

Article

Not peer-reviewed version

---

# Optimization of Sparse Sensor Layouts and Data-Driven Reconstruction Methods for Steady-State and Transient Thermal Field Inverse Problems

---

[Qingyang Yuan](#) , Peijun Yao , [Wenjun Zhao](#) , [Bo Zhang](#) \*

Posted Date: 9 July 2025

doi: 10.20944/preprints202507.0825.v1

Keywords: Temperature field inverse problem; sparse sensor layout optimization; data-driven methodologies; Gappy POD; Gappy C-POD; POD-Radial Basis Function (RBF); POD-Multilayer Perceptron (MLP)



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

## Article

# Optimization of Sparse Sensor Layouts and Data-Driven Reconstruction Methods for Steady-State and Transient Thermal Field Inverse Problems

Yuan Qingyang <sup>1,3</sup>, Yao Peijun <sup>2</sup>, Zhao Wenjun <sup>1</sup> and Zhang Bo <sup>1,2,\*</sup>

<sup>1</sup> Key Laboratory of Complex Energy Conversion and Efficient Utilization of Liaoning Province, School of Energy and Power Engineering, Dalian University of Technology, Dalian 116081, China

<sup>2</sup> Zhejiang SHIP Electronics Technology Co., Ltd, Ningbo, Zhejiang, China;

<sup>3</sup> Ningbo Research Institute, Dalian University of Technology, Ningbo 315032, China

\* Correspondence: 18842396497@163.com

## Abstract

This paper addresses the inverse problem of temperature fields in both steady-state and transient heat conduction scenarios. A comprehensive approach is proposed, integrating the optimization of sparse sensor layouts with a data-driven reconstruction methodology, aiming to enhance the accuracy and computational efficiency of temperature field reconstruction. Initially, models incorporating various irregular heat source distributions and differing boundary conditions for steady-state and transient heat conduction are constructed, with numerical solutions obtained through the finite difference method. Following this, the study systematically investigates the impact of various sensor layout optimization strategies on reconstruction accuracy, including random point selection, S-OPT method, correlation coefficient filtering method (CCFM), and uniform point selection. Additionally, diverse sampling methods such as Latin hypercube sampling, Sobol sequence sampling, and maximum-minimum distance sampling are employed to establish a multifaceted database construction scheme. Experimental results demonstrate the robustness of the Gappy Proper Orthogonal Decomposition (POD) and Gappy C-POD methods across different modal conditions, particularly exhibiting strong stability when the number of modes is low (1-5 modes). The Gappy C-POD method, when paired with CCFM point selection and maximum distance sampling, effectively mitigates reconstruction deviations. Conversely, while the POD-MLP and POD-RBF methods maintain high overall accuracy at greater modal numbers (>10 modes), they display significant sensitivity to sensor placement and database size. Based on the findings of this study, the Gappy C-POD method is recommended as a mainstream approach for tackling the inverse problem of temperature fields, particularly when coupled with well-considered sensor layout optimization strategies to enhance both the accuracy and efficiency of field reconstruction. This research provides a novel perspective on solving heat conduction inverse problems, particularly highlighting the theoretical importance of integrating data-driven methods with sensor layout optimization.

**Keywords:** temperature field inverse problem; sparse sensor layout optimization; data-driven methodologies; Gappy POD; Gappy C-POD; POD-Radial Basis Function (RBF); POD-Multilayer Perceptron (MLP)

## 1. Introduction

Heat conduction problems are prevalent across various engineering applications, including mechanical engineering, architectural engineering, and the thermal management of electronic devices. The inverse problem of temperature fields [1,2], which involves deducing the distribution of

heat sources and boundary conditions from known temperature data, is a critical technique for addressing these challenges. Traditional numerical simulation methods, particularly those based on finite difference methods (FDM)[3] and finite element methods (FEM)[4], often rely on comprehensive sensor networks and detailed boundary conditions. However, in practical scenarios, sensor deployment frequently suffers from sparsity, making it difficult to accurately capture the complete temperature field information, thereby complicating the solution of the temperature field inverse problem.

Classical methods for solving thermal inverse problems primarily encompass analytical methods[5], regularization techniques[6], optimization strategies[7], Bayesian approaches[8], and iterative inversion methods[9]. Analytical methods yield precise solutions through mathematical derivation; in the fields of fluid flow and heat transfer, Fourier transforms, Green's functions, and Laplace inversion[10] are widely utilized. These methods typically apply to linear problems with simple models and regular boundary conditions[12]. For instance, in steady-state conduction problems where thermal conductivity varies linearly with temperature, inverse analysis can be employed for parameter estimation. Although analytical solutions can offer accurate results with clear physical interpretations, they are mainly constrained by stringent requirements regarding the linearity of the problem and geometric regularity, making them inadequate for complex nonlinear problems and irregular geometries.

Ill-posedness is an inherent challenge of inverse problems, where even minor data noise can lead to significant deviations in solutions. Regularization methods aim to stabilize solutions by introducing additional constraints, thereby enhancing robustness. Tikhonov regularization is one of the most established techniques, which smooths solutions by adding a penalty term to the least squares objective function, achieving a balance between data fitting error and the "smoothness" or "simplicity" of the solution[13]. In the context of heat transfer, regularization methods are extensively employed to identify unknown parameters within mathematical models. For instance, when dealing with ill-posed problems, various regularization techniques can be selected with appropriately adjusted regularization parameters[14]. However, the selection of these parameters often depends on empirical judgment, and inappropriate choices may lead to erroneous results.

Optimization methods convert inverse problems into optimization tasks, estimating unknown parameters by minimizing the residuals between observed data and model predictions. Prominent approaches include nonlinear least squares, constrained optimization, and variational methods[15]. These methods are versatile and theoretically sound, finding broad applications in parameter estimation, imaging, and system identification. For example, in intricate heat transfer scenarios, they can optimize thermal performance designs[16]. Nevertheless, these methods are susceptible to local optima and may struggle with multimodal problems, while computational efficiency is restricted by the problem's scale and complexity[17]. Bayesian methods, rooted in probability theory, treat unknown parameters as random variables, updating posterior probability distributions by synthesizing prior knowledge with observational data to quantify uncertainty. Techniques such as Markov Chain Monte Carlo (MCMC) and variational Bayes are commonly employed[18]. This approach excels in managing complex models and quantifying uncertainties, as exemplified by predicting the energy consumption of HVAC systems using Bayesian networks[19]. The primary challenges arise from the computational demands of high-dimensional parameter spaces and the intricacies involved in model construction. Iterative inversion methods approach the true solution through successive iterations, with prevalent algorithms including Landweber iteration, Conjugate Gradient Method (CGM), and Gauss-Newton methods[20]. These techniques are straightforward to implement, capable of addressing nonlinear problems, and are widely used in image reconstruction, bioelectrical impedance tomography (BIT), and seismic imaging. For example, when solving ill-posed linear inverse problems, a recently proposed second-order dynamic method (SODM) integrates Tikhonov regularization with second-order asymptotic regularization and employs a dual-parameter strategy, demonstrating superiority over conventional methods under noisy conditions[21].

However, it is essential to note that this approach has slower convergence rates, sensitivity to initial values, and demands meticulous parameter tuning[22].

With the advancement of machine learning and data-driven methodologies, data-based field reconstruction techniques have increasingly emerged as effective tools for addressing related challenges. These approaches effectively reconstruct accurate temperature field distributions by leveraging limited sensor data, integrated with dimensionality reduction techniques and model simplification strategies. Specifically, the Proper Orthogonal Decomposition (POD) method[23] facilitates dimensionality reduction of temperature field data, extracting principal feature modes and thereby simplifying complex heat transfer problems. The Gappy POD method[24] is an extension of POD designed to handle incomplete data. In heat transfer scenarios, complete temperature field data may not be obtainable due to sensor malfunctions or data transmission errors. Gappy POD allows for the reconstruction of the full temperature field from available partial data. For example, research conducted by Wang et al.[25] involved deploying a limited number of sensors, combined with Computational Fluid Dynamics (CFD) simulations and the Gappy POD methodology to reconstruct the temperature field distribution across an entire indoor space. Moreover, Xu et al.[26] addressed the severe missing data issues in meteorological observations by generating a POD basis using the WRF model and subsequently employing Gappy POD to fill in the gaps, successfully reconstructing the temperature field in the Tibetan region. The POD-RBF (Proper Orthogonal Decomposition - Radial Basis Function) method effectively reconstructs temperature fields with a small number of measurement points. This technique utilises POD to extract key characteristics of the temperature field and employs RBF interpolation for complete temperature reconstruction, thus striking a balance between computational efficiency and accuracy. Research by Wang et al.[27] demonstrated the use of the POD-RBF method to rapidly reconstruct the temperature field of printed circuit boards (PCBs) based on minimal sensor data, facilitating monitoring and control of thermal behaviour in devices. The Multilayer Perceptron (MLP) represents a type of artificial neural network capable of learning nonlinear relationships between inputs and outputs. The POD-MLP method integrates POD dimensionality reduction with MLP neural networks, aiding in the estimation of physical system states, including thermal and flow conditions. Qi et al.[28] leveraged the POD-MLP technique for real-time monitoring and reconstruction of the surface temperature field of batteries. The accuracy of data-driven field reconstruction methods significantly depends on the sparse sensor layout. A strategically designed sensor arrangement can not only enhance the precision of temperature field reconstruction but also effectively reduce data acquisition costs and computational burdens. However, many existing reconstruction methods overlook the influence of sparse sensor configurations, often resorting to uniform or random placement techniques, which can lead to considerable variability in reconstruction accuracy. Willcox et al.[29] optimized the sparse sensor placement for the Gappy POD method using matrix stability criteria, specifically condition numbers, achieving commendable results; however, this approach demands iterative and repeated matrix computations, resulting in lower time efficiency. It also necessitates integration with optimization methods to accelerate the selection process in complex scenarios. Lauzon et al.[29] introduced the S-OPT method, which quickly determines optimal sensor configurations by scanning orthogonal basis matrices to maximise an S-metric. Yet, this method proves sensitive to the dimensionality of the basis, requiring higher-order modes for accurate results in complex fields such as turbulence. Yuan et al.[30] proposed a method utilising the Pearson correlation coefficients of matrices to gauge the interrelationships between sensors and employed a threshold-based strategy to swiftly identify optimal sensor placement locations, known as the Correlation Coefficient Filtering Method (CCFM). This method enables sparse sensor layout optimization under low-modal and confined regional conditions.

Building upon the clustering-based dimensionality reduction method guided by pod structures proposed (C-POD) by Yuan et al.[30], this paper introduces the Gappy C-POD approach, incorporating principles from inverse problem solving, and develops a multi-scale sample database alongside various sensor layout optimization strategies to facilitate its implementation. This study

systematically compares the reconstruction pathways, accuracy, and robustness of four data-driven reconstruction methods—Gappy POD, Gappy C-POD, POD-RBF, and POD-MLP—across different modal truncation numbers, database sizes, and selection strategies, focusing on steady-state and transient heat conduction problems with multiple internal heat sources in two dimensions. The findings aim to provide theoretical foundations and methodological guidance for the engineering modelling and optimisation design of temperature field inverse problems.

## 2. Overview of Gappy C-POD and Data-Driven Field Reconstruction Methods

### 2.1. Introduction to POD and Gappy POD Methods

Proper Orthogonal Decomposition (POD), introduced by Sirovich et al.[31], is a well-established technique for dimensionality reduction and feature extraction. It has become widely adopted for modal decomposition and data compression of high-dimensional physical fields. The POD method decomposes fields, such as temperature distributions, to identify modal bases that encapsulate where energy is predominantly concentrated, thereby characterising the primary structures and dynamics of the field. For a snapshot matrix of the temperature field containing  $N_s$  samples  $T = [T_1, T_2, \dots, T_{N_s}]$ , its covariance matrix  $R$  can be expressed as follows:

$$R = \frac{1}{N_s} T T^T \quad (1)$$

Next, by addressing the eigenvalue problem presented in the subsequent equation(2):

$$R\phi = \lambda_i \phi \quad (2)$$

one can derive the basis functions  $\phi_i$  along with their corresponding energy  $\lambda_i$ . The temperature field can be expanded in terms of these modes as follows:

$$T \approx \sum_{i=1}^r a_i \phi_i = \phi_r a \quad (3)$$

In this expression, the modal basis matrix  $\phi_r = [\phi_1, \dots, \phi_r]$  retains the first  $r$  modes, with  $a$  being the modal coefficient vector. In practical scenarios, due to limitations in data acquisition, temperature observations are often sparse or incomplete. The Gappy POD method was developed to address this issue, aiming to reconstruct the complete field under sparse observation conditions. Let us denote the sparse observed temperatures as  $T_{obs}$ , with the observation matrix  $C$  (which extracts the relevant rows corresponding to the observed points); the objective of Gappy POD is to minimise the residuals at the observed points as follows:

$$\min_a \|C\phi_r a - T_{obs}\|^2 \quad (4)$$

The least-squares solution can be computed using:

$$a = (C\phi_r)^+ T_{obs} \quad (5)$$

Here, the operation  $(\cdot)^+$  signifies the generalized inverse (pseudo-inverse). Consequently, the complete reconstructed temperature field takes the form:

$$T_{recon} = \phi_r a \quad (6)$$

The Gappy POD method demonstrates robust reconstruction capabilities even in cases of data sparsity or local deficiencies. The accuracy of the reconstruction is influenced by the energy distribution of the modal basis, the configuration of the observation points, and the numerical condition of the observation matrix. Therefore, an effective strategy for optimising the layout of sparse sensors is crucial to ensuring optimal performance of the Gappy POD approach.

## 2.2. Introduction to C-POD and Gappy C-POD Methods

The linear assumptions underlying traditional POD constrain its effectiveness in reconstructing fields characterized by strong nonlinearity. To overcome this limitation, Yuan et al. [30] introduced the C-POD method, which combines the inherent orthogonality of POD with the nonlinear feature extraction capabilities of clustering techniques. Initially, C-POD utilizes the low-dimensional modal basis generated by POD as the starting point for clustering, selecting the first  $K$  bases as the initial cluster centers  $\mu_i^{(0)}$ :

$$\mu_i^{(0)} = \phi_i, i = 1, 2, \dots, K \quad (7)$$

The iterative update of these clustering centers can be expressed by:

$$\mu_i^{(t)} = \frac{1}{|S_i^{(t)}|} \sum_{T_j \in S_i^{(t)}} T_j \quad (8)$$

Here, the notation  $S_i^{(t)}$  represents the collection of temperature samples belonging to the  $i$ th cluster at the  $t$ -th iteration, while denotes the  $j$ -th temperature field sample  $T_j$ . By projecting each group of temperature field samples onto the clustering centres, we derive the modal coefficients associated with the C-POD method, articulated as follows:

$$a_j = (\mu^T \mu)^{-1} \mu^T T_j \quad (9)$$

Given the observation matrix  $C$  that extracts sparse temperature observations, the observed temperature data can be formulated as:

$$T_{obs} = CT \quad (10)$$

This leads to the modal coefficient equations(11) and (12) under sparse observation conditions:

$$a = \arg \min_a \|C\mu a - T_{obs}\|^2 \quad (11)$$

$$a = (C\mu)^+ T_{obs} \quad (12)$$

The resulting reconstructed temperature field can be expressed as:

$$\hat{T} = \mu a \quad (13)$$

The clustered modal basis  $\mu$  in C-POD encapsulates the local nonlinear characteristics inherent in the data, effectively overcoming the limitations of the linear orthogonal basis expansion employed by traditional POD. The Gappy C-POD method further enhances this by capturing local features and multi-state evolution of nonlinear flow fields under conditions of sparse observation.

## 2.3. Solutions to Inverse Problems using POD-RBF and POD-MLP Methods

In the context of data-driven inverse problem solving, both POD-RBF and POD-MLP methods establish nonlinear mapping models between inputs (sparse observational data) and outputs (modal coefficients or temperature fields), thereby circumventing the constraints of traditional linear least squares POD reconstruction techniques in nonlinear, complex boundary, or highly perturbed fields.

### 2.3.1. Solving Inverse Problems with the POD-RBF Method

The Radial Basis Function (RBF) neural network is known for its simplicity and adaptability to small sample datasets. The POD-RBF method begins by extracting the POD modal coefficients  $\mathbf{a}$ :

$$\mathbf{T} \approx \boldsymbol{\phi} \mathbf{a} \quad (14)$$

Following this, a Radial Basis Function (RBF) network is established to create a nonlinear mapping between the sparse observational data  $\mathbf{T}_{\text{obs}}$  and the modal coefficients  $\mathbf{a}$ :

$$\mathbf{a} = \sum_{i=1}^{N_c} \omega_i \boldsymbol{\phi}(\|\mathbf{T}_{\text{obs}} - \mathbf{C}_i\|) \quad (15)$$

In this equation,  $\boldsymbol{\phi}(\cdot)$  stands for the RBF kernel function;  $\mathbf{C}_i$  signifies the centres;  $\omega_i$  represents the weight parameters; and  $N_c$  indicates the number of centres. The reconstructed temperature field is detailed as follows:

$$\hat{\mathbf{T}} = \boldsymbol{\phi} \mathbf{a} \quad (16)$$

### 2.3.2. Solving Inverse Problems with the POD-MLP Method

In comparison to RBF methods, Multi-Layer Perceptron (MLP) neural networks exhibit superior fitting capability and are more suitable for larger sample datasets. Using the sparse observational data  $\mathbf{T}_{\text{obs}}$  as input, a Multi-Layer Perceptron (MLP) neural network is trained:

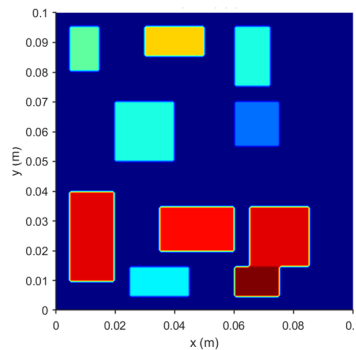
$$\mathbf{a} = f_{\text{MLP}}(\mathbf{T}_{\text{obs}}; \boldsymbol{\theta}) \quad (17)$$

In this expression,  $f_{\text{MLP}}$  refers to the MLP network function, while  $\boldsymbol{\theta}$  denotes the network weights and bias parameters, respectively. The MLP network, through multiple layers of nonlinear activation units, approximates the complex mapping relationship between observation points and modal coefficients. Upon completion of training, it can swiftly predict modal coefficients and reconstruct the complete field.

## 3. Reconstruction of Steady-State Temperature Field

### 3.1. Case Study Overview

To evaluate the applicability and robustness of the methodology for solving inverse problems related to temperature fields, this study presents a classic example of a two-dimensional steady-state heat conduction problem, as illustrated in Figure 1. The model simulates the heat transfer process influenced by multiple irregularly distributed heat sources and certain isothermal boundary conditions, employing the finite difference method for numerical analysis.



**Figure 1.** Two-dimensional steady-state heat conduction problem.

3.1.1. Geometric Parameters and Heat Source Distribution

As depicted in Figure 1, the computational domain in this case has a side length of  $L = 0.1$  m. Within this domain, there are ten rectangular heat sources, the locations, dimensions, and intensity of which are specified in Table 1:

**Table 1.** Properties and intensity parameters of heat sources.

| Heat source<br>identification<br>number | Starting position | Dimensions     | Intensity<br>$\phi$ (W/m <sup>2</sup> ) |
|-----------------------------------------|-------------------|----------------|-----------------------------------------|
| 1                                       | (0.005, 0.005)    | (0.010, 0.015) | $0.5 \times 10^5 \sim 2.0 \times 10^5$  |
| 2                                       | (0.030, 0.005)    | (0.020, 0.010) |                                         |
| 3                                       | (0.060, 0.005)    | (0.012, 0.020) |                                         |
| 4                                       | (0.020, 0.030)    | (0.020, 0.020) |                                         |
| 5                                       | (0.060, 0.030)    | (0.015, 0.015) |                                         |
| 6                                       | (0.005, 0.060)    | (0.015, 0.030) |                                         |
| 7                                       | (0.035, 0.065)    | (0.025, 0.015) |                                         |
| 8                                       | (0.065, 0.065)    | (0.020, 0.020) |                                         |
| 9                                       | (0.025, 0.085)    | (0.020, 0.010) |                                         |
| 10                                      | (0.060, 0.085)    | (0.015, 0.010) |                                         |

3.1.2. Governing Equation

The governing equation is a nonlinear steady-state heat conduction equation, represented as shown in Equation (18):

$$\nabla \cdot (\lambda(T)\nabla T) + \Theta(x, y) = 0, (x, y) \in \Omega$$

(18)

Where  $\Omega = [0, L] \times [0, L]$  is defined as a two-dimensional square area;  $\lambda(T) = 1 + 0.05(T - 298)$  and  $\Theta(x, y)$  is the heat source term is included.

3.1.3. Initial Conditions

The initial temperature field is uniformly distributed, as expressed in Equation (19):

$$T(x, y) = 298K, (x, y) \in \Omega$$

(19)

3.1.4. Boundary Conditions

The boundary conditions are specified as mixed. The left and right boundaries are characterised by adiabatic conditions:

$$\frac{\partial T}{\partial x} = 0$$

(20)

The upper boundary also maintains an adiabatic condition:

$$\frac{\partial T}{\partial y} = 0$$

(21)

For the lower boundary, most of it is adiabatic, except for the central position which is kept at a constant temperature, as illustrated in Equation (22):

$$T = 298K, x \in [\frac{L}{2} - 0.005, \frac{L}{2} + 0.005]$$

(22)

### 3.1.5. Solution Method

The solution is obtained through the finite difference method:

- Grid: A uniform rectangular grid of 100×100 is employed ;
- Spatial Discretisation: A nonlinear five-point difference scheme is used ;
- Nonlinear Solving: The Gauss-Seidel iterative method is applied as per Equation (23):

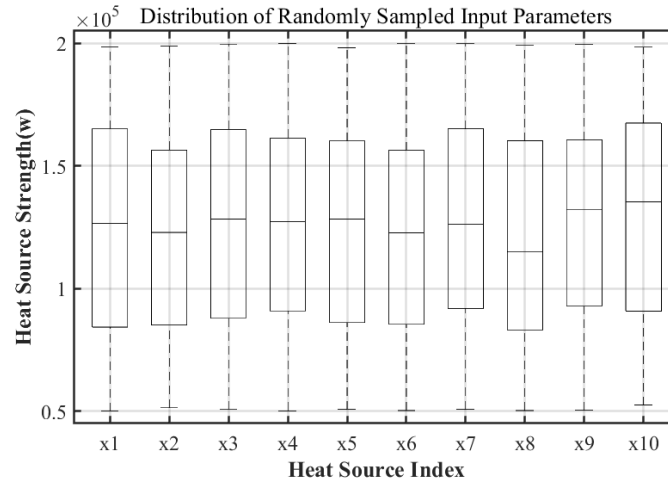
$$T_{i,j} = \frac{\lambda_E T_{i,j+1} + \lambda_W T_{i,j-1} + \lambda_N T_{i-1,j} + \lambda_S T_{i+1,j} + \Delta x^2 \Theta_{i,j}}{\lambda_E + \lambda_W + \lambda_N + \lambda_S} \quad (23)$$

where  $\lambda_E = 1 + 0.05(T_{i,j+1} - 298)$ , the representation of other directions is analogous. The convergence criterion is established in Equation (24), with a maximum iteration limit set to 10,000.

$$\max |T^{(n+1)} - T^{(n)}| < 10^{-5} K \quad (24)$$

### 3.2. Construction of Verification Set

As indicated in section 2.1, the thermal conduction problem incorporates ten internal heat sources as independent variables. A test set comprised of 200 randomly sampled groups has been created, with the distribution depicted in Boxplot 2. The boxplot illustrates the thermal power values for each heat source, uniformly spanning the defined range of  $(0.5-2.0) \times 10^5$  W, devoid of outliers. This indicates sufficient input coverage for both model training and validation purposes.



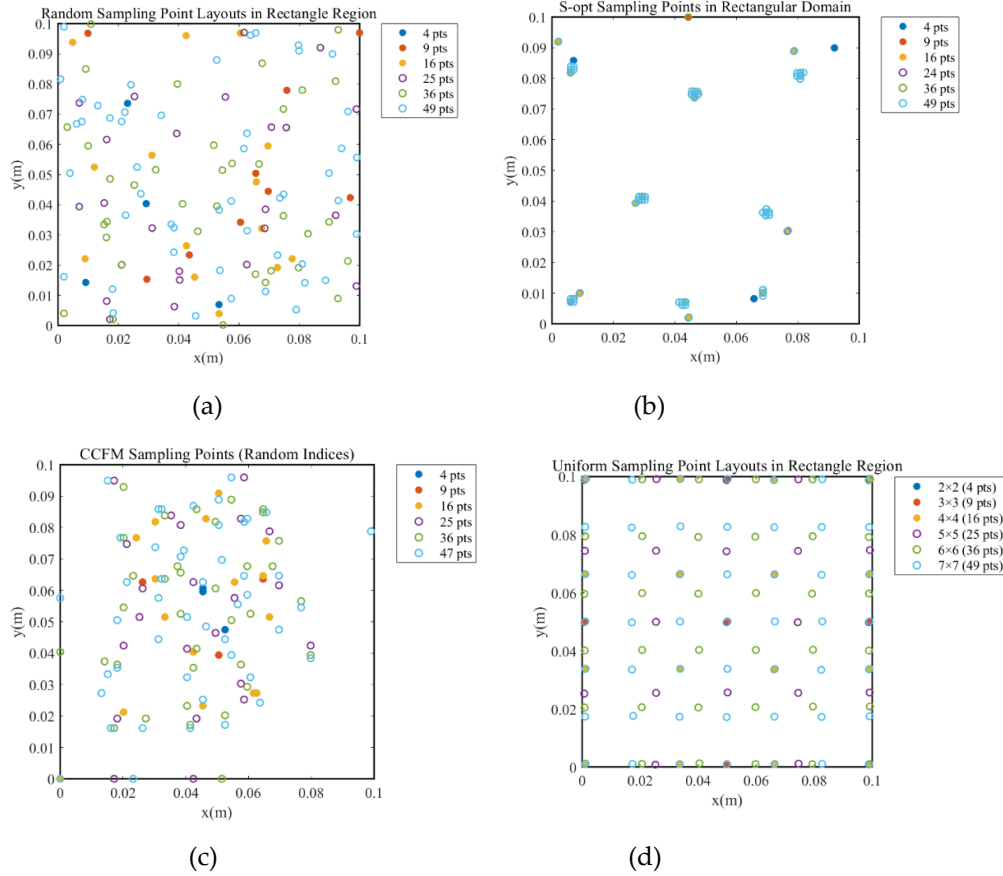
**Figure 2.** Distribution of the test set input parameters (boxplot).

### 3.3. Database Construction

To establish a representative and widely encompassing sample database, this study employs three typical sampling methods: Latin Hypercube Sampling[32], Sobol Sequence Sampling[33], and Maximum Distance Sampling[34]. These methods are applied to sample the combination space of ten thermal power variables. Given the dimensionality of the variables, denoted as  $d=10$ , sample sizes are set at five, ten, twenty, thirty, forty, and fifty times the number of variables, resulting in the generation of 50, 100, 200, 300, 400, and 500 sets of independent variable combinations. This yields a total of 18 multi-scale sample databases that will be utilized for subsequent modeling and analysis.

### 3.4. Sparse Sensor Layout Optimization Methods

To investigate the impact of different sparse sensor layout optimization methods on field reconstruction performance, four representative sensor placement strategies are selected: random sampling, S-OPT method, Correlation Coefficient Filtering Method (CCFM), and uniform sampling. For each strategy, the number of sensors is varied at 4, 9, 16, 25, 36, and 49, resulting in a total of 24 distinct sparse sensor layout configurations, as illustrated in Figure 3.



**Figure 3.** Layouts of four types of sparse sensors: (a) Random point selection; (b) S-OPT method; (c) CCFM method; (d) Uniform point selection.

### 3.5. Reconstruction Methods

To evaluate the reconstruction capabilities of various methods, this study compares four representative approaches: Gappy Proper Orthogonal Decomposition (POD), Gappy Composite Proper Orthogonal Decomposition (C-POD), POD-Radial Basis Function (RBF), and POD-Multilayer Perceptron (MLP). Under the conditions specified in items (3) to (5), a systematic investigation was conducted based on the validation set detailed in Section 2. This research focused on the impact of various factors on reconstruction accuracy, including different dimensional reduction modes, database construction methods, sparse sensor layout optimization strategies, and reconstruction methods. To quantify the performance of each reconstruction approach concerning maximum deviation constraints and overall accuracy, the study employs two key metrics: the minimum maximum relative error (Minimum Maximum Relative Error, MRE) and the minimum average relative error (Minimum Average Relative Error, MARE), defined mathematically by equations (25) and (26):

$$\text{MRE} = \max_i \left( \frac{|x_i^{\text{approx}} - x_i^{\text{true}}|}{|x_i^{\text{true}}|} \right) \tag{25}$$

$$\text{MARE} = \frac{1}{n} \sum_{i=1}^n \left( \frac{|x_i^{\text{approx}} - x_i^{\text{true}}|}{|x_i^{\text{true}}|} \right) \tag{26}$$

where n represents the total number of samples; the reconstruction value for the i-th temperature measurement point is represented as  $x_i^{\text{approx}}$ ; and the actual value for the i-th temperature measurement point is represented as  $x_i^{\text{true}}$ .

The reconstruction approach is denoted as x\_y\_z\_m, where x indicates the reconstruction method (Gappy POD, Gappy C-POD, POD-RBF, POD-MLP), y refers to the database size (5, 10, 20, 30, 40, 50 times the number of independent variables), z denotes the point selection strategy (CCFM, S-OPT, Random, Uniform), and m indicates the number of selected points (4, 5, 16, 25, 36, 49).

3.6. Analysis of Optimal Reconstruction Approaches for Different Modal Counts

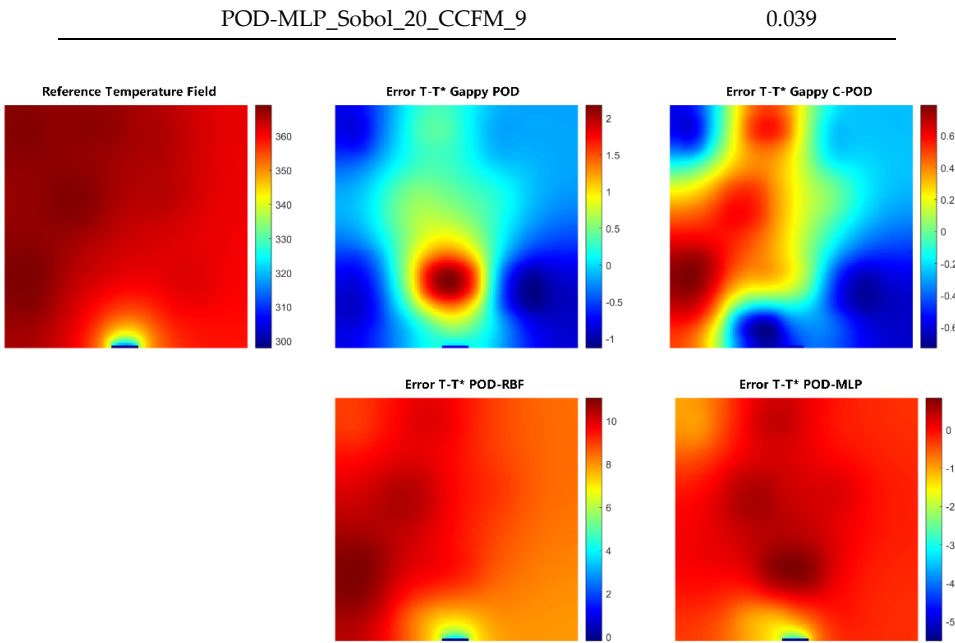
To analyze the performance variations of different reconstruction methods within varying modal ranges, this study collects data on the minimum maximum relative error (Min(MRE)) and minimum average relative error (Min(MARE)) for modes ranging from 1 to 5, 5 to 10, 10 to 15, and 15 to 20. A smaller Min(MRE) indicates a superior capability of the reconstruction scheme in limiting reconstruction deviation within the corresponding mode range (i.e., higher stability), while a smaller Min(MARE) reflects greater overall reconstruction accuracy. The results indicate that the Min(MRE) for all methods does not decrease with an increase in modal count. Notably, at the 15th modal count, all methods achieved the strongest capacity for limiting reconstruction deviation. Table 2 presents the optimal implementation paths and corresponding error values for each reconstruction method when reaching Min(MRE) within the 15-mode range. It is evident from the table that the approach Gappy C-POD\_Maximum\_50\_CCFM\_49 demonstrates the most effective deviation control, suggesting that this scheme is less prone to significant errors during the reconstruction process.

Further observations reveal that under the condition of achieving minimal reconstruction error, the optimal point selection strategy employed by all reconstruction methods is the Correlation Coefficient Filtering Method (CCFM). An analysis combining Figures 1 and 3 suggests that the CCFM method exhibits a higher concentration of selected points in the heat source distribution area, thereby enhancing its ability to capture critical physical features and improving overall reconstruction quality. Additionally, in scenarios with a limited number of modes, the POD-MLP method effectively suppresses deviations even when relying on a small-scale database (20 times the number of independent variables) and a minimal number of sensor points (9). This finding further indicates that utilizing Maximum distance sampling and Sobol sequence sampling for database construction under these conditions yields superior results.

Figure 4 illustrates the distribution of absolute reconstruction errors across various methods on the validation set when the Gappy C-POD\_Maximum\_50\_CCFM\_49 approach achieves Min(MRE). It is evident from the figure that the maximum reconstruction error of the Gappy C-POD method is less than 0.72°C, significantly outperforming the other three reconstruction methods, thereby reaffirming its superiority in low-order modes.

**Table 2.** Minimum MRE values achieved by various reconstruction methods under low-order modal conditions.

| Implementation strategies      | Min (MRE) |
|--------------------------------|-----------|
| Gappy POD_Maximum_10_CCFM_49   | 0.0375    |
| Gappy C-POD_Maximum_50_CCFM_49 | 0.0373    |
| POD-RBF_Sobol_50_CCFM_25       | 0.0455    |



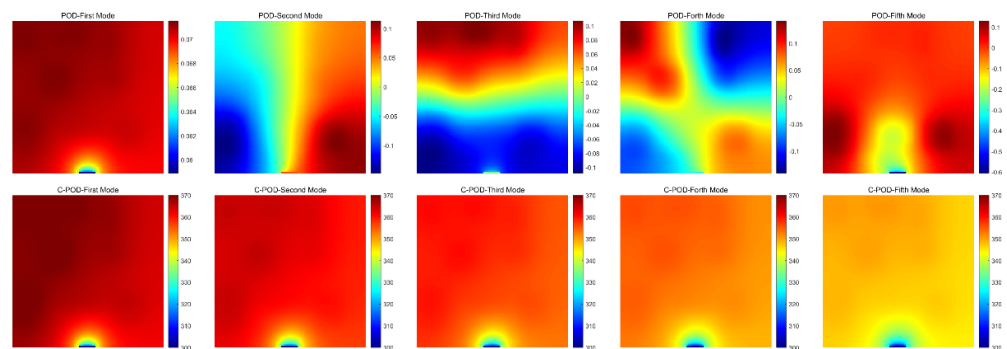
**Figure 4.** Distribution of reconstruction errors for each method on the validation set.

The temperature field modes obtained through the dimensionality reduction techniques of POD and C-POD are illustrated in Figure 5. This figure displays the characteristic modal structures extracted from temperature field samples using both POD and C-POD methodologies. The upper section of the figure presents the first to  $n$ th modes extracted via the POD approach, while the lower section showcases the corresponding clustered modes generated by the C-POD method.

From the observations in Figure 5, the following conclusions can be drawn:

- (1) The POD modes (upper section) exhibit commendable spatial smoothness and energy concentration, with the lower-order modes primarily reflecting the temperature gradient variations aligned with the main direction of the heat sources, thereby encapsulating the overall characteristics of the temperature field. In contrast, the higher-order modes contain more detailed local perturbation information, although their physical interpretability is comparatively limited;
- (2) The C-POD modes (lower section) display a greater complexity in spatial structure, highlighting pronounced local features that enable precise focus on areas of intense heat source concentration or significant temperature fluctuations. Particularly under complex boundary conditions or irregular heat source distributions, the C-POD modes demonstrate enhanced regional adaptability, with modal shapes more closely resembling the nonlinear distribution of actual physical fields.

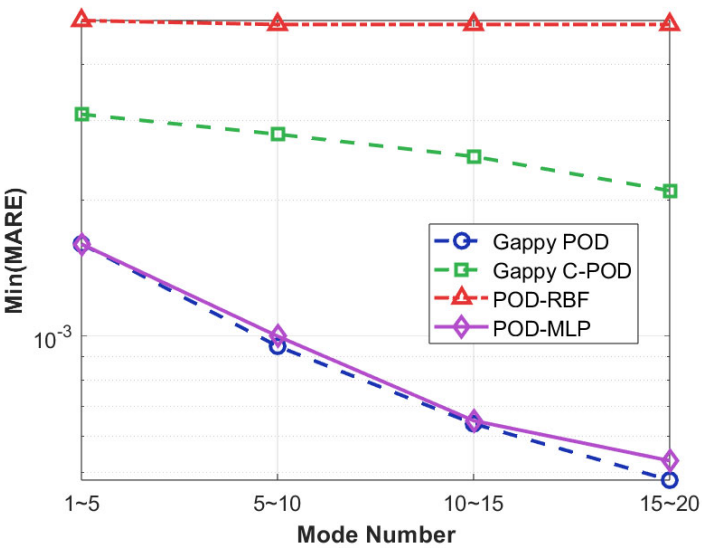
From the perspective of modal structural characteristics, the POD approach is more effective for extracting dominant global features, benefitting from strong mathematical orthogonality and energy spectral decay properties. Conversely, the C-POD modes exhibit superior capabilities in local feature extraction and nonlinear pattern recognition. This distinction is further corroborated in subsequent comparative analyses, which indicate that under conditions of low modal counts, the C-POD method significantly outperforms traditional POD in terms of reconstruction error management and stability, making it particularly suited for the reconstruction of non-uniform temperature fields in scenarios with sparse sensor configurations.



**Figure 5.** POD and C-POD reduced-order modes of each order.

The trend of Min(MARE) values for different reconstruction methods, as the number of modes increases, is illustrated in Figure 6. It is evident from this figure that the Min(MARE) values decrease progressively with an increasing number of modes. The overall reconstruction accuracy can be ranked from lowest to highest as follows: POD-RBF, Gappy C-POD, POD-MLP, and Gappy POD. Notably, both the Gappy POD and POD-MLP methods demonstrate superior overall reconstruction accuracy.

Table 3 summarizes the pathways through which the various methods achieve their Min(MARE) values when the number of modes is between 15 and 20. According to the table, the Gappy POD\_Maximum\_50\_Uniform\_9 approach attains the optimal overall reconstruction accuracy. However, it is important to note that the datasets for the Gappy POD, POD-RBF, and POD-MLP methods are considerably larger, each reaching a scale of 50 times the number of independent variables. In contrast, the Gappy C-POD method achieves high precision in field reconstruction with a dataset only five times the size of the independent variables and relies on 25 sparse sensors, making it more suitable for engineering applications. Furthermore, while the machine-learning-based POD-MLP and POD-RBF methods show potential in terms of accuracy, their performance is heavily reliant on the quality and size of the dataset. The stability and generalization capabilities of these models still require further validation in scenarios with limited samples.



**Figure 6.** Variation of the minimum mean absolute relative error (MARE) values corresponding to different numbers of modalities.

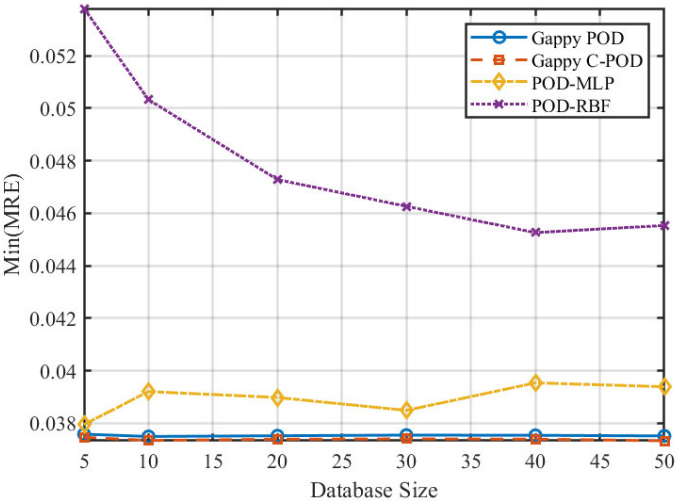
**Table 3.** Approaches for achieving minimum MARE values with various reconstruction methods under high-order modal conditions.

| Implementation strategies      | Min (MARE ) |
|--------------------------------|-------------|
| Gappy POD_Maximum_50_Uniform_9 | 0.00048     |
| Gappy C-POD_Maximum_5_CCFM_25  | 0.0021      |
| POD-RBF_Sobol_50_CCFM_25       | 0.0049      |
| POD-MLP_Maximum_50_SOPT_49     | 0.00053     |

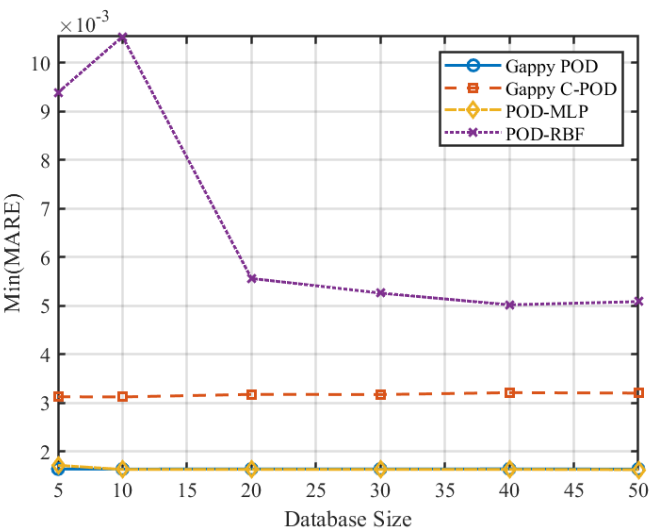
3.7. Analysis of the Impact of Database Size on the Reconstruction Capability of the Temperature Field Inverse Problem

This section investigates the effect of database size on the performance of inverse problem reconstruction by treating "database size" as the sole independent variable. Under varying sample sizes, a comprehensive consideration is given to all modal truncation numbers, diverse sparse sensor layout strategies (including Gappy POD, Gappy C-POD, POD-MLP, and POD-RBF), as well as multiple configurations of selected point quantities. The optimal combination for each database size will be identified to represent its superior reconstruction capability.

Figures 7 and 8 illustrate the variations of the minimum Mean Relative Error (Min MRE) and the minimum Mean Absolute Relative Error (Min MARE) across different database sizes. The data reveal a significant trend of performance enhancement among various reconstruction methods with optimal configurations as the database size increases. Notably, the POD-RBF method demonstrates a remarkable improvement, with its optimal MARE decreasing from 0.00939 at 50 samples to 0.00508 at 500 samples, underscoring the strong dependence of the RBF kernel function model on sample density. Conversely, both Gappy POD and Gappy C-POD reach a performance plateau within a relatively moderate database size range of approximately 200 to 300 samples, indicating efficient data utilisation. The POD-MLP method, constrained by its requirements for network parameter convergence, exhibits instability under smaller sample sizes but tends to stabilise when the sample size increases to 500.



**Figure 7.** Variation of Min(MRE) for four reconstruction methods under different database scale conditions.



**Figure 8.** Variation of Min(MARE) for four reconstruction methods under different database scale conditions.

4. Discussion on Engineering Practice

4.1. Impact of Sparse Sensor Quantity

In engineering applications, the scale of the database and the number of reduced-order modes can significantly influence the overall computational efficiency and deployment capabilities of algorithms. This study examines the effects of database construction methods, the quantity of sparse sensors, and the optimisation of sparse sensor layouts on the accuracy of field reconstruction, considering a database scale of five times the number of independent variables and a mode number ranging from 1 to 5. Using four different field reconstruction methods, we reconstructed the test set to obtain Min(MRE) and Min(MARE) values, which vary with the quantity of sparse sensors, as presented in Table 4. For clearer visualisation, the data from Table 4 is illustrated in Figures 9 and 10.

Figure 9 indicates that the relationship between Min(MARE) values and sensor quantity is relatively weak. Under conditions of low modes and smaller datasets, the overall reconstruction accuracy of the four methods is minimally affected by the number of sensors. In this scenario, the methods are ranked from highest to lowest reconstruction accuracy as follows: Gappy POD, POD-MLP, Gappy C-POD, and POD-RBF. Figure 10 shows that the Min(MRE) values for the Gappy POD, Gappy C-POD, and POD-MLP methods demonstrate a decreasing trend in fluctuation as the number of sensors increases. This suggests that augmenting the number of sensors enhances the stability of these three reconstruction methods. Conversely, the Min(MRE) value for the POD-RBF method increases with the addition of sensors. When the number of sensors reaches or exceeds 25, Gappy C-POD consistently exhibits the lowest Min(MRE) value, indicating its strong capacity to constrain maximum reconstruction error and thus its enhanced stability under conditions of low modal counts and small sample sizes.

**Table 4.** Changes in Min(MRE) and Min(MARE) values with varying numbers of sparse sensors.

| Number of<br>Sparse Sensors | Gappy POD |        | Gappy C-POD |        | POD-MLP |        | POD-RBF |        |
|-----------------------------|-----------|--------|-------------|--------|---------|--------|---------|--------|
|                             | Min(M     | Min(M  | Min(M       | Min(M  | Min(M   | Min(M  | Min(M   | Min(M  |
|                             | RE)       | ARE)   | RE)         | ARE)   | RE)     | ARE)   | RE)     | ARE)   |
| 4                           | 0.0408    | 0.0016 | 0.0482      | 0.0032 | 0.0412  | 0.0019 | 0.0545  | 0.0093 |
| 9                           | 0.0390    | 0.0016 | 0.0398      | 0.0031 | 0.0466  | 0.0022 | 0.0546  | 0.0117 |
| 16                          | 0.0406    | 0.0016 | 0.0473      | 0.0032 | 0.0383  | 0.0017 | 0.0545  | 0.0102 |

|    |        |               |               |        |        |        |        |        |
|----|--------|---------------|---------------|--------|--------|--------|--------|--------|
| 25 | 0.0395 | 0.0016        | 0.0394        | 0.0031 | 0.0397 | 0.0021 | 0.0546 | 0.0119 |
| 36 | 0.0404 | 0.0016        | 0.0384        | 0.0032 | 0.040  | 0.0018 | 0.0545 | 0.0125 |
| 49 | 0.0375 | <b>0.0016</b> | <b>0.0374</b> | 0.0031 | 0.039  | 0.0020 | 0.0606 | 0.0124 |

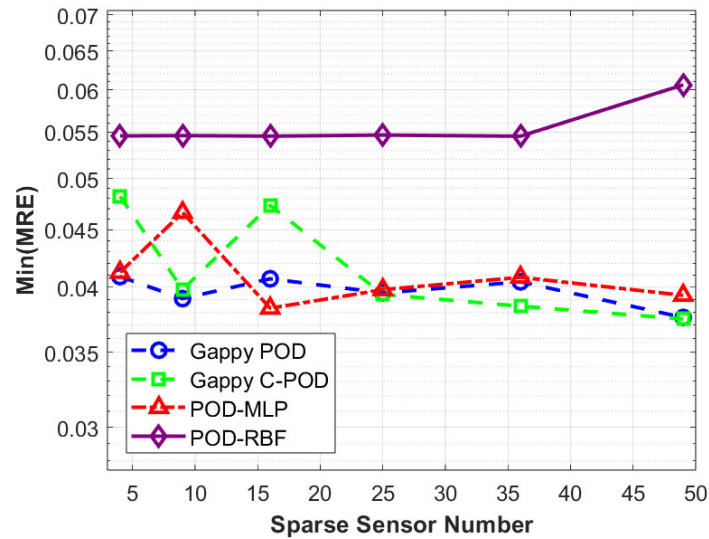


Figure 9. Variation of Min(MRE) values with the number of sensors under four reconstruction methods.

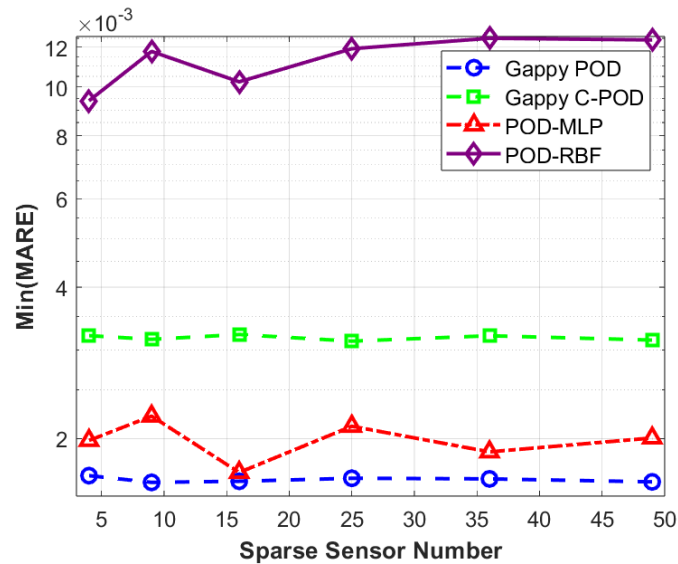


Figure 10. Variation of Min(MARE) values with the number of sensors under four reconstruction methods.

Tables 5 and 6 summarise the database sampling methods and sensor layout optimisation strategies corresponding to the minimum values of Max Relative Error (Min(MRE)) and Mean Absolute Relative Error (Min(MARE)) from Table 3. This organisation facilitates a clearer comparison of the optimal implementation pathways for various reconstruction methods across differing numbers of sensors. From Table 5, it can be observed that under the condition of minimising the maximum relative error (Min(MRE)), the Gappy C-POD method achieves optimal performance by employing Maximum distance sampling combined with the CCFM (Correlation Coefficient Filtering Method) selection strategy when using 49 sensors. Additionally, both the Gappy POD and POD-MLP methods exhibit a preference for Latin Hypercube sampling and the S-OPT layout across varying

sensor counts. In contrast, the POD-RBF method shows a greater reliance on the CCFM selection strategy under either Latin or Maximum sampling conditions.

Table 6 reveals that, under the minimisation of the mean relative error (Min(MARE)), the Gappy POD method demonstrates a stable adaptability to Sobol sampling and Uniform layouts across multiple sensor configurations, particularly exhibiting the highest overall reconstruction accuracy with 49 sensors. Conversely, the Gappy C-POD method maintains relatively low errors at lower sensor counts (such as 4 and 9) while primarily depending on Sobol sampling and either S-OPT or Uniform selection approaches. It is noteworthy that the POD-MLP method generally results in lower overall errors but shows a heightened sensitivity to layout methods, particularly relying on Maximum sampling and S-OPT layouts. The performance of the POD-RBF method on the Min(MARE) metric is comparatively weaker, especially as the number of sensors increases, where the decrease in error is limited.

In conclusion, the comparison of optimal strategies under conditions of minimum maximum error and minimum average error indicates marked differences in the pathways to achieving optimal performance across various reconstruction methods. The synergy between selection strategies and database construction significantly influences model performance. Notably, the Gappy C-POD method demonstrates greater stability in controlling reconstruction deviation, while the Gappy POD and POD-MLP methods exhibit superior performance in terms of overall accuracy.

**Table 5.** Sampling methods and layout methods used by each approach when achieving Min(MRE) under different sensor quantity conditions.

| Number of<br>sparse<br>sensors | Gappy POD   |             | Gappy C-POD |            | POD-MLP     |             | POD-RBF     |            |
|--------------------------------|-------------|-------------|-------------|------------|-------------|-------------|-------------|------------|
|                                | Samplin     | Layout      | Samplin     | Layout     | Samplin     | Layout      | Samplin     | Layout     |
|                                | g<br>method | metho<br>d  | g<br>method | metho<br>d | g<br>method | metho<br>d  | g<br>method | metho<br>d |
| 4                              | Sobol       | Unifor<br>m | Latin       | CCFM       | Latin       | CCFM        | Latin       | SOPT       |
| 9                              | Latin       | SOPT        | Sobol       | SOPT       | Sobol       | Unifor<br>m | Latin       | CCFM       |
| 16                             | Sobol       | SOPT        | Latin       | CCFM       | Latin       | Unifor<br>m | Maximu<br>m | CCFM       |
| 25                             | Latin       | SOPT        | Latin       | SOPT       | Maximu<br>m | Unifor<br>m | Latin       | CCFM       |
| 36                             | Sobol       | Unifor<br>m | Latin       | CCFM       | Maximu<br>m | Unifor<br>m | Latin       | SOPT       |
| 49                             | Latin       | CCFM        | Latin       | CCFM       | Sobol       | SOPT        | Latin       | CCFM       |

**Table 6.** Sampling methods and layout methods used by each approach when achieving Min(MARE) under different sensor quantity conditions.

| Number of<br>sparse<br>sensors | Gappy POD   |             | Gappy C-POD |             | POD-MLP     |             | POD-RBF     |            |
|--------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|------------|
|                                | Samplin     | Layout      | Samplin     | Layout      | Samplin     | Layout      | Samplin     | Layout     |
|                                | g<br>method | metho<br>d  | g<br>method | metho<br>d  | g<br>method | metho<br>d  | g<br>method | metho<br>d |
| 4                              | Maximu<br>m | SOPT        | Sobol       | SOPT        | Maximu<br>m | SOPT        | Maximu<br>m | CCFM       |
| 9                              | Sobol       | Unifor<br>m | Sobol       | Unifor<br>m | Sobol       | Unifor<br>m | Latin       | CCFM       |

|    |             |             |       |             |             |      |       |      |
|----|-------------|-------------|-------|-------------|-------------|------|-------|------|
| 16 | Sobol       | Unifor<br>m | Sobol | Unifor<br>m | Maximu<br>m | SOPT | Sobol | CCFM |
| 25 | Maximu<br>m | Rando<br>m  | Sobol | Unifor<br>m | Maximu<br>m | SOPT | Latin | CCFM |
| 36 | Sobol       | Unifor<br>m | Sobol | Unifor<br>m | Sobol       | SOPT | Sobol | CCFM |
| 49 | Sobol       | Unifor<br>m | Sobol | Unifor<br>m | Sobol       | SOPT | Latin | CCFM |

5. Reconstruction of Transient Thermal Conductivity Temperature Field

5.1. Overview of Transient Thermal Conductivity Issue

In examining the transient heat conduction process, we maintain consistency with the boundary conditions and heat source distribution outlined in Section 2.1. The diffusion of heat over time is simulated using an explicit finite difference method. This study investigates the impact of various reconstruction approaches on the reconstruction capacity of the transient temperature field under the influence of non-uniform internal heat sources.

(1) The governing equation for transient heat conduction is represented in Equation (27).

$$\frac{\partial T(x,y,t)}{\partial t} = \alpha(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}) + \frac{\Theta(x,y)}{\rho c_p}$$

(27)

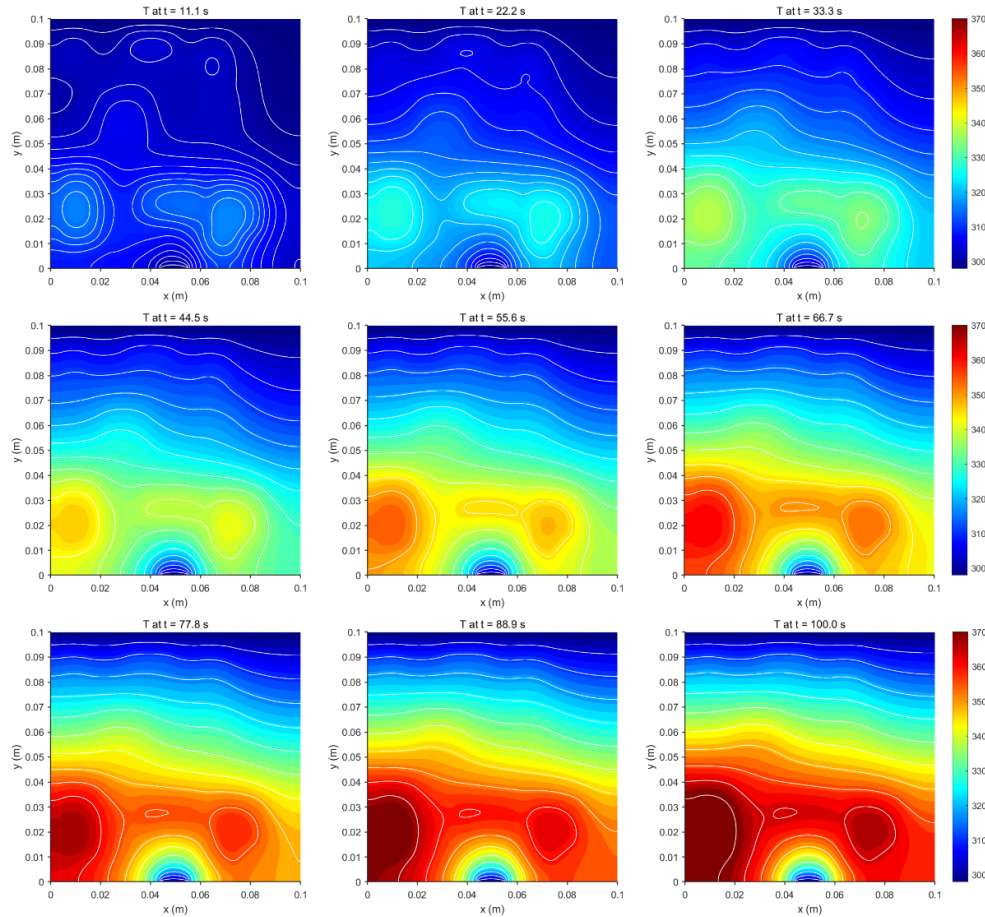
where:  $\alpha$  denotes the thermal diffusivity, while  $\Theta(x,y)$  signifies the heat source density per unit volume.

(2) Materials and numerical parameters.

Table 7. Materials and Numerical Parameters.

| Parameter              | Symbol     | Value              |
|------------------------|------------|--------------------|
| Thermal conductivity   | $k$        | 1                  |
| Density                | $\rho$     | 500                |
| Specific heat capacity | $Cp$       | 100                |
| Thermal diffusivity    | $\alpha$   | $2 \times 10^{-5}$ |
| Grid size              | $dx, dy$   | 1 mm               |
| Time step              | $\Delta t$ | 0.01               |
| Termination time       | tend       | 100                |

Figure 11 illustrates the evolution of the transient thermal field at various time steps. As time progresses, heat gradually disperses from multiple internal heat sources to the surrounding areas. During the initial phase, localized temperature peaks rapidly emerge around the heat sources. Subsequently, influenced by thermal conduction and boundary conditions, the thermal energy continues to diffuse outward. The temperature field exhibits characteristic features of unsteady heat transfer throughout its evolution, with the thermal structure adjusting continuously over time until it achieves stability.



**Figure 11.** Schematic diagram of the evolution of the transient thermal conductivity temperature field distribution.

### 5.2. Influence of Modal Coefficients on the Reconstruction Capability of Transient Thermal Inverse Problems

In the process of constructing the transient thermal dataset and validation set, the combination of independent variables was maintained consistent with conditions observed under steady-state heat conduction. For each set of these independent variable combinations, temperature field snapshots were collected every 10 seconds and incorporated into the dataset. Consequently, in contrast to steady-state thermal scenarios, each sample in the transient heat conduction problem encapsulates multiple time points of temperature field evolution, thereby resulting in datasets and validation sets with matrix dimensions approximately ten times larger than those for steady-state conditions. Apart from the temporal dimension expansion, all other simulation conditions and parameter settings were kept consistent with those in the steady-state heat conduction process to ensure comparability between the two problem types.

To investigate the impact of varying modes on the temperature field reconstruction performance within transient thermal inverse problems, four reconstruction methods (Gappy POD, Gappy C-POD, POD-MLP, and POD-RBF) were employed under conditions of sparse sensor data. The performance of these methods was evaluated by calculating the minimum relative error (Min(MRE)) and the minimum average relative error (Min(MARE)) at optimal configurations, with results plotted against changes in the number of truncated modes (as depicted in Figures 12 and 13). Analysis of these figures reveals that the reconstruction accuracy and stability of the POD-MLP and POD-RBF methods demonstrate a non-monotonic trend, initially increasing and then decreasing. This suggests that both methods struggle to maintain robust predictive performance at low ( $<6$ ) or excessively high ( $>14$ ) mode counts. Conversely, the Gappy POD and Gappy C-POD methods display a rapid

enhancement in reconstruction performance with an increasing number of modes, eventually stabilizing, which indicates their lower sensitivity to mode truncation and greater robustness. Notably, within the range of 6 to 11 modes, all four methods exhibit commendable reconstruction accuracy, with the Gappy C-POD method showcasing exceptional stability in this interval, maintaining a stable error level even as the number of modes rises. In contrast, the POD-RBF method achieves its best reconstruction performance at 6 modes but experiences significant fluctuations in performance, particularly when the mode count is low ( $\leq 4$ ) or high ( $>14$ ), resulting in a marked increase in error and a substantial drop in stability.

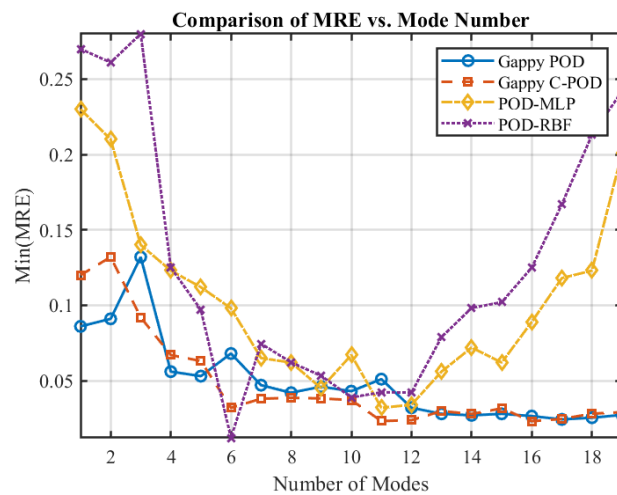
Based on the performance data illustrated in Figures 12 and 13, this study classifies the range of modal counts into three levels: low modes ( $\leq 6$ ), moderate modes ( $>6$  and  $\leq 10$ ), and high modes ( $>10$ ). The stability and overall reconstruction accuracy of the four methods across these intervals have been qualitatively ranked as follows:

(1) Stability Ranking:

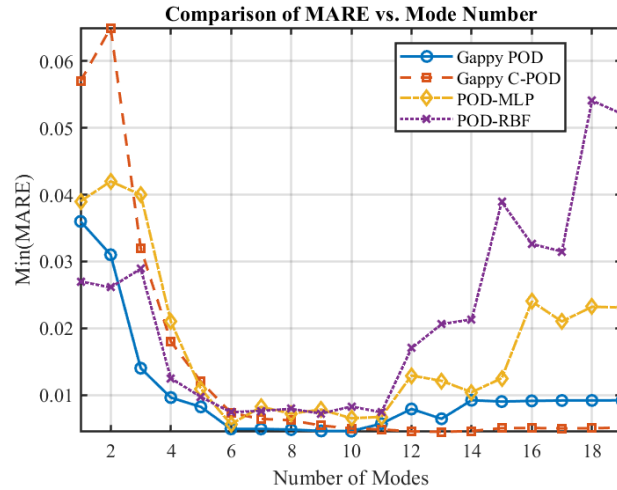
- Low Modes: Gappy POD > Gappy C-POD > POD-MLP > POD-RBF;
- Moderate Modes: Gappy C-POD > Gappy POD > POD-MLP > POD-RBF;
- High Modes: Gappy C-POD > Gappy POD > POD-MLP > POD-RBF.

(2) Overall Accuracy Ranking:

- Low Modes: Gappy POD > POD-RBF > POD-MLP > Gappy C-POD;
- Moderate Modes: Gappy POD > Gappy C-POD > POD-MLP > POD-RBF;
- High Modes: Gappy C-POD > Gappy POD > POD-MLP > POD-RBF.



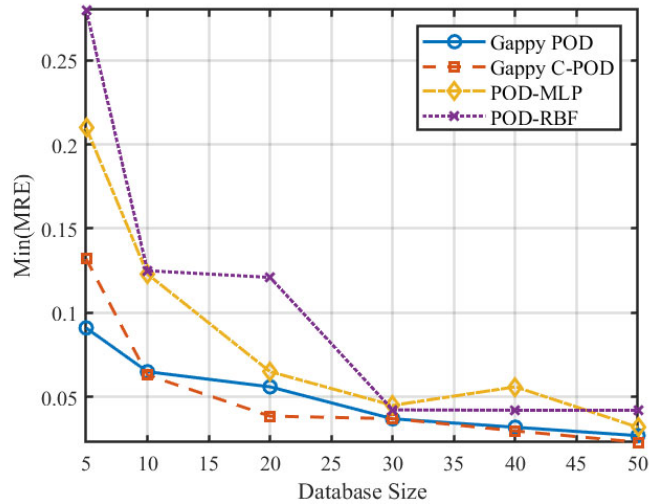
**Figure 12.** The Min(MRE) values achievable by various reconstruction methods under different conditions of sparse sensor quantities.



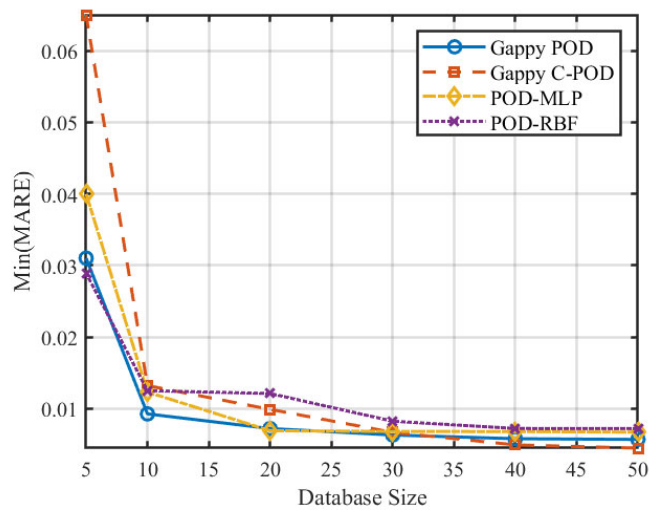
**Figure 13.** The Min(MARE) values achievable by various reconstruction methods under different conditions of sparse sensor quantities.

### 5.3. The Impact of Dataset Scale on Reconstruction Capability

To investigate the influence of varying dataset scales on the reconstruction accuracy of transient thermal temperature fields, we examined dataset sizes of 5, 10, 20, 30, 40, and 50 times the number of independent variables. The changes in Min(MRE) and Min(MARE) values for four different methods across these dataset scales are illustrated in Figures 14 and 15. The results presented in these figures reveal a clear trend: as the dataset size increases, both the precision and stability of the reconstruction methods improve progressively for transient thermal problems.



**Figure 14.** Variation of Min(MRE) values for different reconstruction methods under different dataset scale conditions.



**Figure 15.** Variation of Min(MARE) values for different reconstruction methods under different dataset scale conditions.

5.4. Discussion on Optimal Reconstruction Strategies for Transient Thermal Problems

In conjunction with the discussion in Section 3.2 regarding the number of modes, a 9th order mode was selected for further analysis. Additionally, as shown in Figures 14 and 15, the variations in accuracy and stability become relatively stable when the dataset size exceeds 30 times the number of independent variables. Taking into account computational and storage efficiency, a dataset scale of 30 times the number of independent variables was chosen. Under these conditions, we analysed the reconstruction strategies that yield optimal results for each method, as summarised in Tables 8 and 9.

Table 8 indicates that, with respect to the Min(MRE) metric, the Gappy C-POD method, when combined with the CCFM (Cross-Correlation Filtering Method) for sensor selection and a configuration of 16 sensors, demonstrates the strongest capability for minimising reconstruction error, achieving a Min(MRE) value of 0.0371, which significantly outperforms the other three methods. The Gappy POD method, utilising S-OPT for sensor selection with the same number of sensors, also shows stable performance with a Min(MRE) of 0.0462, indicating its effective error control capabilities. In contrast, the POD-MLP and POD-RBF methods present minimum maximum errors of 0.0487 and 0.0561, respectively, reflecting somewhat lower overall stability due to their susceptibility to sensor selection sensitivity and the instability of nonlinear mapping.

As observed in Table 9, regarding the Min(MARE) metric, both the Gappy C-POD and Gappy POD methods exhibit strong overall precision. The former, using CCFM with 9 sensors, achieves a minimum value of 0.00456, while the latter, utilising S-OPT with 9 sensors, attains a value of 0.00462, demonstrating nearly equivalent performance. This suggests that under the 9th order mode with moderate sample density, both POD-based truncation methods offer high data utilisation efficiency and precision advantages. Although the POD-MLP and POD-RBF methods yield relatively lower precision, they can still deliver results with fewer sensor configurations.

**Table 8.** Achieving Min(MRE) under different reconstruction methods, point selection methods, and point selection quantities.

| Reconstruction methods | Selection method | Number | Min(MRE) |
|------------------------|------------------|--------|----------|
| Gappy POD              | S-OPT            | 16     | 0.0462   |

|             |         |    |        |
|-------------|---------|----|--------|
| Gappy C-POD | CCFM    | 16 | 0.0371 |
| POD-MLP     | Uniform | 9  | 0.0487 |
| POD-RBF     | CCFM    | 9  | 0.0561 |

**Table 9.** Achieving Min(MARE) under different reconstruction methods, point selection methods, and point quantity conditions.

| Reconstruction methods | Selection method | Number | Min(MRE) |
|------------------------|------------------|--------|----------|
| Gappy POD              | S-OPT            | 9      | 0.00462  |
| Gappy C-POD            | CCFM             | 9      | 0.00456  |
| POD-MLP                | S-OPT            | 4      | 0.00726  |
| POD-RBF                | CCFM             | 4      | 0.00823  |

6. Conclusions

This study addresses the precise reconstruction requirements of temperature field inverse problems in two-dimensional steady-state and transient heat conduction scenarios. A systematic investigation was conducted into various data-driven field reconstruction methods, evaluating their applicability and performance across different database sizes, modal truncation numbers, and sparse sensor layout strategies. The primary conclusions are as follows:

For steady-state heat conduction problems:

- (1) In low modal conditions (1-5 modes), all reconstruction methods demonstrated a strong capability for constrained reconstruction deviation. Notably, the Gappy C-POD method, when combined with the CCFM point selection strategy and maximum distance sampling approach, attained the minimum maximum relative error ( $\text{Min(MRE)} = 0.0373$ ) with a configuration of 49 sensors, displaying optimal reconstruction stability.
- (2) Under high modal conditions (15-20 modes), the Gappy POD method achieved the lowest mean relative error ( $\text{Min(MARE)} = 0.00048$ ) with a uniform layout and a sample size of 50 times, illustrating excellent overall reconstruction accuracy.
- (3) Considering reconstruction error, data utilisation efficiency, and deployment complexity, the Gappy C-POD method achieves a performance plateau with medium database sizes (200-300 cases), making it particularly suitable for engineering applications. Although the POD-MLP method shows significant potential for overall accuracy, it is sensitive to the quality and size of the database, thus more suited to scenarios with abundant data and nonlinear modelling requirements.

For transient heat conduction problems:

- (1) The accuracy and stability of various reconstruction methods exhibit significant variability with changes in modal truncation numbers. The Gappy POD and Gappy C-POD methods achieve rapid and stable convergence within the modal range of 6-11, showcasing greater robustness. In contrast, the POD-RBF and POD-MLP methods experience marked performance fluctuations, making them less adaptable to low or high modal conditions.
- (2) The scale of the dataset has a pronounced impact on reconstruction performance. Beyond 30 times the number of independent variables, improvements in reconstruction accuracy tend to plateau. Under this condition, the Gappy C-POD method, in conjunction with CCFM point selection and 16 sensors, achieves the minimum maximum error ( $\text{Min(MRE)} = 0.0371$ ), while also attaining the minimum mean error ( $\text{Min(MARE)} = 0.00456$ ) with just 9 sensors.
- (3) Compared to steady-state problems, transient issues exhibit heightened sensitivity of methods to data density and modal truncation. When selecting modal numbers and sample sizes, a careful balance between computational cost and model accuracy is recommended.

(4) The POD-RBF method can achieve moderate accuracy under conditions of ample samples and reasonable sensor layout, although its overall stability remains inferior to Gappy methods. Conversely, the POD-MLP method shows some adaptability with smaller sensor configurations, making it suited for rapid prediction needs in resource-constrained scenarios.

In summary, the Gappy C-POD method stands out for its excellent error constraint capabilities and robustness in both steady-state and transient heat conduction problems, making it an optimal choice for solving temperature field inverse problems in engineering applications. The Gappy POD method excels in high-accuracy reconstruction tasks, particularly where stringent overall accuracy requirements are essential. Machine learning-based methods can serve as supplementary strategies, particularly applicable in data-rich environments that demand rapid iterative predictions.

## References

1. Grysa K, Ciałkowski Michał, Kamiński H. An inverse temperature field problem of the theory of thermal Stresses [J]. Nuclear Engineering and Design, 1981, 64(2): 169–184.
2. Xu S L. Inverse Heat Conduction Problem of the Temperature Field with Artificial Ground Freezing Method [J]. Advanced Materials Research, 2011, 243–249: 89–92.
3. Wang S, Ni R. Solving of Two-Dimensional Unsteady-State Heat-Transfer Inverse Problem Using Finite Difference Method and Model Prediction Control Method [J]. Complexity, 2019, 2019(1).
4. Bergagio M, Li H, Anglart H. An iterative Finite-element algorithm for solving two-dimensional nonlinear inverse heat conduction Problems [J]. International Journal of Heat and Mass Transfer, 2018, 126: 281–292.
5. Al-Mahdawi H K. Solving of an Inverse Boundary Value Problem for the Heat Conduction Equation by Using Lavrentiev Regularization Method [J]. Journal of Physics: Conference Series, 2021, 1715(1): 012032.
6. Duda P. Solution of inverse heat conduction problem using the Tikhonov regularization Method [J]. Journal of Thermal Science, 2017, 26(1): 60–65.
7. Mesenev P R. Optimization method for solving the inverse problem of complex heat transfer [J]. Dal'nevostochnyi Matematicheskii Zhurnal, 2023, 23(1): 81–84.
8. Rasmussen A K, Seizilles F, Girolami M, et al. The Bayesian approach to inverse Robin problems [Z]. arXiv, 2023(2023).
9. Habermann M, Maxwell D, Truffer M. Reconstruction of basal properties in ice sheets using iterative inverse Methods [J]. Journal of Glaciology, 2012, 58(210): 795–808.
10. Yaparova N. Numerical methods for solving a Boundary-value inverse heat conduction Problem [J]. Inverse Problems in Science and Engineering, 2013, 22(5): 832–847.
11. Janicki M, Napieralski A. Inverse heat conduction problems in electronics with special consideration of analytical analysis methods [C]. 2004 International Semiconductor Conference. CAS 2004 Proceedings (IEEE Cat. No.04TH8748), [no date], 2: 455–458.
12. Mohebbi F, Sellier M, Rabczuk T. Estimation of linearly Temperature-dependent thermal conductivity using an inverse Analysis [J]. International Journal of Thermal Sciences, 2017, 117: 68–76.
13. Tuan N H, Trong D D. Two regularization methods for backward heat problems with new error Estimates [J]. Nonlinear Analysis: Real World Applications, 2011, 12(3): 1720–1732.
14. Sokolov A, Nikulina I. Choice of Regularization Methods in Experiment Processing: Solving Inverse Problems of Thermal Conductivity [J]. Mathematics, 2022, 10(22): 4221.
15. Alonso G, Del Valle E, Ramirez J R. Optimization Methods [M]//Desalination in Nuclear Power Plants. Elsevier, 2020: 67–76.
16. Mlynarczyk P. Numerical optimization methods comparison based on the CFD Conduction-convection heat transfer Case [J]. International Journal of Numerical Methods for Heat & Fluid Flow, 2019, 29(6): 2080–2092.
17. Sun S, Cao Z, Zhu H, et al. A Survey of Optimization Methods from a Machine Learning Perspective [Z]. arXiv, 2019(2019).
18. Wang X, Jin Y, Schmitt S, et al. Recent Advances in Bayesian Optimization [J]. ACM Computing Surveys, 2023, 55(13s): 1–36.

19. Ciampi F G, Rega A, Diallo T M L, et al. Energy consumption prediction of industrial HVAC systems using Bayesian Networks [J]. *Energy and Buildings*, 2024, 309: 114039.
20. Liu R, Cheng S, He Y, et al. On the Convergence of Learning-Based Iterative Methods for Nonconvex Inverse Problems [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(12): 3027–3039.
21. Huang Q, Gong R, Zhang Y. A new Second-order dynamical method for solving linear inverse problems in Hilbert Spaces [J]. *Applied Mathematics and Computation*, 2024, 473: 128642.
22. Talaat M, Tayseer M, Farahat M A, et al. Artificial intelligence strategies for simulating the integrated energy Systems [J]. *Artificial Intelligence Review*, 2024, 57(4).
23. LIANG Y, ZHENG B, GAO X, et al. Reduced order model analysis method via proper orthogonal decomposition for nonlinear transient heat Conduction problems [J]. *SCIENTIA SINICA Physica, Mechanica & Astronomica*, 2018, 48(12): 124603.
24. Nekkanti A, Schmidt O T. Gappy spectral proper orthogonal Decomposition [J]. *Journal of Computational Physics*, 2023, 478: 111950.
25. Wang X, Li D, Zhao J, et al. Indoor environment reconstruction algorithm based on gappy POD and finite Sensors [J]. *Energy and Buildings*, 2023, 297: 113463.
26. Tsering-xiao B, Xu Q. Gappy POD-based reconstruction of the temperature field in Tibet [J]. *Theoretical and Applied Climatology*, 2019, 138(1–2): 1179–1188.
27. Wang Z, Yuan H, Lv J, et al. Decoupling and Reconstruction of Low-Quality Temperature Fields of Printed Circuit Boards [J]. *IEEE Sensors Journal*, 2024, 24(7): 10892–10903.
28. Qi X, Hong C, Gu L, et al. Surface Temperature Field Reconstruction of Lithium-Ion Batteries Toward Lumped Thermal Model-Based KF-MLP Estimation Algorithm [J]. *IEEE Transactions on Industrial Informatics*, 2025, 21(6): 4841–4851.
29. WILLCOX K. Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition [J]. *Computers & fluids*, 2006, 35(2): 208–226.
30. Yuan Q, Guo X, Han J, et al. Correlation Coefficient Filtering Method for Optimal Sensor Layout Strategies in Inverse Problems [J]. *Journal of Engineering*, 2025, 2025(1).
31. Sirovich L. Turbulence and the dynamics of coherent structures. I. Coherent structures [J]. *Quarterly of Applied Mathematics*, 1987, 45(3): 561–571.
32. Stein M. Large Sample Properties of Simulations Using Latin Hypercube Sampling [J]. *Technometrics*, 1987, 29(2): 143–151.
33. Pánek D, Kotlan V, Hamar R, et al. Novel algorithm for modeling combined laser and induction welding respecting keyhole Effect [J]. *Applied Mathematics and Computation*, 2018, 319: 254–263.
34. Zhou Y, Qi H, Zhai Y, et al. Learning to Reconstruct 3D Manhattan Wireframes From a Single Image [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 7697–7706.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.