AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2016JD025013

Key Points:
- Present a spatiotemporal technique termed TFFSRC
- Produce a spatially continuous and daily global total ozone product of about 10 years

Correspondence to:
H. Shen and C. Zeng,
shenhf@whu.edu.cn;
zengchaozc@hotmail.com

Citation:
Peng, X., H. Shen, L. Zhang, C. Zeng, G. Yang, and Z. He (2016), Spatially continuous mapping of daily global ozone distribution (2004–2014) with the Aura OMI sensor, *J. Geophys. Res. Atmos.*, *121*, doi:10.1002/2016JD025013.

Received 29 FEB 2016
Accepted 17 OCT 2016
Accepted article online 19 OCT 2016

Spatially continuous mapping of daily global ozone distribution (2004–2014) with the Aura OMI sensor

Xiaolin Peng[1], Huanfeng Shen[1,2], Liangpei Zhang[2,3], Chao Zeng[3], Gang Yang[4], and Zongyi He[1]

[1]School of Resource and Environmental Sciences, Wuhan University, Wuhan, China, [2]Collaborative Innovation Center for Geospatial Information Technology, Wuhan, China, [3]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China, [4]Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo, China

Abstract Total ozone data from the Aura Ozone Monitoring Instrument (OMI) play an important role in the monitoring of the spatial distribution and temporal change of total ozone. However, since September 2005, and especially after mid-2006, due to row anomalies in the OMI instrument, one third to one half of the OMI total ozone data has been missing. In this study, we generate a spatially continuous and daily global total ozone product (2004–2014) by quantitatively reconstructing the level 3 (gridded) total ozone data via a new two-step method, which takes full advantage of the temporal and spatial correlation characteristics. First, a preliminary prediction is made based on an adaptive weighted temporal fitting method. Residual correction based on an anisotropic kriging method is then proposed to further improve the prediction accuracy. To assess the efficacy of the proposed method, a comparison of different gap filling algorithms through a series of simulated experiments was performed. On this basis, we further evaluated the proposed product with Brewer spectrophotometers' total ozone columns. The evaluation results suggest that the proposed method outperforms the other algorithms, and its product is better able to capture total ozone variation than the MERRA-2 assimilated ozone product. The total ozone product produced in this study can be freely downloaded from http://sendimage.whu.edu.cn/send-resource-download/.

1. Introduction

Ozone, a trace gas in the Earth's atmosphere, has a very important influence on the atmospheric environment and ecological system. First of all, ozone in the stratosphere gathers together to form the ozone layer, which acts as a natural sunscreen to protect the planet's surface by absorbing the harmful wavelengths of incoming solar ultraviolet (UV) radiation [*Hartley*, 1880]. In the early 1980s, scientists began to realize that the Earth's natural sunscreen was thinning dramatically over the Antarctic, which came to be known as the ozone hole [*Herman et al*., 1993; *Loyola et al*., 2009; *Staehelin et al*., 2001]. Countries across the world woke up and negotiated the 1987 Montreal Protocol, an international treaty to prevent a further thinning of the ozone layer and lead to a gradual recovery in the next decades [*Hadjinicolaou et al*., 2005; *Kiesewetter et al*., 2010]. On the other hand, when near to the Earth's surface, relatively high concentrations of ozone are a great threat to human health [*Bell et al*., 2006; *Jerrett et al*., 2009; *Lippmann*, 1989; *White et al*., 1994] and ecosystems [*Heck et al*., 1982; *Reich and Amundson*, 1985]. It is especially true that surface ozone and the impact of ozone pollution are projected to increase in most nations over the next few decades [*Jain et al*., 2005; *Kulkarni et al*., 2010; *Riahi et al*., 2007]. In addition, ozone plays a very important role in global climate and environmental change [*Ramanathan et al*., 1976, 1985] because it is an indispensable component of the atmosphere, and it affects processes both in the troposphere and stratosphere, such as atmosphere dynamics, thermodynamics, and radiation-chemical processes. Therefore, research into the spatial distribution and temporal change of ozone is of great significance.

To monitor the spatial distribution and temporal change of ozone, we need long time series total ozone data sets with a global spatial distribution. Satellite data can meet these requirements, with a high accuracy and precision [*Loyola et al*., 2009]. Since the 1970s, space-based measurement of the ozone column has been performed by the solar backscatter ultraviolet and Total Ozone Mapping Spectrometer (TOMS) series of instruments [*Bhartia et al*., 2004; *Wellemeyer et al*., 2004]. More recently, NASA's Ozone Monitoring Instrument (OMI) [*Levelt et al*., 2006] on board the Aura satellite has replaced TOMS. These instruments provide daily

backscattered solar UV radiation measurements that can be used to retrieve column ozone values and other geophysical atmospheric parameters. The OMI is the first of a new generation of spaceborne spectrometers combining a high spatial resolution with a daily global coverage. Monitoring the ozone layer [*Chipperfield et al*., 2003] and detecting tropospheric ozone pollution at a regional scale [*Fishman et al*., 1990] are the top priorities for the OMI. However, as early as September 2005, and especially after mid-2006, the OMI measurement began to be gradually impacted by "row anomaly" artifacts. On 24 January 2009, the obstruction suddenly increased and now partially blocks an increased proportion of the field of view for certain Aura orbits, which has adversely affected between one third to one half of the ozone measurements [*Yan et al*., 2012]. Since 5 July 2011, row anomalies now affect all the Aura orbits and can be seen as thick white stripes of bad data in the OMI total ozone data. More detailed technical information about the "OMI row anomaly" issue is available at http://www.knmi.nl/omi/research/product/rowanomaly-background.php.

The deterioration in the data quality resulting from the OMI instrument anomalies has become a major obstacle for the monitoring of total ozone data. There is therefore a clear need for an effective methodology to reconstruct the missing information in the degraded data. Data assimilation is an alternative approach for the reconstruction of OMI missing information on synoptic time scales. For example, *Wargan et al*. [2015] performed data assimilation to generate the MERRA-2 assimilated ozone product at a horizontal resolution of $0.5° \times 0.625°$ (latitude × longitude), with a global coverage. Clearly, this assimilated product has a lower spatial resolution, and an underlying difficulty is that models containing atmospheric chemistry and physical processes are complex, and the coupling between them is not easy to operate. Therefore, the data-driven technologies are probably more suitable for retrieving the degraded OMI total ozone data. A large number of interesting algorithms [*Li et al*., 2015, 2014; *Shen et al*., 2015] have been developed to compensate for dead or missing pixels. These approaches can be divided into three categories.

One category is the approaches based on the spatial information in the data sets themselves. The most simple interpolation methods are the nearest-neighbor, bilinear, inverse distance weighting and kriging methods. It is worth mentioning that kriging interpolation [*Goovaerts*, 1997; *Van der Meer*, 2012; *Zhang et al*., 2007] is a well-known geostatistical prediction technique used to analyze the spatial construction of valid data and to provide estimates at missing locations. Furthermore, some inpainting methods [*Cheng et al*., 2014; *Shen and Zhang*, 2009; *Siravenha et al*., 2011] have also been proposed to remove the dead pixels existing in remotely sensed images. However, a major drawback of this category of methods is the spatial discontinuity of the filling result that appears in the situation of wide missing areas.

For the reconstruction of wide missing areas, the second category of methods employing multitemporal auxiliary data is much more attractive. Most of these approaches make use of another clean measurement over a limited time period [*Shen et al*., 2016]. The Savitzky-Golay filter [*Chen et al*., 2004; *Savitzky and Golay*, 1964] and harmonic analysis [*Roerink et al*., 2000; *Yang et al*., 2015] can be used for the reconstruction of missing data due to cloud contamination. However, to eliminate the brightness of the apparent inconsistencies, most of the valid values are also changed. *Zeng et al*. [2013] presented the weighted linear regression (WLR) method and described how to accurately reconstruct the missing gaps, leaving the valid values unchanged. The WLR method fills the degraded pixels using the data from multiple scenes as referable information by building a regression model between the corresponding pixels. It is worth noting that the WLR method is based on the precondition that the changes in the ground features between scenes are regular, and the restored result may be unsatisfactory when the scenes change abruptly. In this study, the OMI total ozone product does contain large changes in the time series.

The collection and processing of spatiotemporal data with a finer temporal resolution calls for sophisticated spatiotemporal interpolation: the third category of methods. Various types of spatiotemporal kriging (STkriging) [*Pebesma and Gräler*, 2015] methods are considered to be the state-of-the-art statistical approaches for spatiotemporal analysis. Some systematic contributions to spatiotemporal geostatistical analysis can be found in the literature [*Cressie and Huang*, 1999; *Gneiting*, 2002; *Ma*, 2002, 2003; *Stein*, 2005]. STkriging can potentially provide more accurate predictions than spatial interpolation because observations taken at other times can be included. However, for the STkriging method, both the computational task and the spatiotemporal data volume are huge, and a global experiment is unrealistic. Another method known as the (a penalized least squares approach (PLS) combines the use of the discrete cosine transform (DCT)) method [*Garcia*, 2010] is recommended as an effective model for gap filling purposes in large spatiotemporal

data sets. The method is fully three-dimensional and thus explicitly utilizes both the spatial and temporal information of the data set to derive the statistical model and predict the missing values. The statistical modeling process is completely controlled by a smoothing parameter which is easy to specify, which eliminates the need for complicated model parameterization. The third category of methods can reveal the significant spatiotemporal dependencies, while a purely spatial or temporal approach ignores these dependencies.

The objective of this study is to develop a combined spatiotemporal technique, according to the characteristics of the OMI total ozone data, to generate a spatially continuous and daily global total ozone product. The proposed technique, which is referred to hereafter as temporal fitting followed by spatial residual correction (TFFSRC), takes full advantage of an adaptive weighted temporal fitting (AWTF) method and spatial residual correction based on anisotropic kriging interpolation technology. That is to say, the proposed method can achieve full space-time coherence by exploiting the values of the neighboring pixels in space and time, while taking into account the autocorrelated error. In order to evaluate the precision of the proposed algorithm, a number of simulated experiments were performed to compare the proposed method with the other effective gap filling methods, such as the spatial kriging, WLR, STKriging, and DCT-PLS methods mentioned above. Based on this, Brewer spectrophotometers' total ozone columns were then employed for the validation of the TFFSRC product.

The rest of this paper is organized as follows. Section 2 describes the data and study areas. Section 3 gives the detailed implementation of the proposed algorithm. Section 4 presents the generated total ozone product, the results of the simulated experiments, and the comparison with ground-based Brewer observations. Finally, Section 5 provides the conclusion and discussion.

2. Data

The OMI instrument [*Levelt et al*., 2006] is on board NASA's Earth Observing System Aura satellite, which was launched on 15 July 2004. At both visible and UV wavelengths, the OMI instrument detects backscattered solar radiance to measure daytime total ozone. The OMI ozone data are available as level 2 (orbit/swath) and level 3 (gridded) data, beginning in late August 2004. The data employed in this study were the L3 OMI total ozone product named OMTO3e, with the spatial resolution of $0.25° \times 0.25°$ (latitude × longitude), which were obtained from the NASA website (http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omto3e_v003.shtml). From October 2004 to the present, the OMTO3e product has been provided by NASA every day, except for a few cases of unavailable data. It is a very important data source for the monitoring of the spatial distribution and temporal change in total ozone and has a high spatial and temporal resolution relative to other ozone satellite products.

It has been mentioned before that the OMI instrument anomalies first appeared in September 2005, caused by physical material which had become dislodged blocking the optical path, resulting in wide missing stripes gradually appearing in the OMI products. The OMTO3e data from late 2004 to September 2005 are almost complete, with no gaps, apart from some sporadic and small gaps caused by factors such as cloud contamination. However, from September 2005, narrow stripes of missing data appear, which then increase from mid-2006. The most serious case is for January 2009, where the missing data account for one third to one half of the data set. Thus, we chose data from late 2004 to mid-2006 with a better quality to perform the simulated experiments, and we selected degraded data from 2007 to the present to simulate the wide gaps, whereby the gaps of the degraded data replaced the values for those same locations in the normal data from 2004 to 2006. The orbital inclination of the Aura satellite is 98.1°, providing latitudinal coverage from 82°N to 82°N. The wide gaps mostly occur from 60°N to 60°S.

3. Methodology

Based on the special properties of the OMTO3e data, we propose a new method termed TFFSRC to implement spatially continuous and daily global mapping of total ozone. The missing data are restored in two steps: a preliminary prediction by an AWTF method, followed by residual correction by a kriging method after investigating and accounting for the anisotropic effects [*Goovaerts*, 1997]. The trend information over time is explained by temporal regression analysis, and the residual representing the spatial structure to account for the local details is interpolated by linear unbiased estimation via the semivariance function. These two steps are introduced in detail below.

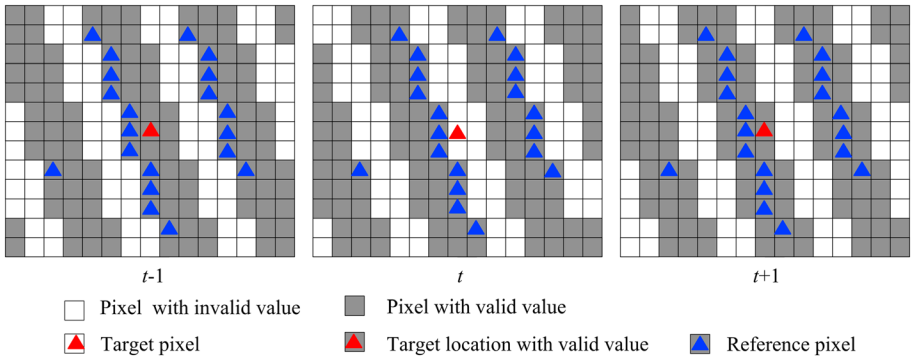


Figure 1. Schematic diagram of the reference pixel selection.

3.1. The First Step Prediction by Adaptive Weighted Temporal Fitting (AWTF)

In this paper, the OMTO3e data set with gaps to be restored is defined as the target data, while the other multitemporal data sets selected to fill the target data are regarded as the auxiliary data. A missing pixel in gaps of the target data is defined as a target pixel, and its location is called the target location.

Figure 1 shows a schematic drawing of three consecutive data sets acquired on $t-1$, t, and $t+1$. It can be seen that the location of the missing areas changes over time, so we can make use of the multitemporal data whose acquisition date is close to the target data as the auxiliary data, ensuring that there are enough overlaps of the pixels in the missing regions to essentially "fill in" the gaps. In Figure 1, it can be further found that the data sets of $t-1$ and $t+1$ can completely overlap the gaps of the t data set, but about 90% of the gaps of the t data set can be overlapped by valid values in common in the $t-1$ and $t+1$ data sets. Due to the large changes in the space-time patterns of total ozone, in order to ensure the least difference between the target and auxiliary data sets, we choose two data sets from different acquisition dates (one day before and after the target data t data set) as auxiliary data sets, and the remaining 10% are filled by either of the auxiliary data sets ($t-1$ or $t+1$) with a valid value in the target location.

3.1.1. Reference Pixel Selection

The proposed algorithm makes use of the pixel at the target location in the auxiliary data to predict the target pixel, while the relationship between them is calculated from the local reference pixels. Accordingly, it is necessary to first determine the reference pixels of the target location in the target and auxiliary data sets. In our study, the target data from the OMI instrument have a lower spatial resolution and wider gaps. Furthermore, due to the synoptic-scale weather patterns, large changes appear in both the space and time scales for the total ozone data, especially at middle and high latitudes in both the Northern and Southern Hemispheres. However, these changes are slight over a short time period and a local space. Thus, it is reasonable to assume that neighboring pixels in close proximity to the missing pixel will share similar characteristics and temporal patterns of variation of total ozone data. To guarantee smaller spatial differences, we select the pixels closest to the target pixel as reference pixels, where valid values exist in both the target and auxiliary data sets. In this study, only pixels with valid values in the unfilled data are used as reference pixels, and they share similar relationships with the target pixels located within the gaps.

For each target pixel, relationships are established by the reference pixels from the target data and the auxiliary data. An adaptive window is used for the reference pixel selection, and the window size is first set to an initial value. The number of selected reference pixels is set as N. If the number of selected reference pixels cannot meet N, the window size is enlarged by two pixels. A maximum window size (W_{max}) should be set, because too large a window will mean that the reference pixels are far away from the target location and will lead to a reduction in the accuracy. For the special situation of when the maximum window size (W_{max}) has been reached but the number of reference pixels is not enough, the target pixel is identified as an outlier. It is worth mentioning that a smaller value of N will increase the unreliability of the result. This is because for a regression algorithm, fewer sample points will make the estimation less stable. However, a larger value of N will increase the selection window size, leading to the reference pixels being farther away from the target pixel. Moreover, a large value of N greatly decreases the computational efficiency. Optimized through a series of comparative experiments, in this study, the number of reference pixels (N) is set to 50, and the initial (W_0)

and maximum window sizes (W_{max}) are set to 7 and 61. In Figure 1, the gray pixels marked by the blue triangle represent the reference pixels.

3.1.2. The Weighted Least Squares Fitting

For every single missing pixel, there are two auxiliary data sets, and each auxiliary data set can offer one prediction through one linear relationship:

$$I'_{t-1}(x,y) = a_{t-1} \times I_{t-1}(x,y) + b_{t-1} \tag{1}$$

$$I'_{t+1}(x,y) = a_{t+1} \times I_{t+1}(x,y) + b_{t+1} \tag{2}$$

where the subscripts $t-1$ and $t+1$ represent the acquisition dates of the auxiliary data sets and $I'_{t-1}(x,y)$ and $I'_{t+1}(x,y)$ are the predictions fitted separately from the different linear relationships built by each auxiliary data set (acquired at the dates of $t-1$ or $t+1$) and the target data set. $I_{t-1}(x,y)$ and $I_{t+1}(x,y)$ are the pixels at the target location in the auxiliary data, and a_{t-1}, b_{t-1}, a_{t+1}, and b_{t+1} are regression coefficients corresponding to the linear relationships calculated from local neighboring reference pixels of the target and auxiliary data. (x,y) represents the location of the target location, and all the reference pixels of the target pixel are expressed as

$$\mathbf{x}^R = \left[x_1^R, x_2^R, x_i^R \cdots, x_N^R\right]^{\mathrm{T}} \tag{3}$$

$$\mathbf{y}^R = \left[y_1^R, y_2^R, y_i^R \cdots, y_N^R\right]^{\mathrm{T}} \tag{4}$$

where $\mathbf{x^R}$ and $\mathbf{y^R}$ represent the coordinate vectors of the x and y axes of the reference pixels. We assume that the target and reference pixels share the same regression relationship, and then a series of equations can be established by substituting the coordinates of the reference pixels into equations (1) and (2). The regression coefficients can be obtained by the use of weighted least squares fitting. For the auxiliary data set on $t-1$, for example, the regression coefficients in equation (1) are given as

$$a_{t-1} = \frac{\sum_{i=1}^{N} T_i \times \left(I_t\left(x_i^R, y_i^R\right) - \overline{I_t(\mathbf{x}^R, \mathbf{y}^R)}\right)\left(I_{t-1}\left(x_i^R, y_i^R\right) - \overline{I_{t-1}(\mathbf{x}^R, \mathbf{y}^R)}\right)}{\sum_{i=1}^{N} T_i \times \left(I_{t-1}\left(x_i^R, y_i^R\right) - \overline{I_{t-1}(\mathbf{x}^R, \mathbf{y}^R)}\right)^2} \tag{5}$$

$$b_{t-1} = \overline{I_t(\mathbf{x}^R, \mathbf{y}^R)} - a_{t-1} \times \overline{I_{t-1}(\mathbf{x}^R, \mathbf{y}^R)} \tag{6}$$

where $I_{t-1}\left(x_i^R, y_i^R\right)$ is the total ozone value of the ith reference pixel on $t-1$ and $\overline{I_t(\mathbf{x}^R, \mathbf{y}^R)}$ and $\overline{I_{t-1}(\mathbf{x}^R, \mathbf{y}^R)}$ are the mean values of all the reference pixels of the target data and auxiliary data on $t-1$, respectively. The weights of each selected reference pixel of the auxiliary data on $t-1$ are defined as

$$T_i = (1/\mathrm{Diff}_i) / \sum_{i=1}^{N} (1/\mathrm{Diff}_i) \tag{7}$$

$$\mathrm{Diff}_i = \left|I_{t-1}\left(x_i^R, y_i^R\right) - I_{t-1}(x,y) + \Delta\right| \times D_i \tag{8}$$

where the range of weight T_i is from 0 to 1 and the sum of all the selected reference pixel weights is 1. Δ is a small value to prevent Diff_i equaling zero. D_i describes the chordal distance between the ith corresponding reference pixel and the target pixel in the target data. It is calculated by the haversine formula [*Sinnott*, 1984]:

$$\mathrm{dlon} = \mathrm{lon}(y) - \mathrm{lon}\left(y_i^R\right), \quad \mathrm{dlat} = \mathrm{lat}(x) - \mathrm{lat}\left(x_i^R\right) \tag{9}$$

$$e = (\sin(\mathrm{dlat}/2))^2 + \cos\left(\mathrm{lat}\left(x_i^R\right)\right)\cos(\mathrm{lat}(x))(\sin(\mathrm{dlon}/2))^2 \tag{10}$$

$$c = 2{*}a\tan(\mathrm{sqrt}(e)), \quad \mathrm{sqrt}(1-e) \tag{11}$$

$$D_i = R{*}c \tag{12}$$

where $\mathrm{lat}(x)$, $\mathrm{lon}(y)$, $\mathrm{lat}\left(x_i^R\right)$, and $\mathrm{lon}\left(y_i^R\right)$ represent the latitude degree and longitude degree at locations (x,y) and $\left(x_i^R, y_i^R\right)$, respectively; c is the great circle distance in radians; and R is the average radius of Earth, at 6371 km. In this study, all the spatial distances are calculated following the haversine formula. T_i is substituted into equations (5) and (6) to calculate a_{t-1} and b_{t-1}. In the same way, a_{t+1} and b_{t+1} can be obtained by using the target and auxiliary data on $t+1$.

3.1.3. The Adaptive Weighted Averaging

The information in the target location of the two auxiliary data sets is used to predict the value of the target pixel. However, the contribution of the auxiliary data may vary because reference pixels from one auxiliary

data set are likely to be more comparable to the target data set than the other. More accurately, for the AWTF method, every single missing pixel can be predicted by the weighted average of the two predictions:

$$\hat{I}_t(x,y) = w_{t-1} \times I'_{t-1}(x,y) + w_{t+1} \times I'_{t+1}(x,y) \tag{13}$$

where $\hat{I}_t(x,y)$ is the predicted value of the pixel at the target location in the target data. w_{t-1} and w_{t+1} are the weights of the two predictions, respectively, and express the similarity of the reference pixels between the auxiliary and target data. We use a Gaussian kernel function to express the adaptive weights as

$$w_{t-1} = 1/C \times \exp\left(-\frac{\overline{\sum d_{t-1}^2}}{g^2}\right) \tag{14}$$

$$w_{t+1} = 1/C \times \exp\left(-\frac{\overline{\sum d_{t+1}^2}}{g^2}\right) \tag{15}$$

where w_{t-1} and w_{t+1} meet the conditions$0 \le w_{t-1}, w_{t+1} \le 1$ and $w_{t-1} + w_{t+1} = 1$. C is the normalized parameter:

$$C = \exp\left(-\frac{\overline{\sum d_{t-1}^2}}{g^2}\right) + \exp\left(-\frac{\overline{\sum d_{t+1}^2}}{g^2}\right) \tag{16}$$

d_{t-1} and d_{t+1} are the difference of the total ozone value between the ith reference pixel in the target and auxiliary data:

$$d_{t-1} = \left\| I_{t-1}\left(x_i^R, y_i^R\right) - I_t\left(x_i^R, y_i^R\right) \right\|^2 \tag{17}$$

$$d_{t+1} = \left\| I_{t+1}\left(x_i^R, y_i^R\right) - I_t\left(x_i^R, y_i^R\right) \right\|^2 \tag{18}$$

g in equations (14) and (15) is determined by the experimental data. For our study, we use g as

$$g = 2 \times \sigma \tag{19}$$

where σ is the standard deviation of all the reference pixels in the target data.

The remaining 10% of the missing pixels which cannot find valid values simultaneously existing at the target location in both the auxiliary data sets are recovered by a single data set acquired on $t-1$ or $t+1$. A regression equation needs to be established, such as equation (1) or (2). We search for reference pixels in the target and auxiliary data, compute the regression coefficients, and then the predicted $I'_{t-1}(x,y)$ or $I'_{t+1}(x,y)$ is the preliminary prediction. If there are no measurements for the target pixel in the days before and after the target data, the target pixel will also be identified as an outlier. After the other missing pixels have been filled, the outliers are recovered by performing kriging using pixels in their neighborhood, where pixels that have already been filled by TFFSRC in previous iterations can be used to establish the interpolation model.

3.2. The Second Step Residual Correction Using the Kriging Method

After the coefficients in equation (13) have been determined, we can obtain a preliminary predicted value of the target pixel. There may, however, be considerable errors because of the dynamic changes caused by the effect of atmospheric circulation. To alleviate this error, the second step of the TFFSRC method is executed: error correction by kriging [*Gundogdu and Guney*, 2007; *Mandallaz*, 2000]. Here we only provide a brief description of kriging and the anisotropic variogram, as these approaches have been widely used and described in the geostatistical literature. Therefore, the interested reader is referred to the works of *Journel and Huijbregts* [1978] and *Cressie* [1993] for further details.

3.2.1. Kriging

Kriging is a method based on regionalized variable theory, which can be used to analyze the spatial continuity and distribution of valid data and provide estimations at missing locations. It is based on a linear combination of observations of the phenomena in the vicinity of the desired location of the estimation, without bias and with minimization of the variance. The main idea of error correction by kriging is to evaluate the residual value in the target location by a linear combination of prediction residuals of all the reference pixels in the target data. Therefore, we should start by computing the prediction residuals of all the reference pixels in the target data by applying the AWTF method.

For each reference pixel in the target data, we can obtain its preliminary prediction using the AWTF method. Similar to equation (13), this can be displayed as

$$\hat{I}_t\left(x_i^R, y_i^R\right) = w_{t-1} \times I'_{t-1}\left(x_i^R, y_i^R\right) + w_{t+1} \times I'_{t+1}\left(x_i^R, y_i^R\right) \tag{20}$$

where $I'_{t-1}\left(x_i^R, y_i^R\right)$ and $I'_{t+1}\left(x_i^R, y_i^R\right)$ are the regression results computed by equations (1) and (2), respectively, with the location of $\left(x_i^R, y_i^R\right)$ instead of (x, y). Because of the similar linear relationships shared with the target pixel and its reference pixels, not all the coefficients in equation (20) are changed. The prediction residual of the ith reference pixel is computed by subtraction of the true and predicted values:

$$\varepsilon_t\left(x_i^R, y_i^R\right) = I_t\left(x_i^R, y_i^R\right) - \hat{I}_t\left(x_i^R, y_i^R\right) \tag{21}$$

where $I_t\left(x_i^R, y_i^R\right)$ is the true value of the ith reference pixel in the target data.

The evaluation of the residual value of the target pixel can be obtained by utilizing the kriging method:

$$\hat{\varepsilon}_t(x_0, y_0) = \sum_{i=1}^{N} \lambda_i \times \varepsilon_t\left(x_i^R, y_i^R\right) \tag{22}$$

where $\varepsilon_t\left(x_i^R, y_i^R\right)$ is the prediction residual of the ith reference pixel obtained by equation (21). λ_i is the weight of the ith reference pixel in the target data associated with each prediction residual and is chosen to make an unbiased estimation with minimum variance [*Wackernagel*, 2003]. However, the literature [*Stein*, 2007] has shown that the intrinsic stationarity assumption, a theoretical premise of the geostatistical methods, is not always valid, especially for heterogeneous landscapes. Concerning this issue, *Chang et al*. [2015] presented a possible solution. In our study, we assume that the prediction residuals satisfy the intrinsic stationarity assumption, and then the coefficients λ_i in equation (22) can be solved by variogram γ.

3.2.2. Variogram Analysis and Modeling

In the kriging method, the variogram is experimentally calculated. Then, in order to apply the kriging to make an estimation, a theoretical variogram model needs to be set up. The appropriate model is chosen by matching the shape of the curve of the experimental variogram to the shape of the curve of the mathematical function. The practice of variogram modeling has been covered in many texts [e.g., *Wackernagel*, 2003], and the common method of modeling is weighted least squares fitting. Among the typical variogram models, a popular model for the prediction residuals is the spherical variogram model.

According to the structural features of spatial variograms, the random fields can be divided into two categories, i.e., isotropic and anisotropic. Isotropic variogram analysis is direction independent, whereas anisotropic variogram analysis is direction dependent. According to the relevant literature [*Stein*, 2007], it has been revealed that the spatial dependence of total ozone varies strongly with latitude. Figure 4 displays the same phenomenon, in that the global distribution of total ozone exhibits a strongly latitudinal dependence, which means that the variation of total ozone is much stronger across latitudes than across longitudes, especially at high latitudes in both the Northern and Southern Hemispheres. Therefore, the isotropic model is clearly unsuitable. However, the theoretical variogram models used for kriging are based on the isotropic model, so correction for anisotropy is necessary to make use of kriging methodology. Anisotropy can be further classified into two types: geometric and zonal anisotropy. Geometric anisotropy occurs when the range, but not the sill, of the variogram changes in different directions, while zonal anisotropy exists when both the sill and range of the variogram change with the direction.

For instance, we choose a missing pixel at high latitude in the Northern Hemisphere, and we obtain the prediction residuals of its reference pixels by performing the AWTF method to diagnose their type of anisotropy. It is a fact that the south-north (S-N) direction and the east-west (E-W) direction are identified as roughly the directions of the maximum and minimum spatial correlation, respectively. Figure 2 plots the separate experimental and theoretical variograms calculated for the maximum (S-N) and minimum directions (E-W) of all the prediction residuals. The tolerance of the angle is $\pm 30°$, and the variogram model is fitted by a spherical model. The E-W direction variogram and the S-N direction variogram are given by experimental variograms (symbols) and fitted models (solid lines) computed from a field of 50 values simulated with a spherical variogram function with zero nugget. As can be seen when comparing the variograms of the two directions, changes in the S-N direction are more rapid, representing higher variance or sill value, compared with relatively smooth changes in the E-W direction, representing a smaller sill value. These phenomena indicate that the variation of all the prediction residuals of the reference pixels presents zonal anisotropy.

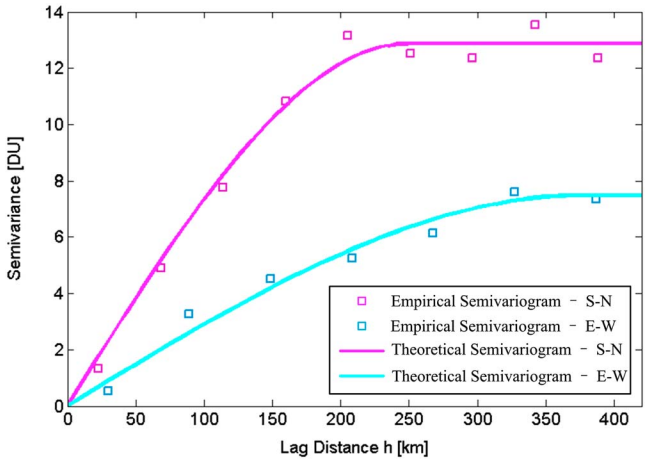


Figure 2. Empirical and theoretical semivariograms in the S-N direction and in the E-W direction.

The practice in modeling a variogram that appears to present zonal anisotropy is to first estimate the axes of the anisotropy and the degree of the anisotropy (i.e., the ratio of the maximum range to the minimum range) and, based on these estimates, to transform the coordinate system to make the two ranges identical. The empirical variogram with a small sill in all directions is then modeled as a geometrically anisotropic variogram and is finally used to create a nested anisotropic model [*Zimmerman*, 1993]:

$$\gamma(\mathbf{h}) = c_{EW} \times \left\{ \frac{3}{2} \times \frac{\sqrt{h_{EW}{}^2 + (K \times h_{SN})^2}}{\text{range}_{EW}} - \frac{1}{2} \times \left[\frac{\sqrt{h_{EW}{}^2 + (K \times h_{SN})^2}}{\text{range}_{EW}} \right]^3 \right\} + (c_{SN} - c_{EW}) \times \left[\frac{3}{2} \times \frac{h_{SN}}{\text{range}_{SN}} - \frac{1}{2} \times \left(\frac{h_{SN}}{\text{range}_{SN}} \right)^3 \right] \tag{23}$$

where c_{EW}, range_{EW}, and h_{EW} are, respectively, the sill, range, and the lag distance in the E-W direction and c_{SN}, range_{SN}, and h_{SN} are the sill, range, and the lag distance in the S-N direction. K is the anisotropy ratio, i.e., the ratio of the maximum range to the minimum range.

3.2.3. The Final Prediction

The target residual prediction is in practice a weighted moving average, in which the weights depend on the variogram and the configuration of the reference pixel residuals within the neighborhood of their targets. Eventually, the final predicted value of the target pixel located at (x, y) in the gaps is obtained by summing the predicted trend and the residual:

$$\hat{I}(x, y) = \hat{I}_t(x, y) + \hat{\varepsilon}_t(x, y) \tag{24}$$

$\hat{I}_t(x, y)$ is the preliminary prediction value fitted by the multitemporal information through the selection of reference pixels to establish the linear regression relationship. $\hat{\varepsilon}_t(x, y)$ represents the target residual predicted from the interpolation of the residuals of the reference pixels by the kriging method. Figure 3 presents a flowchart of the proposed filling method.

4. Results

4.1. The Long-Term Series of the Total Ozone Product

Based on the proposed TFFSRC algorithm, the long-term series of the total ozone product was generated by reconstructing the missing data of OMTO3e from October 2004 to 2014. The coverage of this product is almost worldwide, apart from the regions where the sensor cannot work properly due to the lack of UV radiation at high latitudes in different seasons. The loss at high latitudes, where we were unable to find auxiliary information in another time series, is beyond the scope of our consideration. This total ozone product is hosted on http://sendimage.whu.edu.cn/send-resource-download/ for free download.

As examples, the OMI OMTO3e data sets with missing data dated 15 January, 14 April, 15 July, and 15 October in 2013 and their corresponding retrieved product maps are shown in Figure 4. The data sets reasonably represent the global total ozone characteristics and their time series dynamics. The global distribution of total ozone exhibits substantial seasonal and latitudinal dependence. The minimum value of total ozone is found in the tropics, and the maximum value is found near the polar regions. The significant gradients found in midlatitude regions and the obvious spatiotemporal changes are mainly caused by stratosphere photochemical

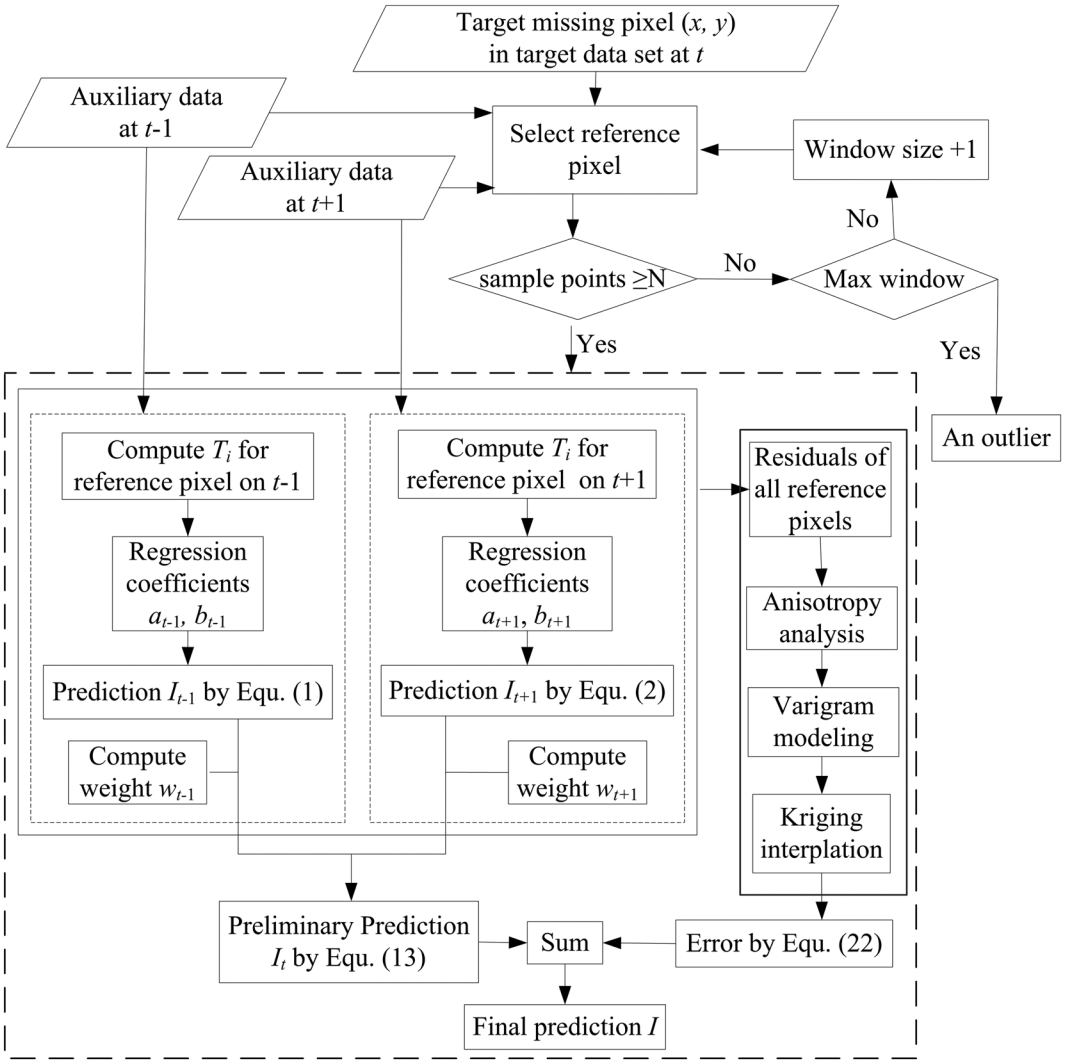


Figure 3. The flowchart of the TFFSRC algorithm.

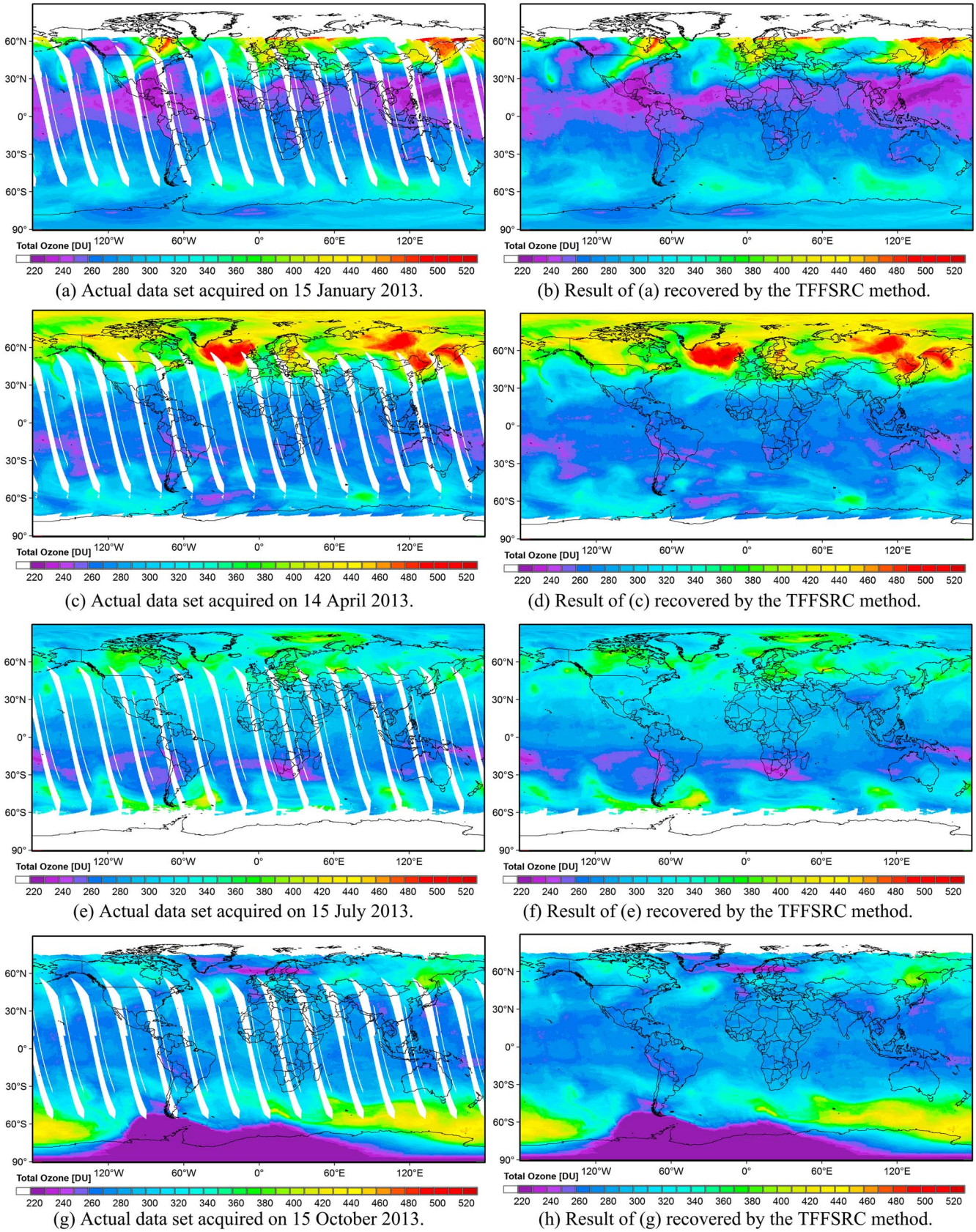
processes and the synoptic-scale weather patterns. Indeed, this significant variation of ozone distribution poses a significant challenge when filling the gaps in ozone maps.

4.2. Evaluation Results

In order to obtain an adequate evaluation, both simulated experiments and validation with ground-based measurements were implemented.

4.2.1. Simulated Experiments

Simulated masks of missing data were defined by degraded data acquired after 2006 and then applied to the selected data with a good quality from 2004 to 2006 to obtain the simulated test data. To explain the process of the simulated tests, an example is shown in Figure 5. The locations of the gaps in Figure 5a were marked as a simulated mask (Figure 5b), which was applied to the target image (Figure 5c), masked out and considered as the missing data in Figure 5d, to provide the data for a quantitative validation. Simulated experiments with global total ozone data sets containing 10 consecutive sets of data in each group were implemented to test and verify the quantitative accuracy of the different methods. Global data sets of different years, seasons, and amounts of missing pixels were selected to form three groups of simulated experiments, thereby ensuring a comprehensive validation of the proposed approach for all the OMTO3e data. Ten consecutive sets of data acquired from 22 to 31 December 2004 were chosen as the first simulated test group. For these data sets, simulated gaps were made by the use of 10 consecutive sets of actual OMTO3e data acquired from 22 to



(a) Actual data set acquired on 15 January 2013.

(b) Result of (a) recovered by the TFFSRC method.

(c) Actual data set acquired on 14 April 2013.

(d) Result of (c) recovered by the TFFSRC method.

(e) Actual data set acquired on 15 July 2013.

(f) Result of (e) recovered by the TFFSRC method.

(g) Actual data set acquired on 15 October 2013.

(h) Result of (g) recovered by the TFFSRC method.

Figure 4. Examples of actual data sets and the corresponding total ozone product.

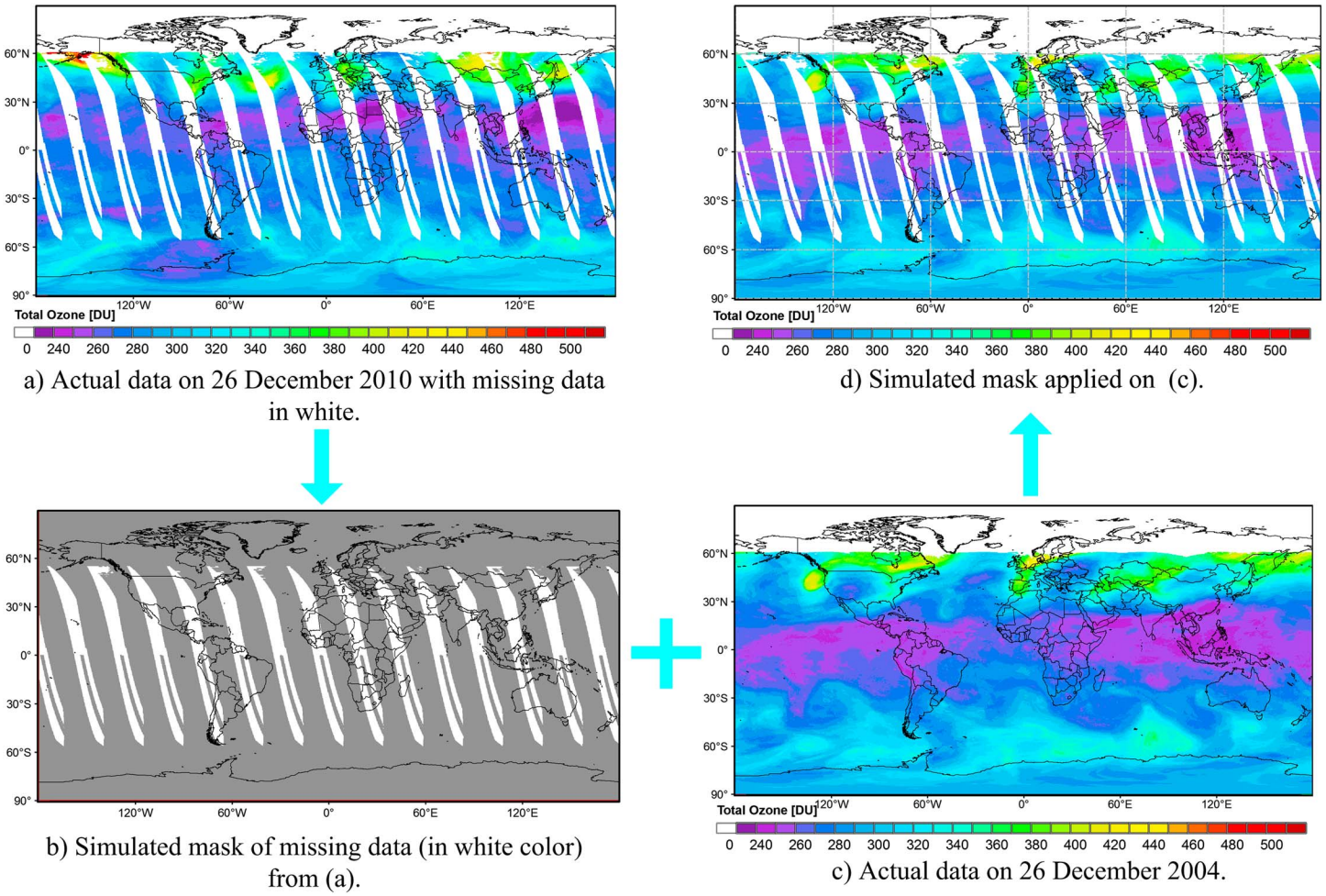
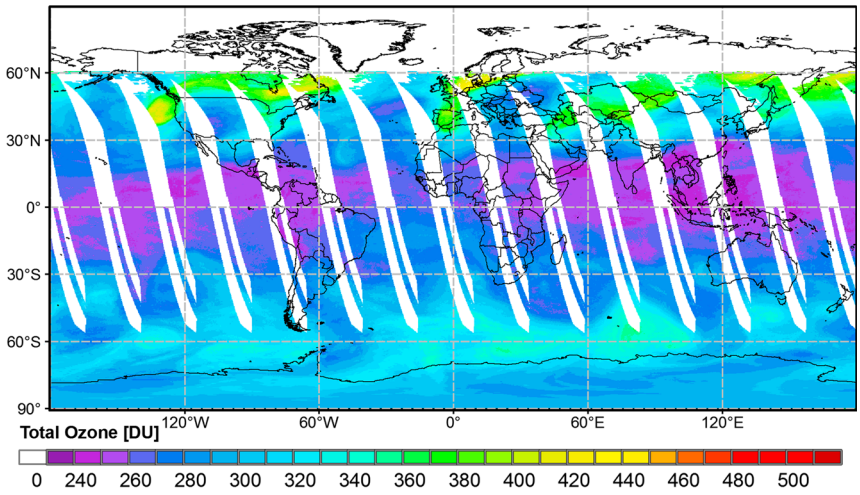


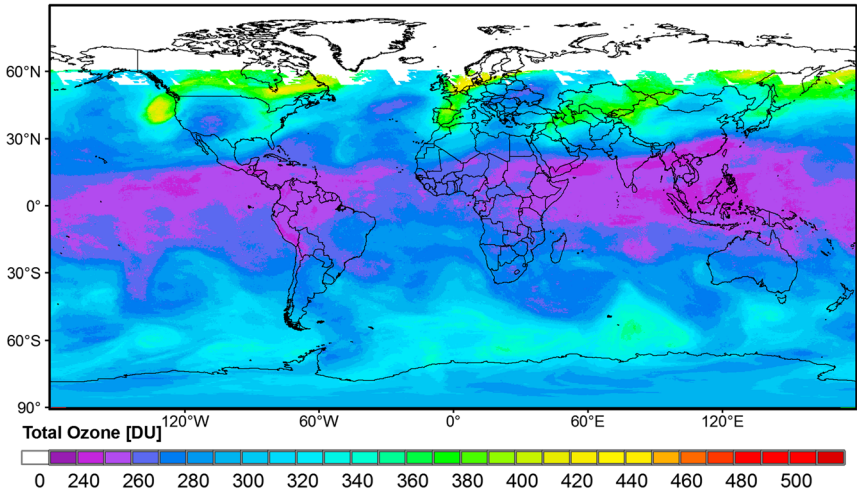
Figure 5. Simulated mask of the degraded data applied to the good-quality data for the quantitative assessment.

31 December 2010. In winter, the seasonal missing region is in the Northern Hemisphere at high latitudes. The data size was 576 × 1440, with about 230,000 missing pixels. The second test group of data was acquired from 20 to 29 July 2005, and the simulated gaps were covered using OMTO3e data acquired from 20 to 29 July 2011. In contrast to the first group, this group of data is from the summer, when the temporal and spatial variation trend of total ozone is different from that of the winter. The total ozone changes are mostly concentrated at middle latitudes in the Southern Hemisphere and are weaker than those in the first group of simulated experiments. The size of the target data was 576 × 1440, with about 180,000 missing pixels. In order to determine the applicability of the proposed method in different times, we selected the data acquired from 28 January to 6 February 2006, and the 28 January to 6 February 2012 data were used to simulate the gaps in the third group of simulated experiments. February can be considered as the end of winter and the beginning of spring. The size of the third simulated data set was 601 × 1440, with about 130,000 missing pixels. These three groups of simulated experiments were performed to validate the proposed method by comparing it with three other effective gap filling methods: kriging, WLR, and DCT-PLS.

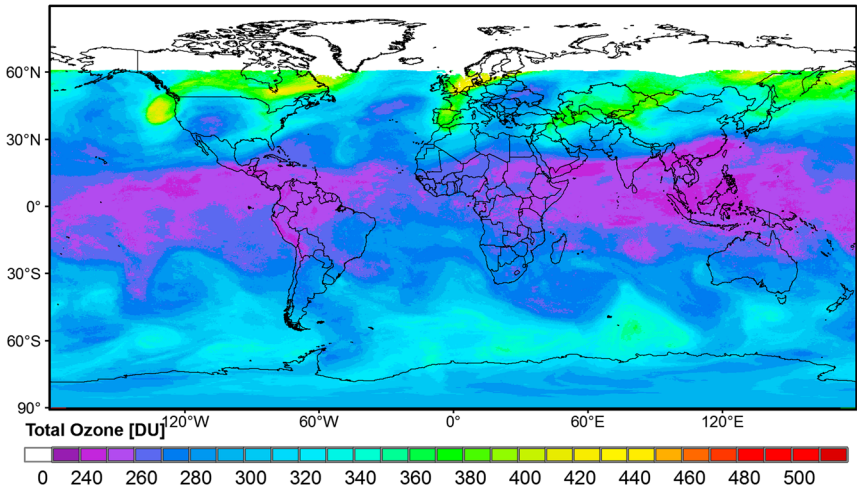
An example of the first group of simulated experiments is shown in Figure 6 for a visual inspection. The target simulated data on 27 December 2004 are shown in Figure 6a, where the width of each gap can reach 10° longitude over about 1000 km. Figure 6b shows the reconstruction result of Figure 6a by the use of TFFSRC. The actual OMTO3e data shown in Figure 6c are used to validate the proposed algorithm by comparing the filled result with the actual data. In order to allow a clear visual inspection, Figure 7 shows an example from each group of simulated experiments, with two zoomed regions to show the difference between the actual data and the reconstructed results by the different methods. The first two rows of the panels show the total ozone distribution in the regions of the South Atlantic Ocean and Mongolia from the first group



(a) Target data on 26 December 2004 simulated using the gaps on 26 December 2010.



(b) Result recovered by the TFFSRC method.



(c) Actual data set acquired on 26 December 2004.

Figure 6. An example for the simulated experiments and the result recovered by TFFSRC compared with the actual data.

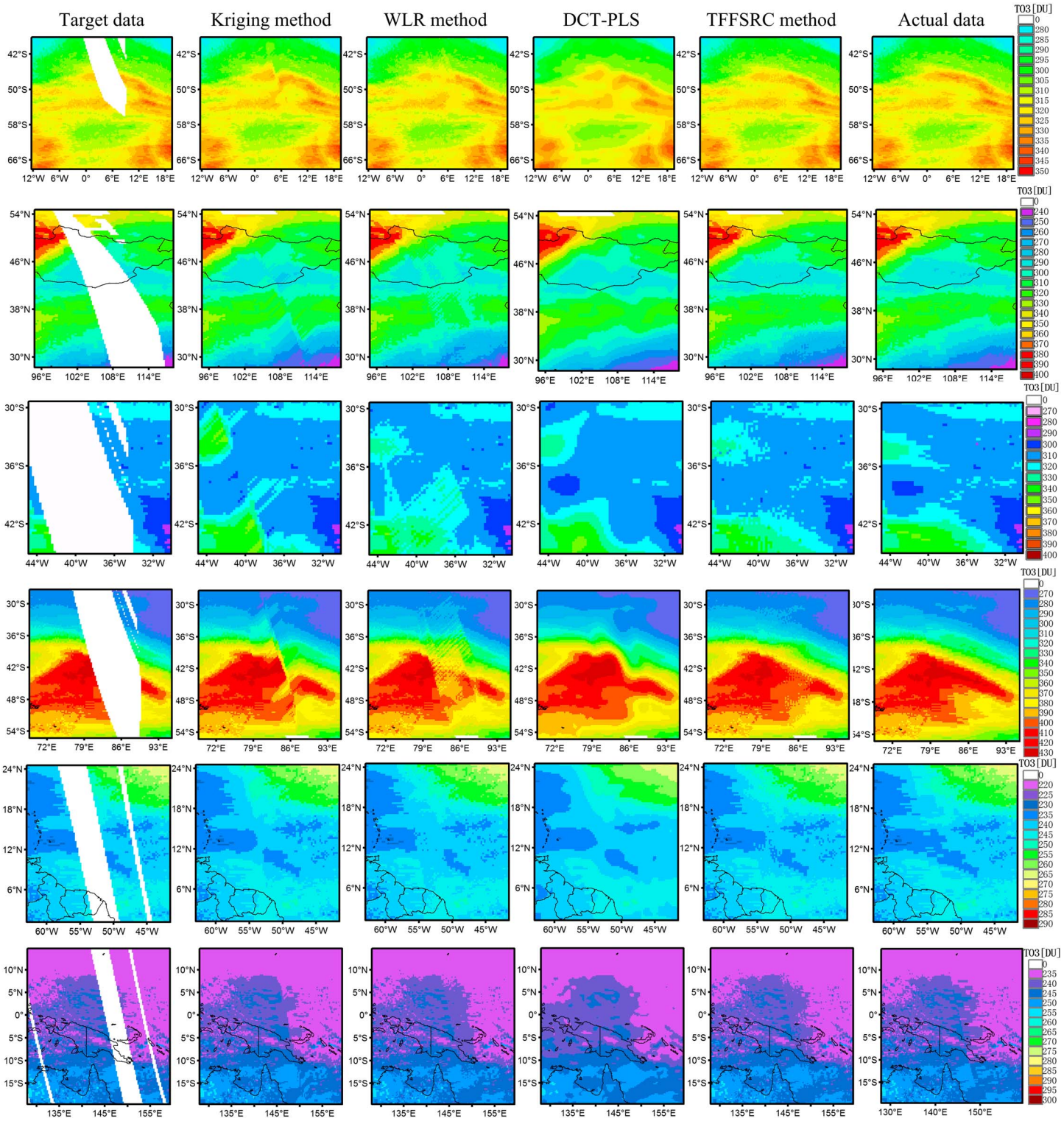


Figure 7. Comparison of the actual total ozone data and the reconstructed results for six detailed regions.

of simulated experiments dated 27 December 2004. The middle two rows of the panels show the regions of the South Atlantic Ocean and Indian Ocean from the second group of simulated experiments dated 26 July 2005. The last two rows of the panels show the regions of the North Atlantic Ocean and the North Pacific Ocean from the third group of simulated experiments dated 3 February 2006. The panels from left to right

show the target data, the reconstructed results by the kriging, WLR, DCT-PLS, and TFFSRC methods, and the actual total ozone data of the corresponding regions. Only using the spatial information cannot lead to an ideal retrieved result, which can be seen in Figure 7 for the result recovered by kriging. Although the kriging method is a relatively mature statistical method for spatial interpolation, which can generally obtain good results, it only uses spatial information, and the weight is mainly determined by the variation function of the distance. For the excessively wide gaps in this experiment, the effect is poor. Compared to the actual data, the result is distorted and shows an obvious spatial discontinuity effect. The reconstructed result is also oversmoothed. For the WLR algorithm, the results also show obvious spatial discontinuity and lots of noise. Because the different auxiliary data sets make the same contribution in the prediction, the spatial correlation from the reference pixels of the target data was not taken into consideration. The results for the DTC-PLS method are mostly well recovered, but an obvious smoothing effect appears. Furthermore, the values of all the valid original data are changed after performing the DTC-PLS method. Through the six zoomed regions, it can be clearly seen that the results retrieved by TFFSRC are the closest to the actual data, when compared to the results of the other methods. In the TFFSRC results, the missing areas are filled more thoroughly, with great spatial continuity and a convincing visual quality.

Figure 8 shows the scatterplots of the relationships between the actual and the predicted values corresponding to the three exampled simulated experiments. It is worth mentioning that only the values of the simulated missing pixels participate in the calculation. In general, the data points in the scatterplots of the proposed method are more concentrated around the diagonal than those of the other three methods. For kriging and WLR, the points in the scatterplots are close to the 1:1 line, with some outliers, while the plots of the DCT-PLS method show the worst result. It is apparent that the scatterplots in Figures 8i–8l appear to have a better agreement with the actual data than the other panels, which is probably because the total number of missing pixels in the third test group is significantly lower than in the other two groups. These plots fully demonstrate that the TFFSRC method can accurately predict the missing data at the pixel level, which is consistent with the results of the visual observation. Because of the greater fluidity of ozone in space-time, the use of only a single piece of information of space or time as a reference will lead to inaccurate results.

The results listed in Table 1 show the quantitative comparisons of these three groups of simulated tests by the means of the values for the three periods. The AWTF method, the first step of the TFFSRC method we propose in this research, is also compared with the four methods mentioned above, to confirm whether error correction by kriging as the second step can improve the accuracy of the prediction process. The root-mean-square error (RMSE) values of the TFFSRC method are the lowest. It is evident that the precision of the restored results using TFFSRC is satisfactory, which is consistent with the analyses implemented above. The AWTF method performs better than WLR and kriging, which provides sufficient proof to show that the residual correction is an effective process to promote the reconstruction results. Through the contrast between the three groups of simulated experiments, it can be seen that the second group has the widest gaps and the largest amount of missing pixels, where the DCT-PLS, WLR, and kriging methods are quite sensitive, producing larger RMSE values, which means greater errors. Meanwhile, TFFSRC is not sensitive to the range of the missing gaps as it obtains a lower mean value of RMSE in the second group than in the other two groups of experiments. The RMSEs of the TFFSRC and AWTF methods are the largest in the third group of experiments, because data in this group were selected in February, and the spatial and temporal liquidity of the total ozone is considerably higher than for the other months, while the kriging method is not affected. This indicates that when total ozone data do not have wide missing gaps, the kriging method can obtain a better effect than the WLR method. It should be mentioned that the DCT-PLS method in this study only used one data set (the target data) as input data to complete the prediction. The reason for this is that, for OMTO3e data, one data set as input data can obtain the best prediction result, which is demonstrated in Figure 9.

To further explore the performance of the TFFSRC method, the other famous spatiotemporal method, the STkriging method [*Cressie and Wikle*, 2011], was also compared. Taking into consideration the computational complexity of the matrix inversion, and to speed up the processing, three local regions were selected to perform the comparisons. The size of the three regions was 200×200, and they respectively comprised the American mainland ((62°N, 12.25°N), (124.75°W, 75°W)), the South Atlantic Ocean ((0.25°S, 50°S), (39.75°W, 10°E)), and Australia ((1°S, 50.75°S), (108.25°E, 158°E)). To guarantee the rationality and universality of the proposed algorithm, the three local regions of data were chosen from different time periods to implement the simulated experiments. The method of generation for the simulated experiment data was similar to the

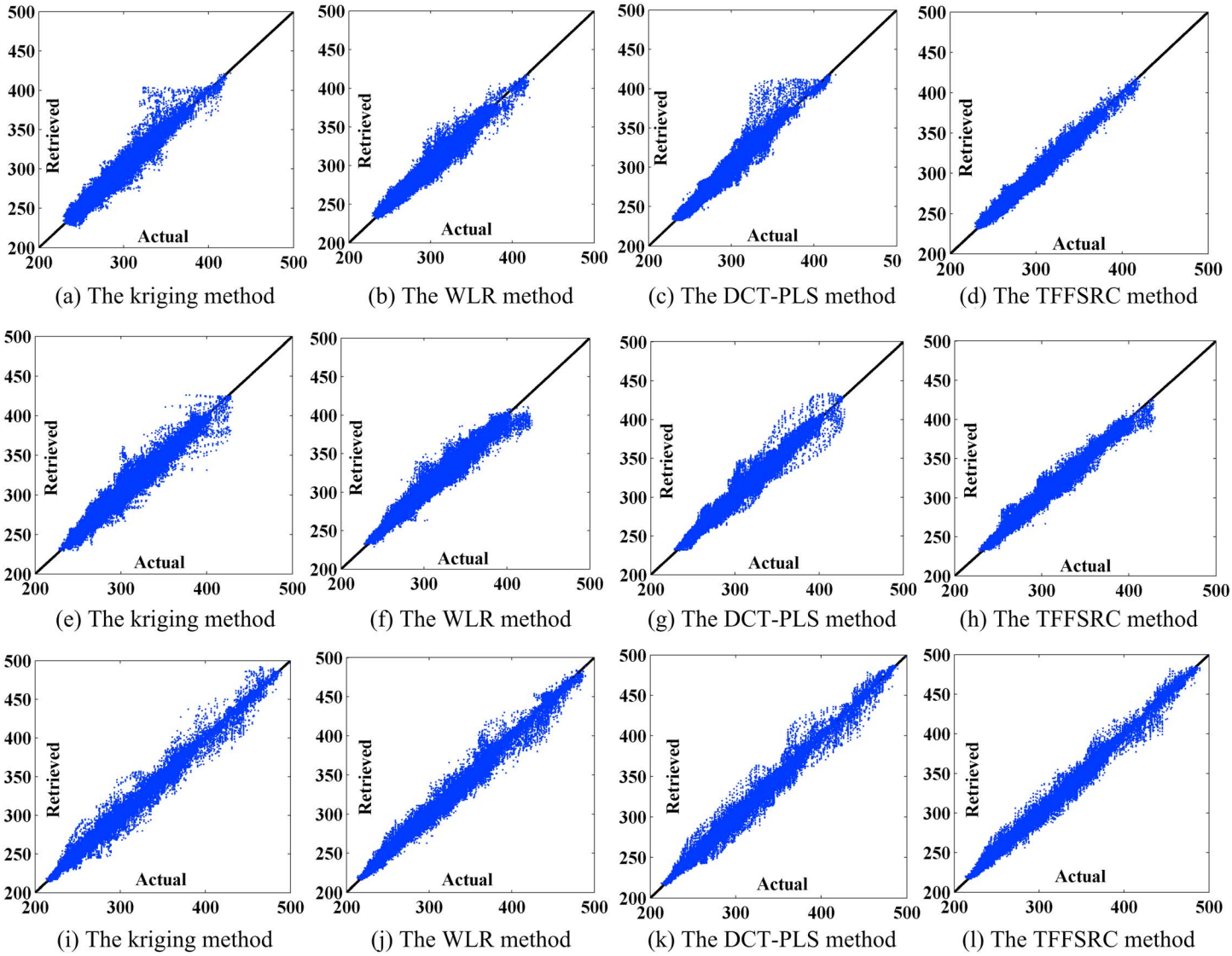


Figure 8. Scatterplots of the actual and the retrieved values of all the missing pixels for the three exampled simulated experiments: (a–d) the first example, (e–h) the second example, and (i–l) the last example.

process described in section 4.2.2. Since the OMTO3e data in 2005 are unique data sets with a good quality over the whole year, data sets from 2011 were used to simulate the missing gaps, which were applied to the data sets in 2005 for the generation of the simulated data. For the American mainland area, the simulated data on 27 January were selected as the target data, and the simulated data from one day before and after the target data were selected as the auxiliary data to perform the TFFSRC method. The simulated data from 21 January to 4 February were selected as the input data to perform the STkriging method. In the same way, for the South Atlantic Ocean area, the simulated data on 12 October were selected as the target data, and for Australia, the simulated data on 8 August were selected as the auxiliary data.

Table 1. The Accuracy by the Means of the RMSE Values for the Three Groups of Simulated Experiments, as Restored by the TFFSRC, DCT-PLS, AWTF, WLR, and Kriging Methods

Date/Method	TFFSRC	DCT-PLS	AWTF	WLR	Kriging
2004	4.0907	4.6055	4.6697	5.0006	5.1930
2005	4.0811	4.9134	4.9125	5.2758	5.4892
2006	4.3945	4.8503	5.1041	5.2450	5.1623

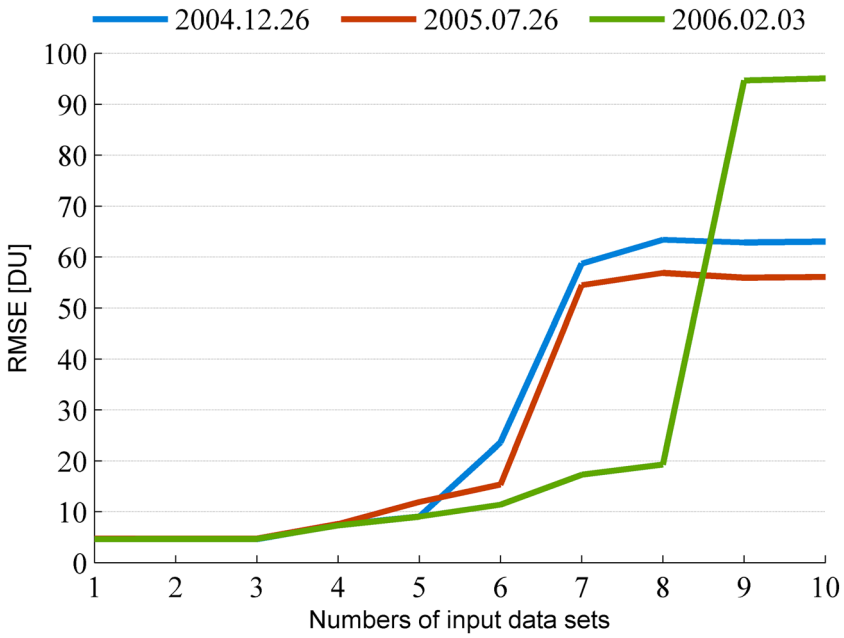


Figure 9. Comparison of the DTC-PLS method with different numbers of input data sets.

As described in the literature [*Chipperfield et al.*, 2003], atmospheric total ozone presents both a temporal trend (seasonal cycle) and a nonconstant spatial trend. Therefore, the spatiotemporal random variables for each local area can be decomposed into a mean component and a residual component [*Cressie and Wikle*, 2011; *De Iaco and Posa*, 2012]. In this study, the deterministic spatiotemporal mean component is further decomposed into a temporal trend and a spatial trend, as described in *Kyriakidis and Journel* [1999]. The estimated spatiotemporal mean component is then subtracted from the full simulated data set to yield the spatiotemporal residual component for each local area, which is used in the following analysis. For all three local areas, the sum-metric model is the most suitable variogram model. Once the best fitting spatiotemporal variogram model is identified, the STkriging interpolation can be executed in universal transverse Mercator projection to obtain the residual prediction at the missing location. Finally, the prediction at the missing location is gained by adding the spatiotemporal trend to the residual prediction. The process is implemented in the *R* environment for geostatistical modeling and interpolation and uses the "gstat" [*Pebesma*, 2004] and "stats" [*Team*, 2014] packages. For all three local areas, the sum-metric model is the most suitable variogram model.

The performances of the TFFSRC and STkriging methods for the three local areas in terms of RMSE are shown in Figure 10. In terms of the histogram in Figure 10, the STkriging method shows hardly any benefit when compared with the pure spatial kriging method. As the number of auxiliary data sets increases in the STkriging method (STkriging5, STkriging7, STkriging9, STkriging11, and STkriging13, where the number

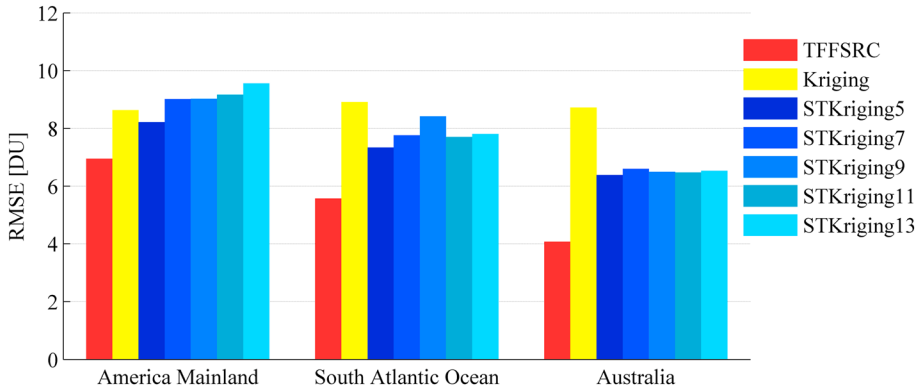


Figure 10. Evaluation of the TFFSRC, STkriging, and spatial kriging methods in the three local areas.

Table 2. Summary of All the Ground-Based Brewer Instruments Used in This Study

Station Name	Latitude	Longitude	Used Coincidences
STN018 (Alert, CAN)	82.45°N	62.51°W	1301
STN021 (Edmonton, CAN)	53.55°N	114.1°W	2933
STN024 (Resolute, CAN)	74.72°N	94.98°W	1331
STN035 (Arosa, CHE)	46.78°N	9.68°E	83
STN065 (Toronto, CAN)	43.78°N	79.47°W	2916
STN076 (Goose Bay, CAN)	53.31°N	60.36°W	2462
STN077 (Churchill, CAN)	58.74°N	94.07°W	2330
STN111 (Amundsen-Scott, ATA)	89.99°S	70.24°E	850
STN282 (Kislovodsk, RUS)	43.73°N	42.66°E	2411
STN290 (Saturna Island, CAN)	48.77°N	123.13°W	3065
STN307 (Obninsk, RUS)	55.12°N	36.3°E	1717
STN315 (Eureka, CAN)	79.988°N	85.94°W	1207
STN316 (DeBilt, NLD)	52.10°N	5.18°E	2188
STN338 (Bratts Lake, CAN)	50.21°N	104.71°W	364
STN468 (Cape D'Aguilar, HKG)	22.21°N	114.26°E	501
STN479 (Aosta, ITA)	45.74°N	7.36°E	1759
STN481 (Tomsk, RUS)	56.48°N	85.07°E	1559
STN512 (University of Toronto, CAN)	43.66°N	79.4°W	1163

means the number of data sets involved in the calculation), the value of the RMSE also increases. This effect can, to a large degree, be explained by the spatiotemporal variograms: a temporal lag of one or a few days leads to a large variability compared to spatial distances of a few hundred kilometers, implying that the temporal correlation is too weak to significantly improve the overall prediction. It should be noted that when the number of data sets is less than five, the STkriging method cannot be performed because there is insufficient information to establish the spatiotemporal variogram. With the lowest value of RMSE, there is no doubt that the proposed TFFSRC method is the most effective for the gap filling purposes in OMTO3e data.

Through all of the above simulated tests selected from different years and seasons, it can be seen that the global total ozone values change with time and space, both in summer and winter, and they change most severely in the middle and high latitudes of the Northern Hemisphere. These great changes result in the reconstructed results not being ideal; however, all the results, including the visual results of the proposed method, are closer to the actual values.

4.2.2. Evaluation Using Ground-Based Ozone Observations

Ground-based instruments measuring long-term total ozone provide an excellent reference and were employed for the validation of the TFFSRC product. The ground-based total ozone observations used in this paper were obtained from the World Ozone and Ultraviolet Radiation Data Center (WOUDC) archive (http://www.woudc.org/). The ground-based stations of the WOUDC independently measure overhead total ozone columns using Dobson, Brewer, and Filter instruments. However, all the Dobson total ozone data should be applied with a temperature correction [*Kerr*, 2002]. The random measurement errors (or "noise") of the Filter instruments are significantly higher than those of the collocated Brewer or Dobson instruments. Therefore, in this study, the data from the Brewer network were adopted as the primary reference. To investigate the quality and the performance of the TFFSRC product, comparisons between the MERRA-2 assimilated product available online through the Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (http://disc.sci.gsfc.nasa.gov/mdisc/) and the ground-based Brewer total ozone columns were also performed. First, the ground-based data in the time series that appeared odd (i.e., sudden jumps) were rejected. The Aura spacecraft has an equatorial crossing time of 13:45 (ascending node), with about 98.8 min per orbit (14.6 orbits per day on average). The time-averaged total ozone observations for each ground-based station

Table 3. The Results of the Statistical Comparisons of the TFFSRC Product and the MERRA-2 Product With the Brewer Spectrophotometers' Total Ozone Column Data Sets

	Spatial Resolution (Latitude × Longitude)	RMSE (DU)	CC	Bias (DU)	RB (%)
TFFSRC product	0.25° × 0.25°	12.51	0.98	5.13	1.46
MERRA-2 assimilated product	0.5° × 0.625°	13.25	0.97	6.49	1.88

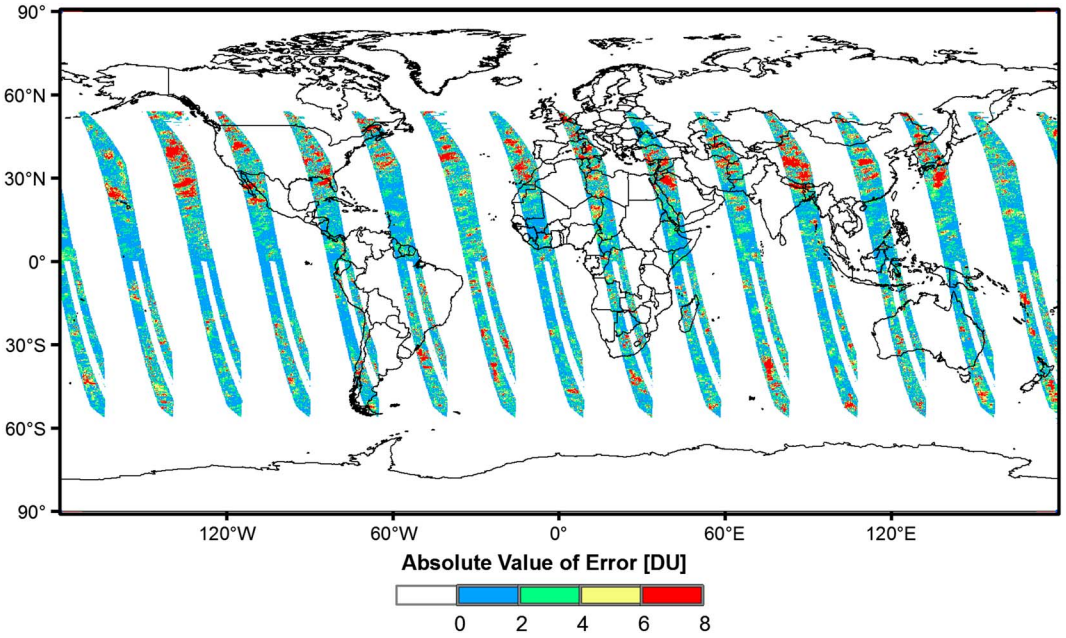


Figure 11. The absolute values of the error.

were then computed by the extraction of the total ozone data when the local time was between 13:00 to 14:00. Meanwhile, the TFFSRC product that has the center of its gridded pixel closest to the ground station was extracted. The UTC time for each station was then obtained and used to extract the corresponding hourly averaged MERRA-2 assimilated data named M2T1NXSLV, also following the principle of nearest space distance. Using these criteria for the October 2004 to December 2014 timeframe, we obtained 30140 sonde-TFFSRC-MERRA-2 assimilation measurement triads. The number of triads considered for the validation was the subset of the data common to the three data sets, which meant that the data in all three data sets were valid. Table 2 give a summary of the ground-based Brewer instruments used in this study, their location, and the number of used coincidences, which means the number of coincidences among the Brewer spectrophotometers' total ozone columns, the TFFSRC product, and the MERRA-2 assimilated product used in the comparison.

The results of the statistical comparisons of the TFFSRC product and the MERRA-2 assimilated product with the Brewer spectrophotometers' total ozone columns are summarized in Table 3 in terms of RMSE, the corresponding correlation coefficient (CC), bias, and relative bias (RB). The ozone bias is defined as the mean value of the ground-based observations minus the TFFSRC product data or the MERRA-2 assimilated product data, and RB (%) is computed with respect to the Brewer total ozone data. The TFFSRC product is biased low by 5.13 DU (Dobson unit) (1.46%) with respect to the Brewer total ozone data. The RMSE and the CC between the TFFSRC product and the Brewer total ozone data are 12.51 DU and 0.98, respectively. However, the four statistical factors listed above between the MERRA-2 assimilated product and the Brewer total ozone data are 6.19 DU, 1.88%, 13.25 DU, and 0.97, respectively. We can see that the TFFSRC product measurements show a better correlation with the Brewer total ozone data than the MERRA-2 assimilated product. The RMSE, bias, and RB between the TFFSRC product and the ground-based Brewer measurements are significantly reduced compared to the MERRA-2 assimilated product. These results show that the TFFSRC product shows no substantial difference with the ground-based Brewer measurements (WOUDC) and is better able to capture total ozone variation than the MERRA-2 assimilated product. Furthermore, it should be noted that the spatial resolution of the TFFSRC product is higher than the MERRA-2 assimilated product.

5. Conclusion and Discussion

Since September 2005, and especially after mid-2006, the OMI instrument row anomalies began to appear, and about one third of the OMI level-3 (gridded) total ozone product (OMTO3e) is missing. Due to the excellent quality of OMI relative to other ozone sensors, OMTO3e is an important data source for many of the

applications of total ozone. Therefore, it is appropriate and necessary to develop a more practical technique for the reconstruction of the wide missing areas in the current OMTO3e product. However, it is challenging to generate such a spatially continuous mapping of the daily global ozone product, because the original OMTO3e data have wide gaps and show dynamic variations in the time series. Considering the properties of the OMTO3e data, this paper has presented a new and effective temporal fitting followed by spatial residual correction (TFFSRC) method, which takes into account the spatial and temporal correlation of the total ozone values to generate a total ozone product with continuity, consistency, integrity, and reliability, in both the spatial and temporal scales. Two temporal auxiliary data sets are employed to provide a preliminary predicted value, which pays sufficient attention to the continuous changes of total ozone in the time scale. Error correction by the anisotropic kriging method, which gives consideration to the spatial autocorrelation, is then employed to correct the error resulting from the preliminary prediction.

When the target pixel cannot find enough reference pixels, or there are no measurements for the target pixel in the days before and after the target data, the target pixel is identified as an outlier. Outliers are recovered by performing kriging using pixels in their neighborhood, where pixels that have already been filled by TFFSRC in previous iterations can be used to establish the interpolation model. Since the outliers are in the minority, we do not distinguish the unfilled reference and the filled reference pixels. Based on the TFFSRC method, we generated a spatially continuous and daily global total ozone product from 2004 to 2014. Meanwhile, we have also provided an additional file for each daily data set, which allows the user to distinguish the TFFSRC-filled pixels flagged as 2 from the real measured pixels with valid values flagged as 1 and the real measured pixels with invalid values flagged as 0. This total ozone product and the flag files are hosted on http://sendimage.whu.edu.cn/send-resource-download/ for free download.

The application of a mask defined by the gaps of the degraded data on the target data allowed us to undertake a series of simulated experiments to implement a precision evaluation through the comparison of different filling algorithms. Global data of different years, seasons, and amounts of missing pixels were selected to implement the simulated experiments and to comprehensively confirm the validity and applicability of the proposed approach. All the simulated experimental results consistently showed that the results reconstructed by the proposed method were closer to the actual values. The RMSE was less than 5 DU, and the relative error was within 2%, reflecting an acceptable accuracy. Furthermore, long-term validation of the TFFSRC product with the Brewer spectrophotometers' total ozone columns from the WOUDC archive was implemented. The validation results show that the TFFSRC product shows no substantial difference with the ground-based Brewer total ozone column data sets and is better able to capture total ozone variation than the MERRA-2 assimilated product.

It should be noted that the above discussion is based on the global average error. Because of the influence of atmospheric circulation, the total ozone changes at middle and high latitudes are greater than at the equator. When examining the absolute values of the errors obtained by the TFFSRC method in the simulated experiment dated 26 December 2004, as shown in Figure 11, we can see that the absolute values of the errors at middle and high latitudes are greater than for the equator. Second, for the same area, the temporal variability of the total ozone is greater, and the precision is lower. As shown in Table 1, the third group of simulated experiments used data from February, and the fluidity of the total ozone in the time series was greater than for the other two groups of experiments; as a result, the RMSE values were the highest. Third, the proposed algorithm is not sensitive to the range of the missing areas when the ratio of the missing area to the whole global area is within 50%. However, with the OMTO3e product, when the continuous missing range is more than 50% (which mostly occurs when the OMI jumps to another measuring mode once a month with much smaller ground pixels and a narrow swath), the proposed TFFSRC method cannot satisfy the precision requirement, and the original OMTO3e data cannot be reconstructed.

Acknowledgments

This research was supported by the National High Technology Research and Development Program (863) of China (2013AA12A301) and the National Natural Science Foundation of China (41271376 and 41422108). The OMI/Aura total ozone data sets were downloaded from the NASA Goddard Earth Sciences Data and Information Services Center (http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omto3e_v003.shtml). The ground-based Brewer total ozone column data sets used in this study were provided by the World Ozone and Ultraviolet Data Centre (WOUDC). The data are publicly available from http://www.woudc.org. The hourly averaged MERRA-2 total ozone data named M2T1NXSLV are available online through the Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (http://disc.sci.gsfc.nasa.gov/mdisc/).

References

Bell, M. L., R. D. Peng, and F. Dominici (2006), The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations, *Environ. Health Perspect.*, *114*(4), 532–536.

Bhartia, P. K., R. McPeters, R. Stolarski, L. E. Flynn, and C. G. Wellemeyer (2004), A quarter century of ozone observations by SBUV and TOMS, paper presented at XX Quadrennial Ozone Symposium, Kos, Greece, 1–8 June.

Chang, K. L., S. Guillas, and V. E. Fioletov (2015), Spatial mapping of ground-based observations of total ozone, *Atmos. Meas. Tech.*, *8*(8), 3967–4009.

Chen, J., P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh (2004), A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter, *Remote Sens. Environ.*, *91*(3), 332–344.
Cheng, Q., H. Shen, L. Zhang, and P. Li (2014), Inpainting for remotely sensed images with a multichannel nonlocal total variation model, *IEEE Trans. Geosci. Remote Sens.*, *52*(1), 175–187.
Chipperfield, M., W. Randel, G. Bodeker, M. Dameris, V. Fioletov, R. Friedl, N. Harris, J. Logan, R. McPeters, and N. Muthama (2003), Global ozone: Past and future, *Sci. Assess. Ozone Depletion*, *2002*, 4.1–4.90.
Cressie, N. A. (1993), *Statistics for Spatial Data*, Wiley, New York.
Cressie, N., and C. K. Wikle (2011), *Statistics for Spatio-Temporal Data*, John Wiley, Hoboken, N. J.
Cressie, N., and H.-C. Huang (1999), Classes of nonseparable, spatio-temporal stationary covariance functions, *J. Am. Stat. Assoc.*, *94*(448), 1330–1339.
De Iaco, S., and D. Posa (2012), Predicting spatio-temporal random fields: Some computational aspects, *C. R. Geosci.*, *41*, 12–24.
Fishman, J., C. E. Watson, J. C. Larsen, and J. A. Logan (1990), Distribution of tropospheric ozone determined from satellite data, *J. Geophys. Res.*, *95*(D4), 3599–3617, doi:10.1029/JD095iD04p03599.
Garcia, D. (2010), Robust smoothing of gridded data in one and higher dimensions with missing values, *Comput. Stat. Data Anal.*, *54*(4), 1167–1178.
Gneiting, T. (2002), Nonseparable, stationary covariance functions for space–time data, *J. Am. Stat. Assoc.*, *97*(458), 590–600.
Goovaerts, P. (1997), *Geostatistics for Natural Resources Evaluation*, Oxford Univ. Press, New York.
Gundogdu, K. S., and I. Guney (2007), Spatial analyses of groundwater levels using universal kriging, *J. Earth Syst. Sci.*, *116*(1), 49–55.
Hadjinicolaou, P., J. Pyle, and N. Harris (2005), The recent turnaround in stratospheric ozone over northern middle latitudes: A dynamical modeling perspective, *Geophys. Res. Lett.*, *32*, L12821, doi:10.1029/2005GL022476.
Hartley, W. (1880), On the probable absorption of solar radiation by atmospheric ozone, *Chem. News*, *42*(268).
Heck, W. W., O. Taylor, R. Adams, G. Bingham, J. Miller, E. Preston, and L. Weinstein (1982), Assessment of crop loss from ozone, *J. Air Pollut. Control Assoc.*, *32*(4), 353–361.
Herman, J., R. McPeters, and D. Larko (1993), Ozone depletion at northern and southern latitudes derived from January 1979 to December 1991 total ozone mapping spectrometer data, *J. Geophys. Res.*, *98*(D7), 12,783–12,793, doi:10.1029/93JD00601.
Jain, S., B. Arya, A. Kumar, S. D. Ghude, and P. Kulkarni (2005), Observational study of surface ozone at New Delhi, India, *Int. J. Remote Sens.*, *26*(16), 3515–3524.
Jerrett, M., R. T. Burnett, C. A. Pope III, K. Ito, G. Thurston, D. Krewski, Y. Shi, E. Calle, and M. Thun (2009), Long-term ozone exposure and mortality, *N. Engl. J. Med.*, *360*(11), 1085–1095.
Journel, A. G., and C. J. Huijbregts (1978), *Mining Geostatistics*, Academic Press, London.
Kerr, J. B. (2002), New methodology for deriving total ozone and other atmospheric variables from Brewer spectrophotometer direct Sun spectra, *J. Geophys. Res.*, *107*(D23), 4731, doi:10.1029/2001JD001227.
Kiesewetter, G., B.-M. Sinnhuber, M. Weber, and J. Burrows (2010), Attribution of stratospheric ozone trends to chemistry and transport: A modelling study, *Atmos. Chem. Phys.*, *10*(24), 12,073–12,089.
Kulkarni, P. S., S. D. Ghude, and D. Bortoli (2010), Tropospheric ozone (TOR) trend over three major inland Indian cities: Delhi, Hyderabad and Bangalore, *Ann. Geophys.*, *28*(10), 1879–1885.
Kyriakidis, P. C., and A. G. Journel (1999), Geostatistical space–time models: A review, *Math. Geol.*, *31*(6), 651–684.
Levelt, P. F., G. H. van den Oord, M. R. Dobber, A. Malkki, H. Visser, J. de Vries, P. Stammes, J. O. Lundell, and H. Saari (2006), The ozone monitoring instrument, *IEEE Trans. Geosci. Remote Sens.*, *44*(5), 1093–1101.
Li, X., H. Shen, L. Zhang, H. Zhang, Q. Yuan, and G. Yang (2014), Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning, *IEEE Trans. Geosci. Remote Sens.*, *52*(11), 7086–7098.
Li, X., H. Shen, L. Zhang, and H. Li (2015), Sparse-based reconstruction of missing information in remote sensing images from spectral/temporal complementary information, *ISPRS J. Photogramm. Remote Sens.*, *106*, 1–15.
Lippmann, M. (1989), Health effects of ozone: A critical review, *Japca*, *39*(5), 672–695.
Loyola, D., R. Coldewey-Egbers, M. Dameris, H. Garny, A. Stenke, M. Van Roozendael, C. Lerot, D. Balis, and M. Koukouli (2009), Global long-term monitoring of the ozone layer—A prerequisite for predictions, *Int. J. Remote Sens.*, *30*(15–16), 4295–4318.
Ma, C. (2002), Spatio-temporal covariance functions generated by mixtures, *Math. Geol.*, *34*(8), 965–975.
Ma, C. (2003), Families of spatio-temporal stationary covariance models, *J. Stat. Plann. Inference*, *116*(2), 489–501.
Mandallaz, D. (2000), Estimation of the spatial covariance in universal kriging: Application to forest inventory, *Environ. Ecol. Stat.*, *7*(3), 263–284.
Pebesma, E., and B. Gräler (2015), Introduction to spatio-temporal variography.
Pebesma, E. J. (2004), Multivariable geostatistics in S: The gstat package, *C. R. Geosci.*, *30*(7), 683–691.
Ramanathan, V., L. Callis, and R. Boughner (1976), Sensitivity of surface temperature and atmospheric temperature to perturbations in the stratospheric concentration of ozone and nitrogen dioxide, *J. Atmos. Sci.*, *33*(6), 1092–1112.
Ramanathan, V., R. J. Cicerone, H. B. Singh, and J. T. Kiehl (1985), Trace gas trends and their potential role in climate change, *J. Geophys. Res.*, *90*(D3), 5547–5566, doi:10.1029/JD090iD03p05547.
Reich, P. B., and R. G. Amundson (1985), Ambient levels of ozone reduce net photosynthesis in tree and crop species, *Science*, *230*(4725), 566–570.
Riahi, K., A. Grübler, and N. Nakicenovic (2007), Scenarios of long-term socio-economic and environmental development under climate stabilization, *Technol. Forecast. Social Change*, *74*(7), 887–935.
Roerink, G., M. Menenti, and W. Verhoef (2000), Reconstructing cloud-free NDVI composites using Fourier analysis of time series, *Int. J. Remote Sens.*, *21*(9), 1911–1917.
Savitzky, A., and M. J. Golay (1964), Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.*, *36*(8), 1627–1639.
Shen, H., and L. Zhang (2009), A MAP-based algorithm for destriping and inpainting of remotely sensed images, *IEEE Trans. Geosci. Remote Sens.*, *47*(5), 1492–1502.
Shen, H., X. Li, Q. Cheng, C. Zeng, G. Yang, H. Li, and L. Zhang (2015), Missing information reconstruction of remote sensing data: A technical review, *Geosci. Remote Sens. Mag. IEEE*, *3*(3), 61–85.
Shen, H., X. Meng, and L. Zhang (2016), An integrated framework for the spatio-temporal-spectral fusion of remote sensing images, *IEEE Trans. Geosci. Remote Sens.*, *54*(12), doi:10.1109/TGRS.2016.2596290.
Sinnott, R. (1984), Virtues of the haversine, *Sky Telesc.*, *68*(2), 159.
Siravenha, A. C., D. Sousa, A. Bispo, and E. Pelaes (2011), Evaluating inpainting methods to the satellite images clouds and shadows removing, in *Signal Processing, Image Processing and Pattern Recognition*, pp. 56–65, Springer, Berlin.

Staehelin, J., N. Harris, C. Appenzeller, and J. Eberhard (2001), Ozone trends: A review, *Rev. Geophys.*, *39*(2), 231–290, doi:10.1029/1999RG000059.

Stein, M. L. (2005), Space–time covariance functions, *J. Am. Stat. Assoc.*, *100*(469), 310–321.

Stein, M. L. (2007), Spatial variation of total column ozone on a global scale, *Ann. Appl. Stat.*, *1*(1), 191–210.

Team, R. C. (2014), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012, edited.

Van der Meer, F. (2012), Remote-sensing image analysis and geostatistics, *Int. J. Remote Sens.*, *33*(18), 5644–5676.

Wackernagel, H. (2003), *Multivariate Geostatistics*, Springer, Berlin.

Wargan, K., S. Pawson, M. A. Olsen, J. C. Witte, A. R. Douglass, J. R. Ziemke, S. E. Strahan, and J. E. Nielsen (2015), The global structure of upper troposphere-lower stratosphere ozone in GEOS-5: A multiyear assimilation of EOS Aura data, *J. Geophys. Res. Atmos.*, *120*, 1208–1218, doi:10.1002/2014JD022493.

Wellemeyer, C. G., P. Bhartia, R. McPeters, S. Taylor, and C. Ahn (2004), A new release of data from the Total Ozone Mapping Spectrometer (TOMS), *SPARC Newslett.*, *22*, 37–38.

White, M. C., R. A. Etzel, W. D. Wilcox, and C. Lloyd (1994), Exacerbations of childhood asthma and ozone pollution in Atlanta, *Environ. Res.*, *65*(1), 56–68.

Yan, H., L. Chen, J. Tao, L. Su, J. Huang, D. Han, and C. Yu (2012), Corrections for OMI SO 2 BRD retrievals influenced by row anomalies, *Atmos. Meas. Tech.*, *5*(11), 2635–2646.

Yang, G., H. Shen, L. Zhang, Z. He, and X. Li (2015), A moving weighted harmonic analysis method for reconstructing high-quality SPOT VEGETATION NDVI time-series data, *Geosci. Remote Sens. IEEE Trans.*, *53*(11), 6008–6021.

Zeng, C., H. Shen, and L. Zhang (2013), Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method, *Remote Sens. Environ.*, *131*, 182–194.

Zhang, C., W. Li, and D. Travis (2007), Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach, *Int. J. Remote Sens.*, *28*(22), 5103–5122.

Zimmerman, D. L. (1993), Another look at anisotropy in geostatistics, *Math. Geol.*, *25*(4), 453–470.