

Article

A Robust Sparse Sensor Placement Strategy Based on Indicators of Noise for Ocean Monitoring

Qiannan Zhang ¹, Huafeng Wu ^{1,*}, Li'nian Liang ¹, Xiaojun Mei ¹, Jiangfeng Xian ²
and Yuanyuan Zhang ³

¹ Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China; zhangqiannan02@stu.shmtu.edu.cn (Q.Z.); lianglinian@stu.shmtu.edu.cn (L.L.); xjmei@shmtu.edu.cn (X.M.)
² Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China; jfxian@shmtu.edu.cn
³ School of Computer and Information Engineering, Changzhou Institute of Technology, Changzhou 213032, China; zhangyy@czu.cn
* Correspondence: hfwu@shmtu.edu.cn

Abstract: A well-performing data-driven sparse sensor deployment strategy is critical for marine monitoring systems, as it enables the optimal reconstruction of marine physical quantities with fewer sensors. However, ocean data typically contain substantial amounts of noise, including outliers (incomplete data) and inherent measurement noise, which heightens the complexity of sensor deployment. Therefore, this study optimizes the sparse sensor placement model by establishing noise indicators, including small noise weight and large noise weight, which are measured by entropy to minimize the prediction bias. Building on this, a robust sparse sensor placement algorithm is proposed, which utilizes the block coordinate update (BCU) iteration method to solve the function. During the iterative updating process, the proposed algorithm simultaneously updates the selection matrix, reconstruction matrix, and noise matrix. This allows for effective identification and mitigation of noise in the data through evaluation. Consequently, the deployed sensors achieve superior reconstruction performance compared to other deployment methods that do not incorporate noise evaluation. Experiments are also conducted on datasets of sea surface temperature (SST) and global ocean salinity, which demonstrate that our strategy significantly outperforms several other considered methods in terms of reconstruction accuracy while enabling autonomous sensor deployment under noisy conditions.

Keywords: sparse sensor placement; robust; noise indicator; ocean monitoring; data reconstruction



Citation: Zhang, Q.; Wu, H.; Liang, L.; Mei, X.; Xian, J.; Zhang, Y. A Robust Sparse Sensor Placement Strategy Based on Indicators of Noise for Ocean Monitoring. *J. Mar. Sci. Eng.* **2024**, *12*, 1220. <https://doi.org/10.3390/jmse12071220>

Academic Editor: Christos Tsabaris

Received: 10 June 2024
Revised: 10 July 2024
Accepted: 16 July 2024
Published: 19 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The establishment of an effective ocean observation system is paramount for comprehensive ocean monitoring. The deployment of sensors plays a critical role in determining the efficiency and accuracy of such a system. Through the strategic placement of these sensors, the ocean observation system is capable of collecting a wide array of oceanic data. The precision of these data is of utmost importance for the study of global oceanic phenomena, including the El Niño and La Niña events [1–3]. However, the placement of oceanic sensors is influenced by a variety of factors, including deployment and maintenance costs, areas of interest, and specific monitoring objectives. Consequently, the investigation of optimal sensor placement strategies, given a limited number of sensors, is of paramount importance and presents a significant challenge in the field of signal processing for global ocean observation.

The main concern regarding the deployment of marine observation sensors is to conduct direct marine measurements with a small number of optimally deployed sensors under constrained conditions, and then accurately describe and predict the process of changes in the marine environment based on the data collected from these measurements.

Traditional research on sensor placement in the ocean mainly includes model-based sensor placement [4–8], interest area-based sensor placement [9–12], and sensor placement based on empirical interpolation [13–15].

As data processing technologies evolve, data-driven sensor deployment is becoming increasingly valuable for real-world physical systems. Data-driven sensor deployment primarily relies on data analysis and machine learning on the sensory data from physical systems, which is generally researched as finding the low-dimensional patterns and features corresponding to sensory attractors for high-dimensional systems. Dimensionality reduction methods are effective ways for discovering low-dimensional map patterns constructed from limited measurements by deployed sensors, such as proper orthogonal decomposition (POD) [16,17], dynamic mode decomposition (DMD) [18,19], and neural networks (NN) [20,21].

In the context of ocean monitoring, various data-driven methods have been developed to optimize sensor deployment. Due to the vastness of the ocean, utilizing a limited number of sensors for monitoring can effectively reduce deployment and maintenance costs. A data-driven sparse sensor deployment strategy can efficiently leverage the intrinsic characteristics of marine data. By employing a constrained number of sensors to collect low-dimensional data, it is possible to achieve a comprehensive reconstruction of the full state of oceanic physical quantities (thereby obtaining high-dimensional data). For example, the POD method has been applied to wind field simulation data [22], while the QR algorithm has been utilized for the reconstruction of the flow field [23]. Additionally, the QR algorithm under cost constraints has been employed for ocean climate data and fluid data [24], and the observation point selection of autonomous underwater vehicles (AUV) has been explored [25]. Furthermore, information entropy-based QR placement has been used for ocean temperature monitoring [26]. These methods illustrate that by processing data, sensor locations can be optimized based on the inherent characteristics and information content of the data. This optimization can lead to improved reconstruction accuracy and meet other monitoring requirements, such as cost, energy consumption, and information volume.

However, data reconstruction using the singular value decomposition (SVD) basis typically involves selecting sensors corresponding to low-rank modes of the data. In such cases, noise in the sensor measurements can significantly impact the global data reconstruction results obtained from a limited number of sensors. Moreover, Peherstorfer et al. [27] demonstrated that utilizing the SVD basis for data reconstruction can lead to noise amplification issues, which become more pronounced as the number of sensors increases.

In ocean monitoring systems, it is common for the data collected by deployed sensors to be incomplete, with outliers frequently and unavoidably occurring [28], or to contain noise, including measurement and process noise [29]. The noise is primarily caused by the constantly changing marine environment and the susceptibility of communication signals to interference from the ionosphere [30], as well as events such as oil spills [31], typhoons, tsunamis, and other similar occurrences. Since data-driven sparse sensor deployment relies on processing and analyzing data to extract intrinsic features, noise within the data (including outliers) significantly impacts the formulation of sparse sensor deployment strategies. Specifically, noise impacts the data training process, subsequently influencing the outcomes of sensor selection in data-driven sensor placement, as well as the robustness of oceanic sensor deployment.

To mitigate the impact of noise, Clark et al. [32] optimized the QR greedy sensor placement strategy by categorizing noise under specific basis choices of SVD and randomized modes. Ghayem et al. [33] proposed a sensor deployment algorithm designed to maximize the probability of achieving a high signal-to-noise ratio, with sensor locations being sequentially selected in a greedy manner. These noise-aware methodologies demonstrate that the robustness of sensor deployment can be enhanced through noise measurement. However, the greedy algorithm may compromise the overall measurement effectiveness of the selected sensor set to a certain extent.

Furthermore, despite these multiple constraints, the aforementioned methods still fall short in simultaneously optimizing both the selected sensor candidate set and the basis used for data reconstruction. This issue can be addressed through feature selection and subspace learning via advanced optimization methods. Zhou et al. [34,35] addressed the feature selection problem as a sparse subspace learning problem in high-dimensional space, employing local preserving regularization terms such as the similarity matrix of features. After relaxing the problem formulation, the block coordinate update (BCU) method was used to alternately iterate between the selection matrix and the corresponding reconstruction matrix. This insight inspires us to address the sparse sensor placement problem through a subspace learning approach, thereby overcoming the limitations associated with the greedy method when solving it from a global perspective.

In this study, we propose a robust sparse sensor placement strategy based on indicators of noise. The key contributions of this work are as follows:

- Establishing noise indicators for the sparse sensor placement model, which encompass both small and large noise weight matrices, and employing the entropy of noise to minimize the prediction bias, thereby deriving a novel optimization objective function for robust sensor deployment;
- The block coordinate update (BCU) iteration method is adopted in the proposed algorithm of robust sparse sensor placement based on indicator of noise (RSSPIN) to solve the optimization objective function, which contains non-convex components. Additionally, the RSSPIN algorithm simultaneously updates the selection matrix, reconstruction matrix, and noise matrix during each iteration;
- During the iterative updating process, continuous noise evaluation is performed, and the selection matrix and reconstruction matrix are updated based on the updated noise matrix. This ensures that the full-state reconstruction capability of the sensor measurement subset corresponding to the obtained selection matrix is minimally affected by noise;
- Experimental verification has been performed to demonstrate the robustness and effectiveness of the RSSPIN algorithm under noisy conditions. Comparative analyses with existing methods reveal that the proposed method achieves superior reconstruction accuracy while facilitating autonomous sensor deployment in the presence of noise.

The remaining part of this paper is arranged as follows. In Section 2, we provide a brief overview of the notations and definitions used in this paper, as well as the sparse sensor deployment optimization problem that arises in existing sparse sensor placement methods. In Section 3, we develop a robust sparse sensor deployment optimization model that incorporates noise regularization methods and propose an algorithm that iteratively solves the model using BCU. This section also includes a theoretical analysis of the convergence of the proposed algorithm. The experimental results and analysis are presented in Section 4. Finally, Section 5 concludes the paper and provides a discussion.

2. Preliminary

Sparse sensor placement based on data reconstruction generally obtains measurement results $Y \in \mathbb{R}^{p \times m}$ by selecting partial data corresponding to selected sensors from the existing full-state measurement $X \in \mathbb{R}^{n \times m}$ through a matrix C [23,28,36], as follows:

$$Y=CX \tag{1}$$

where $C \in \mathbb{R}^{p \times n}$, and it is utilized as a selection matrix to select the optimal partial measurements in X . There are p non-zero variables in C , $p \ll n$. That is, C has p non-zero basis vectors to select from n candidate locations. C has the following structure:

$$C = [e_{\zeta_1}^T \ e_{\zeta_2}^T \ \cdots \ e_{\zeta_p}^T] \tag{2}$$

where e_i represents the i -th location of the selected p locations. Then the measurement selected by C is as follows:

$$Y = CX = [x_{\xi_1}, x_{\xi_2}, \dots, x_{\xi_p}]^T \tag{3}$$

The selection matrix C can be used as a parameter matrix to obtain the complete reconstructed data $\hat{X} = O(Y; C)$ by finding the appropriate mapping relationship, wherein $O(\cdot)$ is the mapping function from Y to \hat{X} . The criterion for selecting p sensors is generally set to choose an appropriate selection matrix C , with the aim of ensuring that the reconstructed data \hat{X} closely mirror the original data X . To quantify this similarity, the L_2 -norm is a common choice (which can also be interpreted as the minimum mean square error, MSE), as follows:

$$C = \underset{C}{\operatorname{argmin}} \|X - \hat{X}\|_2 = \underset{C}{\operatorname{argmin}} \|X - O(Y; C)\|_2 \tag{4}$$

In compress sensing theory, an appropriate basis Ψ from training data X_{tr} can be employed to reconstruct the test data X_{te} . In practice, Ψ can be obtained using SVD decomposition as $X_{tr} = \Psi \Lambda Y \approx \Psi_r \Lambda_r Y_r$. In this context, Ψ_r is the first r columns of Ψ . According to data-driven sparse sensor selection [23], when the test data X_{te} are provided, the reconstructed data are represented as follows:

$$\hat{X}_{te} = \Psi_r \hat{S}, \text{ where } \hat{S} = \begin{cases} (C\Psi_r)^{-1}Y, & p = r \\ (C\Psi_r)^\dagger Y, & p > r \end{cases} \tag{5}$$

where \hat{S} represents the coefficient matrix for data reconstruction. Consequently, Equation (4) can be transformed as follows:

$$\begin{aligned} \hat{C} &= \underset{C}{\operatorname{argmin}} \|X_{te} - \hat{X}_{te}\|_2 = \underset{C}{\operatorname{argmin}} \|X_{te} - \Psi_r (C\Psi_r)^\dagger Y\|_2 \\ &= \underset{C}{\operatorname{argmin}} \|X_{te} - \Psi_r (C\Psi_r)^\dagger C X_{te}\|_2 \\ &= \underset{C}{\operatorname{argmin}} \|X_{te} - A C X_{te}\|_2 \end{aligned} \tag{6}$$

where $A = \Psi_r (C\Psi_r)^\dagger$ denotes the reconstruction matrix corresponding to the selection matrix C . As previously mentioned, the value of A influences the reconstruction accuracy, but the choice of the base matrix Ψ_r , which affects the value of A , will introduce noise. The noise in the original data X will also affect the establishment of the selection matrix C and reconstruction matrix A .

In the following section of this study, a robust sparse sensor placement strategy is proposed for accurate data reconstruction by iteratively updating matrices C and A concurrently, thereby minimizing the impact of noise.

3. Robust Sparse Sensor Placement Based on Indicator of Noise

When measuring general physical quantities, the associated noise typically includes both process noise and measurement noise. The process noise is usually generated by the physical quantity being measured, whereas the measurement noise is usually introduced by the sensors themselves. For the monitoring of physical quantities in the ocean, the measurement noise is often significant due to the instability of long-distance transmission channels and the variability of the measurement environment state (ocean surface measurements are highly dynamic and time-varying under the influence of winds, swells, and currents). In this section, we first establish noise indicators that contain a small noise weight matrix W and a large noise weight matrix \bar{W} (including the outlier entries). Following this, the reconstruction objective function is constrained by the noise weights to derive a new robust sensor deployment strategy objective function. Subsequently, the BCU method is employed for solving the objective function.

3.1. Problem Formulation

In order to enhance the robustness to noise, a small noise distribution weight matrix $W \in \mathbb{R}^{n \times m}$ and large noise distribution weight matrix $\bar{W} \in \mathbb{R}^{n \times m}$ are established as noise indicators for data reconstruction, similar to that in [37], and $W + \bar{W} = \mathbf{1}_{n \times m}$. Then, the objective function of the sparse sensor placement problem in Equation (6) can be redefined as follows:

$$\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^N w_{ij} \|x_{ij} - \hat{x}_{ij}\|_2^2 = \frac{1}{2} \|\sqrt{W} \odot (X - \hat{X})\|_F^2 \tag{7}$$

wherein $w_{ij} \in W$ is an element of the small noise weight matrix of the data X , and \odot is the Hadamard product operator.

If the number of sensors to be selected is p , the model of robust sparse sensor placement problem with small noise and large noise can be represented by combining Equations (6) and (7). Then, we have the following:

$$\begin{aligned} & \min_{A,C,W} \frac{1}{2} \|\sqrt{W} \odot (X - ACX)\|_F^2 \\ & \text{s.t. } W \in \{0, 1\}^{n \times m} \\ & C \in \{0, 1\}^{p \times n}, C^T \mathbf{1}_{n \times 1} = \mathbf{1}_{p \times 1}, \|C \mathbf{1}_{p \times 1}\|_0 = p \end{aligned} \tag{8}$$

However, it is hard to solve the optimal problem in Equation (8) under non-convex constraints. Hence, the maximum rank of the constraint selection matrix C is established as the maximum number p of sensors to be selected, which is represented as $r(C) \leq p$, and then it is incorporated into the objective function as a penalty term. At the same time, due to the sparsity of the weights of outliers caused by large noise, the sparse constraint term $\|\bar{W}\|_1$ is also added as a penalty term in Equation (8), resulting in the following:

$$\begin{aligned} & C \leftarrow \operatorname{argmin}_{A,C,W} \frac{1}{2} \|\sqrt{W} \odot (X - ACX)\|_F^2 + \alpha r(C) + \beta \|\bar{W}\|_1 \\ & \text{s.t. } W + \bar{W} = \mathbf{1}, W \text{ and } \bar{W} \in \{0, 1\}^{n \times m} \\ & C \in \mathbb{R}_+^{p \times n}, \end{aligned} \tag{9}$$

wherein the penalty terms of $\alpha r(C)$ and $\beta \|\bar{W}\|_1$ are used to limit the number of elements in the selected sensor set $J = [\zeta_1, \zeta_2, \dots, \zeta_p]$ and the sparsity of outliers, respectively. The indicators of the noise matrix can be relaxed from binary W and $\bar{W} \in \{0, 1\}^{n \times m}$ into real-valued W and $\bar{W} \in [0, 1]^{n \times m}$ and be measured using entropy $-w_{ij} \log w_{ij}$ and $-\bar{w}_{ij} \log \bar{w}_{ij}$, as in [37]. The maximum entropy criterion indicates that the distribution with higher entropy can more accurately represent the probability distribution of system variables. Consequently, Equation (9) is updated as follows:

$$\begin{aligned} & C \leftarrow \operatorname{argmin}_{A,C,W} \frac{1}{2} \|\sqrt{W} \odot (X - ACX)\|_F^2 + \alpha r(C) + \beta \|\bar{W}\|_1 \\ & \quad + \gamma \sum_{i,j} (w_{ij} \log w_{ij} + \bar{w}_{ij} \log \bar{w}_{ij}) \\ & \text{s.t. } W + \bar{W} = \mathbf{1}, W \text{ and } \bar{W} \in [0, 1]^{n \times m} \\ & C \in \mathbb{R}_+^{p \times n}, \end{aligned} \tag{10}$$

Equation (10) represents the model of our robust sparse sensor placement problem proposed in this study. Following this, we provide the method for solving the model.

3.2. Algorithm Development

It is challenging to solve the problem of robust sparse sensor placement based on the indicators of noise in Equation (10) directly because of the non-convex components. Therefore, the BCU iterative method [38] is introduced here to solve the problem.

Firstly, for the variables C , A , and W to be solved in Equation (10), in order to simplify the calculation process, let the following:

$$F(C, A, W) = \frac{1}{2} \|\sqrt{W} \odot (X - ACX)\|_F^2 \tag{11}$$

$$G_\alpha(C) = \alpha \|C\|_{2,1} \tag{12}$$

$$H(W) = \beta \|\bar{W}\|_1 + \gamma \sum_{i,j} (w_{ij} \log w_{ij} + \bar{w}_{ij} \log \bar{w}_{ij}) \tag{13}$$

Please note that $\bar{W} = \mathbf{1} - W$. Then, the total objective function can be expressed as follows: $O(C, A, W) = F(C, A, W) + G_\alpha(C) + H(W)$. During the k -th iteration, each variable is updated individually according to the following formula:

$$C^{k+1} = \underset{C \in \mathbb{R}_+^{p \times N}}{\operatorname{argmin}} \left\langle \nabla_C F(\hat{C}^k, A^k, W^k), C - \hat{C}^k \right\rangle + \frac{L_C^k}{2} \|C - \hat{C}^k\|_F^2 + G_\alpha(C) \tag{14a}$$

$$A^{k+1} = \underset{A}{\operatorname{argmin}} F(C^{k+1}, A^k, W^k) \tag{14b}$$

$$W^{k+1} = \underset{W}{\operatorname{argmin}} [F(C^{k+1}, A^{k+1}, W^k) + H(W^k)] \tag{14c}$$

where L_C^k is the Lipschitz constant of $\nabla_C F(C, A^k, W^k)$. We also have the following:

$$\hat{C}^k = C^k + \omega_k (C^k - C^{k-1}) \tag{15}$$

where $\omega_k \in [0, 1]$ is the extrapolated weight set according to the BCU method.

The Lipschitz constant L_C^k at the k -th step of the iteration process can be obtained by computing $\nabla_C F(C, A, W)$. Through matrix calculation, the following is easy to derive:

$$\nabla_C F(C, A, W) = -A^T [W \odot (X - ACX)] X^T \tag{16}$$

Assuming that two matrix variables, \hat{C} and \tilde{C} , are provided, then apply the following:

$$\begin{aligned} & \|\nabla_C F(\hat{C}, A, W) - \nabla_C F(\tilde{C}, A, W)\|_F \\ &= \|-A^T [W \odot (X - A\hat{C}X)] X^T + A^T [W \odot (X - A\tilde{C}X)] X^T\|_F \\ &= \|A^T \{W \odot [A(\hat{C} - \tilde{C})X]\} X^T\|_F \\ &\leq \|A^T\|_2 \|W\|_2 \|A\|_2 \|\hat{C} - \tilde{C}\|_F \|X\|_2 \|X^T\|_2 \\ &= \|A\|_2^2 \|X\|_2^2 \|W\|_2 \|\hat{C} - \tilde{C}\|_F \end{aligned} \tag{17}$$

The inequality part in Equation (17) is transformed using the Cauchy–Schwarz inequality, i.e., $\|UV\|_F \leq \|U\|_2 \|V\|_F$. Therefore, according to the definition, the Lipschitz constant of the partial derivative function $\nabla_C F(C, A, W)$ with respect to C is identified as follows:

$$L_C^k = \|A^k\|_2^2 \|X\|_2^2 \|W^k\|_2 \tag{18}$$

According to reference [38], the extrapolated weights in Equation (15) are set as follows:

$$\omega_k = \min(\hat{\omega}_k, \delta_\omega \sqrt{\frac{L_C^{k-1}}{L_C^k}}) \tag{19}$$

where $\delta_\omega < 1$, $\hat{\omega}_k = (t_{k-1} - 1)/t_k$ and $t_0 = 1$, $t_k = (1 + \sqrt{1 + 4t_{k-1}^2})/2$.

References [39,40] indicate that choosing appropriate extrapolation weight values can significantly enhance the BCU method for solving multi-block concave optimization problems.

3.2.1. C-Subproblem Solution

Given that the location selection matrix C is composed of 0 and 1 elements, it is difficult to directly apply the gradient method for its solution. Therefore, referring to [35], the location selection matrix C is updated using the block proximal gradient method. Specifically, apply the following:

First, rewrite Equation (14a) as follows:

$$\operatorname{argmin}_{C \in \mathbb{R}_+^{p \times n}} \frac{1}{2} \|C - S\|_F^2 + \lambda \|C\|_{2,1} \tag{20}$$

with $S = \hat{C}^k - \frac{1}{L_C^k} \nabla_C F(\hat{C}^k, A^k, W^k)$ and $\lambda = \frac{\alpha}{L_C^k}$.

Then, the problem in Equation (20) can be decomposed into n independent subproblems, each corresponding to a column of matrices C and S ; referring to [34], it can be reformulated in the following specific manner:

$$\operatorname{argmin}_{\mathbf{c} \geq 0} \frac{1}{2} \|\mathbf{c} - \mathbf{s}\|_2^2 + \lambda \|\mathbf{c}\|_2 \tag{21}$$

Equation (21) has a closed-form solution, where \mathbf{s} represents the i -th column of the matrix S , Ω represents the index set of the positive elements of \mathbf{s} , and $\|\mathbf{s}_\Omega\|_2 > 0$. Then, apply the following:

$$\mathbf{c}_\Omega = (\|\mathbf{s}_\Omega\|_2 - \lambda) \frac{\mathbf{s}_\Omega}{\|\mathbf{s}_\Omega\|_2} \tag{22}$$

For ease of understanding, we introduce the complete update process of matrix C , as outlined in Algorithm 1, as follows:

Algorithm 1: Nonnegative group Lasso Proximal Operator $C = \operatorname{Prox}\text{-NGL}(S, \lambda)$

- 1: **Input:** S, λ .
 - 2: **Initialize:** $C^0 \in \mathbb{R}^{p \times n} \leftarrow 0, k = 0$.
 - 3: **for** i **in range** (n):
 - 4: $\mathbf{s} = S_{:,i}, \mathbf{c} = 0$.
 - 5: **for** j **in range** (p):
 - 6: **if** $s_j > 0$, **then:**
 - 7: $\Omega_k = j, k = k + 1$.
 - 8: **end if**
 - 9: **end for**
 - 10: **if** $\|\mathbf{s}_\Omega\|_2 > \lambda$, **then:**
 - 11: $\mathbf{c}_\Omega = (\|\mathbf{s}_\Omega\|_2 - \lambda) \mathbf{s}_\Omega / \|\mathbf{s}_\Omega\|_2$.
 - 12: **end if**
 - 13: $C_{i,:} = \mathbf{c}$.
 - 14: **end for**
 - 15: **Output:** C
-

3.2.2. A-Subproblem Solution

The solution for determining A can be obtained by solving Equation (14b), which is as follows:

$$A^{k+1} = \operatorname{argmin}_A \frac{1}{2} \|\sqrt{W^k} \odot (X - AC^{k+1}X)\|_F^2 \tag{23}$$

By taking the first-order partial derivative of the right side of this equation with respect to A and setting it to zero, we have the following: $W^k \odot (C^{k+1}X)^T (X - AC^{k+1}X) = 0$.

Therefore, we can obtain the following:

$$A^{k+1} = X(C^{k+1}X)^\dagger \tag{24}$$

where $(\cdot)^\dagger$ is the pseudoinverse. The update result of this step exactly matches the method proposed in [36], which utilizes the training dataset as the basis library.

3.2.3. $W - \bar{W}$ Subproblem Solution

The solutions for W and \bar{W} can be obtained by solving Equation (14c). Since this subproblem involves two variables, W and \bar{W} , and $W + \bar{W} = \mathbf{1}_{n \times m}$, the Lagrange multiplier method is used for solving it. At the same time, the term (i, j) is taken as an example to convert the variables W and \bar{W} in Equations (11) and (13). The deformed function is then as follows:

$$H_L = \frac{1}{2}w_{ij} \left[X - A^{k+1}C^{k+1}X \right]_{ij}^2 + \beta\bar{w}_{ij} + \gamma(w_{ij} \log w_{ij} + \bar{w}_{ij} \log \bar{w}_{ij}) + \rho_i(w_{ij} + \bar{w}_{ij} - 1) \tag{25}$$

where ρ_i is the Lagrange multiplier.

By taking the first-order partial derivative with respect to w_{ij} , \bar{w}_{ij} , and ρ_i , respectively, in Equation (25) and setting these expressions to zero, we have the following:

$$\frac{\partial H_L}{\partial w_{ij}} = \frac{1}{2} \left[X - A^{k+1}C^{k+1}X \right]_{ij}^2 + \gamma \log w_{ij} + \gamma + \rho_i = 0 \tag{26}$$

$$\frac{\partial H_L}{\partial \bar{w}_{ij}} = \beta + \gamma \log \bar{w}_{ij} + \gamma + \rho_i = 0 \tag{27}$$

$$\frac{\partial H_L}{\partial \rho_i} = w_{ij} + \bar{w}_{ij} - 1 = 0 \tag{28}$$

By solving Equations (26)–(28), the optimal solution for w_{ij} can be obtained as follows:

$$w_{ij}^{k+1} \leftarrow \frac{\exp(\beta/\gamma)}{\exp([X - A^{k+1}C^{k+1}X]_{ij}^2 / (2\gamma)) + \exp(\beta/\gamma)} = \frac{1}{\exp(([X - A^{k+1}C^{k+1}X]_{ij}^2 / 2 - \beta) / \gamma) + 1} \tag{29}$$

Correspondingly, \bar{w}_{ij} can be obtained as follows: $\bar{w}_{ij}^{k+1} \leftarrow 1 - w_{ij}^{k+1}$.

3.3. Algorithm and Computational Complexity

After solving each subproblem, the pseudocode of the entire algorithm can be derived as Algorithm 2, which is named the algorithm of robust sparse sensor placement based on indicator of noise (RSSPIN).

Now, let us analyze the computational complexity of Algorithm 2. For Algorithm 2, the computational complexity is rooted in the size of data matrix X . If matrix X is sparse, the computational complexity will be greatly reduced. For the sensor deployment problem, the number of locations to be selected is $p \ll n$ and $p < m$. The main parts with higher computational complexity during the execution of Algorithm 2 are step 5 (Update C), step 6 (Update A), and step 8 (Update w_{ij}). In this case, when Algorithm 1 is executed at step 5 to update C, the primary task is to complete the calculation of Equation (16). Therefore, when performing gradient calculation according to Equation (16), a total of $2pn^2m + mn$ floating-point operations are executed. When updating A, a total of $pnm + r_{CX}^2(p + m) + n^2m$ float operations are executed according to Equation (24), wherein r_{CX} is the rank of matrix CX . When updating w_{ij} , a total of pn^2m floating-point operations are executed. Step 8 performs floating-point operations mn times. Since $p \ll n$, the computational complexity of each iteration can be approximated as follows: $O(pn^2m + mn + n^2m + pnm) = O((p + 1)(n + 1)nm)$. This result shows that the complexity of Algorithm 2 mainly depends on the number of training samples and candidate locations for sensor placement.

Algorithm 2: Robust Sparse Sensor Placement based on Indicator of Noise (RSSPIN)

```

1: Input: Data matrix  $X$ , number of sensors  $p$ , support  $\Sigma$ , parameter  $\alpha, \beta, \gamma$ .
2: Initialize:  $C^0 \in \mathbb{R}_+^{p \times n}$ ,  $A^0 \in \mathbb{R}^{n \times p}$ ,  $W^0 \in \mathbb{R}^{n \times m} \leftarrow \Sigma \odot \mathbf{1}$ ,  $0 < \delta_\omega < 1$ ,  $k = 0$ .
3: While Not convergent do:
4:     Compute  $L_C^k$  according to Equation (18).
5:     Update  $C^{k+1}$  according to Algorithm 1.
6:     Update  $A^{k+1}$  according to Equation (24).
7:     for  $\forall(i, j) \ \& \ \Sigma_{ij}$ do:
8:         Update  $w_{ij}^{k+1}$  according to Equation (29).
9:     end for
10:    if  $O(C^{k+1}, A^{k+1}, W^{k+1}) \geq O(C^k, A^k, W^k)$ , then:
11:        Set  $\hat{C}^k = C^k$ .
12:    else:
13:        Compute  $\omega_k$  according to Equation (19).
14:        Get  $\hat{C}^k$  according to Equation (15).
15:    end if
16:    Let  $k \leftarrow k + 1$ .
17: end while
18: Normalize each column of  $C' = C^k$ .
19: Sort  $\|C'_{:,i}\|_2$ ,  $i = 1 \cdots n$ , select sensors corresponding to the  $p$  largest ones as  $J$ .
20: Output:  $C^k, A^k, J$ .

```

3.4. Convergence Analysis

In this section, we analyze the convergence of the RSSPIN algorithm. First, define the indicator function of the nonnegative quadrant as follows:

$$l_+(C) = \begin{cases} 0, & \text{if } C \geq 0 \\ +\infty, & \text{otherwise} \end{cases} \quad (30)$$

And let the following:

$$Q(C, A, W) = F(C, A, W) + G_\alpha(C) + H(W) + l_+(C) \quad (31)$$

Then, the optimum problem in Equation (10) is equivalent to the following: $\min_{A, C, W} Q(C, A, W)$.

The first-order optimal condition is as follows: $0 \in \partial Q(C, A, W)$, and any point satisfying this condition is a critical point of Equation (10).

In order to prove that the proposed Algorithm 2 can converge effectively, we first present Lemma 1 and Proposition 1 for the $W - \bar{W}$ subproblem, followed by Theorem 1, which can prove the convergence of Algorithm 2, as follows:

Lemma 1. *At stage k , with a fixed $A^k C^k$, the solutions, which are w_{ij} of the $W - \bar{W}$ subproblem in Equation (29), are globally optimal for their respective intermediary problems.*

Proof. Taking the subproblem in Equations (11) and (13) as examples, having fixed $A^k C^k$, the objective function in Equation (14c) is convex with respect to w_{ij} . The solution in Equation (29) is computed using the Lagrange multiplier method in Equation (25), which guarantees that the solution is feasible and satisfies the KKT conditions of Equation (25). \square

Proposition 1. *The function in Equation (29), which contains two parameters, β and γ , and lets $\eta_{ij} = [X - ACX]_{ij}^2/2$, has the following properties:*

1. $w_{ij}(\beta, \gamma, \eta_{ij})$ is monotonically decreasing with respect to η_{ij} , which holds that the following: $\lim_{\eta_{ij} \rightarrow 0} w_{ij}(\beta, \gamma, \eta_{ij}) = \frac{1}{\exp(-\beta/\gamma)+1}$ and $\lim_{\eta_{ij} \rightarrow \infty} w_{ij}(\beta, \gamma, \eta_{ij}) = 0$;

2. $w_{ij}(\beta, \gamma, \eta_{ij})$ is monotonically increasing with respect to β , which holds that $\lim_{\beta \rightarrow 0} w_{ij}(\beta, \gamma, \eta_{ij}) = \frac{1}{\exp(\eta_{ij}/\gamma)+1}$ and $\lim_{\beta \rightarrow \infty} w_{ij}(\beta, \gamma, \eta_{ij}) = 1$;
3. $w_{ij}(\beta, \gamma, \eta_{ij})$ is an inverse “S”-shaped function, which approximates a binary function when $\gamma \rightarrow 0$ and remains constant at 1/2 when $\gamma \rightarrow +\infty$.

Each statement addresses one target parameter, while the others are held constant.

Proof. It can be easily verified by the definition. \square

Theorem 1. (Iterate subsequence convergence). Let $\{C^k, A^k, W^k\}_{k=1}^{\infty}$ be the sequence generated from Algorithm 2. The sequence of $Q(C^k, A^k, W^k)$ produced by the BCU method converges in a monotonic manner.

Proof. For C , we can deduce according to Lemma 2.1 in [38], as follows:

$$Q(A^k, C^{k+1}, W^k) - Q(A^k, C^k, W^k) \geq \frac{L_C^k}{2} \|C^{k+1} - C^k\|_F^2 \geq 0 \tag{32}$$

where $L_C^k \geq 0$ according to Equation (18).

For A , referencing Lemma 3.1 of [41], we have the following:

$$\begin{aligned} & Q(C^{k+1}, A^k, W^k) - Q(C^{k+1}, A^{k+1}, W^k) \\ &= \frac{1}{2} \|\sqrt{W^k} \odot (X - A^k C^{k+1} X)\|_F^2 - \frac{1}{2} \|\sqrt{W^k} \odot (X - A^{k+1} C^{k+1} X)\|_F^2 \\ &= \frac{1}{2} \|\sqrt{W^k} \odot [(A^k - A^{k+1}) C^{k+1} X]\|_F^2 \geq 0 \end{aligned} \tag{33}$$

For W , according to the aforementioned Lemma 1, we have the following:

$$Q(C^{k+1}, A^{k+1}, W^k) - Q(C^{k+1}, A^{k+1}, W^{k+1}) \geq 0 \tag{34}$$

From Equations (32)–(34), the following can be deduced:

$$Q(C^k, A^k, W^k) \geq Q(C^{k+1}, A^k, W^k) \geq Q(C^{k+1}, A^{k+1}, W^k) \geq Q(C^{k+1}, A^{k+1}, W^{k+1}) \tag{35}$$

In other words, the objective function in Equation (10) has a lowest bound. Thus, Algorithm 2 can be guaranteed to converge. \square

4. Experimental Evaluation and Results

4.1. Dataset and Quality of Reconstruction

4.1.1. Dataset

The algorithm presented in this paper is fundamentally applicable to flow field data characterized by low-rank properties, which can be sparsely represented. As a result, it is feasible to reconstruct the global flow field effectively using data from a limited number of sampling points, thereby enabling efficient global monitoring through sparsely deployed sensors. To rigorously evaluate the performance of the algorithm, two real-world oceanographic datasets were utilized. The specific details of these datasets are introduced as follows:

(a) The sea surface temperature (SST) dataset is used to present and evaluate Algorithm 2. The SST dataset is downloaded from the website of the NOAA Physical Sciences Laboratory, which is part of the NOAA optimum interpolation (OI) sea surface temperature (SST) V2 dataset [42]. The SST dataset comprises temperature values of 360×180 locations, encompassing both global land and sea areas. However, in this instance, only 44,219 candidate locations in the sea area are used for algorithm evaluation. Given that the application scenario is set for ocean monitoring, the data from the remaining $360 \times 180 - 44,219$ land

locations are not taken into consideration. Furthermore, the SST dataset includes 1724 times measurements (once a week) recorded from 1990 up until the date of download.

(b) The global ocean salinity dataset is also used to evaluate Algorithm 2. This salinity dataset was downloaded from the website of the Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences [43,44]. It includes bias-corrected in situ observations from the World Ocean Database, as well as additional data owned by the IAP. The dataset employed in this study has a horizontal resolution of $1^\circ \times 1^\circ$ at vertical levels of 1 m, and it provides monthly data spanning from 2010 to 2019. Furthermore, the dataset contains 43,401 valid seawater salinity data coordinates. These 43,401 coordinates are used as candidate locations for evaluating Algorithm 2.

4.1.2. Quality of Reconstruction

(a) Reconstruction error of subspace learning

Given that the objective function of Equation (6) evaluates the error in subspace learning and is based on the principle of simultaneously obtaining both the feature space and the mapping relationship in subspace learning, the algorithm proposed in this paper iteratively updates both selection matrix C and reconstruction matrix A . Therefore, a viable approach involves comparing the subspace learning error with that of the benchmark methods. The performance of the proposed method is evaluated using reconstruction errors, which are expressed as follows:

$$Error1 = \frac{\|X_{te} - \hat{X}_{te}\|}{\|X_{te}\|} = \frac{\|X_{te} - ACX_{te}\|}{\|X_{te}\|} \quad (36)$$

where A and C are obtained through Algorithm 2 by imputing training data X_{tr} , while X_{te} represents the test data partitioned from the entire dataset X .

(b) Reconstruction error of low-dimensional sampled data

Upon obtaining index J through Algorithm 2, the low-dimensional data sampled by the deployed sensors are as follows: $Y = X_{J,:}$. The low-dimensional data representation obtained from the test set sampling is as follows: $Y_{te} = X_{te J,:}$. The reconstruction error of the low-dimensional data obtained from sampling can be expressed as follows:

$$Error2 = \frac{\|X_{te} - \hat{X}_{te}\|}{\|X_{te}\|} = \frac{\|X_{te} - \hat{A}Y_{te}\|}{\|X_{te}\|} \quad (37)$$

where $\hat{A} = ACX_{tr}Y_{tr}^\dagger$ is the reconstruction matrix derived from the low-dimensional test data.

The method for obtaining \hat{A} is detailed in Appendix A. $Y_{tr} = X_{tr J,:}$ represents the low-dimensional sampling data derived from the training dataset X_{tr} , indexed by J . Both matrices A and C are also derived through Algorithm 2.

4.1.3. Experimental Settings

In this scenario, the SST dataset is partitioned, with 1040 samples used as the training set and the remaining 687 sample data used as the test set. The salinity dataset is partitioned, with 100 samples used as the training set and the remaining 20 samples used as the test set. The robustness considered in this paper includes robustness to noise and outliers. Here, outliers mainly refer to incomplete data, which can be attributed to special events such as sensor failures or certain natural conditions. The robustness is evaluated by introducing a proportion of outliers into the experimental data.

The experiments were conducted using MATLAB R2022a as the programming language. The hardware configuration for the experiments included a computer running Windows 11 Professional, which is a 64-bit operating system. The system was equipped

with 16 GB of RAM and powered by an AMD Ryzen 5 5600G processor (AMD, Santa Clara, CA, USA).

To test the effectiveness of the relevant parameters, a low-rank data matrix was constructed, as sensor deployment methods based on sparse selection are theoretically suitable for flow field data with low-rank structures. During testing, the low-rank data matrix was constructed with dimensions of 100×100 and a rank of five, following the methods referenced in [38,39]. When executing the proposed algorithm, the outlier ratio was set to 0.2 and the signal-to-noise ratio (SNR) was 60. Figure 1 shows the reconstruction error and execution time for different parameters for this low-rank data matrix. The number of iterations is set to 500. The primary objective here is to identify an effective range of parameters rather than the optimal parameters. Effective parameters imply that the algorithm can continuously iterate and update until convergence or the maximum number of iterations is reached. In Figure 1, γ is set to 1×10^{-5} . Figure 1a indicates that the reconstruction error varies with changes in α and β . Figure 1b demonstrates that the execution time of the proposed algorithm significantly differs with changes in α and β . Some unsuitable parameters cause the algorithm to terminate iterations quickly, while parameters within an appropriate range allow the algorithm to continue iterative optimization. For instance, when the value of β is set to 1, the algorithm terminates the iterative process very quickly, resulting in a relatively high reconstruction error. This indicates that $\beta = 1$ does not facilitate effective variable updates and is therefore unsuitable for this algorithm. Conversely, when the value of β ranges from 1×10^{-4} to 1×10^{-32} , the algorithm can perform effective iterative solving, with minimal differences in reconstruction error in most cases.

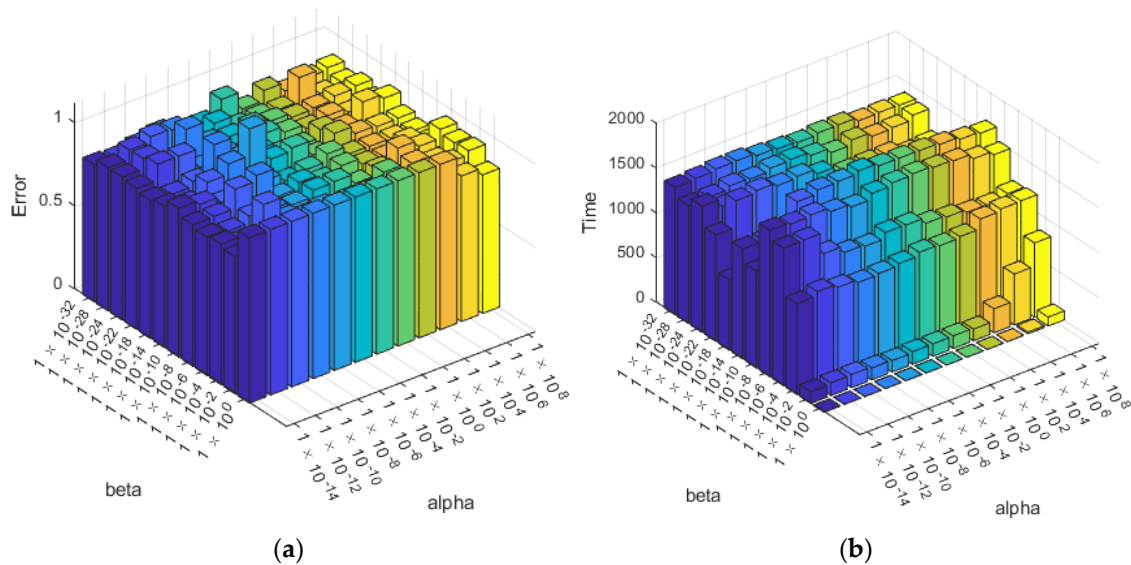


Figure 1. Reconstruction for low-rank data matrix of different parameters in RSSPIN. (a) Reconstruction error of different α and β ; (b) Execution time of different α and β .

Tests conducted on this low-rank matrix indicate that provided the algorithm can effectively converge through iterations, the parameter sizes can be fine-tuned with relatively little impact on the reconstruction error. However, the different data scales of actual datasets still influence the choice of parameters. Therefore, the parameters used for the subsequent two real datasets are slightly different but remain adjustable.

For convenience of calculation, the data were normalized when implementing Algorithm 2. During the implementation of the algorithm in this study, the samples were chosen sequentially, with the first sample from the test dataset used as the test data. Naturally, it can also be adjusted to include multiple samples as required.

The proposed method is evaluated and contrasted with four other data-driven sparse reconstruction methods in [36], as follows: random selection (RS), QR-conditioning [23], POD, and sparse representation (SR). The sensor selection method and reconstruction basis adopted in these methods are listed in Table 1 below. Specifically, \hat{X}_{te} of RS and QR is calculated using the following: $\hat{X}_{te} = (\Phi/\Phi_{J,:}) \cdot X_{te J,:}$, wherein $\Phi = X_{tr}G$, $G \in \mathbb{R}^{m \times 2p}$, and its elements follow a standard normal distribution. \hat{X}_{te} of POD is calculated using the following: $\hat{X}_{te} = \Psi_r \cdot (C\Psi_r/X_{te J,:})$, wherein Ψ_r is the first r columns of matrix Ψ , which is computed using the SVD of training dataset X_{tr} . \hat{X}_{te} of SR is calculated using the following: $\hat{X}_{te} = \Psi_{tr}\hat{s}$, wherein $\Psi_{tr} = X_{tr J,:}$ is the training library, which is set as selected samples in training dataset X_{tr} , and \hat{s} is the coefficient vector acquired by solving the appropriate optimization problem in [36].

Table 1. Different methods for sparse sensor placement.

Method	Sensor Selection	Reconstruction Basis
RS	Random	Randomized rank reduction
QR	Column pivot	Randomized rank reduction
POD	Random	Reduced order mode
SR	Random	Training library
RSSPIN	Iteration of BCU	Iteration of BCU

4.2. Experimental Results of SST

In this section, the proposed algorithm, RSSPIN, is initially executed to demonstrate its convergence. Subsequently, the proposed RSSPIN is compared with the aforementioned methods by setting the outlier ratio at 0, 0.2, and 0.4. Finally, the reconstructed data and sensor locations are provided. The parameters adopted here are as follows: α is set to 1×10^6 , β is set to 3.4263×10^{-29} , γ is set to 1×10^{-4} , and δ_ω is set to 0.5. These parameters are mainly set according to whether effective updating and iteration can be achieved during the execution of Algorithms 1 and 2.

4.2.1. Convergence of RSSPIN

The converge analysis of the RSSPIN algorithm is performed with varying numbers of training samples. It is important to mention that the outlier rate (ra) in Figure 2a,b is set to 0, signifying that the proportion of incomplete data is zero. Figure 2a indicates that the RSSPIN algorithm can effectively converge when varying numbers of samples are input into the algorithm, wherein more samples lead to relatively larger values of all parts in the overall objective $O(C, A, W)$, resulting in a slower convergence of the overall objective. Figure 2b illustrates that when a larger number of training samples are input, the increase in the total data leads to heavier computation requirements. Consequently, the reconstruction error of the normalized data has not been effectively improved within the same number of iterations. In addition, given that there are multiple sub-objective functions within the total objective function, it is logical to expect some fluctuations in the reconstruction error value when examining the reconstruction error without noise weighting. Variables A and W are initialized randomly because L_c in Equation (18) appears as a denominator when solving for variable C in Algorithm 1, and thus its value cannot be zero. This initialization method may also cause instability in the sub-objective function.

Moreover, as the number of iterations increases, the reconstruction error exhibits a trend of continuous decrease. Furthermore, to aid in boosting the overall iteration speed of the proposed algorithm, subsequent comparative experiments are conducted by using 50 samples as a reference, taking into account the proportion of outliers and comparisons with other methods. Certainly, the proposed method can improve reconstruction accuracy by performing more iterations when using a larger number of samples. However, due to the time-consuming nature of unlimited iterations, this approach has not been adopted in the experiments of this study.

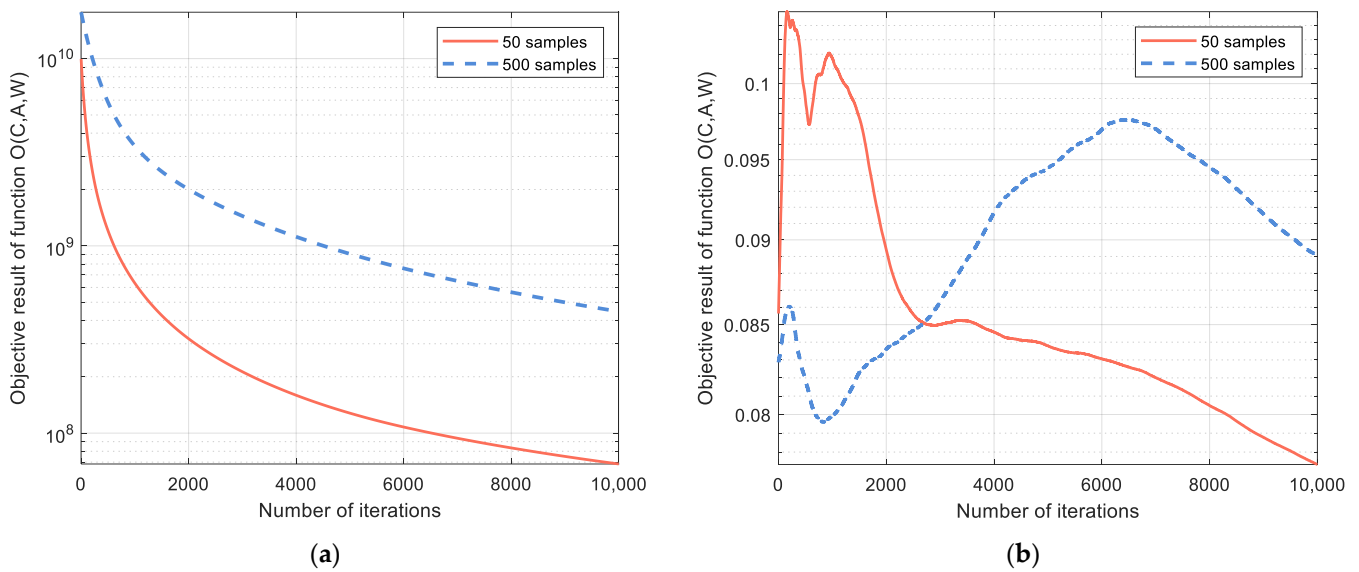


Figure 2. Convergence for different numbers of samples in RSSPIN, in which normalized data are used. (a) Convergence rate of total objective results without outliers in iteration; (b) convergence rate of reconstruction errors without outliers in iteration.

Figure 3a,b contrasts the performance of the algorithm proposed in this study when the input training data contain varying proportions of outliers. Specifically, referring to Figure 3a, it can be seen that as the outlier ratio increases, the convergence speed of the algorithm accelerates due to more outliers (incomplete data) resulting in fewer data participating in the training. However, referring to Figure 3b, it is distinctly noticeable that as the outlier ratio rises, the reconstruction error of the algorithm enlarges when the iteration results tend to stabilize with the increase in the number of iterations. This phenomenon is understandable, primarily because more incomplete data lead to the loss of valuable data information, resulting in a deterioration of the algorithm’s reconstruction accuracy.

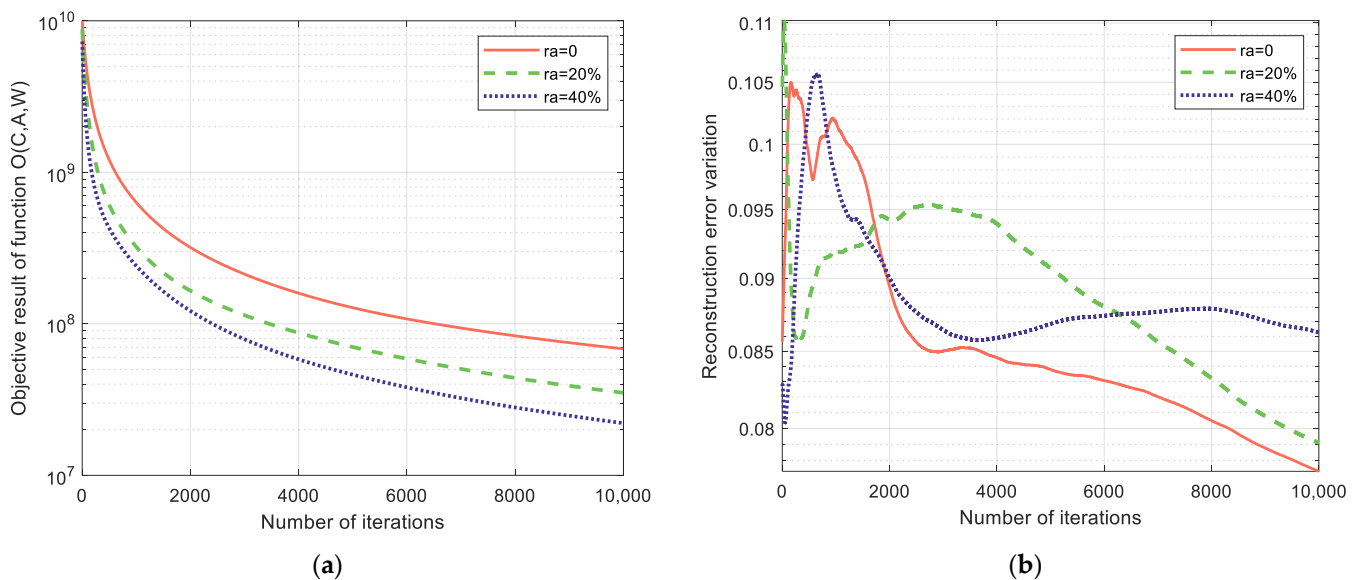


Figure 3. Convergence rate for different outlier ratios in RSSPIN. (a) Convergence rate of total objective results with different outlier ratios in iteration; (b) convergence rate of reconstruction errors with different outlier ratios in iteration.

4.2.2. Reconstruction Error of Different Methods

This section compares the proposed RSSPIN algorithm with other relevant data-driven methods mentioned earlier. Since the sensor locations or base matrices in the comparison methods are randomly determined, increasing the number of executions does not significantly affect the reconstruction error. Therefore, the comparison methods were executed 100 times, and the median reconstruction error of sensor deployment was compared.

After training with the same dataset, the reconstruction error of the test data (using normalized test data for the comparison methods) is illustrated in Figure 4a,b. The error results for our method, as shown in Figure 4a,b, are calculated using Equation (36).

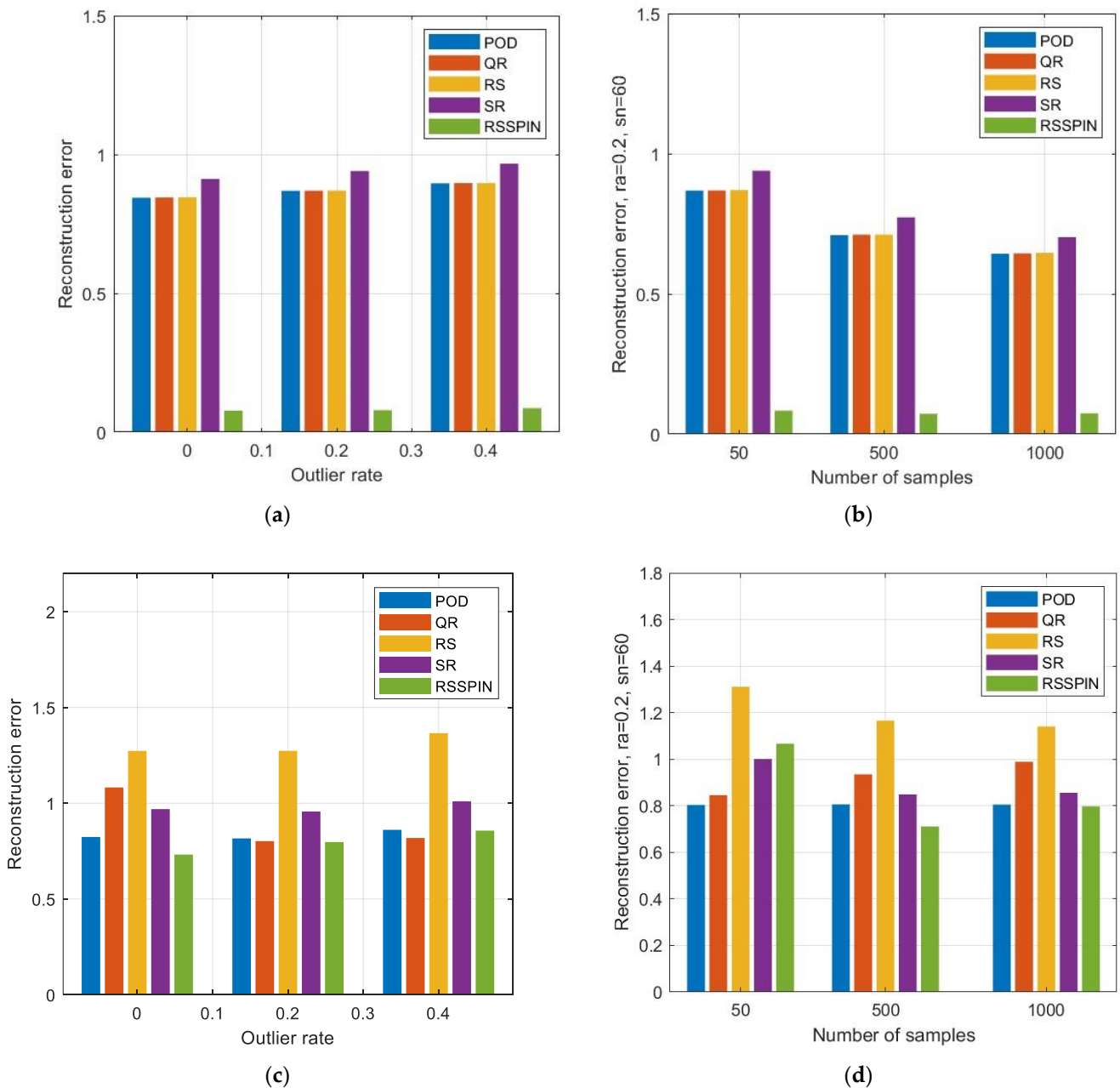


Figure 4. Reconstruction error of different methods for SST. (a) Reconstruction errors of different outlier rates using Equation (36); (b) reconstruction errors of different samples using Equation (36); (c) reconstruction errors of different outlier rates using Equation (37); (d) reconstruction errors of different samples using Equation (37).

Figure 4a demonstrates that the method proposed in this study exhibits a significantly lower reconstruction error across various outlier ratios ($ra = 0, 0.2, 0.4$). Specifically, referring to Figure 4a, when using 50 samples for training, the performance of the comparison methods deteriorates markedly, particularly for the sparse representation (SR) training library, which is more reliant on the number of historical data samples. This indicates that the method proposed in this study can achieve superior data reconstruction through iterative calculations with a limited number of samples.

In Figure 4b, when $ra = 0.2$, additional noise is added to the training data with varying sample sizes. The signal-to-noise ratio, $sn = 60$, is set, so the training data include not only their inherent measurement noise and model noise but also the impact of this additional noise. The comparison results from Figure 4b demonstrate that even after adding additional noise, the method proposed in this study still exhibits significantly smaller reconstruction errors. This further illustrates that the reconstruction performance of the comparison methods improves as the number of training samples increases, whereas the method proposed in this study shows a relatively low dependence on the number of samples with additional noise.

Refer to Figure 4c,d for the comparison of non-normalized reconstruction errors. The reconstruction error is calculated according to Equation (37). As shown in Figure 4c, when there are outliers of varying proportions, the reconstruction results obtained using this method are still superior to other comparison methods, with the only exception being a slightly higher error than the QR method when the outlier proportion is 0.4. As shown in Figure 4d, when both outliers and noise are present, the reconstruction error obtained using this method is still lower than that of other comparison methods when the number of training samples is 500 and 1000. However, when the sample size is 50, the reconstruction results are not ideal. This may be because the presence of both outliers and noise makes it difficult to achieve optimal feature selection with a smaller sample size for subspace learning.

Figure 5 presents the test data and the denormalized reconstruction results obtained using the method proposed in this study. In this context, the reconstructed image is generated from the reconstructed data under the conditions of $ra = 0$ and $sn = 0$. The black dots in Figure 5b,c represent the deployment locations of the 10 sensors obtained by the RSSPIN algorithm, and the full-state data reconstruction of the entire SST is performed using the sampled data from these locations. As observed in Figure 5b, the data reconstructed using our method, even with a limited number of samples, effectively represent the distribution characteristics of the SST test data. However, as shown in Figure 5c, increasing the sample size without changing the number of iterations did not significantly improve the overall reconstruction quality of the test data. When training with a large number of samples, the algorithm proposed in this study can achieve higher reconstruction accuracy by increasing the number of iterations. Nevertheless, this approach is not as cost-effective compared to scenarios in which calculations are performed with a smaller sample size.

The reconstruction results shown in Figure 5b,c indicate that sparse sensor deployment can achieve satisfactory full-state data reconstruction. This is because SST data, being a typical type of fluid flow data, exhibit strong spatiotemporal correlations. Consequently, it is easier to obtain its low-dimensional features through data processing, making it suitable for sparse reconstruction. This further demonstrates that effective ocean monitoring can be achieved with limited sensors, which undoubtedly provides valuable insights for practical sensor deployment.

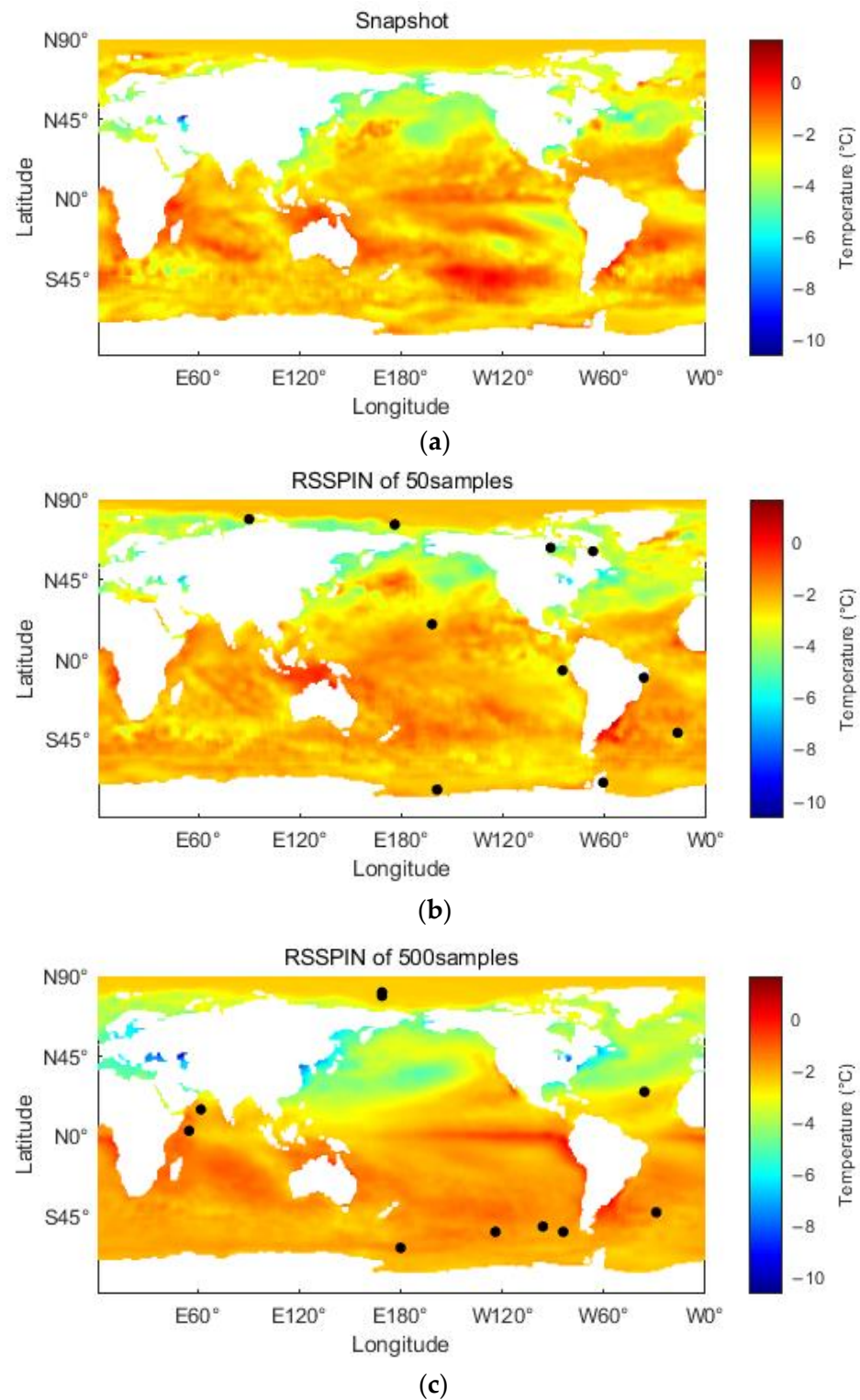


Figure 5. Reconstructed SST of different samples by RSSPIN ($R_a = 0$; $S_r = 0$). (a) Snapshot of test data; (b) reconstructed SST of 50 samples using RSSPIN; (c) reconstructed SST of 500 samples using RSSPIN.

4.3. Experimental Results of Salinity

In this section, the proposed RSSPIN is evaluated against the previously mentioned methods by setting the outlier ratio to 0.2 and 0.4. Subsequently, the reconstructed data and sensor locations are presented. The parameters used in this study are as follows: α is set to 1×10^6 , β is set to 1×10^{-6} , γ is set to 1×10^{-6} , and δ_ω is set to 0.5. Similarly,

these parameters are primarily set based on whether effective updating and iteration can be achieved during the execution of Algorithms 1 and 2. In this experiment, only Equation (37) was used for the reconstruction error comparison, and it was implemented on a non-normalized dataset. It should be understood that during the implementation of Algorithm 2, data normalization is still required. However, the reconstruction error is calculated using non-normalized test data.

4.3.1. Reconstruction Error of Different Methods

Figure 6 presents the results of applying the proposed method and the comparison methods to global ocean salinity data. As shown in Figure 6a, the reconstruction error obtained using the proposed method is significantly lower than that of the comparison methods across various outlier ratios. Figure 6b demonstrates that, even in the presence of both outliers and substantial noise, the reconstruction error achieved by the proposed method remains significantly lower than that of the comparison methods.

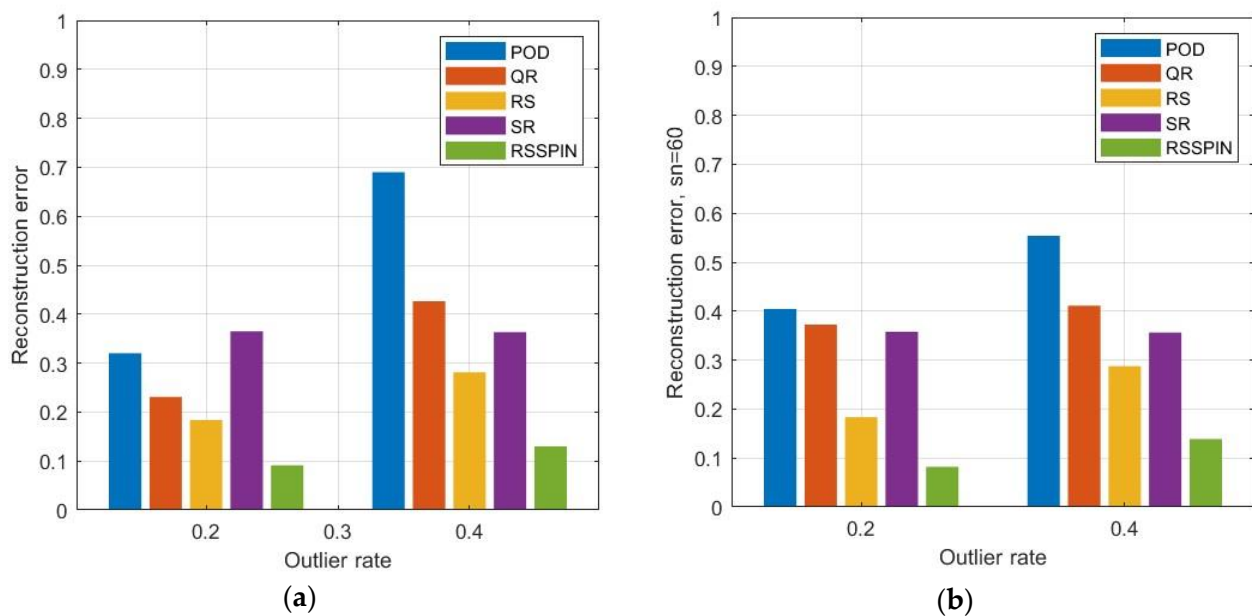


Figure 6. Reconstruction error of different methods for global ocean salinity. (a) Reconstruction errors of different outlier rates; (b) reconstruction errors of different samples.

4.3.2. Reconstruction Salinity of Outlier Ratio

Figure 7 displays the original test data of salinity and the reconstruction results of the method proposed in this study. The reconstructed global ocean salinity field is presented for different proportions of outliers in the original data. Similarly, the black dots in Figure 7b,c represent the deployment locations of the sensor subsets obtained by the proposed RSSPIN algorithm. The high-dimensional SST data reconstruction is performed using the low-dimensional data acquired from these sparsely deployed sensors. As shown in Figure 7b, although the reconstructed salinity field contains more noise points compared to the original salinity distribution, the overall salinity distribution characteristics of different regions remain clearly distinguishable. Figure 7c illustrates that as the proportion of outliers increases, the salinity reconstruction field exhibits more noise points. However, it remains possible, to some extent, to distinguish high-salinity areas (depicted in red) from low-salinity areas (depicted in blue).

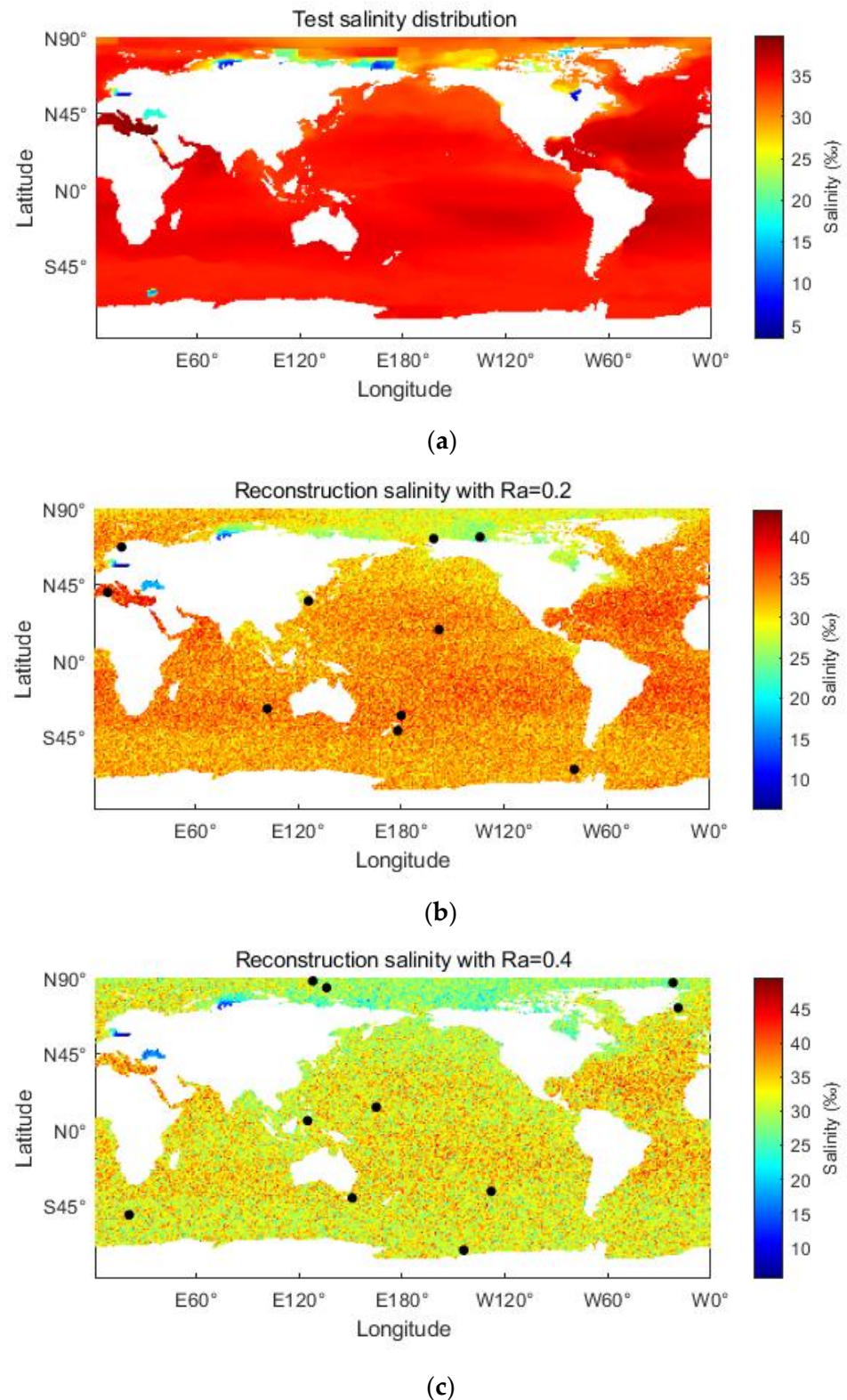


Figure 7. Reconstruction salinity field by RSSPIN. (a) Test salinity; (b) reconstruction salinity with $Ra = 0.2$; (c) reconstruction salinity with $Ra = 0.4$.

5. Conclusions and Discussion

In this study, we develop a robust sparse sensor placement strategy based on noise indicators in ocean observations. This strategy models the sparse sensor placement problem by utilizing sparse regularization terms for selected sensors and assigning a large noise

weight while adding a small noise weight for reconstructed data obtained from the selected sensors. Then, the entropy of noise weight is introduced into the model to minimize the reconstruction bias. Subsequently, a BCU-based algorithm is proposed to compute the reconstruction matrix and selection matrix. During the iterative updating process, the algorithm continuously identifies and evaluates the impact of noise while simultaneously updating the selection matrix, reconstruction matrix, and noise matrix. As a result, the selection and reconstruction matrices gradually become less influenced by data with high noise impact. Experiments demonstrate that the strategy is effective, yielding impressive reconstruction errors compared to the selected benchmark methods, even when the number of training samples is small. In this study, the reconstruction matrix and selection matrix are updated simultaneously, rather than introducing random elements into either the selection matrix or the reconstruction basis, which would result in unstable outcomes in other related data-driven reconstruction methods.

This study optimizes the deployment of sparse sensors using marine data containing noise, achieving better reconstruction of marine data under noisy conditions. This provides a meaningful reference for the selection of sensor locations in marine monitoring systems, enabling more cost-effective and efficient marine monitoring.

Furthermore, while the accuracy has yielded satisfactory results, the complexity of this strategy requires further reduction to expedite convergence. Additionally, a nonlinear reconstruction model should be considered for sparse sensor placement and compared with the linear model. Specifically, in the context of ocean monitoring, further research is necessary to optimize the deployment of multifunctional sensors. These sensors, which monitor various distinct parameters, should be comprehensively considered, taking into account the data characteristics of different physical states.

Moreover, the proposed RSSPIN algorithm in this paper addresses the non-convex problem and is influenced by multiple parameters set during the iterative optimization process. Although the parameter values used in this study enable effective iteration and solution updates, they are not yet optimal. Determining the optimal parameters will be considered in future research. For instance, we will explore the use of parameter learning algorithms from the field of machine learning, such as Bayesian learning and particle filtering, to further refine the algorithm.

Author Contributions: Conceptualization, Q.Z. and H.W.; methodology, Q.Z.; software, Q.Z.; validation, Q.Z.; formal analysis, Q.Z.; investigation, L.L.; resources, X.M. and J.X.; writing—original draft preparation, Q.Z.; writing—review and editing, Q.Z. and Y.Z.; visualization, Q.Z.; supervision, H.W.; project administration, H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant Nos. 52331012, 52071200, 52201401, 52201403, and 52102397), in part by the National Key Research and Development Program (Grant No. 2021YFC2801002), the Shanghai Committee of Science and Technology, China (Grant No. 23010502000), the China Postdoctoral Science Foundation (Grant No. 2022M712027), the Shanghai Post-doctoral Excellence Program (Grant No. 2022767), the Chenguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission (No. 23CGA61), and the Top-Notch Innovative Program for Postgraduates of Shanghai Maritime University under Grant 2022YBR012.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Notations

In this study, matrices are represented in uppercase letters. Please refer to the table below for the associated mathematical symbols and their respective physical interpretations.

Notation	Terminology
$X \in \mathbb{R}^{n \times m}$	Global measurement data with noises
$Y \in \mathbb{R}^{p \times m}$	Measurement data of selected sensors
$C \in \mathbb{R}^{p \times n}$	Selection matrix or sensor selected matrix
$A \in \mathbb{R}^{n \times p}$	Reconstruction matrix
$W \in \mathbb{R}^{n \times m}$	Smaller noise distribution weight matrix
$\bar{W} \in \mathbb{R}^{n \times m}$	Larger noise distribution weight matrix
n	Number of all candidate locations
m	Number of samples
p	Number of selected sensors
$J = [\xi_1, \xi_2, \dots, \xi_p]$	Selected sensors index set
$X_{J,:} \in \mathbb{R}^{p \times m}$	Data of selected sensors
L_C^k	The Lipschitz constant of the k -th iteration
ω_k	Extrapolated weight
α	Penalty parameter for the selected sensors
β	Penalty parameter for the outliers
γ	Penalty parameter for the maximum entropy criterion of outliers
$\ \cdot\ _2$	L_2 -norm of matrix
$\ \cdot\ _F$	The Frobenius norm of matrix
$\nabla_C F(\cdot)$	Partial derivative of function $F(\cdot)$ respect to C
\hat{C}^k	Estimated selected operator of the k -th iteration

Appendix A

In this study, for the selection matrix C , it is transformed into a low-dimensional sampling matrix \widehat{C} , with the columns corresponding to the index J being all-ones vectors, others are zero vectors. The low-dimensional data Y_{tr} sampled from training data X_{tr} can be represented as follows:

$$Y_{tr} = X_{tr J,:} = \widehat{C} X_{tr} \quad (A1)$$

\widehat{A} is the low-dimensional reconstruction matrix for the training data X_{tr} corresponding to \widehat{C} . Therefore, by definition, for the training set we have the following:

$$A C X_{tr} = \widehat{A} \widehat{C} X_{tr} = \widehat{A} Y_{tr} \quad (A2)$$

It is easy to have the following:

$$\widehat{A} = A C X_{tr} Y_{tr}^+ \quad (A3)$$

Therefore, for the test data X_{te} , the reconstructed data \hat{X}_{te} can be obtained from the limited low-dimensional measurements Y_{te} acquired by sparsely deployed sensors, as follows:

$$\hat{X}_{te} = \widehat{A} Y_{te} = A C X_{tr} Y_{tr}^+ Y_{te} \quad (A4)$$

References

- Cheng, L.; Abraham, J.; Trenberth, K.E.; Boyer, T.; Mann, M.E.; Zhu, J.; Wang, F.; Yu, F.; Locarnini, R.; Fasullo, J.; et al. New Record Ocean Temperatures and Related Climate Indicators in 2023. *Adv. Atmos. Sci.* **2024**, *41*, 1068–1082. [\[CrossRef\]](#)
- Akbari, E.; Alavipanah, S.; Jaihouni, M.; Hajeb, M.; Haase, D.; Alavipanah, S. A Review of Ocean/Sea Subsurface Water Temperature Studies from Remote Sensing and Non-Remote Sensing Methods. *Water* **2017**, *9*, 936. [\[CrossRef\]](#)
- Lin, M.; Yang, C. Ocean Observation Technologies: A Review. *Chin. J. Mech. Eng.* **2020**, *33*, 32. [\[CrossRef\]](#)
- Leonard, N.E.; Paley, D.A.; Lekien, F.; Sepulchre, R.; Fratantoni, D.M.; Davis, R.E. Collective Motion, Sensor Networks, and Ocean Sampling. *Proc. IEEE Inst. Electr. Electron. Eng.* **2007**, *95*, 48–74. [\[CrossRef\]](#)
- Yildirim, B.; Chryssostomidis, C.; Karniadakis, G.E. Efficient sensor placement for ocean measurements using low-dimensional concepts. *Ocean Model.* **2009**, *27*, 160–173. [\[CrossRef\]](#)
- Hodgson, Z.; Esnaola, I.; Jones, B. Model-Based Optimal Adaptive Monitoring of Oil Spills. *IEEE Trans. Control Syst. Technol.* **2022**, *30*, 2115–2130. [\[CrossRef\]](#)

7. Liu, K.; Yan, R.-J.; Guedes Soares, C. Optimal sensor placement and assessment for modal identification. *Ocean Eng.* **2018**, *165*, 209–220. [[CrossRef](#)]
8. Krause, A.; Singh, A.; Guestrin, C. Near-Optimal Sensor Placements in Gaussian Processes Theory, Efficient Algorithms and Empirical Studies. *J. Mach. Learn. Res.* **2008**, *9*, 235–284.
9. Mulia, I.E.; Gusman, A.R.; Williamson, A.L.; Satake, K. An Optimized Array Configuration of Tsunami Observation Network Off Southern Java, Indonesia. *J. Geophys. Res. Solid Earth* **2019**, *124*, 9622–9637. [[CrossRef](#)]
10. Benner, P.; Herzog, R.; Lang, N.; Riedel, I.; Saak, J. Comparison of model order reduction methods for optimal sensor placement for thermo-elastic models. *Eng. Optim.* **2018**, *51*, 465–483. [[CrossRef](#)]
11. Nguyen, M.T.; La, H.M.; Teague, K.A. Collaborative and Compressed Mobile Sensing for Data Collection in Distributed Robotic Networks. *IEEE Trans. Control Netw. Syst.* **2018**, *5*, 1729–1740. [[CrossRef](#)]
12. Lin, X.; Chowdhury, A.; Wang, X.; Terejanu, G. Approximate computational approaches for Bayesian sensor placement in high dimensions. *Inf. Fusion.* **2019**, *46*, 193–205. [[CrossRef](#)]
13. Jayaraman, B.; Al Mamun, S.M.A.; Lu, C. Interplay of Sensor Quantity, Placement and System Dimension in POD-Based Sparse Reconstruction of Fluid Flows. *Fluids* **2019**, *4*, 109. [[CrossRef](#)]
14. Yang, X.; Venturi, D.; Chen, C.; Chrysosostomidis, C.; Karniadakis, G.E. EOF-based constrained sensor placement and field reconstruction from noisy ocean measurements: Application to Nantucket Sound. *J. Geophys. Res. Oceans* **2010**, *115*, C12. [[CrossRef](#)]
15. Sirjacobs, D.; Alvera-Azcárate, A.; Barth, A.; Lacroix, G.; Park, Y.; Nechad, B.; Ruddick, K.; Beckers, J.M. Cloud filling of ocean colour and sea surface temperature remote sensing products over the Southern North Sea by the Data Interpolating Empirical Orthogonal Functions methodology. *J. Sea Res.* **2011**, *65*, 114–130. [[CrossRef](#)]
16. Berkooz, G.; Holmes, P.; Lumley, J.L. The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech.* **1993**, *25*, 539–575. [[CrossRef](#)]
17. Xu, Z.; Guo, Y.; Saleh, J.H. Multi-objective optimization for sensor placement: An integrated combinatorial approach with reduced order model and Gaussian process. *Measurement* **2022**, *187*, 110370. [[CrossRef](#)]
18. Schmid, P.; Sesterhenn, J. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **2010**, *656*, 5–28. [[CrossRef](#)]
19. Tu, J.H.; Rowley, C.W.; Luchtenburg, D.M.; Brunton, S.L.; Kutz, J.N. On dynamic mode decomposition: Theory and applications. *J. Comput. Nonlinear Dyn.* **2014**, *1*, 391–421. [[CrossRef](#)]
20. Erichson, N.B.; Mathelin, L.; Yao, Z.; Brunton, S.L.; Mahoney, M.W.; Kutz, J.N. Shallow neural networks for fluid flow reconstruction with limited sensors. *Proc. R. Soc. A* **2020**, *476*, 20200097. [[CrossRef](#)]
21. Sahba, S.; Wilcox, C.C.; McDaniel, A.; Shaffer, B.; Brunton, S.L.; Kutz, J.N. Wavefront sensor fusion via shallow decoder neural networks for aero-optical predictive control. In Proceedings of the SPIE Optical Engineering + Applications 2022, San Diego, CA, USA, 3 October 2022; Volume 12223.
22. Zhang, Z.; Yang, X.; Lin, G. POD-Based Constrained Sensor Placement and Field Reconstruction from Noisy Wind Measurements: A Perturbation Study. *Mathematics* **2016**, *4*, 26. [[CrossRef](#)]
23. 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 Syst.* **2018**, *38*, 63–86.
24. Clark, E.; Askham, T.; Brunton, S.L.; Kutz, J.N. Greedy Sensor Placement with Cost Constraints. *IEEE Sens. J.* **2019**, *19*, 2642–2656. [[CrossRef](#)]
25. Ma, K.-C.; Liu, L.; Sukhatme, G.S. An information-driven and disturbance-aware planning method for long-term ocean monitoring. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Republic of Korea, 9–14 October 2016.
26. Zhang, Q.; Wu, H.; Mei, X.; Han, D.; Marino, M.D.; Li, K.-C.; Guo, S. A Sparse Sensor Placement Strategy based on Information Entropy and Data Reconstruction for Ocean Monitoring. *IEEE Internet Things J.* **2023**, *10*, 19681–19694. [[CrossRef](#)]
27. Peherstorfer, B.; Drmač, Z.; Gugercin, S. Stability of discrete empirical interpolation and gappy proper orthogonal decomposition with randomized and deterministic sampling points. *SIAM J. Sci. Comput.* **2020**, *42*, A2837–A2864. [[CrossRef](#)]
28. Mei, X.; Han, D.; Saeed, N.; Wu, H.; Chang, C.-C.; Han, B.; Ma, T.; Xian, J. Trajectory Optimization of Autonomous Surface Vehicles with Outliers for Underwater Target Localization. *Remote Sens.* **2022**, *14*, 4343. [[CrossRef](#)]
29. Murugan Jaya, M.; Ceravolo, R.; Zanutti Fragonara, L.; Matta, E. An optimal sensor placement strategy for reliable expansion of mode shapes under measurement noise and modelling error. *J. Sound Vib.* **2020**, *487*, 115511. [[CrossRef](#)]
30. Mei, X.; Han, D.; Saeed, N.; Wu, H.; Han, B.; Li, K.-C. Localization in Underwater Acoustic IoT Networks: Dealing with Perturbed Anchors and Stratification. *IEEE Internet Things J.* **2024**, *11*, 17757–17769. [[CrossRef](#)]
31. Chen, X.; Liu, S.; Liu, W.; Wu, H.; Han, B.; Zhao, J. Quantifying Arctic oil spilling event risk by integrating analytic network process and fuzzy comprehensive evaluation model. *Ocean Coast. Manag.* **2022**, *228*, 106326. [[CrossRef](#)]
32. Clark, E.; Brunton, S.L.; Kutz, J.N. Multi-fidelity sensor selection-Greedy algorithms to place cheap and expensive sensors with cost constraints. *IEEE Sens. J.* **2020**, *21*, 600–611. [[CrossRef](#)]
33. Ghayem, F.; Rivet, B.; Farias, R.C.; Jutten, C. Robust Sensor Placement for Signal Extraction. *IEEE Trans. Signal Process.* **2021**, *69*, 4513–4528. [[CrossRef](#)]

34. Zhou, N.; Xu, Y.; Cheng, H.; Fang, J.; Pedrycz, W. Global and local structure preserving sparse subspace learning: An iterative approach to unsupervised feature selection. *Pattern Recognit.* **2016**, *53*, 87–101. [[CrossRef](#)]
35. Zhou, N.; Xu, Y.; Cheng, H.; Yuan, Z.; Chen, B. Maximum Correntropy Criterion-Based Sparse Subspace Learning for Unsupervised Feature Selection. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *29*, 404–417. [[CrossRef](#)]
36. Callaham, J.L.; Maeda, K.; Brunton, S.L. Robust flow reconstruction from limited measurements via sparse representation. *Phys. Rev. Fluids* **2019**, *4*, 103907. [[CrossRef](#)]
37. Guo, X.; Lin, Z. Low-Rank Matrix Recovery Via Robust Outlier Estimation. *IEEE Trans. Image Process.* **2018**, *27*, 5316–5327. [[CrossRef](#)]
38. Xu, Y.; Yin, W. A Block Coordinate Descent Method for Regularized Multiconvex Optimization with Applications to Nonnegative Tensor Factorization and Completion. *SIAM J. Imaging Sci.* **2013**, *6*, 1758–1789. [[CrossRef](#)]
39. Xu, Y.; Yin, W. A globally convergent algorithm for nonconvex optimization based on block coordinate update. *J. Sci. Comput.* **2017**, *72*, 700–734. [[CrossRef](#)]
40. Xu, Y. Alternating proximal gradient method for sparse nonnegative Tucker decomposition. *Math. Program. Comput.* **2015**, *7*, 39–70. [[CrossRef](#)]
41. Xu, Y. On higher-order singular value decomposition from incomplete data. *arXiv* **2014**, arXiv:1411.4324v1.
42. NOAA Optimum Interpolation (OI) SST V2. Available online: <https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html> (accessed on 28 February 2024).
43. Cheng, L.; Trenberth, K.E.; Gruber, N.; Abraham, J.P.; Fasullo, J.; Li, G.; Mann, M.E.; Zhao, X.; Zhu, J. Improved estimates of changes in upper ocean salinity and the hydrological cycle. *J. Clim.* **2020**, *33*, 10357–10381. [[CrossRef](#)]
44. Cheng, L.; Zhu, J. Benefits of CMIP5 multimodel ensemble in reconstructing historical ocean subsurface temperature variation. *J. Clim.* **2016**, *29*, 5393–5416. [[CrossRef](#)]

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.