Instruments and Methods

Generating time-series grid data of sea surface temperature from isotherms in the Northwestern Pacific Ocean using coupled interpolation

Q.Q. Shao*, K. Rong, W.W. Ma, Z.Q. Chen

Institute of Geographical Sciences and Natural Resources Research, CAS, 11A, Anwai, Datun Road, Beijing 100101, China

Received 15 September 2005; received in revised form 31 December 2006; accepted 5 January 2007

Available online 1 February 2007

Abstract

Sea surface temperature (SST) isoline charts that were manually mapped using in situ SST data and satellite-derived SST data are valuable because they incorporate oceanographers’ knowledge and experience. This type of SST data is useful for studying sea conditions of an area, for analyzing environmental factors that could affect fishing grounds, as a parameter for atmospheric or oceanic models, or as a diagnostic tool for comparison with the SSTs produced by ocean models. However, isoline maps must be digitized and interpolated into grid data in order to be used in these applications. Herein, we propose a coupled interpolation (CI), which couples improved multi-section interpolation and single-point change surface interpolation containing orientation, for generating grid data from SST isolines. We interpolated 1049 SST isoline maps (temperature interval 1 °C), which cover an area of the northwestern Pacific Ocean (125°E–180°E, 26°N–50°N) and were published by the Japan Fisheries Information Service Center (JAFIC) during 1990–2000, to grid datasets with 15° grid resolution. We assessed the quality of grid datasets by checking noise points, RMSE analysis, checking offset errors, retrieving percentage of Kuroshio axes and visually comparing inverse isotherms with original isotherms. The quality analysis and comparison with four other interpolators showed the CI interpolator to be a good technique for generating SST grid data from isotherms. We also computed the SST anomaly (SSTA) using the SST grid datasets. The amplitude values of integral SSTA in the area of 31–46°N, 170–180°E were low, whereas they were high in the SW–NE rectangular area of 35–46°N, 142–160°E.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Coupled interpolation; SST; Grid data; Quality analysis; Integral SSTA

1. Introduction

Changes in sea surface temperature (SST) play a fundamental role in the exchange of energy, momentum, and moisture between the ocean and the atmosphere (Wentz et al., 2000). SST is a central determinant of air–sea interactions and climate variability. The recurring El Nino-La Nina cycle is a dramatic manifestation of the coupling of SST to atmospheric circulation (Cane et al., 1986; McPhaden, 1999). SST also influences the development and evolution of tropical storms and hurricanes (DeMaria and Kaplan, 1994; Emanuel, 1999), is correlated with nutrient concentration and primary
productivity (Kamykowski, 1987), and impacts the distribution of fishing grounds (Shao et al., 2004).

Three kinds of SST data are commonly collected and used in SST studies: in situ SST data, satellite-derived SST data, and ocean model-produced SST data. Blended SST analysis (Reynolds, 1988; Reynolds and Marsico, 1993) and optimum interpolation (OI; Reynolds and Smith, 1994) are techniques used with both in situ and satellite-derived SST data to produce monthly, weekly, and daily analyses of data on a 1° lat/long spatial grid. The results have been widely distributed to researchers.

Another type of SST data, isotherm charts, represents data that have been manually mapped using in situ and satellite-derived SST data. Because oceanographers’ knowledge and experience were used to create these maps, this kind of SST data is more reliable than other types in reflecting SST fields, especially in complicated ocean areas (e.g., in perturbed areas such as the transitional area between the Kuroshio and Oyashio fronts in which intense mesoscale variability takes place, Fig. 1) in the Northwestern Pacific Ocean. The Japan Fisheries Information Service Center (JAFIC) has issued time-series SST isoline charts that span 4 and occasionally 5 days during 1980–2005 in the Northwestern Pacific Ocean. These charts are very useful for studying sea conditions of the area, for analyzing environmental factors that could affect fishing grounds, as a parameter for atmospheric or oceanic models, or as a diagnostic tool for comparison with the SSTs produced by ocean models. For example, Sainz-Trápaga and Sugimoto (1998) analyzed the horizontal hydrographic structure of the area and detected the Kuroshio Extension using the SST isoline maps of JAFIC (1990–1994).

In this study, we collected the time-series SST isoline maps for 1990–2000 issued by JAFIC. However, these isoline maps needed to be digitized and interpolated as grid data to be used for other applications. The generation of SST grid data from isotherms has received little attention in the literature. However, the process seems to be similar to the generation of digital elevation models (DEM) from contours.

Two catalog classes of interpolation methods exist for generating DEM from contours: the global method and the local method (Burrough, 1986). The global method uses information from all points in the study area as well as standard statistical ideas of variance analysis and regression (Ardiansyah and Yokoyama, 2002). The global method can generate a smooth surface, but ‘any change of original reference points or accuracy of the original data can cause unsteadiness of the solution, and the structure features and the accuracy of the original data are rather insufficient’ (Weibel and Heller, 1991). The local method uses information from the nearest points defined by a specific window or specific lines. Typical local methods include the triangulated irregular network (TIN)-based method, line method, weighting method, and morphological approach. The local method is used more often for DEM interpolation.

TIN is a surface model used to analyze and display terrain and other types of surfaces (Briggs,
The low- or high-degree function (e.g., linear or quintic) is used to generate grid data from TIN. Herein, we refer to this method as TIN-based interpolation. However, there are some complications with using this method to generate SST grid data from isotherms. For example, the pre-processing workload to generate a sound TIN is terribly large. Furthermore, we cannot define ocean features as hard or soft lines because SST is a dynamic field.

The line method includes: (1) the morphological contour method, the main idea of which is to create intermediate contour lines using dilation and erosion methods (Takagi and Shibasaki, 1996; Taud et al., 1999); (2) the fitted-function interpolation along straight lines in the direction of the steepest slope (Inaba et al., 1988; Ardiansyah and Yokoyama, 2002); and (3) the multi-section interpolation method, which is a fitted-function interpolation along several straight lines that pass an interest point and have a certain angle to each other (Zhang et al., 1998).

The weighting method assigns a value to the target point using a weighted average of all point values within a defined region (Michael, 1997). Many modifications of this approach exist (Hodgson, 1989; Clarke, 1990), including the single-point change surface interpolation containing orientation in DEM proposed by Wu et al. (2001).

The morphological approach for calculating regular grid DEMs from elevation point data, contour data, and stream data (Hutchinson, 1989; Hutchinson and Dowling, 1991) is very popular (Gousie and Franklin, 2005). The incorporation of stream line data in practice is a powerful way to improve accuracy of DEM (Hutchinson, 1989). However, the SST field changes dynamically, and ocean features (such as the Kuroshio path, Kuroshio branches, temperature fronts, meso-eddies, and so on in Northwestern Pacific Ocean), which can be reflected by SST, also change dynamically. Thus, we cannot provide imposed ocean feature condition data for the interpolation of grid datasets from the 1049 isoline coverages. In fact, our objective is to extract the above-mentioned ocean from gridding SST data in contrast.

Finally, we also tried to interpolate our data using multi-section interpolation (commercial software MapGIS) and a single-point change surface interpolation containing orientation method that we developed by ourselves.

Each of these methods had its strong points and shortcomings, but none of them satisfied our needs of retrieving oceanic features such as the Kuroshio path off the south coast of Japan, the Kuroshio branches in the Kuroshio Extension, temperature fronts, meso-eddies, and so on. Thus, we used the method of coupled interpolation (CI) for generating grid data from sea surface isotherms that is proposed in this paper. This method coupled improved multi-section interpolation and single-point change surface interpolation containing orientation. In the sections that follow, we first discuss the data source and data processing, and then we discuss the CI method and analyze the quality of the grid data generated by CI and other popular commercial interpolators. Finally, we provide an example of the application of the time-series SST grid datasets generated from isotherms using the CI method.

2. Data source and data processing

2.1. Data source

For this study, we used the 1049 SST isoline maps (temperature interval 1°C), which cover an area of 125°E–180°E and 26°N–50°N in northwestern Pacific Ocean and were published by JAFIC during 1990–2000 of which 992 SST maps were 4-day span, 52 SST maps were 5-day span, 3 SST maps were 6-day span, and 2 SST maps were 1-day span.

2.2. Data processing

All SST isoline maps were digitized, edited, and processed using ArcGIS software and saved in coverage format with temperature values and spatial topology of isotherms. We refer to these digitized maps as SST e-maps. To detect the effect of different grid resolutions and interpolators on the results, we used five interpolators to interpolate SST grid data from isotherms at three different grid resolutions:

- **TIN-based linear and TIN-based quintic interpolators**: All SST e-maps were interpolated to grid data using TIN-based linear and TIN-based quintic interpolators of ArcGIS software with three grid resolutions: 15', 10', and 5' (the dataset names were correspondingly abbreviated to TIN-L15, TIN-L10, TIN-L5, TIN-Q15, TIN-Q10, and TIN-Q5).
The 102 SST e-maps from 1999 were interpolated with grid resolution of 15' using the multi-section interpolator of MapGIS software (the dataset name was abbreviated to MS15). Sixteen SST e-maps (random samples) were interpolated to grid data using the multi-section interpolator with grid resolutions of 10' and 5' (the dataset names were correspondingly abbreviated to MS10 and MS5).

**TOPOGRID:** All SST e-maps were interpolated with grid resolution 15' using the TOPOGRID interpolator of ArcGIS software (the dataset name was abbreviated to TP15). Sixteen SST e-maps (random samples) were interpolated to grid data using the TOPOGRID interpolator with grid resolutions 10' and 5' (the dataset names were correspondingly abbreviated to TP10 and TP5).

**CI:** All SST e-maps were interpolated to grid data using CI (discussed in Section 3.1) with grid resolution 15' (the dataset name was abbreviated to CI15). Sixteen SST e-maps (random samples) were interpolated to grid data by CI with grid resolution 10' and 5' (the dataset names were correspondingly abbreviated to CI10 and CI15).

To make quality analysis of SST grid data, 16 grid datasets (random samples) interpolated from SST isolines generated by the five interpolators described above were inversely interpolated to SST isolines with the CONTOUR function tool of ArcGIS software (inverse SST isoline map names were abbreviated to IN-TIN-L15, IN-TIN-L10, IN-TIN-L5, IN-TIN-Q15, IN-TIN-Q10, IN-TIN-Q5, IN-MS15, IN-MS10, IN-MS5, IN-TP15, IN-TP10, IN-TP5, IN-CI15, IN-CI10, and IN-C5).

To execute RMSE analysis of SST grid data, an e-map from 27 to 30 June 1994 was randomly selected. Sixty check points were randomly distributed over the whole study area (Fig. 2), and the temperature values of check points on the original isotherm map for IN-TIN-L15, IN-TIN-Q15, IN-MS15, IN-TP15, and IN-CI15 maps were calculated manually.

To compare retrieving percentages of the Kuroshio axis off the coast of Japan from SST grid datasets generated by different interpolators, the Kuroshio axes were retrieved using the STREAM function of ArcGIS from all TIN-L15, TIN-Q15, TP15, and CI15 grid datasets during 1990 and 2000, and from 102 MS15 grid datasets from 1999.

### 3. The CI for interpolating SST grid data from isolerms

#### 3.1. The principle and formulae of CI

The CI method coupled improved multi-section interpolation and single-point change surface interpolation containing orientation for generating grid data from SST isolines. The process is as follows: First, generate grid data from SST isolines using improved multi-section interpolation and synchronously identify the noise points; and second, interpolate the value for each noise point using the normal temperature value of points within the dynamic circle centered at the noise point in the grid dataset. The idea behind multi-section interpolation is to use lines that have a certain angle and
that go through the unknown point to intersect the nearest SST isolines to get known points, and then to use temperature values of known points on each line to calculate the temperature values of the unknown point. The interpolation is either linear or parabolic, depending on the number of known points crossed.

3.1.1. Improved multi-section interpolation

3.1.1.1. The main idea. We improved upon the multi-section interpolation introduced by Zhang et al. (1998). The basic principle of improved multi-section interpolation is (Fig. 3):

- to make a ray PP (positive) from an interest point p(x, y) having angle \( \alpha \) to the horizontal line to intersect isolines to get the nearest point \( PT_1(p_{x1}, p_{y1}) \) and the second nearest point \( PT_2(p_{x2}, p_{y2}) \);
- to make another ray PN (negative) from the interest point p(x, y) in the opposite direction to intersect isotherms to get the nearest point \( NT_1(n_{x1}, n_{y1}) \) and the second nearest point \( NT_2(n_{x2}, n_{y2}) \);
- two rays form one section;
- several sections can be gained by changing the angle \( \alpha \) (which is set to 45° in this paper), and there are four points at most on each section;
- several mathematical curves can be defined on each section, because a mathematical curve (e.g., line, parabola, arc) can be defined by two or three points;
- the temperature value \( (T_i) \) and weight \( (W_i) \) can be calculated using the curve function \( i \); the temperature value \( (T_j) \) and weight \( (W_j) \) of point \( P(x, y) \) on each section \( j \) can be calculated by weighted sum \( (T_i) \) and weighted sum \( (W_i) \);
- the temperature value \( (T) \) of point \( P(x, y) \) can be calculated by weighted sum \( (T_i) \) of sections that fall in normal ranges. The normal temperature value range of point \( P(x, y) \) is dynamically calculated during interpolation to identify the noise points with abnormal temperature values (discussed in Section 3.1.2).

3.1.1.2. Formulae of improved multi-section interpolation. The linear formula for two points is as follows:

\[
y = \left[ \frac{(x - x_2)}{(x_1 - x_2)} \right] y_1 + \left[ \frac{(x - x_1)}{(x_2 - x_1)} \right] y_2, \quad \text{and}
\]

\[
w = 1/|x_1|-|x_2|.
\]

The parabola formula for three points is as follows:

\[
y = \left[ \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} \right] y_1 + \left[ \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} \right] y_2 + \left[ \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} \right] y_3.
\]

If \( x_1 < 0 \) and \( x_3 > x_2 > 0 \), then

\[
w = \left[ \frac{3.0/x_2}{|x_1|} \right] \min[(x_2 + |x_1|)/(x_3 - x_2), (x_3 - x_2)/(x_2 + |x_1|)].
\]

If \( x_1 < x_2 < 0 \) and \( x_3 > 0 \), then

\[
w = \left[ \frac{3.0/x_3}{|x_2|} \right] \min[(|x_2| + x_3)/(x_2 - x_1), (x_2 - x_1)/(|x_2| + x_3)].
\]

where \( x_i \) is the distance between intersect point \( i \) on isolines and an interest grid point; \( y_i \) the temperature value of an isotherm on which intersect point \( i \) lies; \( x \) the coordinate of an interest grid point in the horizontal direction and is always zero; and \( y \) the temperature value of an interest grid point. According to the number and distribution of points on a section, seventeen processing methods exist (see Appendix A).

3.1.2. The method of identifying noise points

It is important to calculate the normal range of temperature values for interest points in order to identify noise points. If the temperature value on one section does not fall in the normal range of values, the weight value of this section will be set to 0. The temperature value of an interest point is calculated using the weighted sum temperature
values \((T_j)\) that fall in the normal ranges on all sections. If temperature values on all sections are 0, this interest point will not be given any value and will be labeled as a noise point.

One criterion for assessing the quality of a DTM interpolated from contours is as follows: ‘In each area bounded by a contour pair, DTM heights must assume values within the elevation range defined by the two contour labels’ (Carrara et al., 1997); this criterion was also used by Gousie and Franklin (2005). We defined the criteria for the normal range of temperature values of an interest point as: In each area bounded by an isotherm pair, grid point temperatures must assume values within the temperature range defined by the two isotherm labels.

Two special conditions were also considered for calculating the normal range of temperature values of an interest point: (1) for each area bounded just by an isotherm for which the value of the isotherm is greater than the value of the nearest isotherm, grid point temperatures must assume values greater than the value of the isotherm; and (2) for each area bounded just by an isotherm for which the value of the isotherm is less than the value of the nearest isotherm, grid point temperatures must assume values less than the value of the isotherm.

For the first assumption, two possible conditions exist: simple and complicated isotherms. If an interest point \(P\) falls in an area of simple isotherms, it is easy to calculate the normal range of temperature values of an interest point. If an interest point \(P\) falls in an area of complicated isotherms (Fig. 4), the high value may be equal to the low value on all sections of interest point \(P\). Thus, we propose the following method of calculation:

From interest point \(P\), make a ray \(l\) along one of the sections to intersect the nearest isotherm to get an intersect point \(a\) and its temperature value \(t_a\) and to intersect the second nearest isotherm whose value is not an equal value \(t_a\) to get intersect point \(b\) and its temperature value \(t_b\).

If value \(|t_b-t_a|\) is equal to temperature difference \((1{\degree}C\) in Fig. 4) of adjacent isotherms, then the upper limitation and lower limitation of the normal range of temperature values of an interest point \(P\) can be calculated as follows:

\( \text{If } t_b > t_a \text{ and the number of intersect points between points } a \text{ and } b \text{ is odd, then } t_{\text{High}} = t_a + 1 \text{ and } t_{\text{Low}} = t_a. \)

\( \text{If } t_b < t_a \text{ and the number of intersect points between points } a \text{ and } b \text{ is even (include 0), then } t_{\text{High}} = t_a + 1 \text{ and } t_{\text{Low}} = t_a. \)

\( \text{If } t_b > t_a \text{ and the number of intersect points between points } a \text{ and } b \text{ is odd, then } t_{\text{High}} = t_a \text{ and } t_{\text{Low}} = t_a - 1. \)

\( \text{If } t_b < t_a \text{ and the number of intersect points between points } a \text{ and } b \text{ is odd, then } t_{\text{High}} = t_a \text{ and } t_{\text{Low}} = t_a. \)

where \(t_{\text{High}}\) is the upper limitation and \(t_{\text{Low}}\) is the lower limitation of the range of normal temperature values of an interest point \(P\).

If the value of \(|t_b-t_a|\) is not equal to the temperature difference of adjacent isotherms, then we get intersect points \(a\) and \(b\) and their temperature values on another section of interest point \(P\) until \(|t_b-t_a|\) is equal to the temperature difference of adjacent isotherms. We used the above algorithms to calculate the upper limitation and lower limitation of the normal range of temperature values of an interest point \(P\).

3.1.3. Point change surface interpolation containing orientation

If temperature values of grid points did not fall within the normal ranges of values, they were not given values and were labeled as noise points. In a grid dataset in which these noise points were not interpolated again, there were many grid points without temperature values. Therefore, we used single-point change surface interpolation containing orientation, which was introduced for generating DEM from contour lines by Wu et al. (2001), to interpolate temperature values of noise points. The
temperature value of a noise point was interpolated from temperature values of 2–4 grid points (hereafter referred to as known points).

3.1.3.1. Formulae of single-point change surface interpolation containing orientation. In this method, every known point is given a weight \( \lambda_i \), which is related to the distance between the known point and the interest point \( A \), and the temperature value \( T_A \) of interest for point \( A \) is estimated using the following formula:

\[
T_A^* = \sum_{i=1}^{n} \lambda_i T_i, \quad \text{where: } \lambda_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{n} \lambda_i = 1 \tag{5}
\]

where \( \lambda_i \) is a weight function of distribution and distance of known points and calculated with the following formula (Wu et al., 2001):

\[
\lambda_i = e^{-kD_i(1 + \beta Q_i)}/\sum e^{-kD_i(1 + \beta Q_i)}. \tag{6}
\]

In formula (6), \( k \) is a constant (usually a positive number less than 1) and \( \beta \) is related to the terrain conditions according to Wu et al. (2001). After running the tests many times and considering the situation of the sea area studied in this paper, we set \( \beta = 1 \) and \( k = 0.5 \). \( Q_i \) is the function of density between known points containing orientation, \( D_i = d_i/d \), \( d_i \) is the distance between a known point and an interest point, and \( d \) is the average distance of known points:

\[
Q_i = \sum_{k=1,k \neq i}^{n} d_k(1 - \cos \angle POA)/\sum_{k=1,k \neq i}^{n} d_k. \tag{7}
\]

In formula (7), \( \angle POA \) is an angle between the line connecting known point \( P \) and another known point and the line connecting known point \( P \) and interest point \( A \); \( d_k \) is the distance between a known point and an interest point; the value range of \( Q_i \) is from 0 to 2 (the higher the value of \( Q_i \), the lower the distance density between known point \( i \) and other known points and the higher the effect of known point \( i \) on the interest point).

3.1.3.2. Searching known points with a dynamic circle. The known points around an interest point were sought using a dynamic circle; its algorithms are described in Michael (1997), Hodgson (1989), Clarke (1990), Wu et al. (2001), and Gousie and Franklin (2005). The initial value of the searching radius \( r \) is the spatial resolution \( d \) of the grid. Because grid resolution is 15" in the geographic coordinate system in this paper, the initial search radius was set to 0.25° \( (r = d = 0.25) \). The search steps are as follows:

(a) seek known points within a circle radius, \( d \), centered at an interest point;
(b) according to the number (\( N \)) of known points within the circle, three conditions existed: If \( 2 \leq N \leq 4 \), then all known points were used for interpolation; if \( N > 4 \), then the four known points nearest to the interest point were used for interpolation; and if \( N < 2 \), the known points were sought again within a circle radius of \( r + d \), centered at the interest point, and this step could be repeated by increasing the radius until \( N \geq 2 \).

3.2. Quality analysis of SST grid data

3.2.1. Checking noise points

The 102 grid datasets of CI15, TIN-L15, TIN-Q15, TP15, and MS15 were checked for noise points, and their total, average, maximum, and minimum numbers of noise points are listed in Table 1. The total number of grid points in all 102 SST grid datasets was \(~1,569,270\). The number of noise points was greatest in the TIN-L15 grid datasets (average of 5392.2) and lowest in the CI15 grid datasets, (average of 33.8). Using criterion-1, the number of noise points was greatest in the TP15 grid datasets (average of 2588.7) and lowest in the CI15 grid datasets (average of 33.7). The difference in the number of noise points between the TIN-L15 and TIN-Q15 grid datasets using the original criterion and criterion-1 was very large, but the difference was small between the CI15 and MS15 grid datasets.

The noise points identified using the criteria described above showed that the accuracy of the grid data generated by CI was greatest, followed by MS, TIN-Q, TP, and TIN-L.

3.2.2. Root mean square error (RMSE) analysis

RMSE analysis is usually used to assess the accuracy of DEMs generated from contours (Gao, 1997; Ardiansyah and Yokoyama, 2002; Gousie and Franklin 2005). We used RMSE in this study to analyze the accuracy of SST grid datasets. The RMSE equation is as follows:

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (T_i - t_i)^2}, \tag{8}
\]
where \( n \) is the number of check points; \( T_i \) the temperature value read interactively from the enlarged original isotherm map at check point \( i \); and \( t_i \) the temperature value of the same check point read interactively from the enlarged inverse interpolated isotherm map.

Sixty check points (see Section 2.2 and Fig. 2) were used to evaluate the accuracy of TIN-L15, TIN-Q15, CI15, TP15, and MS15 grid datasets, and their RMSE and maximum residuals are listed in Table 2. The RMSE of the grid data generated by CI was 0.0599, by MS was 0.0911, by TIN-Q was 0.1184, by TIN-L was 0.1253, and by TP was 0.2779. The accuracy of the grid data generated by CI was greatest, followed by MS, TIN-Q, TIN-L, and TP.

3.2.3. Comparison of inverse interpolated isotherms and initial isotherms

Visualization provides a powerful mechanism for identifying spatial distributions and uncertainty of DEM (Taud et al., 1999). The comparison of initial contours and inverse interpolated contours is usually used for assessing the quality of DEMs (Gao, 1997; Ardiansyah et al., 2002; Carrara et al. 1997), and we used this method to analyze the quality of SST grid data.

The maps of IN-TIN-L15, IN-TIN-L10, IN-TIN-L5, IN-TIN-Q15, IN-TIN-Q10, IN-TIN-Q5, IN-MS15, IN-MS10, IN-MS5, IN-TP15, IN-TP10, IN-TP5, IN-CI15, IN-CI10, and IN-CI15 for 3–6 July 1995 were overlaid with original SST isoline maps. Fig. 5 shows an enlarged area of IN-CI5, IN-MS5, IN-TIN-Q5, IN-TIN-L5, and IN-TP5 overlaid with original SST isoline maps. A comparison of the inverse and original SST isolines in Fig. 5 showed that the inverse SST isolines of the data grid generated by CI and the original isolines were almost identical (Fig. 5a): there were three line-crossed errors and a few offsets in IN-MS5 (Fig. 5b); there were 13 line-crossed errors and some offsets in IN-TIN-Q5 (Fig. 5c), there were 27 line-crossed errors and some offsets in IN-TIN-L5 (Fig. 5d); and there were five line-crossed errors and some offsets in IN-TP5 (Fig. 5e). A comparison of inverse SST isoline in IN-TIN-L15, IN-TIN-Q15, IN-MS15, IN-CI15, and IN-TP15 with original SST isolines shows that the inverse SST isolines in IN-CI15 were still the best, and the inverse SST isolines in IN-TP15 were the worst, which can hardly express the complicated original SST isolines.

Table 3 shows the statistical results of offset errors on inverse SST isoline maps interpolated using grid datasets with 15', 10', and 5' spatial resolutions generated from SST isolines for 3–6 July 1995 by five interpolators. The greater the offset errors, the lower the accuracy of the grid data. This method showed that the accuracy of the grid data generated by CI was greatest, followed by TIN-Q, TIN-L, MS, and TP.

3.2.4. Comparison of retrieving percentages of the Kuroshio axis off the coast of Japan from SST grid datasets generated by different interpolators

One of the most important reasons for generating SST grid data from isolines was to retrieve oceanic features such as the Kuroshio axis, temperature fronts, eddies, Kuroshio branches, Oyashio branches in the Northwestern Pacific Ocean. We used the retrieving percentages of the Kuroshio axis off the coast of Japan from SST grid datasets generated by different interpolators to assess the accuracy of SST grid datasets generated from isolines. The higher the retrieving percentage, the higher the accuracy of the grid data (see Table 4).

The number of Kuroshio axes off the coast of Japan that could be retrieved completely from the

---

<table>
<thead>
<tr>
<th>Dataset name</th>
<th>Total number of NPs</th>
<th>Average number of NPs in a dataset</th>
<th>Max number of NPs in a dataset</th>
<th>Min number of NPs in a dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI15</td>
<td>3441</td>
<td>33.7</td>
<td>248</td>
<td>1</td>
</tr>
<tr>
<td>MS15</td>
<td>5145</td>
<td>50.4</td>
<td>521</td>
<td>14</td>
</tr>
<tr>
<td>TIN-L15</td>
<td>18661</td>
<td>182.9</td>
<td>462</td>
<td>74</td>
</tr>
<tr>
<td>TIN-Q15</td>
<td>48222</td>
<td>472.8</td>
<td>831</td>
<td>322</td>
</tr>
<tr>
<td>TP15</td>
<td>264049</td>
<td>2588.7</td>
<td>3627</td>
<td>1738</td>
</tr>
</tbody>
</table>
1049 SST grid datasets interpolated by CI15 was 649 (retrieving percentage of 61.9%), by TP15 was 474 (45.2%), by TIN-Q15 was 559 (53.3%), by TIN-L15 was 342 (32.6%). These results show that the accuracy of the grid data generated by CI was greatest, followed by TIN-Q, TP, and by TIN-L15. When the retrieving percentages from 1999 were compared, the accuracy of the grid data generated

<table>
<thead>
<tr>
<th>Dataset name</th>
<th>CI15</th>
<th>MS15</th>
<th>TIN-Q15</th>
<th>TIN-L15</th>
<th>TP15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of checking points</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0599</td>
<td>0.0911</td>
<td>0.1184</td>
<td>0.1253</td>
<td>0.2779</td>
</tr>
<tr>
<td>Maximum residual</td>
<td>0.2347</td>
<td>0.5048</td>
<td>0.4003</td>
<td>0.4139</td>
<td>0.8924</td>
</tr>
</tbody>
</table>

Fig. 5. The original SST isolines (black line) and the (a) IN-CI5 (red line); (b) IN-MS5 (red line); (c) IN-TIN-Q5 (red line); and (d) IN-TIN-L5 (red line); (e) IN-TP5 (red line) from 3 to 6 July 1995.
by CI remained greatest, followed by TIN-Q, TP, MS, and TIN-L.

4. Application of time-series SST grid datasets

Here we provide an example of the application of grid datasets with 15° spatial resolution generated by CI from the 1049 SST maps (temperature interval 1°C) that covered an area of 125°E–180°E, 26°N–50°N in the Northwestern Pacific Ocean issued by JAFIC for 1990–2000.

Namias (1972) reported that the area of highest SSTA variability in the North Pacific Ocean was situated at 35°–45°N and 155°E–175°W. Integral SSTA represent the sum of synoptic and interannual anomalies (Moshonkin and Dianskiy, 1995). Chen et al., 1999 calculated the integral monthly SSTA of grid points in the North Pacific Ocean using SST grid data issued by the China Meteorological Administration during 1951–1995 using the following MSTD(t) formula:

$$\text{MSTD}(t) = \int_0^t \Delta T(t) \, dt,$$

$$= \lim_{\Delta t \to 0} \sum_{i=0}^{t_f} \Delta T(t_i) \Delta t_i \approx \sum_{i=0}^{t_f} \Delta T(t_i).$$

(9)

FMSTD(t) is the Fourier transform of MSTD(t) at every grid point. Chen calculated the amplitude of integral monthly SSTA at every grid point from the corresponding curve of FMSTD(t). Fig. 6 shows the spatial distribution of the integral monthly SSTA amplitudes in the North Pacific Ocean calculated by Chen et al. (1999): The amplitude values were more than 50°C in the area of 31–46°N, 170°E–150°W, and maximum amplitude values in the center of the area were up to 90°C.

We used our CI data to calculate the integral monthly SSTA following Chen’s method. Fig. 7 shows the spatial distribution of integral monthly SSTA amplitudes in the Northwestern Pacific Ocean. We found that the amplitude values of integral monthly SSTA were more than 40°C in a rectangular area in the direction SW–NE at 35°–46°N, 142°–160°E, and the maximum amplitude values of the area were up to 70°C. However, the amplitude values of integral monthly SSTA were lower than 50°C in the area of 31–46°N and 170–180°E. The spatial pattern of integral monthly SSTA amplitudes in the Northwestern Pacific Ocean from this study was clearly different from that given by Chen et al. (1999, Fig. 6 left-up). The discrepancies may arise from uncertainty in the grid SST data, from differing spatio-temporal or time-series lengths, or from another source entirely.

5. Conclusion and discussion

In this study, we proposed a CI method for generating SST grid data from isolines. Our method coupled improved multi-section interpolation and single-point change surface interpolation containing orientation. The quality of grid data generated by CI, MS, TIN-L, TIN-Q, and TOPOGRID interpolators was analyzed and compared using five methods (checking noise points, RMSE analysis, checking offset errors, visually comparing inverse isotherms with original isotherms, and retrieving percentage of Kuroshio axes).

All five of the quality analysis results revealed that CI is a very good interpolator for generating grid data from SST isolines (Table 5). Table 5 also shows that CI is superior to other interpolation methods for generating grid data from SST isolines.

We also computed SSTA for 1990–2000 using SST grid datasets and found that the amplitude values of integral SSTA in the area of 31–46°N and 170–180°E were low but that values were high in the rectangular area in the direction SW–NE at
Table 4
The retrieving percentages of the Kuroshio axis off Japan Coast from SST grid datasets by different interpolators

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>CI15</th>
<th>TP15</th>
<th>TIN-Q15</th>
<th>TIN-L15</th>
<th>MS15*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>1990</td>
<td>75</td>
<td>37</td>
<td>49.3</td>
<td>20</td>
<td>26.7</td>
<td>30</td>
</tr>
<tr>
<td>1991</td>
<td>90</td>
<td>43</td>
<td>47.8</td>
<td>23</td>
<td>25.6</td>
<td>35</td>
</tr>
<tr>
<td>1992</td>
<td>91</td>
<td>70</td>
<td>76.9</td>
<td>44</td>
<td>48.4</td>
<td>57</td>
</tr>
<tr>
<td>1993</td>
<td>90</td>
<td>46</td>
<td>51.1</td>
<td>33</td>
<td>36.7</td>
<td>42</td>
</tr>
<tr>
<td>1994</td>
<td>94</td>
<td>56</td>
<td>59.6</td>
<td>48</td>
<td>51.1</td>
<td>44</td>
</tr>
<tr>
<td>1995</td>
<td>101</td>
<td>70</td>
<td>69.3</td>
<td>46</td>
<td>45.5</td>
<td>56</td>
</tr>
<tr>
<td>1996</td>
<td>101</td>
<td>90</td>
<td>89.1</td>
<td>74</td>
<td>73.3</td>
<td>77</td>
</tr>
<tr>
<td>1997</td>
<td>102</td>
<td>66</td>
<td>64.7</td>
<td>53</td>
<td>52</td>
<td>54</td>
</tr>
<tr>
<td>1998</td>
<td>102</td>
<td>69</td>
<td>67.6</td>
<td>57</td>
<td>55.9</td>
<td>73</td>
</tr>
<tr>
<td>1999</td>
<td>102</td>
<td>45</td>
<td>44.1</td>
<td>40</td>
<td>39.2</td>
<td>41</td>
</tr>
<tr>
<td>2000</td>
<td>101</td>
<td>57</td>
<td>56.4</td>
<td>36</td>
<td>35.6</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>1049</td>
<td>649</td>
<td>61.9</td>
<td>474</td>
<td>45.2</td>
<td>559</td>
</tr>
</tbody>
</table>

*The symbol ‘—’ means it is not calculated.

Fig. 6. The spatial pattern of integral monthly SSTA amplitudes in the Northwestern Pacific Ocean (after Chen et al., 1999) during 1951–1999.

Fig. 7. The spatial pattern of integral monthly SSTA amplitudes in the Northwestern Pacific Ocean during 1990–2000.
35–46°N and 142–160°E. These results differed from the results of others. The uncertainty of SST grid data, different sources, different spatial-temporal resolution, or different time series might account for the discrepancy and need to be studied further.

These time-series SST grid data can be used in many applications, such as detecting the surface thermal front, calculating the surface horizontal temperature gradient, and detecting the surface axis of the Kuroshio, all of which need to be studied further.

Although the smaller the \(a\) (angle between section lines), the greater the accuracy of interpolation results, the smaller the \(a\) the greater the computation cost of intersecting section lines and SST isolines to gain known points. The speed of CI procedure is slow now and needs to be improved.

**Acknowledgements**

This study was supported by the National Natural Science Foundation (40671151) of China and the Key Knowledge Innovation Project of the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences. We wish to express our thanks to Prof. Xinjun Chen and JAFIC for providing SST isoline charts of the Northwestern Pacific Ocean and Prof. Jiyuan Liu and Prof. Chenghu Zhou for helpful discussions. We are particularly grateful to the reviewers of the manuscript for their valuable comments and suggestions. We also thank other members of our workgroup (Lu Bai, Zhang Shuai, Kebiao Mao, Zhiming You, Xiaofeng Jia, Tao Wu, Hao Huang, Yan Chi, Xionghao Zhu, and Bo Yang) for processing data.

### Appendix A. Seventeen conditions of multi-section interpolation

According to the number and distribution of points on a section, seventeen processing methods exist (where \(h\) is the temperature value, \(w\) the weight, \(p\) the number of points on ray PP, and \(n\) the number of the points on ray PN):

1. If all temperature values of intersect points are the same, or the number of points \((p + n)\) is less than two on one section, then \(w = 0\).
2. If \(n = 1\) and \(p = 1\), then \(h\) and \(w\) are calculated by formula (1).
3. If \(n = 0\) and \(p = 2\), then \(h\) and \(w\) are calculated by formula (1).
4. If \(n = 2\) and \(p = 0\), then \(h\) and \(w\) are calculated by formula (1).
5. If \(n = 1\) and \(p = 2\), then three conditions exist:
   - If \(n_y\) is not equal to \(p_y\), then \(h_1\) is calculated by the parabola formula, \(w_1\) is calculated by formula (3), and \(h_2\) and \(w_2\) are calculated by formula (1) with two points \((PT_1, PT_2)\) where \(w = w_1 + w_2\), \(h = h_1/(w_1 + w_2)h_1 + w_2/(w_1 + w_2)h_2\).
   - If \(n_y\) equals \(p_y\) and \(n_y\) is not equal to \(p_y\), then \(h_1\) is calculated by the parabola formula, \(w_1\) is calculated by formula (3), and \(h_2\) and \(w_2\) are calculated by formula (1) with points \((NT_1, PT_1)\) where \(w = w_1 + w_2\), \(h = h_1/(w_1 + w_2)h_1 + w_2/(w_1 + w_2)h_2\).
   - In other cases, \(h\) is calculated by the parabola formula and \(w\) is calculated by formula (3).
6. If \(n = 2\) and \(p = 1\), then three conditions exist:
   - If \(n_y\) equals \(p_y\), then \(h_1\) is calculated by the parabola formula, \(w_1\) is calculated by formula (3), and \(h_2\) and \(w_2\) are calculated by formula (4), and \(h_2\) and \(w_2\) are calculated by formula (3).

### Table 5

The accuracy of SST grid datasets by different interpolators

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Criterion</th>
<th>Criterion-1</th>
<th>RMSE</th>
<th>Comparing inverse and initial isolines</th>
<th>Retrieving percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very good</td>
<td>CI/33.8</td>
<td>CI/33.7</td>
<td>CI/0.0599</td>
<td>CI</td>
<td>CI/61.9</td>
</tr>
<tr>
<td>Good</td>
<td>MS/85.0</td>
<td>MS/50.4</td>
<td>MS/0.0911</td>
<td>MS</td>
<td>MS/91</td>
</tr>
<tr>
<td>General</td>
<td>TIN-Q/5392.2</td>
<td>TIN-L/182.9</td>
<td>TIN-Q/0.1184</td>
<td>TIN-Q</td>
<td>TIN-Q/53.3</td>
</tr>
<tr>
<td>Low</td>
<td>TP/1725.3</td>
<td>TIN-Q/472.8</td>
<td>TIN-L/0.1253</td>
<td>TIN-L</td>
<td>TP/45.2</td>
</tr>
<tr>
<td>Bad</td>
<td>TIN-L/2988.2</td>
<td>TP/2588.7</td>
<td>TP/0.2779</td>
<td>TP</td>
<td>TP/136</td>
</tr>
</tbody>
</table>

*Just calculated 102 datasets in 1999.*
by formula (1) with two points (NT1 and NT2). \( w = w_1 + w_2, \ h = w_1/(w_1 + w_2)h_1 + w_2/(w_1 + w_2)h_2. \)

(B) If \( n_1 \neq p_1 \) and \( n_1 = n_2 \), then \( h_1 \) is calculated by the parabola formula, \( w_1 \) is calculated by formula (4), and \( h_2 \) and \( w_2 \) are calculated by formula (1) with two points (NT1 and PT1). \( w = w_1 + w_2, \ h = w_1/(w_1 + w_2)h_1 + w_2/(w_1 + w_2)h_2. \)

(C) In other cases, \( h \) is calculated by the parabola formula and \( w \) is calculated by formula (4).

(7) If \( n = 2 \) and \( p = 2 \), then seven conditions exist:

(A) If \( n_2 = n_1, p_2 = p_1, \) and \( n_1 \neq p_1 \), then \( h \) and \( w \) are calculated by formula (1) with two points (NT1 and PT1).

(B) If \( n_2 \neq n_1, n_1 \neq p_1, \) and \( p_1 = p_2 \), then \( h_1 \) is calculated by the parabola formula with three points (NT1, PT1, PT2), \( w_1 \) is calculated by formula (3), and \( h_2 \) and \( w_2 \) are calculated by formula (1) with two points (NT1 and PT2). \( w = w_1 + w_2, \ h = w_1/(w_1 + w_2)h_1 + w_2/(w_1 + w_2)h_2. \)

(C) If \( n_2 = n_1, n_1 \neq p_1, \) and \( p_1 \neq p_2 \), then \( h_1 \) is calculated by the parabola formula with three points (NT1, NT2, and PT1), \( w_1 \) is calculated by formula (4), and \( h_2 \) and \( w_2 \) are calculated by formula (1) with two points (NT1 and PT2). \( w = w_1 + w_2, \ h = w_1/(w_1 + w_2)h_1 + w_2/(w_1 + w_2)h_2. \)

(D) If \( n_2 \neq n_1, n_1 = p_1, \) and \( p_1 \neq p_2 \), then \( h_1 \) is calculated by the parabola formula with three points (NT1, NT2, and PT1), \( w_1 \) is calculated by formula (4), \( h_2 \) is calculated by the parabola formula with three points (NT1, PT1, and PT2), and \( w_2 \) is calculated by formula (3). \( w = w_1 + w_2, \ h = w_1/(w_1 + w_2)h_1 + w_2/(w_1 + w_2)h_2. \)

(E) If \( n_2 \neq n_1, n_1 \neq p_1, \) and \( p_1 \neq p_2 \), then \( h_1 \) is calculated by the parabola formula with three points (NT1, NT2, and PT1), \( w_1 \) is calculated by formula (4), \( h_2 \) is calculated by the parabola formula with three points (NT1, PT1, and PT2), and \( w_2 \) are calculated by formula (3), and \( h_3 \) and \( w_3 \) are calculated by formula (1) with two points (NT1, and PT1). \( W = w_1 + w_2 + w_3, \ h = w_1/(w_1 + w_2 + w_3)h_1 + w_2/(w_1 + w_2 + w_3)h_2 + w_3/(w_1 + w_2 + w_3)h_3. \)

(F) If \( n_2 = n_1, n_1 = p_1, \) and \( p_1 \neq p_2 \), then \( h \) is calculated by the parabola formula with three points (NT1, PT1, and PT2) and \( w \) is calculated by formula (3).

(G) If \( n_2 \neq n_1, n_1 = p_1, \) and \( p_1 = p_2 \), then \( h \) is calculated by the parabola formula with three points (NT1, NT2, and PT1) and \( w \) is calculated by formula (4).

References


