OPEN ACCESS

Energy Storage PERSPECTIVE

Machine learning-driven paradigm shift in thermochemical energy storage

Wei Li,[1,2,*] Siyi Wang,[1] Shengguan Xu,[1] and Qiuwang Wang[2]

[1]School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 211816, China

[2]Key Laboratory of Thermo–Fluid Science and Engineering, Ministry of Education, Xi'an Jiaotong University, Xi'an 710049, China

*Correspondence: wli@njtech.edu.cn

Received: April 15, 2025; Accepted: May 26, 2025; Published Online: January 9, 2026; https://doi.org/10.59717/j.xinn-energy.2026.100127

Citation: Li W., Wang S., Xu S., et al. (2025). Machine learning-driven paradigm shift in thermochemical energy storage. *The Innovation Energy* **3**:100127.

Thermochemical energy storage (TCES) technology, which enables high-density energy storage and flexible dispatch through reversible chemical reactions, is a critical solution to address the spatiotemporal mismatch of renewable energy. However, TCES technology has been difficult to widely diffuse because of the difficulty in discovering TCES materials with even better properties and the difficulty of TCES systems in accurately predicting the dynamic heat load demand of users in smart energy systems. Traditional research methods relying on experimental trial and error and simplified numerical simulations face challenges such as high costs, prolonged experimental cycles, and insufficient prediction accuracy. It is imperative that novel research tools be employed as soon as possible to address these two major issues. Machine Learning (ML), as a data-driven intelligent tool, transforms the exploration of causal relationships into correlational analyses, demonstrating exceptional potential in TCES for novel material screening and system optimization due to its advanced data processing capabilities, low hardware cost requirements, and high prediction accuracy. This perspective intensively describes the methodology for embedding ML algorithms in TCES, as well as describing the characteristics of the various algorithms. Furthermore, it outlines strategies to enhance the application of ML algorithms in TCES technologies, proposing pathways such as establishing standardized TCES databases through extensive data collection to broaden ML applicability, implementing weighted parameter optimization for input variables or combining physical theory in the input process of the underlying data to improve algorithmic stability, integrating transfer learning with ML frameworks to maintain accuracy under data scarcity, and training ML models with region-specific operational data to explore multi-dimensional and multi-objective optimization across the 4E criteria (Energy, Environment, Economic, Exergy) and interdisciplinary convergence. While current applications of ML in the TCES domain remain in their nascent stages, the technology holds significant potential to drive transformative advancements in future research and industrial deployment.

INTRODUCTION

As the global energy crisis intensifies, developing renewable and sustainable energy sources has become imperative to mitigate energy shortages and global warming. Solar energy, industrial waste heat, and off-peak electricity are considered promising alternatives due to their abundance and environmental benefits. However, their intermittency, geographical constraints, and temporal variability lead to mismatches with user demands, which conventional energy utilization methods fail to address efficiently.[1] Thermochemical Energy Storage (TCES) technology, which achieves high-density energy storage and controlled release through reversible chemical reactions, emerges as a core solution to resolve these spatiotemporal mismatches. At present, the wide application of the TCES system faces two problems: one is the screening of more promising new materials, and the other is that, in terms of energy supply and demand, the model predictive control strategy of the TCES system is difficult to satisfy the dynamic of heat load of the building. Therefore, it is imperative to adopt more flexible research tools to screen materials with better apparent properties and build more resilient systems with accurate energy prediction at less cost.

Currently, research methodologies for TCES technologies primarily fall into two categories: experimental studies and numerical simulations. Experimental studies involve constructing thermal storage devices and conducting real-time monitoring to accurately capture material thermodynamic behaviors and heat charge/discharge characteristics, such as thermal capacity and cyclical stability. However, their effort-intensive, time-consuming, and reliance on specialized equipment limit large-scale technological validation. Numerical simulations, based on multi-physics coupled models including heat transfer, mass transfer, and energy conversion, enable rapid prediction of system performance and parameter optimization. Yet, their strong assumptions regarding boundary conditions and material property parameters may lead to deviations from real-world operating conditions. As a result, the potential and advantages of TCES to save energy or operational costs fail to be fully utilized.

The rise of artificial intelligence has significantly advanced the modeling of complex phenomena and reduced computational costs.[2] Machine Learning (ML), as a key branch of AI, can automatically extract features and patterns from data to construct nonlinear mapping models. In contrast to experimental studies and numerical simulation methods that rely on specific physical laws and require extensive computational resources, ML provides a novel, entirely data-driven approach to accurate prediction, optimization, and real-time control of complex TCES systems. In a case study, Scapino et al.[3] developed two types of artificial Neural Network models to predict the performance of a sorption TCES system, such as the state of charge, outlet temperature, and thermal power output of a sorption storage reactor. Moreover, their proposed algorithm evaluated 32 distinct configurations of hydration and dehydration processes to determine the optimal configuration for each system architecture and operating mode, achieving prediction errors of less than 5% compared to experimental measurements. They found that the Artificial Neural Network is much faster in dynamic systems analysis than physics-based models. By leveraging data-driven methodologies, ML transforms the exploration of causal relationships into correlational analyses, shifting the focus from understanding "what" to uncovering "why" in problem-solving frameworks.[4]

There have been a number of recent studies integrating ML into TCES technology. However, a systematic synthesis of ML's multifunctional capabilities in addressing the complexities of TCES technologies remains absent, impeding researchers' efforts to advance subsequent exploration. Given that current TCES research primarily focuses on the synthesis and characterization of novel sorbent materials, development of new operating cycles, and design of reactors and systems, This article answers the question of what potential can be realized by combining ML with TCES technology and how to combine ML with TCES technology in a deeper way. Spanning from single-objective to multi-objective optimization, bridging energy engineering with interdisciplinary convergence, and transitioning from current research landscapes to future directions, this work provides a comprehensive overview of the field.

This perspective summarizes the methodology of using ML for novel material screen and system optimization in the field of TCES, and provides a summary of the characteristics of various algorithms in ML, and points out the challenges of embedding ML in the field of TCES, further provides an enriched outlook on the future use of ML to make the TCES technology more advantageous. Despite the challenges associated with embedding ML algorithms into TCES research, it is evident that TCES technology holds significant potential to unlock transformative advancements in the future.

METHODOLOGY FOR THE APPLICATION OF ML IN THE FIELD OF TCES

Variable correlation analysis

Since ML algorithms are data-driven black-box models, correlation analysis of various known variables is necessary to improve the prediction accuracy of ML algorithms before training them. In TCES systems, such as the hydration-dehydration cycles for materials, critical variables include reaction enthalpy, temperature gradients, material porosity, and gas-phase water partial pressure, all of which exhibit complex interdependencies. Correlation Analysis can be used to quantify the degree of association between two or more variables, to determine whether the variables are linearly or nonlinearly dependent, and to measure the strength of this relationship through numerical indicators, such as the Pearson Correlation Coefficient. For instance, in TCES, the Pearson Correlation coefficient between hydration reaction rates and water vapor pressure often reveals a strong positive correlation, guiding the optimization of operating conditions. When the Pearson Correlation Coefficient between two variables is positive, it indicates that there is some kind of hidden association between these two variables. Identifying such associations helps prioritize key parameters for system design in TCES, such as reaction kinetics, cyclability. On the one hand, characteristic variables of high correlation nature can be understood as key factors that help researchers to understand the mechanisms behind them during the prediction process of black-box modeling. On the other hand, if during the process of data collection gives the finding is that the correlation between target variables or input characterization variables in the model is weak, it indicates that the quality of the data is insufficient or the key variables have been omitted, then other variables need to be introduced. Overall, the high correlation of the input variables needs to be eliminated to simplify the complexity of the model and improve the accuracy of the model predictions. For TCES, this entails refining feature selection to focus on dominant physicochemical interactions, thereby enhancing ML-driven predictions of charge-discharge dynamics or long-term cyclability.

For example, when screening TCES materials with excellent performance properties using the ML algorithm, the key material screening criteria to be considered are:[5]

- High energy storage density and low charging temperature.
- High thermal conductivity and high charge/discharge efficiency.
- Low cost and environmental friendliness.
- Excellent chemical stability and mechanical strength.

Kiyabu et al.[6] adopted the enthalpy of dehydration per mole H_2O as a key metric for training ML models to screen high-performance salt hydrates. Their correlation analysis revealed a strong negative Pearson correlation ($r = -0.75$) between cation electronegativity and dehydration enthalpy, demonstrating that low-electronegativity cations favor the formation of hydrates with superior energy storage density. This empirical relationship critically guided their feature selection strategy in subsequent ML model development.

Variable correlation analysis can help researchers to calculate the correlation of many variables within the same level, while for variables in different levels, this paper recommends classifying the variables through the 4E dimensions. For example, the variables representing economy are specific investment cost, total cost, levelized cost of storage, etc., the variables representing environment are total equivalent warming impact, annual emission reduction, etc., the variables representing energy are energy storage density, power storage density, energy efficiency, etc., and exergy efficiency, etc. for exergy. Therefore, the 4E dimensions can be referred to when inputting variables to ensure that the ML algorithm can get a more complete and comprehensive dataset as much as possible when being trained.

Suitable ML algorithm

ML algorithms are categorized into three types: supervised learning, unsupervised learning, and reinforcement learning. In TCES research, studies typically focus on material properties and system parameters, such as energy storage density, enthalpy of dehydration, cycle stability, reaction kinetics, heat and mass transfer rates, and low-cost requirements. These target variables can be directly obtained through physical computing, forming well-defined labeled datasets. Coincidentally, the core principle of supervised learning lies in utilizing well-defined labeled data to train models, enabling them to learn the mapping relationships between input features and output labels. Consequently, supervised learning algorithms, including Random Forest, Support Vector Regression, XGBoost, and Neural Networks, have been predominantly adopted in existing TCES studies.

For instance, Wei et al.[7] integrated TCES into district heating networks using ML models—Support Vector Regression, Regression Trees, and long short-term memory—to optimize charge/discharge cycles in Nottingham, UK. Their findings reveal that each ML model achieved comparable performance, with Coefficient of Variation of the Root Mean Square Error values within the 9%–11% range. Among these, Neural Networks exhibit superior capabilities in fitting complex nonlinear functions and benefit from mature technical frameworks, making them particularly prevalent in TCES applications. Colak et al.[8] used an artificial Neural Network to predict the discharging performance of sorption heat storage materials, utilizing experimental data from a laboratory-scale fixed-bed thermochemical heat storage unit. Input parameters included inlet temperature, average relative humidity, discharging air mass flow rate, and sorption rate, while outputs estimated heat output rate, exergy output rate, and energy storage density. The model achieved remarkable accuracy, with an R-value of 0.99987 and a mean squared error of 3.66×10^{-6}. The results demonstrated that Artificial Neural Networks serve as a powerful tool for determining optimal operating parameters of candidate materials. Similarly, Wang et al.[9] developed a BP Neural Network algorithm to optimize the decarbonation process in $MgO/MgCO_3$ systems, achieving prediction errors below 5% and guiding reactor design with minimal computational overhead.

Notably, in the aforementioned studies, without physical constraints, such models are prone to overfitting due to inherent data noise. To address this limitation, Praditia et al.[10] proposed Physics-Inspired Neural Networks to ensure physical plausibility and enhance the accuracy of dynamic internal state predictions in TCES systems. Unlike conventional Neural Networks, Physics-Inspired Neural Networks integrate governing physical equations into the training process, thereby improving prediction reliability even in highly dynamic and nonlinear scenarios. These hybrid models demonstrated comparable performance to purely data-driven approaches.

Model assessment

In general research, the collected data is divided into a training set and a test set. To verify that the obtained algorithms are capable, the results obtained by different ML algorithms need to be evaluated using the same quantitative metrics. There are two common metrics called Mean Square Error as well as the Coefficient of Determination were chosen to evaluate the results of ML models in different scenarios. For example, Tasneem et al.[2] used the Bayesian ridge and the k-nearest neighbors, respectively, to create the formula and make the model for the temperature of NH_4NO_3 by KOH and water temperatures, and using XGBoost can predict the future temperature of NH_4NO_3. The three algorithms can be compared well by the mean square error value. The mean square error values of the three are 16.900704, 5.355852 and 0.009042 in that order, so using XGBoost can be computed the future of the KOH.

Outlook

In summary, ML algorithms demonstrate extraordinary efficiency in predicting the performance of TCES systems and the thermochemical reaction properties of materials beyond traditional research methods. Based on existing studies, the following directions are critical to fully unlock the potential of TCES:

The stability of ML algorithms heavily depends on extensive, high-quality data, underscoring the urgency to establish comprehensive TCES databases covering material properties and system optimization metrics. For scenarios with limited experimental data, transfer learning and deep learning techniques should be leveraged to refine model performance. Despite the nascent stage of ML applications in TCES, substantial untapped potential exists.

The black-box nature of ML algorithms shifts focus from physically derived relationships between dependent and independent variables to data-driven input-output mappings. However, real data are noisy with measurement errors; thus, combining physical theory with observational data will help to improve the prediction and plausibility of the model.

The success of an algorithm in one scenario does not guarantee universal applicability. Therefore, investigating diverse TCES systems and materials will improve ML efficiency, while real-world implementation studies are essential to ensure scalability and robustness, key factors for validating ML algorithms

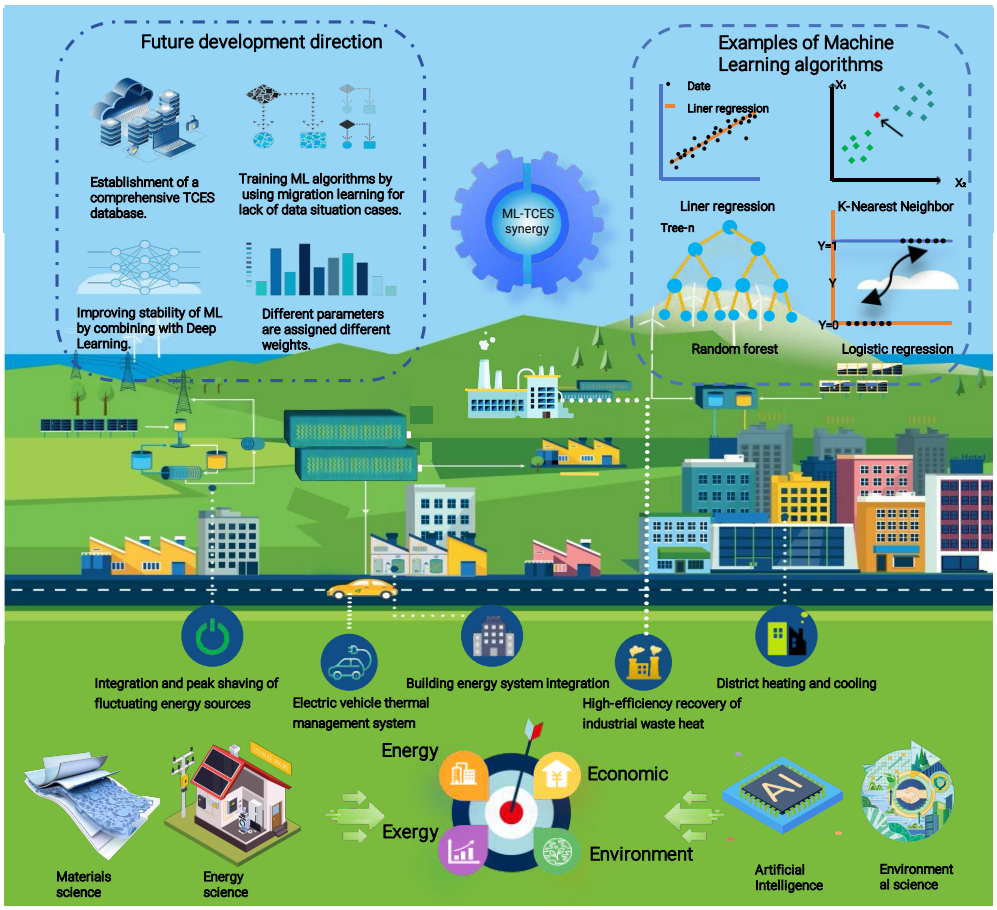

Figure 1. Schematic illustration of the steps of ML application in TCES.

in the TCES domain. Furthermore, ML's strong generalization capability for complex problems provides a versatile tool for building resilient energy networks and optimizing multi-dimensional objectives, such as economic efficiency through optimal scheduling strategies under electricity price fluctuations and environmental sustainability by balancing energy security with carbon neutrality goals.

The integration of ML with TCES technologies inherently represents an interdisciplinary convergence, demanding deep synergy of knowledge from energy science, materials science, environmental science, and artificial intelligence. Future research should adopt multidisciplinary perspectives, transitioning algorithmic optimization objectives from single-target to multi-objective frameworks, thereby achieving collaborative synergy across diverse domains. For the future, the methodology of embedding ML in TCES technology is schematically shown in Figure 1. We believe that following these directions will accelerate the integration of ML-driven TCES technologies into practical applications, unlocking their transformative potential in sustainable energy systems.

CONCLUSION

TCES technology plays an important role in renewable energy utilization due to its ultrahigh energy density, low thermal losses, and long-term storage potential. Compared to traditional experimental and numerical methods, In order to achieve more efficient screening of TCES materials with better apparent properties and optimize the TCES system to better match the dynamic thermal loade requirements, combining ML algorithms, which are tolerant of uncertainty, imprecision and partial truth, with TCES technology is a powerful solution to meet the challenges of the complex real world. In this paper, we propose a methodology that combines ML and TCES by firstly performing correlation analysis on various known variables to ensure the input variables are representative, secondly selecting an appropriate ML algorithm, and finally determining the accuracy of the data in the test set by calculating the Mean Square Error and the Coefficient of Determination. ML algorithms facilitate comprehensive multi-dimensional optimization across 4E criteria within resilient and intelligent energy networks, enabling precise prediction of thermal demand to enhance the alignment of TCES systems with dynamic thermal load requirements.

For the next step in the study, this paper proposing pathways such as establishing standardized TCES databases through extensive data collection to broaden ML applicability, implementing weighted parameter optimization for input variables or combining physical theory in the input process of the underlying data to improve algorithmic stability, integrating transfer learning with ML frameworks to maintain accuracy under data scarcity, and training ML models with region-specific operational data to explore multi-dimensional and multi-objective optimization across the 4E criteria and interdisciplinary convergence. While current applications of ML in the TCES domain remain in their nascent stages, the technology holds significant potential to drive transformative advancements in future research and industrial deployment.

REFERENCES

1. Zhang Y. and Wang R. (2020). Sorption thermal energy storage: Concept, process, applications and perspectives. *Energy Stor. Mater.* **27**:352–69. DOI:10.1016/j.ensm.2020.02.024
2. Tasneem S., Sultan H.S., Ageeli A.A., et al. (2023). Machine learning modeling of reversible thermochemical reactions applicable in energy storage systems. *J. Taiwan Ins. Chem. Eng.* **148**:104926. DOI:10.1016/j.jtice.2023.104926
3. Scapino L., Zondag H.A., Diriken J., et al. (2019). Modeling the performance of a sorption thermal energy storage reactor using artificial neural networks. *Appl. Energy* **253**:113525. DOI:10.1016/j.apenergy.2019.113525
4. Chen Z., Pan S., Wang J., et al. (2024). Machine learning will revolutionize perovskite solar cells. *The Innovation* **5**:100602. DOI:10.1016/j.xinn.2024.100602
5. Li W., Klemeš J.J., Wang Q., et al. (2022). Salt hydrate–based gas-solid thermochemical energy storage: Current progress, challenges, and perspectives. *Renew. Sustain. Energy Rev.* **154**:111846. DOI:10.1016/j.rser.2021.111846
6. Kiyabu S., Girard P. and Siegel D.J. (2022). Discovery of salt hydrates for thermal energy storage. *J. Am. Chem. Soc.* **144**:21617–27. DOI:10.1021/jacs.2c08993
7. Wei Z., Tien P.W., Calautit J., et al. (2024). Investigation of a model predictive control (MPC) strategy for seasonal thermochemical energy storage systems in district heating networks. *Appl. Energy* **376**:124164. DOI:10.1016/j.apenergy.2024.124164
8. Colak A.B., Aydin D., Al-Ghosini A., et al. (2022). Discharging performance prediction of experimentally tested sorption heat storage materials with machine learning method. *J. Energy Storage* **56**:106159. DOI:10.1016/j.est.2022.106159
9. Wang Y., Wang R., Guo Y., et al. (2024). The optimization of the MgO/MgCO3 decarbonation process and machine learning-based improved reactor design approach. *Energy* **305**:132326. DOI:10.1016/j.energy.2024.132326
10. Praditia T., Walser T., Oladyshkin S., et al. (2020). Improving thermochemical energy storage dynamics forecast with physics-inspired neural network architecture. *Energies* **13**:3873. DOI:10.3390/en13153873

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of Jiangsu Province (No. BK20230319) and the China Postdoctoral Science Foundation (2024M762590).

AUTHOR CONTRIBUTIONS

Wei Li: Conceptualization, Methodology, Investigation, Formal analysis, Writing- Original draft, Writing-Review & Editing, Project administration, Supervision; **Siyi Wang:** Investigation, Formal analysis, Software, Writing-Original & Editing; **Shengguan Xu:** Funding acquisition, Resources; **Qiuwang Wang:** Writing -Review Editing.

DECLARATION OF INTERESTS

The authors declare no competing interests.