


INNOVATION
– PRESS –

Energy Use

Energy Use 3 (2026) 100073

Contents lists available at INNOVATION PRESS

Thermodynamics-Grounded AI-Collaboration Trend in Systematically Exploring Models and Interpretably Predicting Thermophysical Properties for Working Fluids

Xiayao Peng,¹ Yiyang Jin,¹ Jian Li,² Zhen Yang,¹ and Yuanyuan Duan^{1,*}
¹Department of Energy and Power Engineering, Tsinghua University, Beijing 100084, China

²School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China

*Correspondence: yyduan@tsinghua.edu.cn

Received: September 12, 2026; Accepted: September 21, 2026; Published Online: September 23, 2026; <https://doi.org/10.59717/ipj.energy-use.2026.100073>

© 2026 The Author(s). This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Peng X., Jin Y., Li J., et al. (2026). Thermodynamics-Grounded AI-Collaboration Trend in Systematically Exploring Models and Interpretably Predicting Thermophysical Properties for Working Fluids. *Energy Use* 2:100073.

Dear Editor,

Reliable thermophysical property data of working fluids provide an essential thermodynamic foundation in energy utilization systems. Conventional regression-based thermophysical modeling methods face a tension between semi-empirical mathematical forms and rigorous physical constraints, making it difficult to balance accuracy and applicability. Artificial intelligence (AI) technologies offer distinct advantages in characterizing complex thermophysical behaviors through their systematic exploration and high-dimensional association abilities, thus their use in thermophysical modeling has grown rapidly over the past decade. This paper aims to clarify the significance and role of AI technologies in theory-intensive thermophysical research, as well as provides directions for developing highly interpretable and broadly applicable AI methods for predicting thermophysical properties of working fluids in energy utilization systems. Machine learning methods are compared in thermophysical characterization, examining their modeling mechanisms, performance advantages, and potential limitations. With supervised learning as the basic workflow, collaboration between AI and thermophysics is gradually changing the modeling paradigm through physically grounded inputs, thermodynamic laws guidance, a priori physical baseline models, post hoc explanation, and the direct construction of explicit models. But concerns always persist in the interpretability and reliability of the resulting data. Future work should embed thermodynamic criteria throughout model exploration algorithm and establish analytical links between molecular characteristics, thermophysical properties, and cycle performance. These developments could support autonomous thermophysical modeling and joint optimization of working fluids properties and operating parameters for energy use.

Working fluids are the fundamental carriers of heat-work conversion in thermodynamic cycles. Their thermophysical properties describe the existence states and energy characteristics, which serve as the thermodynamic basis for performance optimization, component configuration, and working fluid design in energy systems. For example, phase equilibria and enthalpy determine the cycle parameters, density affects equipment size, whereas viscosity and thermal conductivity related to flow resistance and heat transfer performance. Especially in novel low-carbon energy systems with limited operating experience, accurate thermophysical property data directly determine the performance reliability and operational safety.

In complex new-energy applications, thermodynamic systems operate across wider temperature ranges and under increasingly variable conditions. These applications require an increasing type and property range of new

working fluids, indicating that experimentally measuring the thermophysical properties of all the fluids and compositions is impractical. Whereas thermophysical models can integrate scattered high-precision experimental data into systematic thermodynamic descriptions. Yet a longstanding tension remains between the semi-empirical modeling approaches and the rigorous thermophysical theories developed over the past century. Most traditional modeling procedures first assume a mathematical form and then fit some empirical parameters. The process often requires both extensive exploration and a flash of insight, which may be time-consuming. Even such efforts cannot ensure that the optimal model form is found, since incomplete observations and empirical judgments are unavoidable. For new working fluids with scarce data, accurately describing their physical characteristics remains difficult, thus the resulting models may struggle to balance accuracy, applicability, and modeling cost.¹

Recent advances in artificial intelligence (AI) offer advantages in depth, breadth, and speed in exploration, providing new perspectives and systematic approaches to thermophysical modeling.² Machine learning can integrate multiple influencing factors and high-dimensional variables, to comprehensively and efficiently establish mappings between known parameters and target data, which reduces reliance on prolonged empirical trial and error by researchers. However, interpretability and reliability remain major obstacles and risk sources in theory-intensive disciplines and engineering applications. Recent research therefore seeks effective connections and complementary roles between thermodynamic theories, empirical criteria, and AI mechanisms, so that both the process and results can be analyzed and tested.

TRENDS OF AI'S FUNCTIONAL ROLES IN WORKING-FLUID THERMOPHYSICAL PROPERTIES

Studies in thermophysical property modeling by machine learning methods became increasingly common after 2018. Figure 1 shows their development and physical integration process, distinguishing several independent but complementary techniques.

Physically grounded inputs: diverse parameters representing fluid characteristics

Most related studies currently use data-driven supervised learning to obtain target thermophysical data. The conventional approach takes state parameters and basic thermophysical parameters as inputs and uses various algorithms to map them to target properties. It enables rapid estimates for complex interaction mixtures that are difficult to model accurately. Such as H₂O–CO₂–H₂ ternary-mixture, the maximum absolute relative deviation

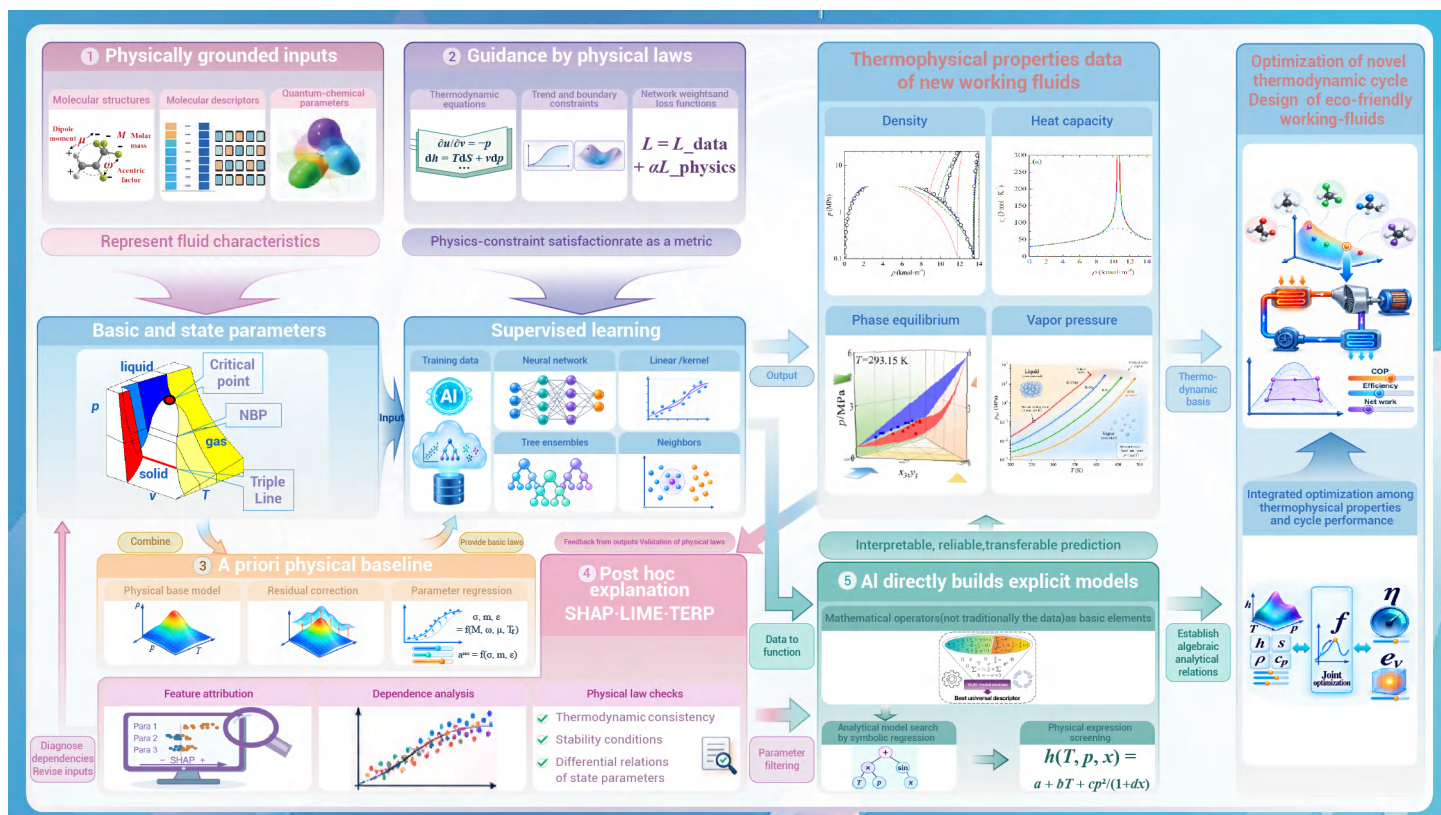


Figure 1. Physically interpretable machine learning methods in thermophysical property modeling of working fluids in energy utilization systems

from the AI obtained data to the reference data was reduced by over 90% compared with three traditional physical equations of state.³ To represent the distinct characteristics of various molecules, the atomic group, molecular descriptors, and quantum-chemical parameters are introduced as inputs, which makes molecular structure and interaction characteristics available for learning. The method enables a common correlation across a thermophysical dataset of 343 fluorine/chlorine-based refrigerants.⁴ The central to this approach is the parameter selection, which should describe microscopic interactions, be readily obtainable, and avoid inter-parameter coupling.

Process guidance by thermodynamic laws: improving the interpretability of algorithmic exploration

As for the black-box search processes in data-driven algorithms, thermodynamic laws provide an important means of improving interpretability. Neural network architectures, for example, allow physical laws to be encoded directly. The index related to the physical equation consistency can be incorporated into training loss functions, while physical relationships can also be imposed through network structure. This enables the treatment of higher-dimensional and more complex parameter relationships. In the model-fluid study, physics-guided neural networks reproduced thermodynamic differential relationships across vapor, liquid, and solid phases, and the test-set mean squared error of the compressibility factor was only 5.86×10^{-3} .⁵ Neural networks have also simultaneously predicted multiple thermophysical properties of new hydrofluorocarbons and hydrofluoroethylenes refrigerants,⁶ although these examples do not establish the embedding of physical constraints. This approach requires careful control of data quality, physical constraints, and evaluation metrics. For real fluids, complex molecular interactions, mixture behavior, and near-critical properties must be considered. The transferability of complex networks is another challenge that remains to be addressed.

A priori physical baseline models: preserving theoretical bounds

To improve the efficiency and predictive capability of AI methods, many studies start from high-performance thermophysical models, where AI is used only to obtain key model parameters or correction terms. These techniques retain the theoretical core of existing thermophysical modeling

approaches, thus improving generalization and predictive capability and reducing the dependence on experimental data. For example, taking the comprehensive equations of state as the basis, learning 5 equation parameters from 6 molecular descriptors enabled simultaneous prediction of multiple properties for 18 refrigerants over broad ranges.⁷ The central to this approach is deeply bound by the accuracy and applicability of the baseline thermophysical model.

Post hoc explanation: supporting checks of thermodynamic laws

To assess the consistency with multiscale thermophysical constraints and mathematical parameter relationships, post hoc explanation methods such as LIME, SHAP, and TERP are often added. They examine how input parameters influence the output results and assess the relative importance of the basic parameters or molecular descriptors. For the 6 molecular descriptors discussed above, post hoc analysis identified molecular weight and the surface-charge descriptor S_5 as the most important overall for predicting soft-SAFT parameters.⁷ For organic refrigerants, SHAP also revealed how atom groups contribute to critical properties and boiling points,⁴ as well as the corresponding relationship with microscopic interaction. These results provide clues for interpreting intermolecular interactions and reference information for working fluid molecular design, rather than independently verified relationships with microscopic interaction potentials. Such methods provide quantitative tools for explaining the black-box process, identifying dependencies, and selecting features. They establish exploration priorities for property relationships and targeted parameter design. Their application requires attention to the simplicity and fidelity of explanations and their sensitivity to data partitioning. Particular attention is needed to assess whether the inferred correlations are consistent with thermodynamic laws.

AI directly builds explicit models: improving reliability and transferability

Meeting the requirements of theory-intensive research for model transparency, controllable uncertainty, and physical analyzability remains a central challenge for AI. With increasing computing power, symbolic regression has recently provided a means of generating explicit models. It performs supervised learning by searching over algebraic structures, changing the conventional data-to-data approach into a data-to-explicit-function approach. This

provides an interface compatible with the thermodynamic theories and criteria that have clear algebraic characteristics. An explicit gaseous sound-speed model generated by symbolic regression achieved a relative root-mean-square deviation of only 0.3% for 17 binary working fluid mixtures. In high-density gas regions, deviations were reduced by one to two orders of magnitude compared with existing models.⁸ A symbolic-regression-aided equation of state also improved liquid-density prediction relative to conventional cubic equations, within its recommended range up to 100 MPa. These methods could incorporate physical laws into interpretable, analyzable, and reliable mathematical models to address the black-box problem. The resulting explicit models can also be transferred directly to subsequent applications, including the exploration of thermophysical relationships and thermodynamic performance analysis. But these methods require a thorough understanding of established thermophysical theory and the definition of algorithm-acceptable theoretical criteria. Such criteria must guide and constrain the entire search process, otherwise the resulting models may be disorganized and difficult to interpret.

FUTURE RESEARCH DIRECTIONS

Homogeneously embedding thermodynamic knowledge into algorithmic cores for autonomous modeling throughout the workflow

For theory-intensive thermophysical research, the first priority is to embed accumulated knowledge and experience within the AI core, combining broad exploration mechanism with rigorous thermodynamic constraints. For example, dimensional consistency, ideal-gas limit, component-exchange symmetry, and pure-component limits can serve as boundary conditions; thermodynamic differential relationships, phase equilibrium conditions, and stability criteria can constrain the direction of exploration; previously established empirical parameters and model fragments can form an element library. Recent large language models can also be used to improve the AI algorithms. The eventual aim is an exploratory modeling method that accommodate the commonness and individuality of substances as well as the consistency between microscopic mechanisms and macroscopic properties. AI would not only be an aid to conjecture, but also act as an independent agent with highly reliable modeling capable, even when available data are scarce.

Analytical joint optimization linking molecular characteristics, working-fluid properties, and cycle evaluation

When relating to energy systems, future research should focus on whole-chain parametric design from molecular characteristics to cycle performance. Traditional working fluid design proceeds through sequential screening. For a specified cycle configuration and parameter requirements, the thermodynamic performance of each candidate is evaluated using data or models and the best fluid is then selected from a finite set of candidates,⁹ which may limit both efficiency and scope. While explicit thermophysical models based on molecular characteristic can instead be coupled with specific cycle models. This coupling can directly relate performance index, such as efficiency and energy density, to basic substance parameters, thereby providing general optimization criteria and enabling performance-oriented joint design. Such approaches would define a working fluid through a small number of parameters, establishing broadly applicable design criteria and optimization methods at the molecular level.¹⁰ This avoids the cumbersome process of screening numerous molecules sequentially, thus substantially improving the speed and reliability of energy system parameter design as well as allowing flexible responses to variable operating conditions.

CONCLUSIONS

Whereas rapid responses and language understanding are central to generative large language models, theoretical sciences place greater demands on interpretability and algebraic computation of AI technologies. In

strong-theory and high-certainty thermophysical research, current AI challenges primarily concern the limited interpretability of the resulting data and substantial engineering risks. Future breakthroughs require to translate the accumulated thermophysical theories into the algorithmic core. This would support AI-driven exploration with capabilities for theoretical derivation and algebraic computation, improving the reliability of the resulting data. Then AI's systematic exploration abilities could establish a highly interpretable modeling paradigm, overcoming the potential subjectivity and limited perspectives of traditional methods. The new modeling method would provide broadly applicable thermophysical descriptions for novel working fluids with scarce data. Further, relationships between thermophysical models and cycle performances could be analytically integrated and jointly optimized, enabling AI to autonomously configure cycle parameters and design working fluids. This would support the intelligent development of novel energy systems covering diverse configurations and broad operating ranges.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used OpenAI Codex to improve the manuscript's language and readability. All AI-assisted text was subsequently checked and revised by the authors, who assume full responsibility for the published content.

REFERENCES

- Jirasek F. and Hasse H. (2023). Combining Machine Learning with Physical Knowledge in Thermodynamic Modeling of Fluid Mixtures. *Annu. Rev. Chem. Biomol. Eng.* **14**:31–51. DOI:10.1146/annurev-chembioeng-092220-025342
- Xu Y., Shao Z., Liu X., et al. (2026). AI for science: Progress, challenges, and perspectives. *The Innovation* 101530 (in press). DOI: 10.1016/j.xinn.2026.101530
- Liu Y., Hong W., and Cao B. (2019). Machine learning for predicting thermodynamic properties of pure fluids and their mixtures. *Energy* **188**:116091. DOI:10.1016/j.energy.2019.116091
- He Y., Feng Y., Qiu L., et al. (2024). Data-driven approach augmented by attention mechanism in critical and boiling thermophysical properties prediction of fluorine/chlorine-based refrigerants. *Energy* **306**:132490. DOI:10.1016/j.energy.2024.132490
- Chaparro G. and Müller E.A. (2024). Development of a Helmholtz free energy equation of state for fluid and solid phases via artificial neural networks. *Commun. Phys.* **7**:406. DOI:10.1038/s42005-024-01892-3
- Wang X., Li Y., Yan Y., et al. (2020). Prediction on the viscosity and thermal conductivity of HFC/HFO refrigerants with artificial neural network models. *Int. J. Refrig.* **119**:316–325. DOI:10.1016/j.ijrefrig.2020.07.006
- Alkhatib I.I.I., Albà C.G., Darwish A.S., et al. (2022). Searching for Sustainable Refrigerants by Bridging Molecular Modeling with Machine Learning. *Ind. Eng. Chem. Res.* **61**:7414–7429. DOI:10.1021/acs.iecr.2c00719
- Peng X., Qing K., Tan Y., et al. (2026). A physically interpretable machine-learning method in efficiently and explicitly exploring predictive thermodynamic models for mixture working fluids. *Energy* **342**:139577. DOI:10.1016/j.energy.2025.139577
- McLinden M.O., Brown J.S., Brignoli R., et al. (2017). Limited options for low-global-warming-potential refrigerants. *Nat. Commun.* **8**:14476. DOI:10.1038/ncomms14476
- Yan Y., Yang F., Zhang H., et al. (2023). Study on performance evaluation framework and design/ selection guidelines of working fluids for subcritical organic Rankine cycle from molecular structure perspective. *Energy* **282**:128582. DOI:10.1016/j.energy.2023.128582

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the Beijing Natural Science Foundation (Grant No. 3264044) and the China Postdoctoral Science Foundation (Grant No. 2026M790642). Xiayao Peng acknowledges support from the Shuimu Tsinghua Scholar Program (Grant No. 2025SM040).

DECLARATION OF INTERESTS

The authors declare no competing interests.