AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2013JC009577

Special Section:
Pacific-Asian Marginal Seas

Key Points:
- Three sources of uncertainty in gridded pCO_2 data are quantified
- Remote sensing-derived and underway pCO_2 data are combined
- The uncertainties in gridded pCO_2 data from the East China Sea are estimated

Correspondence to:
M. Dai,
mdai@xmu.edu.cn

Citation:
Wang, G., M. Dai, S. S. P. Shen, Y. Bai, and Y. Xu (2014), Quantifying uncertainty sources in the gridded data of sea surface CO_2 partial pressure, *J. Geophys. Res. Oceans*, *119*, 5181–5189, doi:10.1002/2013JC009577.

Received 5 NOV 2013
Accepted 21 JUL 2014
Accepted article online 24 JUL 2014
Published online 15 AUG 2014

Quantifying uncertainty sources in the gridded data of sea surface CO_2 partial pressure

Guizhi Wang[1,2], Minhan Dai[1,2], Samuel S. P. Shen[3], Yan Bai[4], and Yi Xu[1]

[1]State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen, China, [2]College of Ocean and Earth Science, Xiamen University, Xiamen, China, [3]Department of Mathematics and Statistics, San Diego State University, San Diego, California, USA, [4]State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, State Oceanic Administration, Hangzhou, China

Abstract The bulk uncertainty in the gridded sea surface pCO_2 data is crucial in assessing the reliability of the CO_2 flux estimated from measurements of air-sea pCO_2 difference, because atmospheric pCO_2 are relatively homogeneous and well defined. The bulk uncertainty results from three different sources: analytical error (E_m), spatial variance (σ_s^2), and the bias from undersampling (σ_u^2). Common uncertainty quantification by standard deviation may mix up the different sources of uncertainty. We have established a simple procedure to determine these three sources of uncertainty using remote sensing-derived and field-measured pCO_2 data. E_m is constrained by the analytical method and data reduction procedures. σ_s^2 is derived from the remotely sensed pCO_2 field. σ_u^2 is determined by spatial variance and the effective number of observations, considering, for the first time, the geometric bias introduced by pCO_2 sampling. This approach is applied to $1^\circ \times 1^\circ$ gridded pCO_2 data collected from the East China Sea. We demonstrate that the spatial distribution of these biases is uneven and that none of them follow the same spatial trend as the standard deviation. σ_s^2 contributes the most to the uncertainty in gridded pCO_2 data over those grid boxes with good sampling coverage, while σ_u^2 dominates the total uncertainty in the grid boxes with poor sampling coverage. Application of this procedure to other parts of the global ocean will help to better define the inherent spatial variability of the pCO_2 field and thus better interpolate and/or extrapolate pCO_2 data, and eventually better constrain air-sea CO_2 fluxes.

1. Introduction

Closing the carbon budget and reducing the uncertainty in estimations of air-sea CO_2 fluxes, to thereby better constrain the temporal-spatial variability of the ocean carbon sink and ultimately better predict climate system changes, have been the goals of many international efforts [e.g., *Le Quéré et al*., 2009; *Wanninkhof et al*., 2013]. One of the most promising methods in estimating air-sea CO_2 fluxes is based on measurements of air-sea pCO_2 (partial pressure of carbon dioxide) difference, or ΔpCO_2 [*Takahashi et al*., 2009, and references therein], where pCO_2 in the ocean surface water is the primary determinant, because pCO_2 in the atmosphere is relatively homogeneous and well defined.

After *Tan et al*. [1990], the first compilation of sea surface pCO_2 data measured since 1972, tremendous efforts have been devoted to acquiring surface ocean pCO_2 data, primarily through underway measurements around the world's oceans [e.g., *Bakker et al*., 2013; *Pfeil et al*., 2013; *Sabine et al*., 2013; *Wanninkhof et al*., 2013]. When estimating the air-sea CO_2 fluxes, observational CO_2 data are typically aggregated onto grid boxes to establish gridded data [e.g., *Takahashi et al*., 1997, 2002, 2009]. Essentially, three error sources with different implications may contribute to uncertainty in gridded pCO_2 data. One is the analytical error in the pCO_2 determination and the associated environmental parameters used in deriving pCO_2. Second, the ocean surface pCO_2 may be spatially inhomogeneous, and thus may have nonuniform variances within an investigation area, especially in the grid boxes over coastal regions and other ocean boundaries or fronts [*Sweeney et al*., 2002]. Any interpolation within a particular spatial domain/grid box can thus be subject to error. Finally, errors can also result from extrapolation to areas without any actual measurements or undersampling, since sampling sites may be unevenly distributed over the entire investigation area, and the scarcity of sampling is common. In the Takahashi climatology, these error terms have been carefully examined [*Takahashi et al*., 2009]. They estimated that one standard deviation around a mean ΔpCO_2 for their $4^\circ \times 5^\circ$

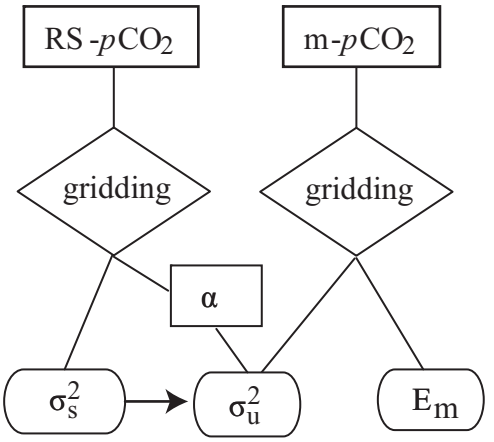

Figure 1. Flowchart of the uncertainty estimate for pCO_2 data. The flow direction is downward if not indicated by an arrow. RS-pCO_2 represents remote sensing-derived pCO_2 data, and m-pCO_2 represents field-measured pCO_2 data. E_m is the analytical error, σ_s^2 is the spatial variance, and σ_u^2 represents the undersampling variance. See the text for details.

boxes was about ±10 μatm, on average, across global oceans. Based on the total number of measurements across the number of grid boxes, the average errors for monthly mean values of seawater pCO_2 were estimated to be ±0.8 μatm. Additionally, they evaluated the systematic biases in surface water pCO_2 due to undersampling, and their interpolation method was based on a two-dimensional diffusion-advection transport equation [*Takahashi et al.*, 1995]. A composite of the biases was evaluated by comparing the SST in the study and the climatological SST field. However, it remains challenging to differentiate the three identified sources of error in any spatial horizon. It is even more difficult to quantify the error sources at given temporal and spatial scales with which the observational data are associated. In some other studies, particularly over the coastal ocean, because of insufficient data coverage, data on CO_2 fluxes are often reported with an uncertainty of the average pCO_2 in each grid box quantified by the standard deviation of all the data in the box [e.g., *Zhai et al.*, 2013]. This simplified uncertainty quantification may mix up the different sources of uncertainty and, moreover, lend difficulty to data interpolation/extrapolation beyond the sampling sites, or to optimizing strategies for designing field observations that reduce the uncertainty. This study is a step forward from the previous simplified uncertainty quantification and explores the feasibility of estimating the contribution of each source of uncertainty in gridded pCO_2 data and hence helps assess the reliability of the CO_2 fluxes. Our aim is to explore three sources of uncertainty in gridded pCO_2 data: (i) analytical error (E_m), determined by the analytical method and data reduction procedures, (ii) spatial variance (σ_s^2), which represents characteristics of pCO_2 spatial variations in a given region, and (iii) the bias resulting from undersampling (σ_u^2), determined by spatial variations and the effective number of sampling observations. Using the gridded pCO_2 measurement data over the East China Sea, we attempt to demonstrate an approach to quantitatively differentiate the three sources of uncertainty. Remote sensing pCO_2 data are used to calculate the spatial variability and the correlation among observations in individual grid boxes. Given the rapidly growing applications of remote sensing techniques in mapping pCO_2 and constraining pCO_2 spatial variability [e.g., *Stephens et al.*, 1995; *Nelson et al.*, 2001; *Ono et al.*, 2004; *Sarma et al.*, 2006; *Zhu et al.*, 2009; *Hales et al.*, 2012; *Jo et al.*, 2012], our approach to estimating uncertainty in gridded pCO_2 data should have numerous applications.

2. Method to Estimate Sources of Error in Gridded pCO_2 Data

For a given pCO_2 data set, the three sources of error in its gridded pCO_2 can be estimated with aid of a concurrent high-spatial-resolution satellite-derived pCO_2 data set for the same region. Our procedure for estimating the three sources of uncertainty is summarized in a simplified way in the flowchart in Figure 1. First, the analytical error, E_m, is assessed by aggregating all the errors introduced in the measurement and data reduction of pCO_2. Next is the spatial variance, σ_s^2, which is determined using pCO_2 data retrieved from satellite remote sensing that has relatively high spatial coverage. Last is the bias resulting from undersampling, σ_u^2, determined by the spatial variance, a correlation factor, and the effective number of observational stations in a desired grid box. Here, the correlation factor is estimated by a least squares regression using the satellite remote sensing pCO_2 data. The effective number of observations reflects the spatial evenness of the sampling and is estimated by a subgridding technique. The calculations to determine σ_u^2 are similar to those used in calculating the sampling error variance for surface air temperature data [*Shen et al.*, 2007, 2012],

but differ in four aspects: (a) instead of one data set, we combine two independent pCO_2 data sets, one retrieved from high-spatial-resolution satellite data that enable better constraints of the spatial variations, and the other from in situ measurements, to describe σ_u^2; (b) we consider the spatial variation, not the temporal variation; (c) smoothing techniques are not applied, since spatial variations in pCO_2 may be great, especially in coastal oceans; and (d) the effective number, not the actual number, of observations is applied to reflect the uneven distribution of observations.

Described in detail, the following steps are taken to estimate the three sources of error in gridded in situ pCO_2.

2.1. Calculation of the Analytical Error (E_m)

The analytical error calculation follows the error propagation technique of *Taylor* [1982, 1997], which has been used in many studies [e.g., *Goñi et al*., 1998; *Bhattacharya et al*., 2002; *Lee et al*., 2005; *Wang et al*., 2010; *Han et al*., 2012]. Briefly, the analytical error for gridded pCO_2 data is estimated by transferring the analytical errors of all the in situ pCO_2 data in the grid box using the error propagation.

2.2. Calculation of the Spatial Variance (σ_s^2)

In a grid box with pCO_2 data retrieved from satellite data, the spatial variance is estimated by

$$\sigma_s^2 = \frac{1}{N}\sum_{i=1}^{N}\left(P_i - \overline{P}\right)^2, \tag{1}$$

where N is the number of satellite pCO_2 data in the grid box, P_i is the ith pCO_2, and $\overline{P}$ is the spatial simple average in the grid box. The minimum number for N is 4, to make the statistics meaningful and to make it consistent with the requirement in estimating the correlation factor in the calculation of σ_u^2 (see below).

2.3. Estimation of the Undersampling Variance (σ_u^2)

The uncertainty is given as

$$\sigma_u^2 = \left(\alpha \times {\sigma_s}^2\right)/(f \times K), \tag{2}$$

where α is a correlation factor, indicating the correlation among true pCO_2 in a desired grid box and better represented by the correlation among high-spatial-resolution satellite pCO_2 data, determined by a least squares regression. K is the number of underway pCO_2 observations in the grid box, and f is the fraction of subgrids with underway pCO_2 observations in the grid box. Here, $f \times K$ can be interpreted as the effective number of underway pCO_2 observations. Since the errors are estimated only for the grid boxes with the observed data, $f \times K$ is thus never equal to zero. This uncertainty reflects the geometric bias introduced by pCO_2 sampling in a grid box. For the same number of samples, the one with the widest coverage has the smallest errors while the samples from the same location are redundant and would result in large errors. This is an improvement on the method of *Shen et al*. [2007, 2012], which did not consider this geometric distribution factor.

2.3.1. Estimation of the Correlation Factor (α)

The correlation factor is

$$\alpha = 1 + \frac{1}{N}\sum_{\substack{i,j=1 \\ i\neq j}}^{N}\left\langle \frac{P_i - \overline{P}}{\sigma_s}\frac{P_j - \overline{P}}{\sigma_s} \right\rangle, \tag{3}$$

where $\langle\rangle$ represents the ensemble mean, and i and j are the observation number. σ_s is determined in equation (1). This formula shows that the correlation factor is determined by an average correlation between each pair of in situ observation points. Because the true pCO_2 values are unknown, the correlation factor cannot be directly calculated from equation (3) and is estimated by regression using high-resolution satellite-derived pCO_2 data, as follows. For a grid box with N observations ($N \geq 4$), the pCO_2 data of the N observations are taken as a statistical population. The population mean of the pCO_2 data in the grid box is

$$\overline{P}_N = \frac{1}{N}\sum_{i=1}^{N} P_i. \tag{4}$$

We randomly pick n observations from the population 5000 times ($1 \le n < N$). The sample mean of the n observations for each pick is

$$\overline{P}_n = \frac{1}{n}\sum_{i=1}^{n} P_{n,i}. \tag{5}$$

The mean square difference between the population mean and the sample mean is

$$D_n^2 = \frac{1}{5000}\sum_{n=1}^{5000}\left(\overline{P}_N - \overline{P}_n\right)^2. \tag{6}$$

The random sampling for 5000 times of pCO_2 data may have many repeated samples for a limited number of sampling observations in a grid box, so equation (6) represents well the exhaustive sample mean. Repeating this random picking for $n = 1, 2, 3, \ldots, N-1$ and two $(N-1)$-dimension vectors can be set up,

$$\mathbf{X} = \left(\frac{1}{n}\right), (n = 1,\ 2,\ 3, \ldots,\ N-1), \tag{7}$$

and

$$\mathbf{Y} = \left(\frac{D_n^2}{\sigma_s^2}\right), (n = 1,\ 2,\ 3, \ldots,\ N-1). \tag{8}$$

The correlation factor α is estimated by the least squares regression between **X** and **Y**.

2.3.2. Estimation of *f*

The purpose of introducing f in the calculation is to quantitatively reflect the uneven distribution of in situ pCO_2 observations, so the greater the number of subgrids, the better f would represent the real unevenness. However, due to the calculation load imposed by large numbers of subgrids and limitation by the resolution of satellite data, the size of the subgrid can be compromised to be somewhat greater than the resolution of the satellite-derived pCO_2 data. Divide the grid box into a number of equal-sized subgrids. The fraction of subgrids with underway pCO_2 observations in the grid box is

$$f = \sum_{i=1}^{M} L_i / M, \tag{9}$$

where M is the number of equal-sized subgrids in the grid box, and for the subgrid i,

$$L = \begin{cases} 1 & \text{if there is at least one } pCO_2 \text{ observation in the sub-grid} \\ 0 & \text{if there is no } pCO_2 \text{ observation in the sub-grid} \end{cases}.$$

2.3.3. Estimation of σ_u^2

For a grid box having at least four satellite pCO_2 data points, the correlation factor and spatial variance are calculated. If there are fewer than four satellite pCO_2 data points for a grid box, the correlation factor and spatial variance are not calculated. If there are pCO_2 in situ measurement data in the grid box, that is, $K > 0$, then σ_u^2 is calculated based on equation (2). If there are no pCO_2 in situ measurement data in the grid box, then σ_u^2 is not calculated.

2.4. The Total Uncertainty σ_T

Integrating the three errors (i.e., the analytical error (E_m), the spatial variance (σ_s^2), and the undersampling variance (σ_u^2)) and assuming their independence, the total uncertainty (σ_T) is

$$\sigma_T = \sqrt{E_m^2 + \sigma_s^2 + \sigma_u^2}. \tag{10}$$

Note that σ_u^2 is not really independent of σ_s^2, as shown in equation (2); however, the introduction of the total uncertainty in equation (10) is helpful in evaluating the relative significance of these sources of error.

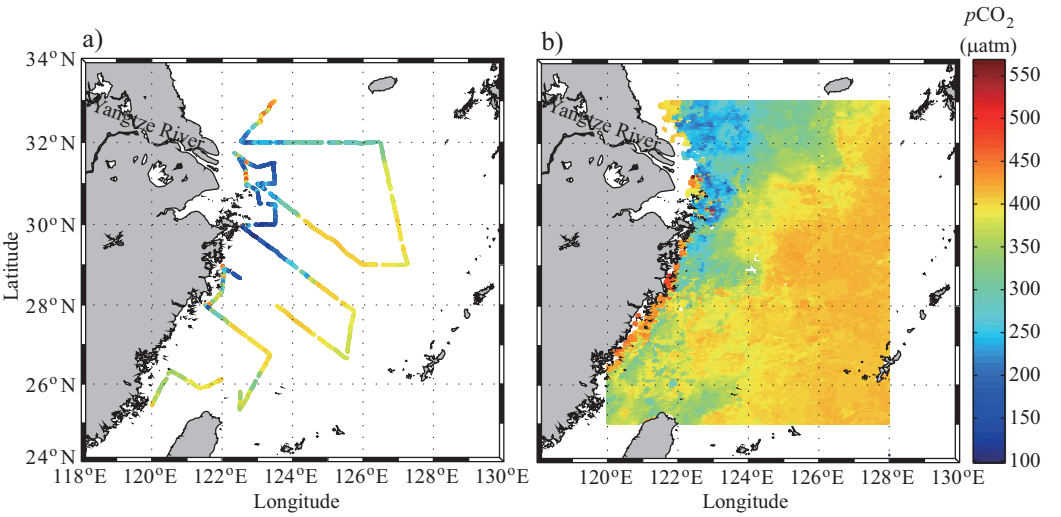

Figure 2. Distributions of sea surface pCO_2 in the East China Sea in August 2009; (a) field-measured pCO_2, and (b) remote sensing-derived sea surface pCO_2 plotted using data from Bai et al. (manuscript in preparation, 2014).

The estimation of the spatial variance, the correlation factor, and subsequently, the undersampling variance requires a concurrent high-spatial-resolution pCO_2 field. In the present study, we used satellite remote sensing-derived pCO_2 data generated from a parallel study (Y. Bai et al., manuscript in preparation, 2014). We must point out that this satellite-derived pCO_2 field was used to define the spatial variance and the correlation factor, but not for the purpose of defining their absolute values. Optionally, sea surface temperature that is more readily available from field measurements or remote sensing techniques could also serve this specific purpose, when a proper correlation with pCO_2 field is achievable.

In summary, the analytical error is estimated using the in situ pCO_2 data by error propagation. The spatial variance and the correlation factor are quantified using a high-spatial-resolution pCO_2 data set. The undersampling variance is subsequently determined by the spatial variance, the correlation factor, and an effective number of in situ pCO_2 data in the grid (Figure 1). The choice of grid box size is case dependent. One rule is to follow the grid size that is used to average pCO_2 data for air-sea flux estimations.

3. Application of the Error Quantification to the East China Sea

The procedure to differentiate the sources of uncertainty was applied to the underway surface pCO_2 data from the East China Sea collected in August 2009. In the summer, the East China Sea is mainly influenced by four water masses: the relatively low-temperature Yellow Sea water in the north; the Yangtze River plume water; and in the mid-shelf, the Taiwan Warm Current from the south, which often combines with a branch of the Kuroshio current in the outer shelf and at the shelf break, characterized by high temperature and salinity [*Su and Pan*, 1989]. The way to determine underway sea surface pCO_2 is referred to *Zhai and Dai* [2009] and the overall uncertainty in the pCO_2 measurement and data reduction is less than 1%. Figure 2a shows that the underway sampling locations are not evenly distributed in the area, with high-spatial-resolution data along several transects. The spatial range of the underway pCO_2 observation is (25°N–33°N, 120°E–127.3°E). Figure 2b shows the distribution of the satellite remote sensing-derived pCO_2, using the 16 day composite data in August 2009 reported in Bai et al. (manuscript in preparation, 2014). Briefly, these satellite pCO_2 data were estimated using a semianalytical algorithm, which analytically expressed pCO_2 variations as the sum of individual components contributed by major controlling factors such as temperature, mixing, and biology. These factors were based on satellite products of sea surface temperature, chlorophyll *a*, and satellite-derived salinity using the salinity algorithm developed for the East China Sea by *Bai et al.* [2013]. On average, these data have much higher spatial resolution, $5' \times 5'$, than the underway pCO_2 data. The selected grid size was $1° \times 1°$, with 100 subgrids, as limited by the resolution of the satellite pCO_2 in each grid box, to have a good spatial resolution of the sampling coverage. Although this remote sensing-based data set primarily served as a constraint on the spatial variance, the data retrieved have been well validated using in situ measurements (Bai et al., manuscript in preparation, 2014).

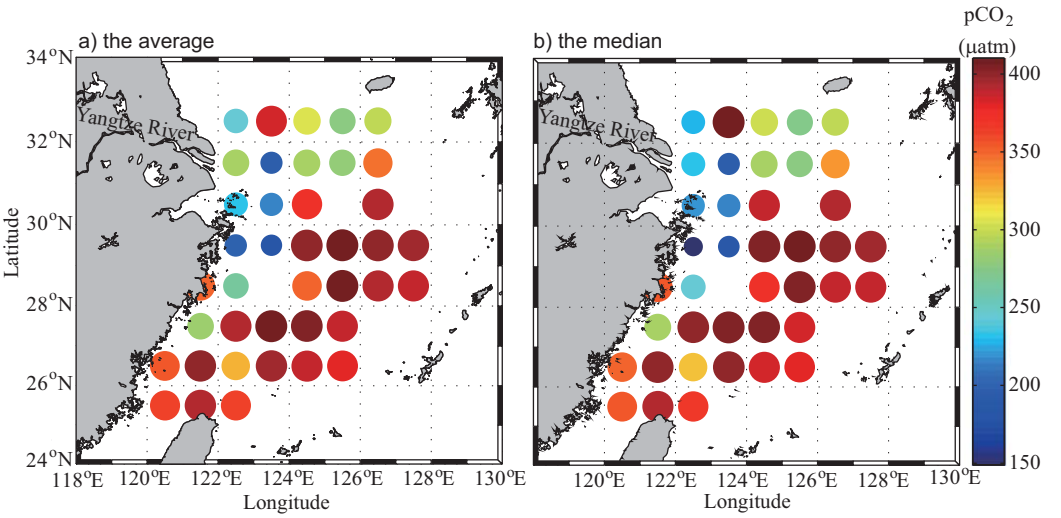

Figure 3. The average and median partial pressure of CO_2 in 1° × 1° grids in the East China Sea in August 2009, collected by field measurements: (a) the average pCO_2 and (b) the median pCO_2. The grid starts from (25°E, 120°N) and the filled circles representing each grid are located at the center of each grid. Note that the value of each parameter is shown by the color bar, the relative magnitude of which is illustrated by the size of the circle.

After gridding, starting from the point of (120°E, 25°N), with a gridding step of 1° northward and eastward, the average surface pCO_2 in the grids in the East China Sea in August 2009 ranges from 185 to 410 μatm, while the median in the grids is 149–413 μatm (Figure 3). The latitudinal trends of the average and the median pCO_2 are similar—relatively low near the coast, and high in regions affected by warm currents. A local maximum is present for the average pCO_2 off the Yangtze River estuary, as demonstrated previously by *Zhai and Dai* [2009].

The analytical error for the underway pCO_2 data is estimated to be less than 1% [*Zhai and Dai*, 2009; *Zhai et al.*, 2013]. After gridding, the analytical error in each grid varies from 2.9 to 5.8 μatm (Figure 4a). The analytical error for each grid is determined by the maximum pCO_2 in the grid. The analytical errors for most observations are approximately 4 μatm. The spatial distribution of the analytical error shows no similarity to that of the standard deviation, which, in general, is relatively high near the coast and much smaller on the outer shelf (Figures 4a and 4f).

As for the spatial variance, the grid boxes closest to the coast have the maximum along latitudes, except for the grid boxes along 25.5°N, where the grid closest to Taiwan shows the maximum spatial variance (Figure 4b). σ_s decreases with distance offshore, with the rate of decrease much higher on the inner shelf than on the outer shelf of the East China Sea. The latitudinal trend is similar to the general trend of the standard deviation (Figures 4b and 4f). On the outer shelf from north to south, the spatial variance in the grid boxes along 32.5°N and 31.5°N, which are affected by the Yellow Sea water, is at least twice as high as that of other grid boxes, indicating physical controls on the pCO_2 variation. The longitudinal distribution of the standard deviation, however, shows no trend. The greatest σ_s on the East China Sea shelf, 77.8 μatm, appears outside the Yangtze River estuary, and the lowest σ_s, 3.2 μatm, is on the outer shelf of the East China Sea.

The undersampling variance is also the highest off the Yangtze River estuary, with the range of σ_u from 0.9 to 45.5 μatm (Figure 4c). The maximum is determined by both the maximum spatial variance ($\sigma_s = 77.8$ μatm), a relatively small number of underway observations ($K = 59$), and relatively poor sampling coverage in the grid. At the same latitude, the grid on its right has σ_s of 62.1 μatm, but due to a much greater number of observations ($K = 511$) in this grid, its σ_u is about one tenth of the maximum. The second largest undersampling variance appears in the grid centered at (26.5°N, 124.5°E), in which grid there are only three observations. The undersampling variance is affected by both the spatial variation and the effective number of observations; thus, it shows no latitudinal trend as the spatial variance does. The undersampling variance and the standard deviation do not distribute in a similar way (Figures 4c and 4f). In the grid boxes with similar spatial variances, their undersampling variances reflect the relative magnitude of the inverse of the

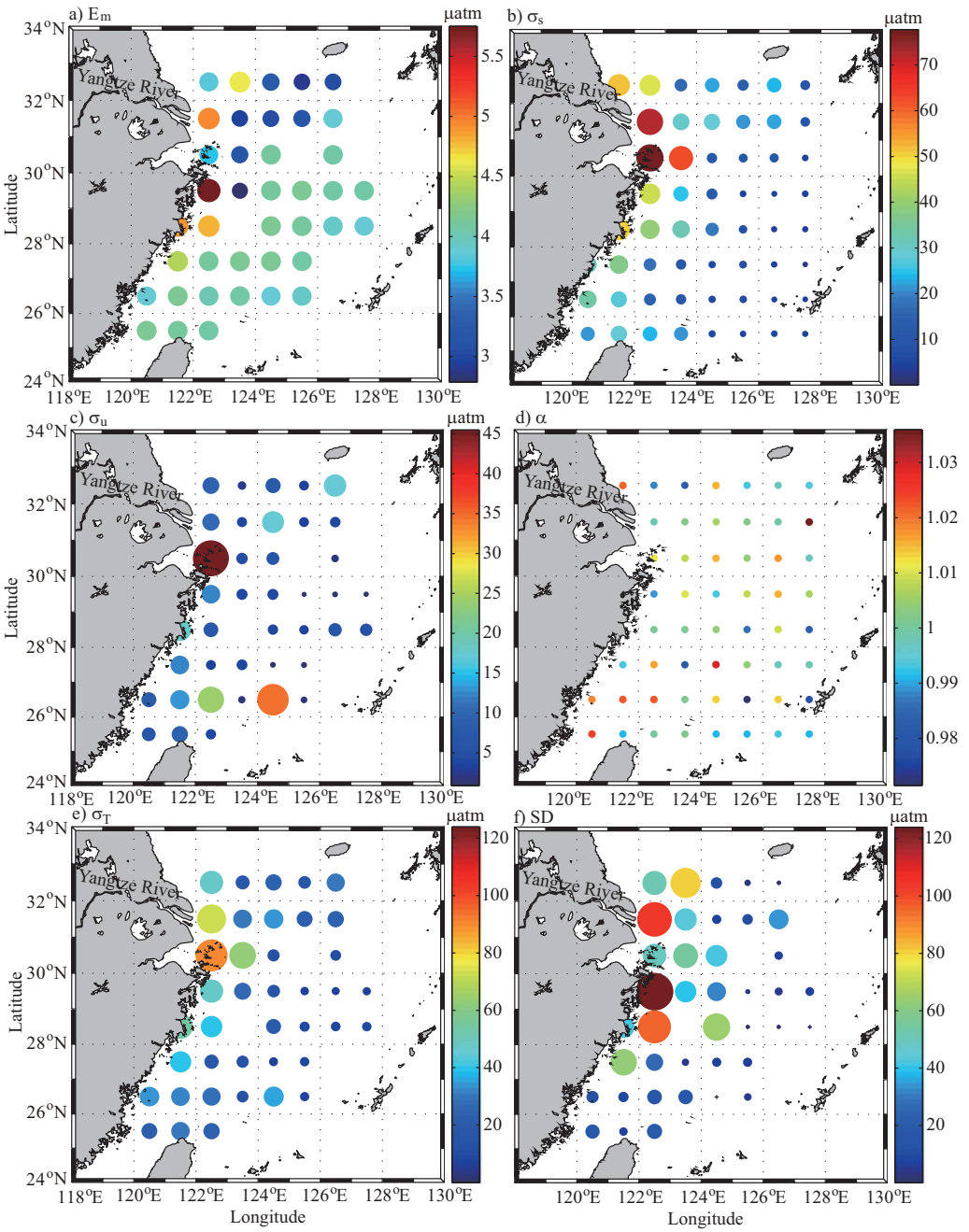

Figure 4. The three sources of uncertainty (i.e., analytical error, E_m, spatial variance of pCO_2, σ_s^2, and the bias from undersampling, σ_u^2), the correlation factor (α), the total uncertainty (σ_T), and the standard deviation (SD) calculated for the 1° × 1° gridded underway pCO_2 data in the East China Sea collected in August 2009; (a) E_m, (b) σ_s, (c) σ_u, (d) α, (e) σ_T, and (f) SD. The filled circles representing each grid are located at the center of each grid. Note that the value of each parameter is shown by the color bar, the relative magnitude of which is illustrated by the size of the circle.

effective number of sampling observations in each grid, since the correlation factor for all the grids is approximately 1 (Figure 4d).

The total uncertainty derived from the three sources of uncertainty is the highest off the Yangtze River estuary. The total uncertainty apparently differs greatly from the standard deviation of the gridded underway pCO_2 data, especially near the coast, where greater spatial variance is present (Figures 4b, 4e, and 4f). For example, in the grid box centered at (29.5°N, 122.5°E), the biggest standard deviation, 124.0 μatm, appears,

while σ_T is 47.9 μatm. In the grid centered at (25.5°N, 121.5°E), σ_T is 30.2 μatm, while the standard deviation is 6.4 μatm. This difference of the standard deviation from the total uncertainty is mainly caused by the uneven spatial coverage of the underway data.

For the three sources of uncertainty in the gridded pCO_2 data, the spatial variance is dominant in the grid boxes with relatively good sampling coverage, equivalent to more than 90% of the total uncertainty (Figure 4). For the grid boxes with relatively poor sampling coverage, the spatial variance is equivalent to as low as 16.9% of the total uncertainty, which appears in the grid centered at (26.5°N, 124.5°E), where the undersampling variance dominates the total uncertainty. The analytical error is relatively uniform for the gridded underway surface pCO_2 data collected in August 2009, while the range of the spatial variance is the greatest in the three uncertainties. None of the three uncertainties show a similar spatial pattern to that of the standard deviation. In further investigations, we recommend increasing sampling coverage in the grid boxes with relatively large undersampling variances and/or spatial variances.

4. Conclusions

Our approach successfully differentiates and quantifies the three sources of uncertainty in gridded surface pCO_2 data: analytical error, spatial variance, and undersampling variance. Application of this approach to the underway surface pCO_2 data collected from the East China Sea demonstrates that none of the sources of uncertainty follow the same spatial trend as the standard deviation and that the total uncertainty derived from the three sources is controlled by the spatial variance in grids with relatively good sampling coverage. The undersampling variance contributes the most to the total uncertainty in the grid boxes with poor sampling coverage. The quantification of the sources of uncertainty in gridded pCO_2 data differentiates the inherent spatial variability of the pCO_2 field from real errors associated with the analytical methods and induced from the undersampling. Such differentiation would thus ease the difficulties in data interpolation/extrapolation beyond the sampling sites, and in optimizing strategies for designing field observations to reduce the uncertainty, and eventually better constrain air-sea CO_2 fluxes.

Acknowledgments
The data for this paper are available upon request from the corresponding author. This study was funded by the National Basic Research Program of China (2009CB421200), the Public Science and Technology Research Funds Projects for Ocean Research (200905012), and the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (41121091). Discussions with Xianqiang He helped to better define the uncertainty. We would also like to thank the anonymous reviewers for his/her critical comments, which helped improve the manuscript.

References

Bai, Y., D. Pan, W.-J. Cai, X. He, D. Wang, B. Tao, and Q. Zhu (2013), Remote sensing of salinity from satellite-derived CDOM in the Changjiang River dominated East China Sea, *J. Geophys. Res.*, *118*, 227–243, doi:10.1029/2012JC008467.

Bakker, D. C. E., et al. (2013), An update to the surface ocean CO_2 Atlas (SOCAT version 2), *Earth Syst. Sci. Data*, *6*, 465–512, doi:10.5194/essdd-6-465-2013.

Bhattacharya, A., V. V. Calmidi, and R. L. Mahajan (2002), Thermophysical properties of high porosity metal foams, *Int. J. Heat Mass Transfer*, *45*, 1017–1031.

Goñi, M. A., K. C. Ruttenberg, and T. I. Eglinton (1998), A reassessment of the sources and importance of land-derived organic matter in surface sediments from the Gulf of Mexico, *Geochim. Cosmochim. Acta*, *62*, 3055–3075.

Hales, B., P. G. Strutton, M. Saraceno, R. Letelier, T. Takahashi, R. Feely, C. Sabine, and F. Chavez (2012), Satellite-based prediction of pCO_2 in coastal waters of the eastern North Pacific, *Prog. Oceanogr.*, *103*, 1–15, doi:10.1016/j.pocean.2012.03.001.

Han, A., M. Dai, S.-J. Kao, J. Gan, Q. Li, L. Wang, W. Zhai, and L. Wang (2012), Nutrient dynamics and biological consumption in a large continental shelf system under the influence of both a river plume and coastal upwelling, *Limnol. Oceanogr.*, *57*(2), 486–502.

Jo, Y.-H., M. Dai, W. Zhai, X.-H. Yan, and S. Shang (2012), On the variations of sea surface pCO_2 in the Northern South China Sea—A remote sensing based neural network approach, *J. Geophys. Res.*, *117*, C08022, doi:10.1029/2011JC007745.

Le Quéré, C., et al. (2009), Trends in the sources and sinks of carbon dioxide, *Nat. Geosci.*, *2*, 831–836, doi:10.1038/ngeo689.

Lee, P.-S., S. V. Garimella, and D. Liu (2005), Investigation of heat transfer in rectangular microchannels, *Int. J. Heat Mass Transfer*, *48*, 1688–1704.

Nelson, N. B., N. R. Bates, D. A. Siegel, and A. F. Michaels (2001), Spatial variability of the CO_2 sink in the Sargasso Sea, *Deep Sea Res., Part II*, *48*, 1801–1821.

Ono, T., T. Saino, N. Kurita, and K. Sasaki (2004), Basin-scale extrapolation of shipboard pCO_2 data by using satellite SST and Chla, *Int. J. Remote Sens.*, *25*(19), 3803–3815, doi:10.1080/01431160310001657515.

Pfeil, B., et al. (2013), A uniform, quality controlled Surface Ocean CO_2 Atlas (SOCAT), *Earth Syst. Sci. Data*, *5*, 125–143, doi:10.5194/essd-5-125-2013.

Sabine, C. L. et al. (2013), Surface Ocean CO_2 Atlas (SOCAT) gridded data products, *Earth Syst. Sci. Data*, *5*, 145–153, doi:10.5194/essd-5-145-2013.

Sarma, V. S. S., T. Saino, K. Sasaoka, Y. Nojiri, T. Ono, M. Ishii, H. Y. Inoue, and K. Matsumoto (2006), Basin-scale pCO_2 distribution using satellite sea surface temperature, Chl a, and climatological salinity in the North Pacific in spring and summer, *Global Biogeochem. Cycle*, *20*, GB3005, doi:10.1029/2005GB002594.

Shen, S. S. P., H. Yin, and T. M. Smith (2007), An estimate of the sampling error variance of the gridded GHCN monthly surface air temperature data, *J. Clim.*, *20*, 2321–2331, doi:10.1175/JCLI4121.1.

Shen, S. S. P., C. K. Lee, and J. Lawrimore (2012), Uncertainties, trends, and hottest and coldest years of U.S. surface air temperature since 1895: An update based on the USHCN V2 TOB data, *J. Clim.*, *25*, 4185–4203, doi:10.1175/JCLI-D-11-00102.1.

Stephens, M. P., G. Samuels, D. B. Olson, R. A. Fine, and T. Takahashi (1995), Sea–air flux of CO_2 in the North Pacific using shipboard and satellite data, *J. Geophys. Res.*, *100*, 13,571–13,583.

Su, J., and Y. Pan (1989), A preliminary study on circulation dynamics of continental shelf to the north of Taiwan, *Acta Oceanol. Sin.*, *11*, 1–14.

Sweeney, C., T. Takahashi, and A. Gnanadesikan (2002), Spatial and temporal variability of surface water pCO_2 and sampling strategies, in *A Large-Scale CO_2 Observing Plan: In Situ Oceans and Atmosphere (LSCOP)*, pp. 155–175, NOAA Off. of Global Programs, Washington, D. C.

Takahashi, T., T. T. Takahashi, and S. C. Sutherland (1995), An assessment of the role of the North Atlantic as a CO_2 sink, *Philos. Trans. R. Soc. London B*, *348*, 143–152.

Takahashi, T., R. A. Feely, R. F. Weiss, R. H. Wanninkhof, D. W. Chipman, S. C. Sutherland, and T. T. Takahashi (1997), Global air-sea flux of CO_2: An estimate based on measurements of sea-air *p*CO2 difference, *Proc. Natl. Acad. Sci. U. S. A.*, *94*, 8292–8299, doi:10.1073/pnas.94.16.8292.

Takahashi, T., et al. (2002), Global sea-air CO_2 flux based on climatological surface ocean *p*CO2, and seasonal biological and temperature effects, *Deep Sea Res., Part II*, *49*, 1601–1622.

Takahashi, T., et al. (2009), Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, *Deep Sea Res., Part II*, *56*, 554–577, doi:10.1016/j.dsr2.2008.12.009.

Tan, P. P., I. Y. Fung, and T. Takahashi (1990), Observational constraints on the global atmospheric CO_2 budget, *Science*, *247*, 1431–1438.

Taylor, J. R. (1982), *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, Univ. Sci. Books, Mill Valley, Calif.

Taylor, J. R. (1997), *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 2nd ed., Univ. Sci. Books, Sausalito, Calif.

Wang, G., A. J. Spivack, and S. D'Hondt (2010), Gibbs energies of reaction and microbial mutualism in anaerobic deep subseafloor sediments of ODP Site 1226, *Geochim. Cosmochim. Acta*, *74*, 3938–3947.

Wanninkhof, R., et al. (2013), Global ocean carbon uptake: Magnitude, variability and trends, *Biogeosciences*, *10*, 1983–2000, doi:10.5194/bg-10-1983-2013.

Zhai, W., and M. Dai (2009), On the seasonal variation of air–sea CO_2 fluxes in the outer Changjiang (Yangtze River) Estuary, East China Sea, *Mar. Chem.*, *117*, 2–10.

Zhai, W., M. Dai, B. Chen, X. Guo, Q. Li, S. Shang, C. Zhang, W.-J. Cai, and D. Wang (2013), Seasonal variations of air-sea CO_2 fluxes in the largest tropical marginal sea (South China Sea) based on multiple-year underway measurements, *Biogeosciences*, *10*, 7775–7791.

Zhu, Y., S. Shang, W. Zhai, and M. Dai (2009), Satellite-derived surface water pCO_2 and air-sea CO_2 fluxes in the northern South China Sea in summer, *Prog. Nat. Sci.*, *19*(6), 775–779.