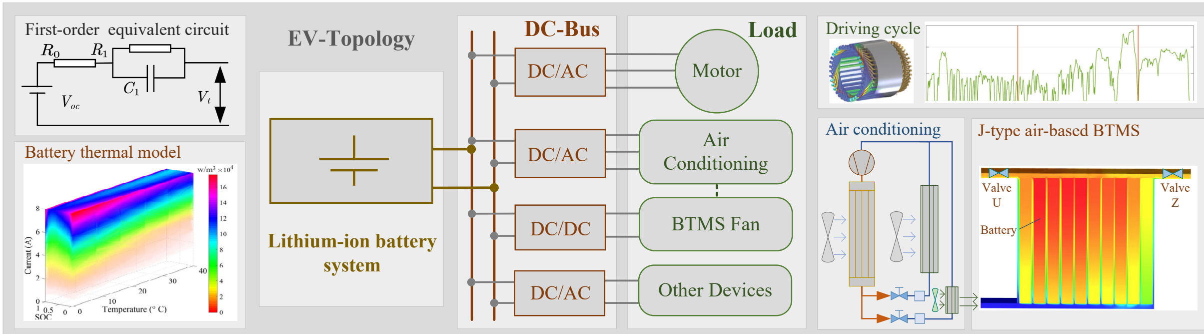


Figure 5.1: The battery topology and main loads

panel and windows, or through window glasses. As regards the heat conduction via the roof panel or windows, it is observed that the surface temperature may probably be higher than the ambient or cabin internal temperature because of solar radiation. The heat conduction from body shell to the cabin is estimated using a heat balance method, as given by:

$$Q_{cr} = \alpha I A - h_{ex} A((\alpha I + T_{ex} h_{ex} + \frac{T_{in}}{\sum \frac{\delta}{\lambda} + \frac{1}{h_{in}}})/(h_{ex} + \frac{1}{\sum \frac{\delta}{\lambda} + \frac{1}{h_{in}}}) - T_{ex}) \tag{5.2}$$

where h_{ex} denotes the convective heat transfer coefficients between the roof panel and external ambient. h_{in} denotes the convective coefficient between the roof panel and internal cabin. I is the solar radiation, and α denotes the absorptivity. T_{in} and T_{ex} are the cabin internal air temperature and the ambient temperature, respectively. The heat conduction via windows Q_{cw} has a similar expression except the value differences of the radiation absorptivity, thickness, and thermal conductivity of glasses [56].

A previous study [24] have found that the equivalent heat transfer coefficient between the roof panel and the external ambient is highly related to the vehicle velocity, while the solar radiative thermal load through windows highly depends on both the operation time in a day and its relative whether condition. The radiation I is selected as 1,200 W/m^2 in this study. Part of the parameter settings in this model come directly from the AC simulation toolbox *coolsim* developed by the National Renewable Energy Laboratory [33] and Ref. [56].

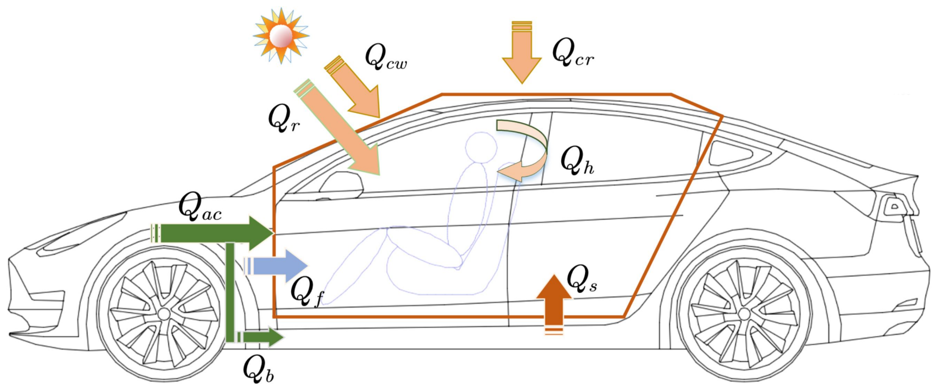


Figure 5.2: The transient thermal model of a vehicle's cabin

It is worth mentioning that the pre-cooling load serves as an accessibility option for BTMS. When it encounters with extreme ambient temperatures or severe operation conditions such as super fast charging and high-speed cruising, this feature is activated to provide a stronger cooling capability towards thermal control via cooling down the approaching air. In general, the control-oriented dynamic temperature response of the vehicle cabinet can be formulated as follows.

$$T_{in(k+1)} = T_{in(k)} + \frac{(Q_{cr} + Q_{cw} + Q_r + Q_h + Q_f + Q_s + Q_b) - Q_{ac}}{\rho_{air} V_{in} C_{air}} \delta t \tag{5.3}$$

where the cabinet volume V_{in} equals to 3 m^3, and δt denotes the time step in seconds. Note that the hysteresis effects from both the power train and the liquid-loop are not considered here. The adjustment of cabin air temperature is controlled by adjusting the total power input for the AC system with basic control logistics. Detailed modeling and controlling of the air conditioning system, including the compressor, the fan, and the condenser, are beyond the research scope of this dissertation.

Table 5.1: Vehicle cabin thermal modeling

Physical term	Symbol	Heat source	Temperature	Estimation (W)	Descriptions and highlights
Conduction/ convection load via roof panel	Q_{cr}	External air solar radiation	Dependent	Eq. 5.2 or without radiation $Q_{cr} = k_{cr} A (T_{ex} - T_{in})$ $k_{cr} = (\sum \frac{\delta}{\lambda} + \frac{1}{h_{in}} + \frac{1}{h_{ex}})^{-1}$	δ: thickness; λ:conductivity; A=3.6 m^2; $h_{in}=25\ W/(m^2K)$; $h_{ex}=4.65 + 13.96\sqrt{v}$
Conduction/ convection load via windows	Q_{cw}	External air solar radiation	Dependent	Eq. 5.2 or without radiation $Q_{cw} = \sum k_{cw} A (T_{ex} - T_{in})$ $k_{cw} = (\frac{\delta}{\lambda} + \frac{1}{h_{in}} + \frac{1}{h_{ex}})^{-1}$	A=1.5 m^2; $h_{in}=25\ W/(m^2K)$; $h_{ex}=4.65 + 13.96\sqrt{v}$
Solar radiation through windows	Q_r	Solar flux	Independent	$Q_r = \sum_{i=1}^{n=4} \eta I A_i sin\theta_i \beta$	Four windows: windshield, rear, left, right; β: shading factor; θ: installation angle; η: penetration rate; I: incident radiation
Human body thermal load	Q_h	Driver & Passenger(s)	Independent	$Q_h = 145 + 116n$ [24]	n: passengers number; n=3
Fresh air thermal load	Q_f	Fresh air	Dependent	$Q_f = \xi \dot{m}_{air} C_{air} (T_{ex} - T_{in})$	Ventilation fresh air portion ξ=12%; $\dot{m}_{air}$=0.186 kg/s
Sensible heat load	Q_s	Cabin interior	Dependent	$Q_s = h_c A_c (T_c - T_{in})$ $T_c(k+1) = T_c(k) - Q_s/(C_c m_c)$	$h_c = 20\ W/(m^2K)$; A_c=8 m^2; m_c=200 kg; C_c=1500 $J/(kgK)$
BTMS pre-cooling load	Q_b	Pre-cool BTMS	Dependent	$Q_b = \dot{m} C_{air} (T_{ex} - T_{air})$	$\dot{m}$: BTMS cooling air flow rate T_{air}: pre-cooled temperature

Control Model

Based on the cabin thermal load model, a control-oriented AC thermal system can be established. The cabin temperature is selected as the system output and the AC cooling capability is chosen as the control variable, as formulated by:

$$\begin{cases} \dot{x} = \dfrac{(Q_{cr}+Q_{cw}+Q_r+Q_h+Q_f+Q_s+Q_b)-u}{\rho_{air}V_{in}C_{air}} \\ y = x \end{cases} \tag{5.4}$$

where both x and y denote the cabin temperature T_{in}, and u represents the cooling capacity Q_{ac}. A basic proportional-integral (PI) controller is developed to control the temperature of the cabin temperature with a targeted value of 294 K. Parameters of the PI controller are tuned based on the situation without the pre-cooling function. For the sake of energy efficiency, limitations other than the targeted value are not imposed on the thermal control.

It is worth mentioning that the pre-cooling thermal load for BTMS, Q_b, only activates at a certain time, and the cabin climate control may probably be affected, since the cooling capacity of the AC system is limited and the BTMS has a higher priority over the cabin thermal management. When the maximum temperature of the battery pack exceeds 311.8 K, the pre-cooling function is activated, whereas it is deactivated after the temperature declines back to 310 K. The constraints are formulated based on the limitations proposed in Ref. [24], including an operation boundary and a changing rate limitation as follows:

$$\begin{aligned} &0 \leq Q_{ac} \leq 4500\eta_{cop} \\ &|\Delta Q_{ac}| \leq 1000 \end{aligned} \tag{5.5}$$

5.1.2 BTMS Energy Consumption Model

In the concept design of the J-type BTMS, the cooling air is actuated by an air fan under normal conditions. When the ambient temperature is higher than the predefined threshold or encountering severe operating conditions, a heat exchanger will be activated to pre-cool the air, in which the coolant comes directly from the AC system, as shown in Fig. 5.1. The fan power is estimated by:

$$P_{btms} = \frac{\dot{m} P_c}{\rho \eta_c} \tag{5.6}$$

where P_c is the pressure augment of the compressor, ρ is the air density, and η_c is the flow efficiency. All the properties and parameters are obtained via CFD simulations, and the relationship is approximated using a support vector regression model, as shown in Fig. 5.3.

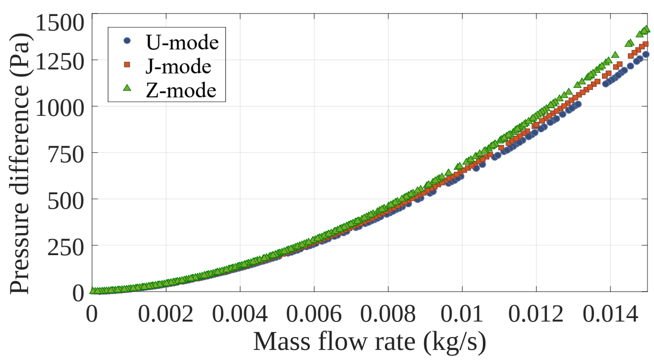


Figure 5.3: The relationship between the pressure augment and the mass flow flow rate for U-, J-, and Z-type structures

Besides the aforementioned systems, other major subsystems and devices include the power steering, the braking system, the lights, and the entertainments. Due to the complexities under dynamic conditions, it is challenging to establish a comprehensive dynamic model that consists of all these subsystems. In this paper, as a trade-off, only the driving motor with regenerative function, the air conditioner, and the thermal control system are considered and modeled in detail, while other auxiliary devices are assigned with an estimated fluctuated power that follows a normal distribution as follows.

$$P_{aux} \sim \mathcal{N}(\mu, \alpha^2) = \mathcal{N}(1000, 250) \tag{5.7}$$

5.1.3 Energy Management Strategy

Based on the driven power model, the AC model, and the BTMS model, this study seeks to develop an energy management strategy with high efficiency. The energy management strategy aims to enhance the energy efficiency as well as the battery health performance via an MPC algorithm, while retaining the constraints from the perspectives of thermal limitations and electrical requirements. The considerations and approaches are predefined in a twofold manner: (i) to avoid potential overlapping peaks by rescheduling the operation of different devices; (ii) to mitigate the negative effects regarding the battery cycle-life in recharging the battery by distributing the regenerative energy to auxiliary systems. By adopting an MPC algorithm, the thermal control and centralized optimization framework is formulated as:

$$\begin{aligned} \underset{Q_{ac},\dot{m}}{\arg\min} \quad & J = \sum_{k=n}^{n+N} (\alpha_k (P_{drv_k} + P_{aux_k} + P_{ac_k} + P_{bs_k})^2 \\ & \quad + \beta_k (T_{ref_k} - T_{b_k})^2 + \xi_k (T_{tar_k} - T_{in_k})^2) \\ \text{subject to} \quad & 0 \leq \dot{m} \leq 0.012 \\ & |\Delta \dot{m}| \leq 0.003 \\ & 0 \leq Q_{ac} \leq 4500 \eta_{cop} \\ & |\Delta Q_{ac}| \leq 1000 \end{aligned} \tag{5.8}$$

where n is the current step, and N is the control horizon. Based on specific control purposes, penalty coefficients α_k, β_k, and ξ_k are preset to attribute weights of the overall power consumption, the control target temperatures of the cabin and the battery, respectively. The soft constraints of the two subsystems are implemented via a real-time adjustment of the corresponding coefficient, i.e., the temperature biases should be constrained within 0.5 K and 1 K for BTMS and AC, respectively. Since the second term about the thermal control of the battery system involves safety concerns, a larger weighting, β_k, is assigned to BTMS compared with the AC system.

A particle swarm optimization (PSO) algorithm is employed to solve the problem, in which PSO starts with the original direct control solutions of the two subsystem. It is worth mentioning that the stochastic PSO approach is promising to obtain a reasonable solution instead of a global optimum within limited calculation time by tuning different convergence criteria.

5.2 Case Studies and Results

An integrated driving cycle that consists of the EPA urban dynamometer driving schedule (UDDS), the world-harmonized light-duty vehicles test cycle (WLTC), and the highway fuel economy driving schedule (HWFET) is directly utilized to test and validate the energy management strategy. Combined with the power consumption from auxiliary devices, the uncontrollable power consumptions, i.e., as the management system inputs, are presented in Fig. 5.4. The sample time is set to be 5 s.

5.2.1 No Energy Management Strategy

A comparative case without any energy management strategy is conducted as the benchmark, in which the BTMS and AC systems are operated separately based on their own conditions without a global energy management strategy. For BTMS, the power consumption of the last step is taken into consideration to calculate the heat generation rate for the current step, and thus determine the mass flow rate, as shown in Fig. 5.5. The temperature distribution of the battery pack along the dynamic process is presented in Fig. 5.6. It is seen that the temperatures follow close behind the targeted reference trajectory via switching among the J-, U-, and Z-mode. Due to the flow characteristic of Z-mode, the battery cells near the right outlets have a stronger heat dissipation capability under large flow rate conditions.

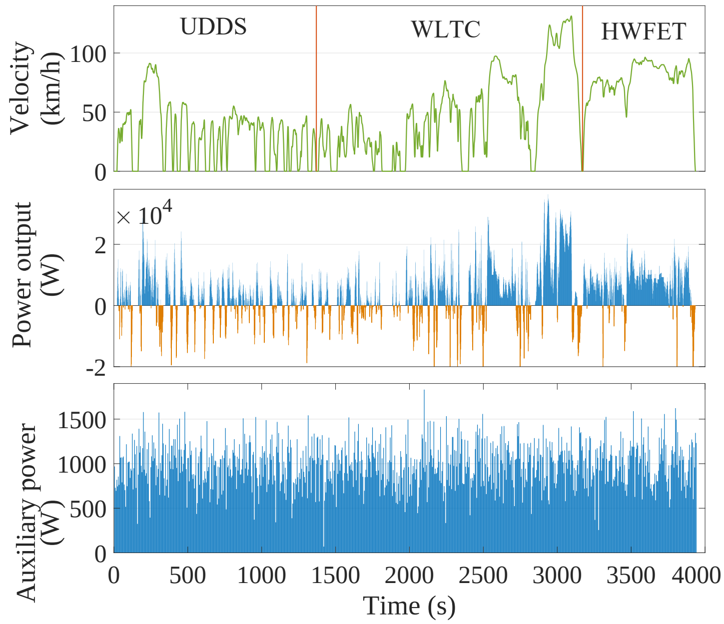


Figure 5.4: The power consumptions of tested driving cycles and auxiliary devices

The temperature uniformity deteriorates after 2,100 s, while the maximum temperature in the 7^{th} battery cell also decreases. At around 2,950 s, the pre-cooling module activates to provide extra cooling capacity to cope with the large amount of heat generation. Overall, the BTMS is capable to retain the temperature within an expected range.

For the AC, though the temperature of the internal air goes down in a fast manner, the base temperature decreases very slowly because of limited convection heat transfer, as shown in Fig. 5.7. The velocity variations may inevitably bring about fluctuations to the cabin temperature as well as the power usage. Only a base load is required to balance the sensible heat from solar radiation and the human body under normal operations.

5.2.2 Model Predictive Control-Based Energy Management Strategy

Based on the MPC algorithm, the forecast horizon is set to be 5 steps or 25 seconds. Aiming to reduce the overall power usages, the power consumptions of BTMS and AC are scheduled flexibly according to the real-time driving and auxiliary power usage. Both the BTMS and

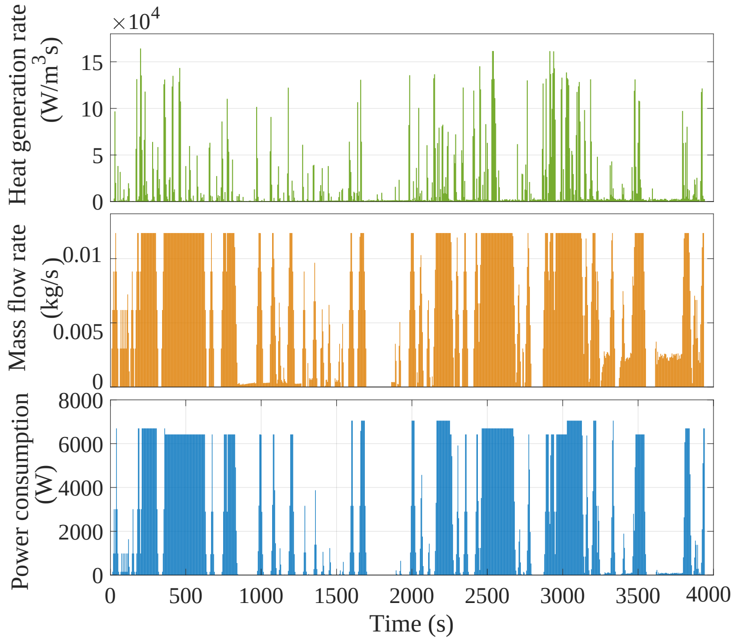


Figure 5.5: The BTMS properties without energy management

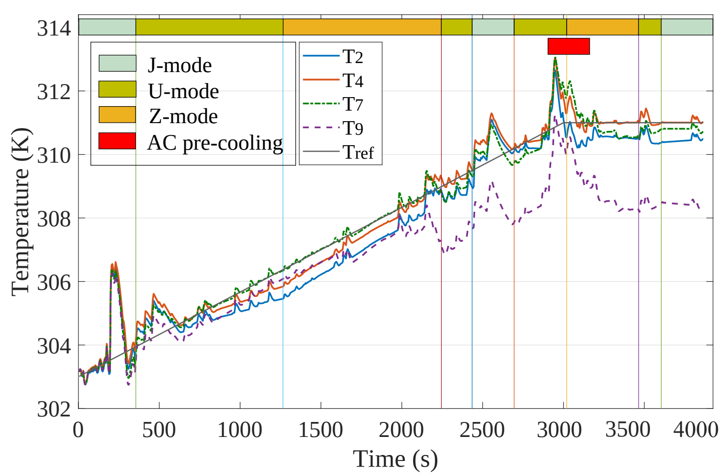


Figure 5.6: The BTMS temperature distribution without energy management

AC have similar and accepted performances compared to that of no energy management, as presented in Figs. 5.8 and 5.9. However, it is observed that the power consumption of BTMS is reduced for the reason that the overall heat generation is reduced by lowering the peak load, which is implemented by shifting the operation with AC, as shown in Fig. 5.10.

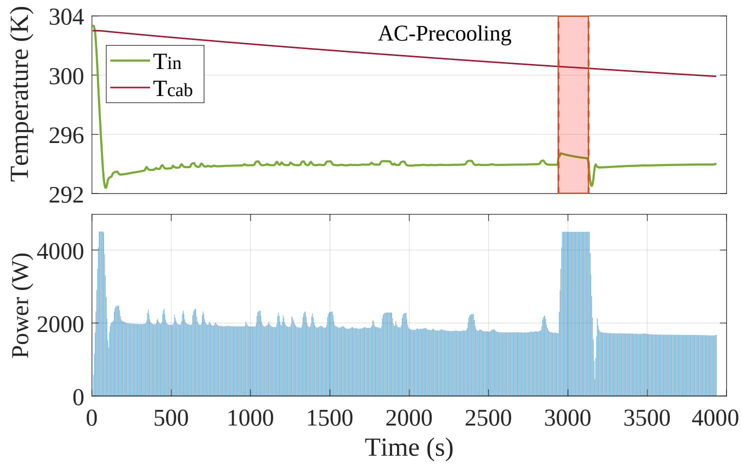


Figure 5.7: The AC performance without energy management

For instance, the operation of AC almost completely terminates at around 200-300 s and 2,600-2,650 s to avoid overlapping with existing load peaks.

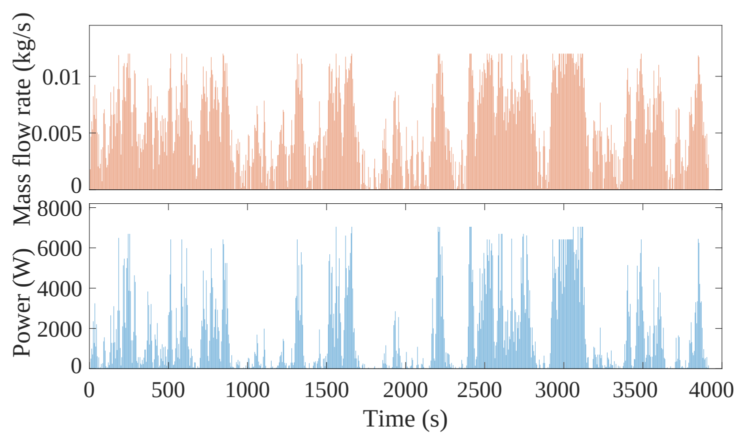


Figure 5.8: The BTMS properties with MPC-based energy management

It is also seen that the regeneration power that recharges back to the battery system is reduced by 4.3% from 8,809 J to 8,430 J per battery cell, at the cost of bringing about extra fluctuations in the temperature controls of BTMS and AC system. The total energy consumption also has a 6.5% improvement from MPC with a value of 23,800 J compared to 25,500 J without any management strategy per battery. Moreover, the final stage SoC with MPC is 0.5864 compared to 0.5653 for the system without control, which has a 3.8%

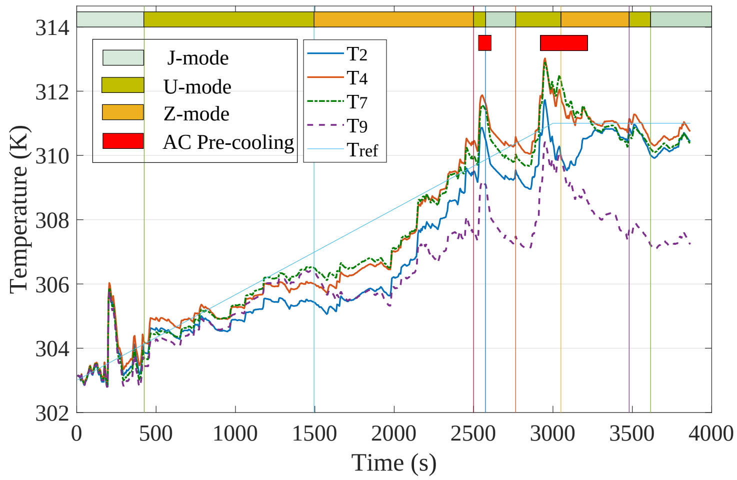


Figure 5.9: The battery temperature distribution with MPC-based energy management

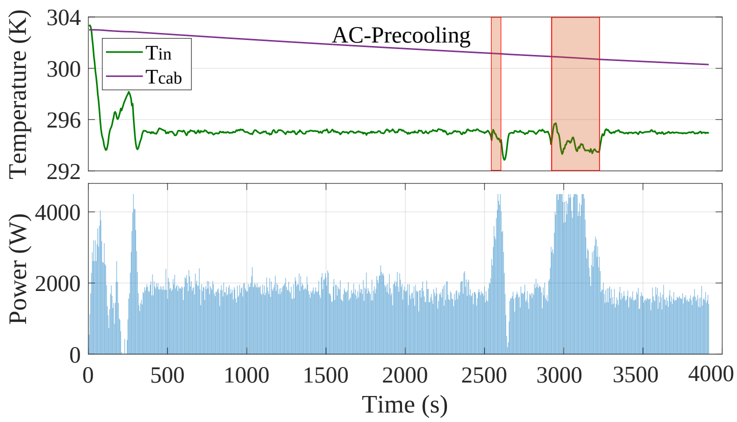


Figure 5.10: The AC performance with MPC-based energy management

improvement. Note that the developed algorithm and control framework are also applicable to liquid-based battery cooling system, since they share similarities in terms of fluidity and controllability.

5.3 Summary

This chapter developed an MPC-based energy management strategy to control the electric vehicle cabin climate system and battery thermal management system, simultaneously. A battery thermal control model was developed using neural network, while the cabin air conditioning system was established with a proportional-integral control method. The energy management strategy aims to reduce the values of load peaks, while retaining the constraints of the BTMS and AC systems.

The MPC-based energy management strategy was tested using an integrated driving cycle. Compared to the system with no energy management, no significant differences were observed in terms of thermal properties and dynamic balance. However, from the perspective of energy efficiency, simulation results revealed that there were a 4.3% reduction for the recharging energy, and a 6.5% improvement for the overall energy consumption. It is shown that the MPC-based energy management is a promising solution to enhance the overall efficiency of EVs.

CHAPTER 6

ENERGY MANAGEMENT-ORIENTED SHORT-TERM VEHICLE VELOCITY FORECASTING[1]

To further improve the performance of individual vehicle velocity forecasting, a hybrid velocity forecasting algorithm is developed in this chapter, by leveraging the fact that most of the newly-released electric or hybrid vehicles are equipped with on-board GPS devices. To validate the proposed algorithm, a repeated urban driving cycle dataset is first generated by the same driver (with the same driving habits) in the area of Dallas, Texas. The driving patterns between weekdays and weekends are investigated, and road segments are also identified. A forecasting pool that consists of multiple base forecasting algorithms is established, and a localized model selection and ensemble framework is developed to dynamically choose the appropriate forecasting models for each road segment. This chapter seeks to enhance the forecasting accuracy with the currently available urban transportation infrastructure, and the contributions are threefold: i) generate a publicly available commuting dataset with repeated driving cycles for energy management and control Co-design; ii) develop a segment-based vehicle speed forecasting model, in conjunction with a localized model selection and ensemble framework.

6.1 Data Collection and Analyses

6.1.1 Data Collection

There exist several repeated driving cycles based on a fixed route published in the literature, such as the Connected Ann Arbor (A2) dataset [47], the Gothenburg driving dataset [30], the CSU driving cycle [70], the MTU driving cycle [8], and the Fort Collins repeated

[1]©2021 ASME. Reprinted, with permission, from Y. Liu and J. Zhang (2021), A Repeated Commuting Driving Cycle Dataset with Application to Short-term Vehicle Velocity Forecasting, *Journal of Autonomous Vehicles and Systems*, 1-44

driving cycle dataset [20]. However, few datasets are publicly available at the current stage. Another piratical way to obtain repeated driving cycles is to extract the repeated routes from large-scale traffic or vehicle energy consumption dataset that covers a certain period of time, such as the Geolife Trajectories[1], the vehicle energy dataset (VED) [64][2], the performance measurement system (PeMS) dataset[3], the Roma taxi dataset[4], and the San Francisco bay area taxi dataset[5]. These datasets are readily available online, but one of the prominent drawbacks is that the extracted cycles are usually inconsecutive and scattered with a very short period of time, making it challenging to be integrated with any predictive control implementation. Moreover, a large dataset with repeated routes is also beneficial to route planning and decision-making for autonomous driving development. By capturing the surrounding environment and analyzing dynamic traffic conditions via image recognition, the autonomous driving system with improved confidence can be potentially advanced to a higher level regarding reliability and safety for a specific repeated route in a limited area.[6]

To better understand the impacts of vehicle speed forecasting on energy management, we have generated a Dallas repeated driving cycle (DRD) dataset[6]. In this dataset, dozens of driving cycle tests have been performed on a fixed route in the Dallas area to simulate a typical commuting route for passenger vehicles, which consists of an expressway test of 5 kilometers and a local urban road test of 20 kilometers, as shown in Fig. 6.1. The dataset was acquired between December 2020 to January 2021 at around 4:45 PM to 6:00 PM, in which each cycle takes approximately 30 minutes using a conventional internal combustion

[1]https://www.microsoft.com/en-us/download/details.aspx?id=52367

[2]https://github.com/gsoh/VED/blob/master/README.md

[3]https://archive.ics.uci.edu/ml/datasets/PEMS-SF

[4]https://crawdad.org/roma/taxi/20140717/

[5]https://crawdad.org/epfl/mobility/20090224/cab/

[6]Available at: https://github.com/UTD-DOES/Dallas-Repeated-Driving-Cycle-Dataset

engine passenger vehicle. All the testing cycles were conducted with a fixed route by the same driver with similar driving manners. It needs to be noted that some of the traffic signals in this area are adaptively controlled in a daily or continuous real-time manner, according to the progress report of a regional traffic retiming program [63]. More repeated commuting cycles by different drivers are expected to be included in the DRD dataset in the coming future.

Several critical dynamic driving parameters and indices are recorded using a GPS logger (brand: Garmin, model: eTrex 10) and a camera. All the data are time-stamped with a time interval of 1 second for the sake of accuracy and consistency, including the vehicle velocity, altitude and longitude information, elevation, heading direction, and traffic light picture/video. Compared with other datasets, one of the remarkable merits of this DRD dataset is that it covers a broader set of road types with tens of intersections, and traffic light images can be used as a potential tool for intersection waiting time prediction.

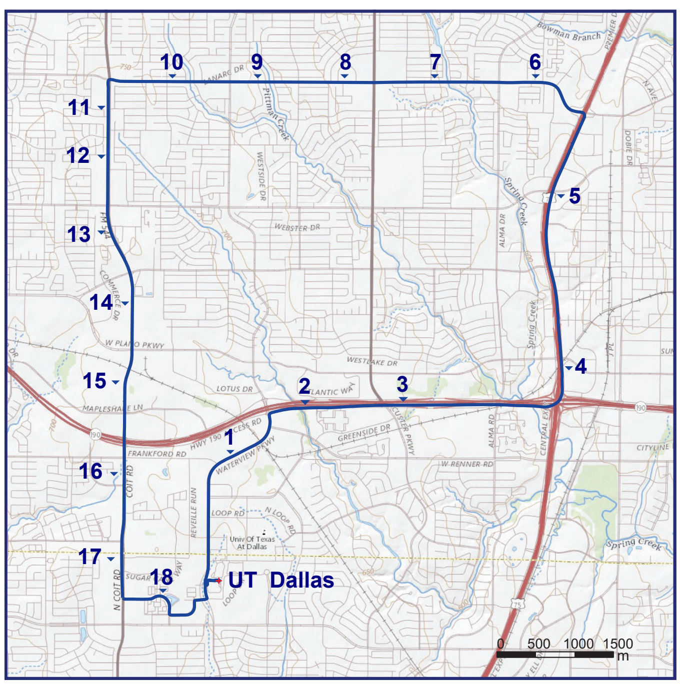


Figure 6.1: The testing route of DRD repeated driving cycles, which is close to the University of Texas at Dallas (The numbers indicate the positions for traffic congestion analyses.)

6.1.2 Data Processing

As a preprocessing step, clustering plays an important role in improving forecasting accuracy. For velocity forecasting using historical data, there are two general approaches to cluster the data and identify major spatial-temporal patterns: i) cluster the cycles for pattern recognition, split the grouped cycles into segments, perform another round of clustering for the segments, and then establish the forecasting models; ii) skip the first-round clustering, and follow the rest steps [15, 38]. Note that unsupervised clustering can be transformed into supervised classification by manually labeling the data with expert knowledge, e.g., the cycles can be classified directly into weekday and weekend/holiday conditions based on daily driving experiences.

To quantify the traffic discrepancies between weekdays and weekends, a congestion index ε_{it} (of a location i at time t) is modified here by comparing the measured floating vehicle velocity with the free-flow velocity [62], as defined in Eq. 6.1.

$$\varepsilon_{it} = \frac{s_{it}}{v_{it}} - 1 \tag{6.1}$$

where the parameter S_{it} represents the free-flow velocity defined by the maximum value recorded, while v_{it} represents the current vehicle velocity. A larger index ε_{it} indicates relatively more severe traffic congestion. Here, a total of 18 locations away from the intersections are selected randomly using the stratified sampling algorithm, as shown in Fig. 6.1. The discrepancies between weekdays and weekends/holidays are illustrated in Fig. 6.2. It is seen that there do not exist many significant differences between weekdays and weekends/holidays, except for the saturated section ranging from location 2 to 5 on the upper-right corner, which illustrates an improvement of the highway traffic on weekends and holidays. The possible reason behind this phenomenon is the current remote-work environment due to the COVID-19 pandemic. It is also noticed that the whole trip takes an averaged time of 1,801 seconds on weekends and holidays, which is 193 seconds shorter on weekdays. By investigating the details of route segments, it is observed that moving through the intersections on weekdays requires extra time, as shown in Fig. 6.3.

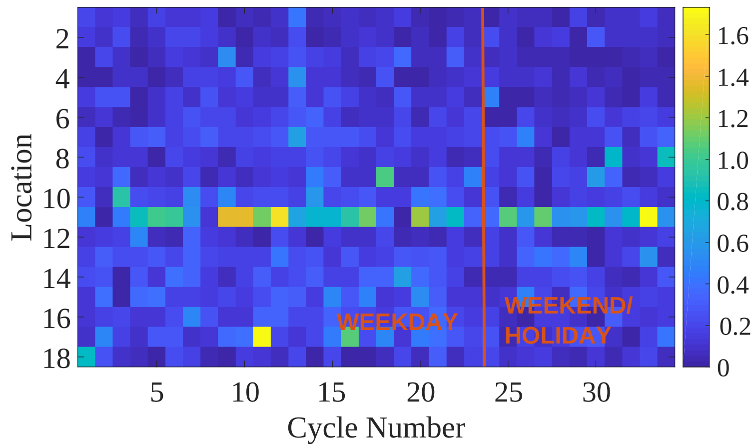


Figure 6.2: A comparison of the congestion index between weekdays and weekends/holidays

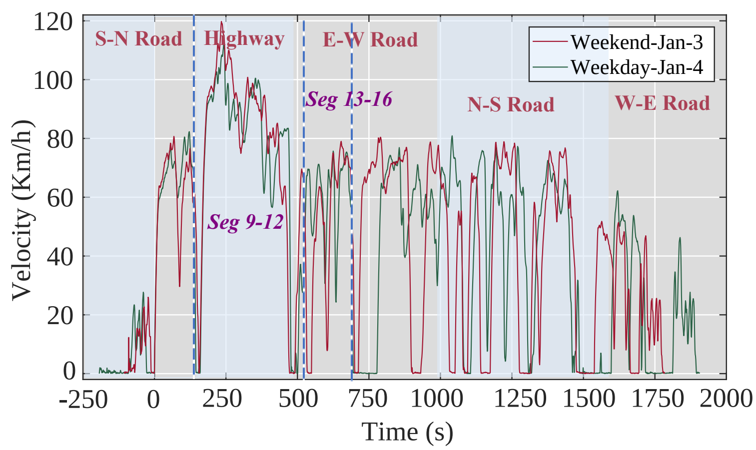


Figure 6.3: The velocity trajectories of typical driving cycles on weekdays and weekends/holidays (The location annotations are based on the weekday cycle.)

6.1.3 Intersection/stop Identification

Given the preceding analysis, the second clustering approach discussed in Section 6.1.2 (i.e., split the grouped cycles into segments, perform another round of clustering for the segments, and then establish the forecasting models) is employed here for feature identification in the study. The driving cycles are directly divided into segments, followed by a time sequence clustering of the segments if necessary. Specifically, all the intersections are extracted from the route as separated segments because of their significant impacts on the whole driving time. A location is identified as an intersection or a T-junction with stop/yield signs if it is detected with a complete stop or low velocity (i.e., 10 km/h) more than twice. Once the intersections are located, the routes in between will also be defined as independent cycle segments. It is worth mentioning that there are two major reasons why traffic signal identification is necessary rather than using the labeled data from public map sources directly: i) the vehicles have a high probability of moving through some labeled intersections or T-junctions without any interruption; ii) the vehicles need to wait for two rounds at some traffic lights due to the heavy traffic conditions, resulting in another indirect hidden stop.

As can be seen from Fig. 6.4, the final location of a vehicle is scattered at an intersection, depending on the traffic volume and its arrival time. The furthest point downstream among all is treated as the location of the intersection. The primary step here is to organize the stop points that belong to the same intersection or potential stop into a group. General density or centroid-based machine learning-based clustering algorithms like K-means have been tested however with unsatisfied performance, since it is challenging to determine the appropriate number of clusters. In this study, we attempt to cluster the stop points using a connectivity-based single linkage hierarchical clustering algorithm, in which only connections with a coordinate distance smaller than the maximum threshold will be considered to form a same group, i.e., 40 m for this DRD dataset. Given the specific velocity profile, the single linkage approach has a great superiority regarding the accuracy over other clustering methods for a relatively small dataset, as it determines a group in a more straight-forward manner by merely using a hard distance threshold. The clustering results are also highlighted in Fig. 6.4. A group consisting of fewer than three points is ignored and will not be treated as an intersection or a stop, given a low stopping probability of 5.8% (2/34).

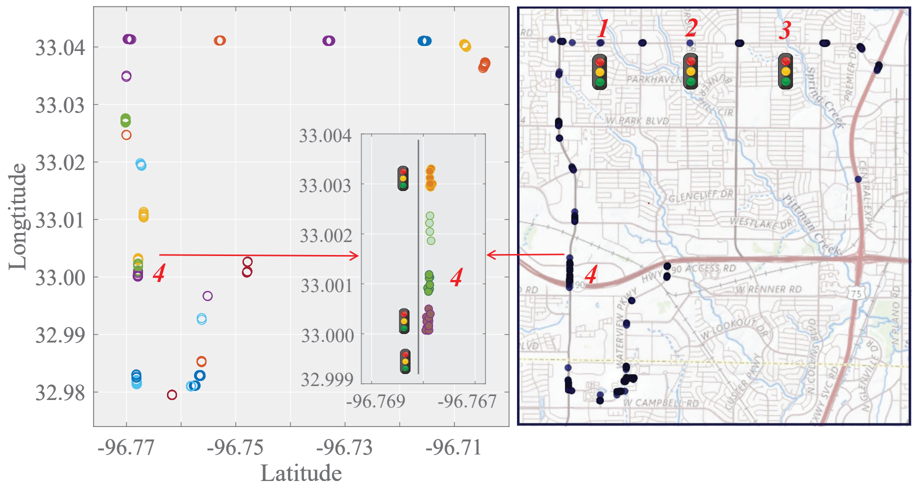


Figure 6.4: Intersection detection using the velocity profile and its location on a map. The left figure shows the locations of intersections after identification. The onward route before the first stop sign has been removed, so has the return route. Annotations 1 and 2 illustrate that these two intersections are equipped with traffic lights but will be ignored as a normal straight road due to a low stopping probability. Annotation 3 shows no stops have been ever detected even though there is a traffic light. Annotation 4 indicates that vehicles are very likely to stop moving somewhere between two intersections because of heavy traffic, which may be regarded as a stop.

However, the computational complexity may increase considerably when applying the method to a large repeated driving cycle dataset. Given its temporal and spatial nature, a semi-supervised iterative approach is employed aiming to improve the computation efficiency. The clustering process begins with a single driving cycle, which is then clustered using a maximum distance threshold as discussed above. By adding new cycles, the new clustering outcomes get updated iteratively using a divide and conquer method. The proposed method seeks to avoid unnecessary coordinate distance calculations, as illustrated in Fig. 6.5.

6.1.4 Road Segment

To better model the velocity trajectory passing through an intersection, besides the final stopping location, it is also crucial to identify the deceleration and reacceleration processes

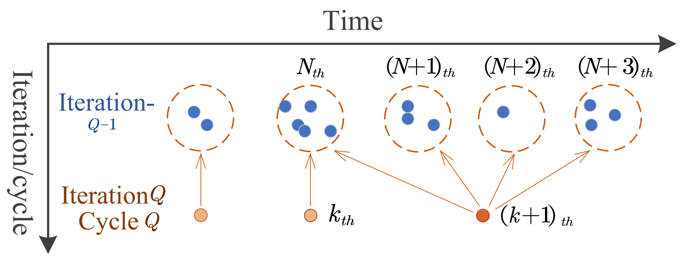


Figure 6.5: The sketch of an iterative clustering method for a large dataset. (Detailed processes: after clustering the k_{th} point in *Cycle–Q* into the existing N_{th} group, the $(k{+}1)_{th}$ point only needs to calculate its distance with groups starting from N_{th}. Once it is clustered, i.e., into the $(N+i)_{th}$ group, an additional distance with the $(N+i+1)_{th}$ group needs to be conducted to confirm that the new point does not belong to the next group, otherwise, the $(N+i)_{th}$ group will be combined with the $(N+i+1)_{th}$ group to form another group.)

and divide the routes into varying segments, as illustrated in Fig. 6.6. An intersection segment consists of a deceleration, a waiting, and a reacceleration process, while a normal road segment refers to a continuous move at a steady speed. When it comes to the local street with a lower speed limit and a smaller traffic volume, i.e., a stop sign, the deceleration and acceleration prepossess are captured with a very similar pattern, i.e., same stopping locations and almost equal waiting time.

Following the principle discussed above, the whole cycle is divided into 42 segments using location coordinates, and a portion of the segments are presented in Figs. 6.7 and 6.8. It is seen from the figures that within a same road segment, the velocity patterns may still differ, especially for intersection segments. However, it is observed from Fig. 6.8 that the trends of velocity trajectories versus location for intersection segments are more uniformed than that versus time in Fig. 6.7. The possible discrepancies are mainly due to the waiting time for traffic lights. Thus, no further clustering is implemented in this study. We also assume that the final stopping location may have considerable influences on the waiting time at an intersection, which is crucial to vehicle energy management and will be further investigated in the following section.

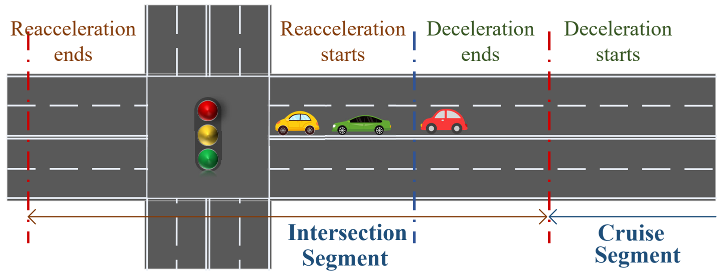


Figure 6.6: The schematic diagram for road segment division

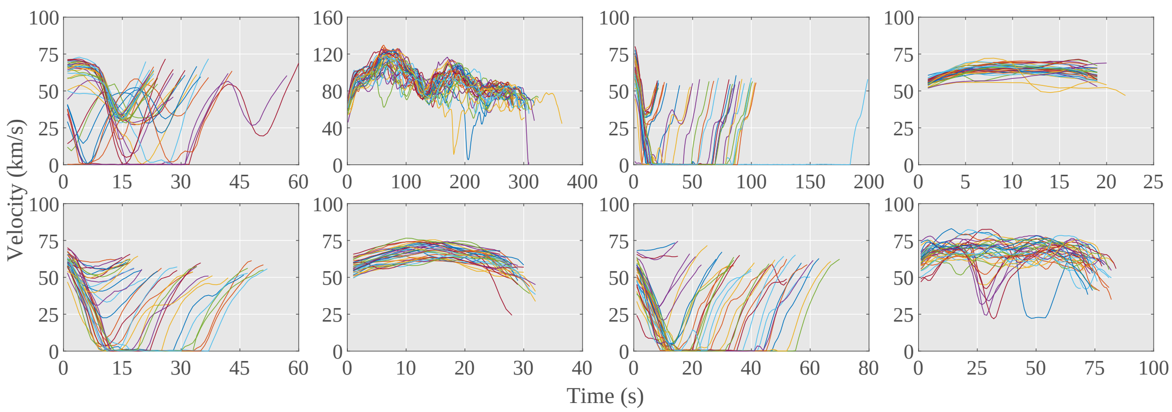


Figure 6.7: Segment velocity trajectory *vs.* time (top: Segments 9-12; bottom, Segments 13-16 (see Fig. 6.3), cycle size: 34)

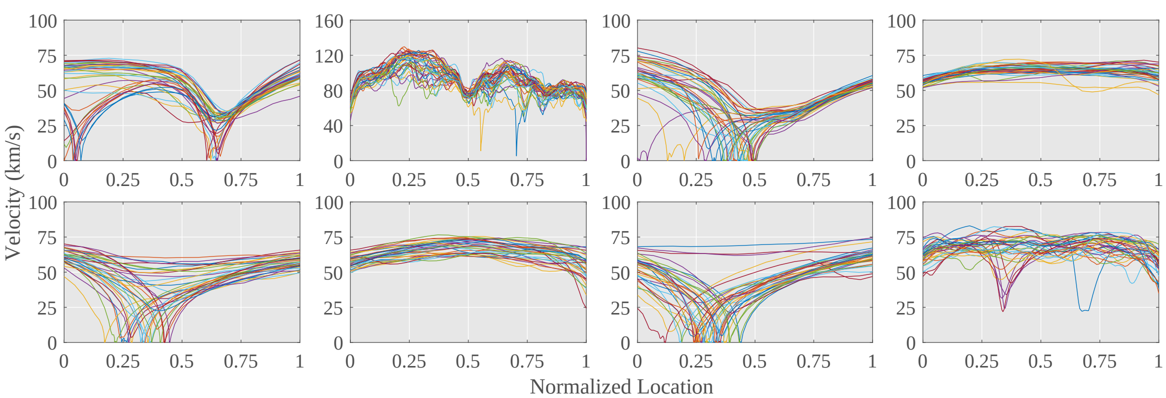


Figure 6.8: Segment velocity trajectory *vs.* location (top: Segments 9-12, bottom: Segments 13-16 (see Fig. 6.3), cycle size: 34)

6.2 Base Forecasting Methods and Results

From the perspective of energy management with a specific driving route, it is expected that accurate traffic forecasting, including the averaged velocity, the deceleration and acceleration processes could significantly enhance the efficiency via energy scheduling and planning. In this study, we construct a forecasting model pool that consists of a collection of stochastic and deterministic methods based on their popularity and performance, i.e., LSTM, HMM, ANN, SVR, and a similarity-based method, with varying kernels, training algorithms, and hyper parameters. However, due to a very short prediction and action time interval (e.g., 5 seconds for most of the predictive energy management implementations [50, 5, 4, 65]), it remains challenging to include all the submodels simultaneously for short-term velocity forecasting. The best subset of models are preselected in the training-validation stage using two evaluation metrics, i.e., the mean absolute error (MAE) and the root-mean-square error (RMSE), expressed as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{6.2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{6.3}$$

where $\hat{y}_i$ and y_i are the forecasted and actual value of sample index i, respectively.

Via the model preselection, only five models are employed to forecast the velocity for each road segment. In this process, a total of 23 randomly-selected driving cycles (i.e., renamed as Cycle 1-23) are utilized for training, and another 5 cycles (i.e., Cycle 24-28) are used for validation.

6.2.1 Stochastic Approach

Hidden Markov chain is a discrete-time stochastic memoryless process to model a sequence of events, in which the future state or action only depends on the current state. The chaining process is characterized by a set of implicit hidden states S and its transition probabilities matrix A, and each a_{ij} represents the probability of moving from a state i to a new state j, s.t. $\sum_{j=1}^{N} a_{ij} = 1$, $\forall i$, which is also referred to as the Markov assumption, expressed as:

$$P(s_i|s_{i-1}, s_{i-2}, \cdot, \cdot, \cdot, s_1) = P(s_i|s_{i-1}) \tag{6.4}$$

Another fundamental assumption of HMM is that the explicit observation O only relies on the state that generates the observation with a probability of $B = [b_j(O_t)]$. Starting with an initial probability distribution over states π, an HMM process can be modeled as $\lambda = (A, B, \pi)$.

To apply HMM for velocity prediction, it is required to transform the historical speed records into a set of observations indexed by integers. If the numbers of states and observations are known and set to be equal, the probability matrix A and B can be solved by calculating the frequency counts of a labeled state transition among all the transitions or a specific observation among the observations. In this study, the DRD dataset contains a very limited number of data points, posing challenges in directly solving this problem, as the matrix might be singular.

To enhance the accuracy, two data augmentation techniques for time series are leveraged to expand the existing dataset, including the weighted dynamic time warping barycenter averaging (w-DBA) and Gaussian noise injection. The w-DBA is accomplished via assigning varying weight factors to a set of similar temporal sequences measured using the dynamic time warping (DTW) [71], as given by:

$$T_{new}(n) = \omega_b T_{base}(n) + \sum_{i=1}^{4} \omega_i \mathbb{D}[T_{base}(n), T_i(\overline{m})] \tag{6.5}$$

where T_{base} and T_i are the basic sequence to be augmented by its selected i_{th} sequence with similarity. $[\omega_b, \omega_i]$ denotes the assigned weight factor consistently with the DTW normalized distances, s.t. $\sum \omega_j = 1$. Here, it is predefined as [0.4, 0.18, 0.16, 0.14, 0.12] through trials for simplification. $\mathbb{D}[\cdot]$ represents the DTW alignments between two time sequences, i.e., the n_{th} element of the base sequence may align with the m_{th} element or several continuous m_{th}-p_{th} elements of a similar sequence, where the value of $\mathbb{D}[\cdot]$ equals to $T_i(m)$ or the average of sequence $T_i(m) - T_i(p)$. The main motivation of adopting the w-DBA approach is that the new augmented sequence keeps the same length as the original base sequence but with reasonable variations.

Gaussian noise injection is another popular and alternative augmentation approach for time sequences, defined as:

$$
\begin{aligned}
T_{new} &= T_{base} + \mathcal{N}(\mu, \sigma^2) \\
s.t. \mid \sum_{i=1}^{N} &\mathcal{N}(\mu, \sigma^2) \mid \leq 0.02\mathcal{L}
\end{aligned}
\tag{6.6}
$$

where $\mathcal{N}(\mu, \sigma^2)$ represents a zero-mean normal distribution with a variance of 0.8 ($\mu = 0$, $\sigma^2 = 0.8$), and $\mathcal{L}$ denotes the total length of the sequence. The motivation why a variance of 0.8 is chosen is to generate a series of noises, 98% of which lie on an appropriate range of $[-2, 2]$. The constraint here is to confine the total length of augmented noises within an acceptable range, making the total distance remains unchanged for each segment after data augmentation. Figure 6.9 illustrates the comparison between the original sequence and the augmented sequences. It is observed that adding the Gaussian noise may bring in a larger fluctuation than the w-DBA method, but overall, the augmented series stays in step with the general trend. Through these processes, the training dataset has tripled to 69 sequences compared to the original dataset.

Regarding the HMM training, we obtain a model by employing the Baum-Welch algorithm that treats the hidden states as implicit variables using an expectation maximization

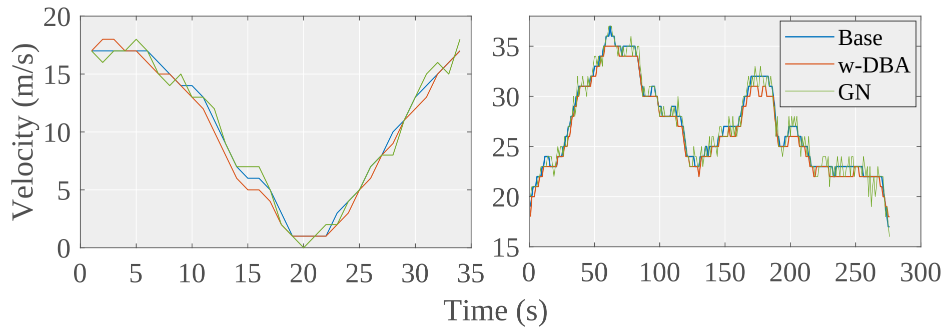


Figure 6.9: A comparison among the base, the w-DBA augmented, and the Gaussian noise (GN) augmented sequences (left: Segment 9, right: Segment 10)

algorithm, and the detailed derivation and explanation can be found in Ref. [10]. After the model is trained, the prediction procedure uses a sequence of velocity as the observation input, and a dynamic programming (i.e., Viterbi) algorithm as the solver to acquire the most probable state path. The prediction results are achieved one step ahead along the Markov chain within the variant constraints.

The forecasting task in this study is to predict the vehicle velocity 10 seconds ahead, where approaches with different step windows can be used, e.g., 5 steps ahead with a 2-second interval or 2 steps ahead with a 5-second interval. The reason to adopt these recursive multi-step approaches is that most of the decelerations and accelerations in the DRD dataset occur within 12 seconds, and it is challenging for a one-step 10-second ahead direct prediction to capture these processes using HMM.

Table 6.1 tabulates the validation performances across the original and the augmented datasets. The recursive 2-step approach achieves an approximately 3% higher accuracy than the 5-step method using these two datasets. The accuracy discrepancies are mainly due to error accumulations, where a double recursion performs better. Moreover, considerable enhancements are also observed after leveraging the data augmentation approaches. However, there still exist drastic fluctuations in the HMM forecasting as shown in Fig. 6.10, especially during the deceleration and stop states, which may lead to undesired disturbances when applying to the energy management system. The possible reasons for causing these drastic fluctuations and discrepancies are threefold: (i) the transformation of original data into discrete positive integers may potentially lead to non-ignorable accuracy loss; (ii) both the transition and observation matrices are sparse in nature, which can be mitigated by accumulating a much larger number of driving cycles as the implementation of data augmentation indicates; (iii) for the 2-step method, only 3 previous observations/velocities (i.e., [O_{t-10}, O_{t-5}, O_t]) are used as the model input. Additionally, the single HMM model built on the whole cycle is investigated, yielding an undesirable outcome with an MAE of 3.49 m/s and an RMSE of 5.38 m/s based on the original dataset, which is approximately 30% worse than the segment-based approach. The results also further validate the necessities and effectiveness of the segment analysis discussed in Section 6.1. It is worth noting that there are multiple ways to implement HMM in velocity forecasting, e.g., treating the acceleration/deceleration speed as the observable states, real-time updated transition matrix, higher-order Markov chain model, and subdivided HMM models for varying driving conditions [109, 82, 47]. However, due to the huge diversity of driving datasets, no comparative study has been reported yet regarding the performance of these different approaches in the literature.

Table 6.1: Segment-based HMM accuracy analysis

	Original Dataset		Augmented Dataset	
HMM Model	5-step	2-step	5-step	2-step
MAE (m/s)	2.68	2.55	2.61	**2.47**
RMSE (m/s)	4.39	4.22	4.30	**4.07**

6.2.2 Deterministic Approaches

Long short-term memory model (LSTM) is an improvement over the recurrent neural network (RNN) with feedback connections designed to address the long-term dependency challenge when modeling sequential events by propagating through time. In addition to the

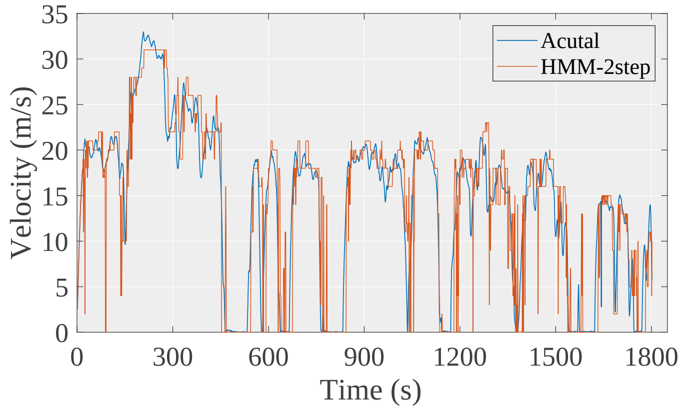


Figure 6.10: 2 steps ahead HMM forecasting based on a 5-second interval for Cycle-28

existing structures of RNN like hidden states, LSTM employs a novel layer, named the cell states, to selectively store the previous event information, making it capable of alleviating the vanishing/exploding gradient issue. In this study, a typical LSTM structure with a forget gate is employed to forecast velocity 10 steps/seconds ahead, which takes approximately 300 epochs to converge. Given the sequence length of the dataset, the LSTM model adopts a deep learning structure with one hidden layer after random search. The number of hidden units is narrowed down within the range of 180 to 260, varying from segment to segment. To avoid potential overfitting, the model also employs the settings of bidirectional layer and dropout layer with a dropout rate of 0.5. Three popular training algorithms are considered here, i.e., stochastic gradient descent (SGD), adaptive moment estimation (Adam), and root mean square propagation (RMSprop), which provides important properties of spatial and temporal locality.

Compared to LSTM, feed forward artificial neural network with back-propagation and SVR are other two popular machine learning-based estimation methods that mainly consider spatial properties. For temporal forecasting, we utilize the past 10 sequential steps/seconds record as the input to estimate the velocity 10 seconds ahead. We take into account three training algorithms for ANN, i.e., the standard Levenberg-Marquardt (LM), variable learning rate gradient descent (GDX), and resilient back-propagation (RP), in which a fully connected two-layer structure is empirically predefined with a maximum size of 30 neurons for each layer. Similar to the ANN settings, we diversify and examine the model by utilizing three different kernels for SVR based on their popularity, i.e., linear, polynomial, and Gaussian kernels. Other hyperparameters like the penalty weight factor and intensive parameter are optimized and determined based on the validation dataset.

In addition, given the temporal and spatial repetition nature of the DRD dataset as discussed in Section 6.1, a similarity-based estimation approach is proposed here by comparing the similarities. This method takes the previous multiple steps and the real-time position (i.e., the GPS longitude and latitude, and the elevation data as supplementary) as the inputs to retrieve the most similar historical sequences near the location. In contrast to the DTW algorithm, since all the sequences are extracted in the same length, the sequence similarity can be attained directly by calculating the Euclidean distance expressed as:

$$Dist = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \alpha_i (S_i - T_i)^2} \tag{6.7}$$

where S and T are the test sequence and the ranked historical sequences, respectively. n denotes the total length of the two sequences, and here, n is set to be 10. α is a set of unequal weighted factor assigned to different steps, where the latest steps have a larger weight. According to the similarity ranking, the most similar sequences are selected, and the new forecast can be integrated as:

$$T_{fore}(t+10) = \sum_{i=1}^{3} \beta_i T_i(t+10) \tag{6.8}$$

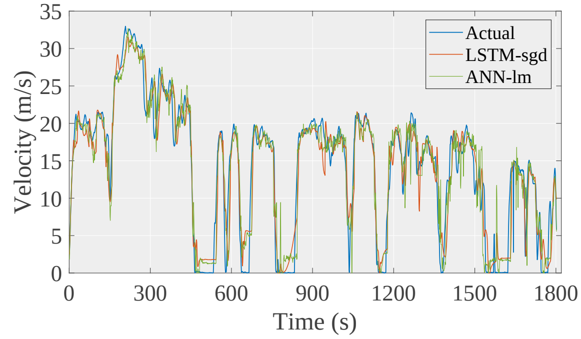
where β is an equal weight factor for generalization, s.t. $\sum \beta_j = 1$. The selected cycle number is narrowed down to 3 with the best performance.

The global performance using the original dataset is compared and tabulated in Table 6.2. Prior to outcome analyses, it is worth mentioning that RMSE is more sensitive than

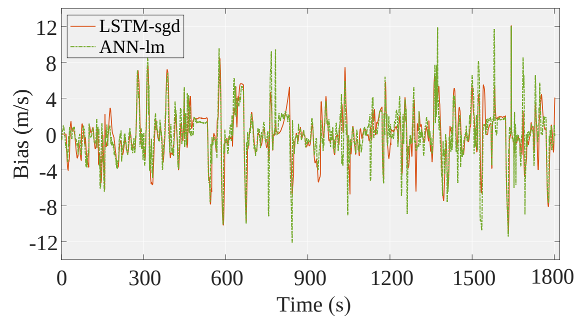
MAE to large bias or outliers due to the squaring operation. When it comes to this driving cycle dataset, the cruising segments tend to yield relatively smaller absolute differences, while larger biases/residuals normally occur at intersections. Accurate forecasting at intersections is more likely to produce a smaller RMSE. To evaluate the model performance, it is very necessary to report both metrics.

As a piece-wise approach, the segment-based approach dominates the whole cycle-based method for all machine learning models. It is also seen that the diversified submodels with different kernels or training algorithms vary in global accuracy. Given the limited processing time, only the LSTM-sgd, ANN-lm, SVR-Gaussian, and the similarity-based method are selected to perform velocity forecasting. It is worth noting that LSTM models have relatively lower accuracies comparing to the ANN approaches, especially in the piece-wise segment-based forecasting, while LSTM dominates other algorithms in similar studies [47, 70]. There are two possible reasons that may account for this discrepancy: i) The majority of the sequence lengths in this study merely range from 15 to 60 after segmentation, while the sequence lengths in Ref. [47] are around 700; ii) The LSTM models are pre-trained offline based on the historical records, which will not be updated during the cycle velocity forecasting given the on-board computational limitations. The detailed forecast results of the best submodels are shown in Fig. 6.11. Less fluctuations are observed using deterministic methods compared to the stochastic HMM method in Fig. 6.10.

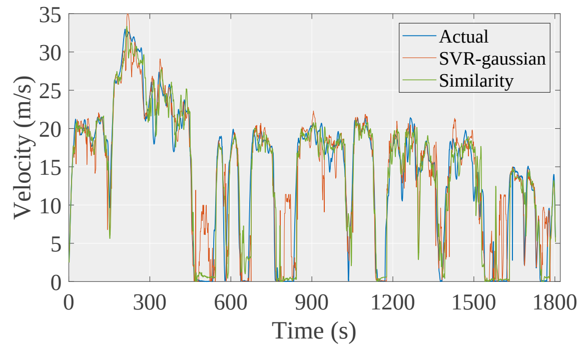
It needs to be noted that all the deterministic models are based on the original dataset, since no significant improvements are observed by using the augmented dataset. For example, the MAE of ANN-lm and SVR-Gaussian methods using augmented dataset are 2.19 m/s and 2.21 m/s, respectively, which are slightly worse than those using the original dataset. Considering the unique nature of the DRD dataset, one of the possible reasons is that the data augmentation techniques employed in this study are not forecasting-oriented [7].



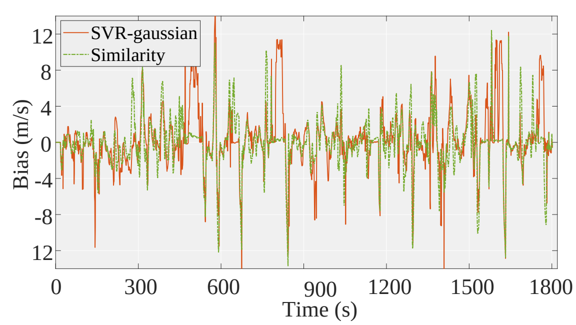
(a) Cycle-based forecasting results using LSTM and ANN



(b) Forecasting biases using LSTM and ANN



(c) Cycle-based forecasting results using SVR and similarity



(d) Forecasting biases using SVR and similarity

Figure 6.11: Forecasting results of Cycle-28 using deterministic methods

Table 6.2: Deterministic forecasting model accuracy

	Segment-based		Cycle-based	
Metrics (m/s)	MAE	RMSE	MAE	RMSE
LSTM-adam	2.36	4.01	2.79	3.89
LSTM-sgd	**2.26**	**3.23**	2.91	4.18
LSTM-rmsprop	2.29	3.74	2.78	3.83
ANN-lm	***2.15***	3.17	2.69	3.78
ANN-gdx	2.35	3.31	3.15	4.34
ANN-rp	2.19	***3.14***	2.79	3.89
SVR-linear	2.21	**3.38**	2.74	4.32
SVR-poly	2.49	3.73	3.14	4.87
SVR-Gaussian	**2.16**	3.42	2.65	3.91
Similarity	**2.25**	**3.71**	2.25	3.71

Note: **Bold values** indicate the best MAE or RMSE within each category; ***Bold italic values*** indicate the best MAE or RMSE among all models.

6.3 Localized Model Selection and Ensemble Approach

As illustrated in Figs. 6.10, 6.11, and 6.12, we have observed that the accuracies of the forecasting models differ from segment to segment. To further improve the forecast performance based on these individual models, ensemble models and dynamic model selections are two widely-used second-stage approaches. As a higher-level forecast algorithm, an ensemble model aggregates the predictions of certain diverse base models and results in a final output using averaged or weighted methods, while an online dynamic model selection could deploy an optimal individual model via methods such as reinforcement learning or Bayesian updating [57]. This section will comparatively investigate ensemble and online model selection methods, aiming to develop an enhanced forecasting model with a higher generalization and accuracy. Due to the lack of available data, we reuse the validation dataset (Cycle 24-28) together with a set of new data (Cycle 29-31) for the model training if needed, and the remaining portion of the DRD dataset (Cycle 32-34) is used for testing.

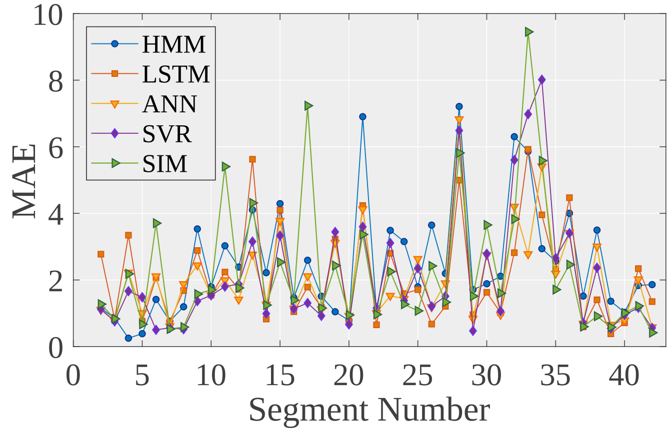


Figure 6.12: The segment-based model error distribution of Cycle-32

6.3.1 Single Model Selection

The segment-dependent probability approach is an offline individual model selection method that only considers the prior probabilities for different segments, and directly utilizes the base model with the maximum likelihood for prediction. One of the prominent advantages of this method is that no online training is required and the model can be updated offline. Only the selected base model is implemented for prediction, which saves computing time. The probability distribution calculated via frequency count is illustrated in Fig. 6.13, which results in an improved MAE of 1.99 m/s and a deteriorated RMSE of 3.29 m/s, compared to the best single base model (i.e., ANN) with an MAE of 2.07 m/s and an RMSE of 3.11 m/s for this DRD dataset. The results indicate that this offline model selection method may not be an appropriate solution for this problem.

In contrast to the offline probability-aided method, a persistence approach directly utilizes the dominating model of the last segment as the forecasting model for the current segment. A training process is unnecessary but it entails the implementations of all the submodels pre-selected in Section 6.2. As shown in Fig. 6.12, the accuracy ranking differs in both cycles

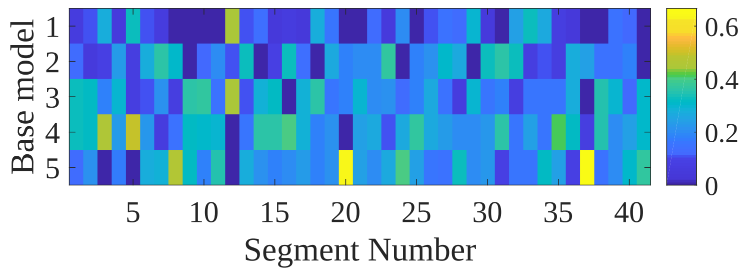


Figure 6.13: The normalized discrete prior probability distribution of base models for Cycle-32 . (The model numbers 1-5 indicate the HMM, LSTM, ANN, SVR, and SIM model, respectively. The models with an MAE difference threshold of 0.2 m/s are counted as the top models.)

and segments, the intersection segments in particular. However, the performance of a prior dominating model is not likely to change dramatically in its next step prediction. The base model ranking and the dynamic model selection are illustrated in Fig. 6.14, which successfully chooses the top two models with a ratio of 59% (23/39). Considerable improvements are observed with an averaged MAE of 2.02 m/s and an RMSE of 3.07 m/s.

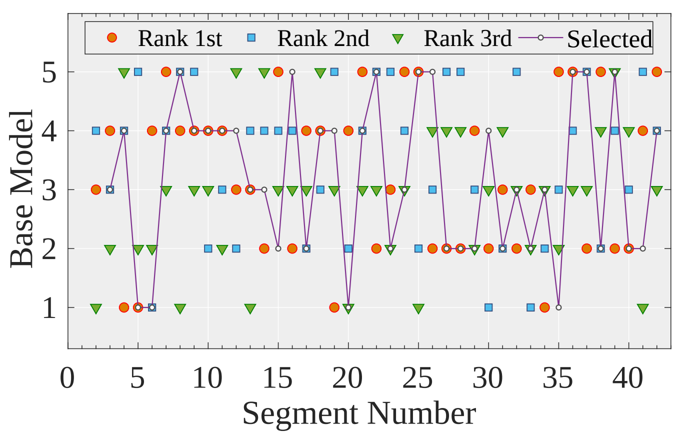


Figure 6.14: The base model ranking and selected models using the persistence method for Cycle-32

An enhanced online model selection is accomplished via extending the evaluation interval several steps backward rather than only a single step. The optimal base model is selected and updated dynamically and continuously in a certain rolling window. Reinforcement learning-based algorithms like Q-learning are one of the widely-used solutions that have been successfully developed in the literature to determine the optimal base model for wind and solar forecasting [19]. However, one of the barriers is that the Q-learning based model selection requires a large amount of dataset and intensive online model updating. Another implementable alternative is possibility-based algorithms like Bayesian model selection [74]. Similarly, in this study, given the piece-wise nature of the segment-based velocity forecasting, we develop a probability-based second-order Markov chain model that takes into account the two previous states to determine the optimal base model, as expressed by:

$$P(s_i|s_{i-1}, s_{i-2}, \cdot, \cdot, \cdot, s_1) = P(s_i|(s_{i-1}, s_{i-2})) \tag{6.9}$$

where the state s_i represents the best base model. In contrast to the HMM model discussed in Section 6.2.1, the transition matrix is achieved explicitly via frequency counting. Only the best-performing model of each segment is considered, and if the transition probability is a null set, it can be replaced with the best model of the last step as described previously. Similar to the probability approach, this method witnesses a significant enhancement on MAE (1.93 m/s), but a worse RMSE (3.16 m/s), meaning more larger variations have been brought into the forecasting.

6.3.2 Ensemble Model

Instead of utilizing a single prediction, ensemble forecasting combines a set of diverse models to mitigate the forecasting fluctuations and improve the robustness, and a linear ensemble model is given by:

$$M_{ensemble} = \sum_{i=1}^{N} \omega_i M_i \tag{6.10}$$

where ω_i denotes the weighting factor, s.t. $\sum \omega_i = 1$. Note that the weighting factor can be determined via optimization, and this study adopts an equal-weighting scheme for a generalization purpose. Combining all the five base models, the achieved MAE and RMSE are 1.95 m/s and 2.89 m/s, respectively, and a remarkable improvement has been observed in terms of RMSE compared to base models.

6.3.3 Hybrid Approach

From the aforementioned comparative studies, it is found that the online and offline model selection methods tend to improve the MAE but deteriorate the RMSE, while the ensemble approach tends to enhance the RMSE and decrease the fluctuations. As illustrated in Fig. 6.15, in this subsection, we integrate the offline probability-aided and online Markov chain model (MM)-based model selection methods with the equal-weighted ensemble approach, aiming to further improve the performance of velocity forecasting.

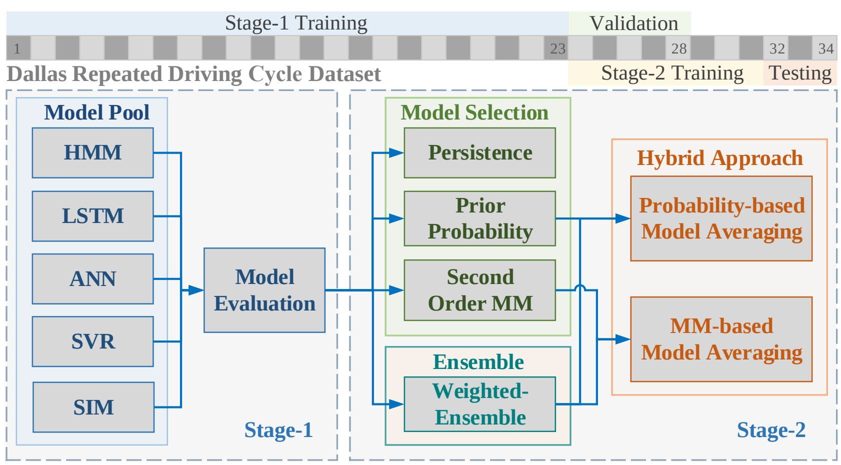


Figure 6.15: The prediction framework for short-term velocity forecasting

Based on the probability-based model, the top three base models rather than the best single one are integrated together using an ensemble method. The combined offline model yields an MAE of 1.89 m/s and an RMSE of 2.93 m/s. It is also noticed that there are non-negligible discrepancies for the performance among different cycles. The reason is that the offline probability model heavily relies on the historical data, making it challenging to adapt itself to the real-time driving conditions.

Regarding the Markov chain-based model, we further modify this approach by taking into account the top 2 models in each state, $s_t = \{M_{ti}, M_{tj}\}, i, j \in [1, 2, 3, 4, 5]$. Determined by maximum likelihood, the newly adopted models are highly related to the performances of the previous two model, as illustrated in Fig. 6.16 and expressed as:

$$M_t = argmax_{(M_t)} P(M_t | [M_{t-2}, M_{t-1}]) \tag{6.11}$$

During the training process, we explore 8 scenarios for training and 4 scenarios for testing

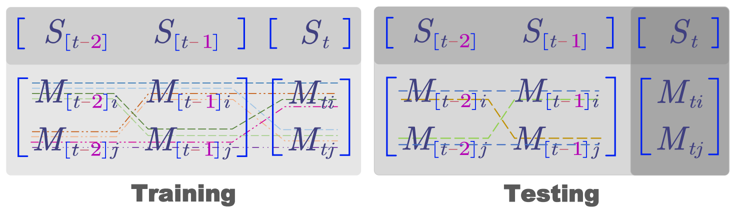


Figure 6.16: The dynamic model selection sketch using second order Markov Chain

(only 2 states are considered for testing, and the third state is the one to be predicted) in forming a Markov chain $[s_{t-2}, s_{t-1}, s_t]$. Especially, a larger factor is assigned to emphasize the sequences with top models. The final forecasts are aggregated by averaging the outputs of the four testing scenarios, which produces dominating results with an MAE of 1.87 m/s and an RMSE of 2.92 m/s. This is also the best forecasting output obtained. Compared to the single ANN model, both the MAE and RMSE have been significantly improved by 9.7% and 6.1%, respectively.

By comparison, it is found that the hybrid approaches combining the individual model selection and ensemble methods perform better than other approaches, as illustrated in Table 6.3. Though the Markov chain-based model averaging algorithm tends to produce more desirable results, given its training and updating expenses, it is more reasonable to implement the offline probability-based model averaging method in practice as a trade-off between the forecasting accuracy and computational efficiency within a control interval of 5 seconds. However, its feasibility in practical applications still requires further on-board testing and validation, when the control interval is shorten to 1 or 2 seconds, aiming to achieve more accurate and sensitive controls. Moreover, due to the very limited volume of the DRD dataset, its generalization also needs to be further verified with more accumulated cycles.

Table 6.3: A performance comparison among the individual model selection, ensemble, and hybrid approaches

Data source	Cycle-32		Cycle-34		Averaged	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE
ANN-base	1.94	2.87	2.13	3.37	2.07	3.11
LSTM-base	1.97	2.80	2.16	3.17	2.09	3.00
Persistence	1.86	2.79	2.03	3.19	2.02	3.07
Probability	1.67	2.60	2.04	3.44	1.99	3.29
2nd HMM	1.71	2.76	2.10	3.49	1.93	3.16
Ensemble	1.81	2.61	1.94	**2.97**	1.95	**2.89**
Prob-averaging	1.75	2.59	**1.90**	3.05	1.89	2.93
MM-averaging	**1.67**	**2.56**	1.99	3.14	**1.87**	2.92

Note: **Bold values** indicate the best MAE or RMSE among different approaches.

6.4 Estimation of Intersection Waiting Time

The most challenging forecast task comes from the intersection segments, where the traffic conditions as well as the operations of individual vehicles are complicated, and it is challenging to fully capture the stochastic natures. The forecasting accuracy can be significantly improved if the waiting time could be accurately predicted. According to the discussions in Section 6.1.4, it is assumed that the waiting time at an intersection is highly related to the final stopping location at roughly the same period of a day.

Given this consideration, all the intersections as well as the waiting time are extracted from the DRD dataset, and ANN models are established to further estimate the waiting time.

This approach utilizes 21 cycles for training, 7 cycles for validation, and the rest 6 cycles for testing. The hyperparameters like the number of layers and neurons are determined via grid search. The actual and estimated waiting times of Cycle 31-34 are compared in Fig. 6.17, yielding an averaged MAE of 19.08 seconds and an RMSE of 27.22 seconds. This is an acceptable outcome with such a limited set of data samples, given the purpose of waiting time estimation is to optimize the short-time scheduling of onboard systems, such as the air conditioning system and battery cooling/heating system, to avoid overlapping with the energy demand from motor start and reacceleration.

The waiting time highly depends on the arrival time or the remaining time in a traffic light cycle. The assumption behind the ANN approach is that the arrival time can be estimated based on the vehicle's location under a constant traffic volume condition. However, the vehicle locations are heavily affected by the actual accumulated in-between space, making the waiting time challenging to be accurately estimated. Without the implementation of any connectivity devices, An alternative way to further improve the estimation accuracy is to analyze the traffic signal via image detection embedded in the auto-driving module, which is actually mounted available in most newly-released electric vehicles.

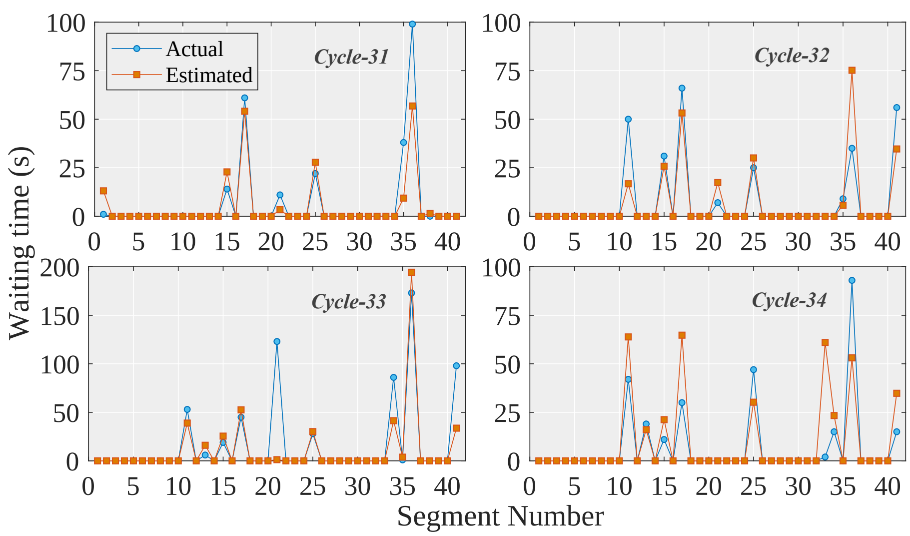


Figure 6.17: The estimated waiting time *vs.* the actual waiting time

6.5 Summary

This chapter generated a repeated urban driving cycle dataset at a fixed route in the Dallas area. Based on the data preprocessing and intersection identification, a cycle segmentation was conducted to provide location-dependent segmental data for improving velocity forecasting. A segment-based velocity forecasting model pool was developed to perform 10 seconds ahead forecasting, which takes into account the HMM, LSTM, ANN, SVR, and similarity methods. Results showed that the segment-based forecast dominated the whole cycle-based approach with great advantages. Especially, significant improvements have been observed using the ANN method with a 24% reduction for MAE and a 15% decrease for RMSE.

To further improve the forecasting accuracy, a comparative study regarding the individual model selection, ensemble approach, and a combination of them was performed. Results showed that a 9.7% improvement was obtained by leveraging the localized second-order MM-based averaging methods. However, it is more reasonable to implement the offline probability-based model averaging method in practice due to its high computational efficiency. An ANN-based intersection waiting time estimation model was also established and validated with acceptable accuracy. It is foreseeable that the improvements in both velocity forecasting and waiting time estimation will lead to better energy management, especially for electric or hybrid vehicles.

CHAPTER 7

VEHICLE ENERGY MANAGEMENT VIA TRAFFIC LIGHT DETECTION AND SEGMENTAL VELOCITY FORECASTING

The segmental velocity forecasting approach developed in Chapter 6 reveals the challenges to forecast the velocity at intersection segments only using the velocity data. To mitigate this concern, this chapter seeks to develop a YOLO-V2-based object detection deep network to recognize the traffic lights in advance, and leverage the detected signals to establish a forecasting model that integrates with the probability-based hybrid forecasting approach. Results prove that the traffic light detection-based forecasting model can significantly improve the forecasting accuracy for intersection segments. Based on the forecasting velocity 5-20 seconds ahead, the effectiveness of the MPC-based energy management strategy is further evaluated with a liquid-based battery thermal control system. Moreover, a traffic light-based real-time energy management framework is developed to directly contro l the power demand from the AC system. Simulation results suggest that this method could be a competitive alternative compared to these predictive energy management strategies.

7.1 Short-term Velocity Forecasting via Traffic Light Detection

It is worth noting that the forecasting horizon for predictive energy management varies with the control algorithm as well as the computing capability. Especially, for multi-horizon and hierarchical predictive control, the controllers utilize varying forecasting horizons for different control layers and devices. Another important factor that has great impacts on predictive controls is the control interval, which usually equals to the forecasting time step. Given these considerations, the short-term velocity forecasting in this chapter mainly focuses on 5-20 seconds ahead forecasting.

7.1.1 Localized Hybrid Model for Short-term Velocity Forecasting

Building on the aforementioned segmentation and basic forecasting approaches, a forecasting framework has been constructed with a two-stage structure to perform the short-term velocity forecasting. In the first stage, a forecasting sub-model pool that consists of a collection of stochastic and deterministic methods is established, including LSTM, ANN, SVR, HMM trained with augmented data, and a similarity-based method, as illustrated in Fig. 7.1.

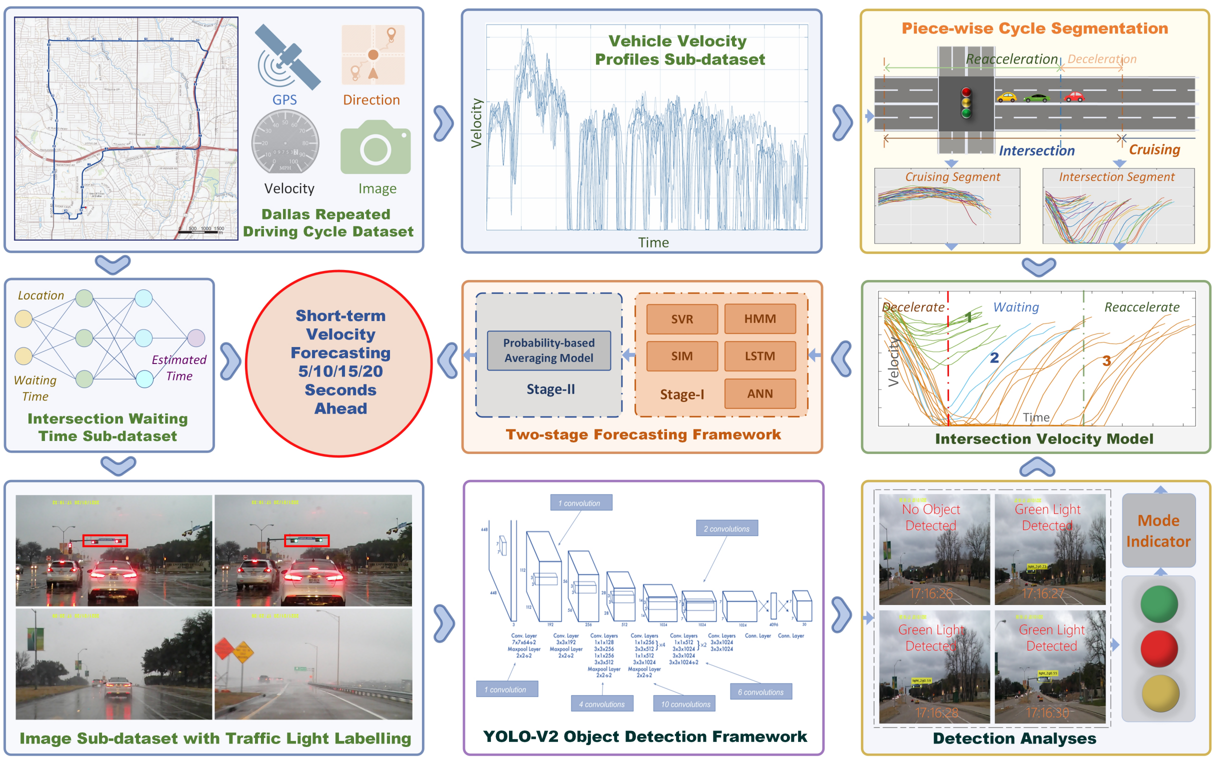


Figure 7.1: The framework of short-term vehicle velocity forecasting based on traffic light recognition and driving cycle segmentation

Similarly, as the model evaluation indicated in Table 6.3, the second stage hybrid approach only integrates the offline probability-aided ensemble model due to the limitation in computation cost, aiming to further improve the forecasting accuracy via mitigating potential fluctuations and uncertainties. The probability-aided averaging model takes advantage of the historical statistic model ranking and directly ensembles the best three base models for a specific driving segment, given as:

$$M_{prob} = \sum_{i=1}^{N} \omega_i M_i \tag{7.1}$$

where ω_i denotes an equal weighting factor with a generalization purpose for all the divided segments, s.t. $\sum \omega_i = 1$. Compared with the Markov chain-based dynamic model selection method discussed in Chapter 6, the biggest difference between these two hybrid models lies on the number of base model that is required for each segment. The dynamic Markov Chain model needs a real-time online ranking among all the base models, while the probability-aided model can update the models offline based on the recorded velocity data. This stage reuses the Stage-I validation dataset combining a collection of new data (Cycle 29-31) for training, while the rest cycles (Cycle 32-34) are used for testing.

A performance comparison among these methods is tabulated in Table 7.1. Two matrices including MAE and RMSE are employed here for performance evaluation. Compared with the base models, it is observed from the table that the hybrid method can further improve the accuracy of multi-horizon forecasting. However, it is also noticed that the improvements tend to decline as the forecasting horizon increases, which results in a worsen RMSE for 20 seconds ahead forecasting.

Table 7.1: Comparisons among different forecasting methods for 5-20 seconds ahead forecasting

Lead Time	5 seconds		10 seconds		15 seconds		20 seconds	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
ANN	1.22	1.88	2.07	3.11	2.50	3.61	2.81	3.93
LSTM	1.34	1.97	2.09	3.00	2.36	3.37	2.74	3.82
Prob-averaged	1.16	1.87	1.89	2.93	2.26	3.34	2.72	3.85

Besides the hybrid model attempts, we also recognize that the velocity profiles at intersections can be generally classified into three groups, i.e., moving forward at a constant velocity, passing through with deceleration, and completely stop, as illustrated in the intersection velocity model in Fig. 7.1. As a widely used data mining technique, extensive surveys have implied that forecasting can be significantly enhanced via accurate classification [58]. A traditional physical model-based classification method is developed in this study to divide the velocity sequences with unequal length, which directly utilizes the predefined deceleration, reacceleration, and stop processes as a classification threshold. We have also tested the hierarchical classification algorithm integrated with dynamic time warping and obtained similar results. Using the base models to forecast the velocity 5 seconds ahead, an averaged improvement of 0.11 m/s regarding MAE is observed for all the intersection segments compared to an original MAE of 1.80 m/s. Possible reasons for achieving such a limited enhancement can be attributed to the dataset size and the unbalance of varying groups, which also means the limitations could be mitigated by enlarging the dataset.

7.1.2 Short-term Velocity Forecasting via Traffic Light Detection

As discussed above, intersection velocity classification could be leveraged to improve velocity forecasting. However, in practice, it is extremely challenging to classify the unknown future velocity sequence by merely using the velocity data. For the vast majority of situations, we notice that the aforementioned three scenarios at intersections are strongly associated with the traffic light signals, as presented in the intersection velocity model. Moreover, it is observed that a green light usually leads to a steady running with high velocity, while a vehicle tends to stop at red/yellow lights. For the deceleration and reacceleration scenario, it can either be a green light at heavy traffic conditions or an ending red light followed by a green light. Given this consideration, in this chapter, we aim to develop an image-based indication framework for velocity forecasting by detecting and identifying the extra traffic control light signals via object detection.

Extensive studies have been conducted on traffic light detection by leveraging convolution neural networks (CNNs) in the emerging field of autonomous driving, covering a broad range of deep learning structures such as the R-CNN family (Fast R-CNN, Faster R-CNN, and Mask R-CNN), the YOLO (you only look once) family (v1-v5), the single shot detection (SSD) family, and the Retina-net family [28]. Since the motivations of this study emphasize on the impact of traffic light detection on vehicle velocity forecasting, a one-stage YOLO-v2 network with pre-trained network structures is directly employed here after modification. Compared with other networks, YOLO-V2 possesses an effective classification backbone with 19 convolutional layers and 5 max-pooling layers, which provides an accurate detecting precision while maintaining a high processing rate.

For this DRD dataset, all the traffic lights are horizontally installed with the same light arrangements for different colors, making it possible to directly detect the traffic light and its corresponding colors. Approximate 1,100 images are extracted from the driving cycles for labeling. To prevent potential recognition errors during the cycle, we not only label a single group of lights but also combine and label two nearby groups of lights as a whole. As a result, a traffic light and its color are recognized only when both objects have positive feedbacks. The model receives an averaged precision of 0.845 and an averaged recall of 0.567. Compared to the reported accuracies around 94% in the literature [101], this basic model still needs further improvements. Possible reasons and potential improvements are twofold: i) we have empirically labeled all the vague images for forecasting-oriented purposes to identify traffic lights as early as possible, which brings in a large amount of misleading noises; ii) the images in this study were taken in a 3X optical zoom using a household camera device, making the images in an undesirable relative low quality. However, we still obtain inspiring results in traffic light detection: the model is able to detect the traffic light with its correct color at a distance of approximately 100 meters away in a straight road with no slopes, which is also 5 to 6 seconds ahead of an intersection at a constant cruising speed. It is worth noting that the confidence score threshold in this study is set as 0.38 as a trade-off between detection accuracy and exploration. Although there may be potential misidentifications, e.g., detect other objects as a single group of traffic lights, the object detection model is still able to yield desirable outcomes with the implementation of labeling other groups of lights nearby.

Given the intersection velocity model and classification discussed above, the detected traffic light signal can act as a classification and mode indicator for the short-term velocity forecasting. Three different scenarios and their corresponding models are predefined according to the indicator outcome, as illustrated in Algorithm 2. For simplification, we only use ANN to establish these models. Especially, we find out that the reacceleration delay is highly related to the stopping location, which can be empirically calculated by a sum of the drivers' reaction time:

$$t_d = \alpha\mathcal{D}/6.5 + \beta \tag{7.2}$$

where $\mathcal{D}$ represents the distance between the vehicle and the intersection. α and β are two coefficients depending on drivers' patterns and vehicle performance. Here, α and β are selected as 1.18 and 2, respectively. Note that this model varies with location, time, and whether condition in particular.

The forecasting differences between the probability-based hybrid method and traffic light detection-assisted method are compared in Fig. 7.2. By comparing the 5 seconds ahead forecasting, it is seen that the traffic light detection-assisted method results in a significant improvement from 1.56 m/s to 0.78 m/s regarding MAE for this specific intersection. The improvement mainly comes from the deceleration process where the red light signal acts as an indicator to directly determine the ongoing stopping scenario, while there are still unavoidable forecasting delay in the reacceleration process. For other scenarios, only very small enhancements are observed by adopting this traffic light detection-assisted method. The overall improvement for the whole driving cycle is limited with an averaged 0.02 m/s reduction for MAE. Note that there is no significant improvement for the 10 seconds ahead forecasting using the light detection-assisted approach, since most of the deceleration and reacceleration processes occur within or around 10 seconds. Multiple potential approaches

Algorithm 2: Intersection Velocity Model via traffic Light Detection - Mode Indicator

Scenario-1: Moving forward at a constant velocity, Model: M1
Scenario-2: Passing through with deceleration, Model: M2
Scenario-3: Completely stop, Model: M3-1 for deceleration, M3-2=0 for waiting, M3-3 for reacceleration
Inputs: Detected traffic light: $\{green, red, yellow\}$
Velocity input: $[v_{t-1}, v_t]$, $a = v_t - v_{t-1}$
Definition: $a <= -1.2\ m/s^2$ deceleration ; $a >= 1.2\ m/s^2$ acceleration (once detected, the status will be memorized)
Switch *light color*
Case *green*
If no red signal detected previously & deceleration detected **Then** Scenario-2: M2 **end**
If no red signal detected previously & deceleration undetected **Then** Scenario-1: M1 **end**
If red signal detected previously **Then** Scenario-3: M3-3 **end**
Case *yellow*
If no green signal detected previously **Then** Scenario-3: M3-1 **end**
If green signal detected previously & deceleration undetected **Then** Scenario-1: M1 **end**
If green signal detected previously & deceleration detected **Then** Scenario-3: M3-1 **end**
Case *red*
If deceleration detected **Then** Scenario-3: M3-1 **end**
If $v_t <= 1\ m/s$ **Then** Scenario-3: M3-2 **end**

could be employed to further enhance 10-second ahead forecasting, e.g., enlarge the traffic light detection distance with high-quality images, detect and analyze driving behavior of the vehicles ahead, and develop more accurate reacceleration models. Overall, as a promising indicator, the image-based traffic light detection could be leveraged to improve the energy efficiency in a twofold manner: i) traffic light detection tends to increase the forecasting accuracy for short-term velocity forecasting; ii) it also has the potential to act as a mode trigger to activate or terminate the functions and operations of devices in advance in a predictive energy management strategy.

The short-term velocity forecasting of 5-20 seconds ahead are readily available for energy management, as illustrated in Fig. 7.3. The velocity forecasts will be regarded as the system input to implement the MPC-based energy management. For predictive energy management,

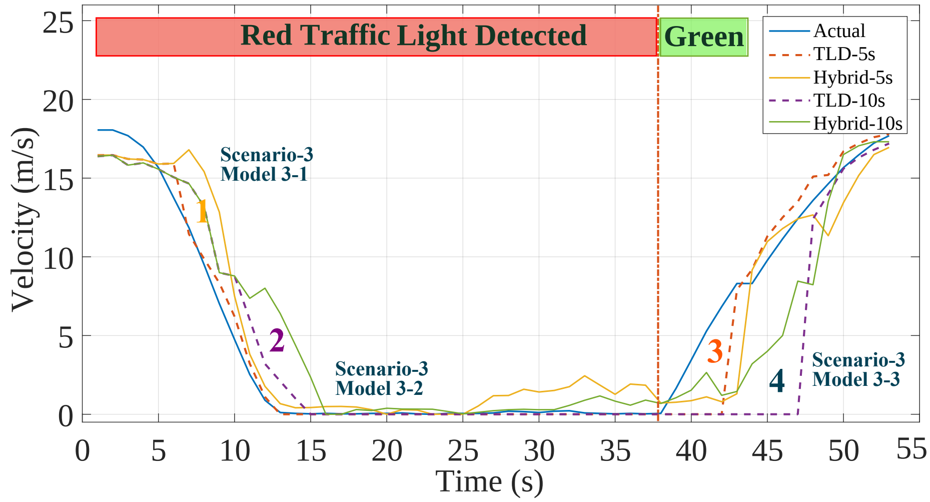


Figure 7.2: A comparison between the traffic light detection-based (TLD) method and the hybrid forecasting method at an intersection (labels 1-3 mark the improved sections, label 4 indicates a worsen section for 10 seconds ahead forecasting)

the forecasting results are extracted based on the control interval. For example, only the velocity data of 5-20 seconds ahead are required for a 5-second interval, and other velocity data can be calculated via explicit fractal interpolation.

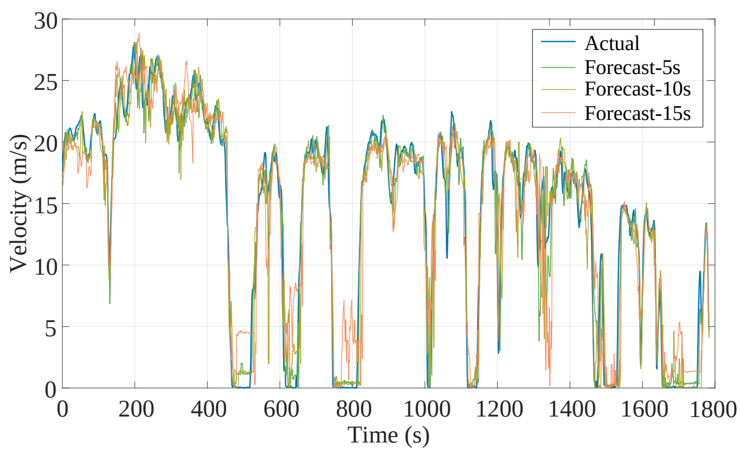


Figure 7.3: Multi-horizon velocity forecasting for the whole driving cycle

7.2 Vehicle System Modeling

There are various systems integrating and working as a whole in an electric vehicle from the perspective of electric, thermal, and energy control, including: the main battery system, the vehicle motor system, the air conditioning system, the battery thermal control system, and the cooling functions for varying components. This section attempts to develop control-oriented models for the systems discussed above, aiming to provide a comprehensive overview for further energy management. Note that some of the models have been established with more details in Chapter 5, and this section tends to emphasize more on the improvements on different parts.

7.2.1 Vehicle Battery System

For an electric vehicle, the battery pack serving as the only energy source is designed with the capability to provide sufficient energy to satisfy the power demand from varying devices and subsystems. At the battery pack level, based on the first-order lumped equivalent circuit model, the effective power output from the battery pack to the power bus can be calculated as:

$$P_{b2bus} = U_b I = (\mathcal{V}_{ocv} - I \sum_{i=0}^{N} R_i) I \tag{7.3}$$

where $\mathcal{V}_{ocv}$ denotes the open-circuit voltage, $R_b = \sum R_i$ represents the total internal resistance of a battery pack that consists of the polarization resistance and Ohmic resistance. The internal resistance R should be dynamically determined, which highly depends on the operating temperature T, state of charge (SoC), and current I. These parameters can be extracted using hybrid pulse power characterization (HPPC) [98]. For the sake of simplification, within a short time of period at a steady operating condition, Both the temperature and SoC can be treated as a fixed constant. Equation 7.3 is regarded as a quadratic equation with only one variable I, and the current can be calculated as:

$$I(SoC, T)) = \frac{\mathcal{V}_{ocv} - \sqrt{\mathcal{V}_{ocv}^2 - 4R_b P_{b2bus}}}{2R_b} \tag{7.4}$$

From the perspective of power users, the total effective power output also equals to the sum of all the power demands from subsystems, as given by:

$$P_{b2bus} = P_{sum} = P_{drv} + P_{aux} + P_{ac} + P_{bs} \tag{7.5}$$

where P_{drv}, P_{aux}, P_{ac}, and P_{bs} are the power for driving motor, auxiliary devices, air conditioning system, and battery thermal control system, respectively.

The open-circuit voltage $\mathcal{V}_{ocv}$ has a positive correlation relationship with SoC, which can also be estimated using a polynomial regression model:

$$\begin{aligned} \mathcal{V}_{ocv} &= \mathcal{P} \cdot \mathcal{S}_{soc} \\ \mathcal{S}_{soc} &= [SoC^6 \cdots SoC^2, SoC, 1]^T \end{aligned} \tag{7.6}$$

where $\mathcal{P}$ denotes a regression matrix that equals to [-40.2, 138.6, -186.2, 123.5, -42.4, 7.5, 3.3] for a specific Lithium manganese oxide (LMO) battery based on the study in Ref. [99]. The battery SoC changes as the battery discharges and charges dynamically, as defined using the Coulomb counting by:

$$SoC_{t+1} = SoC_t \pm \frac{1}{Q_c} \int_t^{t+1} I dt \tag{7.7}$$

where Q_c is the original nominal battery capacity.

Moreover, from the perspective of energy balance, it is claimed that the waste energy or power loss in the form of sensible heat can be attributed to the internal resistances, expressed as:

$$\dot{\mathcal{G}}(T, SoC, I) = \frac{I^2 R_b(T, SoC, I)}{V} \tag{7.8}$$

where $\dot{\mathcal{G}}$ and V are the volumetric heat generation rate and total battery volume, respectively. Based on the raw experimental data, the electro-thermal model is established using a Kriging-based model, and the details of the thermal model can be found in Chapters 3 and 5. Based on the above models, a battery system with a capacity of 50 kWh is investigated in this study, which consists of a total of 7200 LMO battery cells. Each battery has a nominal capacity of 1.8 Ah with a nominal voltage of 3.75 v. All the data is adopted from the studies in Refs. [98, 99].

7.2.2 Vehicle Dynamic System

For a running vehicle, the equivalent traction power consists of four major terms: the rolling friction, gravitational potential, air friction, and vehicle acceleration, defined as:

$$P_{drv} = \frac{1}{\eta_m}(mgv\mu + \frac{1}{2}\rho A_f C_d v^3 + m\frac{dv}{dt}v + mg\frac{dh}{dt}) \tag{7.9}$$

where $\mu = 0.01$ and $C_d = 0.24$ are the rolling resistance and air friction coefficient, respectively. The windward area of a vehicle A_f is set as 2.22 m/s, and the vehicle mass m equals to 1, 875 kg for a specific passenger vehicle. h denotes the altitude, which is non-negligible for an urban driving cycle with varying terrain and highway overpasses. It is also revealed that the motor power efficiency η_m is highly dependent on the motor rotating speed and torque output, ranging from 0.8 to 0.95 for varying conditions [26]. Here, η_m is also set as a constant with an average efficiency of 0.90.

During deceleration or downhill driving conditions, the regenerative braking system is activated to harvest the extra kinetic energy, with a predefined averaged efficiency of 80% (η_r) of the kinetic energy back to the power bus in this study. The driving power is negative when regeneration occurs according to the definition in Eq. 7.3, expressed as:

$$P_{r2bus} = \eta_m \eta_r (mgv\mu + \frac{1}{2}\rho A_f C_d v^3 + m\frac{dv}{dt}v + mg\frac{dh}{dt}) \tag{7.10}$$

The regenerative energy may either be utilized directly for AC and BTMS, or be recharged back to the battery pack. How and when to optimally distribute the regenerative energy needs prompt solutions. This is also one of the major contributions in this study.

7.2.3 Air Conditioning System

Compared to conventional internal combustion engine (ICE) vehicles, the air conditioning system in EVs differs in multiple perspectives: i) In winter conditions, a heat pump type AC system for EVs has predominant advantages on energy efficiency toward conventional AC with positive temperature coefficient heaters (PTC), while ICE can directly utilize the waste heat from the engine without any auxiliary devices; ii) Compared with ICE, there are several extra thermal loads for EVs , e.g., battery thermal control, motor cooling, and converter cooling; iii) The AC compressor in ICE is propelled by the engine, while it is powered by the battery pack in EVs. All these unique characteristics have led to a heat pump type AC with more complicated structures and larger cooling/heating capacity for EVs, as illustrated in Fig. 7.4.

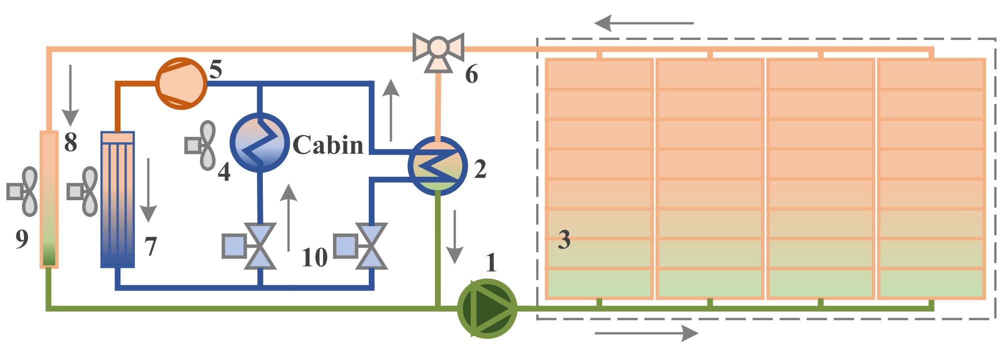


Figure 7.4: A simplified AC cooling mode for EVs. 1: BTMS pump, 2: BTMS chiller/evaporator, 3: battery pack, 4: AC evaporator, 5: compressor, 6: BTMS three-way value, 7: condenser, 8: BTMS radiator, 9: cooling fan/blower for different systems, 10: expansion valve

There are dozens of simplified conventional AC models developed in the literature using component-based dynamic thermal-fluid models or energy-based mathematical methods [35].

For vehicle energy evaluation, linear control-oriented models are constructed using numerical approaches [5, 24]. In our previous study [50], we have also established a linear mathematical AC model only with cooling mode based on the simplified AC model developed by Pino et al. [67] and the Simulink-based *CoolSim* platform developed by the National Renewable Energy Laboratory [33]. This control-friendly model consists of three parts, i.e., the AC cooling/heating capacity model, the vehicle cabin thermal load model, and the AC energy control model. From the view of energy, the relationship between the overall power consumption and the cooling/heating capacity provided can be modeled using the coefficient of performance (CoP) (η_{cop}), defined as:

$$\eta_{cop} = \frac{Q_{ac}}{P_{ac}} = \mathcal{F}(T_{in}, T_{ex}, P_{lr}) \tag{7.11}$$

where Q_{ac} and P_{ac} are the cooling/heating capacity and the power consumed, respectively. For a specific AC system, the CoP η_{cop} highly depends on the operating parameters, including the internal temperature, the external temperature, and the partial load ratio P_{lr}. The CoP for cooling ranges from 1.8 to 4.5, when the partial load ratio is larger than 0.2; the maximum CoP of the heating mode that can be achieved is only around 1.5 at a sub-zero temperature. The heat pump tends to lose its advantages in energy efficiency at a -20 $°C$ low temperature compared to PTC [105]. We also notice that besides the two-phase refrigerant, the other heat transfer medium of the evaporator is cabin air, while it is the cooling liquid for the BTMS heat exchanger. According to the refrigerant-based BTMS study conducted in [79], the CoP discrepancies for different external heat transfer mediums are limited and neglectable for energy-level modeling. In this study, the same CoP model is utilized for the sake of simplification.

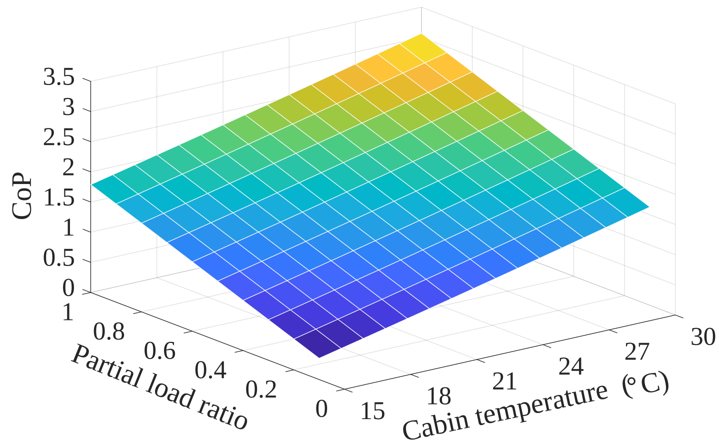


Figure 7.5: The coefficient of performance of the cooling mode

For the cabin thermal load model, there are several external and internal heat loads identified after assumptions. The detailed modeling can be found in Chapter 5. The transient thermal model of the vehicle cabin is formulated as:

$$T_{in(k+1)} = T_{in(k)} + \frac{(Q_{cr} + Q_{cw} + Q_r + Q_h + Q_f + Q_s) - (Q_{ac} - Q_{bat})}{\rho_{air} V_{in} C_{air}} \delta t \tag{7.12}$$

where δt is the time step. $T_{in(k)}$ denotes the cabin air temperature and is regarded as the state variable, and Q_{ac} is treated as the control variable. It is worth noting that the thermal inertia of the AC system has not been considered from a long-run energy balancing perspective, i.e., the sensible heat of pipe or devices and the thermal inertia of the chiller. As a result, a basic proportional-integral (PI) controller can be developed to maintain the cabin temperature within a comfortable zone.

7.2.4 Battery Liquid-based Thermal Management System

Battery Pack Thermal Model

As discussed in the literature review chapter, there are multiple thermal control technologies developed aiming to maintain the battery temperature within an appropriate range. Based on the widely-used liquid cooling technology, we have developed a phase change material

(PCM)-assisted plate cooling battery thermal control system, as illustrated in Fig. 7.6. Each battery module consists of three liquid cooling plates vertically, and in between are two battery layers. Multiple modules can be arranged to form a whole battery pack in the horizontal direction. The flow field consists of two sets of symmetric S-shape cooling channels, and the spaces in between are filled with PCM, aiming to improve the thermal performance of the cooling plate. The predefined paraffin-based RT42 PCM material has a phase-transition temperature range of T_s (41 $°C$) to T_l (42 $°C$), meaning that the PCM is solid below T_s, liquid above T_l, and mushy (solid/phase) in between the temperature range [78]. The liquid fraction β of the PCM is defined as:

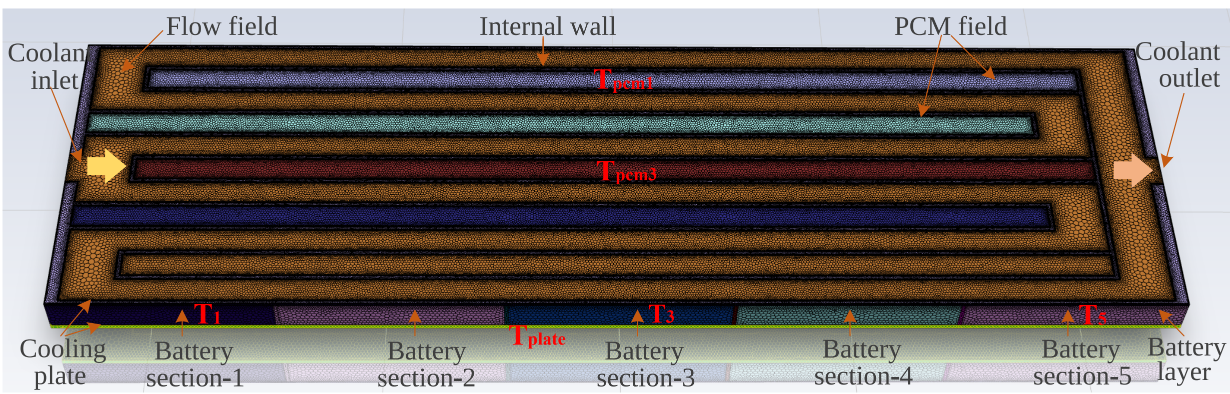


Figure 7.6: The sketch of liquid cooling structure. The whole battery pack consists of three battery modules. Each module has two battery layers and three cooling plates in the vertical direction. The battery layer has a size of 40 mm in thickness, consisting of 5 battery bricks. Each brick is 0.39 m in length and 0.26 m in width, and the spaces in between have thermal insulation materials with a thickness of 5 mm.

$$\beta = \begin{cases} 0 & T_p < T_s \\ \dfrac{T_p - T_s}{T_l - T_s} & T_s \leqslant T_p \leqslant T_l \\ 1 & T_p \geqslant T_l \end{cases} \tag{7.13}$$

where T_p denotes the averaged temperature of the PCM, which can be used to estimate the status of PCM. The motivations of adding the PCM as a supplementary here are twofold:

i) PCM is used as a thermal storage buffer to prevent any potential severe thermal impacts under intense driving conditions; ii) PCM can store a large amount of latent heat under cold temperature environment, saving considerable amount of energy from thermal preservation, especially after long time driving.

Multiple parameters are required to present a specific dynamic state of the battery pack system, such as the temperature of battery bricks (T_{b1}, T_{b3}, T_{b5}), the temperature of PCM sections (T_{pcm1}, T_{pcm3}), and the temperature of cooling plate (T_{plate}), as indicated in Fig. 7.6. Three parameters are regarded as the system inputs, including the battery inlet temperature $T_{cool\_btm\_in}$, the mass flow rate $\dot{m}_{cool}$ of the coolant, and the volumetric heat generation rate $\dot{\mathcal{G}}$. The system state updates by iteration following new system inputs, as illustrated in Fig. 7.7.

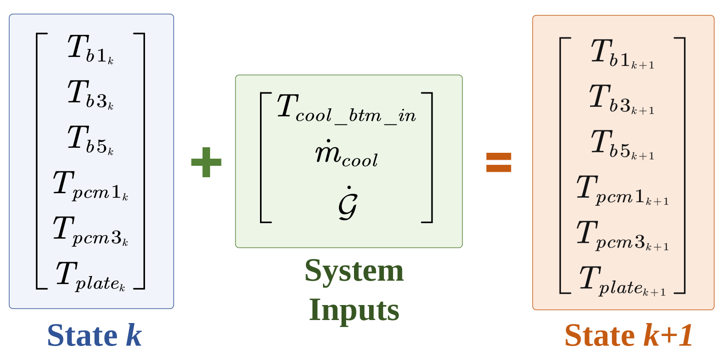


Figure 7.7: The states and inputs of the liquid-based thermal system

To evaluate and identify the dynamic response of the system, a total of 1,000 computational fluid dynamic (CFD) simulations are performed via the commercial software ANSYS Fluent with the $k-\varepsilon$ turbulence model. The CFD model has a meshing size of 3,500,000 after mesh dependency analysis, taking approximately 25 minutes to simulate a 5-second transient thermal behavior on a 12-core 1.6 GHz workstation. Based on the simulated data, a feed-forward multi-input multi-output (MIMO) neural network with three hidden layers is employed here to model the dynamic response. Determined by Bayesian optimization, the numbers of each layer are set as [100, 79, 41], which yields an RMSE of 1.5% and an MAE of 1.2%. The next system state estimated via neural network-based black-box is expressed as:

$$[T_{b1}, T_{b3}, T_{b5}, T_{pcm1}, T_{pcm3}, T_{plate}]_{k+1} =$$

$$[T_{b1}, T_{b3}, T_{b5}, T_{pcm1}, T_{pcm3}, T_{plate}]_k + \mathcal{N}_{1-1}([T_{b1}, T_{b3}, T_{b5}, T_{pcm1}, T_{pcm3}, T_{plate}]_k, [T_{cool\_btm\_in}, \dot{m}_{cool}, \dot{\mathcal{G}}]) \tag{7.14}$$

In addition, the BTMS outlet coolant temperature can also be estimated as another output using the same MIMO model, given as:

$$T_{cool\_btm\_out} = T_{cool\_btm\_in} + \mathcal{N}_{1-2}([T_{b1}, T_{b3}, T_{b5}, T_{pcm1}, T_{pcm3}, T_{plate}]_k, [T_{cool\_btm\_in}, \dot{m}_{cool}, \dot{\mathcal{G}}]) \tag{7.15}$$

BTMS Control Model

As illustrated in Fig. 7.4, two kinds of heat sink are designed for different conditions, including an AC-based chiller that works in high temperature environments or intense charging/discharging conditions, and a fluid-air radiator that operates in an environment with a relatively low temperature. Compared with the AC-connected chiller, a radiator has significantly higher energy efficiency by directly dissipating the waste heat into environment where there exists a big temperature difference.

Under high temperature conditions, the chiller is the only alternative for thermal control. As demonstrated in Eq. 7.12, the chiller model is defined using the energy balance equation, given as:

$$Q_{bat} = \dot{m}_{cool} \left(T_{cool\_chl\_in} - T_{cool\_chl\_out} \right) C_{cool} \tag{7.16}$$

where T_{cool_in} and T_{cool_out} are the inlet and outlet temperature of the BTMS coolant, respectively. From the view of efficiency, the coolant mass flow rate $\dot{m}_{cool}$ is supposed to be adaptively adjusted via the speed control of coolant pump for varying operating conditions. It is worth noting that controlling the coupled variables is very challenging, which requires a detailed dynamic thermal-fluid model to determine the mass flow rate and the coolant temperature at the chiller outlet simultaneously. Here in this study, for the sake of simplification, the mass flow rate is fixed at its maximum volume ($\dot{m}$=0.9 kg/s) via an on-off control. The coolant temperature at the chiller outlet can then be determined by Eq. 7.16.

As regards the BTMS, a thermal control strategy is expected to address the temperature gap between the real-time temperature and the pre-defined temperature control trajectory, defined as:

$$\Delta T_{gap} = ave([T_{b1}, T_{b3}, T_{b5}]) - T_{bat\_ref} \tag{7.17}$$

From the CFD simulations, we notice that the battery temperatures between two states change when the coolant is running at a specific mass flow rate and inlet temperature, as described in Eq. 7.14. The relationship can be extracted and expressed as:

$$\begin{aligned} \Delta T_{bat\_dif} &= ave([T_{b1}, T_{b3}, T_{b5}]_k) - ave([T_{b1}, T_{b3}, T_{b5}]_{k+1}) \\ \Delta T_{bat\_dif} &= \mathcal{N}_{1-3}([T_{b1}, T_{b3}, T_{b5}, T_{pcm1}, T_{pcm3}, T_{plate}]_k, [T_{cool\_btm\_in}, \dot{m}_{cool}, \dot{\mathcal{G}}]) \end{aligned} \tag{7.18}$$

In a control process, ΔT_{bat_dif} is treated as an obtained value after control actions, while ΔT_{gap} is the control target. The anticipated coolant temperature at the BTMS inlet is predicted as follow:

$$T_{cool\_bat\_in} = \mathcal{N}_{2-1}([ave(T_{b1}, T_{b3}, T_{b5}), T_{pcm1}, T_{pcm3}, T_{plate}]_k, [\dot{m}_{cool}, \dot{\mathcal{G}}, \Delta T_{gap}]) \tag{7.19}$$

where the multi-input single-output (MISO) controller model can also be regarded as an inverse function of Eq. 7.14. Similar to the plant model, the neuron numbers of each layer are determined as [53, 28,12] via Bayesian optimization, yielding an RMSE of 3.4% and an MAE of 2.8%. By substituting Eq. 7.2.4 into the chiller model as depicted in Eq. 7.16, the cooling demand of BTMS can be determined. Additionally, the power consumption of the coolant pump can be estimated based on Eq. 7.20:

$$P_{pump} = \dot{m}_{cool} \Delta p_{pre} / \rho_{cool} \eta_{pump} \tag{7.20}$$

where η_{pump} and Δp_{pre} represent the pump efficiency and the pressure drop of the battery system, respectively. ρ_{cool} is the coolant density.

7.3 Energy Management and Case Study

7.3.1 MPC-based Energy Management for Daily Commute

Similar to the MPC-based energy strategy investigated in Chapter 5, based on the subsystem models established above, an MPC strategy is developed here with a cost function given as:

$$\begin{aligned}
\underset{P_{cab}, P_{bat}}{\arg\min}\ J = & \sum_{k=n}^{n+N} (\alpha_k (P_{drv_k} + P_{aux_k} + P_{cab_k} + P_{bat_k} + P_{pump_k})^2 \\
& + \beta_k (T_{ref_k} - T_{bat_ave_k})^2 + \xi_k (T_{tar_k} - T_{cab_in_k})^2) \\
\text{subject to} \quad & Q_{cab} = P_{cab} \eta_{cop} \\
& Q_{bat} = P_{bat} \eta_{cop} \\
& 0 \leq Q_{cab} \leq 5000 \\
& 0 \leq Q_{bat} \leq 5000 \\
& |\Delta Q_{cab}| \leq 1000 \\
& |\Delta Q_{bat}| \leq 500
\end{aligned} \tag{7.21}$$

where P_{cab} and P_{bat} are the AC power consumption for the cabin thermal control system and battery thermal management system, respectively. The power allocations within the AC system can be implemented via value and flow direction controls, and here, we only focus on the optimization at the system level. P_{bat} is the power demand from battery coolant pump, which is relatively small compared to the AC system. Note that the coefficients are adaptive to different conditions and are determined using other driving cycles. This problem is also solved via the particle swarm optimization algorithm, due to the highly non-convex and nonlinear characteristics, and a local optimum instead of a global optimum is expected given a limited computational time.

Based on the 5-20 seconds ahead velocity forecasting for a daily commute route, the initial system parameters are set with an exterior temperature of 310.15 K and a battery SoC of 0.95. The cabin temperature is targeted at 294.15 K, while the battery control temperature aims at 313.15 K. The upper bound of the battery is set to be 317.15 K, leaving a large margin for thermal impacts considering the usage of PCM.

The simulation results of the real-time control and the MPC-based approach are presented in Fig. 7.8 and Fig. 7.9, respectively. The PCM remains fluid status for most of the time during the driving cycles since the temperature is well constrained around the target value. We observe that the final value of SoC decreases from 0.95 to 0.8818 with the MPC, while the real-time control yields an SoC of 0.8806. Compared with the real-time control, only limited improvement is obtained via MPC for this specific driving cycle, i.e., less than 2% regarding the energy efficiency calculated by SoC. We also verified the results based on the real velocity data rather than forecasting. It is known that the MPC strategy is less effective for a steady driving stage with cruising speed, but performs well for varying conditions. Given this consideration, four potential reasons may account for this observation: (i) Compared to the hybrid driving cycles that consist of UDDS, WLTC, and HWFET cycles, this Dallas driving cycles are considerably smoother and have fewer intersections and complete stops. (ii) The waste heat dissipation from the battery system lies on the AC system, and combining the cooling demand from the cabin side, the AC system is running at a high load ratio. The AC load is also very close to the regenerative power, leaving limited space for load shifting. (iii) The control action interval is 5 seconds, making the AC system less responsive to the varying powertrain demands, so does the coolant pump with on-off control. (iv) The solutions are local optima rather than global optima for a nonlinear non-convex problem due to the very limited computational time of 4 seconds.

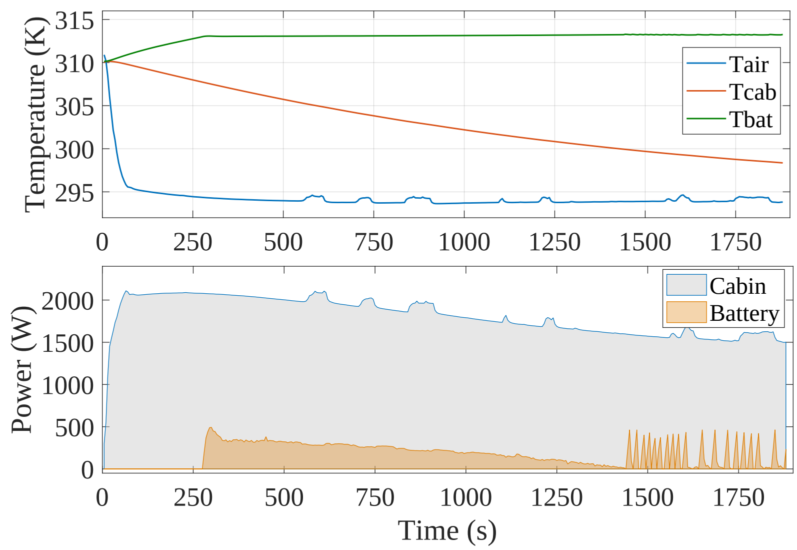


Figure 7.8: The performance of real-time control based energy management

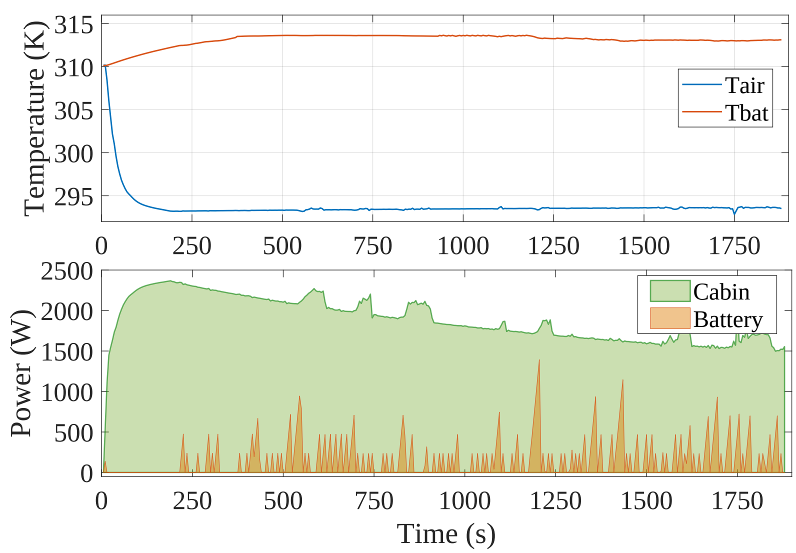


Figure 7.9: The performance of MPC-based energy management

7.3.2 Real-time Energy Management Using Traffic Light Detection

As discussed above, the MPC-based energy management has its drawbacks during steady driving conditions, especially when using nonlinear and non-convex models for complicated systems. It is also identified that the load shifting mainly occurs on changing conditions, i.e., the deceleration and reacceleration processes at intersections. Given these considerations, we aim to develop a real-time energy control framework to avoid the overlapping among peaks and to reuse the regenerative power instead recharging back to the battery pack. For example, if a deceleration is known, the system is motivated to decrease the power demand of the AC system by defining a higher target temperature prior to that, then the AC system can be powered by the regenerative energy during the deceleration. On the other hand, when a reacceleration is predicted, the system tends to lower down the manipulated temperature to avoid overlapping of different loads.

The image-based traffic light detection method developed in Section 7.1, in conjunction with real-time acceleration signals, is utilized here as the mode indicator to update the control parameter settings. At an intersection, the vehicle is able to activate the low-demand mode based on the vehicle locations. Given the largest detection threshold of 100 meters in our model, the distance for low-demand mode activation is set as 250 meters, which is approximately 12-13 seconds in prior to an intersection at a cruising speed. It is worth noting that accurate traffic light recognition can be achieved as far as 130-150 meters away as reported in the literature [101], which is about 7-8 seconds ahead for a cruising scenario and more than 20 seconds ahead for a completely stop scenario. The following implementation of this detection-based energy management method is based on the aforementioned obtained detection results in this chapter.

Similar to the pseudo-code described in Algorithm 2, base on the detected traffic light signals, two major scenarios are predefined with a sequence of energy allocation actions: (i) For a green light signal, the vehicle tends to the low-demand model is deactivated and switched back to the normal control. (ii) For red-light conditions, the AC system is expected to reuse the regenerative power as more as possible instead of recharging back to the battery system during deceleration. For the waiting section, the system needs to determine an averaged AC power demand based on the waiting time estimation. The AC system continues working in a relative higher-power manner to lower down the temperature in prior to the intersection. As a result, the AC system avoids overlapping with the power demand from the driving system during reacceleration. The system switches back to the normal control mode when leaving an intersection, as indicated in Fig. 7.10. It is worth noting that the traffic light can act as an early termination signal, i.e., a green light detected in front of an intersection suggests switching back to normal control directly.

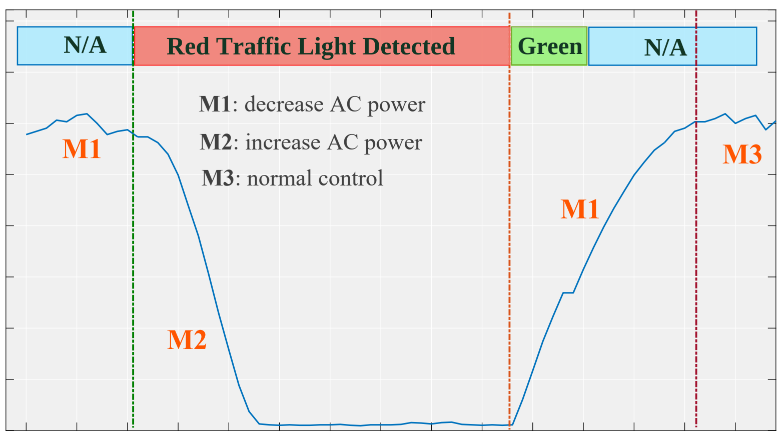


Figure 7.10: The sketch of a traffic light detection-based energy management strategy for intersection

In this study, the AC energy consumed by the battery thermal system remains unchanged due to its limited total volume. The AC energy for the cabin is adjusted based on the aforementioned principle. The simulation results of the modified control are presented in Fig. 7.11. Compared with the real-time energy distribution, the main differences come from the intersections. As a trade-off, the final stage SoC is improved by approximately 2.5% to 0.8823, at the cost of introducing more variations to the cabin temperature. Compared to the MPC-based energy management, the traffic light detection-based energy management

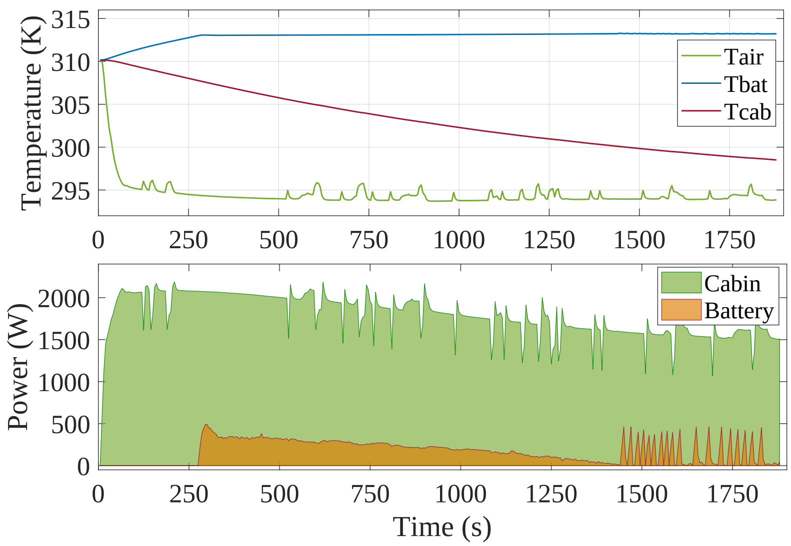


Figure 7.11: The performance of traffic light detection-based energy management

has very similar performance for a smooth driving cycle. Given its computational efficiency, it is anticipated that the traffic light detection-based method could be a potential alternative for vehicle energy management.

7.4 Summary

This chapter developed a deep learning-based YOLO V2 traffic light detection framework, to improve the accuracy of short-term vehicle velocity forecasting, especially for 5-seconds ahead forecasting. According to the model evaluation conducted in Chapter 6, a probability-based offline hybrid model was employed to perform 5-20 seconds ahead velocity forecasting. Besides the thermal model established in Chapter 5, this chapter also established a liquid-based battery thermal management system via transient CFD simulations. An MPC approach was employed to optimize the energy efficiency. Moreover, we also evaluated the feasibility of leveraging image-based traffic light detection to modify real-time energy distribution. Results revealed that the traffic light detection-aided real-time control could be a potential alternative to vehicle energy management for urban commuting routes.

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

In this dissertation, we have comprehensively discussed the challenges in electric vehicles and developed corresponding solutions, from the perspective of thermal control and energy optimization. To address the temperature uniformity problem at the battery pack level, we have developed a J-type air-based cooling structure via optimization and its corresponding thermal control method with an operation mode switching mechanism. To improve the energy efficiency, we have also further developed model predictive control-based and real-time traffic light detection-based energy management strategies to schedule power allocation at the device level.

In Chapter 3, we developed a novel J-type air-based cooling structure based on the existing conventional U- and Z-type structures. The J-type structure has two outlets with two control valves, while the conventional U-type and Z-type only have one outlet. By changing the opening degrees of two control valves, the J-type structure can be flexibly switched to either U-type or Z-type. To quantify the electric-thermal inputs for cooling structures, an electric-thermal model for Lithium-ion battery was first established using experimental data via a Kriging surrogate model. Based on the established electric-thermal model, comparative studies among the U-, Z-, and J- type structures were then performed to evaluate the impacts of basic structural and control parameters via CFD simulations. Aiming to optimize the thermal performances, we developed a two-stage cluster-based resampling optimization framework via a collection of surrogate models and successfully applied it to structure optimization. The results revealed that the J-type structure has prevailing advantages towards the U- and Z-type structure regarding thermal performances and pressure drop. Especially, we found that the J-type structure is flexible to adapt the settings by controlling the opening degree of the two control valves under varying working conditions.

In Chapter 4, inspired by flexible settings of the J-type structure, we established data-driven neural network-based plant and controller models via transient CFD simulations to validate the dynamic performance of the whole thermal management system. To balance the temperature distribution within a battery pack, we also developed an operation mode switching module to quasi-periodically switch the operation mode by controlling the two outlet valves. Moreover, an MPC-based thermal control strategy was employed aiming to improve the energy efficiency in thermal management. Simulation results revealed significant improvements by comparing with the original benchmark control implementations.

In Chapter 5, we extended the MPC-based approach to energy management by simultaneously controlling the vehicle cabin climate system and the J-type air-based BTMS. The energy management aimed to reduce the load peaks by rescheduling the operations of varying devices and directly reusing the regenerative power rather than recharging back to the battery system, while retaining the thermal constraints of BTMS and AC systems. As the second largest power consumer, the air conditioning system model was established based on both the cabin thermal mode and the BTMS pre-cooling model. By comparing with the conventional real-time energy allocation, an overall improvement of 6.5% was observed using the MPC-based approach regarding the energy efficiency under a known dynamic driving model, and the MPC-based management strategy was proved as a promising solution to enhance the battery efficiency for electric vehicles.

In Chapter 6, to evaluate the effectiveness of the MPC-based energy management approach to practical velocity forecasting, we generated a repeated commuting driving cycle dataset in the Dallas area, aiming to simulate a typical urban commuting route and provide insights for individual vehicle velocity forecasting. During the pre-processing, a piece-wise segmentation approach was employed using the identified intersections and potential stops. Then based on the piece-wise data, a velocity forecasting pool with a collection of forecasting algorithms was established, including ANN, SVM, LSTM, HMM, and similarity-based methods. To further improve the forecasting accuracy, multiple higher-level probability-based ensemble models were developed using off-line and on-line dynamic model selection techniques.

In Chapter 7, the forecasted multi-horizon velocity data was utilized as the system input to validate a liquid-based predictive energy management approach. An imaged-based traffic light detection framework was developed to improve the forecasting performance for intersection segments based on a deep learning-based YOLO V2 network. To model the dynamic systems of an EV, a liquid-based battery thermal control system with PCM was established via CFD simulations. Then an MPC-based approach was employed to optimize the energy efficiency. Moreover, the feasibility of leveraging image-based traffic light detection to modify real-time energy distribution was also evaluated. Results revealed that both the traffic light detection-aided real-time control and MPC-based energy management could be applied to urban commuting routes.

The potential extension of this dissertation includes: (i)reduce the computational complexity by linearizing the system or employing decentralized MPC approaches, (ii) adopt a hierarchical multi-system vertical model predictive control algorithm to strengthen the control for the BTMS and cabin AC system, (iii) develop advanced traffic light recognition systems to improve the vehicle energy efficiency.

APPENDIX

TECHNICAL REMARKS OF POPULAR EVS (2018)

Model	*EVs*	*Cooling strategy*	*Battery*	*Chemistry*	*Type*	*N. voltage(V) /capacity(Ah)*	*Arrangement*	*E. density (Wh/kg)*	*Capacity (KWh)*
BAIC EV200	BEV	passive cooling	SKI (BESK)	NMC	pouch	3.3/30.5	-	184	30.4
NISSAN Leaf 2018	BEV	passive cooling	AESC	NMC 622	pouch	3.65/56.3	$\boxed{2P\times2S\times2S}\times24S$	224	40
TESLA Model S	BEV	active liquid cooling	Panasonic	NCA	cylindrical 18650	3.6/3.35	$\boxed{74P\times6S}\times16S$	265	85
JAC IEV 5	BEV	active liquid& air cooling	SINOEV	NMC	cylindrical 18650	3.7/2.2	$\boxed{32P}\times92S$	-	23
TESLA Model X	BEV	active liquid cooling	Panasonic	NCA	cylindrical 18650	3.6/-	-	265	90/100
RENAULT Zoe40 2017	BEV	active air cooling	LG Chem	NMC	pouch	3.75/65	$\boxed{2P\times8S}\times12S$	240	41
CHERY EQ1 2017	BEV	passive cooling	CATL	NMC	prismatic	-	-	240	32.3
TESLA Model 3	BEV	active liquid cooling	Panasonic	NCA	cylindrical 2170	3.6/5.0	$\boxed{46P\times23S}\times2S + \boxed{46P\times25S}\times2S$	265	50
VOLKSWAGEN E-Golf 2017	BEV	passive cooling	SDI	NMC	prismatic	3.7/37	$\boxed{3P\times2S}\times10S + \boxed{3P\times4S}\times17S$	236	35.8
ZOTYE E30 2017	BEV	passive cooling	JEVE (Mitsui)	NMC	pouch	3.2/22	-	210	16
CHEVROLET Bolt EV	BEV	active liquid cooling	LG Chem	NMC 622	pouch	3.65/55	$\boxed{3P\times10S}\times8S + \boxed{3P\times8S}\times2S$	-	57
TOYOTA Prius prime	PHEV	active air cooling	Panasonic	LIB	-	3.7/25	$\boxed{19S}\times5S$	-	8.8
BYD Qin EV300	PHEV	passive cooling	BYD	LFP	prismatic	-	-	-	-

REFERENCES

[1] Afram, A. and F. Janabi-Sharifi (2014). Theory and applications of hvac control systems–a review of model predictive control (mpc). *Building and Environment 72*, 343–355.

[2] Alfi, A., M. Charkhgard, and M. H. Zarif (2014). Hybrid state of charge estimation for lithium-ion batteries: design and implementation. *IET Power Electronics 7*(11), 2758–2764.

[3] Amini, M. R., I. Kolmanovsky, and J. Sun (2021). Hierarchical mpc for robust eco-cooling of connected and automated vehicles and its application to electric vehicle battery thermal management. *IEEE Transactions on Control Systems Technology 29*(1), 316–328.

[4] Amini, M. R., J. Sun, and I. Kolmanovsky (2018). Two-layer model predictive battery thermal and energy management optimization for connected and automated electric vehicles. In *2018 IEEE Conference on Decision and Control (CDC)*, pp. 6976–6981. IEEE.

[5] Amini, M. R., H. Wang, X. Gong, D. Liao-McPherson, I. Kolmanovsky, and J. Sun (2019). Cabin and battery thermal management of connected and automated hevs for improved energy efficiency using hierarchical model predictive control. *IEEE Transactions on Control Systems Technology*.

[6] Bagloee, S. A., M. Tavana, M. Asadi, and T. Oliver (2016). Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of modern transportation 24*(4), 284–303.

[7] Bandara, K., H. Hewamalage, Y.-H. Liu, Y. Kang, and C. Bergmeir (2020). Improving the accuracy of global forecasting models using time series data augmentation. *arXiv preprint arXiv:2008.02663*.

[8] Barik, B., P. K. Bhat, J. Oncken, B. Chen, J. Orlando, and D. Robinette (2018). Optimal velocity prediction for fuel economy improvement of connected vehicles. *IET Intelligent Transport Systems 12*(10), 1329–1335.

[9] Bassiouny, M. and H. Martin (1984). Flow distribution and pressure drop in plate heat exchangers—i u-type arrangement. *Chemical Engineering Science 39*(4), 693–700.

[10] Blunsom, P. (2004). Hidden markov models. *Lecture notes, August 15*(18-19), 48.

[11] Brunton, S. L., J. L. Proctor, and J. N. Kutz (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences 113*(15), 3932–3937.

[12] Burban, G., V. Ayel, A. Alexandre, P. Lagonotte, Y. Bertin, and C. Romestant (2013). Experimental investigation of a pulsating heat pipe for hybrid vehicle applications. *Applied Thermal Engineering 50*(1), 94–103.

[13] Chen, B., D. Robinette, M. Shahbakhti, K. Zhang, J. Naber, J. Worm, C. Pinnow, and C. Morgan (2017). Connected vehicles and powertrain optimization. *Mechanical Engineering 139*(09), S12–S18.

[14] Chen, K., S. Wang, M. Song, and L. Chen (2017). Structure optimization of parallel air-cooled battery thermal management system. *International Journal of Heat and Mass Transfer 111*, 943–952.

[15] Cheng, S., F. Lu, and P. Peng (2020). Short-term traffic forecasting by mining the non-stationarity of spatiotemporal patterns. *IEEE Transactions on Intelligent Transportation Systems*.

[16] Cui, M. and J. Zhang (2018). Estimating ramping requirements with solar-friendly flexible ramping product in multi-timescale power system operations. *Applied Energy 225*, 27–41.

[17] Dai, H., X. Wei, Z. Sun, J. Wang, and W. Gu (2012). Online cell soc estimation of li-ion battery packs using a dual time-scale kalman filtering for ev applications. *Applied Energy 95*, 227–237.

[18] Fan, X., C. Xiang, L. Gong, X. He, Y. Qu, S. Amirgholipour, Y. Xi, P. Nanda, and X. He (2020). Deep learning for intelligent traffic sensing and prediction: recent advances and future challenges. *CCF Transactions on Pervasive Computing and Interaction*, 1–21.

[19] Feng, C., M. Sun, and J. Zhang (2020). Reinforced deterministic and probabilistic load forecasting via q -learning dynamic model selection. *IEEE Transactions on Smart Grid 11*(2), 1377–1386.

[20] Gaikwad, T., A. Rabinowitz, F. Motallebiaraghi, T. Bradley, Z. Asher, A. Fong, and R. Meyer (2020). Vehicle velocity prediction using artificial neural network and effect of real world signals on prediction window. Technical report, SAE Technical Paper.

[21] Gao, X., Y. Ma, and H. Chen (2018). Active thermal control of a battery pack under elevated temperatures. *IFAC-PapersOnLine 51*(31), 262–267.

[22] Guanetti, J., Y. Kim, and F. Borrelli (2018). Control of connected and automated vehicles: State of the art and future challenges. *Annual Reviews in Control 45*, 18–40.

[23] Guo, J. and B. M. Williams (2010). Real-time short-term traffic speed level forecasting and uncertainty quantification using layered kalman filters. *Transportation Research Record 2175*(1), 28–37.

[24] He, H., H. Jia, C. Sun, and F. Sun (2018). Stochastic model predictive control of air conditioning system for electric vehicles: Sensitivity study, comparison, and improvement. *IEEE Transactions on Industrial Informatics 14*(9), 4179–4189.

[25] He, H., R. Xiong, and H. Guo (2012). Online estimation of model parameters and state-of-charge of lifepo4 batteries in electric vehicles. *Applied Energy 89*(1), 413–420.

[26] Huynh, T. A. and M.-F. Hsieh (2018). Performance analysis of permanent magnet motors for electric vehicles (ev) traction considering driving cycles. *Energies 11*(6), 1385.

[27] Jiang, B. and Y. Fei (2015). Traffic and vehicle speed prediction with neural network and hidden markov model in vehicular networks. In *2015 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1082–1087. IEEE.

[28] Jiao, L., F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, and R. Qu (2019). A survey of deep learning-based object detection. *IEEE access 7*, 128837–128868.

[29] Jing, J., D. Filev, A. Kurt, E. Özatay, J. Michelini, and Ü. Özgüner (2017). Vehicle speed prediction using a cooperative method of fuzzy markov model and auto-regressive model. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pp. 881–886. IEEE.

[30] Johannesson, L., M. Asbogard, and B. Egardt (2007). Assessing the potential of predictive control for hybrid vehicle powertrains using stochastic dynamic programming. *IEEE Transactions on Intelligent Transportation Systems 8*(1), 71–83.

[31] Khayyam, H., A. Z. Kouzani, E. J. Hu, and S. Nahavandi (2011). Coordinated energy management of vehicle air conditioning system. *Applied thermal engineering 31*(5), 750–764.

[32] Kim, J., J. Oh, and H. Lee (2018). Review on battery thermal management system for electric vehicles. *Applied Thermal Engineering*.

[33] Kiss, T., L. Chaney, and J. Meyer (2013). New automotive air conditioning system simulation tool developed in matlab/simulink. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

[34] Laboratory, I. N. Ev auxiliary systems impacts. [Online] Available at: `https://avt.inl.gov/sites/default/files/pdf/fsev/auxiliary.pdf`.

[35] Lajunen, A., Y. Yang, and A. Emadi (2020). Review of cabin thermal management for electrified passenger vehicles. *IEEE Transactions on Vehicular Technology 69*(6), 6025–6040.

[36] Lana, I., J. Del Ser, M. Velez, and E. I. Vlahogianni (2018). Road traffic forecasting: Recent advances and new challenges. *IEEE Intelligent Transportation Systems Magazine 10*(2), 93–109.

[37] Lazrak, A., J.-F. Fourmigué, and J.-F. Robin (2018). An innovative practical battery thermal management system based on phase change materials: Numerical and experimental investigations. *Applied Thermal Engineering 128*, 20–32.

[38] Le Rhun, A., F. Bonnans, G. De Nunzio, T. Leroy, and P. Martinon (2019). A bi-level energy management strategy for hevs under probabilistic traffic conditions.

[39] Lee, H., C. Song, N. Kim, and S. W. Cha (2020). Comparative analysis of energy management strategies for hev: Dynamic programming and reinforcement learning. *IEEE Access 8*, 67112–67123.

[40] Lee, K., M. Eo, E. Jung, Y. Yoon, and W. Rhee (2020). Short-term traffic prediction with deep neural networks: A survey. *arXiv preprint arXiv:2009.00712*.

[41] Lemieux, J. and Y. Ma (2015). Vehicle speed prediction using deep learning. In *2015 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pp. 1–5.

[42] Li, M., Y. Liu, X. Wang, and J. Zhang (2019). Modeling and optimization of an enhanced battery thermal management system in electric vehicles. *Frontiers of Mechanical Engineering 14*(1), 65–75.

[43] Ling, Z., J. Chen, X. Fang, Z. Zhang, T. Xu, X. Gao, and S. Wang (2014). Experimental and numerical investigation of the application of phase change materials in a simulative power batteries thermal management system. *Applied Energy 121*, 104–113.

[44] Ling, Z., F. Wang, X. Fang, X. Gao, and Z. Zhang (2015). A hybrid thermal management system for lithium ion batteries combining phase change materials with forced-air cooling. *Applied Energy 148*, 403–409.

[45] Ling, Z., Z. Zhang, G. Shi, X. Fang, L. Wang, X. Gao, Y. Fang, T. Xu, S. Wang, and X. Liu (2014). Review on thermal management systems using phase change materials for electronic components, li-ion batteries and photovoltaic modules. *Renewable and Sustainable Energy Reviews 31*, 427–438.

[46] Liu, H., Z. Wei, W. He, and J. Zhao (2017). Thermal issues about li-ion batteries and recent progress in battery thermal management systems: A review. *Energy conversion and management 150*, 304–330.

[47] Liu, K., Z. Asher, X. Gong, M. Huang, and I. Kolmanovsky (2019). Vehicle velocity prediction and energy management strategy part 1: Deterministic and stochastic vehicle velocity prediction using machine learning. Technical report, SAE Technical Paper.

[48] Liu, Y., P. Ghassemi, S. Chowdhury, and J. Zhang (2018). Surrogate based multi-objective optimization of j-type battery thermal management system. In *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection.

[49] Liu, Y., M. Li, and J. Zhang (2017). An experimental parametric study of air-based battery thermal management system for electric vehicles. In *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pp. V02AT03A024–V02AT03A024. American Society of Mechanical Engineers.

[50] Liu, Y. and J. Zhang. Electric vehicle battery thermal and cabin climate management based on model predictive control. *Journal of Mechanical Design 143*(3).

[51] Liu, Y. and J. Zhang (2019). Design a j-type air-based battery thermal management system through surrogate-based optimization. *Applied Energy 252*, 113426.

[52] Lophaven, S., H. Nielsen, and J. Sondergaard (2016). Dace a matlab kriging toolbox [eb/ol].

[53] Lv, M., Z. Hong, L. Chen, T. Chen, T. Zhu, and S. Ji (2020). Temporal multi-graph convolutional network for traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 1–12.

[54] Ma, X., Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang (2017). Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. *Sensors 17*(4), 818.

[55] Mahamud, R. and C. Park (2011). Reciprocating air flow for li-ion battery thermal management to improve temperature uniformity. *Journal of Power Sources 196*(13), 5685–5696.

[56] Marcos, D., F. J. Pino, C. Bordons, and J. J. Guerra (2014). The development and validation of a thermal model for the cabin of a vehicle. *Applied Thermal Engineering 66*(1-2), 646–656.

[57] Mark, C., C. Metzner, L. Lautscham, P. L. Strissel, R. Strick, and B. Fabry (2018). Bayesian model selection for complex dynamic systems. *Nature communications 9*(1), 1–12.

[58] Masdari, M. and A. Khoshnevis (2020). A survey and classification of the workload forecasting methods in cloud computing. *Cluster Computing 23*(4), 2399–2424.

[59] Masoudi, Y. and N. L. Azad (2017). Mpc-based battery thermal management controller for plug-in hybrid electric vehicles. In *2017 American Control Conference (ACC)*, pp. 4365–4370. IEEE.

[60] Mengzhang, L. and Z. Zhanxing (2020). Spatial-temporal fusion graph neural networks for traffic flow forecasting. *arXiv preprint arXiv:2012.09641*.

[61] Moser, D., H. Waschl, R. Schmied, H. Efendic, and L. del Re (2015). Short term prediction of a vehicle's velocity trajectory using its. *SAE International Journal of Passenger Cars-Electronic and Electrical Systems 8*(2015-01-0295), 364–370.

[62] Nair, D. J., F. Gilles, S. Chand, N. Saxena, and V. Dixit (2019). Characterizing multicity urban traffic conditions using crowdsourced data. *PLoS one 14*(3), e0212845.

[63] NCTCG. Government report transportation system management. [Online] Available at: `https://www.nctcog.org/trans/manage/its/transportation-systems-management`.

[64] Oh, G., D. J. Leblanc, and H. Peng (2019). Vehicle energy dataset (ved), a large-scale dataset for vehicle energy consumption research. *arXiv preprint arXiv:1905.02081*.

[65] Oncken, J. and B. Chen (2020). Real-time model predictive powertrain control for a connected plug-in hybrid electric vehicle. *IEEE Transactions on Vehicular Technology 69*(8), 8420–8432.

[66] Park, H. (2013). A design of air flow configuration for cooling lithium ion battery in hybrid electric vehicles. *Journal of power sources 239*, 30–36.

[67] Pino, F. J., D. Marcos, C. Bordons, and F. Rosa (2015). Car air-conditioning considerations on hydrogen consumption in fuel cell and driving limitations. *international journal of hydrogen energy 40*(35), 11696–11703.

[68] Qi, Z. (2014). Advances on air conditioning and heat pump system in electric vehicles–a review. *Renewable and Sustainable Energy Reviews 38*, 754–764.

[69] Qu, Z., W. Li, and W. Tao (2014). Numerical model of the passive thermal management system for high-power lithium ion battery by using porous metal foam saturated with phase change material. *International Journal of Hydrogen Energy 39*(8), 3904–3913.

[70] Rabinowitz, A., F. M. Araghi, T. Gaikwad, Z. D. Asher, and T. H. Bradley (2021). Development and evaluation of velocity predictive optimal energy management strategies in intelligent and connected hybrid electric vehicles. *Energies 14*(18), 5713.

[71] Rakthanmanon, T., B. Campana, A. Mueen, G. Batista, B. Westover, Q. Zhu, J. Zakaria, and E. Keogh (2012). Searching and mining trillions of time series subsequences under dynamic time warping. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 262–270.

[72] Rama, N., H. Wang, J. Orlando, D. Robinette, and B. Chen (2019). Route-optimized energy management of connected and automated multi-mode plug-in hybrid electric vehicle using dynamic programming. *Society of Automotive Engineers Technical Paper Series 1*.

[73] Rao, L. and J. Newman (1997). Heat-generation rate and general energy balance for insertion battery systems. *Journal of the Electrochemical Society 144*(8), 2697–2704.

[74] Sato, M.-A. (2001). Online model selection based on the variational bayes. *Neural computation 13*(7), 1649–1681.

[75] Saw, L. H., H. M. Poon, H. San Thiam, Z. Cai, W. T. Chong, N. A. Pambudi, and Y. J. King (2018). Novel thermal management system using mist cooling for lithium-ion battery packs. *Applied energy 223*, 146–158.

[76] Saxén, H., C. Gao, and Z. Gao (2012). Data-driven time discrete models for dynamic prediction of the hot metal silicon content in the blast furnace—a review. *IEEE transactions on industrial informatics 9*(4), 2213–2225.

[77] Shahid, S. and M. Agelin-Chaab (2017). Analysis of cooling effectiveness and temperature uniformity in a battery pack for cylindrical batteries. *Energies 10*(8), 1157.

[78] Sharma, A., V. V. Tyagi, C. Chen, and D. Buddhi (2009). Review on thermal energy storage with phase change materials and applications. *Renewable and Sustainable energy reviews 13*(2), 318–345.

[79] Shen, M. and Q. Gao (2020). System simulation on refrigerant-based battery thermal management technology for electric vehicles. *Energy Conversion and Management 203*, 112176.

[80] Sóbester, A., S. J. Leary, and A. J. Keane (2004). A parallel updating scheme for approximating and optimizing high fidelity computer simulations. *Structural and multidisciplinary optimization 27*(5), 371–383.

[81] Song, X., L. Lv, J. Li, W. Sun, and J. Zhang (2018). An advanced and robust ensemble surrogate model: Extended adaptive hybrid functions. *Journal of Mechanical Design 140*(4), 041402.

[82] Sun, C., X. Hu, S. J. Moura, and F. Sun (2014). Velocity predictors for predictive energy management in hybrid electric vehicles. *IEEE Transactions on Control Systems Technology 23*(3), 1197–1204.

[83] Tao, X. and J. Wagner (2016). A thermal management system for the battery pack of a hybrid electric vehicle: modeling and control. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 230*(2), 190–201.

[84] Torre-Bastida, A. I., J. Del Ser, I. Laña, M. Ilardia, M. N. Bilbao, and S. Campos-Cordobés (2018). Big data for transportation and mobility: recent advances, trends and challenges. *IET Intelligent Transport Systems 12*(8), 742–755.

[85] Vatanparvar, K. and M. A. Al Faruque (2016). Otem: Optimized thermal and energy management for hybrid electrical energy storage in electric vehicles. In *2016 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 19–24. IEEE.

[86] Vlahogianni, E. I., M. G. Karlaftis, and J. C. Golias (2014). Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies 43*, 3–19.

[87] Wang, H., F. He, and L. Ma (2016). Experimental and modeling study of controller-based thermal management of battery modules under dynamic loads. *International Journal of Heat and Mass Transfer 103*, 154–164.

[88] Wang, J. (2010). Pressure drop and flow distribution in parallel-channel configurations of fuel cells: Z-type arrangement. *International Journal of Hydrogen Energy 35*(11), 5498–5509.

[89] Wang, J. and Q. Shi (2013). Short-term traffic speed forecasting hybrid model based on chaos–wavelet analysis-support vector machine theory. *Transportation Research Part C: Emerging Technologies 27*, 219–232.

[90] Wang, T., K. Tseng, J. Zhao, and Z. Wei (2014). Thermal investigation of lithium-ion battery module with different cell arrangement structures and forced air-cooling strategies. *Applied energy 134*, 229–238.

[91] Wang, X., M. Li, Y. Liu, W. Sun, X. Song, and J. Zhang (2017). Surrogate based multidisciplinary design optimization of lithium-ion battery thermal management system in electric vehicles. *Structural and Multidisciplinary Optimization 56*(6), 1555–1570.

[92] Wang, X., Y. Liu, W. Sun, X. Song, and J. Zhang (2018). Multidisciplinary and multifidelity design optimization of electric vehicle battery thermal management system. *Journal of Mechanical Design 140*(9).

[93] Wu, Y. and H. Tan (2016). Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework. *arXiv preprint arXiv:1612.01022*.

[94] Xun, J., R. Liu, and K. Jiao (2013). Numerical and analytical modeling of lithium ion battery thermal behaviors with different cooling designs. *Journal of Power Sources 233*, 47–61.

[95] Yang, N., X. Zhang, G. Li, and D. Hua (2015). Assessment of the forced air-cooling performance for cylindrical lithium-ion battery packs: A comparative analysis between aligned and staggered cell arrangements. *Applied thermal engineering 80*, 55–65.

[96] Yang, X.-G., T. Liu, and C.-Y. Wang (2021). Thermally modulated lithium iron phosphate batteries for mass-market electric vehicles. *Nature Energy 6*(2), 176–185.

[97] Yang, Z., K. Li, and A. Foley (2015). Computational scheduling methods for integrating plug-in electric vehicles with power systems: A review. *Renewable and Sustainable Energy Reviews 51*, 396–416.

[98] Yang, Z., D. Patil, and B. Fahimi (2018). Online estimation of capacity fade and power fade of lithium-ion batteries based on input–output response technique. *IEEE Transactions on Transportation Electrification 4*(1), 147–156.

[99] Yang, Z., D. Patil, and B. Fahimi (2019). Electrothermal modeling of lithium-ion batteries for electric vehicles. *IEEE Transactions on Vehicular Technology 68*(1), 170–179.

[100] Ye, Q., W. Y. Szeto, and S. C. Wong (2012). Short-term traffic speed forecasting based on data recorded at irregular intervals. *IEEE Transactions on Intelligent Transportation Systems 13*(4), 1727–1737.

[101] Yoneda, K., A. Kuramoto, N. Suganuma, T. Asaka, M. Aldibaja, and R. Yanase (2020). Robust traffic light and arrow detection using digital map with spatial prior information for automated driving. *Sensors 20*(4), 1181.

[102] Yu, H., Z. Wu, S. Wang, Y. Wang, and X. Ma (2017). Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. *Sensors 17*(7), 1501.

[103] Zhang, F., J. Xi, and R. Langari (2016). Real-time energy management strategy based on velocity forecasts using v2v and v2i communications. *IEEE Transactions on Intelligent Transportation Systems 18*(2), 416–430.

[104] Zhang, J., S. Chowdhury, and A. Messac (2012). An adaptive hybrid surrogate model. *Structural and Multidisciplinary Optimization 46*(2), 223–238.

[105] Zhang, Z., W. Li, J. Shi, and J. Chen (2016). A study on electric vehicle heat pump systems in cold climates. *Energies 9*(11), 881.

[106] Zhao, L., Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li (2019). T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*.

[107] Zhao, Z., W. Chen, X. Wu, P. C. Chen, and J. Liu (2017). Lstm network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems 11*(2), 68–75.

[108] Zhou, Y., A. Ravey, and M.-C. Péra (2019). A survey on driving prediction techniques for predictive energy management of plug-in hybrid electric vehicles. *Journal of Power Sources 412*, 480–495.

[109] Zhou, Y., A. Ravey, and M.-C. Pera (2019). A velocity prediction method based on self-learning multi-step markov chain. In *IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, Volume 1, pp. 2598–2603. IEEE.

BIOGRAPHICAL SKETCH

Yuanzhi Liu received a BS degree in Power Engineering from Wuhan University, Wuhan, China, in 2012. He is currently a PhD candidate in the Department of Mechanical Engineering at The University of Texas at Dallas. His research interests include multidisciplinary surrogate-based optimization, battery thermal and energy management, and vehicle velocity data analytics and forecasting. He has published multiple papers and received two best paper awards.

CURRICULUM VITAE

Yuanzhi Liu

November 05, 2021

Contact Information:

Department of Computer Science
The University of Texas at Dallas
800 W. Campbell Rd.
Richardson, TX 75080-3021, U.S.A.

Email: yuanzhi.liu@utdallas.edu

Educational History:

B.S., Power Engineering, Wuhan University, 2012
Ph.D., Mechanical Engineering, The University of Texas at Dallas, 2021

Employment History:

Engineer, China General Nuclear Power Group (CGN), China, July 2012 – July 2016

Professional Recognitions and Honors:

Outstanding Student Award, The Department of Mechanical Engineering, UT Dallas, 2021
Betty & Gifford Johnson Travel Award, UT Dallas, 2021
Paper of Distinctions, ASME-DAC, 2021
Paper of Distinctions, ASME-DAC, 2020
CGN Training Scholarship, CGN & Wuhan University, 2012
Outstanding Student Scholarship, Wuhan University, 2011
Outstanding Student Scholarship, Wuhan University, 2010

Patent:

Zhang, Jie, Liu, Yuanzhi, "J-type Air-Cooled Battery Thermal Management System and Method", Utility patent, Application No.: 16/996,475, Filed date: 08/18/2019. Publication NO. US 2021/0046869 A1, Publication date: 02/18/2021.

Recent Publications:

Liu, Y., Zhang, J., Electric Vehicle Battery Thermal and Cabin Climate Management Based on Model Predictive Control, *Journal of Mechanical Design*, 2020, pp. 1-11.

Liu, Y., Zhang, J. , Self-adapting J-type air-based battery thermal management system via model predictive control, *Applied Energy*, Vol. 263, 2020, pp. 114640.

Liu, Y., Zhang, J., Design A J-type Air-based Battery Thermal Management System through Surrogate-based Optimization, *Applied Energy*, Vol. 252, 2019, pp. 113426.

Liu, Y., Zhang, J. A Repeated Commuting Driving Cycle Dataset with Application to Short-term Vehicle Velocity Forecasting, Journal of Autonomous Vehicles and Systems, 2021.

He, L., **Liu, Y**., Zhang, J., Peer-to-Peer Energy Sharing with Battery Storage: Energy Pawn in Smart Grid, *Applied Energy*, Vol.297, 2021, pp. 117129.

Feng, C., **Liu, Y.**, Zhang, J., A Taxonomical Review on Recent Artificial Intelligence Applications to Solar Photovoltaic SystemGrid Integration, *International Journal of Electrical Power and Energy Systems*, Vol. 132, 2021, pp. 107176.

Li, M., **Liu, Y.**, Wang, X. and Zhang, J., , Modeling and Optimization of An Enhanced Battery Thermal Management Systemin Electric Vehicles, *Frontiers of Mechanical Engineering*, Vol. 14, Issue 1, 2019, pp. 65-75.

Wang, X., **Liu, Y.**, Sun, W., Song, X. and Zhang, J., , Multidisciplinary and Multifidelity Design Optimization of ElectricVehicle Battery Thermal Management System, *Journal of Mechanical Design*, Vol. 140, Issue 9, 2018, pp. 094501.

Wang, X., Li, M., **Liu, Y.**, Sun, W., Song, X. and Zhang, J., , Surrogate based Multidisciplinary Design Optimization of Lithium-ion Battery Thermal Management System in Electric Vehicles, *Structural and Multidisciplinary Optimization*, Vol. 56,Issue 6, 2017, pp. 1555-1570.

He, L., **Liu, Y.**, Zhang, J., An Occupancy-Informed Customized Price Design for Consumers: A Stackelberg Game Approach, (under review).

Conference Publications:

Liu, Y., Zhang, J. Short-term Vehicle Velocity Forecasting Using Dynamic Model Selection. *ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection. (**Papers of Distinction**, ASME-DAC-2021)

Liu, Y., Zhang, J. A Model Predictive Control-based Energy Management Strategy Considering Electric Vehicle Battery Thermal and Cabin Climate Control. *ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection. (**Papers of Distinction**, ASME-DAC-2020)

Liu, Y., Zhang, J. Self-Adapting Intelligent Battery Thermal Management System via Artificial Neural Network Based Model Predictive Control. *ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection.

Liu, Y., Ghassemi, P., Chowdhury, S., and Zhang, J.. Surrogate Based Multi-Objective Optimization of J-Type Battery Thermal Management System. *ASME 2018 International*

Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. V02BT03A034-V02BT03A034). American Society of Mechanical Engineers.

Liu, Y., Li, M. and Zhang, J. , An Experimental Parametric Study of Air-Based Battery Thermal Management System for Electric Vehicles. *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 56, Issue 6, 2017, pp. 1555-1570.

Li, M., Wang, X., **Liu, Y.**, Zou, Z., Cui, M. and Zhang, J. , Multidisciplinary Design Optimization of Air-based Battery Thermal Management System in Electric Vehicles, *AIAA Science and Technology Forum and Exposition*, Grapevine, Texas, January 9-13, 2017.

Book and Book Chapter:

Book Chapter: Feng, C., Sun, M., Dabbaghjamanesh, M., **Liu, Y.**, Zhang, J., "Advanced Machine Learning Applications to the Modern Power Systems", Book Title: "New Technologies for Power System Operation and Analysis," Elsevier, 2019.

Professional Services:

Journal Reviewer: IEEE Transactions in Vehicular Technology, Applied Thermal Engineering, International Journal of Energy Research, Energy Storage, Sustainable Energy Technologies and Assessments, Engineering Science and Technology, Applied Sciences, Forecasting, Data in Brief

Conference Reviewer: ASME-IDETC-2020, ASME-IDETC-2021, IEEE-IES-2021