Research Article

Lijun Zhang*

Nonlinear parameter optimization method for high-resolution monitoring of marine environment

https://doi.org/10.1515/nleng-2022-0219
received December 15, 2021; accepted July 3, 2022

Abstract: In this article, marine environment detection has been studied for improving the high resolution of the environment. The problem of low resolution of marine environment detection is caused by data synthesis defects. The supply chain management (SCM) technology is used to optimize related data to improve the resolution. The main procedure is to first preprocess the obtained hydro-logical data and eliminate the unreasonable amount represented by extreme values, and then the SCM method was used to estimate the results. Finally, the accuracy of the estimation is evaluated by the cross-validation algorithm. In the example verification, the comparison between the SCM method and the traditional optimal interpolation (OI) method in data integration accuracy has been done. This article compares mean square error, mean absolute error (MAE), root-mean-square error (RMSE), and $R^2$ parameters. SCM provides better results than OI. Mean error (ME) = 0.6°C/month, MEA = 1.6°C/month, RMSE = 42.3°C/month, and ME and MAE values are lower in summer. It shows that it is sensitive to the lack of data and has a better ability to provide high-resolution and accurate marine environmental data in real time.

Keywords: marine environmental monitoring, resolution, parameter optimization

1 Introduction

The diversified development of marine monitoring methods has laid a massive marine data foundation for the exploitation of marine resources, the development of marine economy, and the protection of marine environment. It also causes ocean data to exhibit multi-modal, superdimensional, massive, and other characteristics. The emergence of massive ocean complex data brings new challenges to the processing, analysis and application of ocean data. Algorithms for merging data sets from different sources continue to evolve and improve, and this is mainly due to advances in many fields, including marine science, especially hydrology and digital fisheries forecasting. Technological developments in these areas are based on basic assumptions about the reliability of remote sensing data for the marine environment, such as satellite remote sensing data (commonly known as background fields). Oceans cover more than seven-tenths of the earth’s surface and contain extremely rich resources. Existing study findings demonstrate that marine resources such as marine species, solid minerals, oil and natural gas, renewable energy, and other resources are abundant, with enormous development potential. The specific process is shown in Figure 1.

Al-Dhabi et al. systematically presented the validation process of the “adaptation” analysis method for the determination of mercury in marine organisms and sediments by solid sampling HR-CS-atomic absorption spectrometry based on the rapid temperature program and solid sampling calibration. The detection of the marine environment by using parameters such as selectivity, linearity, working range, repeatability, reproducibility, recovery of the calibration curve detection, and quantification limit [1]. Marine environmental monitoring programs require accurate high-resolution maps of seabed habitats. Wang et al. calculated the minimum distance between samples to ensure spatial independence, object-based image analysis on marine biomedicine and environmental sciences data products. The clustering algorithm is also used to map the large spatial variability of seabed structure. Generate a sampling scheme consisting of the locations and supplementary locations in the original video data for sampling. The
obtained ground truth data is integrated into a supervised task decision tree to obtain the habitat map [2]. Yang et al. aimed at the initial application of the unmanned aerial vehicle (UAV) remote sensing system used in dynamic and sea areas monitoring. The research shows that UAV low-altitude remote sensing system has the advantages of strong flexibility, fast response speed, and high resolution, can reach complex or dangerous airspace that no one can reach. The system can be equipped with special equipment such as spectrometer, collect relevant information, and ensure better monitoring [3]. Zuo et al. stated that in recent years, with the development and utilization of the ocean, the pollution of the marine environment has gradually increased, and natural marine disasters have occurred from time to time, which has caused a serious impact on the marine ecological environment and brought huge economic losses. Therefore, it is urgent to protect the marine environment [4]. Catania et al. believed that the monitoring of the marine environment is an important part of protecting the marine environment. The technical level and capability of marine environment monitoring directly affect the degree and effect of marine resource development and marine environmental protection [5]. Aullybux et al. researched and designed a marine environment monitoring system based on the Internet-of-thing technology and realized intelligent monitoring of the marine environment. The goal of monitoring service integration is to accomplish the integration of measurement parameters, which includes system modularization, and real-time data transmission for marine environment monitoring [6]. Ghabez and Harrison believed that the measurement of marine environmental parameters is to obtain the temporal and spatial distribution and variation laws of various marine environmental parameters. It provides basic marine environmental data for marine scientific research, marine resource development, and marine engineering [7]. Ding and Li stated that various specific marine parameter measurement systems are the premise for the implementation of accurate measurement of various marine parameters. The scientific onsite detection method of the measurement system is an indispensable condition to ensure the quality and effective use of the measurement system [8]. Zhu et al. stated that the marine environmental parameter measurement system has the following outstanding advantages in the field of inspection. First, it can quickly find various defects of the measurement system in design, components, parts, raw materials, and processes. Second, it can provide information on the integrity of the measurement system, improve the success of the task, reduce the cost of maintenance personnel and support costs. Third, it confirms whether it meets the design performance requirement of the measurement system [9]. Wang et al. believed that a large number of facts prove that only onsite testing can ensure that various indicators of the measurement system meet the requirements. Therefore, it is necessary to study the onsite detection method of the marine environmental parameter measurement system [10]. Tian et al. believed that at present, the detection methods of measurement systems are individually designed for specific measurement equipment, lacking general and normative methods [11]. Xavier believed that the detection methods of various marine environmental parameters of measurement scientific systems should be standardized and generalized to improve and complete the offshore detection task of the measurement system, that is, a scientific and urgent need for an in-depth study on the measurement system’s detecting methods and technologies [12]. González-Morales et al. used an example of an above-water and underwater measurement system to describe the inspection and statistical method of the marine environmental parameter measurement system in detail, which might serve as a valuable reference for future acceptance of the marine environmental parameter measuring system [13]. Jiang et al. believed that the necessary and correct detection methods are the premise of the quality assurance of the measurement system. The standardized detection
method can reduce test costs and inspection errors, improve the scientificity, increase the accuracy of the test and inspection, and also shorten the test and inspection cycle to promote economic efficiency [14]. Abbasi et al. believed that with the increasing role of measurement systems in marine engineering, the research on detection methods is becoming more and more important, and the requirements are getting higher and higher [15]. The main detection contents and methods of a typical marine environmental parameter measurement system introduced by Zhao et al. and its successful application in practice can provide a reference for the design and acceptance of similar measurement systems in the future [16].

Currently, optimal interpolation (OI) and its improved methods (Kalman filter, 3D variational assimilation, or spatial analysis) are the best methods for data synthesis. However, these methods are restricted by their own samples in terms of real-time performance and applicability, so this article adopts the supply chain management (SCM) method of minimum variance estimation [17].

2 Data synthesis algorithm

SCM is an iterative empirical method and is used in many engineering fields such as global meteorological space analysis. In classic SCM [18,19], the first test value of the grid point \( x_i^0 \) is calculated from the background \( x_i^0 \) estimate at the grid \( x_i^0 \). Before the second estimate, iterative formula (1) is obtained through SCM [20].

\[
x_i^{n+1} = x_i^n + \frac{\sum_{k=1}^{n} \omega_{ik}(x_k^n - x_i^n)}{\sum_{k=1}^{n} \omega_{ik} + \varepsilon_i^2},
\]

where \( x_i^n \) is the value at \( i \) of the \( n \)th iteration estimation of grid points, \( x_i^n \) is the \( k \)th observation around the grid point \( i \), \( \omega_{ik} \) is the weight value of the \( n \)th background estimation of the observation point \( k \) around the grid point \( i \), and \( \varepsilon_i^2 \) is the ratio \( (\varepsilon_i^2 = \varepsilon_i^2 / \varepsilon_i^2) \) between the observed error variance and the field error variance [21].

The weight is defined as formula (2).

\[
\omega_{ik} = \begin{cases} R_n^2 - r_{ik}^2, & r_{ik}^2 \leq R_n^2, \\ R_n^2 + r_{ik}^2, & r_{ik}^2 > R_n^2, \\ 0, & \end{cases}
\]

where \( R_n \) is the influence radius and \( r_{ik}^2 \) is the square distance between the value of grid point \( i \) and the observed value \( k \) [22]. The weight is defined as formula (3).

\[
\omega_{ik} = \exp\left(-\frac{r_{ik}^2}{2R_n^2}\right).
\]

Here, \( R \) is a constant. Both methods depend on the weight between the analysis grid point \( i \) and the observed values within the influence radius \( R_n \). \( R_n \) is fixed in the first iteration, as the region of influence changes with each iteration, and \( r \) varies between 1 and 2 [23,24]. In the first iteration, this radius is set to a large value \((r = 1)\) to capture the correlation of large-scale background fields [25]. The choice of radius values depends on a number of factors, for example, check the spatial distribution of data (a small number of points may affect the variability of neutron grid points in an unrepresentative sample) and the distance of observed data. Moreover, comparison of error variances \( \varepsilon_i^2 \) plays a critical role [26]. If \( \varepsilon_i^2 = 0 \) is in a small radius, the analysis field reflects \( k \) observed values in the small radius at the grid point \( i \). If the observed results are noisy or simply represent small-scale variability, a “bull’s eye” phenomenon can be observed in the field of analysis [27]. To avoid this problem, the error \( \varepsilon_i^2 > 0 \) of the observed values is assumed, and so it has some weight on the background field [28].

3 Data processing and experimental verification

3.1 Data preprocessing

The data obtained from the 16 ocean buoys of the State Oceanic Administration of H contain very few measurements of poor quality, so some adjustments have been made before the monthly mean is obtained [29]:

1) Delete extreme values (>42°C, <10°C) from daily records, which may be caused by human factors or instrument failure.
2) Replace missing or incorrect values in the temperature map with nearest neighbor interpolation at full resolution.
3) The data resolution was increased by sampling, the ocean buoy measurement results were compared with the remote sensing database, and the bilinear interpolation method was used to detect extreme differences (>42°C).
The differences were compared one by one, and the spatial distribution around the difference locations in the two datasets was evaluated.

Temperature estimation is prone to an error caused by horizontal space resolution limitations, simplification of numerical calculation incomplete solutions to ocean systems, and instrument deviations [30, 31]. Before merging the two data sets, the systematic biases of the two types of data must be eliminated.

With the advent of remote sensing satellite measurements, many different bias correction algorithms have been developed over the past decade [32]. Most of these methods fall into four categories.

First, average deviation correction includes estimating the mean deviation of all ocean buoys over a given period of time and using this value to correct remote sensing data. This method can be used in the case of uniform bias field [33]. Otherwise, the region is divided into a smaller region with uniform deviation.

Second, regression equation [34] includes estimating regression equation coefficients and correcting remote sensing data using historical time series and average coefficients for each buoy. The regression equation usually obtained in the literature is \( y = ax + b, \ y = ax^2 + bx, \) or \( y = ax^b \). This method is used when there is a good spatial correlation between ocean buoy data and remote sensing estimation [35].

Third, distribution transformation is a simplest method to use parameters estimated from two statistical distributions (mean \( \mu \) and standard deviation \( \sigma \)), the first is derived from the ocean buoy and the second is derived from the remote sensing estimate (at the location of the ocean buoy). Formula (4) is used to transform the second distribution into the first distribution [36].

\[
R_c = (R_0 - \mu_R) \frac{\sigma_B}{\sigma_R} + \mu_R + \frac{\sigma_B}{\sigma_R},
\]  

(4)

where \( R_c \) is the remote sensing estimation from deviation correction, \( R_0 \) is the uncorrected remote sensing estimate, and \( B \) and \( R \) are ocean buoy and remote sensing data, respectively.

Fourth, space transformation method involves using a determined deviation between ocean buoys, and remote sensing calculates the position of each buoy, generates a two-dimensional deviation curve, and in general uses a spline interpolation algorithm. Finally, the difference value is added to the remote sensing estimation [37].

To evaluate the performance of the aforementioned four methods, the gamma distribution is consistent with the ocean data, the uncorrected remote sensing data (for visualization only), and the corrected remote sensing data, and gamma distribution parameters of maximum likelihood estimation (\( \alpha \) is the shape parameter and \( \beta \) is the scale parameter) are used. The gamma distribution is widely used to represent shear stress transfer (SST) at different temporal resolutions because it is nonnegative, is positively skewed, and has shape flexibility. In addition, it can be defined by only two parameters [38]. Gamma goodness-of-fit evaluation formulas of ocean buoy and remote sensing SST are Eqs. (5), (6), and (7) in order:

\[
\delta = [(\hat{\mu}_G - \hat{\mu}_R)^2 + (\hat{\alpha}_G - \hat{\alpha}_R)^2]^{1/2},
\]

(5)

\[
\hat{\mu}_{G,R} = \hat{\alpha}_{G,R} \cdot \hat{\beta}_{G,R},
\]

(6)

\[
\hat{\alpha}_{G,R} = (\hat{\alpha}_{G,R} \cdot \hat{\beta}_{G,R})^{1/2},
\]

(7)

where \( \hat{\mu} \) and \( \delta \) are mean and standard deviation, by two parameters (\( \hat{\alpha}, \hat{\beta} \) of the gamma distribution, respectively, and the subscripts “\( G \)” and “\( R \)” represent buoy data and remote sensing data, respectively. The nonparametric Kolmogorov–Smirnov (K–S) test was used to verify whether the two samples were from the same distribution at the significance level of 10%. Table 1 presents examples of the results of each of the above fixes in March 2010. The data for this time period were chosen because the gamma probability density function studies for ocean buoy values, uncorrected remote sensing data, and corrected remote sensing data estimates were calibrated during this time. The power function \( y = ax^b \) is selected in the regression equations, and the coefficients \( a \) and \( b \) were estimated by the least square method [39]. Formula (4) is used in the distribution transformation group, and T spline regression algorithm is used in the spatial transformation group.

Table 2 summarizes parameters \( \delta \) and \( P \) for each of the summer and winter implementations over the entire time span. The smaller \( \delta \) value represents two gammas.

A better fit between distributions and a \( P \)-value greater than 0.1 mean that the assumption that samples are drawn from the same distribution is invalid. Distribution transformation and spatial transformation are the best, followed by mean deviation correction and regression equation, respectively. In the last two, the \( P \) value represents rejecting the null hypothesis in summer, not in winter. The optimal \( \delta \) value in summer is obtained by the distribution transformation method and in winter by the spatial transformation method [40].

As shown in Table 2, regression and spatial transformation correction algorithms may improve the mean bias results when performed in regions with uniform bias values. As for regression correction techniques, other types of regression equations can be used, but their
success is largely related to the time scale chosen for the data set. Therefore, considering the similarity of the winter results ($\delta$) and the variability of the results obtained by the distribution transformation method to the summer $\delta$ value, the distribution transformation method will be used.

### 3.2 Data processing

Final SST estimates were obtained using the proposed SCM method. To calculate the spatial correlation distance parameter $R$ in Eq. (3), the fitting of ocean buoy data is based on the model given in Eq. (3) to estimate the spatial correlation graph. The semi-variogram analysis has proved that the degree of anisotropy measured by the ocean buoy is negligible, so the isotropic function in Eq. (3) can be applied. Figure 2 shows two mean correlations, one for summer and one for winter. The correlation graph is only based on the semi-variogram and excludes the approximation of the bulking effect, using the data of the 6 months of summer and 6 months of winter (randomly selected) to calculate the average and using the exponential variogram model to describe the spatial correlation between the observed values. The distance corresponding to the spatial correlation of 0.5 is about 100 km in summer and 66 km in winter. Two seasons ($R = 0.5^\circ$) of correlation distance will be used for the maximum 100 km as the difference in distance difference is small [41]. The background field is a remote sensing SST from national minority supplier development council (NMSDC) with a horizontal resolution of 21 km $\times$ 21 km.

Only one correlation distance and one iteration are used in SCM, where $r = 1$. (i) Only one correlation distance is selected quality control procedures, (ii) set the observations to include a sample of subgrid scale variability (due to loss of measurement records), (iii) in the case of special SST spatial distribution, the final field should only reflect the small-scale background field, and (iv) the background field (remote SST) should be the best solution above the ocean buoy data. Otherwise, using statistical parameters (such as $R^2$, mean error (ME), and others) and visual inspection, after iteration, the best result is obtained.

### Table 1: Gamma probability density functions of ocean buoy values, uncorrected remote sensing data, and corrected Remote sensing data estimates (March 2010)

<table>
<thead>
<tr>
<th>Method</th>
<th>$\delta_{\text{sum}}$</th>
<th>$P$ value</th>
<th>$\delta_{\text{win}}$</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean deviation correction</td>
<td>26.22</td>
<td>0.10</td>
<td>9.20</td>
<td>0.12</td>
</tr>
<tr>
<td>Regression equation</td>
<td>36.10</td>
<td>0.06</td>
<td>11.52</td>
<td>0.32</td>
</tr>
<tr>
<td>Distribution transformation</td>
<td>6.76</td>
<td>11.52</td>
<td>6.21</td>
<td>0.13</td>
</tr>
<tr>
<td>Spatial transformation</td>
<td>10.54</td>
<td>0.32</td>
<td>4.17</td>
<td>0.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uncorrected remote sensing</th>
<th>Buoy</th>
<th>Corrected remote sensing</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 0.001</td>
<td>0.0012</td>
<td>0.001</td>
</tr>
<tr>
<td>15 0.0017</td>
<td>0.0023</td>
<td>0.0027</td>
</tr>
<tr>
<td>20 0.0028</td>
<td>0.0032</td>
<td>0.0035</td>
</tr>
<tr>
<td>25 0.0027</td>
<td>0.0035</td>
<td>0.0031</td>
</tr>
<tr>
<td>30 0.0015</td>
<td>0.0027</td>
<td>0.0017</td>
</tr>
<tr>
<td>35 0.0010</td>
<td>0.0007</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uncorrected remote sensing</th>
<th>Buoy</th>
<th>Corrected remote sensing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.004</td>
<td>0.003</td>
<td>0.0003</td>
</tr>
<tr>
<td>0.0013</td>
<td>0.0012</td>
<td>0.0018</td>
</tr>
<tr>
<td>0.003</td>
<td>0.0024</td>
<td>0.0022</td>
</tr>
<tr>
<td>0.0028</td>
<td>0.0035</td>
<td>0.0035</td>
</tr>
<tr>
<td>0.0017</td>
<td>0.0018</td>
<td>0.0020</td>
</tr>
<tr>
<td>0.0012</td>
<td>0.0003</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

### Table 2: Parameters $\delta$ and $P$ of various methods in summer and winter over the time span

<table>
<thead>
<tr>
<th>Method</th>
<th>$\delta_{\text{sum}}$</th>
<th>$P$ value</th>
<th>$\delta_{\text{win}}$</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean deviation correction</td>
<td>26.22</td>
<td>0.10</td>
<td>9.20</td>
<td>0.12</td>
</tr>
<tr>
<td>Regression equation</td>
<td>36.10</td>
<td>0.06</td>
<td>11.52</td>
<td>0.32</td>
</tr>
<tr>
<td>Distribution transformation</td>
<td>6.76</td>
<td>11.52</td>
<td>6.21</td>
<td>0.13</td>
</tr>
<tr>
<td>Spatial transformation</td>
<td>10.54</td>
<td>0.32</td>
<td>4.17</td>
<td>0.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uncorrected remote sensing</th>
<th>Buoy</th>
<th>Corrected remote sensing</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 0.0016</td>
<td>0.0012</td>
<td>0.0012</td>
</tr>
<tr>
<td>15 0.0031</td>
<td>0.0027</td>
<td>0.0026</td>
</tr>
<tr>
<td>20 0.003</td>
<td>0.0033</td>
<td>0.0034</td>
</tr>
<tr>
<td>25 0.0022</td>
<td>0.0036</td>
<td>0.0037</td>
</tr>
<tr>
<td>30 0.0015</td>
<td>0.0017</td>
<td>0.0020</td>
</tr>
<tr>
<td>35 0.001</td>
<td>0.0003</td>
<td>0.0007</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uncorrected remote sensing</th>
<th>Buoy</th>
<th>Corrected remote sensing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.004</td>
<td>0.003</td>
<td>0.0003</td>
</tr>
<tr>
<td>0.0013</td>
<td>0.0012</td>
<td>0.0018</td>
</tr>
<tr>
<td>0.003</td>
<td>0.0024</td>
<td>0.0022</td>
</tr>
<tr>
<td>0.0028</td>
<td>0.0035</td>
<td>0.0035</td>
</tr>
<tr>
<td>0.0017</td>
<td>0.0018</td>
<td>0.0020</td>
</tr>
<tr>
<td>0.0012</td>
<td>0.0003</td>
<td>0.0005</td>
</tr>
</tbody>
</table>
Finally, the accuracy of the estimation is evaluated by the leave-one-out cross-validation algorithm. Using 120 sets of records (at least 98% of ocean buoys had complete data records over the study time span), leaving one for each consecutive month, each algorithm had 1,920 estimates (a total of 3,840 estimates). To evaluate the performance of different SSTs, ME, mean absolute error (MAE), the root-mean-square error (RMSE), and the determination coefficient $R^2$ are calculated according to Eqs. (8)–(11).

$$\text{ME} = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i),$$  \hspace{1cm} (8)

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\hat{Y}_i - Y_i|,$$  \hspace{1cm} (9)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2},$$  \hspace{1cm} (10)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (\bar{Y} - Y_i)^2},$$  \hspace{1cm} (11)

where $n$ is the number of observations per month, $\hat{Y}_i$ and $Y_i$ are estimated and observed SST values measured at position $i$, respectively, and $\bar{Y}_i$ is the average value of the estimated values.

To analyze the data synthesis effect of the SCM method, it was applied to the SST data synthesis of a sea area in Country H, and the results were compared with those of the OI method. An average of 120 months (2010–2019) was calculated to make a statistical comparison of the spatial distributions obtained by the two algorithms. Figure 3(a) shows the location of ocean buoys...
used to calculate the monthly mean. Note that ocean buoy data are only valid if more than 27 days of data are available per month. Figure 3(b) shows the background field cloud map generated by remote sensing SST estimation after the elimination of offset. Figure 3(c) and (d) shows the comparison results of the SCM method and the OI method, in which all cloud graphs are the monthly average values of the span from 2010 to 2019.

According to the map detection results of the two methods, the results of the SCM method and the OI method have similar spatial distribution, but it can also be observed that the measurements of individual ocean buoys are not significantly corrected for the background field, and by comparing Figure 3(a) and (b), it can be seen that most of the uncorrected ones are located in the center and northwest of the sea area, and the “bull’s eye” effect can be observed in the map, as shown in Figure 3(c) and (d). The “bull’s eye” effect is more pronounced in the center of the South China Sea, where there are differences between SST data from some ocean buoys and background fields, but these differences are not errors in daily and monthly verification procedures. Compared with the OI algorithm, the SCM method has smoother and more detailed SST cloud Figure 3(d). On the surface, both methods appear to incorporate ocean buoy data and remote sensing data, showing similar results in Figure 3(a). Therefore, it is difficult to see which method is better by visual inspection without spatial statistical analysis.

Figure 4 summarizes the results of the statistical analysis. All values were calculated during a single month between 2010 and 2019. The statistical parameters considered here are as follows: ME (Figure 4(a)), MAE (Figure 4(b)), RMSE (Figure 4(c)), and determination coefficient $R^2$ Figure 4(d). The statistical methods were OI and SCM. The OI method shows the intermediate values of remote SST and $R^2$, but the ME and MAE values obtained using SCM are closer.

Table 3 presents the summer, winter, and overall data set statistics. The mean sizes of ME, MAE, and RMSE decreased slightly in winter (light gray) and in summer increased when they. As expected, $R^2$ values decrease in summer, but increase in winter. The OI method shows the intermediate values of remote SST and $R^2$, but the ME and MAE values obtained using SCM are closer.

To evaluate the performance difference between the two methods, the leave-one-out cross-validation technique was applied in this study. Each method was applied 120 times for the remaining set of 16 selected ocean buoy data sets. All 120 values are added to the remote sensing data grid corresponding to the location of the ocean buoy before calculating the next ocean

![Figure 4: Monthly statistics of total time span (2010–2019). (a) Mean error, (b) mean absolute error, (c) root mean square error, and (d) determination coefficient.](image-url)
buoy. This method has been applied 3,840 times (16 ocean buoys, 120 months, and 2 algorithms). Figure 5(a) shows the scatter diagram between the ocean buoy data and the cross-validation of OI results, Figure 5(b) shows a scatter plot between the ocean buoy data and the cross-validation results using the SCM approach. In the end, the results of the two methods are very similar. Compared with the SCM method, the OI method presents a better value $R^2 = 0.77$, but the SCM method produces a better RMSE value.

### 4 Conclusions

In this study, ocean buoys and remote SST from the NMSDC data set over a 12-month period were combined using the SCM method, and the results were evaluated using the leave-one-out cross-validation technique to obtain a better data fusion method. The method takes into account the unique Marine environmental data, the density of Marine buoys, and the spatial and temporal resolution of marine environmental factors. Compared with traditional data analysis method OI, SCM provided better results (ME = 0.8°C/month, MEA = 1.8°C/month, RMSE = 41.7°C/month, $R^2 = 0.87$°C). In contrast, the OI method was less accurate (ME = 0.9°C/month, MEA = 1.8°C/month, RMSE = 37.3°C/month, $R^2 = 0.85$°C). Compared with the OI method, SCM has better implementation, has stronger universality, has faster calculation speed, and can iteratively increase the smoothness of correction, the ability to provide high-resolution, accurate Marine environmental data in real time, especially in the South China Sea, where surface installations are sparse. It can improve the accuracy of fishery simulation and prediction so as to better plan and manage marine resources.

With the increasing development and utilization of ocean space, the ocean itself is far away from the living environment of human beings and has complex and changeable characteristics, and the high-resolution detection of the marine environment has a positive and far-reaching significance for the utilization of marine resources and navigation and traffic.

**Funding information:** Author states no funding involved.

**Author contributions:** Author has accepted responsibility for the entire content of this manuscript and approved its submission.

**Conflict of interest:** The author states no conflict of interest.

**References**


