Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A knowledge driven few shot transfer learning framework for large scale thermal field applications

Yang Yu

yu_yang@tjutcm.edu.cn

Tianjin University of Traditional Chinese Medicine

Yue Zhang

Tianjin University of Traditional Chinese Medicine

Qilong Xue

Tianjin University of Traditional Chinese Medicine

Jingxuan Zhang

Tianjin University of Traditional Chinese Medicine

Wensong Huang

Tianjin University of Traditional Chinese Medicine

Zheng Li

Tianjin University of Traditional Chinese Medicine

Article

Keywords: Cross tank, Small sample, Transfer learning, Temperature field reconstruction, Empirical equation

Posted Date: June 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9867547/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

A knowledge driven few shot transfer learning framework for large scale thermal field applications

Yue Zhang[a, b,#], Qilong Xue[a, b,#], Jingxuan Zhang[a, b],Wensong Huang[a, b], Yang Yu [a, b, c, d, e,*], Zheng Li [a, b, c, d, e,*]

a. College of Pharmaceutical Engineering of Traditional Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, 301617, China.

b. State Key Laboratory of Component Traditional Chinese Medicine, Tianjin, 301617, China.

c. Haihe Laboratory of Modern Chinese Medicine, Tianjin, 301617, China.

d. Tianjin Key Laboratory of Intelligent and Green Pharmaceuticals for Traditional Chinese Medicine, China.

e. State Key Laboratory of Chinese Medicine Modernization, Tianjin, 301617, China.

#. These authors contributed equally to this work.

*. Corresponding author. Tianjin University of Traditional Chinese Medicine, Tianjin, 301617, China.

*E-m*ail addresses:lizheng@tjutcm.edu.cn(Z. Li),yu_yang@tjutcm.edu.cn(Y.Yu).

Abstract

Cross-Device transfer learning aims to rapidly generalize acquired physical knowledge and model capabilities to new industrial equipment, reducing repeated modeling costs and enhancing process intelligence. In large scale industrial heating systems, differences in geometry, liquid levels, boundary heat transfer, operating conditions, and sensor layouts can cause temperature field variations, making direct data driven temperature fitting prone to capturing device specific scale features rather than transferable heat transfer behavior. To address this, we propose a knowledge driven cross-device temperature field transfer learning framework. Using only spatiotemporal measurements from 14 industrial heating tanks, the framework extracts shared heat transfer knowledge via geometric normalization, local control volume energy balance, and the Gr/Ra response pathway, reconstructing sparse observations into physically meaningful intermediate heat transfer states. From the logarithmic effective Nusselt state, we further construct a gradient projected response along the dimensionless temperature gradient direction, capturing local sensitivity in buoyancy driven heat transfer while mitigating device scale, boundary, and operational amplitude effects. A deep learning model then learns the Nusselt

intermediate state, and a physical projection operator forms a transferable heat transfer representation. Industrial experiments show that this representation compresses inter tank distributional discrepancies and establishes a stable cross domain similarity space in heat transfer active regions. Tests on unseen tanks demonstrate robust active region response prediction. These results indicate that few sample industrial temperature field transfer should rely on interpretable intermediate heat transfer variables, enabling a closed loop from observational data to physical knowledge and rapid field prediction.

Keywords: Cross tank, Small sample, Transfer learning, Temperature field reconstruction, Empirical equation

1.Introduction

In large scale industrial heating systems, the rapid transfer of modelling knowledge acquired from existing equipment to new devices and operating scenarios remains a central challenge in spatial field modelling and intelligent industrial control. Conventional data driven models can achieve high accuracy under specific equipment configurations and operating conditions, but they generally require large amounts of representative data. When equipment geometry, liquid level, boundary heat transfer conditions, heating mode, or sensor layout changes, pronounced inter domain discrepancies may arise in the temperature field. Under such circumstances, models tend to learn device specific scale features or condition dependent empirical patterns rather than transferable heat transfer laws. Physics based numerical simulation, on the other hand, can describe heat transfer processes by discretizing conservation equations. However, its practical use in real industrial scenarios is often limited by uncertain boundary conditions, sensor errors, device to device variability, and the complex coupling among heat loss, convection, and equipment geometry. To obtain a usable model, repeated cycles of simulation, validation, parameter correction, and re-simulation are often required in engineering practice, resulting in high computational cost, long development cycles, and limited repeatability. This limitation is particularly critical in industrial production, where physical processes

evolve continuously whereas high fidelity simulations often require computational times that exceed the timescale allowed for practical decision making. Therefore, cross-device temperature field modelling cannot rely solely on repeated simulations or direct temperature field fitting. Instead, a knowledge transfer strategy is required that preserves shared heat transfer mechanisms while suppressing device specific discrepancies[1-6].

In recent years, physics informed machine learning[7-12], neural operators[13], sparse sensor reconstruction[14-18], and generative field prediction methods have provided new opportunities for the rapid modelling of complex flow and heat transfer systems. Physics informed neural networks constrain empirical losses using partial differential equations, neural operators accelerate the solution of parameterized governing equations by learning mappings between function spaces, and sparse reconstruction and generative models have demonstrated the potential to recover full field dynamics from limited observations. However, when these methods are applied to real industrial cross-device transfer, several limitations remain. First, physical laws are often incorporated mainly as penalty terms or external constraints, rather than being used to transform raw observations into scale invariant heat transfer variables. As a result, the physical basis for cross-device generalization is not sufficiently explicit. Second, end to end temperature field prediction is sensitive to changes in sensor layout, equipment size, and boundary intensity, and its accuracy may deteriorate as discrepancies between devices increase. Third, although some generative or pretrained models can improve field reconstruction, their sampling cost can be high and their internal representations often lack clear heat transfer meaning. For few sample industrial scenarios, the key challenge is therefore not simply to fit a new temperature field, but to extract reusable heat transfer knowledge from limited experimental phenomena and convert it into a learnable, transferable, and interpretable intermediate representation.

A knowledge driven cross-device transfer learning method for industrial thermal processes is proposed here, in which transferable similarity principles, conservation relationships, and response laws from classical theory are reorganized into learnable representations. Distributed temperature observations are first obtained from

industrial equipment to describe the spatiotemporal phenomena of the thermal process[19-23]. Geometric normalization is then performed to map the radii, heights, liquid levels, and spatial coordinates of different devices into a unified dimensionless space, thereby reducing the influence of device size differences. On this basis, an effective Nusselt state is inferred from temperature evolution through local control volume energy balance, transforming raw temperature observations into a unified representation of local heat transfer intensity. Furthermore, the Gr/Ra response pathway is introduced[24], and a gradient projected state is constructed along the dimensionless temperature gradient direction to characterize the sensitivity of local thermal responses to variations in the thermal driving force. Finally, this intermediate heat transfer state is learned by an artificial intelligence model and refined through a physical projection operator into a transferable knowledge representation for thermal processes. Through this structure, the learning target is shifted from device dependent temperature amplitudes to response laws associated with the underlying heat transfer mechanism.

To evaluate the industrial applicability of this method, traditional Chinese medicine extraction is selected as a representative demonstration scenario[25,26]. The manufacture of traditional Chinese medicine is characterized by multiple product varieties, manufacturers, equipment platforms, and production batches. The same process knowledge often needs to be transferred across different extraction tanks, liquid levels, heating conditions, and production platforms[27-30]. If a separate model must be rebuilt for each product, device, or operating condition, the costs of data acquisition and model maintenance will increase substantially, and the requirements of production scale up, process transfer, and process optimization will be difficult to satisfy. Therefore, thermal processes in traditional Chinese medicine manufacturing provide a representative setting for cross-device knowledge transfer. In this study, spatiotemporal temperature field data from 14 industrial extraction tanks with different sizes, liquid levels, heating powers, and environmental conditions are used to validate the proposed method. By transforming temperature field observations into dimensionless heat transfer response representations, the study examines whether a

stable similarity space can be formed across different tanks and further evaluates the capability of the method for temperature field reconstruction and heat transfer response prediction in unseen devices[31-35].

The main contributions of this study are fourfold. First, a cross-device transfer methodology for industrial thermal processes is proposed, shifting the modelling objective from direct temperature field fitting to the learning of shared heat transfer knowledge. Second, a physically interpretable intermediate variable chain is established by integrating geometric normalization, local energy balance, the effective Nusselt state, and gradient projected response representation. Third, the method is demonstrated in a traditional Chinese medicine manufacturing scenario involving multiple product categories, equipment platforms, and operating conditions, thereby validating its feasibility in industrial settings that require frequent transfer. Fourth, a closed loop framework is established from limited temperature observations to full field reconstruction and unseen device prediction, providing a generalizability route for combining classical heat transfer theory with artificial intelligence to enable reusable knowledge representation in industrial thermal processes.

2. Methods

This study focuses on transfer learning of temperature fields in vertical cylindrical extraction tanks. Water was used as the working fluid in all experiments, and all tanks were heated by jacketed steam to minimize additional physical variability arising from different heat source configurations. Temperature field data were acquired using a distributed temperature sensing (DTS) based reconstruction. Each experimental data was defined by the tank identifier, liquid level or effective sensing height, steam-valve opening and acquisition period, and was treated as an independent data set for training, validation or testing.

2.1 Experimental system and multi-tank data acquisition

The temperature field was measured using a DTS system. The DTS measurement principle is based on Raman scattering: when a laser pulse propagates along an optical fibre, Stokes and anti-Stokes scattering components are generated, and their intensity

ratio is temperature dependent. The system acquired the back-scattered signals along the fibre in real time through two channels, thereby providing continuous spatial temperature information inside the tank. The experimental platform consisted of an OTS3-20-50-2CH distributed fibre-optic temperature measurement system, temperature-sensing fibers, a fibre fusion splicer, an industrial metal sensing frame, a local-heating positioning device and auxiliary fixtures.

2.1.1 Sensor Array Configuration, Spatial Registration, and Temperature Field Coordinate System

The sensing frame was installed inside the tank and assembled from multiple cross-shaped bracket with a height of 0.5 m and diameters of 1.0 or 1.4 m, matching the dimensions of different tanks. As shown in the installation schematic diagram in Fig.1,the frame was numbered clockwise: the four corner points on the bottom layer were labeled 1-1, 1-2, 1-3 and 1-4; those on the second layer were labeled 2-1, 2-2, 2-3 and 2-4; and the same rule was applied up to the top layer. The optical fibre was wound around the frame following a predefined route, ensuring a one-to-one correspondence between the DTS signal and the physical sensing locations on the frame.

After fibre installation, local thermal positioning was performed by heating prescribed locations with a hot-air gun. The resulting local temperature peaks in the DTS signal were used to register the fibre-length coordinate to the corresponding physical coordinate. The sampling interval was 30 s. A total of 14 temperature rise datasets were collected, covering various areas, tank types, liquid levels, and valve openings under different seasonal ambient temperatures, as shown in Tab.1:

After spatial registration, the temperature field at time step n is denoted as:

$$T_i^n = T(x_i, t_n) \tag{1}$$

where i indexes a control volume or grid point, and the cylindrical coordinate is:

$$x_i = (r_i, \theta_i, z_i) \tag{2}$$

where r_i、θ_i、z_i denote the radial, azimuth and vertical coordinates.

2.1.2 Multi tanks datasets and unified cylindrical-coordinate representation

the raw DTS measurements formed an irregular temperature point cloud in cylindrical coordinates. To enable cross-tank learning and reduce computational cost, all temperature fields were mapped onto a unified cylindrical grid. The final training grid had a resolution of 3 × 5 × 36, comprising three radial nodes, five axial nodes and 36 azimuth planes. Each azimuth plane was unfolded into a rectangular surface along the radial and vertical directions, and adjacent planes were separated by 10 degrees.

Because the tank radius and effective sensing height varied across dataset, geometric dimensionless was first applied. Let R denote the effective tank radius and H the height of the sensing frame. The dimensionless coordinates are defined as:

$$\eta_i = \frac{r_i}{R}, \varphi_i = \frac{\theta_i}{2\pi}, \zeta_i = \frac{z_i}{H} \tag{3}$$

,where η_i、φ_i、ζ_i denote the dimensionless radial, azimuth and vertical coordinates, respectively. Periodic boundary conditions were applied in the azimuth direction, symmetry was imposed at the central axis, and the thermal boundary conditions at the wall, top and bottom surfaces were retained.

To reduce the influence of different global heating rates, the raw temperature field was converted to a dimensionless temperature state:

$$C_i^n = \frac{T_i^n - T_i^0}{\bar{\Delta T}^n} \tag{4}$$

Here, T_i^n is the temperature at grid point i and time step n, T_i^0 is the temperature at grid point i at the initial time. $\bar{\Delta T}^n$ is the spatially averaged temperature over the sensing domain at the current time. This normalization separates the overall heat-up level from local thermal redistribution, making the gradient structures comparable across tanks.

The gradient operator in the dimensionless coordinate system is written as:

$$\nabla_* = \left(\frac{\partial}{\partial \eta}, \frac{\partial}{\partial \phi}, \frac{\partial}{\partial \xi}\right) \tag{5}$$

2.2 Mechanism-based calculation of Nusselt number

Local Nusselt numbers were inferred from the measured temperature sequence through energy conservation. The purpose of this step was to introduce an

interpretation heat-transfer intermediate before data-driven learning, so that the transferable state learned later had a well-defined physical origin.

2.2.1 System-scale heat balance

At the system scale, the effective heat supplied by the steam jacket primarily increases the internal energy of the water, while part of the heat is dissipated to the environment through the side wall and bottom. The system heat balance can be written as:

$$\dot{Q}_{in}^{n} = c\rho V \frac{\bar{T}^{n+1} - \bar{T}^{n}}{\Delta t} + \dot{Q}_{wall}^{n} + \dot{Q}_{bottom}^{n} \tag{6}$$

,where $\dot{Q}_{in}^{n}$ denotes the heat input from the steam jacket,ρ is the density of water, c is the specific heat capacity of water, V is the water volume,$\bar{T}^{n}$ is the average temperature in the temperature measurement range, Δt is the sampling interval，$\dot{Q}_{wall}^{n}$ and $\dot{Q}_{bottom}^{n}$ are the heat losses through the side wall and bottom, respectively. Both side wall and bottom losses include natural convection and thermal radiation. The boundary heat losses were therefore written in an equivalent heat-transfer form:

$$\dot{Q}_{loss}^{n} = h_{eq}^{n} \mathrm{A}_s (T_s^n - T_{env}^n) \tag{7}$$

$$h_{eq}^{n} = h_{conv}^{n} + h_{rad}^{n} \tag{8}$$

,where $A_s$is the heat-transfer area, T_s^n is the boundary surface temperature, T_{env}^n is the environment temperature, and h_{conv}^n and h_{rad}^n are the natural-convection and linear radiation heat-transfer coefficients, respectively.

2.2.2 Control-volume heat balance

At the local scale, each cell in the unified grid was treated as a control volume. The energy balance of control volume is:

$$c\rho V_i \frac{T_i^{n+1} - T_i^n}{\Delta t} = \sum_{f \in F_i} \dot{Q}_{f,i}^{n} + \dot{Q}_{bc,i}^{n} \tag{9}$$

,where V_i is the volume of the control volume,T_i^n is the temperature of control volume i at time step n, F is the set of neighbouring faces, $\dot{Q}_{f,i}^{n}$ is the heat rate

entering the control volume through face f, and $\dot{Q}_{bc,i}^{n}$ represents the heat source or heat loss term introduced at the wall, top, or bottom boundary. For internal neighbouring faces, heat exchange was approximated using an equivalent-conduction form:

$$\dot{Q}_{f,i}^{n} = \lambda\, A_f \frac{T_{f,i}^{n} - T_i^{n}}{\delta_f} \tag{10}$$

Here λ is the thermal conductivity of water, A_f is the area of the control volume face, δ_f is the distance between neighbouring control volume centre, and $T_{f,i}^{n}$ is the temperature of the neighbouring control volume across face f.

2.2.3 Effective Nusselt inversion for wall-adjacent control volumes

When control volume was adjacent to the heated wall, the wall normal heat exchange could not be obtained directly from the internal finite-difference stencil. It was therefore inferred from the control-volume energy balance. The wall-normal heat-exchange rate is:

$$\dot{Q}_{w,i}^{n} = c\,\rho V_i \frac{{T_i}^{n+1} - {T_i}^{n}}{\Delta t} - \sum_{f \neq w} \dot{Q}_{f,i}^{n} - \dot{Q}_{bc,i}^{n} \tag{11}$$

Substituting the internal heat-exchange approximation gives:

$$\dot{Q}_{w,i}^{n} = \rho c_p V_i \frac{T_i^{n+1} - T_i^{n}}{\Delta t} - \sum_{f \in F_i, f \neq w} \lambda A_{f,i} \frac{T_{f,i}^{n} - T_i^{n}}{\delta_f} - \dot{Q}_{bc,i}^{n} \tag{12}$$

The wall heat flux density, local heat-transfer coefficient and local Nusselt number are then obtained as:

$$q_{w,i}^{n} = \frac{\dot{Q}_{w,i}^{n}}{A_{w,i}} \tag{13}$$

$$h_i^{n} = \frac{q_w^{n}}{T_{w,i}^{n} - T_i^{n} + \varepsilon_T} \tag{14}$$

$$Nu_i^{n} = \frac{h_i^{n} L}{\lambda} \tag{15}$$

Organize the above formula:

$$Nu_i^{n} = \frac{L_i}{\lambda A_{w,i}(T_{w,i}^{n} - T_i^{n} + \varepsilon_T)} \left| c\rho V_i \frac{T_i^{n+1} - T_i^{n}}{\Delta t} - \sum_{f \in F_i, f \neq w} \lambda A_{f,i} \frac{T_{f,i}^{n} - T_i^{n}}{\delta_f} - \dot{Q}_{bc,i}^{n} \right| \tag{16}$$

,where $A_{w,i}$ is the wall heat transfer area, $T_{w,i}^{n}$ is the local wall or near-wall temperature, L_i is the characteristic length, and ε_T is introduced to avoid numerical

divergence when the wall-fluid temperature difference becomes too small. Nu_i^n is interpreted as an effective grid-scale heat-transfer indicator for wall-adjacent control volumes, rather than a boundary-layer-resolved local Nusselt number.

The logarithmic effective Nusselt state is then defined as:

$$U_i^n = \frac{ln\,(Nu^n + \varepsilon_{Nu})}{ln\ 10} \tag{17}$$

where ε_{Nu} is a numerical stabilization term that prevents logarithmic instability caused by zero or extremely small values. Use U^n to represent the local $ln\ Nu$ state.

2.3 Grashof derived transferable heat-transfer state

The transferable heat transfer state was defined from the relationship between Grashof driven buoyancy and the Nusselt response in natural convection. The Grashof number measures the relative strength of temperature-induced buoyancy to viscous resistance. For a non-uniform temperature field, a local Grashof number was used:

$$Gr^n = \frac{g\beta(T_i^n)\Delta T_i^n L^3}{\nu^2(T_i^n)} \tag{18}$$

Here g is the gravitational acceleration, β is the thermal expansion coefficient, ν is the kinematic viscosity, and ΔT_i^n is the local thermal driving temperature difference. Near the wall, ΔT_i^n is given by the wall-fluid temperature difference; in the interior fluid region, it can be approximated from the local temperature gradient. The corresponding local Rayleigh number is:

$$Ra^n = Gr^n Pr^n \tag{19}$$

,where Pr^n is the local Prandtl number. In natural-convection heat transfer, the local Nusselt response can be written in a local power-law form:

$$Nu^n = C(Ra^n)^m \tag{20}$$

,where C and m vary with position, time, liquid level and wall condition. Taking the logarithm gives:

$$ln\ Nu^n = ln\ C + m ln\ Ra^n \tag{21}$$

This suggests that a transferable heat-transfer representation should not only describe the magnitude of the effective Nusselt response, but also how the logarithmic heat-transfer response changes along the dominant thermal-driving direction.

In the present experimental system, accurately computing local Rayleigh number everywhere is difficult because wall temperature, local thermal properties and boundary heat losses are not fully observed at every point. Therefore, the dimensionless temperature gradient is used as an experimentally observable proxy for the local buoyancy-driving direction.

We define the transferable heat-transfer state G_{eff} as:

$$G_{eff}^{n} = \frac{\nabla_* U^n \cdot \nabla_* \Theta^n}{\|\nabla_* \Theta^n\|_2^2 + \varepsilon_G} \tag{22}$$

Here, G_{eff}^{n} is the effective rate of change of the local logarithmic Nusselt response along the dominant dimensionless temperature gradient direction.

2.4 Active heat transfer region

Regions with weak temperature gradients usually have low thermal driving force and are more sensitive to measurement noise and numerical regularization. Therefore, the active heat-transfer region was defined using the magnitude of the dimensionless temperature gradient.

The dimensionless temperature-gradient magnitude at grid point i and time step n is defined as:

$$\chi_{\Theta,i}^{n} = \| \nabla_* \Theta_i^n \|_2 \tag{23}$$

The active-region threshold is determined from the training set:

$$\tau_{act} = \mathrm{q}_{0.70}(\{\chi_{\Theta,i}^{n} : (i,n) \in I_{train}\}) \tag{24}$$

Here, $\mathrm{q}_{0.70}$ denotes the 70th percentile, and $\mathrm{I}_{\mathrm{train}}$ is the set of all spacetime samples in the training dataset. The same threshold is fixed and applied to the validation and test cases. No validation or test information is used to redefine the active region.

The active region mask is:

$$M_{act,i}^{n} = 1(\chi_{\Theta,i}^{n} \geq \tau_{act}) \tag{25}$$

Equivalently, the active set is:

$$A = \{(i,n) : \chi_{\Theta,i}^{n} \geq \tau_{act}\} \tag{26}$$

To increase the relative contribution of high-gradient heat transfer regions while retaining weakly driven regions in the loss function, a continuous sample weight is assigned to each point:

$$w_i^n = 1 + \alpha min(\frac{\chi_{\Theta,i}^n}{\tau_{act}+\varepsilon_\chi}, \gamma) \tag{27}$$

Here, α controls the strength of active-region weighting, γ limits the maximum weight, and ε_χ prevents division by a small threshold. The weights are normalized within each mini-batch:

$$\tilde{w}_i^n = \frac{w_i^n}{\frac{1}{|B|}\sum_{(j,m)\in B} w_j^m} \tag{28}$$

,where B is the mini-batch. This normalization changes the relative importance of samples without changing the overall scale of the loss.

2.5 Hybrid local-global learning model

The model predicts the logarithmic effective Nusselt state U_i. The transferable response state $G_{eff}$is then derived from the predicted U_i using the same projection operator as that used for label construction.

For each grid point i at time step n, the model input is:

$$x_i^n = [\Theta_i^n, \eta_i, \phi_i, \zeta_i, \nabla_* \Theta_i^n, b_i, r_i^n] \tag{29}$$

,where b_i contains boundary indicators, and $\mathrm{r_i^n}$ denotes local-neighbourhood and field-level thermal-regime features. The local neighbourhood tensor is denoted by $\mathrm{N_i^n}$.The model prediction is:

$$\hat{U}_i^n = \mathrm{f}_\psi(x_i^n, N_i^n) \tag{30}$$

,where f_ψ is the hybrid local-global learning model with trainable parameters ψ.

The predicted transferable response is computed as

$$\hat{G}_{eff,i}^n = \frac{\nabla_* \hat{U}_i^n \cdot \nabla_* \Theta_i^n}{\|\nabla_* \Theta_i^n\|_2^2 + \varepsilon_G} \tag{31}$$

The same discrete gradient operator, boundary treatment and active-region definition are used for both G_{eff}^n and $\hat{G}_{eff,i}^n$. This ensures that the training target and predicted response are mathematically consistent.

The network contains two complementary branches. The local branch uses three-dimensional convolution to capture short range conductive exchange, near-wall structures and local thermal gradients. The global branch uses a Transformer encoder with positional encoding to capture long-range coupling among radial, axial and azimuth regions. The two branches are fused through a gated module, and the final output is $\hat{U}$.The neural network architecture is shown in Fig.2:

2.5.1 Baseline model comparison

To evaluate the necessity of the proposed transferable heat transfer state, we compared four model configurations using the same dataset partition, mesh representation, and evaluation metrics. The field calibration baseline incorporates field-level scaling correction to test whether global calibration alone enhances cross-domain transfer performance; the dimensionless prior baseline retains dimensionless coordinates and temperature states but omits the active heat transfer constraint based on G_{eff}; the ensemble robust baseline introduces source domain balance or a robust penalty term to verify whether domain balance alone suffices; the balanced two-stage model integrates dimensionless inputs, an effective Nusselt intermediate layer, a supervision mechanism based on G_{eff}, active region weighting, and source domain balance training.

The model selection is not based solely on the global RMSE but comprehensively considers the active region G_{eff} error, the global G_{eff} error, the U error, and stability across unseen test cases.

The training set, validation set, and test set are all partitioned at the dataset level. All images and grid points from the same heating experiment are retained within the same partition group. This prevents data from adjacent time steps or neighboring spatial points from being mixed in, ensuring that the test performance accurately reflects the generalization capability across cases.

2.5.2 Training objective

Training uses U_i as an auxiliary effective heat-transfer target and G_{eff} as the

main transferable response target. The loss contains three terms: the logarithmic effective Nusselt loss, the weighted all field G_{eff} loss and the active-region G_{eff} loss.

The U_i loss is

$$L_U = \frac{1}{|B|} \sum_{(i,n)\in B} \ell(\hat{U}_i^n, U_i^n) \tag{32}$$

The weighted all field G_{eff} loss is

$$L_{G,all} = \frac{\sum_{(i,n)\in B} \tilde{w}_i^n \ell(\hat{G}_{eff,i}^n, G_{eff,i}^n)}{\sum_{(i,n)\in B} \tilde{w}_i^n} \tag{33}$$

The active region G_{eff} loss is

$$L_{G,act} = \frac{\sum_{(i,n)\in B} M_{act,i}^n \ell(\hat{G}_{eff,i}^n, G_{eff,i}^n)}{\sum_{(i,n)\in B} M_{act,i}^n + \varepsilon_M} \tag{34}$$

Here, ℓ denotes the point wise regression loss, which may be the squared error, Smooth L1 loss or Huber loss.

The total loss is

$$L = \lambda_U L_U + \lambda_{G,all} L_{G,all} + \lambda_{G,act} L_{G,act} \tag{35}$$

,where λ_U, $\lambda_{G,all}$ and $\lambda_{G,act}$ are loss weights.

In the first training stage, the model learns a stable mapping to U_i and G_{eff}. In the second stage, $\lambda_{G,act}$ is increased to emphasize heat-transfer-dominant regions while retaining the auxiliary $U_i$constraint. This prevents the model from improving G_{eff} at the expense of the effective Nusselt intermediate layer.

2.7 Coupling with control-volume temperature advancement

After training, the predicted logarithmic effective Nusselt state can be converted back to an effective Nusselt indicator:

$$\hat{\mathrm{Nu}}_\mathrm{i}^\mathrm{n} = 10^{\hat{\mathrm{U}}_\mathrm{i}^\mathrm{n}} - \varepsilon_{\mathrm{Nu}}, 0 \tag{36}$$

The corresponding effective heat-transfer coefficient is:

$$\hat{\mathrm{h}}_\mathrm{i}^\mathrm{n} = \frac{\lambda\, \hat{\mathrm{Nu}}_\mathrm{i}^\mathrm{n}}{\mathrm{L_i}} \tag{37}$$

The predicted wall heat flux is:

$$\hat{q}^{n}_{w,i} = \int_{i}^{n}(T^{n}_{w,i} - T^{n}_{i}) \tag{38}$$

The wall heat rate entering the control volume is:

$$\hat{\dot{Q}}^{n}_{w,i} = A_{w,i}\hat{q}^{n}_{w,i} \tag{39}$$

For internal faces, heat exchange is computed using the same equivalent-conduction form:

$$\hat{\dot{Q}}^{n}_{f,i} = \lambda A_{f,i}\frac{T^{n}_{N_f(i)} - T^{n}_{i}}{d_{f,i}} \tag{40}$$

The next-step temperature is then advanced by the control-volume balance:

$$\hat{T}^{n+1}_{i} = T^{n}_{i} + \frac{\Delta t}{\rho c_p V_i}\left(\sum_{f\in F_i,\, f\neq w}\hat{\dot{Q}}^{n}_{f,i} + \hat{\dot{Q}}^{n}_{w,i} + \hat{\dot{Q}}^{n}_{bc,i}\right) \tag{41}$$

For control volumes that are not adjacent to the heated wall, the wall term $\hat{\dot{Q}}^{n}_{w,i}$ is omitted. In this way, the learned U_i state provides the boundary heat transfer input, while the control-volume equation enforces local energy conservation during temperature-field advancement.

For multi-step prediction, the updated field $\hat{T}^{n+1}$is used to recompute $\hat{\Theta}^{n+1}$, $\hat{\nabla_*\Theta}^{n+1}$, the active mask and the model input for the next step. The prediction loop is therefore:

$$\hat{T}^{n} \rightarrow \hat{\Theta}^{n} \rightarrow \hat{U}^{n} \rightarrow \hat{Nu}^{n} \rightarrow \hat{\dot{Q}}^{n} \rightarrow \hat{T}^{n+1} \tag{42}$$

This closes the loop between temperature-state sensing, transferable heat-transfer prediction and physics-based temperature-field advancement.

3 Results

3.1 Large-scale long-horizon cross-tank temperature-field dataset

Transferable heat transfer learning requires data foundation that contains genuine equipment and operating-condition variation. The reconstructed 3D temperature field dataset covers spring, summer, and winter seasons, involving six different tank types

and a total of fourteen industrial cylindrical storage tanks. It encompasses datasets with varying liquid levels, valve openings, and environmental conditions, enabling the model to account for differences in geometry, thermal storage capacity, and boundary heating input. Fig.3 (a) By partitioning the data through independent heating experiments rather than random spatial or temporal samples, the leakage of relevant thermal historical information can be prevented, and a rigorous validation standard for cross-domain generalization capability is established.

Analysis of the original temperature field in Fig. 3(b) reveals significant spatial and temporal heterogeneity. As the heating process progresses, the thermal front propagates at varying rates and exhibits distinct spatial propagation patterns across different tanks. This phenomenon indicates that the initial temperature variation is strictly constrained by the equipment's specific heating method and boundary conditions.

The temperature variation curves of the interconnected storage tanks in Fig. 3(c) further reveal significant regional differences among various tanks and operating conditions. Despite similar overall heating trends, datasets with varying liquid levels or valve openings exhibit distinct average temperature variation patterns.

3.2 Discovery of transferable common heat-transfer knowledge

Transferable heat-transfer knowledge must remain stable when tank geometry, liquid level, heating intensity and boundary conditions change. The variables in Fig. 4 form a mechanistic progression from observed temperature T_i^n, to the dimensionless thermal state Θ, to the logarithmic effective Nusselt state ${U_i}^n$, and finally to the gradient-projected response state G_{eff}^n. This progression progressively separates equipment-specific thermal history from the local buoyancy-driven response shared across tanks.

In Fig.4 (a), raw temperature ${T_i}^n$ exhibits strong separation among different tanks and test domains, indicating that absolute temperature is highly case dependent. After normalization, Fig.4 (b) shows that Θ increases distributional overlap, but residual domain differences remain, especially in low gradient regions. The U^n

better reflects local heat transfer intensity but still displays domain-dependent spreading due to differences in boundary conditions and local driving temperature differences(Fig.4(c)). In contrast, Fig.4 (d) shows that G_{eff}^{n} produces the strongest overlap across tanks, particularly in the active heat-transfer region.

From Fig.4 (e),compared to the original temperature and reference condition achieves a more significant reduction in domain spacing within the heat transfer active region. Specifically, the domain spacing decreased significantly - from about 0.176 at the original temperature to about 0.011 at the current level, by 93.6% (compared with the original temperature) and 64.2% (compared with the reference value).The more critical insight revealed by this figure is that the compression effect of is most pronounced in the most actively heated regions. This indicates that it is not merely a standardization transformation of data but rather an integration of logarithmic formulations for effective Nusselt numbers with dimensionless temperature gradient directions, effectively filtering out temperature amplitude and geometric scale effects imposed by specific operating conditions while successfully extracting the intrinsic, transferable components inherent in responses driven by local Gr/Ra .

The physical consistency of G_{eff} was examined using the Gr/Ra–Nu response pathway. Representative wall-adjacent maps Fig.5.(a) shows that the thermal driving proxy $\chi_{\Theta,i}^{n}$, derived from the dimensionless temperature-gradient field, identifies regions with strong local thermal driving. As shown in Fig.5 (b), the spatially localized distribution of the input effect magnitude $\left|G_{eff}^{n}\right|$ within local sub regions of the heat transfer active zone indicates that the gradient field not only provides a clear driving direction for the transmissible response but also spatially defines the effective support range of its active response region.

However, Fig.5 (c) shows that $\chi_{\Theta,i}^{n}$ alone does not determine the response magnitude. In the active region, the point wise association between $\log_{10}(1+\chi_{\Theta,i}^{n})$ and $\log_{10}(\left|G_{eff}^{n}\right|+\varepsilon_{G_{eff}^{n}})$ is moderate rather than strongly monotonic. This result is

physically reasonable because G_{eff} is not defined as the temperature-gradient magnitude. Instead, it represents the projected variation of the logarithmic effective Nusselt state U_i^n along the dominant dimensionless temperature-gradient direction. The gradient proxy therefore identifies where and along which direction the heat-transfer response should be evaluated, whereas the final response also depends on the spatial variation of U_i^n.

The internal consistency of the proposed response state is verified in Fig.5 (d). The response sensitivity computed from U_i^n and the dimensionless temperature-gradient direction agrees with G_{eff}, with a correlation coefficient of 1.000 and a normalized error of 0.000. This confirms that G_{eff} is not an unconstrained latent variable, but a derived heat-transfer response state defined by the same gradient-projection operator used in label construction. Representative maps Fig.5(e) further shows that the active regions selected by the fixed dimensionless-gradient threshold remain identifiable across different tanks and heating stages, although their shapes vary with operating conditions. Together, these results support the physical and mathematical consistency of G_{eff} with the proposed Gr/Ra–Nu response pathway, while avoiding the incorrect interpretation that the gradient magnitude alone monotonically determines G_{eff}.

Overall, Figs. 4 and 5 identify G_{eff}^n as the transferable common heat-transfer knowledge in this study. It satisfies three requirements simultaneously: it reduces domain shift across tanks, remains stable in active heat-transfer regions, and has a mechanistic origin in the Gr/Ra–Nu response chain. Therefore, subsequent prediction experiments use G_{eff}^n not as an empirical latent variable, but as a physically interpretable transferable state.

3.3 Dimensionless representation and cross-tank similarity

We then evaluated whether the proposed dimensionless representation increased dynamic similarity between tanks. The similarity matrices in Fig.6 provide a direct comparison of the domain structure before and after the proposed representation

change. Similarity was calculated from z-scored distribution histograms for each independent dataset, so that higher values indicate more similar field-state distributions. In the original temperature similarity matrix Fig.6 (a), data points originating from the same manufacturing domain or closely related operational series cluster into clearly structured blocks, reflecting that the original temperature field still retains tank-specific and operational-scale information. After temperature normalization Fig.6 (b), the block structure becomes less pronounced, suggesting that part of the global heating-scale effect is removed .

The G_{eff}-based similarity matrix shows a further reduction in domain-identity clustering (Fig.6 (c)). Datasets are less separated by manufacturer or tank identity and show higher cross-domain similarity when their heat-transfer response distributions are comparable. This trend is quantified in Fig.6 (d). The intra-domain minus inter-domain similarity gap decreases from 0.134 in the raw-temperature space to 0.013 in the G_{eff} space, corresponding to a 90.2% reduction in domain-identity clustering. These results indicate that G_{eff} forms a more transferable similarity space than raw temperature. Fig.6 demonstrates that the dataset does not become identical after transformation, and that the amplitude differences between tanks are reduced, thereby preserving physically meaningful heat transfer variations.

The active-region analysis in Fig.7 further examines where the transferable heat-transfer information is concentrated. The active masks were defined using the dimensionless temperature-gradient threshold determined from the training set and were then applied directly to validation and test cases without re-selecting the threshold. Representative maps Fig.7 (a)， the transferred masks identify localized heat-transfer-active zones in both validation and test datasets. The temporal curves further Fig.7 (b) shows that the fraction of active points evolves during heating, indicating that the selected zones follow the development of thermal fronts rather than remaining fixed in space or time.

The G_{eff} distributions in Fig.7 (c) compare all-field and active-region samples. The active-region distributions remain broadly overlapped across datasets, but the current data show that active-only analysis should not be interpreted as a simple

increase in absolute inter-domain similarity. Quantitatively, Fig.7 (d) shows the inter-domain G_{eff} similarity changes from 0.968 in the all-field analysis to 0.949 in the active-only analysis. More importantly, the intra-domain minus inter-domain similarity gap decreases from 0.011 to −0.005, corresponding to a 144.4% reduction in domain-identity clustering. The inter-domain to intra-domain transfer ability ratio increases from 0.989 to 1.005. These results indicate that active regions reduce the residual clustering by tank or manufacturer identity in the G_{eff} space, supporting the interpretation that the most transferable heat-transfer response is concentrated in thermally active regions.

Therefore, it provides the physical basis for the transfer-learning framework: the model is trained not to memorize tank-specific temperature patterns, but to learn a common heat-transfer response in a dimensionless similarity space.

3.4 Transfer learning framework for full-field thermal prediction

Finally, we assessed whether the learned heat-transfer representation improves response-state prediction under unseen tanks and operating conditions. The model takes the current dimensionless thermal state, spatial coordinates, boundary indicators and local neighbourhood information as inputs to predict the logarithmic effective Nusselt state U_i^n. The transferable response G_{eff} is then obtained through the same gradient-projection operator used in the label construction. This design ensures that the prediction is mediated by a physical heat-transfer layer rather than by direct black-box temperature regression.

The model comparison in Fig.8 (a) and (b) shows that the balanced two-stage model exhibits the strongest testing performance.It achieves an active-region G_{eff} RMSE of 0.721 with $R^2 = 0.978$, and a U_i^n RMSE of 0.135 with $R^2 = 0.950$. Compared with the prior-dimensionless baseline, the active-region G_{eff} RMSE is reduced by 64.0% and the U_i^n RMSE is reduced by 59.2%. Compared with the group-robust baseline, the corresponding reductions are 69.4% and 62.2%, respectively. The field-level calibration model shows no improvement in predicting response states from current data, indicating that while it reduces certain global

560 deviations, it may fail to maintain consistency in local heat transfer responses.

561 The domain-specific results in Fig.8 (c) and (d) further show that the balanced
562 two-stage model performs consistently on both test domains. It obtains active-region
563 G_{eff} $R^2 = 0.965$ on DRT.T1 and $R^2 = 0.986$ on HR.T2.H.rep, while the
564 corresponding U_i^n R^2 values are 0.935 and 0.980. The composite model-selection
565 score in Fig.8 (e) is also highest for the balanced two-stage model.

566 We further evaluated whether the learned heat-transfer state can generalize to
567 unseen or weakly represented tank conditions. Instead of closed-loop temperature
568 rollout, Fig.9 examines the prediction of the dimensionless heat-transfer response
569 state on representative JC.T1 and DRT.T1 cases. Fig.9 (a) shows that the axial error
570 profiles show that the all-field G_{eff} RMSE varies strongly along the height direction,
571 especially in the DRT.T1 case, where large errors appear in several axial layers. In
572 contrast, the substantially lower and axially more stable active-region RMSE reveals
573 that the learned representation preferentially captures the physically consistent
574 heat-transfer dynamics within dominant regions, whereas the higher and more
575 variable all-field RMSE arises from non-dominant zones where tank-specific or
576 amplitude-driven information persists and hinders generalization.

577 The temporal error profiles further show that all-field G_{eff} errors fluctuate
578 during heating and may contain large transient peaks, particularly for the external
579 DRT.T1 case in Fig.9 (b). With the passage of time, the active region error always
580 decreases, which suppresses the instability caused by different tank structures, and
581 retains the transportable part of the heat transfer response.Fig.9 (c) shows that the
582 true-versus-predicted scatter plots provide additional evidence for this interpretation.
583 For all valid points, G_{eff} predictions are affected by a small number of
584 extreme-response points, whereas the active-region scatter is more concentrated
585 around the ideal agreement line. The predicted logarithmic effective Nusselt state U_i^n
586 also shows a clear near-diagonal trend, indicating that the model captures the main
587 heat-transfer state before the gradient-projection step.

588 Together, these results show that the learned G_{eff}-based representation does not

uniformly remove all cross-domain errors, especially in challenging unseen DRT conditions. Nevertheless, it provides more stable and physically meaningful predictions in active heat-transfer regions, where the thermal driving is strongest. This supports the use of active-region G_{eff} prediction as a transferable heat-transfer-state constraint, while also indicating that full-field extrapolation to strongly shifted external domains remains a limitation of the current model.

The prediction results confirm that the proposed transfer learning knowledge framework achieves robust cross-tank generalization with limited samples by integrating transferable heat transfer knowledge with AI. The physically interpretable state vectorG_{eff} not only reduces individual variations among reactors but also preserves the structural characteristics of heat transfer in active regions, enabling reliable predictions for unseen systems even with minimal sample data.

4. Discussion

This study addresses the challenge of large-scale, long term temperature field prediction in industrial cylindrical extraction tanks by proposing a transfer learning framework centered on shared heat transfer knowledge across tanks. This study extracts recognizable, interpretable, and universally applicable general heat transfer knowledge from raw temperature data with limited samples through geometric dimensionless transformation, temperature state normalization, reconstruction of the Gr/Ra response chain, and artificial intelligence-based processing. They are grounded in physical relationships between local buoyancy-driven forces, dimensionless temperature gradients, and Nusselt responses. Traditional transfer methods rely on single-sample calibration or empirical correlations under fixed conditions; when applying the calibration model to systems with unknown geometric structures or different operating conditions, substantial data and a re-identification process are required, which hinders the application of small-sample transfer. The research results demonstrate that transforming the temperature field from the original physical quantity space to a dimensionless heat transfer response space not only significantly reduces variations between different storage tanks but also preserves structural information related to local heat transfer mechanisms. This reveals that cross tank

transfer requires only similar local response patterns under appropriate dimensionless heat transfer conditions, rather than identical temperature distributions across systems. These insights provide a new physical foundation for cross device modeling in complex industrial thermal processes.

This study fundamentally differs from single-device mechanism models, purely supervised deep learning models, and traditional field reconstruction methods. Single-device mechanism model typically relies on explicit boundary conditions, geometric parameters, and heat transfer coefficient specifications, offering good interpretability under specific operating conditions. However, in complex industrial environments with diverse operational conditions, this model is susceptible to unknown boundary disturbances, multi physical field coupling coupling, and drift in operating parameters.Pure supervised deep learning models are capable of fitting complex nonlinear relationships; however, when trained using raw temperature fields, they tend to over fit device-specific identity features and operational scale information. Consequently, their prediction results lack transferability.In contrast, the research framework integrating control volume energy balance, the Nu inter layer, dimensionless state variables, and active region constraints shifts the model's learning objective from temperature amplitude fitting to learning local heat transfer responses, thereby bridging data-driven performance with physical interpretability.

The concept of G_{eff} proposed during the research is derived from the Gr concept, aiming to further extract more transfer-efficient response information from Nu-related states. Nu reflects local heat transfer intensity and serves as a critical physical intermediate layer linking the temperature field and heat transfer boundaries; However, Nu itself remains influenced by geometric scale, wall temperature difference, and local low-drive noise. G_{eff} describes the effective rate of change of the $ln\ Nu$ response along the dimensionless temperature gradient direction, providing a closer approximation of heat transfer sensitivity under local buoyancy-driven conditions. Consequently, the model first predicts $ln\ Nu$, then uses G_{eff} to constrain the heat transfer structure within the active region, ensuring the prediction retains Nu physical intermediate layer while enhancing cross-system

transfer capability. This design avoids the risk of interpreting deep learning outputs as direct physical laws.

Build on research, future work can further explore real-time full-field reconstruction and rolling prediction driven by a limited number of measurement points. By selecting optimal measurement points based on transferred knowledge and utilizing identified active regions along with the G_{eff} representation, sensor layout strategies can be developed to maximize information gain, domain stability, and physical sensitivity, enabling a small set of measurement points to accurately capture the complete temperature field. Specifically, this method combines sparse measured temperature data with general knowledge of heat transfer acquired through learning, enabling reliable reconstruction of temperature fields in unmeasured regions. From a computational perspective, future closed-loop systems can integrate generative prior models with the physical constraints of PINN to develop digital twin models that combine high accuracy, rapid inference capabilities, and physical consistency. By collecting sparse measurement data from various tank compartments, the acquired information is fed into a generative temperature field reconstruction model for restoration; the reconstructed temperature field data is then input into a fast-computation model, thereby achieving continuous temperature field prediction across large spatial and temporal scales within the closed-loop system.

5.Conclusion

This study develops a knowledge framework that integrates empirical heat transfer equations with artificial intelligence, enabling transferable predictions of large-scale, long-term thermal fields in industrial cylindrical storage tanks.By conducting distributed temperature measurements across multiple storage tanks, liquid levels, and heating conditions, the original temperature field undergoes geometric normalization, dimensionless temperature representation, control volume energy balance analysis, and Gr/Ra-Nu response projection processing to restructure the measurement data. This transforms the observed thermal evolution process into easily interpretable heat transfer knowledge.This study identifies G_{eff}as a transferable heat-transfer response state. Both the original temperature and dimensionless

679 temperature only partially account for cross-domain variability, whereas this response
680 state most effectively eliminates inter-domain differences in heat-transfer active
681 regions. This indicates that heat transfer across storage tanks does not require
682 identical temperature distributions; a shared local response mechanism can be
683 employed once the global heating scale, geometric configuration, and thermal driving
684 direction are uniformly represented in a dimensionless space.

685 The active heat-transfer region provides the physical carrier of this transferable
686 knowledge. A fixed training-set threshold based on the dimensionless
687 temperature-gradient magnitude identifies evolving thermal-front regions without
688 using validation or test information. Within these regions, the response distributions
689 show reduced tank-identity clustering and improved inter-domain comparability,
690 demonstrating that transferable heat-transfer information is spatially concentrated
691 rather than uniformly distributed across the full field.

692 The hybrid local-global learning model operationalizes this knowledge by
693 predicting U_i^n and deriving G_{eff} through the same projection operator used in label
694 construction. This design embeds physical information into the learning target and
695 loss structure, while allowing AI to learn nonlinear cross-domain mappings under
696 unseen tank and operating conditions. The balanced two-stage strategy improves
697 active-region response prediction and preserves the Nusselt intermediate layer,
698 confirming that physical interpretability and predictive performance can be coupled
699 rather than treated as competing objectives.

700 The proposed framework establishes a comprehensive workflow encompassing
701 industrial low-sample thermal field observation, effective transferable representation
702 discovery, dimensionless similarity construction, and AI-driven transferability of
703 unknown fields. Its current limitation lies in the potential presence of domain-specific
704 errors in strongly offset or weakly driven regions; therefore, future research should
705 extend this method to areas such as sparse sensor reconstruction, closed-loop rolling
706 prediction, and adaptive active region updating. In summary, this study deeply
707 integrates classical heat transfer theory with artificial intelligence knowledge
708 extraction techniques to construct a reconfigurable and learnable knowledge

representation system, thereby developing an industrial thermal process knowledge model with enhanced interpretability and transferability.

References

Brunton, S. L., Noack, B. R. & Koumoutsakos, P. Machine Learning for Fluid Mechanics. 52, 477-508, doi:https://doi.org/10.1146/annurev-fluid-010719-060214 (2020).

Fukami, K., Goto, S. & Taira, K. Data-driven nonlinear turbulent flow scaling with Buckingham Pi variables. Journal of Fluid Mechanics 984, R4, doi:10.1017/jfm.2024.211 (2024).

Tang, H., Liao, Y., Yang, H. & Xie, L. A transfer learning-physics informed neural network (TL-PINN) for vortex-induced vibration. Ocean Engineering 266, 113101, doi:https://doi.org/10.1016/j.oceaneng.2022.113101 (2022).

Zhao, X. et al. Temperature field spatiotemporal modeling of lithium-ion battery pack configured sparse temperature sensors. Journal of Energy Storage 97, 113005, doi:https://doi.org/10.1016/j.est.2024.113005 (2024).

Zhou, K., Liu, Z., Qiao, Y., Xiang, T. & Loy, C. C. Domain Generalization: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 45, 4396-4415, doi:10.1109/TPAMI.2022.3195549 (2023).

Zhuang, F. et al. A Comprehensive Survey on Transfer Learning. Proceedings of the IEEE 109, 43-76, doi:10.1109/JPROC.2020.3004555 (2021).

Eivazi, H., Tahani, M., Schlatter, P. & Vinuesa, R. Physics-informed neural networks for solving

739 Reynolds-averaged Navier–Stokes equations. Physics of Fluids 34, 075117,
740 doi:10.1063/5.0095270 (2022).
741 Gao, H., Zahr, M. J. & Wang, J.-X. Physics-informed graph neural Galerkin networks: A unified
742 framework for solving PDE-governed forward and inverse problems. Computer Methods in
743 Applied Mechanics and Engineering 390, 114502,
744 doi:https://doi.org/10.1016/j.cma.2021.114502 (2022).
745 Haghighat, E., Raissi, M., Moure, A., Gomez, H. & Juanes, R. A physics-informed deep learning
746 framework for inversion and surrogate modeling in solid mechanics. Computer Methods in
747 Applied Mechanics and Engineering 379, 113741,
748 doi:https://doi.org/10.1016/j.cma.2021.113741 (2021).
749 Mao, Z., Jagtap, A. D. & Karniadakis, G. E. Physics-informed neural networks for high-speed flows.
750 Computer Methods in Applied Mechanics and Engineering 360, 112789,
751 doi:https://doi.org/10.1016/j.cma.2019.112789 (2020).
752 Prantikos, K., Chatzidakis, S., Tsoukalas, L. H. & Heifetz, A. Physics-informed neural network with
753 transfer learning (TL-PINN) based on domain similarity measure for prediction of nuclear
754 reactor transients. Scientific Reports 13, 16840, doi:10.1038/s41598-023-43325-1 (2023).
755 Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning
756 framework for solving forward and inverse problems involving nonlinear partial differential
757 equations. Journal of Computational Physics 378, 686-707,
758 doi:https://doi.org/10.1016/j.jcp.2018.10.045 (2019).
759 Lu, L., Jin, P., Pang, G., Zhang, Z. & Karniadakis, G. E. Learning nonlinear operators via DeepONet
760 based on the universal approximation theorem of operators. Nature Machine Intelligence 3,
761 218-229, doi:10.1038/s42256-021-00302-5 (2021).
762 Chen, L., Li, H.-X. & Yang, H.-D. Spatiotemporal Modeling for Distributed Parameter System under
763 Sparse Sensing. Industrial & Engineering Chemistry Research 59, 16321-16329,
764 doi:10.1021/acs.iecr.0c02120 (2020).
765 de Silva, B. et al. PySensors: A Python package for sparse sensor placement. Journal of Open Source
766 Software 6, doi:10.21105/joss.02828 (2021).
767 Erichson, N. B. et al. Shallow neural networks for fluid flow reconstruction with limited sensors.
768 Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 476,

20200097, doi:10.1098/rspa.2020.0097 (2020).

Fang, Z., Wang, M., Hu, W., Chang, K. & Zhang, B. Temperature-Field Sparse-Reconstruction of Lithium-Ion Battery Pack Based on Artificial Neural Network and Virtual Thermal Sensor Technology. Energy Technology 9, 2100258, doi:https://doi.org/10.1002/ente.202100258 (2021).

Manohar, K., Brunton, B. W., Kutz, J. N. & Brunton, S. L. Data-Driven Sparse Sensor Placement for Reconstruction: Demonstrating the Benefits of Exploiting Known Patterns. IEEE Control Systems Magazine 38, 63-86, doi:10.1109/MCS.2018.2810460 (2018).

Bense, V. F. et al. Distributed Temperature Sensing as a downhole tool in hydrogeology. Water Resources Research 52, 9259-9273, doi:https://doi.org/10.1002/2016WR018869 (2016).

Law, R. et al. Thermodynamics of a fast-moving Greenlandic outlet glacier revealed by fiber-optic distributed temperature sensing. Science Advances 7, eabe7136, doi:10.1126/sciadv.abe7136.

Li, J. & Zhang, M. Physics and applications of Raman distributed optical fiber sensing. Light: Science & Applications 11, doi:10.1038/s41377-022-00811-x (2022).

Li, J. et al. High-accuracy distributed temperature measurement using difference sensitive-temperature compensation for Raman-based optical fiber sensing. Opt. Express 27, 36183-36196, doi:10.1364/OE.27.036183 (2019).

Pieracci, A., Nanni, J., Tartarini, G. & Lanzoni, M. A Novel Approach to Raman Distributed Temperature-Sensing System for Short-Range Applications. Sensors 24, 2669 (2024).

Plumley, M. & Julien, K. Scaling Laws in Rayleigh-Bénard Convection. Earth and Space Science 6, 1580-1592, doi:https://doi.org/10.1029/2019EA000583 (2019).

Bhadange, Y. A., Carpenter, J. & Saharan, V. K. A Comprehensive Review on Advanced Extraction Techniques for Retrieving Bioactive Components from Natural Sources. ACS Omega 9, 31274-31297, doi:10.1021/acsomega.4c02718 (2024).

Onyebuchi, C. & Kavaz, D. Effect of extraction temperature and solvent type on the bioactive potential of Ocimum gratissimum L. extracts. Scientific Reports 10, 21760, doi:10.1038/s41598-020-78847-5 (2020).

Ponphaiboon, J. et al. Advances in Natural Product Extraction Techniques, Electrospun Fiber Fabrication, and the Integration of Experimental Design: A Comprehensive Review. 28, 5163 (2023).

Silva, L. C. B., Segatto, M. E. V. & Castellani, C. E. S. Raman scattering-based distributed temperature sensors: A comprehensive literature review over the past 37 years and towards new avenues. Optical Fiber Technology 74, 103091, doi:https://doi.org/10.1016/j.yofte.2022.103091 (2022).

Yan, Y. et al. Rapid temperature field prediction in natural product extraction. Journal of Food Engineering 404, 112782, doi:https://doi.org/10.1016/j.jfoodeng.2025.112782 (2026).

Sridhar, A., Vaishampayan, V., Senthil Kumar, P., Ponnuchamy, M. & Kapoor, A. Extraction techniques in food industry: Insights into process parameters and their optimization. Food and Chemical Toxicology 166, 113207, doi:https://doi.org/10.1016/j.fct.2022.113207 (2022).

Bakarji, J., Callaham, J., Brunton, S. L. & Kutz, J. N. Dimensionally consistent learning with Buckingham Pi. Nature Computational Science 2, 834-844, doi:10.1038/s43588-022-00355-5 (2022).

Kapjor, A., Durcansky, P. & Vantuch, M. Effect of Heat Source Placement on Natural Convection from Cylindrical Surfaces. Energies 13, 4334 (2020).

Shishkina, O. & Lohse, D. Ultimate Regime of Rayleigh-B\'enard Turbulence: Subregimes and Their Scaling Relations for the Nusselt vs Rayleigh and Prandtl Numbers. Physical Review Letters 133, 144001, doi:10.1103/PhysRevLett.133.144001 (2024).

Xie, X., Samaei, A., Guo, J., Liu, W. K. & Gan, Z. Data-driven discovery of dimensionless numbers and governing laws from scarce measurements. Nature Communications 13, 7562, doi:10.1038/s41467-022-35084-w (2022).

Yuan, Y. & Lozano-Durán, A. Dimensionless learning based on information. Nature Communications 16, 9171, doi:10.1038/s41467-025-64425-8 (2025).

Figure Legend:

Fig.1 (a) Installation diagram of optical fiber on temperature measuring frame. (b) Installation diagram of actual optical fiber temperature measuring device. (c) Perspective view of frame in tank.

Fig.2 Hybrid Neural Network Architecture

Fig.3 (a) Distribution of training dataset, test dataset and verification dataset. (b) 3D temperature distribution of HR, HRT and DRT extraction tanks under low temperature, medium temperature and high temperature during the heating process. (c) The figure above shows the average temperature curve of the heating process under different liquid levels of the same tank, and the figure below shows the average temperature change curve of the same tank under different valve opening.

Fig.4 (a) Distribution of the original temperature T after z-score normalization. (b) Distribution of the dimensionless temperature state θ. (c) Distribution of the logarithmic effective Nusselt state. (d) Distribution of the transferable heat-transfer state G_{eff}.(e) Mean squared maximum mean discrepancy (MMD^2), across domain pairs for different representations.Grey and blue bars denote all-field and active-region statistics, respectively.

Fig.5 (a) Representative thermal driving proxy field $\chi_{\Theta,i}^{n}$ derived from the dimensionless temperature-gradient magnitude. (b) Corresponding projected response magnitude $\left|G_{eff}^{n}\right|$.(c) Statistical association between $\log_{10}(\left|G_{eff}^{n}\right| + \varepsilon_{G_{eff}^{n}})$ and $\log_{10}(1 + \chi_{\Theta,i}^{n})$ within the active region.(d) Consistency between G_{eff} and the response sensitivity computed from $U_i{}^{n}$along the dimensionless temperature-gradient direction.(e) Representative active heat-transfer regions selected by the fixed training-set gradient threshold across different tanks and heating stages. Grey points indicate inactive grid points, and coloured points indicate active grid points.

Fig.6 (a) Dataset similarity matrix calculated from raw temperature distributions. (b) Dataset similarity matrix calculated from dimensionless temperature state distributions. (c) Dataset similarity matrix calculated from the G_{eff} response state distributions. (d) Intra-domain and inter-domain similarity statistics for T, θ, $U_i{}^{n}$, and G_{eff}.

Fig.7 (a) Representative active masks in validation and test cases obtained using the fixed threshold determined from the training set. (b) Temporal evolution of the active-region fraction during heating.Thin curves represent individual datasets and thick curves represent split-level averages. (c) Distributions of G_{eff} for all-field and active-region samples.Thin curves denote individual datasets and thick curves denote mean distributions. (d) Quantitative comparison between all-field and active-only G_{eff} similarity statistics.

Fig.8 (a) Overall test R^2 for active-region G_{eff} and $U_i{}^{n}$ prediction. (b) Overall test RMSE for active-region G_{eff} and $U_i{}^{n}$. (c) Domain-specific active-region G_{eff} R^2 on DRT.T1 and HR.T2.H.rep. (d) Domain-specific $U_i{}^{n}$ R^2 on DRT.T1 and HR.T2.H.rep. (e) Composite model selection score combining response-state and

Nusselt state criteria.

Fig.9 (a) Axial variation of G_{eff} RMSE for all-field and active-region points in JC.T1 and DRT.T1. (b) Temporal variation of G_{eff} RMSE during heating. (c)True-versus-predicted scatter plots for all valid G_{eff} points, active-region G_{eff} points and the logarithmic effective Nusselt state ${U_i}^n$.

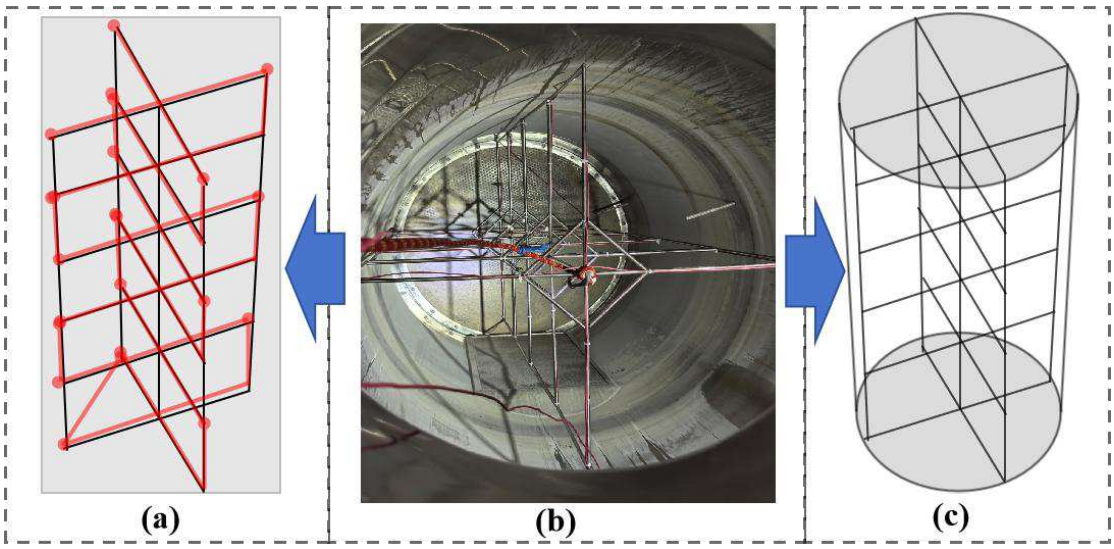


Fig.1(a) Installation diagram of optical fiber on temperature measuring frame. (b)

Installation diagram of actual optical fiber temperature measuring device. (c) Perspective view of frame in tank.

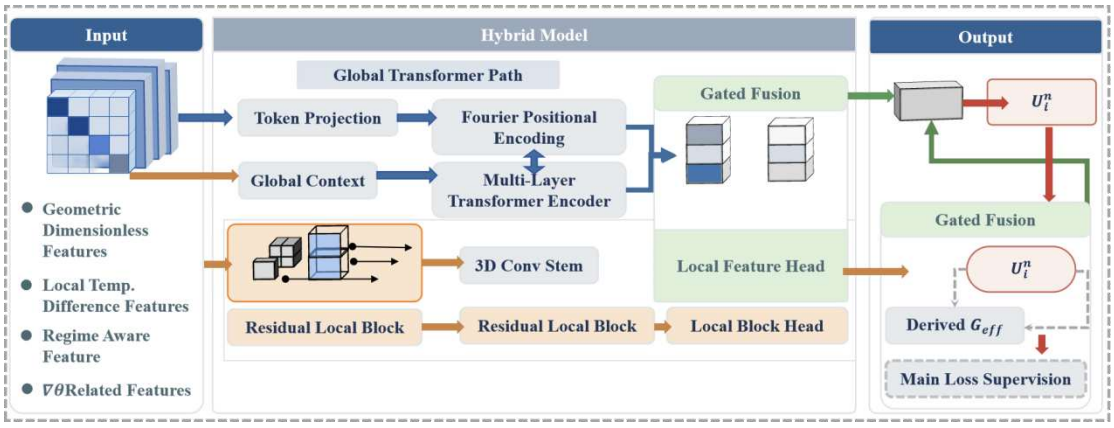
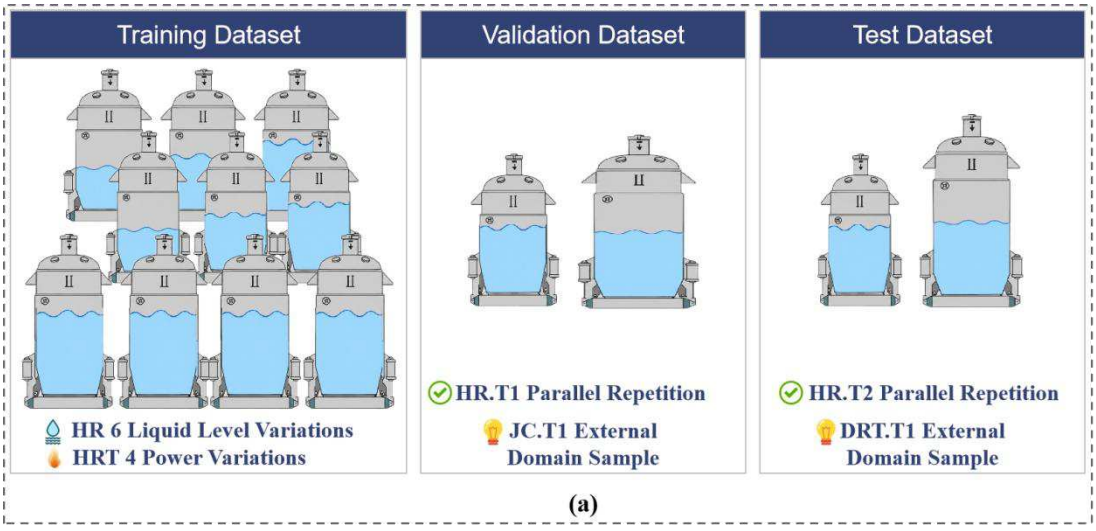
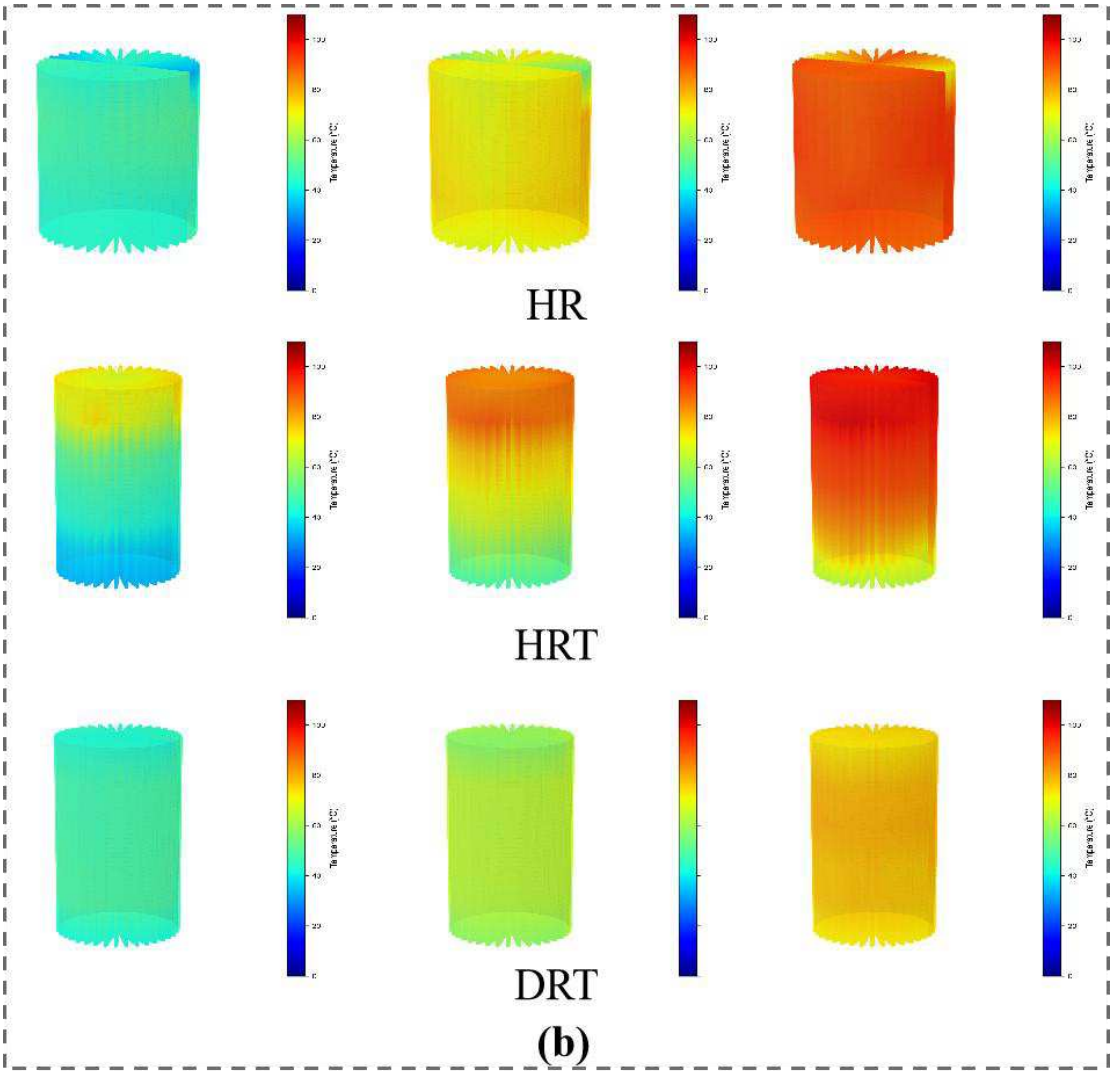


Fig. 2 Hybrid Neural Network Architecture


Training Dataset
Validation Dataset
Test Dataset
HR 6 Liquid Level Variations
HRT 4 Power Variations
HR.T1 Parallel Repetition
JC.T1 External Domain Sample
HR.T2 Parallel Repetition
DRT.T1 External Domain Sample

(a)


HR
HRT
DRT
(b)

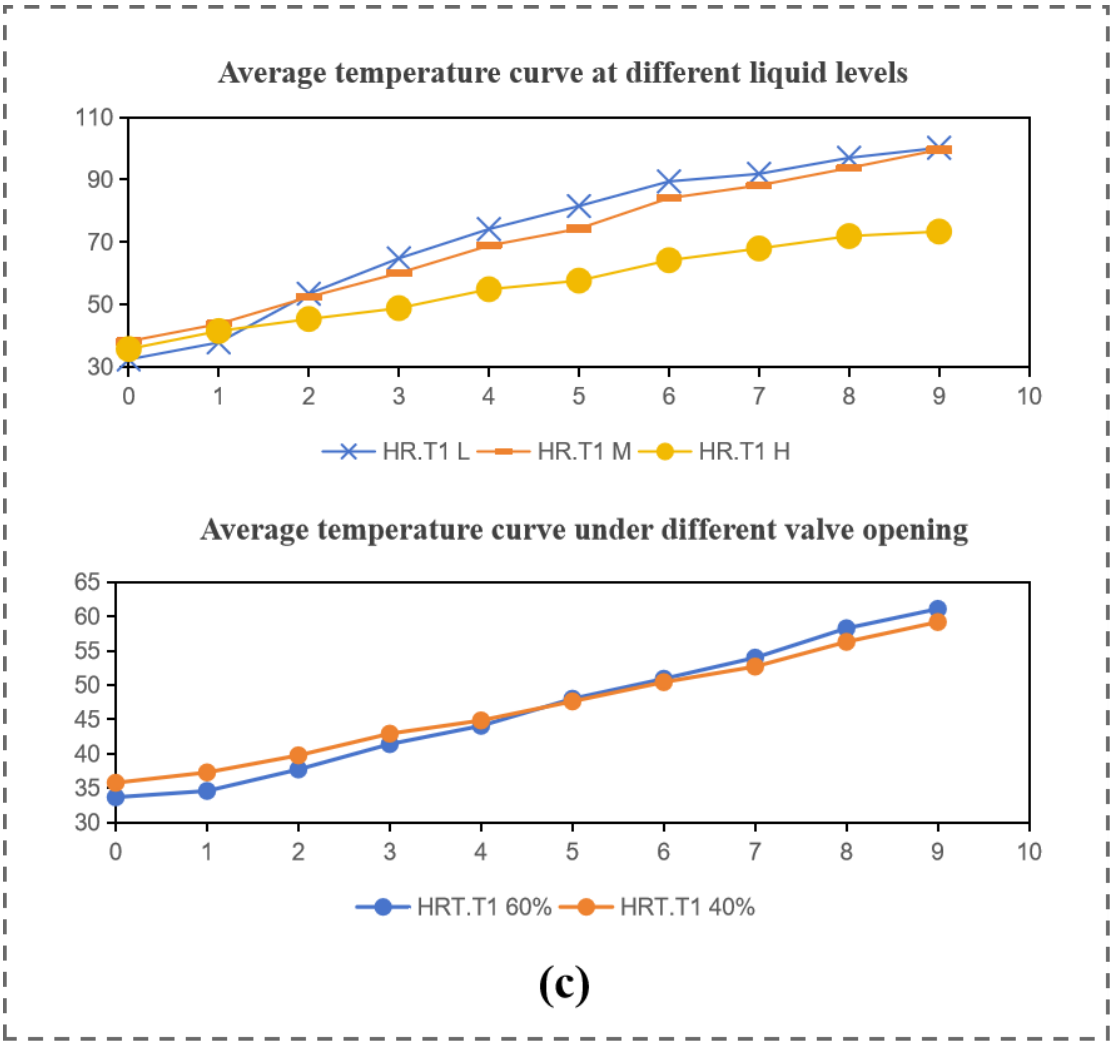


Fig. 3(a) Distribution of training dataset, test dataset and verification dataset. (b) 3D temperature distribution of HR, HRT and DRT extraction tanks under low temperature, medium temperature and high temperature during the heating process. (c) The figure above shows the average temperature curve of the heating process under different liquid levels of the same tank, and the figure below shows the average temperature change curve of the same tank under different valve opening.

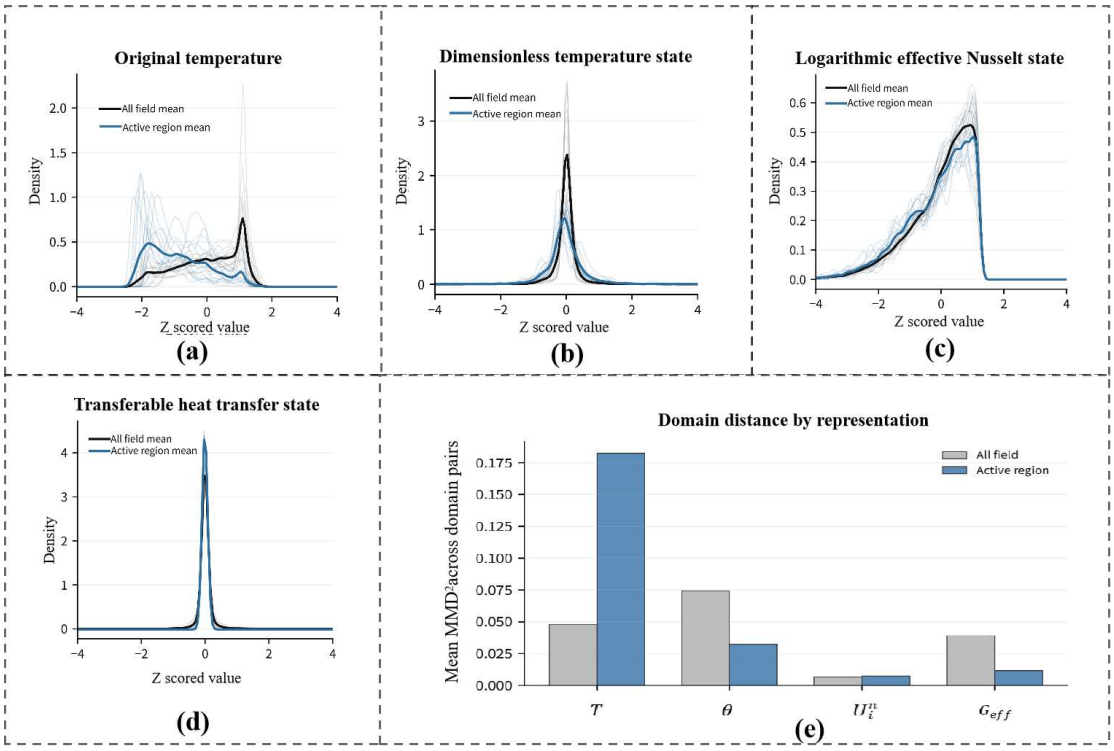


Fig. 4(a) Distribution of the original temperature T after z-score normalization. (b) Distribution of the dimensionless temperature state θ. (c) Distribution of the logarithmic effective Nusselt state. (d) Distribution of the transferable heat-transfer state G_{eff}.(e) Mean squared maximum mean discrepancy (MMD^2), across domain pairs for different representations.Grey and blue bars denote all-field and active-region statistics, respectively.

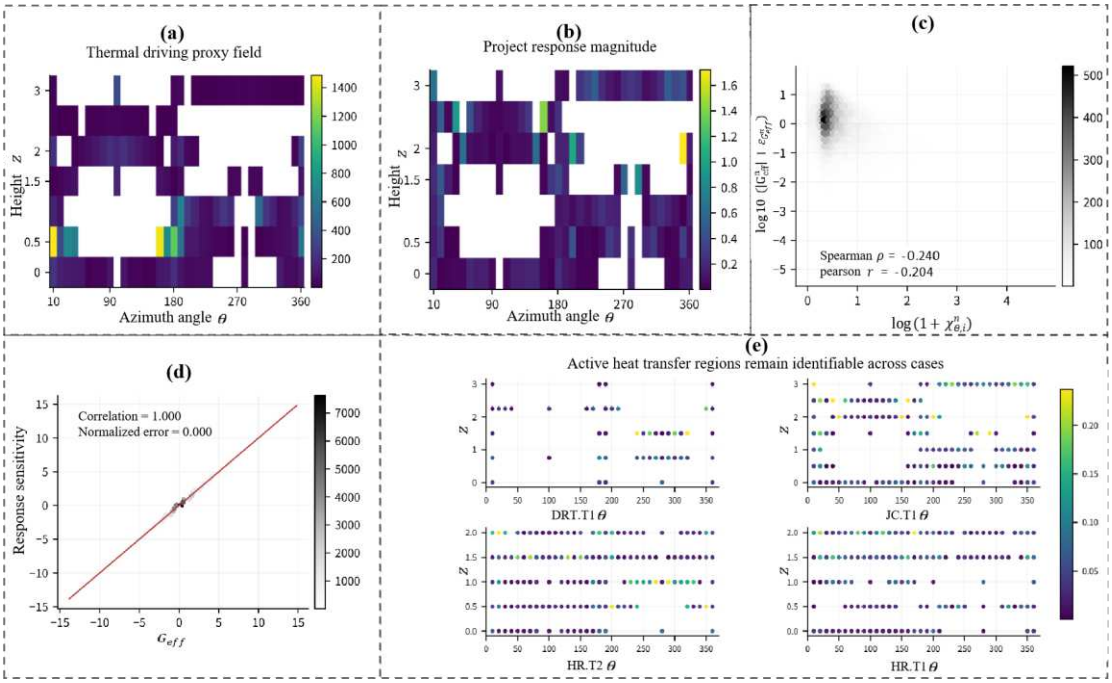


Fig.5(a) Representative thermal driving proxy field $\chi_{\Theta,i}^{n}$ derived from the dimensionless temperature-gradient magnitude. (b) Corresponding projected response magnitude $\left|G_{eff}^{n}\right|$.(c) Statistical association between $\log_{10}(\left|G_{eff}^{n}\right| + \varepsilon_{G_{eff}^{n}})$ and $\log_{10}(1 + \chi_{\Theta,i}^{n})$ within the active region.(d) Consistency between G_{eff} and the response sensitivity computed from $U_i{}^{n}$along the dimensionless temperature-gradient direction.(e) Representative active heat-transfer regions selected by the fixed training-set gradient threshold across different tanks and heating stages. Grey points indicate inactive grid points, and coloured points indicate active grid points.

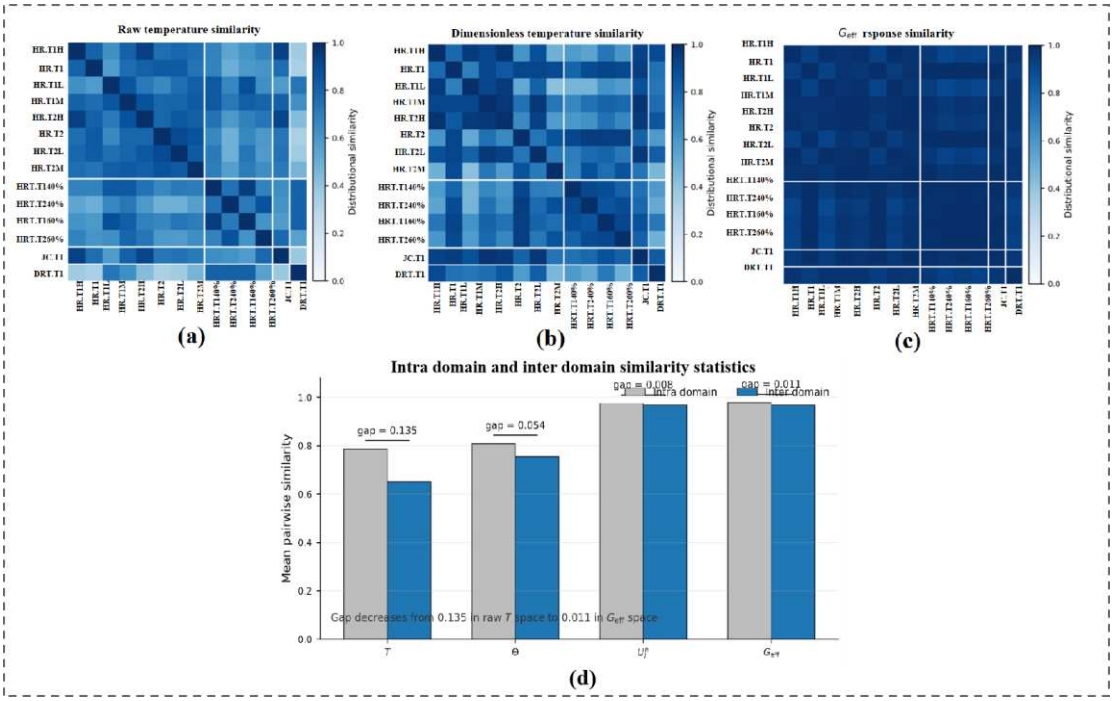


Fig. 6 (a) Dataset similarity matrix calculated from raw temperature distributions. (b) Dataset similarity matrix calculated from dimensionless temperature state distributions. (c) Dataset similarity matrix calculated from the G_{eff} response state distributions. (d) Intra-domain and inter-domain similarity statistics for T, θ, U_i^n, and G_{eff}.

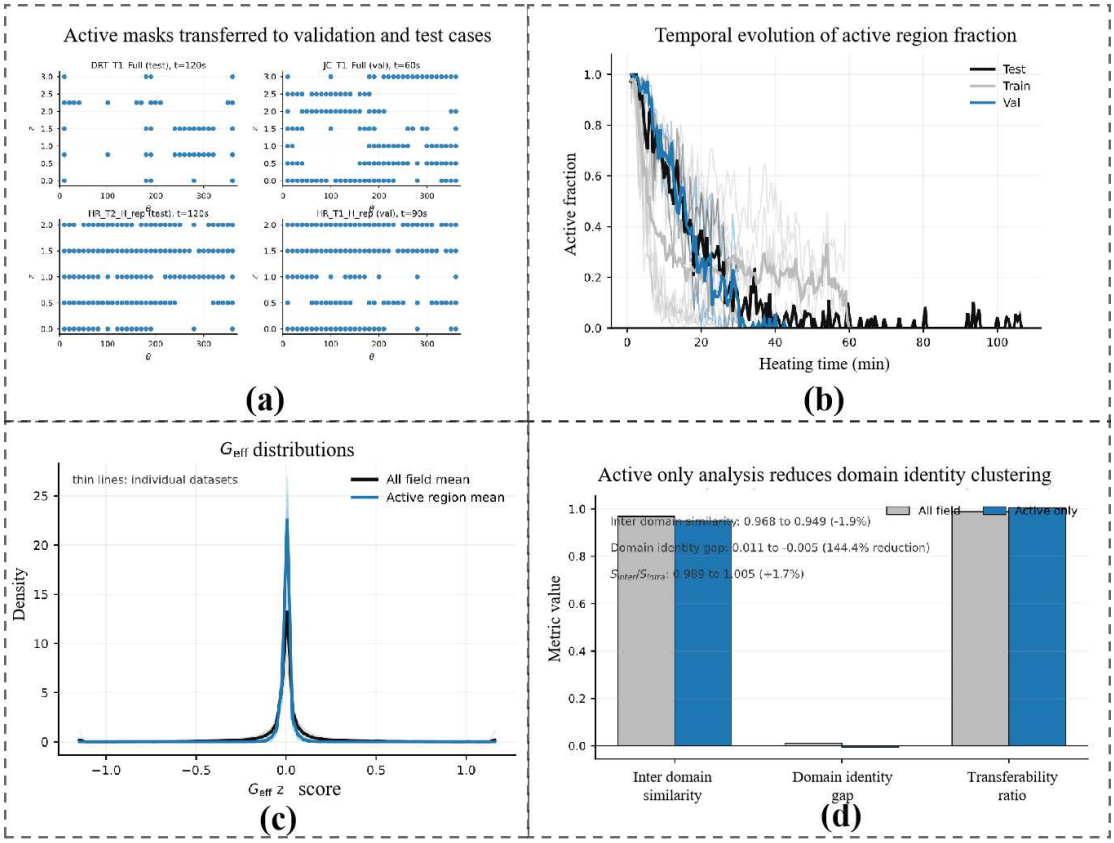


Fig. 7 (a) Representative active masks in validation and test cases obtained using the fixed threshold determined from the training set. (b) Temporal evolution of the active-region fraction during heating.Thin curves represent individual datasets and thick curves represent split-level averages. (c) Distributions of G_{eff} for all-field and active-region samples.Thin curves denote individual datasets and thick curves denote mean distributions. (d) Quantitative comparison between all-field and active-only G_{eff} similarity statistics.

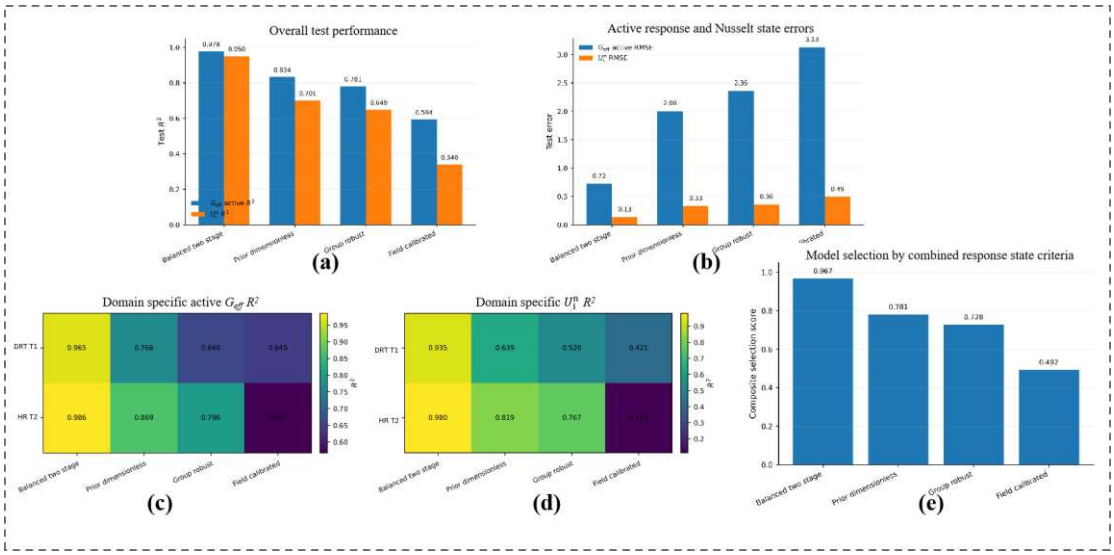


Fig. 8 (a) Overall test R^2 for active-region G_{eff} and U_i^n prediction. (b) Overall test RMSE for active-region G_{eff} and U_i^n. (c) Domain-specific active-region G_{eff} R^2 on DRT.T1 and HR.T2.H.rep. (d) Domain-specific U_i^n R^2 on DRT.T1 and HR.T2.H.rep. (e) Composite model selection score combining response-state and Nusselt state criteria.

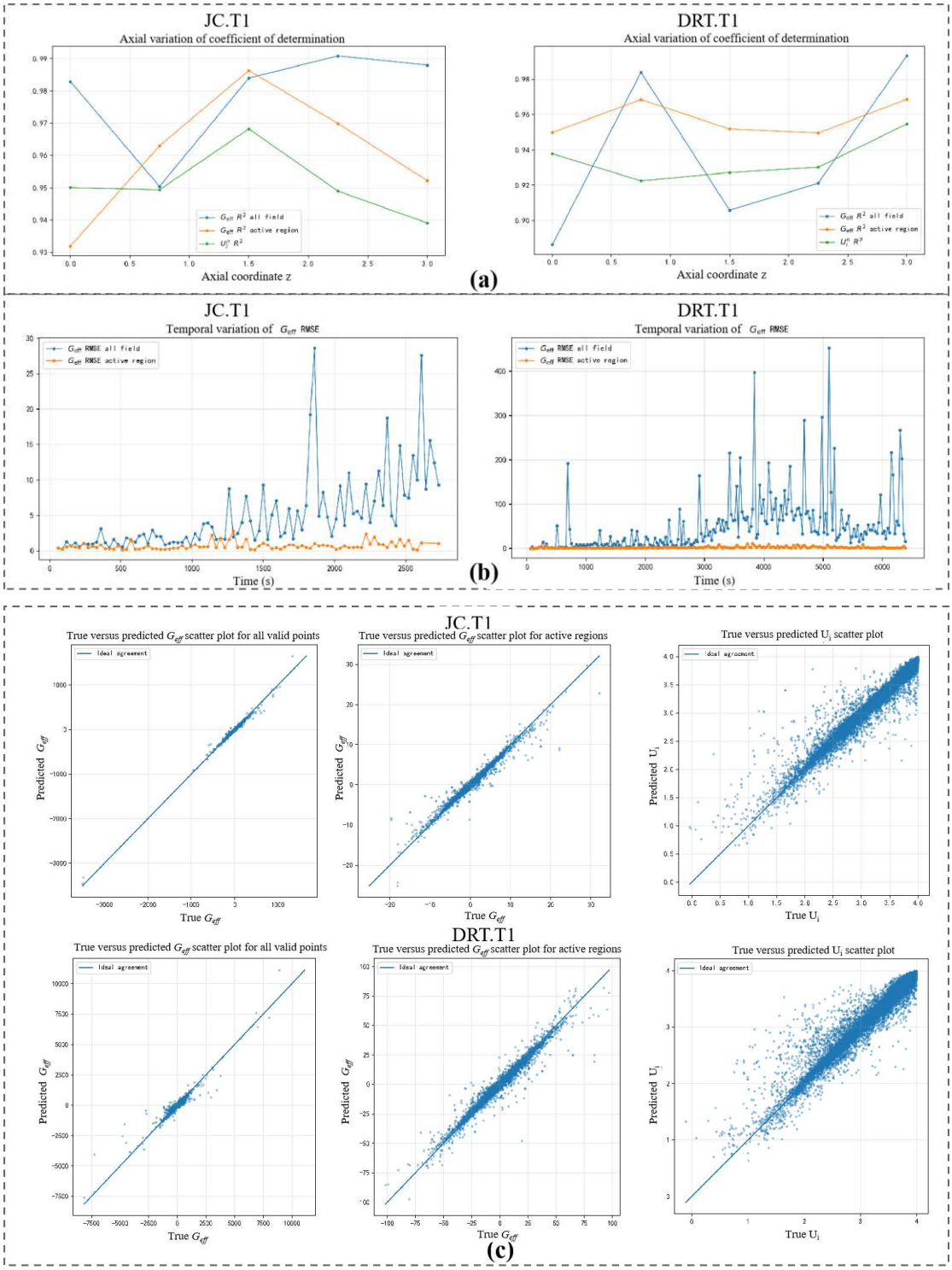


Fig. 9 (a) Axial variation of G_{eff} RMSE for all-field and active-region points in JC.T1 and DRT.T1. (b) Temporal variation of G_{eff} RMSE during heating. (c)True-versus-predicted scatter plots for all valid G_{eff} points, active-region G_{eff} points and the logarithmic effective Nusselt state U_i^n.

Tab.1 Summary Table of Sample datasets for Multiple Tank Bodies and Various Operating Conditions.

Manufacturer ID	Acquisition time	Ambient temperature	Diameter of the stem	Height of the stem	Fluid	Liquid volume	Heating duration	Valve opening
HR.T1	2025.9	32℃	1.3m	2.9m	Water	2700L	28min	Auto control
HR.T2	2025.9	32	1.1m	3m	Water	2700L	30min	Auto control
HR.T1 H	2026.2	2℃	1.3m	2.9m	Water	2700L	27min	Auto control
HR.T2 H	2026.2	2℃	1.1m	3m	Water	2700L	41.5min	Auto control
HR.T1 M	2026.2	2℃	1.3m	2.9m	Water	2036.3 L	21.5min	Auto control
HR.T2 M	2026.2	2℃	1.1m	3m	Water	2224.8 L	18min	Auto control
HR.T1 L	2026.2	2℃	1.3m	2.9m	Water	1372.7 L	27.5min	Auto control
HR.T2 L	2026.2	2℃	1.1m	3m	Water	1749.6 L	27min	Auto control
HRT.T1 40%	2024.8	30℃	1.4m	3m	Water	3500L	47.5min	40%
HRT.T2 40%	2024.8	30℃	1.4m	3m	Water	3500L	60min	40%
HRT.T1 60%	2024.8	32℃	1.4m	3m	Water	4500L	60min	60%
HRT.T2 60%	2024.8	32℃	1.4m	3m	Water	4500L	60min	60%
JC.T1	2025.4	15℃	1.4m	3m	Water	4015.1 L	40min	80%
DRT.T1	2024.7	27℃	1.4m	3m	Water	2703L	120min	40%

Tab. 1 Instrument parameters measured

Instrument	Parameter
DF-101T Heating Kettle	Inner diameter：345 mm Outer diameter：350 mm High：210 mm
Round-Bottom Bottle	Outer diameter：230 mm

Tab. 2 Results of specific heat capacity measurements

XIAMITER PMX-200 100cs Silicone Fluid	
Temperature (℃)	Specific Heat (J/kg·k)
28.62	1566.5
30.96	1523.1
54.78	1878.3
72.46	2003.7
80.38	1543.2

Tab. 3 Error between predicted and actual values of oil temperature at 200 steps

200 steps Oil temperature error			
Time（min）	A Error (℃)	B Error (℃)	C Error (℃)
1	-1.0	1.0	-1.5
2	0.8	0.9	0.4
3	1.9	0.8	1.9
4	2.0	0.7	1.3
5	-0.9	0.6	-2.4
6	2.9	0.5	-0.8
7	2.9	0.4	0.4
8	-0.9	0.4	1.1
9	1.0	0.3	-2.4
10	2.8	0.2	0.1
11	0.6	0.2	2.7

12	4.3	0.1	1.5
13	4.4	-0.2	-2.3
14	3.6	-0.5	-1.5
15	1.3	-0.9	1.0
16	1.3	-1.3	0.3
17	-1.6	-1.5	-0.2
18	1.5	-1.6	-2.9
19	2.6	-1.6	-0.8
20	0.7	-1.8	0.0

Tab. 4 Error between predicted and actual values of extraction liquid temperature at 200 steps

200 steps Extraction liquid temperature error			
Time（min）	A Error (℃)	B Error (℃)	C Error (℃)
1	0.5	-1.0	-0.6
2	0.5	-1.0	0.1
3	0.6	-1.0	0.1
4	0.7	-1.0	0.3
5	0.7	-1.1	-0.4
6	0.5	-1.0	-0.1
7	0.6	-1.0	0.5
8	0.4	-1.1	0.7
9	0.4	-1.1	-0.2
10	0.5	-1.1	0.0
11	0.4	-1.1	-0.1
12	0.7	-1.1	0.0
13	0.9	-1.0	-0.4
14	0.9	-0.9	0.1
15	0.3	-0.8	-0.3

16	0.4	-0.7	-0.2
17	0.3	-0.6	-0.7
18	0.6	-0.6	-1.0
19	0.6	-0.6	-0.6
20	0.8	-0.6	-0.4

Tab. 5 Extracting the 90 minute changes in Paeoniflorin content from PID and LSTM-PID

Paeoniflorin content（mg/ml）		
Time(min)	PID	LSTM-PID
15	1.85	1.70
30	2.05	2.16
45	2.05	2.30
60	2.40	2.50
75	2.82	2.48
90	2.41	2.46