Proceedings Paper:

Zhang, M., Wilson, J., Roger, T.J. et al. (1 more author) (2025) State of charge estimation approaches for domestic thermal storage using phase change material. In: IOP Conference Series: Materials Science and Engineering. UK Energy Storage Conference, UKES 2024, 10-12 Apr 2024, Nottingham, United Kingdom. IOP Publishing. Article no: 012005. ISSN: 1757-8981. EISSN: 1757-899X.

https://doi.org/10.1088/1757-899x/1331/1/012005

IOP Conference Series: Materials Science and Engineering

PAPER • **OPEN ACCESS**

State of charge estimation approaches for domestic thermal storage using phase change material

To cite this article: Manyu Zhang *et al* 2025 *IOP Conf. Ser.: Mater. Sci. Eng.* **1331** 012005

STATE OF CHARGE ESTIMATION APPROACHES FOR DOMESTIC THERMAL STORAGE USING PHASE CHANGE MATERIAL

Manyu Zhang[1*], James Wilson[1], Timothy J. Roger[1], Robert J. Barthorpe[1]

[1]Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK
**email address of corresponding author*
Manyu.zhang@sheffield.ac.uk

Keywords: THERMAL ENERGY STORAGE, PHASE CHANGE MATERIAL (PCM), STATE OF CHARGE (SOC) ESTIMATION, GAUSSIAN PROCESS (GP)

Abstract

This paper discusses the development and implementation of State of Charge (SoC) estimation methods for a domestic-scale Phase Change Materials (PCM) system, aimed at enhancing energy storage system management and efficiency. Addressing challenges such as the complexity of physical processes and the need for real-time algorithm operation, the study integrates physics-based insights with statistical analysis of training data to accurately predict the thermal energy stored. The estimation process relies on a historical dataset of temperature measurements collected from multiple points within the PCM system, providing a comprehensive understanding of thermal behaviour over time. The research involves an experimental program for data collection, starting with linear regression analysis of temperature data to assess thermal performance and standing loss, followed by piecewise linear approximation to characterise the transition between sensible and latent heat processes. Advanced regression techniques, specifically Gaussian Process (GP) models trained on diverse operational data, demonstrate strong predictive capability in forecasting system performance. By leveraging historical temperature profiles, this approach significantly enhances SoC estimation accuracy for PCM systems and optimises charging and discharging cycles in domestic settings. The findings contribute valuable insights and methodologies to the field of thermal energy storage, supporting the development of more efficient and intelligent energy management strategies.

1 Introduction

In the evolving energy landscape, thermal energy storage has emerged as a cornerstone for enhancing renewable energy integration and ensuring a resilient grid. By efficiently capturing and storing heat during surplus periods for later use, it bridges the gap between supply and demand, mitigates variability, and improves the overall efficiency of energy systems. Thermal storage, particularly through innovative materials and methods, plays a pivotal role in shifting towards sustainable, low-carbon energy solutions, making it indispensable in the quest for energy security and environmental sustainability [1]. Among various technologies, Phase Change Material (PCM) stands out for its efficient energy storage capacity and stable temperature control performance, marking a critical evolution in storage technology [2]. PCM is a substance utilised for latent thermal storage, relying on changes in its physical state to absorb and release heat [3, 4]. These materials store energy through the transition from solid to liquid state [5].

Realising the full potential of PCM-based systems hinges on accurately estimating their State of Charge (SoC), which is critical to the operational efficiency and reliability of energy storage systems, informing the decision-making process related to charging and discharging cycles. However, the complexity of phase change mechanisms in PCM-based systems presents unique challenges [6], and traditional in-store temperature measurements cannot capture the essence of phase change dynamics. Such measurements struggle to reflect the true energy storage levels during phase transitions, where the temperature of PCM remains almost constant despite the absorption or release of substantial heat. Additionally, external environmental temperature fluctuations further complicate the accuracy of temperature-based SoC assessments, blurring the lines between changes caused by phase transitions and those included by ambient conditions. Moreover, the intricate phase change processes, possibly involving solid-solid alongside the typical solid-liquid transformations, introduce further complexity into SoC estimation using temperature data alone. These disparity highlights the need for innovative SoC estimation methods that can navigate the complexities of PCM technology and improve energy management strategies [1, 2].

The importance of developing sophisticated, real-time SoC estimation algorithms cannot be overstated [3]. Such advancements promise not only to enhance the operational efficiency of PCM-based systems but also to facilitate proactive energy management strategies. The need for innovation is underscored by the limitations of existing methodologies, which often rely on oversimplified assumptions and fail to account for the intricacies of phase change processes [4, 5]. In response to these challenges, this paper introduces a novel approach to SoC estimation that integrates physics-based understanding with statistical analysis of empirical data. By meticulously analysing temperature data from a range of operational scenarios, this method seeks to provide a more nuanced understanding of thermal dynamics within PCM systems, thereby improving the accuracy of SoC predictions. Grounded in practical experimentation and data collection from operational systems, our methodology is designed to be applicable in real-world settings.

Our experimental programme serves as the foundation for this endeavour, designed to collect comprehensive datasets that inform the development and refinement of our estimation algorithms. Initially, the unit was charged completely and subsequently disconnected from its heat source. Following this, it was left idle for a certain duration before being completely discharged. Central to our investigation is the focus on the idle period of PCM. During this phase, a PCM unit, after being fully charged and detached from its heat source, enters a state of dormancy for a predetermined duration before its complete discharge, which is shown in Fig. 1. This interval is crucial; the difference in energy between the initial charge and the amount recovered upon discharge quantifies the standing loss. By systematically conducting tests over varied lengths of idle time, we are able to model standing loss as a direct function of the idle period's duration. Our emphasis on the idle phase stems from its inherent challenges in observability and its consequential impact on the efficiency and reliability of energy storage management. This phase is critically underexplored yet vital, as it presents a substantial opportunity to enhance SoC estimation techniques by addressing the gaps in current understanding and methodologies. The inability to directly observe or measure changes within the PCM during this quiescent phase complicates the accurate assessment of energy losses and storage dynamics. By delving into the nuances of the idle phase, our research aims to illuminate these obscured aspects of PCM behaviour, thereby enabling more precise control strategies.

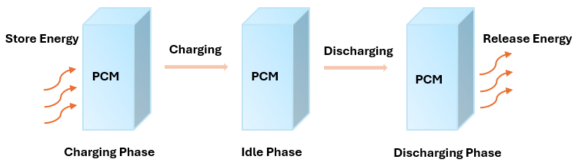


Fig. 1 Process of charging, idle and discharging in PCM

In our study, we employed piecewise linear regression analysis to assess temperature data collected from various locations within and on the surface of the PCM storage system. This initial step illuminates the system's thermal performance and the characteristics of standing loss, setting the stage for the application of more advanced analytical techniques. The transition from sensible heat storage to latent heat storage, a key feature of PCM, is critically evaluated using piecewise linear approximation. This approach provides a structured analysis of the shift in thermal storage mechanisms. This method provides an enhanced understanding of the dynamics of energy storage, essential for improving the efficiency of PCM systems. Advancing our analysis, we explore the use of Gaussian Process (GP) models, which draw upon a vast dataset covering a broad spectrum of operational conditions. The predictive power of GP models, capable of forecasting future performance, is particularly valuable for the development of predictive control strategies. These models offer a strategic advantage in managing energy storage systems by allowing for more informed and proactive decision-making. The predictive insights gained enable operators to optimise charging and discharging cycles, significantly enhancing system efficiency and reliability. Our methodical progression from piecewise linear regression to the sophisticated application of Gaussian Process models demonstrates an incremental approach to understanding PCM systems' thermal behaviour. By meticulously analysing thermal patterns and correlations, our research uncovers insights that are instrumental for the formulation of predictive control strategies. These strategies, underpinned by the detailed thermal analysis provided by our investigation, are pivotal for revolutionising the management of PCM-based energy storage systems. The value of integrating predictive control into energy management practices is underscored, highlighting its role in achieving more efficient, reliable, and proactive energy storage solutions.

In conclusion, by navigating the complex landscape of thermal energy storage with innovative tools and perspectives, this paper aims to make a significant contribution to the field. Our study not only provides a detailed methodology for SoC estimation in PCM systems, but also sets the stage for future research, highlighting new pathways for enhancing the performance and efficiency of thermal storage technologies. The insights gained from this research, particularly the advanced algorithmic approaches developed, have the potential to markedly improve the performance and efficiency of domestic-scale energy storage solutions. Through this work, we not only address the existing challenges in SoC estimation but also set a new standard for future research in the field.

2. Methodology

This section delineates the methodology employed to implement SoC estimation for domestic-scale PCM systems, addressing the challenges identified in accurately estimating SoC for efficient energy storage management. Our approach combines physics-based insights with advanced statistical analysis techniques, utilising datasets acquired from experimental setups. The methodology encompasses the implementation of piecewise linear regression and Gaussian Process (GP) models, each chosen for their unique strengths and capabilities in modelling complex systems.

Piecewise linear regression is selected for its ability to model the nonlinear behaviour of PCM systems during phase transitions effectively. This approach allows us to approximate the thermal performance and energy storage dynamics across different states of material change, providing a piecewise linear analysis. The primary advantage of this method is offering clear insights into different operational states of the PCM. However, its disadvantage lies in the potential for oversimplification, where the model may not fully capture the intricate dynamics of phase changes in all scenarios. On the other hand, GP models are chosen for their flexibility and robustness in capturing the complex and stochastic nature of thermal processes within PCM systems. The main advantage of GP models is their ability to adapt to unknown functions and their capability to provide a detailed uncertainty measure. However, their complexity and computational intensity can be seen as drawbacks, potentially limiting their scalability to large or real-time applications.

These approaches are expected to offer a comprehensive understanding of PCM thermal behaviour, enhancing SoC estimation accuracy and informing effective energy management and predictive control strategies.

2.1 Data Collection and Processing

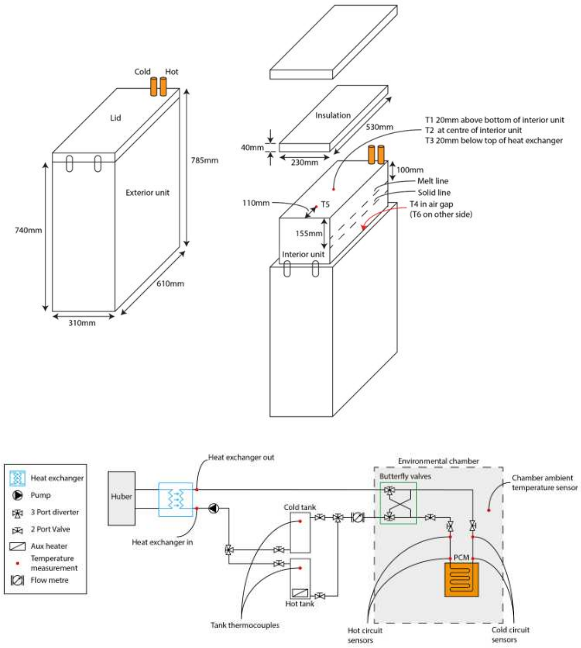


Fig. 2 Diagram of the locations of the temperature sensors and a detailed sketch of the PCM unit

The data recorded for each test included time stamps, the flow rate through the rig, and series of temperature measurements, with sensor locations illustrated in Fig. 2. These measurements encompassed the environmental chamber temperature, temperature of hot and cold water tanks, inlet and outlet temperatures at the Huber heat exchanger, and the PCM unit's inlet and outlet temperatures. Additionally, thermocouples were installed on the unit to capture localised temperature variations. The maximum storage capacity of the PCM unit is 4.5 kWh.

For all tests, the PCM unit was heated to 70℃ during the charging phase and cooled to 20℃ during discharge. The criteria for determining a fully charged or discharged state were based on the qualitative observation that the internal and outlet temperature sensors reached a steady state around the target charge or discharge temperature. Although the ambient temperature was not actively controlled during testing, it was monitored and recorded. Observations indicated that ambient temperature remained stable throughout the tests, with minor fluctuations of 2-3℃, considered negligible in influencing the results presented in this section.

Our experiments were designed to simulate various operational conditions, including charging, discharging, and idle periods. Temperature data were collected from multiple points within and on the surface of the PCM storage system, forming the basis for developing and testing algorithmic approaches. Preprocessing steps, such as data cleaning and normalisation ensured that the dataset accurately reflected the system's operational dynamics. A total of 10 tests (Test 1 to Test 10) were conducted, each differing in idle period duration, as detailed in Table 1. Temperature sensors (T1–T6) were strategically placed to capture a comprehensive thermal profile:

T1: 20mm above the bottom of the interior unit; T2: At the centre of the interior unit; T3: 20mm below the top of the heat exchanger; T4: In the air gap; T5: 100mm from the back of the interior unit and T6: On the opposite side.

Together, these measurements provided a detailed map of temperature distribution and dynamics, enabling a thorough analysis of the PCM system's thermal behaviour.

Table 1 Number of tests and idle period of experiments

Test no.	Idle period (hours:mins:sec)
Test 1	14:48:00
Test 2	0:03:00
Test 3	18:37:00
Test 4	19:10:00
Test 5	116:26:00
Test 6	163:52:46
Test 7	16:10:00
Test 8	69:27:00
Test 9	44:00:00
Test 10	91:58:00

2.2 Piecewise Linear Regression Model

Piecewise linear regression is a form of linear regression that allows for the prediction of a dependent variable (y) based on the value of an independent variable (x), with the relationship between the two variables being represented by different linear segments within different intervals of the independent variable. This method is particularly useful when the relationship between the variables changes at certain points, which can't be accurately modelled using a single linear regression line across the entire range of data [6].

The piecewise linear regression can be defined as a series of linear equations, each applicable to a specific interval of x. For k segments, the general model can be described as in Equation (1):

$$\begin{cases} y_1 = \beta_{01} + \beta_{11}x & for\ x\,\epsilon[a_1, b) \\ y_2 = \beta_{02} + \beta_{12}x & for\ x\,\epsilon[a_2, b) \\ y_k = \beta_{0k} + \beta_{1k}x & for\ x\,\epsilon[a_k, b) \end{cases} \tag{1}$$

Where y is the dependent variable, x is the independent variable, β_{01} is the intercept of the i-th segment, β_{11} is the slope of the i-th segment, a_i and b_i are the lower and upper bounds, respectively, of the i-th interval of x.

The initial phase of our analysis utilises piecewise linear regression to model standing loss behaviour during idle periods at different temperatures in a PCM storage system. This approach is chosen for the differentiation of system behaviour across various operational phases, particularly during the transition between sensible and latent heat storage processes. To assess the predictive performance and fit of our regression models, two critical statistical measures are utilised: the coefficient of determination R^2, and the Mean Squared Error (MSE). The R^2 metric is integral to regression analysis, signifying the proportion of variance in the dependent variable explained by the independent variable, with a value of 1 denoting a perfect predictive model with no residual variability. Meanwhile, MSE offers a measure of the average squared differences between observed and predicted values, with lower values indicating a closer fit to the data points. In the context of our piecewise linear regression, these metrics served as a testament to the model's ability not only to delineate the system's behaviour during various phases but also to reflect the accuracy with which it captures the underlying temperature patterns, as evidenced by the high R^2 values and low MSEs across different temperature datasets.

2.3 Gaussian Process Regression

Gaussian process regression (GPR), is a powerful and flexible non-parametric statistical modelling technique used for supervised learning tasks, most commonly for regression and sometimes for classification [7]. Rooted in Bayesian statistics, GPR offers a

IOP Conf. Series: Materials Science and Engineering **1331** (2025) 012005 doi:10.1088/1757-899X/1331/1/012005

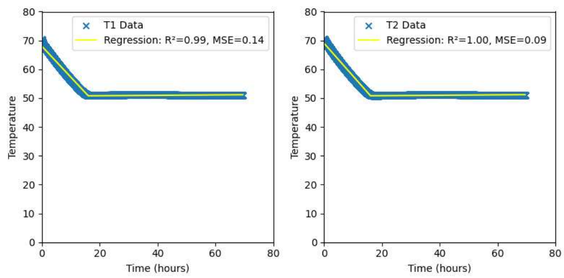
probabilistic approach to prediction, providing not only estimates of the mean of the predicted values but also their uncertainties (variances). Instead of inferring a distribution over the parameters of a parametric function, Gaussian processes can be used to infer a distribution over the function of interest directly. A GP defines a prior function, which is converted to a posterior function after having observed some values from the prior distribution. This makes GPR particularly appealing for tasks where understanding the confidence in predictions is crucial, such as in optimisation, scientific experiments, and machine learning.

A critical step in GP modelling for PCM systems involves selecting an appropriate covariance function, or kernel, based on the temperature data characteristics and the physical processes at play within the PCM system. This selection underpins the model's ability to capture the complex dynamics governing system behaviour accurately. Once the kernel is chosen, the GP model undergoes a rigorous training process using the gathered datasets, during which its parameters are fine-tuned to maximise the likelihood of the observed data. This parameter adjustment is crucial for ensuring that the kernel parameters accurately reflect the underlying process dynamics. After training, the GP model is equipped to make predictions on new, unseen data, providing not only the expected temperature values but also confidence intervals that shed light on the reliability of these predictions. This integrated approach combines careful kernel selection, meticulous model training, and insightful prediction capabilities, offering a robust framework for understanding and forecasting the performance of PCM systems.

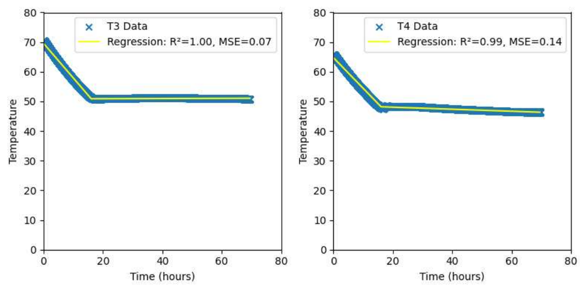
Applying GPR to model the round-trip efficiency during idle periods across various tests has proven to be a highly effective method for capturing the nuanced behaviour of energy storage systems. Utilising GPR enables not only an accurate fitting of the efficiency curves but also a quantification of the uncertainty associated with these fits. This approach makes a detailed analysis of how the round-trip efficiency behaves under different conditions, accounting for the complex interactions within the system that traditional models might overlook. The inherent flexibility of GPR, with its capacity to incorporate a range of kernel functions tailored to the specific dynamics observed in the efficiency data, allows for a more sophisticated understanding of the efficiency trends. Moreover, the provision of confidence intervals alongside the efficiency predictions offers invaluable insights into the predictability and reliability of the system's performance during the idle period, facilitating more informed decision-making regarding system operation and optimisation.

3 Results

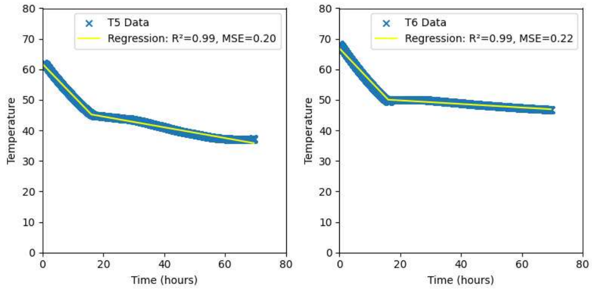
The use of piecewise linear regression to analyse the behaviour of T1-T6 during idle periods and the application of a Gaussian process regression model to evaluate round-trip efficiency across various datasets are discussed in this section.



a



b



c

Fig. 3 Piecewise linear regression analysis for temperature (T1-T6) in different locations of PCM (a) T1 and T2 (b) T3 and T4 (c) T5 and T6

A series of piecewise linear regression analyses is illustrated in Fig. 3, applied to the temperature data from Test 8, which corresponds to samples T1 through T6. In each plot, the blue line with marked 'x' points depicts the experimental temperature data over the idle period, measured in hours. According to the positional information mentioned earlier, T1-T3 behave as expected during the idle period, maintaining an almost horizontal state. This observation suggests stability in temperature readings for these samples, indicating no significant changes or activities during this phase. The regression line is shown in yellow, demonstrating how the piecewise linear model approximates the changes in the temperature over the duration of the experiment. The breakpoint in the model is set at hour 16. The models boast high coefficients of determination R^2 across all tests, ranging from 0.99 to 1.00, indicating that the piecewise linear regression closely fits the experimental data. The Mean Square Error (MSE) values in the analysis serve as indicators of the average squared differences between the observed data points and the predictions made by the model. These values are considerably low, with a range starting from 0.07 and going up to 0.22. The low MSE scores suggest that the predictions of the piecewise linear regression model are accurate, and the error margin is minimal. The level of accuracy in the model's predictions emphasises its effectiveness in capturing the true relationship between time and the recorded temperature data in the experiments. The nearly perfect R^2 values alongside low MSE scores reflect the efficacy of the piecewise linear regression model in capturing the relationship between time and temperature in this dataset. Each segment of the piecewise regression delineated a distinct phase in temperature change, reflecting potential shifts in experimental conditions or system responses over time. The consistent high-quality fit across multiple datasets, T1-T6, indicates that the observed temperature patterns are systematic and potentially predictable, which can be critical for understanding and controlling the processes being tested.

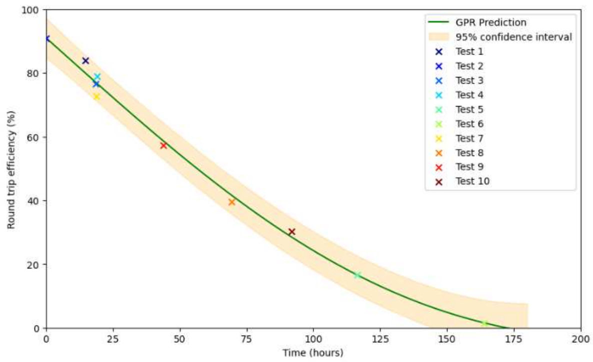


Fig. 4 Round-trip efficiency versus idle periods

The heat energy involved in the system's charging and discharging processes is determined by Equation (2), and the round-trip efficiency predicted by the GPR is calculated using Equation (3).

$$Q = C_p * \dot{m} * \Delta T \tag{2}$$

$$Round\ trip\ efficiency = \frac{Discharging\ energy}{Charging\ energy} * 100\% \tag{3}$$

Where Q is the power, $\dot{m}$ is the mass of the water, C_p is the heat capacity of the water and ΔT is the difference between the inlet and outlet temperature.
Fig. 4 presents a GPR analysis of round-trip efficiency during idle periods across 10 different tests. The central curve, presenting the GPR prediction, indicates a decreasing trend in efficiency as time progresses, suggesting a system's declining performance over the observation period. Surrounding this prediction is a shaded region delineating the 95% confidence interval, which notably broadens with time, reflecting increasing uncertainty in the model's predictions as one moves further out in time. The individual tests, marked by distinct symbols and colours, generally cluster close to the GPR prediction line, implying that the GPR model provides a reasonable fit to the observed data. Notably, there are no significant outliers, and most of the observed efficiency values from the tests fall within the prediction confidence interval. This close fit suggests that the GPR model captures the underlying trend effectively. The efficiency values range from nearly 100% to less than 20%, with the higher efficiencies observed during tests with short standing periods (0-20 hours), and the lower efficiencies recorded in tests with longer standing periods (Tests 5 and Test 6, approaching 200 hours). This gradient of efficiency underscores the system's diminishing ability to retain energy during the charging and discharging processes over longer standing periods, as defined by Equation (3).
After training, the GP model takes real-time temperature measurements from multiple points in the PCM system as input. Using its trained covariance function, it predicts the thermal energy stored and outputs an estimate of the SoC along with confidence intervals. This approach effectively models complex phase change behaviours, improving SoC accuracy beyond simple temperature-based estimation methods.

Additionally, the underlying heat transfer Equation (2), potentially informs the GPR model's interpretation, relating the physical process of energy loss to the thermal properties of the material, the mass, and the temperature difference. The graph communicates a clear narrative of degradation over time, and such data could be crucial for planning maintenance schedules or predicting the lifespan of the system under test.

4 Conclusion

In conclusion, this study presents an advanced State of Charge (SoC) estimation methodology for PCM-based thermal energy storage systems, integrating physics-based modelling with statistical techniques to enhance accuracy and efficiency. The research demonstrates that Gaussian Process (GP) models outperform traditional approaches by effectively capturing complex phase change dynamics and providing uncertainty quantification, leading to more reliable SoC predictions. Additionally, piecewise linear regression successfully models the transition between sensible and latent heat storage, allowing for a structured understanding of standing loss behaviour. Experimental validation confirms the robustness of these methods, with high predictive accuracy (R^2 values between 0.99 and 1.00) and low Mean Squared Errors (MSE), reinforcing their applicability to real-world thermal storage systems. Furthermore, results indicate that round-trip efficiency declines over extended idle periods, with the GP model effectively forecasting this trend. The study also highlights the importance of multi-point temperature measurements in improving SoC estimation accuracy. While the proposed methodology enhances predictive performance, potential challenges such as model overfitting and environmental variability remain, necessitating further refinement. Future work will focus on expanding datasets, improving model generalization, and exploring real-time implementation to optimize energy management strategies. Overall, this research provides a novel framework for predictive thermal storage control, contributing to the broader goal of sustainable and intelligent energy systems.

5 References

[1] A. Sarı, H. Sarı, and A. Önal: 'Thermal properties and thermal reliability of eutectic mixtures of some fatty acids as latent heat storage materials, ' Energy Conversion and Management, 2004, 45, (3), pp. 365-376
[2] L. F. Cabeza, A. Castell, C. Barreneche, A. de Gracia, and A. I. Fernández: 'Materials used as PCM in thermal energy storage in buildings: A review, ' Renewable and Sustainable Energy Reviews, 2011,15, (3), pp. 1675-1695
[3] A. de Gracia and L. F. Cabeza: 'Phase change materials and thermal energy storage for buildings, ' Energy and Buildings, 2015, 103, pp. 414-419
[4] B. Cárdenas and N. León: 'High temperature latent heat thermal energy storage: Phase change materials, design considerations and performance enhancement techniques, ' Renewable and Sustainable Energy Reviews, 2013, 27, pp. 724-737
[5] M. F. Demirbas: 'Thermal Energy Storage and Phase Change Materials: An Overview, ' Energy Sources, Part B: Economics, Planning, and Policy, 2006, 1, (1), pp. 85-95
[6] M. A. Hannan, M. S. H. Lipu, A. Hussain, and A. Mohamed: 'A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations, ' Renewable and Sustainable Energy Reviews, 2017, 78, pp. 834-854
[7] G. Zsembinszki, C. Orozco, J. Gasia, T. Barz, J. Emhofer, and L. F. Cabeza: 'Evaluation of the State of Charge of a Solid/Liquid Phase Change Material in a Thermal Energy Storage Tank, ' Energies, 2020, 13, (6)

[8] D. N. T. How, M. A. Hannan, M. S. Hossain Lipu, and P. J. Ker: 'State of Charge Estimation for Lithium-Ion Batteries Using Model-Based and Data-Driven Methods: A Review, ' IEEE Access, 2019, 7, pp. 136116-136136
[9] X. Hu, F. Feng, K. Liu, L. Zhang, J. Xie, and B. Liu: 'State estimation for advanced battery management: Key challenges and future trends, ' Renewable and Sustainable Energy Reviews, 2019, 114
[10] M. S. Hossain Lipu et al.: 'Data-driven state of charge estimation of lithium-ion batteries: Algorithms, implementation factors, limitations and future trends, ' Journal of Cleaner Production, 2020, 277
[11] Z. Onder, A. Degirmenci, and O. Karal: 'Estimating Breakpoints in Piecewise Linear Regression Using Machine Learning Methods, ' presented at the 2022 Innovations in Intelligent Systems and Applications Conference (ASYU), 2022.
[12] E. Schulz, M. Speekenbrink, and A. Krause: 'A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions, ' Journal of Mathematical Psychology, 2018, 85, pp. 1-16