A Sparse Sensor Placement Strategy based on Information Entropy and Data Reconstruction for Ocean Monitoring

Qiannan Zhang, Huafeng Wu, *Senior Member, IEEE,* Xiaojun Mei, Dezhi Han, *Member, IEEE,* Mario Donato Marino, *Member, IEEE,* Kuan-Ching Li, *Senior Member, IEEE,* and Song Guo, *Fellow, IEEE*

***Abstract*—Sparse sensor placement strategies are applied to reconstruct a region's full-state data conditioned to a limited number of sensors, particularly crucial to ocean monitoring systems. In maritime systems, existing sparse sensor placement methods consider the reconstruction error of data or rely on specific requirements. Considering how sensors acquire essential information for monitoring systems, the utilization of entropy from information theory becomes quite interesting. In this article, we show that entropy measurements on different quantities of information are sensitive to indicate the border areas, thus requiring a balance between the number of sensors needed and the amount of information collected by them in coastal areas. Due to such, we propose (i) a novel sparse sensor placement strategy based on entropy, where the entropy measurements in temporal dimension are utilized for sample selection, so portions of samples selected are utilized for training data, significantly improving the training efficiency without sacrificing accuracy of subsequent data reconstruction. In the proposed strategy, (ii) we use orthogonal triangle decomposition from linear algebra, where a low-cost sensor is employed as a pivot. In terms of spatial dimension, the entropy of each location is adopted as entropy weight to reconstruct full-state data. Additionally, (iii) the strategy employs a greedy algorithm of weighted column pivoting for the orthogonal triangle decomposition, which is designed to suit yet effectively seek additional information and minimal reconstruction error in each iteration processing step. Experimental results using Sea Surface Temperature (SST) data show that the proposed strategy outperforms existing methods, acquiring more information, ensuring higher efficiency, and reducing costs while minimizing reconstruction errors.**

***Index Terms*—Sensor placement, ocean monitoring, entropy, data reconstruction, orthogonal triangle decomposition.**

Manuscript received March 8, 2023; accepted May 23, 2023. Date of publication ; date of current version . This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 52071200, 52201401, 52201403, and 52102397), in part by the National Key Research and Development Program (Grant No. 2021YFC2801002), in part by the Shanghai Committee of Science and Technology, China (Grant No. 23010502000), in part by the China Postdoctoral Science Foundation (Grant No. 2022M712027), in part by the Shanghai Post-doctoral Excellence Program (Grant No. 2022767).

Qiannan Zhang, Huafeng Wu, Xiaojun Mei, and Dezhi Han are with the Shanghai Maritime University, China (e-mail: zhangqiannan02@stu.shmtu.edu.cn; hfwu@shmtu.edu.cn; xjmei@shmtu.edu.cn; dzhan@shmtu.edu.cn).

Mario Donato Marino is with Leeds Beckett University, UK (e-mail: M.D.Marino@leedsbeckett.ac.uk).

Kuan-Ching Li is with the Department of Computer Science and Information Engineering, Providence University, Taiwan (e-mail: kuancli@pu.edu.tw).

Song Guo is with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong (e-mail: song.guo@polyu.edu.hk).

(Corresponding Authors: Kuan-Ching Li, Huafeng Wu)

(Co-first Authors: Qiannan Zhang, Xiaojun Mei and Dezhi Han)

I. Introduction

sensors are used in a wide range of applications today, such as intelligent transportation, distributed data security transmission, etc [1], [55]. Many research technologies cannot do without sensors [3], [4].However, different locations of sensors have different impacts; it is well known that different placement of the sensors in the ocean observation network directly affects the monitoring efficiency and the costs of deployment and maintenance, so selecting appropriate sensor locations is an essential task [5]. Still, a suitable sensor placement strategy is crucial to accurately reconstruct the global atmosphere and hydrologic data with a limited number of sensors [6].

Nowadays, the data collection process in effect for the global ocean is implemented through the oceanographic buoys [7] or Unmanned Aerial Vehicles (UAVs) [9], equipped with different sensors. The more locations sensors place, the more data availability, and accuracy are achieved.Monitoring systems are needed here, and monitoring technology has always been highly valued and widely used [8]. However, due to the spatial and temporal sparseness of marine monitoring data [5] and the cost constraints, the full-state measurements are not feasible because monitoring locations of the entire ocean is impractical. Mainly, longer offshore distance means higher deployment and maintenance costs and signifies that the more limited network resources, the more limited the use of monitoring ocean sensors. Consequently, an efficient sparse deployment strategy of sensors based on the monitoring requirements is needed, which potentially well balances the data reconstruction errors for the full-state measurements while keeping the monitoring system costs at lower levels [10]. Herein to summarize, the sensor placement problem in the ocean means placing the buoys or UAVs to obtain the collected data effectively.

To formalize the sparse sensor placement problem based on ocean data, k sensors are selected in a state space with n sampling points (corresponding to the locations of all sensors) aiming to achieve optimal full state reconstruction, i.e., to sample yet obtain all required information. Important research [11] has concluded that this is an NP-hard problem as the number of combination possibilities is as much as $\binom{n}{k} =$

$\frac{n!}{(n-k)!k!}$: since sensor locations should be precisely selective from global coordinates, the total amount of calculations is highly demanding due to the ocean's vastness; that is, larger k, much larger the number of combination possibilities.

Many sensor deployment strategies have been proposed to obtain the optimal locations automatically and the number of sensors. For example, compressed sensing [12] and information entropy-based on information theory [13], [14], classification based on dimensionality reduction [15], [16], among others. Nevertheless, dimensional reduction only takes account of the intrinsic feature of data, while compressive sensing and information entropy cannot balance either accuracy or information. Inspired by the previous deployment methods mentioned, it is crucial to consider the sparse reconstruction errors and obtain information about the monitoring system using measured data characteristics.

Different from compressive sensing [12] that considers the reconstruction error, the complex calculation of conditional information entropy in modal identification [13], and the hybrid information-entropy approach [14] for identifying leakage water distribution networks, we propose a novel optimal sparse sensor placement strategy where sparse sensing theory is combined with entropy to process ocean temperature data. Using entropy in temporal sampling can better characterize areas with significant temperature gap changes, thus lowering costs for deployment and maintenance. The proposed strategy advances the state of the art in the area of sensors placement via the following contributions:

- The entropy of temporal dimension measurements is utilized for sampling selection, where partial samples are used to train data, significantly improving training efficiency without sacrificing accuracy. The use of entropy aims to sort the samples and improve the training efficiency, thus better characterizing the boundaries of the ocean and areas with noticeable changes.
- A sparse sensor placement based on the optimal information entropy is used to find the optimal sensor placement strategy. The entropy of each location in the spatial dimension is adopted as the entropy weight for the matrix orthogonal triangle decomposition (also known as QR decomposition). The weighted singular values of the determinant in QR pivoting are shown to be submodular, so a greedy algorithm of entropy-weighted QR column pivoting is proposed, suitable and effective to seek more information and minimal reconstruction error in each iteration.
- The performance of the sparse sensor placement based on the entropy strategy is analyzed in terms of information acquisition efficiency, reconstruction error, and deployment costs for training the Sea Surface Temperature (SST) data and analyzing mechanisms designed to excel their capabilities.

The remainder of this article is organized as follows. Section II presents a review of related works, the proposed methodology and corresponding algorithm are depicted in Section III, the experimental results and analysis are shown in Section IV, and finally, concluding remarks and future directions are presented in Section V.

II. Related work and Background

Sensor placement is significant for effective data acquisition in a monitoring system, which concretely affects data collection quality. However, data collection has always been one of the focuses of various research field [17].For such, the sensor placement problem is detailed in this section, and existing solutions are also presented.

Padula et al. [18] defined the sensor placement problem as focused on a set of n possible locations and targets to locate a subset of k locations ($k \ll n$) with the best performance, depending on the specific problem to be solved and the application scenario, such as accuracy of information reconstruction [11], costs [19], [20], utilization of network resources [21], [22], the capability of events detection [14], [23], and others. The purpose is to determine the location and quantity of sensors in the target area that meets a set of given constraints (such as costs, lifetime, and connectivity), making the target optimal.

Whilst the next subsections approach the background involved in this investigation, they discuss previous related work.

A. Data-driven Sparse Sampling

Compressed or Sparse Sampling is a technique that aims to mitigate the pros and cons of more sampling by reducing data reconstruction errors and employing fewer sensors (to reduce deployment costs). As previously stated, Compressive Sensing (CS) technology [12] is a well-known technique that recovers the information of a signal $x \in \mathbb{R}^n$ with significantly fewer number of samples compared with Nyquist sampling theory under the condition of x is k-sparse, and it can be expressed by the product of universal transform basis $\psi \in \mathbb{R}^{n\times n}$ and sparse vector $s \in \mathbb{R}^n$ (where k values are non-zero). Which means that the measurement of a system is treated as a high dimensional state representation, and it could be approximatively recovered by a low-dimensional measurement which is given by some selected sensors [15], [24].

In fact, recovering a high-dimensional state based on CS with random measurements on a universal basis can be suitable for situations where the content of the sampled signal is unknown. For example, for a physical system, if the type of signal is known, it is possible to optimize sensor placement by tailored strategies for specific targets. Among them, the Singular Value Decomposition (SVD) is an efficient dimensional reduction method for sparse representation of nature-based phenomena measurements such as from ocean sensors [25]. Therein, SVD explores the dominant low-dimensional modes of high-dimensional data based on proper orthogonal decomposition (POD) [26], [27], which also have been employed to guide the placement and to reconstruct data collected from of a limited number of ocean-placed velocity and temperature sensors [28]. To approach the reconstruction measurements, the ℓ_2 optimal POD is adopted to represent the low-rank structure of the system and the ℓ_1 sparse sampling method in CS theory is combined to correct the bifurcation regime from noisy measurements in [29]. In subsequent studies, B. W.

Brunton et al. [30] proposed an algorithm to solve sparse sensor placement optimization for classification (SSPOC), where the sparsity-promoting ℓ_1 minimization is adopted to find the sparse sensor locations that correspond to the fewest nonzero entries of the sparse measurement vector to reconstruct the measurement vector of high-dimensional space in feature space.

Sparse sensor placement based on POD has also been developed and discussed by other researches [11], [31], and one of the disadvantages of this technique is to achieve the target of modal approximation with a full state reconstruction. However, due to the fact that the number of POD modes are increased, the presence of more modes would lead to an over-fitting and lower level of reconstruction (uninformative) considering a wide scale of SST. Therefore, in this case, sparse sensor placement based on QR columns pivot is employed to better reconstruct the measurements of the entire system by filtering uninformative features. QR pivot is a fast and efficient method to find near-optimal data-driven POD basis for dimensionality reduction [32]. However, it only focuses on reconstruction errors of the data without considering the obtained information compared, which is the focus of the novel sparse sensor placement strategy proposed here in this report.

The previous sparse sensor placement methods based on CS are data-driven and likely leverage the intrinsic values of the measured data. Likewise, the impact of the high-level information features of the data on sensor placement is not considered as entropy in our approach.

B. *Entropy Maximization*

For real monitoring systems, it is necessary to consider the amount of knowledge/information that can be acquired. In general, the amount of information obtained is measured by the entropy of information theory. Shewry et al. [33] first proposed maximum entropy sampling and applied it to the human population to find an optimal resolution size of potential observation sites. The sampled population measurement had the maximum quantity of variability conditioned on the unsampled one, itself with minimum variability. Similarly, Ramakrishnan et al. [34] have proposed an active data mining mechanism for qualitative analysis of spatial data sets that utilize entropy-based functions defined on spatial aggregation in order to optimize sample selection and reduce uncertainty in the process of selections of targets of these geographical locations. In the case of the previous study, the neighborhoods around each location were considered to define its distribution and scale the variance-based design criterion, which could instinctively distinguish the amount of high-level information in the acquired data. These methods mainly focus on local data variability without considering the state reconstruction of the overall data, which is considered in this paper.

For ocean monitoring, maximization of mutual information was utilized to plan the navigation waypoints of Autonomous Underwater Vehicles (AUVs), which could maximize the information-gain and save time and energy when collecting ocean data [35]. To reconstruct the distribution of density for Antarctic krill, maximum-entropy method was iterative and combined with the Bayesian approach of evaluating the posterior probability of candidate solutions, which had proved to be powerful in biological systems [36]. The maximum entropy method also was applied in reconstructing the ocean by using radial basis functions [37], and this method performed better than the least squared method that was subject to appropriate constraint information [38]. Due to the limitation of these application scenarios, the methods do not take into account the entropy representation of ocean boundaries in the process of data reconstruction.

Entropy-based maximization methods also appear in some particular monitoring environments. For sensor placement, Christodoulou S.E. et al. [39] considered the sensor placement optimization problem as an entropy maximization problem, in which the number and locations of sensors are placed by maximizing the knowledge/information. Furthermore, sensor placement was carried out by maximizing all possible sensor sets of conditional information entropy indexes [13]. Another example of entropy-based methods can be found in leak detections, where a pressure sensor placement method based on the value of information and transformation entropy was proposed to optimize the entire decision space [14].

One of the disadvantages of the entropy criterion is not to consider the predicted quality of the placement of sensors in predetermined locations: traditional entropy prefers to place the sensor in the area (boarders) of interest in the actual monitoring of the particular physical phenomena. These methods either do not require data reconstruction in their applicable scenarios, or only consider the information properties of the data without considering the reconstruction quality, which is one of the aspects considered in our approach.

There is a requirement to place a sensor, considering both the reconstruction error and the knowledge/information to be acquired to optimize the monitor efficiency. To measure the impact of information, the entropy weight method is adopted to investigate the weight of each index, which is derived from the statistical information of measured data [40]. In different applications, the index could indicate different features, conditions, influences, types, or other factors of analyzed objects, such as the ship types and regions in evaluating marine accidents [41], different influence parameters, for example, the ones associated with tidal flat resources [42], or even for analyzing eco-environment vulnerability [43].

Notably, the entropy criterion tends to deploy sensors near the boundary of the monitoring area, and this method shows significant variability. Especially for ocean monitoring, the ocean boundary is undoubtedly an area with relatively dense human activities and lower deployment and maintenance costs. Due to the high density, the monitoring of this area refers to the entire monitoring data distribution valuable, because the adequate information obtained by monitoring this particular area will affect the application of more data to be monitored. The entropy weight method only considers the influence of the amount of information on selecting certain indexes.

Because of the entropy criterion has the property to deploy sensors near the boundary of monitoring area and the QR decomposition has been proved to be able to reconstruct data accurately, the strategy proposed here takes into account the accuracy of data reconstruction and obtain more information at the same time, as well as the deploy and maintenance costs.

Next, we describe in section III the proposed strategy to leverage the effect of entropy when placing sensors for ocean monitoring.

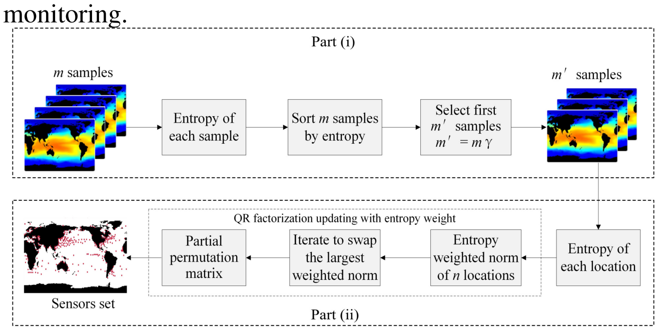


Fig. 1. Flow diagram of the sparse sensor placement based on entropy method.

III. Proposed Strategy

This section introduces the proposed novel sparse sensor placement strategy based on the entropy method and data reconstruction for ocean monitoring. We aim to show that the greedy algorithm is suitable for finding more information in each iteration of the weighted determinant in the sparse sensor placement problem and also that the complexity (order) can minimize costs.

Terminologies used throughout the text are depicted in Table I.

TABLE I
Terminology and Notation

Notation	Description
m	Number of measurements
n	Number of candidate sensors, corresponding to locations
γ	Predefined rates of selected samples
$m^{'}$	Number of measurements with higher entropy $m^{'} = m\gamma$
p	Number of selected sensors
J	Index set of selected sensors $\lvert J \rvert = p$
$Er_{\{J\}}$	Reconstruction error of index set J
e_j	The entropy of j-th sensor, $j = 1, 2, \ldots, n$
ω_j	The entropy weight of j-th sensor
$En_{\{J\}}$	The entropy of selected sensors in the index set J

A. Organization of the Proposed Strategy

Fig. 1 depicts the two composing parts of the proposed strategy as follows:

- In part (i), the selection of the samples is performed based on information entropy in the temporal dimension of measurement data. Therein, the entropy method evaluates each sample's information for all samples and then sorts the samples by the evaluated entropy values. After that, the most valuable measurements which account for γ in the total number of samples are selected as training data. In this paper, the entropy of each sample in SST [44] is calculated in temporal dimension for measurements of the data selection (as further explained and depicted in Eq. 1). Next, the samples are ordered by the calculated entropy values and utilized the first $m^{'}$ samples most informative in the temporal dimension to improve the training efficiency, as in [45]. That is, $m^{'} = m\gamma$. Herein, m is the total number of samples in the temporal dimension, γ is the proportionality coefficient.
- In part (ii), the locations of the sensor are selected by QR factorization under the effect of entropy weight for all candidate locations in the spatial dimension of measurement data. Therein, the entropy of each location is obtained from the selected samples in the first part. Then, QR factorization is carried out to update each step of the decomposition procedure in which the entropy is applied to select the largest entropy-weighted norm of candidate locations so as to obtain the partial permutation matrix, which contains the locations of selected sensors. Each of the blocks in Fig1 is described further.

We start with the description of the entropy sensor placement method. Therein, the entropy method is introduced firstly; the theory of sparse sensor placement under entropy weight is introduced secondly; the detailed algorithm is given finally.

B. Entropy method for sensor placement

The entropy of evaluated objects in temporal dimension and spatial dimension can be calculated in the same way but separately. The calculation procedure of candidate locations' entropy in spatial dimensions is taken as an example as follows:

For sparse sensor placement in this work, we define that the entropy of selected sensors in $J(J = \{J_1, J_2, \ldots, J_p\} \subset \{1, \ldots, n\})$ is represented as $En_{\{J\}} = [e_{J_1}, e_{J_2}, \ldots, e_{J_p}]$ and the corresponding entropy weight as $\omega_{\{J\}} = [\omega_{J_1}, \omega_{J_2}, \ldots, \omega_{J_p}]$, wherein, according to [41], [46]–[48], e_j can be expressed as:

$$e_j = -k\sum_{i=1}^{m} p_{ij} ln(p_{ij}) \tag{1}$$

And we have:

$$\omega_j = \frac{1 - e_j}{\sum\limits_{j=1}^{n} (1 - e_j)} \tag{2}$$

wherein, $j = J_1, J_2, \ldots, J_p$. ω_j is the entropy of each evaluation object (the selected sensor in this paper), also called the entropy weight method that is practical for evaluating the importance of real monitor objects [41], [46]–[48].

Generally, $k = 1/ln(m)$ and m is the number of evaluating objects (the selected samples or sensors in this paper). For the positive indicator, the larger the value of e_j is, the more significant the effect of the j-th indicator for the results. The smaller the indicator entropy value, the greater the degree of data dispersion and therefore, the more significant the impact on the comprehensive evaluation of the indicator.

Then, the next step is to calculate the proportion equation p_{ij} of the i-th sample of the evaluation object under the j-th indicator as follows:

$$p_{ij} = \frac{z_{ij}}{\sum\limits_{i=1}^{m} z_{ij}} \tag{3}$$

For each indicator, after deleting null values in the original sparse samples, normalization of the value of indicators can be done in two ways:

- Positive indicator:

$$z_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \tag{4}$$

- Or negative indicator:

$$z_{ij} = \frac{\max(x_j) - x_{ij}}{\max(x_j) - \min(x_j)} \tag{5}$$

wherein, z_{ij} is the normalized value of the i-th evaluating object on the j-th indicator. $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$.

According to Caselton et al. investigation [49], monitoring data of different locations in a natural environment can be treated as a multivariate normal distribution.The entropy of evaluated objects has the important feature of additivity when the evaluated objects are independent variables in the normal distribution. As a result, the total system entropy of all selected sensors in J can be calculated as the sum of each entropy in $En_{\{J\}}$, as:

$$En_{\{J\}} = \sum_{r=1}^{p} e_{J_{(r)}}, J(r) \in [1, n] \tag{6}$$

Analogously, the greater the entropy weight, the greater the overall impact of the selected sensor is. In this article, the entropy of sensors in the index set J directly affects the knowledge to be acquired, as explained in the following subsections.

If it is in the sample selection case, the entropy of samples in temporal dimension are calculated by converting the index j to temporal dimension, in which $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$. In addition, $j = J_1, J_2, \ldots, J_{m'} \subset \{1, \ldots, m\}$ and $k = 1/ln(n)$.

C. Sparse sensor placement based on entropy method

The novel proposed method leverages the effect of entropy to establish the sparse sensor placement strategy.

In this paper, the locations of the SST data [44] are indicated by entropy weight for the training process of the QR-column pivot for sparse data reconstruction. Specifically in the method, each iteration of QR decomposition depends on maximizing the determinant and entropy value together, which is a variation of traditional D-optimality. And, importantly, the total entropy value of the system with selected sensors is also determined and analyzed.

Generally, the more sensors are selected by QR factorization, the lower the reconstruction error could be achieved, so more knowledge/information can be implemented. Nevertheless, the higher the cost is spawned in equipment, maintenance, communication, and others. Therefore, when the number of selected sensors is given, the locations of QR decomposition for the data matrix are influenced by the entropy weight method in this part to obtain more knowledge/information.

The entropy of each location can index the variation of measurement data that corresponds to the size of information. Although there are often significant variations at regional boundaries due to the coastal region with much lower cost, the entropy method is used to influence the process of selecting sensor positions for QR decomposition column transformation, retrieving more information at a lower cost.

The detailed theory of the proposed method is furtherly explained in the following description, and the corresponding algorithm is given in the following section.

The deployment of the sparse sensor in this research is based on the principles of sparse sampling [12], given as follows: For a system with multiple samples, the measurement set can be represented as a data matrix $X = [x_1 x_2 \ldots x_n]^T$, $x_i \in \mathbb{R}^n$ is corresponding to all sensors data of i-th sample, $i \in [1, m]$. According to Manohar et al. [11], sparse sensor placement means to select p ($p \ll n$) sensors from n candidate sensors in order to reconstruct the measurements of the full-state vector from the actual state vector. Generally, the sparse sensor placement problem seeks the index set J so that the reconstruction error $Er_{\{J\}}$ can be minimized:

$$\begin{aligned} \underset{\{J\}}{\arg\min} \quad & Er_{\{J\}} \\ & s.t. |J| = p \end{aligned} \tag{7}$$

The reconstruction error of selected sensors in J can be represented as:

$$Er_{\{J\}} = \frac{\|X - \hat{X}_J\|_F}{\|X\|_F} \tag{8}$$

wherein $\|\cdot\|_F$ is the Frobenius norm. $\hat{X}_J$ is estimated reduced measurements and $\hat{X}_J = C_J \Psi_{\cdot,J} \hat{A}$, $C_J \in \mathbb{R}^{n\times p}$ is the select matrix for measurements X, $\Psi_{\cdot,J} \in \mathbb{R}^{n\times p}$ is the reduced basis, $A \in \mathbb{R}^{p\times m}$ is the coefficients matrix of X in the constructed modal basis $\Psi_{\cdot,J}$ under the decomposition of $X = \Psi_{\cdot,J} A$. Additionally, according to the Eckart Young theorem [50], the matrix X can be approximately determined by the first p columns of U(U is the left $m \times m$ unitary matrix of singular value decomposition of the matrix X). Thus, the basis mode $\Psi_{\cdot,J}$ for reduced-order modeling is taken as: $\Psi_{\cdot,J} = U_{:,1:p}$. When X is unknown, the estimated coefficients of the matrix $\hat{A}$ can be set as:

$$\hat{A} = (C_J \Psi_{\cdot,J})^{\dagger} X = \Theta^{\dagger} X \tag{9}$$

when let $\Theta = C_J \Psi_{\cdot,J}$, wherein $(\cdot)^{\dagger}$ is the Moore-Penrose pseudo-inverse.

To make $Er_{\{J\}}$ of Eq. (8) minimized, it means to make $\|X - \Theta\Theta^{\dagger} X\|_F$ minimized. When $X - \Theta\Theta^{\dagger} X = 0$, we have $XX^{\dagger} = \Theta\Theta^{\dagger}$. Let $M_J = \Theta^T \Theta$ so that Eq. (7) can be solved by calculating the spectral radius as in the report [11]:

$$\begin{aligned} J_* &= \underset{J,|J|=p}{\arg\max} |\det M_J| = \underset{J,|J|=p}{\arg\max} \prod_j |\lambda_j(M_J)| \\ &= \underset{J,|J|=p}{\arg\max} \prod_j \sigma_j(M_J) \end{aligned} \tag{10}$$

Therefore, J_* can be acquired by data matrix Y. QR factorization with column pivoting of $\Psi_{\cdot,J}$ based on approximate greedy strategy [32] is proved efficiently to solve Eq. (10) and interpreted later considering the entropy of each column.

Considering the entropy, the sparse sensor placement problem in Eq. (10) can be re-defined as:

$$J_* = \underset{J,|J|=p}{\arg\max} \prod_j \sigma_j(M_J)\omega_j \tag{11}$$

where the reconstruction contribution of j-th sensor is the principal factor, but if entropy weight is more significant, the higher the possibility it is selected, and thus, more knowledge can be obtained.

According to the D-optimal experiment design [51], the column selection can be given by first p weighted singular values of $\Theta\Theta^T$ to increase the $\det M_J$. The determinant maximizing subset J is submodular according to Clark et al. [10], which means that it follows the nature of diminishing marginal returns. Referring to the sub-modular property of entropy of random variables [52], we have the entropy weighted of selected sensors by Eq. (11), obeying the diminishing returns property as Theorem 1.

Theorem 1. *For submodular function f_1,f_2 and $f_1 \geq 0$,$f_2 \geq 0$, then the function $f_1 f_2$ is also submodular.*

Proof. According to the definition of the submodular function, for sets $A \subseteq B \subseteq V$ and an element $e \in V - B$, we have:

$$f_1(A \cup \{e\}) - f_1(A) \geq f_1(B \cup \{e\}) - f_1(B) \tag{12}$$

$$f_2(A \cup \{e\}) - f_2(A) \geq f_2(B \cup \{e\}) - f_2(B) \tag{13}$$

Given $f_1 \geq 0$, $f_2 \geq 0$, multiplying both sides of Eq. (12) by $f_2(A \cup \{e\})$ and multiply both sides of Eq. (13) by $f_1(B \cup \{e\})$ as:

$$\begin{aligned} & f_1(A \cup \{e\}) f_2(A \cup \{e\}) - f_1(A) f_2(A \cup \{e\}) \\ & \geq f_1(B \cup \{e\}) f_2(A \cup \{e\}) - f_1(B) f_2(A \cup \{e\}) \end{aligned} \tag{14}$$

$$\begin{aligned} & f_2(A \cup \{e\}) f_1(B \cup \{e\}) - f_2(A) f_1(B \cup \{e\}) \\ & \geq f_2(B \cup \{e\}) f_1(B \cup \{e\}) - f_2(B) f_1(B \cup \{e\}) \end{aligned} \tag{15}$$

The two sides of Eq. (14) and Eq. (15) are correspondingly added, then adjust the added items to obtain:

$$\begin{aligned} & f_1(A \cup \{e\}) f_2(A \cup \{e\}) - f_2(B \cup \{e\}) f_1(B \cup \{e\}) \\ & \geq f_1(A) f_2(A \cup \{e\}) + f_1(B \cup \{e\}) f_2(A \cup \{e\}) - \\ & f_1(B) f_2(A \cup \{e\}) - f_2(A \cup \{e\}) f_1(B \cup \{e\}) + \\ & f_2(A) f_1(B \cup \{e\}) - f_2(B) f_1(B \cup \{e\}) \\ & = f_1(B \cup \{e\})(f_2(A) - f_2(B)) + \\ & f_2(A \cup \{e\})(f_1(A) - f_1(B)) \end{aligned} \tag{16}$$

According to the nature of the submodular function, we have $f_1(B \cup \{e\}) \geq f_1(B)$, $f_2(A \cup \{e\}) \geq f_2(A)$, the Eq. (16) can be organized as:

$$\begin{aligned} & f_1(A \cup \{e\}) f_2(A \cup \{e\}) - f_2(B \cup \{e\}) f_1(B \cup \{e\}) \\ & \geq f_1(B)(f_2(A) - f_2(B)) + f_2(A)(f_1(A) - f_1(B)) \\ & = f_1(A) f_2(A) - f_1(B) f_2(B) \end{aligned} \tag{17}$$

Next, we have:

$$\begin{aligned} & f_1(A \cup \{e\}) f_2(A \cup \{e\}) - f_1(A) f_2(A) \\ & \geq f_1(B \cup \{e\}) f_2(B \cup \{e\}) - f_1(B) f_2(B) \end{aligned} \tag{18}$$

Formula 18 satisfies the definition of the submodular function, so that the function $f_1 f_2$ is also submodular, and the proposition is proved. □

According to Theorem 1, entropy-weighted determinant maximizing is also submodular because both the entropy and singular values are non-negative and submodular. Therefore, a greedy algorithm is suitable for finding more information in each iteration of the weighted determinant of the Eq. (11). The detailed greedy algorithm based on data reconstruction and entropy is described in next section.

D. Greedy algorithm based on data reconstruction and entropy method

In the greedy algorithm, QR factorization with column pivoting as a data reconstruction method is adopted here to find p sensors (pivots) so that base modes $\Psi_{\cdot,J}$ under entropy weighted could be best sampled. The procedure is described as follows. For a given data matrix $X \in \mathbb{R}^{m \times n}$, $m \leq n$, QR factorization with column pivoting is represented as:

$$XP = QR \tag{19}$$

where $P \in \mathbb{R}^{n \times n}$ is a column permutation matrix for matrix X, and $Q \in \mathbb{R}^{m \times m}$ is unitary, $R \in \mathbb{R}^{m \times n}$ is upper triangular. Among them, each column corresponds to a candidate location of a sensor. Householder transformation is performed to reflectively permutate the columns of matrix Y and initialized permutation matrix $P_0 = I_{n \times n}$. Divide the matrix X into columns: $X = (x_1, x_2, \ldots, x_n)$. Then the column x_j with the largest entropy weighted norm $\|x_j\|\omega_j = \sigma_j \omega_j$ will be selected to be swapped with the first column x_1 and the current permutation matrix is:

$$P_1 = (e_j \quad e_2 \quad \ldots \quad e_{j-1} \quad e_1 \quad e_{j+1} \quad \ldots \quad e_n) \tag{20}$$

where j is the first location index of selected sensors. The Householder reflection vector of column x_j is $H_1 x_j$, which is on the hyperplane $u = x_j - \|x_j\| e_j$. Therein:

$$H_1 x_j = \begin{pmatrix} \|x_j\| \\ 0 \\ \vdots \end{pmatrix} \tag{21}$$

Take $w = \frac{u}{\|u\|}$ and H_1 are given by:

$$H_1 = I_{n \times n} - 2ww^* \tag{22}$$

We can obtain the Householder transformed matrix of XP_1:

$$H_1 X P_1 = \begin{pmatrix} \|x_j\| & * & * & \ldots \\ 0 & & & \\ 0 & & X^{(1)} & \\ \vdots & & & \ddots \end{pmatrix} \tag{23}$$

where $X^{(i)}$ is a submatrix of i-th Householder transformation, $i = 1, 2, \ldots, p$. In each transformation, the column with the largest entropy-weighted norm will be reflected to the first column of the submatrix $X^{(i)}$. For the Householder transformed matrix of XP_2 and $P_2 = (e_j \quad e_{j^{(2)}} \quad \ldots \quad e_{j^{(2)}-1} \quad e_2 \quad e_{j^{(2)}+1} \quad \ldots \quad e_n)$, steps above are repeated:

$$H_2 H_1 X P_2 = \begin{pmatrix} \|x_j\| & * & * & \ldots \\ 0 & \|x^{(1)}_{j^{(2)}}\| & * & \ldots \\ 0 & 0 & X^{(2)} & \\ \vdots & \vdots & & \ddots \end{pmatrix} \tag{24}$$

The rest transformation can be done in the same manner by leveraging the Householder reflection until p the order Householder matrix H_{p-1} is obtained. That is, the decomposition is achieved as:

$$H_{p-1} \ldots H_2 H_1 X P_p = R \tag{25}$$

where P_p is the column permutation matrix P in Eq. (11) and $Q = {H_1}^* {H_2}^* \ldots {H_{p-1}}^*$.

During each iteration, QR factorization updating j^q to maximize entropy weighted $|\det X^{q-1}|$, in which the optimal solution is greedy-approximate. The computational complexity of Algorithm 1 is $O(mnp)$, determined by counting operations of the pseudocode in Algorithm 1.

Algorithm 1 QR pivoting with entropy weight

Input: data matrix $X \in \mathbb{R}^{m \times n}$, entropy weight $\omega \in \mathbb{R}^n$, number of sensors $p \leq m$

Output: Partial permutation matrix P

```
Initial P ⟵ I_{n×n}
for q = 1 to n do
    J_q ⟵ arg max_{j≥q} ||X(q : m, j) · ω_j||
    Swap (P_{:,q}, P_{:,J_q})
    Calculate Householder reflection H_q such that(X_{q,J_q} X_{q+1,J_q} ... X_{m,J_q})^T = (||X_{q:m,J_q}|| 0 ... 0)^T
    X ⟵ diag(I_{(q−1)×(q−1)})XP
end for
return P_{:,1:p}
```

As mentioned above, QR factorization with column pivoting is equivalent to D-optimal experimental design, seeking to maximize the determinant and volume as well the entropy for the measurement matrix. Taking ${C_J}^T = P_{:,1:p}$ and reviewing the previous $\Theta = C_J \Psi_{\cdot,J}$ then combining Eq. (10), the optimal sensor selection for Eq. (11) can be obtained, because rows in $\Psi_{\cdot,J}$ with largest norms (such that the determinant of Θ) correspond to the optimal sensor locations. Then, Algorithm 1 can consider both the reconstruction errors and information simultaneously for the solutions of the sparse sensor placement problem proposed in Eq. (11).

In addition, Algorithm 1 of QR pivoting with entropy weight considers the construction error and acquired information greedily, which has been proved previously to obey the diminishing returns property by Theorem 1. Meanwhile, the entropy of sensor locations can better characterize the ocean boundary described above, which means that the proposed Algorithm 1 prefers to select locations near the coast with more variation of monitoring data and lower costs of deployment and maintenance.

IV. Experimental Evaluation and Results

Sparse sensor placement strategy based on entropy is analyzed in this section in several aspects, including reconstruction error, deployment costs, and information acquisition rate in temporal and particularly in dimensions, by comparing it with QR-decomposition and random selection from Clark et al. [10].

A. Methodology

We have implemented Algorithm 1 in Matlab (Table II). To showcase the effect of the aforementioned sparse sensor placement based on the entropy method, we use the Sea Surface Temperature (SST) dataset to evaluate and present the temperature change [44]. The SST dataset is included in the collected data by the Physical Sciences Laboratory, and the obtained data is high-resolution data (weekly and monthly time resolutions) and located in the NOAA Optimum Interpolation (OI) Sea Surface Temperature (SST) V2 dataset. Downloaded global temperature data here contains 360*180 locations (corresponding to the longitude from 180°E to 180°W and latitude 90°S to 90°N respectively) and 1,600 times measurements (once a week) that are published from 1990 till present.

TABLE II
Experimental Environment

Hardware	Memory	8G
	CPU	Intel(R) Core(TM) i7-6700 @3.4GHz
Software	Programming Language	Matlab
	Operating System	Windows 10

TABLE III
Parameters for Experiment Processing

Parameters	Value
Number of candidate locations in the ocean	44219
Number of total samples	1600
Sample selection ratio γ	0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1
Number of selected samples	1600*γ
Number of training samples	1600*γ*80%
Number of validation samples	1600*γ*20%
Number of selected sensors: p	5, 10, 25, 50, 75, 100, 150, 200, 250, 300
Number of cross-validation	10

Importantly, the measurement and number of locations adopted in this paper follow other well-known reports and their methodologies [10] [32]. Among them, we similarly adopt 44219 locations from the SST dataset are candidate locations of the global sea, wherein the remaining 360*180-44219 locations of data on land are not considered, whilst their entropy default values are set to 0. Table II shows the features of the environment used to develop the algorithm and perform the experiments. Table III lists the main parameters employed for Algorithm 1's processing.

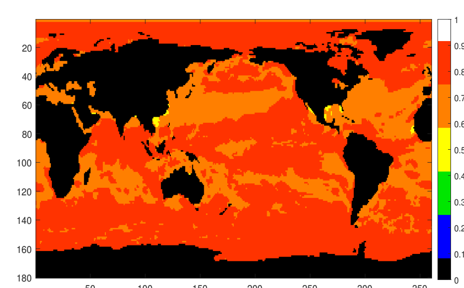


Fig. 2. Entropy of global sea surface temperature

Before the training of Algorithm 1, the entropy of each measurement is calculated first, comprising 44,219 locations of the global sea as the first part in Fig. 1. Then the most valuable measurements selected by γ are taken as training data. At the same time, the entropy weight of each location is also calculated, and the entropy weight values of the global

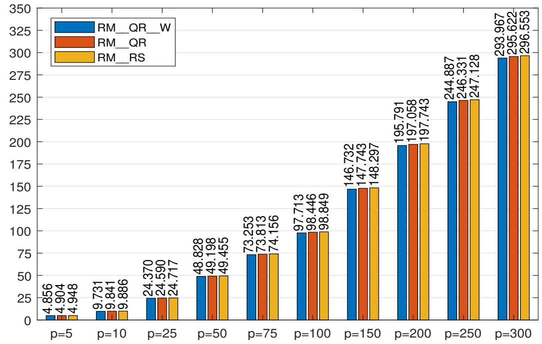


Fig. 3. Total entropy of a different number of sensors.

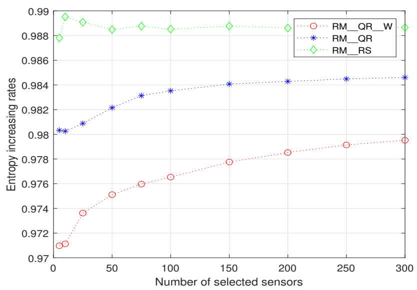


Fig. 4. Entropy increasing rates with a different number of sensors.

sea are displayed in Fig. 2. As per the negative indicator in Eq. (5) and Eq. (2), the entropy weight of each location comprises 1,600 times temperature data, in which the negative indicator adopted in Fig. 2 conveys the lower entropy that represents more information except for the locations of land. If the entropy weight is calculated as the positive indicator in Eq. (4) and Eq. (2), the higher entropy magnitude will represent more information.

It is worth noting that, to validate our measurements, Fig. 2 shows a significant variation of the temperature around the coast near the equator and mid-latitudes, and these temperatures match the distribution of El Niño.

Importantly, in this article, 80% of the samples are selected and applied as training data, while the remaining as verification data. Because random interpolation is superior to extrapolation, the random interpolation method is adopted to select the training data and to cross-validate them.

Once set the reduced-order mode $\Psi_{\cdot,J}$ as a random mode $\Psi_{\cdot,J\_RM} = GX$ as per randomized linear algebra, $G \in \mathbb{R}^{2p\times m}$ and its elements are standard normal distribution. We also set the reduced order mode $\Psi_{\cdot,J}$ as the first p singular vectors of U from the selected training data matrix X for comparison purposes.

B. Analyses of the Sparse sensor placement based on entropy weight method

Sparse sensor placement based on the entropy weight method is carried out after data pre-processing, as per Algorithm 1 in Section III-D. Next, this method is compared with the unweighted QR decomposition and random selection method.

1) Total entropy of different methods:

Fig. 3 shows three different methods to illustrate the different number of sensors for SST data reconstruction and the total information/knowledge varying with the number of sensors and the different methods. In this case, all samples are conducted to be calculated, which means $\gamma = 1$. RM_QR_W represents the sparse sensor placement based on the entropy weight method and data reconstruction in this paper, wherein the random mode basis is used for QR-decomposition, and the entropy weight method is used to affect the procedure of D optimization under QR-decomposition. Correspondingly, the unweighted QR algorithm is expressed as RM_QR within a random mode basis, while RM_RS represents the random selection method based on a random mode basis.

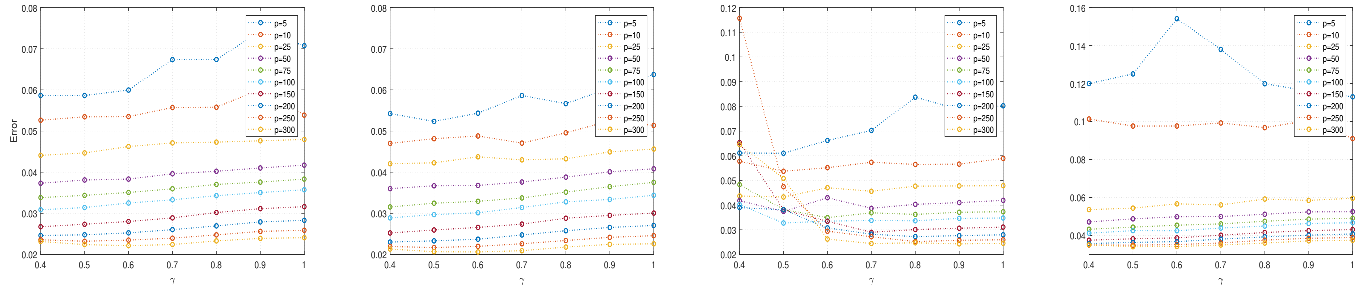
As shown in Fig. 3, the entropy obtained by RM_QR_W is the smallest among the three methods when deploying the different number of sensors. Due to negative indicators, the smaller the entropy, the greater the amount of information. That is, the sensor deployed by the method proposed in this work can obtain the most information compared to the other two methods.

To highlight, Fig. 4 shows the entropy growth rate of different numbers of sensor nodes; that is, the average amount of information that each sensor node obtains when different numbers of sensors are deployed. Similarly, the smaller the entropy, the greater the amount of information. We can see that the method proposed in this work can obtain the most information when deploying a different number of sensors.

2) Effects of different γ on error and entropy:

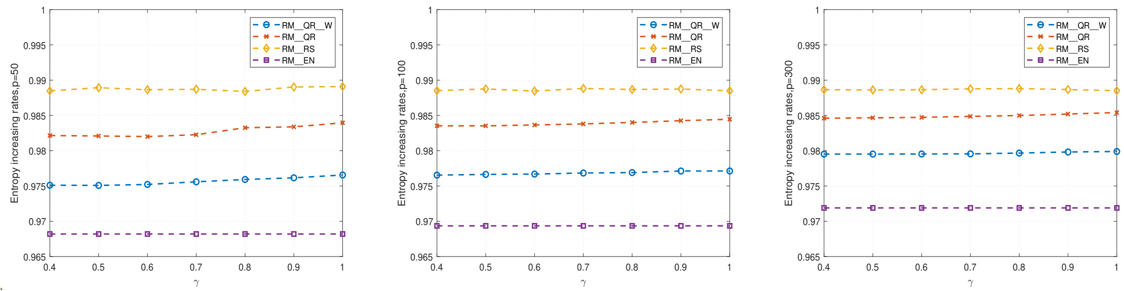
As depicted in Figs. 5a, 5b, 5c and 5d, when γ is different, the reconstruction error varies with the number of selected sensors for the three methods. It has been shown that sample selection based on entropy value can reduce the amount of training data while improving the reconstruction accuracy. In the first case, when the selected number increases, the corresponding reconstruction error decreases, and when the parameter γ increases, the reconstruction errors also increase when the number of sensors is smaller than 200 for RM_QR_W in Fig. 5a and RM_QR in Fig. 5b. This behavior shows that the selected samples with larger entropy values are more conducive to data reconstruction in these figures, which means the smaller γ will bring lower reconstruction error for a smaller number of sensors. In the second case, reconstruction errors also increase as the number of selected samples with smaller entropy values increases with larger γ. However, when the number of sensors is larger (when it reaches 200, 250, and 300, as shown in Figs. 5a and 5b, the value of γ can be 0.5 to 0.7 to reduce the impact on reconstruction errors positively). This indicates that, when the number of sensors selected is large, partial samples with larger entropy can be used to better reconstruct the data without affecting the reconstruction accuracy, significantly reducing the computational complexity of the data training algorithm for selecting sparse sensors. Furthermore, as shown in Fig. 5c, when the number of sensors is small, the reconstruction error is high and unstable for RM_RS. In addition, in the following comparison, we also compare the entropy maximization method, named RM_EN, which is similar to [53] but here the total value of entropy adopts the Eq. (6), to compare the reconstruction error and cost. As shown in Fig. 5d, when the number of sensors is large (p=300), a γ of 0.5 or 0.6 can obtain a lower reconstruction error, while the error value is much larger than the method of RM_QR_W in Fig. 5a and RM_QR in Fig. 5b.

For the relationship between entropy growth rate and γ,



(a) RM_QR_W (b) RM_QR (c) RM_RS (d) RM_EN

Fig. 5. Reconstruction error with different γ.



(a) p=50 (b) p=100 (c) p=300

Fig. 6. Entropy increasing rates with different γ.

we turn attention to Figs. 6a, 6b and 6c. When the number of sensors deployed is different, the entropy growth rates of the four methods are also different. For instance, the entropy increasing rates of RM_EN is the lowest because it only considers maximizing entropy. However, the increasing entropy rates of RM_QR_W is much lower than the other two methods, which shows that RM_QR_W can gain more information for 50, 100, and 300 sensors. The RM_QR method also seeks a better amount of information to a certain extent, although its primary purpose is to focus on the internal characteristics of the data.

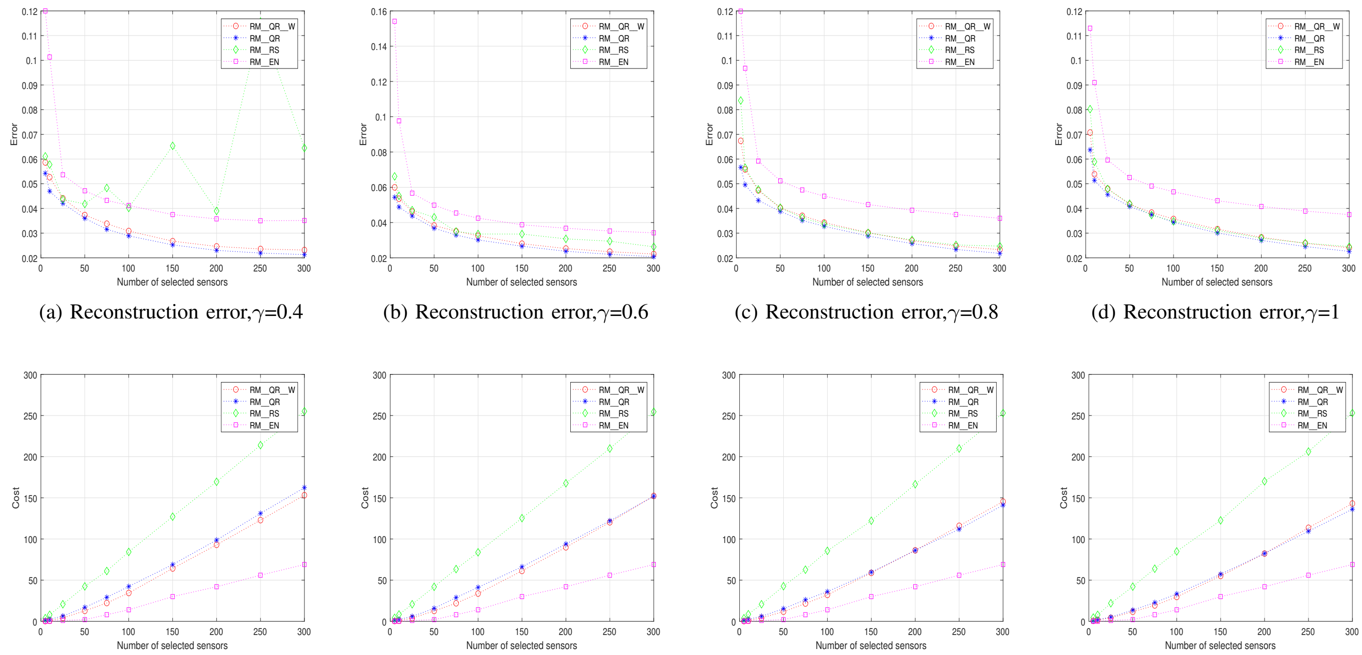
Again, if the value of γ is 0.5 to 0.7, then a relatively good result of the entropy increase rate of RM_QR_W and RM_QR can be obtained, which undoubtedly saves the amount of calculation in data training by Algorithm 1 significantly. Although γ has little effect on the entropy increase rate when the number of sensors is relatively large (p=300), it also shows that the sample screening based on the entropy value can drastically reduce the amount of calculation while obtaining better benefits. Because of the greedy nature of the RM_QR_W and RM_QR, when the number of sensors is small, the entropy increase rate is more prominent. Nevertheless, due to the randomness of the RM_RS method, different γ and different numbers of sensors have almost no influence on the information growth rate.

3) Reconstruction errors and costs for different numbers of sensors:

Reconstruction errors of the three methods are also compared in Figs. 7a, 7b, 7c and 7d. We observe that, the smaller the number of sensors, the larger the reconstruction error for all four methods. Among them, the QR algorithm based on D optimization (RM_QR) has the slightest error because it only considers minimizing the data reconstruction error. The method RM_QR_W sacrifices reconstruction accuracy during the procedure of data reconstruction when seeking the maximum information. In addition, the reconstruction error of the randomly selected method is slightly lower than the method proposed in this work when the number of sensors is large, but much worse than the proposed method when the number of sensors is small. The reconstruction error of RM_EN performs the worst because the reconstruction error is not considered.

Figs. 7e, 7f, 7g and 7h show the costs of all four methods, which gradually increase as the number of sensors increases. However, the cost of RM_RS grows much faster than the other two methods, because the selected sensors are randomly distributed in the global sea, so there are a large number of sensors far away from shore, increasing the total cost. When a different number of sensors is selected, the cost of the proposed method RM_QR_W is relatively close to that of RM_QR. Among them, when γ is small ($\gamma = 0.4$), the costs of the proposed method RM_QR_W are lower than the method RM_QR due to the samples selection operation, as depicted in Fig. 7e. Referring to Fig. 7f to Fig. 7h, the comparison results of cost change slightly when γ is increased. Specifically, when $\gamma = 0.6$ and the number of sensors is 300, the cost of the proposed method RM_QR_W is almost the same as the method RM_QR, although the cost of other numbers of sensors is lower than RM_QR. For the case $\gamma = 0.8$ and 1, when the number of sensors is less than 200, the cost of the proposed method RM_QR_W is lower, which is comparable to the method RM_QR. Due to the sensitivity of the entropy property to ocean boundaries, RM_EN prefers to deploy sensors near ocean boundaries, and thus, the costs are significantly less than other methods.

Overall, it can be seen that the method proposed in this work sacrifices a portion of data reconstruction accuracy when greedily seeking more information, but the reconstruction error is not much different from that of method RM_QR. However, the method proposed in this work has advantages



(a) Reconstruction error,γ=0.4 (b) Reconstruction error,γ=0.6 (c) Reconstruction error,γ=0.8 (d) Reconstruction error,γ=1

(e) Cost,γ=0.4 (f) Cost,γ=0.6 (g) Cost,γ=0.8 (h) Cost,γ=1

Fig. 7. reconstruction error and costs varies with γ and sensors number.

both in the entropy and the cost, especially when γ is small and the number of sensors is small. Additionally, smaller γ means a considerable increase in computational efficiency while reducing the cost of selected sensors. Although RM_EN has the lowest cost, it also has the highest reconstruction error and is unsuitable for deploying sensors in ocean monitoring.

4) Sets of Sensors for placement results:

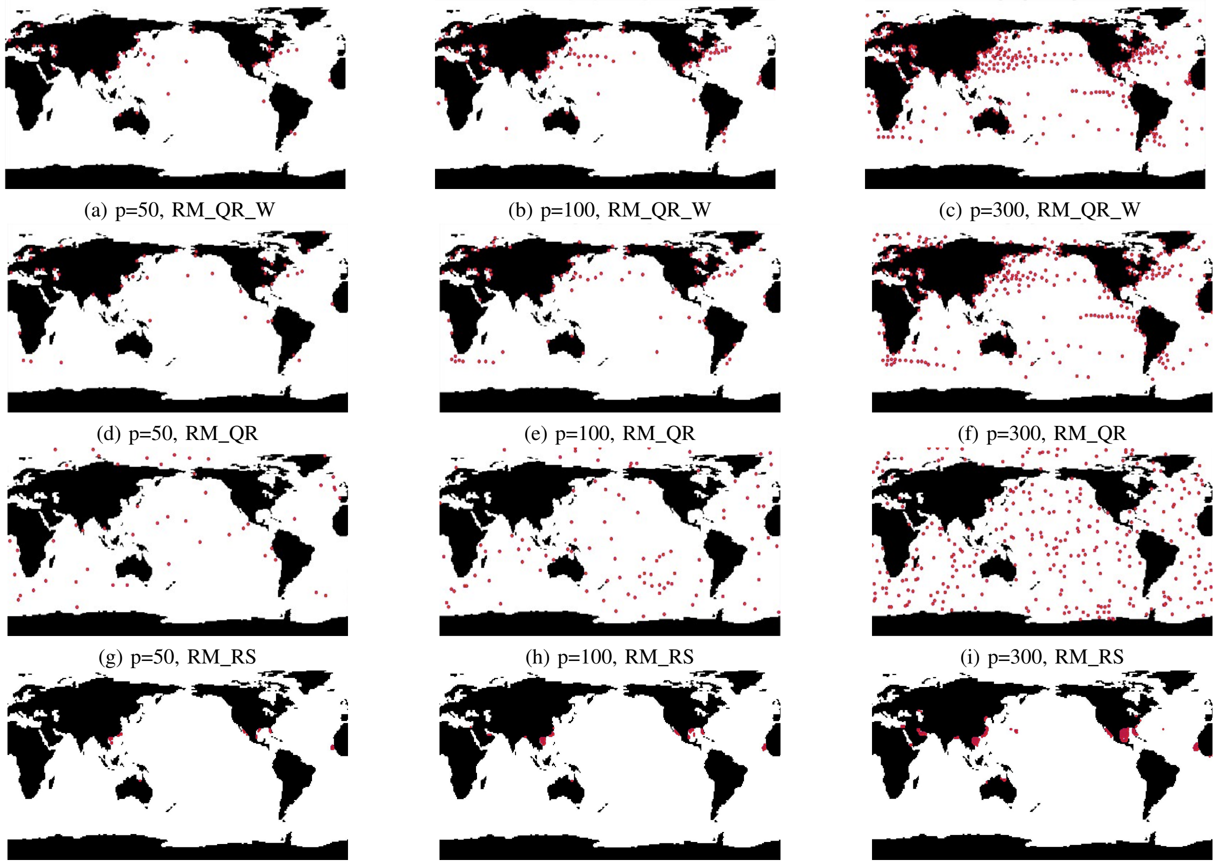
The average sensor locations are shown in Figs. 8a, 8b, 8c, 8d, 8e, 8f, 8g, 8h, 8i, 8j, 8k, and 8l where the number of sensors is set as p=50, 100, and 300. We can see that both the RM_QR_W and RM_QR prefer to place sensors near the coast area and equator, the main area affected by El Niño and the area with more human activities. However, the sensor locations of the random selection method are much more scattered than the two methods. Further, the proposed RM_QR_W is inclined to deploy sensors in areas with drastic temperature changes, which will lead to acquiring more information corresponding to the unweighted reconstruction method. Figs. 8j, 8k and 8l again verifies that RM_EN deploys sensors centrally near the coastline at the equator, which is not conducive to the reconstruction of global ocean temperature data.

The above results show that the strategy proposed in this paper affects the selection of traditional sparse sensors through the value of entropy so as to obtain more information, and can better take into account the reconstruction error and cost at the same time. The proposed strategy not only exploits the sensitivity of entropy to ocean boundary regions, but also makes the deployed sensor locations more in line with the need for monitoring more coastal economic zones. Since all the global ocean locations are used as candidate locations, the strategy in this paper also avoids the defect of the entropy maximization method that deploys sensors in areas with high entropy values.

V. Conclusions and Future Work

In this article, the need for more information on marine monitoring systems is raised and considered. When deploying using sensors, the location where more information can be obtained is preferred to be selected. The entropy from information theory is used to measure information obtained from a region, such as an ocean. Specifically, the entropy method is used to measure the possibility of use of each sample of temporal dimension and each location of the spatial dimension selected. In our novel method, the entropy of each sample is calculated so that samples with the highest entropy are selected to be trained. Second, the entropy weight method is used to influence the solving process of the D optimization problem based on QR decomposition, to greedily select a sensor deployment strategy that considers both the minimal reconstruction errors and more information, wherein the sample selection based on entropy could effectively reduce the amount of data processing without affecting the processing results. At the same time, the entropy can better characterize the region boundary used for sparse sensor placement, so that the selected locations are more in line with the application requirements of ocean monitoring systems. By using ocean surface temperature data to verify the implementation of this proposed strategy, results demonstrate that the strategy can achieve more information while acquiring lower reconstruction errors and deployment costs. Moreover, training selected samples with higher entropy can achieve approximate or better results than training all samples, improving efficiency.

Nevertheless, each sensor must collect multiple environmental parameters for sensors deployed in the ocean. Therefore, it is required to be further resolved how to consider the amount of information and reconstruction errors of different environmental parameters simultaneously. In the future, multiple different types of ocean data that will affect sensor placement strategy will be tackled. For instance, it will be better to consider the influence of different data sets on sensor selection simultaneously, such as the fusion efficiency [54], the reconstruction accuracy, the obtained information, and other factors of multiple data sets, wherein the data categories can include wind, waves, currents, air pressure, visibility,



(a) p=50, RM_QR_W (b) p=100, RM_QR_W (c) p=300, RM_QR_W

(d) p=50, RM_QR (e) p=100, RM_QR (f) p=300, RM_QR

(g) p=50, RM_RS (h) p=100, RM_RS (i) p=300, RM_RS

(j) p=50, RM_EN (k) p=100, RM_EN (l) p=300, RM_EN

Fig. 8. Average sensor locations.

ice distribution, and other parameters that impact the global natural environment and human activities. In addition, sensor placement in the ocean also needs to consider further the influence of network resources [21], [49], sparse data recovery of spatial-temporal dependencies [55] and other factors involved in the sensor placement in ocean monitoring systems. Thereafter, sensor placement under the influence of various factors will be more in line with practical applications for ocean monitoring.

REFERENCES

[1] W. Liang, Y. Huang, J. Xu, and S. Xie, "A distributed data secure transmission scheme in wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 13, no. 4, p. 1550147717705552, 2017.

[2] W. Liang, Y. Li, K. Xie, D. Zhang, K.-C. Li, A. Souri, and K. Li, "Spatial-temporal aware inductive graph neural network for c-its data recovery," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[3] W. Liang, S. Xie, J. Cai, J. Xu, Y. Hu, Y. Xu, and M. Qiu, "Deep neural network security collaborative filtering scheme for service recommendation in intelligent cyber–physical systems," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 22 123–22 132, 2021.

[4] W. Liang, S. Xie, J. Long, K.-C. Li, D. Zhang, and K. Li, "A double puf-based rfid identity authentication protocol in service-centric internet of things environments," *Information Sciences*, vol. 503, pp. 129–147, 2019.

[5] W. Schmidt, D. Raymond, D. Parish, I. G. Ashton, P. I. Miller, C. J. Campos, and J. D. Shutler, "Design and operation of a low-cost and compact autonomous buoy system for use in coastal aquaculture and water quality monitoring," *Aquacult. Eng.*, vol. 80, pp. 28–36, Jan. 2018.

[6] G. Liu and W. Kang, "Idma-based compressed sensing for ocean monitoring information acquisition with sensor networks," *Math. Probl. Eng.*, vol. 2014, pp. 1-13, Jan. 2014.

[7] A. M. Durán-Rosal, C. Hervás-Martínez, A. J. Tallón-Ballesteros, A. Martínez-Estudillo, and S. Salcedo-Sanz, "Massive missing data reconstruction in ocean buoys with evolutionary product unit neural networks," *Ocean Eng.*, vol. 117, pp. 292–301, May. 2016.

[8] K. X. D. Z. S. X. K.-C. L. Yuhui Li, Wei Liang, "Lightnestle: Quick and accurate neural sequential tensor completion via meta learning," in *IEEE INFOCOM 2023*, vol. 1. New York, USA, 2023.

[9] J. Sánchez-García, J. García-Campos, M. Arzamendia, D. G. Reina, S. Toral, and D. Gregor, "A survey on unmanned aerial and aquatic vehicle multi-hop networks: Wireless communications, evaluation tools and applications," *Comput. Commun.*, vol. 119, pp. 43–65, Apr. 2018.

[10] E. Clark, T. Askham, S. L. Brunton, and J. N. Kutz, "Greedy sensor placement with cost constraints," *IEEE Sens. J.*, vol. 19, no. 7, pp. 2642–2656, Dec. 2018.

[11] K. Manohar, B. W. Brunton, J. N. Kutz, and S. L. Brunton, "Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns," *IEEE Control Syst. Mag.*, vol. 38, no. 3, pp. 63–86, May. 2018.

[12] D. L. Donoho, "Compressed sensing," *IEEE T. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[13] X.-Y. Pei, T.-H. Yi, C.-X. Qu, and H.-N. Li, "Conditional information entropy based sensor placement method considering separated model error and measurement noise," *J. Sound Vib.*, vol. 449, pp. 389–404, Mar. 2019.

[14] M. S. Khorshidi, M. R. Nikoo, N. Taravatrooy, M. Sadegh, M. Al-Wardy, and G. A. Al-Rawas, "Pressure sensor placement in water distribution networks for leak detection using a hybrid information-entropy approach," *Inf. Sci.*, vol. 516, pp. 56–71, Apr. 2020.

[15] B. W. Brunton, S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Optimal sensor placement and enhanced sparsity for classification," 2013. [Online]. Available: arXiv:1310.4217

[16] K. Manohar, S. L. Brunton, and J. N. Kutz, "Environment identification in flight using sparse approximation of wing strain," *J Fluids Struct.*, vol. 70, pp. 162–180, Apr. 2017.

[17] W. LIANG, Z. RUAN, M. TANG, and P. LI, "A secure-efficient data collection algorithm based on self-adaptive sensing model in mobile internet of vehicles," *ÊÕ^{2}Ø*, vol. 2, 2016.

[18] S. L. Padula and R. K. Kincaid, "Optimization strategies for sensor

and actuator placement," Langley Research Center, Hampton, VA, USA, Tech. Rep., NASA/TM-1999-209126, Apr. 1999.
[19] B. Tang, F. Guo, B. Cao, M. Tang and K.-C. Li, "Cost-aware Deployment of Microservices for IoT Applications in Mobile Edge Computing Environment," *IEEE Trans. Netw. Service Manag.*, early access, doi: 10.1109/TNSM.2022.3232503.
[20] M. Cui, D. Han, J. Wang, K.-C. Li and C.-C. Chang, "ARFV: An Efficient Shared Data Auditing Scheme Supporting Revocation for Fog-Assisted Vehicular Ad-Hoc Networks," *IEEE Trans. Veh. Technol.*, vol. 69, pp. 15815–15827, Dec. 2020.
[21] S. Khezri, K. Faez, and A. Osmani, "Modified discrete binary pso based sensor placement in wsn networks," in *Int. Conf. CICN*, Bhopal, India, 2010, pp. 200–204.
[22] D. Han, Y. Zhu, D. Li, W. Liang, A. Souri, and K.-C. Li, "A blockchain-based auditable access control system for private data in service-centric IoT environments," *IEEE Trans. Industr. Inform.*, vol. 18, no. 5, pp. 3530-3540, May. 2022.
[23] Q. Tian, D. Han, K.-C. Li, X. Liu, L. Duan and A. Castiglione, "An intrusion detection approach based on improved deep belief network," *Appl. Intell.*,vol. 50, pp. 3162-3178, May. 2020.
[24] E. Kaiser, M. Morzy¨½ski, G. Daviller, J. N. Kutz, B. W., Brunton, and S. L. Brunton, "Sparsity enabled cluster reduced-order models for control," *SIAM J. Appl. Dyn. Syst.*, vol. 352, pp. 388–409, Jan. 2018.
[25] B. Jayaraman, S. Al Mamun, and C. Lu, "Interplay of sensor quantity, placement and system dimension in pod-based sparse reconstruction of fluid flows," *Fluids*, vol. 4, no. 2, p. 109, June. 2019.
[26] G. Berkooz, P. Holmes, and J. L. Lumley, "The proper orthogonal decomposition in the analysis of turbulent flows," *Annu. Rev. Fluid Mech*, vol. 25, no. 1, pp. 539–575, Jan. 1993.
[27] K. Taira, S. L. Brunton, S. T. Dawson, C. W. Rowley, T. Colonius, B. J. McKeon, O. T. Schmidt, S. Gordeyev, V. Theofilis, and L. S. Ukeiley, "Modal analysis of fluid flows: An overview," *Aiaa J.*, vol. 55, no. 12, pp. 4013–4041, Oct. 2017.
[28] B. Yildirim, C. Chryssostomidis, and G. Karniadakis, "Efficient sensor placement for ocean measurements using low-dimensional concepts," *Ocean Model.*, vol. 27, no. 3-4, pp. 160–173, Jan. 2009.
[29] Brunton, Steven L and Tu, Jonathan H and Bright, Ido and Kutz, J Nathan, "Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems," *SIAM J. Appl. Dyn. Syst.*, vol. 13, pp. 1716–1732, Dec. 2013.
[30] B. W. Brunton, S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Sparse sensor placement optimization for classification," *SIAM J. Appl. Math.*, vol. 76, no. 5, pp. 2099–2122, Oct. 2016.
[31] S. Sargsyan, S. L. Brunton, and J. N. Kutz, "Nonlinear model reduction for dynamical systems using sparse sensor locations from learned libraries," *Phys. Rev. E*, vol. 92, no. 3, p. 033304, Sep. 2015.
[32] E. Clark, J. N. Kutz, and S. L. Brunton, "Sensor selection with cost constraints for dynamically relevant bases," *IEEE Sens. J.*, vol. 20, no. 19, pp. 11674–11687, Oct. 2020.
[33] M. C. Shewry and H. P. Wynn, "Maximum entropy sampling," *J. Appl. Stat.*, vol. 14, no. 2, pp. 165–170, 1987.
[34] N. Ramakrishnan, C. Bailey-Kellogg, S. Tadepalli, and V. N. Pandey, "Gaussian processes for active data mining of spatial aggregates," in *Proc. SDM*, SIAM, 2005, pp. 427–438.
[35] K.-C. Ma, L. Liu and G.S. Sukhatme, "An information-driven and disturbance-aware planning method for long-term ocean monitoring," *2016 IEEE/RSJ Int. Conf. IROS*, Daejeon, Korea (South), 2016, pp. 2102-2108.
[36] A. S. Brierley, S. F. Gull, M. H. Wafy, "A Bayesian maximum entropy reconstruction of stock distribution and inference of stock density from line-transect acoustic-survey data," *ICES J. Mar. Sci.*, Vol. 60, Issue 3, pp. 446-452, Jan. 2003.
[37] S. Shentu, H. Zhao, and Y. Wu and X. Zhu, "Maximum entropy method for ocean acoustic tomography," *2017 IEEE ICSPCC*, Xiamen, China, 2017, pp. 1-6.
[38] I. Csiszar, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *Ann. Statist.* vol. 19, no. 4, pp. 2032–2066, Dec. 1991.
[39] S. E. Christodoulou, A. Gagatsis, S. Xanthos, S. Kranioti, A. Agathokleous, and M. Fragiadakis, "Entropy-based sensor placement optimization for waterloss detection in water distribution networks," *Water Resour. Manag.*, vol. 27, no. 13, pp. 4443–4468, Sep. 2013.
[40] Z.-H. Zou, Y. Yi, and J.-N. Sun, "Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment," *J. Environ. Sci*, vol. 18, no. 5, pp. 1020–1023, Sep.-Oct. 2006.
[41] J. Chen, W. Bian, Z. Wan, Z. Yang, H. Zheng, and P. Wang, "Identifying factors influencing total-loss marine accidents in the world: Analysis and evaluation based on ship types and sea regions," *Ocean Eng.*, vol. 191, p. 106495, Nov. 2019.
[42] T. Zhang and X. Niu, "Analysis on the utilization and carrying capacity of coastal tidal flat in bays around the bohai sea," *Ocean & Coast. Manag.*, vol. 203, p. 105449, Apr. 2021.
[43] X. Zhang, C. Wang, E. Li, and C. Xu, "Assessment model of ecoenvironmental vulnerability based on improved entropy weight method," *Sci. World J.*, vol. 2014, Jul. 2014.
[44] "NOAA Optimum Interpolation (OI) SST V2 data," from NOAA PSL, Boulder, Colorado, USA, https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html.
[45] H. Foroozand and S. V. Weijs, "Entropy ensemble filter: A modified bootstrap aggregating (bagging) procedure to improve efficiency in ensemble model simulation," *Entropy*, vol. 19, no. 10, p. 520, Sep. 2017.
[46] B. Han, H. Liu, and R. Wang, "Urban ecological security assessment for cities in the beijing–tianjin–hebei metropolitan region based on fuzzy and entropy methods," *Ecol. Modell.*, vol. 318, pp. 217–225, Jan. 2015.
[47] V. Amiri, M. Rezaei, and N. Sohrabi, "Groundwater quality assessment using entropy weighted water quality index (EWQI) in Lenjanat, Iran" *Environ. Earth Sci.*, vol. 72, no. 9, pp. 3479–3490, Apr. 2014.
[48] J. Zhao, G. Ji, Y. Tian, Y. Chen, and Z. Wang, "Environmental vulnerability assessment for mainland china based on entropy method," *Ecol. Indic.*, vol. 91, pp. 410–422, Aug. 2018.
[49] W. F. Caselton and J. V. Zidek, "Optimal monitoring network designs," *Stat. Probab. Lett.*, vol. 2, no. 4, pp. 223–227, Aug. 1984.
[50] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, Sep. 1936.
[51] P. Y. Papalambros and D. J. Wilde, *Principles of optimal design: modeling and computation*. Cambridge University Press, 2000.
[52] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *J. Mach. Learn. Res.*, vol. 9, no. 2, Feb. 2008.
[53] S. Y. Lee, I. B. Lee, U. H. Yeo, R. W. Kim, and J. G. Kim, "Optimal sensor placement for monitoring and controlling greenhouse internal environments," *Biosyst. Eng.*, vol. 188, pp. 190–206, Dec. 2019.
[54] W. Liang, L. Xiao, K. Zhang, M. Tang, D. He and K.-C. Li, "Data fusion approach for collaborative anomaly intrusion detection in blockchain-based systems," *IEEE Internet Things J.*, vol. 9, no. 16, pp. 14741–14751, Aug. 2022.
[55] W. Liang, Y. Li and K. Xie, D. Zhang, K.-C. Li, A. Souri and K. Li, "Spatial-temporal aware inductive graph neural network for C-ITS data recovery," *IEEE Trans. Intell. Transport. Syst.*, early access, doi: 10.1109/TITS.2022.3156266.

Qiannan Zhang is currently working toward a PhD degree in traffic information engineering and control at Shanghai Maritime University. Her main research interests include sensor placement and data processing in sensor networks.

Huafeng Wu (Senior Member, IEEE) received a Ph.D. from Fudan University. He is currently a Professor and a Ph.D. Supervisor with the Shanghai Maritime University. His research interests include the Internet of Things, wireless sensor networks and marine search and rescue.

Xiaojun Mei (Member, IEEE) received a Ph.D. in 2021 and he is currently a Postdoctoral Researcher with the College of Information Engineering at Shanghai Maritime University. His current interests include target localization in wireless sensor networks and path planning for localization.

Dezhi Han (Member, IEEE) received a Ph.D. from the Huazhong University of Science and Technology. He is currently a Professor at Shanghai Maritime University. His research interests include network security, cloud computing, mobile networking, and wireless communication.

Mario Donato Marino (Member, IEEE) received a Ph.D. from the University of Sao Paulo, Brazil. He is currently a Senior Lecturer at Leeds Beckett University, UK. His research interests include distributed computing, parallel computing, and performance evaluation.

Kuan-Ching Li (Senior Member, IEEE) received a Ph.D. from the University of Sao Paulo, Brazil. He is currently a Professor at Providence University, Taiwan. He is also a fellow of the IET. His research interests include parallel and distributed computing, Blockchain, and emerging technologies.

Song Guo (Fellow, IEEE) is currently a Professor at The Hong Kong Polytechnic University, Hong Kong. He is also FCAE, FAAIA, and MAE. His research interests include Edge AI and Federated Learning, AI empowered Internet-of-Things, and Edge Computing Driven Ubiquitous Blockchain.