A Modified Interpolation Method for Surface Total Nitrogen in the Bohai Sea

YONGZHI LIU

Physical Oceanography Laboratory, Ocean University of China, Qingdao, China

JIE YU

Physical Oceanography Laboratory, Ocean University of China, Qingdao, and Guangdong Province Key Laboratory for Coastal Ocean Variation and Disaster Prediction Technologies, College of Ocean and Meteorology, Guangdong Ocean University, Guangzhou, China

YOULI SHEN

Physical Oceanography Laboratory, Ocean University of China, Qingdao, and Beihai Marine Environmental Monitoring Center Station, State Oceanic Administration, BeiHai, China

XIANQING LV

Physical Oceanography Laboratory, Ocean University of China, Qingdao, China

(Manuscript received 30 November 2015, in final form 2 March 2016)

ABSTRACT

A modified Cressman interpolation method (MCIM) is presented for the routine monitoring data of total nitrogen (TN) in the Bohai Sea to reduce interpolation errors by decreasing the influence radius and introducing background value. In twin experiments, two prescribed distributions are successfully estimated by MCIM with lower interpolation errors than the traditional Cressman interpolation method (TCIM) and the kriging method. In practical experiments, cross validation is applied to evaluate the interpolation results for four quarters in 2009 and 2010. Practical experimental results show that the interpolation results obtained with MCIM are greatly improved and can describe the spatial distribution characteristics of TN in the Bohai Sea with lower mean absolute error than the kriging method.

1. Introduction

Sufficient and accurate pollutant data are essential for spatial analysis and management of pollution (Jeffrey et al. 2001). The fact that pollutant data are typically insufficient, however, makes it difficult to analyze the environmental problems. In particular, such data may 1) be recorded for discrete periods, not spanning the entire time period; 2) contain short intermittent periods where data have not been recorded; 3) contain either systematic or random errors (Peck 1997), and 4) contain some sparsely distributed monitor stations. Sparse data should be integrated to obtain regular grid data.

Therefore, spatial interpolation techniques are essential for estimating biophysical variables for unsampled locations (Li and Heap 2011). Many interpolation methods such as inverse distance weighting and kriging interpolation have been widely used in multidisciplinary subjects (Goodin et al. 1979; Franke 1982; Willmott et al. 1985; Biau et al. 1999; Orùs et al. 2005; Largueche 2006; Kebaili Bargaoui and Chebbi 2009). They all share the same general estimation formula as follows:

$$\widehat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i), \tag{1}$$

where $\widehat{z}(x_0)$ is the estimated value of an attribute at the point of x_0, $z(x_i)$ is the observed value at sampled point x_i, λ_i is the weight function assigned to each $z(x_i)$ value, and n represents the number of ambient sampled data points used for the estimation (Webster and Oliver 2001).

Corresponding author address: Xianqing Lv, Physical Oceanography Laboratory, Ocean University of China, 238 Songling Road, Qingdao 266003, China.
E-mail: xqinglv@ouc.edu.cn

DOI: 10.1175/JTECH-D-15-0250.1

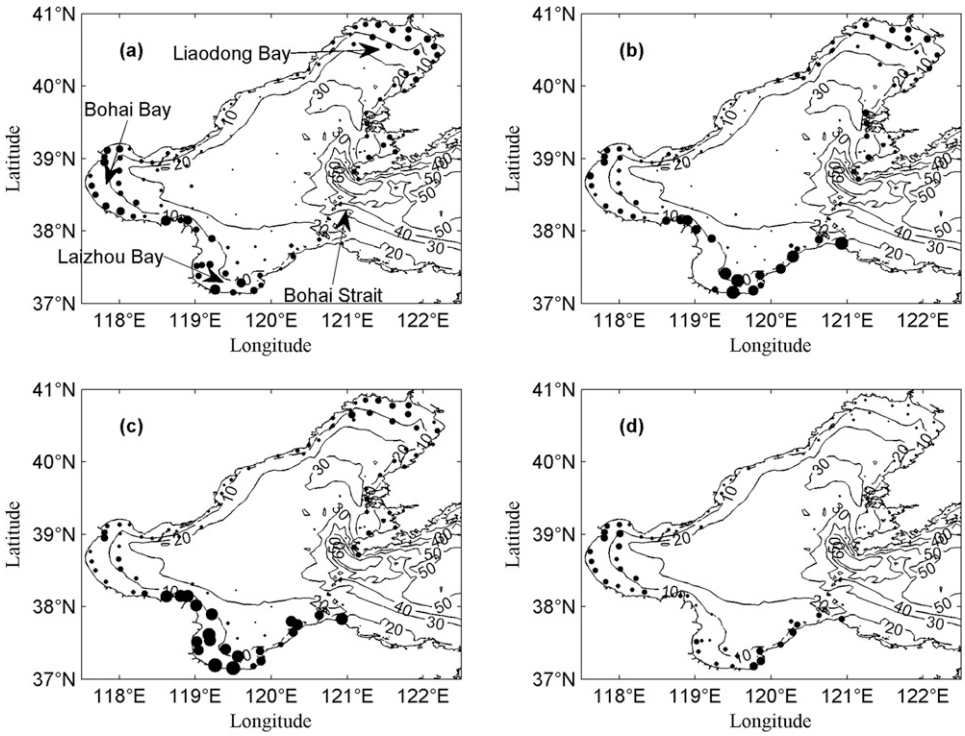

FIG. 1. Distribution of routine monitoring stations in the Bohai Sea in (a) May 2009, (b) August 2009, (c) October 2009, and (d) May 2010. The size of the dot is in direct proportion to the value of TN concentration.

With the increasing application of spatial interpolation methods, there is a growing concern about their accuracy and precision (Hartkamp et al. 1999). The kriging and cokriging methods provide a description of data spatial structure and variance estimation, but they are time consuming and cumbersome (Kravchenko and Bullock 1999) because both need inverse-matrix calculations.As a modified inverse distance weighting method, Cressman interpolation (Cressman 1959) is flexible and easy to implement. However, when applied to the datasets with large distances between grid points, the traditional Cressman interpolation method (TCIM) should be modified to reduce interpolation errors (Gu 2003; Kravchenko 2003; Tongsuk and Kanok-Nukulchai 2004; Physick et al. 2007; Sampson et al. 2013; Huang et al. 2014). In our study, a modified Cressman interpolation method (MCIM) is presented to interpolate total nitrogen (TN) data in the Bohai Sea.

This paper is organized as follows. In section 2 we introduce the routine monitoring of TN data and the interpolation methods in the Bohai Sea. Then twin

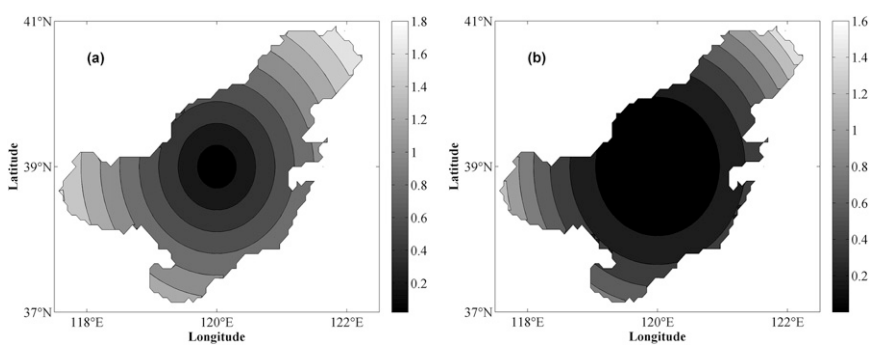

FIG. 2. Two prescribed distributions (mg L^{-1}): (a) conical surface distribution and (b) revolution parabolic surface distribution.

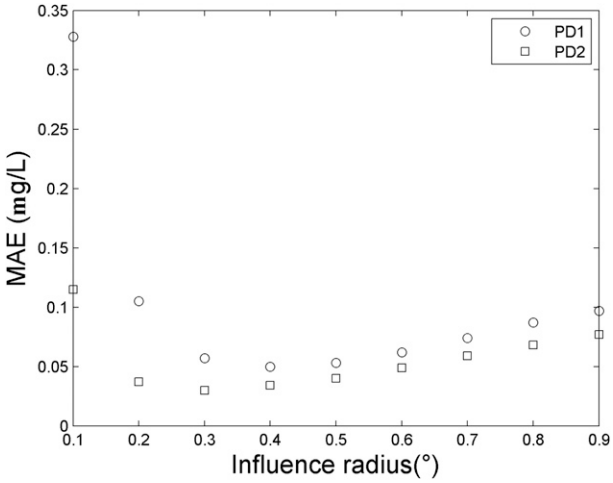

FIG. 3. Difference between interpolation results with MCIM and the prescribed distributions in the twin experiments among different influence radii.

experiments are carried out in section 3. In section 4 practical experiments are performed in four different quarters. Section 5 provides the summary and main conclusions.

2. Data and method

The observation data used in this paper were obtained from the marine environment monitoring data in the Bohai Sea and the north Yellow Sea from 2009 to 2010. The marine environment monitoring is conducted by the North China Sea Environmental Monitoring Center, State Oceanic Administration. The monitoring is implemented four times a year: February, May, August, and October every year. The purpose is to investigate the temporal/spatial distribution and variation of temperature, salinity, and various pollution factors in the Bohai Sea and the north Yellow Sea, to assess and evaluate the current situation and the variation trend of the marine environment, and to diagnose the environmental pollution problems. The monitoring items include pH, phosphate, nitrate, chemical oxygen demand (COD), petroleum, and so on.

TABLE 1. MAEs between interpolation results and prescribed distributions.

Distribution	MAE (mg L^{-1}) MCIM	TCIM	Kriging	Kriging with background value
PD1	0.050	0.187	0.071	0.071
PD2	0.030	0.164	0.069	0.069

The routine monitoring stations are shown in Fig. 1, in which the size of dot is in direct proportion to the value of TN concentration in several months, such as May 2009, August 2009, October 2009, and May 2010. From Fig. 1 it can be concluded that stations in the central part of Bohai Sea are sparsely distributed and that the concentrations of TN over there are relatively low, while in coastal areas there are more stations with much higher concentrations.

The weight function of Cressman interpolation is expressed as follows:

$$\lambda_i = \frac{\varpi_i}{\sum \varpi_i}, \tag{2}$$

where

$$\varpi_i = \frac{R^2 - r_i^2}{R^2 + r_i^2}. \tag{3}$$

Herein R is the influence radius and r_i is the distance from x_0 to the sampling point x_i. Conventionally, because of the sparse distribution of routing monitoring stations, the influence radius should be large enough to guarantee that the concentration of TN in each grid is interpolated, which may increase the interpolation errors in many regions. In the present work, a modified Cressman method is proposed to reduce the interpolation errors, and the procedure is implemented as follows.

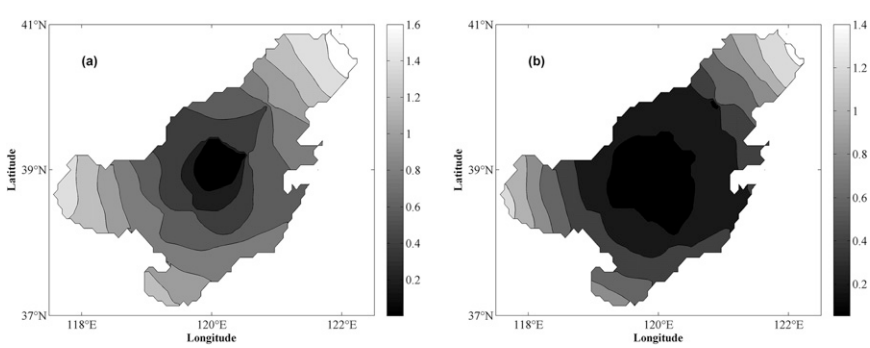

FIG. 4. Interpolation results with MCIM in twin experiments (mg L^{-1}): (a) PD1and (b) PD2.

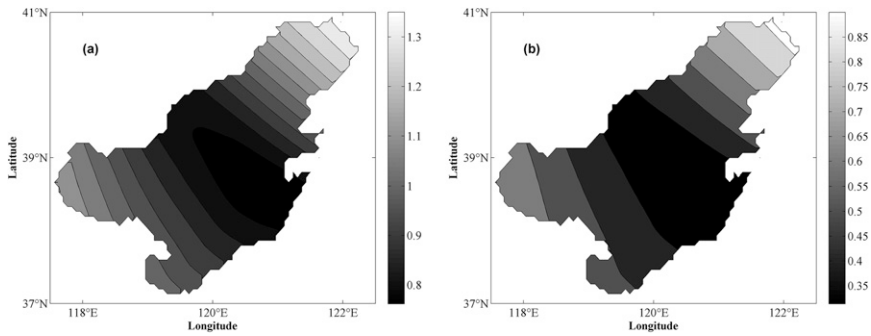

FIG. 5. Interpolation results with TCIM in twin experiments (mg L^{-1}): (a) PD1 and (b) PD2.

1) Concentrations with low values (≤ 0.1 mg L^{-1}) in the central part of the Bohai Sea are averaged as the background value.
2) The background value is subtracted from all observed concentrations.
3) The aforementioned processed data are used for traditional Cressman interpolation (the specific influence radius will be described in section 3).
4) The background value is added to concentrations calculated by the traditional Cressman method, and then concentrations in the whole field are obtained.

The TCIM and the kriging method are employed for comparison. Developed by Matheron (1963), the kriging method calculates the optimal unbiased estimate at every interpolated point using the structural properties of the semivariogram and initial set of data values (David 1977). In the cokriging method, auxiliary (secondary) variables are introduced to improve the estimation of primary variables. Because of the lack of auxiliary variables, only the kriging method is discussed in this paper.

The kriging method is implemented with the help of the mGstat toolbox (Hansen 2004), with parameters set as follows:

1) A semivariogram model is specified as the spherical semivariogram model, with a range of 10° and a sill of 1°.
2) A universal kriging method is selected, with a polynomial trend for each dimension (direction) set as one order.

We also introduce the background value in the kriging method similar to MCIM for comparison in the twin experiments.

3. Twin experiments

To evaluate the effectiveness of the modified Cressman interpolation method, two types of distribution, including the conical surface distribution and the revolution parabolic surface distribution, are prescribed according to the characteristics of pollutant distribution in the Bohai Sea. Hereafter, the two distributions, which are shown in Figs. 2a and 2b, will be referred to as PD1 and PD2, respectively.

In TCIM, the influence radius is set as 1.5° initially. If there are at least two stations within the influence radius of a grid, the value of this grid is calculated; otherwise, 0.1° is added to the influence radius. As mentioned by Gandin (1965) and Eddy (1967), the choice of weight functions and influence radius depends on the field statistics. Stephens and Stitt (1970) find that the optimum influence radius primarily depends upon the average station separation. In MCIM, since the TN concentration in the central part of

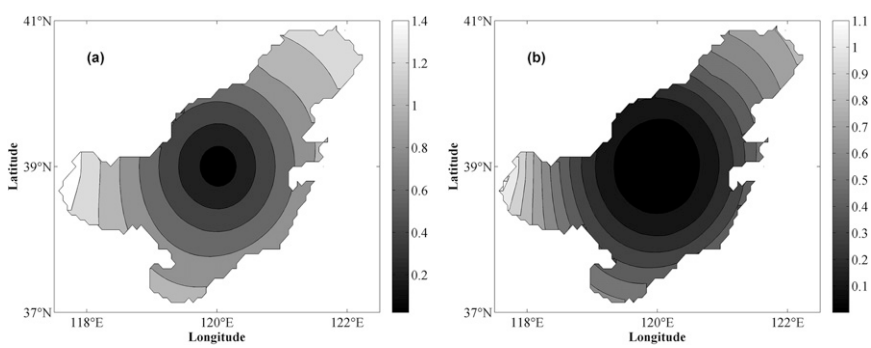

FIG. 6. Interpolation results with the kriging method in twin experiments (mg L^{-1}). (a) PD1 and (b) PD2.

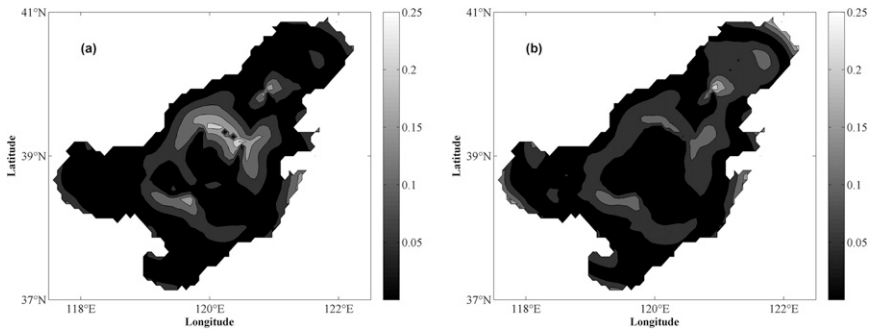

FIG. 7. Spatial distribution of absolute error with MCIM in twin experiments (mg L^{-1}). (a) PD1and (b) PD2.

the Bohai Sea is relatively low with rare routine monitoring stations located there, the TN concentration there can be treated as the background value. Therefore, the influence radius can be selected relatively small, ranging from 0.1° to 0.9°, which is much smaller than that of TCIM, making it possible to reduce the interpolation errors.

The mean absolute error (MAE) between the observed and predicted values at sampling points is often used as a precision indicator of interpolation methods (Burrough and McDonnell 1998). The MAE between the interpolation results with MCIM and the prescribed distributions among different influence radii is depicted in Fig. 3. It can be deduced that the optimum influence radius with MCIM varies among different distributions. For the conical surface distribution, the optimum influence radius is 0.4°, while for the revolution parabolic surface distribution it is 0.3°.

Table 1 shows the MAE between the interpolation results and the prescribed distributions. The influence radius with MCIM is selected as the optimal one as shown in Fig. 3. From Table 1, it can be deduced that the MCIM can yield a much better result than TCIM and the kriging method. Although the background value is introduced in the kriging method, the interpolation results with the kriging method are not improved.

Figure 4 shows the interpolation results calculated by MCIM with the optimum influence radius, and those with TCIM and kriging are depicted in Figs. 5 and 6, respectively. With the prescribed distributions depicted in Fig. 2 as reference, it can be deduced that estimation with MCIM is close to the prescribed distribution, though deviations exist in some regions. The interpolation results with TCIM present a gradual and smooth distribution that, however, is incapable of recovering the prescribed spatial distributions. The interpolation results with the kriging method present a similar spatial distribution to the prescribed distributions, but on the whole the interpolated values are relatively low and the MAE is relatively large. The spatial distribution of absolute error with MCIM and the kriging method are depicted in Figs. 7 and 8, respectively. It can be deduced that the interpolation errors with MCIM are smaller than those with the kriging method.

4. Practical experiments

As described in section 3, MCIM is an effective method to recover the prescribed distribution designed in accordance with the characteristics of pollutant distribution in

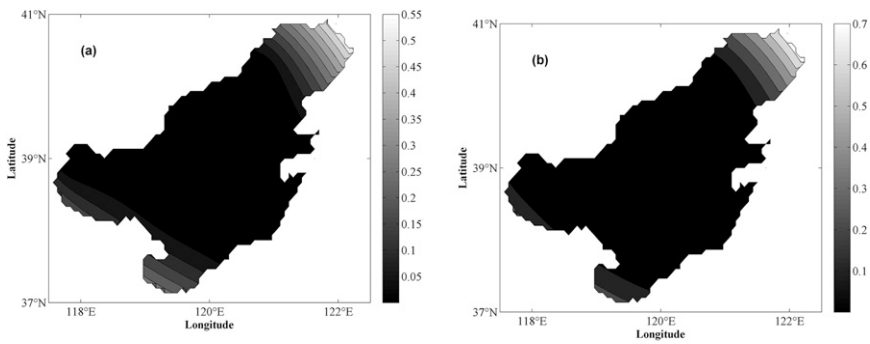

FIG. 8. Spatial distribution of absolute error with the kriging method in twin experiments (mg L^{-1}). (a) PD1 and (b) PD2.

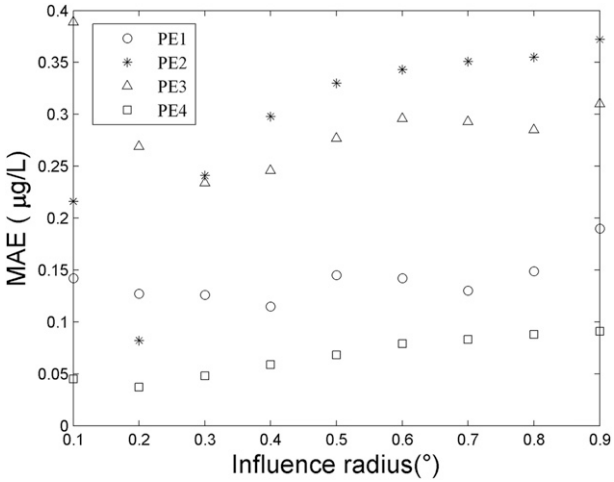

FIG. 9. Difference between interpolation results and data with MCIM in practical experiments.

the Bohai Sea. This encourages us to further apply this method to practical cases in the Bohai Sea.

Cross validation is applied to assess the interpolation result (Holdaway 1996). In this study, 10-fold cross validation (Kohavi 1995) is applied. The routine monitoring dataset is randomly split into 10 subsets, of which each subset is selected as set A in sequence, while the other data are referred to as set B. Data in set B are used for interpolation, while the interpolation results are assessed with the data in set A. The cross validation is repeated 10 times in each practical experiment. The MAE between the interpolation results with MCIM and data in set A is depicted in Fig. 9, in which PE1–PE4 represent the results for May 2009, August 2009, October 2009, and May 2010, respectively.

TABLE 2. Interpolation errors for test datasets in four quarters in 2009 and 2010.

Quarter	MAE (mg L^{-1}) MCIM	TCIM	Kriging
PE1	0.115	0.288	0.283
PE2	0.082	0.151	0.190
PE3	0.234	0.301	0.236
PE4	0.037	0.115	0.098

The surface TN data in four quarters in 2009 and 2010 are interpolated to the horizontal grids, with the same influence radius set in section 3. The interpolation errors for data after interpolation are given in Table 2, in which PE1–PE4 represent the results for May 2009, August 2009, October 2009 and May 2010, respectively. The influence radius with MCIM in Table 2 is set as the optimal one as shown in Fig. 9. The interpolation results with MCIM are depicted in Fig. 10, and those with TCIM and the kriging method are depicted in Figs. 11 and 12, respectively.

From Figs. 10–12 and Table 2, the following conclusions can be deduced:

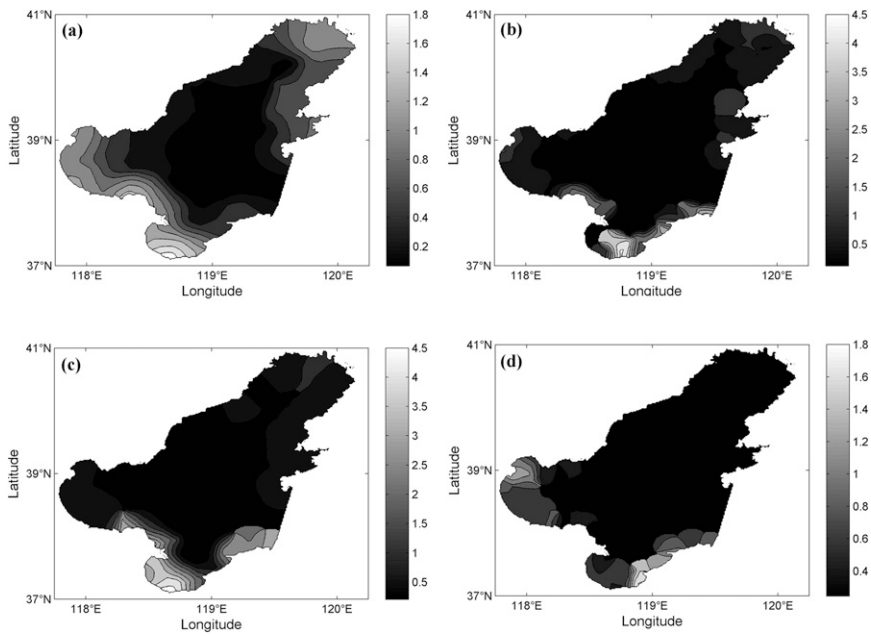

FIG. 10. Interpolation results in four quarters in 2009 and 2010 with MCIM in practical experiments (mg L^{-1}): (a) PE1, (b) PE2, (c) PE3, and (d) PE4.

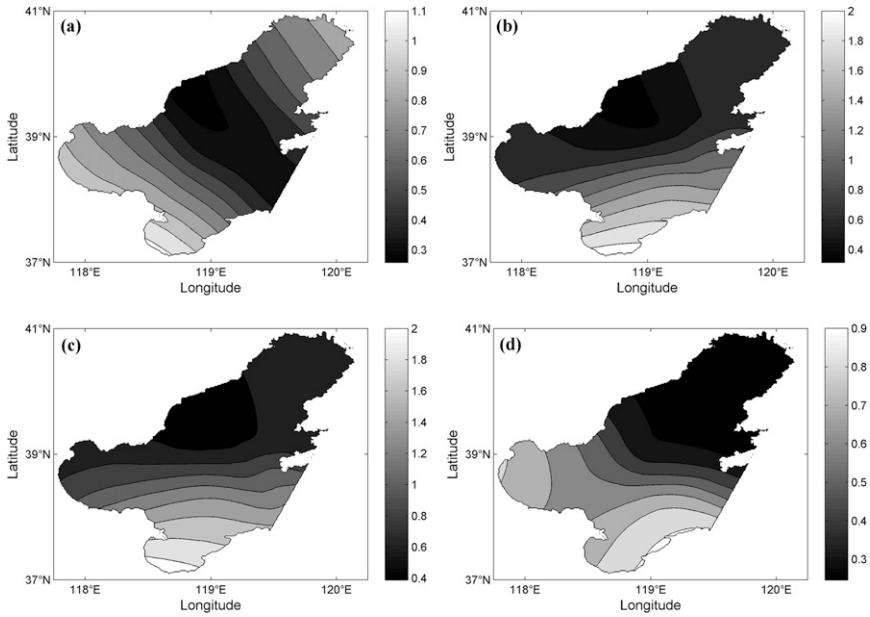

FIG. 11. Interpolation results in four quarters in 2009 and 2010 with TCIM in practical experiments (mg L^{-1}): (a) PE1, (b) PE2, (c) PE3, and (d) PE4.

1) The interpolation results with MCIM present lower errors than those with TCIM and the kriging method. Taking PE1 as an example, the MAE with MCIM is 0.115 mg L^{-1}, while those with TCIM and the kriging method are 0.288 and 0.458 mg L^{-1}, respectively.
2) The optimum influence radius of MCIM is different among different quarters. For example, it is 0.2° for August 2009, while it is 0.3° for October 2009.
3) In spatial pattern analysis, TCIM presents relatively higher concentrations in most regions due

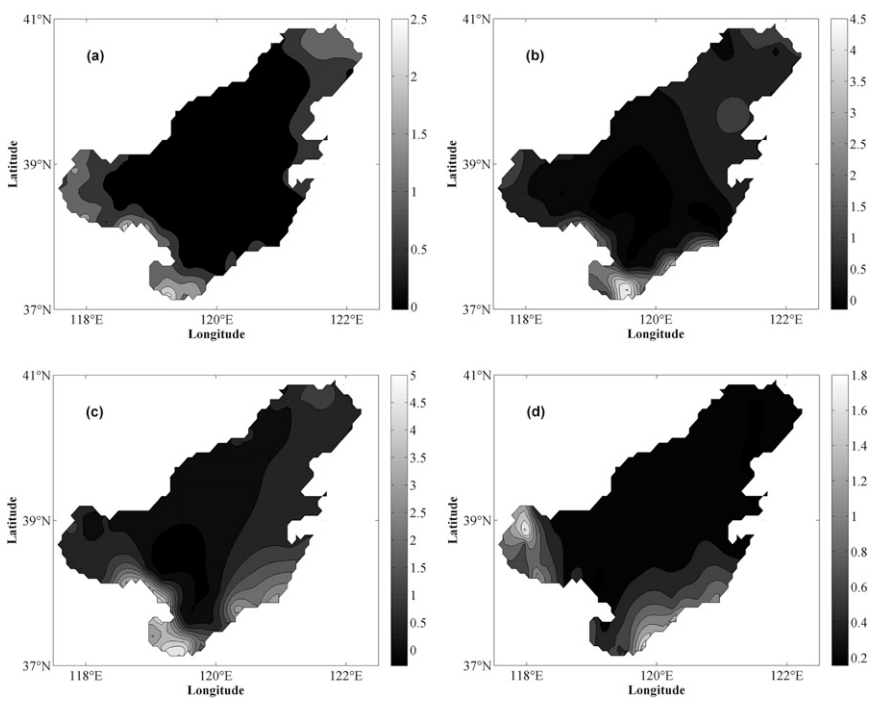

FIG. 12. Interpolation results in four quarters in 2009 and 2010 with the kriging method in practical experiments (mg L^{-1}): (a) PE1, (b) PE2, (c) PE3, and (d) PE4.

to the overlarge influence radius. The kriging method presents a smooth distribution with a larger MAE than MCIM, while MCIM presents a smaller MAE

5. Conclusions

This paper presents a modified Cressman interpolation method for routine monitoring data of TN in the Bohai Sea. In this method, the influence radius is decreased by introducing background value to reduce interpolation errors. In twin experiments, two prescribed distributions designed in accordance with characteristics of pollutant distribution in the Bohai Sea are successfully recovered using MCIM, in which MAEs are much lower than those with TCIM and the kriging method. Therefore, this method can be feasible and effective in calculating the spatial distribution in the Bohai Sea.

This study applied 10-fold cross validation (Kohavi 1995; Holdaway 1996) to assess the interpolation result in practical experiments. Experimental results of surface TN concentration indicate that the interpolation results with MCIM could accurately capture the spatial distribution characteristics of TN in the Bohai Sea, with a much lower MAE than those with TCIM and the kriging method.

Acknowledgments. Partial support for this research was provided by the Natural Science Foundation of Shandong Province of China through Grant ZR2014DM017, the State Ministry of Science and Technology of China through Grant 2013AA122803, and the Ministry of Education's 111 Project through Grant B07036.

REFERENCES

Biau, G., E. Zorita, H. von Storch, and H. Wackernagel, 1999: Estimation of precipitation by kriging in the EOF space of the sea level pressure field. *J. Climate*, **12**, 1070–1085, doi:10.1175/1520-0442(1999)012<1070:EOPBKI>2.0.CO;2.

Burrough, P. A., and R. A. McDonnell, 1998: *Principles of Geographical Information Systems*. Oxford University Press, 333 pp.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, doi:10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

David, M., 1977: *Geostatistical Ore Reserve Estimation*. Developments in Geomathematics, Vol. 2, Elsevier, 364 pp.

Eddy, A., 1967: The statistical objective analysis of scalar data fields. *J. Appl. Meteor.*, **6**, 597–609, doi:10.1175/1520-0450(1967)006<0597:TSOAOS>2.0.CO;2.

Franke, R., 1982: Scattered data interpolation: Test of some methods. *Math. Comput.*, **38**, 181–200.

Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields*. R. Hardin, Ed., Israel Program for Scientific Translations, 242 pp.

Goodin, W. R., G. J. McRa, and J. H. Seinfeld, 1979: A comparison of interpolation methods for sparse data: Application to wind and concentration fields. *J. Appl. Meteor.*, **18**, 761–771, doi:10.1175/1520-0450(1979)018<0761:ACOIMF>2.0.CO;2.

Gu, L., 2003: Moving kriging interpolation and element-free Galerkin method. *Int. J. Numer. Methods Eng.*, **56**, 1–11, doi:10.1002/nme.553.

Hansen, T. M., 2004: mGstat: A geostatistical Matlab toolbox. [Available online at http://mgstat.sourceforge.net.]

Hartkamp, A. D., K. De Beurs, A. Stein, and J. W. White, 1999: Interpolation techniques for climate variables. NRG-GIS Series 99-01, CIMMYT, 26 pp.

Holdaway, M. R., 1996: Spatial modeling and interpolation of monthly temperature using kriging. *Climate Res.*, **6**, 215–225, doi:10.3354/cr006215.

Huang, C., X. Zheng, A. Tait, Y. Dai, C. Yang, Z. Chen, T. Li, and Z. Wang, 2014: On using smoothing spline and residual correction to fuse rain gauge observations and remote sensing data. *J. Hydrol.*, **508**, 410–417, doi:10.1016/j.jhydrol.2013.11.022.

Jeffrey, S. J., J. O. Carter, K. B. Moodie, and A. R. Beswick, 2001: Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environ. Modell. Software*, **16**, 309–330, doi:10.1016/S1364-8152(01)00008-1.

Kebaili Bargaoui, Z., and A. Chebbi, 2009: Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. *J. Hydrol.*, **365**, 56–73, doi:10.1016/j.jhydrol.2008.11.025.

Kohavi, R., 1995: A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI-95: Proceedings of the 14th International Joint Conference on Artificial Intelligence*, C. S. Mellish, Ed., Vol. 2, Morgan Kaufmann Publishers Inc., 1137–1145.

Kravchenko, A. N., 2003: Influence of spatial structure on accuracy of interpolation methods. *Soil Sci. Soc. Amer. J.*, **67**, 1564–1571, doi:10.2136/sssaj2003.1564.

——, and D. G. Bullock, 1999: A comparative study of interpolation methods for mapping soil properties. *Agron. J.*, **91**, 393–400, doi:10.2134/agronj1999.00021962009100030007x.

Largueche, F. Z. B., 2006: Estimating soil contamination with kriging interpolation method. *Amer. J. Appl. Sci.*, **3**, 1894–1898, doi:10.3844/ajassp.2006.1894.1898.

Li, J., and A. D. Heap, 2011: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecol. Inf.*, **6**, 228–241, doi:10.1016/j.ecoinf.2010.12.003.

Matheron, G., 1963: Principles of geostatistics. *Econ. Geol.*, **58**, 1246–1266, doi:10.2113/gsecongeo.58.8.1246.

Orùs, R., M. Hernández-Pajares, J. M. Juan, and J. Sanz, 2005: Improvement of global ionospheric VTEC maps by using kriging interpolation technique. *J. Atmos. Sol.-Terr. Phys.*, **67**, 1598–1609, doi:10.1016/j.jastp.2005.07.017.

Peck, E. L., 1997: Quality of hydrometeorological data in cold regions. *J. Amer. Water Resour. Assoc.*, **33**, 125–134, doi:10.1111/j.1752-1688.1997.tb04089.x.

Physick, W. L., M. E. Cope, S. Lee, and P. J. Hurley, 2007: An approach for estimating exposure to ambient concentrations. *J. Exposure Sci. Environ. Epidemiol.*, **17**, 76–83, doi:10.1038/sj.jes.7500523.

Sampson, P. D., M. Richards, A. A. Szpiro, S. Bergen, L. Sheppard, T. V. Larson, and J. D. Kaufman, 2013: A regionalized national universal kriging model using Partial Least Squares regression for

estimating annual $PM_{2.5}$ concentrations in epidemiology. *Atmos. Environ.*, **75**, 383–392, doi:10.1016/j.atmosenv.2013.04.015.

Stephens, J. J., and J. M. Stitt, 1970: Optimum influence radii for interpolation with the method of successive corrections. *Mon. Wea. Rev.*, **98**, 680–687, doi:10.1175/1520-0493(1970)098<0680:OIRFIW>2.3.CO;2.

Tongsuk, P., and W. Kanok-Nukulchai, 2004: Further investigation of element-free Galerkin method using Moving Kriging interpolation. *Int. J. Comput. Methods*, **1**, 345, doi:10.1142/S0219876204000162.

Webster, R., and M. A. Oliver, 2001: *Geostatistics for Environmental Scientists*. Wiley, 271 pp.

Willmott, C. J., C. M. Rowe, and W. D. Philpot, 1985: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *Amer. Cartogr.*, **12**, 5–16, doi:10.1559/152304085783914686.