A high-accuracy method for filling voids on remotely sensed XCO₂ surfaces and its verification

Tian-Xiang Yue^{a, c, *}, Ming-Wei Zhao^{a, b, *}, Xing-Ying Zhang^b

^a State Key Laboratory of Resources and Environment Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, PR China

^b National Satellite Meteorological Center, China Meteorological Administration, Beijing 100081, PR China

^c Beijing Normal University, Zhuhai, Guangdong 519087, PR China

ARTICLE INFO

Article history:

Received 30 April 2014

Received in revised form

2 August 2014

Accepted 23 August 2014

Available online 3 September 2014

Keywords:

HASM

GOSAT

SCIAMACHY

XCO₂ surface

Void filling

Accuracy

ABSTRACT

The method for high accuracy surface modeling (HASM), inverse distance weighting (IDW) and ordinary Kriging (OK) are used to fill voids on XCO₂ surfaces of GOSAT and SCIAMACHY. Inner voids and boundary voids are artificially made by cleaning out the downloaded data, where there are no voids, in different latitude belts of the northern hemisphere and the southern hemisphere at random. The cleaned-out data in the artificial voids are selected as the 'true' values for verification. The results demonstrated that HASM always has the highest accuracy compared with the classical methods of IDW and OK, whether voids are inner ones or boundary ones, data sets are from GOSAT or from SCIAMACHY, and the void areas are larger or smaller. HASM is an alternative approach to filling voids on XCO₂ surfaces from the satellites.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

More than 100 countries have adopted a global surface warming limit of 2 °C or below as a guiding principle for mitigation efforts to reduce climate change risks, impacts and damages (Meinshausen et al., 2009). The rising atmospheric CO₂ concentration is believed to be the primary cause of global climate change (West and Marland, 2002; Buchwitz et al., 2006). To keep the increase of global mean temperature on a 2 °C level and not to risk extensive negative impacts of climate change, CO₂-concentration in the Earth's atmosphere should be stabilized at least on a 400–450 ppm level to reach the 2 °C with a probability of between 50% and 75% (Moss et al., 2008). Unfortunately, the average CO₂-concentration was 385 ppm in 2007 and CO₂-emissions have still been growing; on 9 May 2013, for the first time since measurement began in 1958, the daily mean concentration of CO₂ in the atmosphere of Mauna Loa, and Hawaii surpassed 400 ppm (Oberheitmann, 2010, 2013).

In the last 100 years, the European average temperature increased by 0.95 °C and it is projected to further increase 2.0 °C–6.3 °C by the year 2100 (Mimikou and Baltas, 2013). During the period from 1961 to 2010, mean annual temperature (MAT) of China increased by 1.44 °C (Yue et al., 2013b); from 2010 to 2099, MAT was estimated to be 2 °C–4.62 °C warmer (Yue, 2011). However, the projected results of current climate models are questionable on a regional level (Raisanen, 2007; Makarieva et al., 2014).

To accurately project climate change, various investigations have been conducted to understand sources and sinks of CO₂. For instances, it was demonstrated that the yearly average carbon stock change was a net increase of 1.71 Mt C/year in Thailand's total rubber planting area, while the yearly average greenhouse gas emission from cultivation of rubber tree was 0.32 Mt C/year (Peters et al., 2013). A study indicated that total carbon emissions in coastal Jiangsu amounted to 822.17×10^4 t in 1985 and increased to 2931.52×10^4 t in 2010, an increasing of 2.57 times from 1985 to 2010 (Chuai et al., 2015). A finding indicated that the conservation tillage practice had a greater contribution to the atmospheric CO₂ reduction (Chen et al., 2015). An input–output based methodology was proposed to identify key CO₂ emission sectors and the optimized production structure with respect to emission reduction targets in an economy (Chang, 2015). Effects of post-earthquake reconstruction on carbon cycle were investigated in Guangyuan, a

* Corresponding authors. State Key Laboratory of Resources and Environment Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, PR China.

E-mail addresses: yue@lreis.ac.cn (T.-X. Yue), zhaomw@lreis.ac.cn (M.-W. Zhao).

mountainous city in China, by means of an integrated accounting framework for urban carbon cycle (Hao et al., 2015).

However, insufficient knowledge of CO₂ leads to large uncertainties in future climate predictions because the observation of CO₂ is spatially and temporally limited on the globe (Yoshida et al., 2011). Satellite measurement is one of the most effective approaches to monitoring the global distributions of greenhouse gases at high spatiotemporal resolution and is expected to improve the accuracy of source and sink estimates of these gases (Rayner and O'Brien, 2001). Two satellites have first been designed specifically to measure the column-averaged dry air mole fraction of CO₂ (XCO₂). They are Japan's Greenhouse Gases Observing Satellite (GOSAT) and NASA's Orbiting Carbon Observatory (OCO). Both missions try to identify cloud-free scenes for their retrievals, since clouds can cause large errors in the retrieved CO₂ concentrations (Baker et al., 2010). GOSAT was successfully launched in January 2009. Although Orbiting Carbon Observatory 1 (OCO-1) was lost because of its launch failure in February 2009, Orbiting Carbon Observatory 2 (OCO-2) was successfully launched in July 2014, with the instrument that is almost an exact copy of the OCO-1 instrument with minor changes.

GOSAT carried a thermal and near infrared sensor for carbon observation-Fourier transform spectrometer (TANSO-FTS) for collecting information about the concentrations of CO₂ in the upper-troposphere (Chevallier et al., 2009). Column-averaged CO₂ was retrieved from spectra in the shortwave infrared (SWIR) band. The GOSAT observing strategy is to observe CO₂ column-averaged concentrations under cloud-free conditions, with little disturbance from aerosols. Bias and random error were simulated under assumption that the bias depended on the aerosol optical thickness (AOT) and surface albedo. The number of successfully retrieved GOSAT data was corrected under assuming a global mean clear-sky probability of 11% (Kadyrov et al., 2009).

Clouds and aerosols are major sources of disturbance in greenhouse gases observations from space because they strongly modify the equivalent optical path length (Mao and Kawa, 2004). 3.2% of measurement scenes are suitable for retrieval analysis for July 2009 and 2.7% for January 2010. The monthly mean fraction of the measurement scenes suitable for retrieval analysis is about 3% (Yoshida et al., 2011). In other words, no valid observations were available for GOSAT retrieval analysis on 97% of the earth surface.

The scanning imaging absorption spectrometer for atmospheric chartography (SCIAMACHY) spectrometer on board of the European Space Agency ENVISAT-1 was successfully launched in March 2002. SCIAMACHY first showed the global distribution of the CO₂ column abundance from space, although it was not specialized for CO₂ observations. SCIAMACHY is a passive hyper-spectral spectrometer designed to investigate tropospheric and stratospheric composition and processes (Bovensmann et al., 1999). The near-infrared nadir spectra of reflected solar radiation measured by SCIAMACHY contain information on the vertical columns of these gases, which we retrieve using the scientific algorithm, weighting function modified differential optical absorption spectroscopy (WFM-DOAS) approach. The horizontal resolution of the nadir measurements depends on orbital position and spectral interval but is typically 60 km for CO₂.

Calculations for a few typical aerosol classes, such as mineral dust, soot, sea salt, and sulfate aerosol, and a broad range of atmospheric conditions show that the impact of aerosols on SCIAMACHY retrieved CO₂ reaches significant levels. The largest errors are found over the deserts, owing to the high surface albedo in combination with relatively large aerosol loads. For instance, 33% of the data have to be ignored over deserts and the remaining data are used to obtain improved flux estimates. In other words, the voids for SCIAMACHY account for 33% of the total data set under cloud free conditions over deserts. The aerosol induced errors in the

measured CO₂ total column are too large to allow meaningful source and sink estimates (Houweling et al., 2005).

Any areas of missing data that exist in a CO₂ surface are defined as voids. Void filling is the process of filling all no-data areas using the best-performing algorithm available (Reuter et al., 2007). In this paper, the method for high accuracy surface modeling (HASM), inverse distance weighting (IDW) and Ordinary Kriging are comparatively used to fill voids on XCO₂ surfaces of GOSAT and SCIAMACHY in order to identify the best-performing method.

2. Methods

2.1. HASM

In terms of the fundamental theorem of surfaces, a surface is uniquely defined by the first fundamental coefficients and the second fundamental coefficients (Henderson, 1998). The first fundamental coefficients are used to express the intrinsic geometric properties that do not depend on the shape of the surface, but only on measurements that we can carry out while on the surface itself (Toponogov, 2006). The second fundamental coefficients reflect the local warping of the surface, namely its deviation from a tangent plane at the point under consideration, and observed from outside the surface (Liseikin, 2004).

If the surface can be expressed as $z = (x, y, f(x, y))$, then the first fundamental coefficients can be formulated as.

$$\begin{cases} E = 1 + f_x^2 \\ F = f_x f_y \\ G = 1 + f_y^2 \end{cases} \quad (1)$$

The second fundamental coefficients be formulated as.

$$\begin{cases} L = \frac{f_{xx}}{\sqrt{1 + f_x^2 + f_y^2}} \\ M = \frac{f_{xy}}{\sqrt{1 + f_x^2 + f_y^2}} \\ N = \frac{f_{yy}}{\sqrt{1 + f_x^2 + f_y^2}} \end{cases} \quad (2)$$

The first fundamental coefficients and the second fundamental coefficients must satisfy the following Gauss equation set.

$$\begin{cases} f_{xx} = \Gamma_{11}^1 f_x + \Gamma_{11}^2 f_y + \frac{L}{\sqrt{E + G - 1}} \\ f_{xy} = \Gamma_{12}^1 f_x + \Gamma_{12}^2 f_y + \frac{M}{\sqrt{E + G - 1}} \\ f_{yy} = \Gamma_{22}^1 f_x + \Gamma_{22}^2 f_y + \frac{N}{\sqrt{E + G - 1}} \end{cases} \quad (3)$$

where $\Gamma_{11}^1 = \frac{1}{2}(GE_x - 2FF_x)/(EG - F^2)$, $\Gamma_{12}^1 = \frac{1}{2}(GE_y - FG_x)/(EG - F^2)$, $\Gamma_{12}^2 = \frac{1}{2}(2GF_y - GG_x - FG_y)/(EG - F^2)$, $\Gamma_{11}^2 = \frac{1}{2}(2EF_x - EE_y - FE_x)/(EG - F^2)$, $\Gamma_{12}^2 = \frac{1}{2}(EG_x - FE_y)/(EG - F^2)$ and $\Gamma_{22}^2 = \frac{1}{2}(EG_y - 2FF_y + FG_x)/(EG - F^2)$ are the second kind of the Christoffel symbols.

If $\{(x_i, y_j)\}$ is an orthogonal division of a computational domain and h represents simulation step length, the central point of lattice (x_i, y_j) could be expressed as $(0.5h + (i-1)h, 0.5h + (j-1)h)$, in which $i = 0, 1, 2, \dots, I, I+1$ and $j = 0, 1, 2, \dots, J, J+1$. If $f_{ij}^{(n)}$ ($n \geq 0$) represents the iterants of $f(x, y)$ at (x_i, y_j) in the n th iterative step, in which $\{f_{ij}^{(0)}\}$ are interpolations based on sampling values $\{f_{ij}\}$, in terms of

numerical mathematics (Quarteroni et al., 2000), the iterative formulation of the HASM master equation set can be expressed as (Yue et al., 2013b; Zhao and Yue, 2014).

where $E_{ij}^{(n)}$, $F_{ij}^{(n)}$ and $G_{ij}^{(n)}$ are the iterants of the first fundamental

$(f_{1,1}^{(n+1)}, \dots, f_{1,J}^{(n+1)}, \dots, f_{I,1}^{(n+1)}, \dots, f_{I,J}^{(n+1)})^T = (z_1^{(n+1)}, \dots, z_J^{(n+1)}, \dots, z_{(I-1)J+1}^{(n+1)}, \dots, z_{IJ}^{(n+1)})^T$, $f_{ij}^{(n)}$ is the value of the n th iteration of $f(x,y)$ at grid cell (x_i, y_j) ; $z_{(i-1)J+j}^{(n+1)} = f_{ij}^{(n+1)}$ for $1 \leq i \leq I, 1 \leq j \leq J$.

$$\frac{-f_{i+2,j}^{(n+1)} + 16f_{i+1,j}^{(n+1)} - 30f_{ij}^{(n+1)} + 16f_{i-1,j}^{(n+1)} - f_{i-2,j}^{(n+1)}}{12h^2} = (I_{11}^{(n)})_{ij} \frac{f_{i+1,j}^{(n)} - f_{i-1,j}^{(n)}}{2h} + (I_{11}^{(n)})_{ij} \frac{f_{i,j+1}^{(n)} - f_{i,j-1}^{(n)}}{2h} + \frac{L_{ij}^{(n)}}{\sqrt{E_{ij}^{(n)} + G_{ij}^{(n)} - 1}} \quad (4)$$

$$\frac{-f_{i,j+2}^{(n+1)} + 16f_{i,j+1}^{(n+1)} - 30f_{ij}^{(n+1)} + 16f_{i,j-1}^{(n+1)} - f_{i,j-2}^{(n+1)}}{12h^2} = (I_{22}^{(n)})_{ij} \frac{f_{i+1,j}^{(n)} - f_{i-1,j}^{(n)}}{2h} + (I_{22}^{(n)})_{ij} \frac{f_{i,j+1}^{(n)} - f_{i,j-1}^{(n)}}{2h} + \frac{N_{ij}^{(n)}}{\sqrt{E_{ij}^{(n)} + G_{ij}^{(n)} - 1}} \quad (5)$$

$$\frac{f_{i+1,j+1}^{(n+1)} - f_{i+1,j}^{(n+1)} - f_{i,j+1}^{(n+1)} + 2f_{ij}^{(n+1)} - f_{i-1,j}^{(n+1)} - f_{i,j-1}^{(n+1)} + f_{i-1,j-1}^{(n+1)}}{2h^2} = (I_{12}^{(n)})_{ij} \frac{f_{i+1,j}^{(n)} - f_{i-1,j}^{(n)}}{2h} + (I_{12}^{(n)})_{ij} \frac{f_{i,j+1}^{(n)} - f_{i,j-1}^{(n)}}{2h} + \frac{M_{ij}^{(n)}}{\sqrt{E_{ij}^{(n)} + G_{ij}^{(n)} - 1}} \quad (6)$$

coefficients at the n th iterative step; $L_{ij}^{(n)}$, $M_{ij}^{(n)}$ and $N_{ij}^{(n)}$ represent the iterants of the second fundamental coefficients at the n th iterative step; $(I_{11}^{(n)})_{ij}$, $(I_{11}^{(n)})_{ij}$, $(I_{22}^{(n)})_{ij}$ and $(I_{22}^{(n)})_{ij}$ the iterants of the Christoffel symbols of the second kind at the n th iterative step, which depend only upon the first fundamental coefficients and their derivatives.

The matrix formulation of HASM master equations can be respectively expressed as.

$$\mathbf{A} \cdot \mathbf{z}^{(n+1)} = \mathbf{d}^{(n)} \quad (7)$$

$$\mathbf{B} \cdot \mathbf{z}^{(n+1)} = \mathbf{q}^{(n)} \quad (8)$$

$$\mathbf{C} \cdot \mathbf{z}^{(n+1)} = \mathbf{p}^{(n)} \quad (9)$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ represent coefficient matrixes of the first equation, the second equation and the third equation; $\mathbf{d}^{(n)}$, $\mathbf{q}^{(n)}$ and $\mathbf{p}^{(n)}$ are right-hand side vectors of the three equations respectively; $\mathbf{z}^{(n+1)} =$

If $\bar{f}_{ij}$ is value of $z = f(x,y)$ at the p th sampled point (x_i, y_j) , $s_{p,(i-1) \times J+j} = 1$, $k_p = \bar{f}_{ij}$. There is only one non-zero element, 1, in every row of the coefficient matrix, $\mathbf{S}$, making it a sparse matrix. The solution procedure of HASM, taking the sampled points as its constraints, can be transformed into solving the following linear equation set in terms of least squares principle.

$$\begin{bmatrix} \mathbf{A}^T & \mathbf{B}^T & \mathbf{C}^T & \lambda \cdot \mathbf{S}^T \end{bmatrix} \mathbf{z}^{(n+1)} = \begin{bmatrix} \mathbf{A}^T & \mathbf{B}^T & \mathbf{C}^T & \lambda \cdot \mathbf{S}^T \end{bmatrix} \begin{bmatrix} \mathbf{d}^{(n)} \\ \mathbf{q}^{(n)} \\ \mathbf{p}^{(n)} \\ \lambda \cdot \mathbf{k} \end{bmatrix} \quad (10)$$

The parameter λ is the weight of the sampling points and determines the contribution of the sampling points to the simulated surface. λ could be a real number, which means all sampling points have the same weight, or a sector, which means every sampling point has its own weight. An area affected by a sampling point in a

Fig. 1. Location of the artificial voids.

Table 1
Mean absolute Errors of void filling on XCO₂ surfaces of GOSAT (unit: ppmv).

Time	Errors of inner void filling			Errors of boundary void filling		
	IDW	OK	HASM	IDW	OK	HASM
January in 2010	0.152	0.126	0.121	0.487	0.484	0.476
February in 2010	0.288	0.275	0.255	0.515	0.496	0.488
March in 2010	0.321	0.351	0.26	0.313	0.362	0.149
April in 2010	0.329	0.36	0.215	0.362	0.411	0.343
May in 2010	0.395	0.28	0.262	0.829	0.567	0.432
June in 2010	0.445	0.309	0.301	0.752	0.64	0.508
July in 2009	0.271	0.228	0.156	0.517	0.495	0.333
August in 2009	0.351	0.301	0.22	0.746	0.531	0.485
September in 2009	0.229	0.212	0.116	0.397	0.379	0.297
October in 2009	0.214	0.219	0.209	0.399	0.436	0.361
November in 2009	0.099	0.01	0.099	0.461	0.488	0.457
December in 2009	0.139	0.083	0.053	0.513	0.491	0.336
On an average	0.269	0.230	0.189	0.524	0.482	0.389

complex region is smaller than in a flat region. Therefore, a smaller value of λ is selected in a complex region and a bigger value of λ is selected in a flat region.

Table 2
Mean relative Errors of void filling on XCO₂ surfaces of GOSAT (unit: %).

Time	Errors of inner void filling			Errors of boundary void filling		
	IDW	OK	HASM	IDW	OK	HASM
January in 2010	0.04	0.033	0.031	0.128	0.129	0.125
February in 2010	0.076	0.072	0.067	0.136	0.131	0.129
March in 2010	0.085	0.092	0.068	0.082	0.094	0.039
April in 2010	0.086	0.094	0.056	0.095	0.107	0.092
May in 2010	0.104	0.074	0.068	0.217	0.149	0.116
June in 2010	0.117	0.082	0.08	0.199	0.169	0.113
July in 2009	0.072	0.061	0.041	0.136	0.13	0.088
August in 2009	0.094	0.08	0.058	0.198	0.141	0.129
September in 2009	0.061	0.056	0.031	0.105	0.1	0.079
October in 2009	0.057	0.058	0.055	0.106	0.115	0.096
November in 2009	0.026	0.026	0.026	0.122	0.128	0.121
December in 2009	0.037	0.022	0.014	0.135	0.129	0.089
On an average	0.071	0.063	0.050	0.138	0.127	0.101

Table 3
Mean absolute errors of void filling on XCO₂ surfaces of SCIAMACHY (unit: ppmv).

Time	Errors of inner void filling			Errors of boundary void filling		
	IDW	OK	HASM	IDW	OK	HASM
January in 2009	1.152	1.388	1.135	3.667	3.824	3.62
February in 2009	1.858	1.928	1.775	3.889	3.906	2.19
March in 2009	1.473	1.591	1.358	3.66	3.794	3.65
April in 2009	1.669	1.708	1.593	2.643	2.975	2.4
May in 2009	1.516	1.467	1.472	4.172	4.122	4.09
June in 2009	1.573	1.495	1.433	3.57	3.644	2.82
July in 2009	1.979	2.117	1.832	5.159	5.638	4.7
August in 2009	1.651	1.768	1.583	3.816	3.824	2.85
September in 2009	1.267	1.232	1.183	3.481	3.604	3.37
October in 2009	1.376	1.336	1.333	3.032	3.001	2.92
November in 2009	1.706	1.829	1.544	3.645	3.732	4.07
December in 2009	1.401	1.56	1.356	3.363	3.565	2.39
On an average	1.552	1.618	1.466	3.675	3.802	3.256

$$\text{Let } \mathbf{W} = [\mathbf{A}^T \quad \mathbf{B}^T \quad \mathbf{C}^T \quad \lambda \cdot \mathbf{S}^T] \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \lambda \cdot \mathbf{S} \end{bmatrix} \text{ and, } \mathbf{v}^{(n)} = [\mathbf{A}^T \quad \mathbf{B}^T \quad \mathbf{C}^T \quad \lambda \cdot \mathbf{S}^T] \begin{bmatrix} \mathbf{d}^{(n)} \\ \mathbf{q}^{(n)} \\ \mathbf{p}^{(n)} \\ \lambda \cdot \mathbf{k} \end{bmatrix}, \text{ then HASM has the following formulation.}$$

Table 4
Mean relative Errors of void filling on XCO₂ surfaces of SCIAMACHY (unit: %).

Time	Errors of inner void filling			Errors of boundary void filling		
	IDW	OK	HASM	IDW	OK	HASM
January in 2009	0.297	0.358	0.293	0.948	0.987	0.936
February in 2009	0.478	0.496	0.456	0.995	0.999	0.564
March in 2009	0.379	0.408	0.348	0.935	0.969	0.935
April in 2009	0.429	0.438	0.409	0.673	0.758	0.667
May in 2009	0.391	0.378	0.379	1.065	1.052	1.051
June in 2009	0.403	0.383	0.369	0.92	0.939	0.731
July in 2009	0.51	0.545	0.471	1.32	1.439	1.224
August in 2009	0.429	0.46	0.413	0.985	0.986	0.738
September in 2009	0.328	0.319	0.306	0.907	0.939	0.874
October in 2009	0.356	0.345	0.344	0.791	0.76	0.758
November in 2009	0.44	0.472	0.397	0.953	0.976	1.055
December in 2009	0.364	0.405	0.352	0.871	0.923	0.618
On an average	0.400	0.417	0.378	0.947	0.977	0.846

Fig. 2. The land surfaces of XCO₂ from GOSAT in August of 2009, a) the original surface with voids, b) the surface on which voids have been filled.

$$\mathbf{W} \cdot \mathbf{z}^{(n+1)} = \mathbf{v}^{(n)} \quad (11)$$

In terms of Gauss-Codazii equation set, iteration stopping criterion of HASM is formulated as.

with available data, and boundary voids that are surrounded by both grids with available data and grids without data. We design a verification procedure for these two types of voids (Fig. 1) as follows:

$$(\varphi_{1y} - \varphi_{2x} - \varphi_2 P - \varphi_1 Q)^2 + (\varphi_{2x} - \varphi_{1y} - \varphi_1 Q - \varphi_2 P)^2 (Q_x + P_y + \varphi_1 \varphi_2 - \varphi_1 \varphi_2)^2 < EI \quad (12)$$

where $\varphi_1 = L/\sqrt{E}$; $\varphi_2 = N/\sqrt{G}$; $P = \sqrt{E_y}/\sqrt{G}$; $Q = \sqrt{G_x}/\sqrt{E}$; $\varphi_1 = M/\sqrt{G}$; $\varphi_2 = M/\sqrt{E}$; EI is the iteration stopping criterion of HASM determined by an application requirement for accuracy.

HASM has been successfully applied to simulating climate change (Yue et al., 2013a, 2013b), constructing DEM (Yue et al., 2007, 2010a, 2010b; Yue and Wang, 2010) and interpolating soil properties (Yue, 2011; Shi et al., 2009, 2011).

3. Verification

The voids on XCO₂ surfaces of GOSAT and SCIAMACHY can be classified into two types: 1) inner voids that are embraced by grids

- 1) data sets of monthly mean XCO₂ from GOSAT on 2.5° latitude–longitude grid and from SCIAMACHY on 0.5° latitude–longitude grid were downloaded respectively from the websites of <http://www.gosat.nies.go.jp> and <http://www.iup.uni-bremen.de/sciamachy>;
- 2) SCIAMACHY data are up-scaled into a dataset at a spatial resolution of 2.5° × 2.5° using a bilinear re-sampling approach for its comparability with the GOSAT dataset;
- 3) according to the data availability, we selected the dataset of monthly mean XCO₂ from SCIAMACHY during the period from January to December of 2009 and the one from GOSAT during the period from July of 2009 to June of 2010 so that the two

Fig. 3. The land surfaces of XCO₂ from SCIAMACHY in August of 2009, a) the original surface with voids, b) the surface on which voids have been filled.

datasets have an overlap for 6 months to compare the performances of GOSAT data and SCIAMACHY data;

- 4) nine inner voids in size of $2.5^\circ \times 2.5^\circ$ are artificially made by cleaning out the downloaded data on land, where there are no voids, in the high and middle latitude belts of the northern hemisphere ($30^\circ - 70^\circ\text{N}$), the lower latitude belts of the northern hemisphere ($0^\circ - 30^\circ\text{N}$) and the southern hemisphere ($0^\circ - 30^\circ\text{S}$) at random respectively; the cleaned-out data in the nine artificially made voids are selected as the 'true' values for verification;
- 5) six boundary voids on land in different sizes, of which the maximum size is $22.5^\circ \times 5^\circ$ and the minimum size is $7.5^\circ \times 5^\circ$, are artificially made by cleaning out the downloaded data in the three latitude belts and the cleaned-out data are selected as the 'true' values for verification;
- 6) the voids on land are filled, taking available XCO₂ data over oceans as auxiliary data, by respectively using IDW, OK and HASM; errors of void filling results are comparatively analyzed in terms of mean absolute error (MAE) and mean relative error (MRE).

MAE and MRE are respectively formulated as.

$$MAE(t) = \frac{1}{n} \sum_i^n |o_i(t) - s_i(t)| \quad (13)$$

$$MRE(t) = 100\% \times MAE(t) / \left(\frac{1}{n} \sum_i^n o_i(t) \right) \quad (14)$$

where o_i represents the original value at time t on grid i ; s_i the simulated value on grid i .

For the inner voids of GOSAT data set (Tables 1 and 2), MAEs of the inner void filling from IDW, OK and HASM are 0.269 ppmv, 0.23 ppmv and 0.189 ppmv respectively on the average; their MREs are respectively 0.071%, 0.063% and 0.05%. The maximum MAE of IDW is 0.445 ppmv, appearing in June of 2010; the one of OK is 0.351 ppmv, happening in March of 2010; and the one of HASM is 0.301 ppmv, in June of 2010. The minimum MAE of IDW is 0.099 ppmv in November of 2009; the one of OK is 0.01 ppmv in November of 2009; and the one of HASM is 0.053 ppm in December of 2009.

For the boundary voids of GOSAT data set (Tables 1 and 2), MAEs of IDW, OK and HASM are 0.524 ppmv, 0.482 ppmv and 0.389 ppmv

Table 5Effects of void filling on mean values and variances of XCO₂ surfaces of GOSAT by using HASM.

Time	Before void filling		After void filling	
	Mean value (ppmv)	Standard deviation	Mean value (ppmv)	Standard deviation
January in 2010	378.745	1.63	379.14	1.659
February in 2010	379.102	1.886	379.514	1.815
March in 2010	380.384	1.91	380.597	1.622
April in 2010	381.38	2.004	381.532	1.984
May in 2010	381.708	1.956	381.349	1.933
June in 2010	380.473	2.277	380.117	2.258
July in 2009	377.649	2.623	377.22	2.629
August in 2009	375.403	3.305	374.562	3.303
September in 2009	376.096	2.093	375.566	2.299
October in 2009	376.949	1.839	375.76	2.428
November in 2009	377.905	1.281	377.256	1.913
December in 2009	378.179	1.422	377.948	1.946
On an average	378.664	2.019	378.380	2.149

respectively on average; their MREs are respectively 0.138%, 0.127% and 0.101%. The maximum MAE of IDW is 0.829 ppmv appearing in May of 2010; the minimum one is 0.313 ppmv in March of 2010. The maximum MAEs of OK and HASM are respectively 0.64 ppmv and 0.508 ppmv, both happening in June of 2010. The minimum MAEs of IDW, OK and HASM are respectively 0.313 ppmv, 0.362 ppmv and 0.149 ppmv, all happening in March of 2010.

For inner voids on SCIAMACHY surfaces of XCO₂ (Tables 3 and 4), MAEs of IDW, OK and HASM are respectively 1.552 ppmv, 1.618 ppmv and 1.466 ppmv on the average; their MREs are respectively 0.4%, 0.417% and 0.378%. Their maximum MAEs are respectively 1.858 ppmv, 1.928 ppmv and 1.775 ppmv, all appearing in February. The minimum MAEs of IDW and HASM are respectively 1.152 ppmv and 1.135 ppmv, both happening in January; the one of OK is 1.232 ppmv in September.

According to the tests of filling boundary voids on SCIAMACHY surfaces of XCO₂ (Tables 3 and 4), the average MAEs of IDW, OK and HASM are respectively 3.675 ppmv, 3.802 ppmv and 3.256 ppmv in 2009; their MREs are respectively 0.947%, 0.977% and 0.846%. The biggest MAEs of IDW, OK and HASM are respectively 5.159 ppmv, 5.638 ppmv and 4.7 ppmv, appearing in July. The smallest MAEs of IDW and OK are respectively 2.643 ppmv and 2.975 ppmv, both appearing in April; the one of HASM is 2.19 ppmv in February.

The verification results indicated that HASM always has the highest accuracy compared with the classical methods of IDW and OK, whether voids are inner ones or boundary ones, data sets are from GOSAT or from SCIAMACHY, and the void areas are larger or smaller. For all the methods of IDW, OK and HASM, the void filling errors on XCO₂ surfaces of SCIAMACHY are much bigger than the ones on the surfaces of GOSAT, especially for the boundary voids, which may be mainly caused by the greater spatial variability of SCIAMACHY surfaces. In terms of HASM, the average MAE and MRE of GOSAT are 0.269 ppmv and 0.07% respectively under consideration of both inner voids and boundary voids, while the ones of SCIAMACHY are 2.567 ppmv and 0.663% in the 12 months.

4. Results

According to the original data, taking the surfaces in August of 2009 as an example (Figs. 2 and 3a), spatial distribution of voids on land surfaces of XCO₂ from GOSAT are different comparing with the ones from SCIAMACHY. After GOSAT data, the voids mainly distribute in east-middle Africa, northwestern South America, southern South America, eastern North America, northwestern North America, southern Asia and eastern Asia, but after SCIAMACHY, voids mainly scattered over northwestern South America, southern South America, north North America, western Tibet plateau, eastern Africa, northern Europe and Northern Asia.

The surfaces of XCO₂, in which voids have been completely filled, show that spatial variability of SCIAMACHY surface is much higher than GOSAT surface (Figs. 2b and 3b). Standard deviation of GOSAT surfaces was 2.094 ppmv while the one of SCIAMACHY was 4.825 ppmv during the period from July to December in 2009 (Tables 5 and 6). The results from subtracting GOSAT surface from SCIAMACHY surface of XCO₂ indicates that almost all grids on SCIAMACHY surfaces have bigger values of XCO₂ than the ones on GOSAT surfaces (Tables 5 and 6), except for a few grids on the boundaries of land (Fig. 4). The maximum difference is 28.255 ppmv. The grids, on which the difference is greater than 15 ppmv, distribute in middle Africa, India, western Tibet plateau, northern Russia and northern Alaska.

In terms of the GOSAT surface (Fig. 2b), XCO₂ on grids in southern hemisphere is generally much bigger than the ones on grids in northern hemisphere. The reason might be that southern hemisphere is in the season of winter and northern hemisphere in summer. However, this phenomenon could not be found from the SCIAMACHY surface. The difference of XCO₂ between SCIAMACHY and GOSAT is 11.463 ppmv on an average in Eurasian continent, 11.274 ppmv in African continent, 9.106 ppmv in the North American continent, 7.281 ppmv in mainland South America, and 6.861 ppmv in Oceania mainland (Table 7).

Table 6Effects of void filling on mean values and variances of XCO₂ surfaces of SCIAMACHY by using HASM.

Time	Before void filling		After void filling	
	Mean value (ppmv)	Standard deviation	Mean value (ppmv)	Standard deviation
January in 2009	387.465	5.15	387.194	6.546
February in 2009	388.728	6.072	390.249	7.031
March in 2009	389.823	6.632	390.798	6.681
April in 2009	389.883	6.486	390.391	5.988
May in 2009	389.173	6.017	388.315	7.898
June in 2009	386.352	5.481	386.761	4.754
July in 2009	384.754	6.004	385.606	5.9
August in 2009	384.887	4.956	385.158	4.115
September in 2009	385.334	4.593	385.361	4.051
October in 2009	386.639	4.451	386.097	3.594
November in 2009	387.151	4.434	386.949	6.443
December in 2009	387.519	4.509	386.052	5.702
On an average	387.309	5.399	387.411	5.725

Fig. 4. The difference map of XCO₂ surfaces from GOSAT and SCIAMACHY.

5. Conclusions and discussion

The carbon satellites are providing valuable data to improve global models of the carbon cycle and to make more accurate predictions of global climate change. Measurements by the carbon satellites allow scientists to monitor the geographic distribution of CO₂ sources and to map the natural and human emissions. It becomes possible to simulate interaction mechanisms between cleaner production, the carbon cycle and climate change.

A simulation system for XCO₂ surfaces is being developed on the basis of HASM (simply termed as HASM system), which takes satellite remote sensing data as its driving field and ground observation data as its optimum control constraints. The XCO₂ surfaces from satellite remote sensing are difficult to be directly used as the driving fields because of a large number of voids caused by clouds, aerosols and the high surface albedo. It is necessary to conduct void filling for constructing the HASM system.

The verification results showed that HASM always has the highest accuracy compared with the classical methods of IDW and OK, whether voids are inner ones or boundary ones, data sets are from GOSAT or from SCIAMACHY, and the void areas are larger or smaller.

In terms of HASM, the MAEs on monthly XCO₂ surfaces of GOSAT are less than 0.6 ppmv for both inner voids and boundary voids. Although the void filling errors on the XCO₂ surfaces from SCIAMACHY are much bigger than the ones on the surfaces from GOSAT, the MREs on the SCIAMACHY surfaces are less than 0.9% on the average. HASM is an alternative approach to void filling for GOSAT, SCIAMACHY, Orbiting Carbon Observatory –2 (OCO-2) as well as TanSat, to be launched in 2015.

It is essential to validate the XCO₂ surfaces for understanding qualities of data sets from different satellites. The Total Carbon

Column Observing Network (TCCON) has provided an essential validation resource for assessing the accuracy of XCO₂ surfaces from the satellites. However, the TCCON needs to be improved because the ground-based Fourier Transform Spectrometers are too sparsely scattered over the Earth's surface to give sufficiently believable conclusion, especially when TCCON would be also worked as the optimum control constraints for the simulation system for XCO₂ surfaces.

Acknowledgments

This work is supported by National High-tech R&D Program of the Ministry of Science and Technology of the People's Republic of China (2013AA122003), by the Key Program of National Natural Science of China (91325204), and by National Fundamental R&D Program of the Ministry of Science and Technology of the People's Republic of China (2013FY111600-4). We would like to acknowledge the two anonymous reviewers.

References

- Baker, D.F., Bosch, H., Doney, S.C., O'Brien, D., Schimel, D.S., 2010. Carbon source/sink information provided by measurements from the Orbiting Carbon Observatory. *Atmos. Chem. Phys.* 10, 4145–4165.
- Bovensmann, H., Burrows, J.P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V., Chance, K.V., Goede, A., 1999. SCIAMACHY –mission objectives and measurement modes. *J. Atmos. Sci.* 56, 127–150.
- Buchwitz, M., de Beek, R., Noel, S., Burrows, J.P., Bovensmann, H., Schneising, O., Khlystova, I., Bruns, M., Bremer, H., Bergamaschi, P., Koerner, S., Heimann, M., 2006. Atmospheric carbon gases retrieved from SCIAMACHY by WFM-DOAS: version 0.5 CO and CH₄ and impact of calibration improvements on CO₂ retrieval. *Atmos. Chem. Phys.* 6, 2727–2751.
- Chang, N., 2015. Changing industrial structure to reduce carbon dioxide emissions: a Chinese application. *J. Clean. Prod.* 103, 40–48.
- Chen, Z.D., Dikgwathle, S.B., Xue, J.F., Zhang, H.L., Chen, F., Xiao, X.P., 2015. Tillage impacts on net carbon flux in paddy soil of the Southern China. *J. Clean. Prod.* 103, 70–76.
- Chevallier, F., Maksyutov, S., Bousquet, P., Breon, F.-M., Saito, R., Yoshida, Y., Yokota, T., 2009. On the accuracy of the CO₂ surface fluxes to be estimated from the GOSAT observations. *Geophys. Res. Lett.* 36, L19807. <http://dx.doi.org/10.1029/2009GL040108>.
- Chuai, X.W., Huang, X.J., Wang, W.J., Zhao, R.Q., Zhang, M., Wu, C.Y., 2015. Land use, total carbon emissions change and low carbon land management in Coastal Jiangsu, China. *J. Clean. Prod.* 103, 77–86.
- Hao, Y., Su, M.R., Zhang, L.X., Cai, Y.P., Yang, Z.F., 2015. Integrated accounting of urban carbon cycle in Guangyuan, a mountainous city of China: the impacts of earthquake and reconstruction. *J. Clean. Prod.* 103, 231–240.
- Henderson, D.W., 1998. *Differential Geometry*. Prentice-Hall, Inc., London.

Table 7

Difference of XCO₂ from GOSAT and SCIAMACHY on the average in the five continents (unit: ppmv).

	Eurasian continent	African continent	The north American continent	Mainland South America	Oceania mainland
GOSAT	373.289	375.972	374.114	377.729	378.2
SCIAMACHY	384.752	387.246	383.22	385.01	385.061
Difference	11.463	11.274	9.106	7.281	6.861

- Houweling, S., Hartmann, W., Aben, I., Schrijver, H., Skidmore, J., Roelofs, G.-J., Breon, F.M., 2005. Evidence of systematic errors in SCIAMACHY-observed CO₂ due to aerosols. *Atmos. Chem. Phys.* 5, 3003–3013.
- Kadygrov, N., Maksyutov, S., Eguchi, N., Aoki, T., Nakazawa, T., Yokota, T., Inoue, G., 2009. Role of simulated GOSAT total column CO₂ observations in surface CO₂ flux uncertainty reduction. *J. Geophys. Res.* 114, D21208. <http://dx.doi.org/10.1029/2008JD011597>.
- Liseikin, V.D., 2004. *A Computational Differential Geometry Approach to Grid Generation*. Springer, Berlin.
- Makarieva, A.M., Gorshkov, V.G., Sheil, D., Nobre, A.D., Bunyard, P., Li, B.L., 2014. Why does air passage over forest yield more rain? examining the coupling between rainfall, pressure, and atmospheric moisture content. *J. Hydrometeorol.* 15, 411–426.
- Mao, J., Kawa, S.R., 2004. Sensitivity studies for space-based measurement of atmospheric total column carbon dioxide by reflected sunlight. *Appl. Opt.* 43, 914–927.
- Meinshausen, M., Meinshausen, N., Hare, W., Raper, S., Frieler, K., Knutti, R., Frame, D., Allen, M., 2009. Greenhouse gas emission targets for limiting global warming to 2°C. *Nature* 458, 1158–1163.
- Mimikou, M.A., Baltas, E.A., 2013. Assessment of climate change impacts in Greece: a general overview. *Am. J. Clim. Change* 3, 46–56.
- Moss, R., Babiker, M., Brinkman, S., Calvo, E., Carter, T., Edmonds, J., Elgizouli, I., Emori, S., Erda, L., Hibbard, K., Jones, R., Kainuma, M., Kelleher, J., Lamarque, J.F., Manning, M., Matthews, B., Meehl, J., Meyer, L., Mitchell, J., Nakicenovic, N., O'Neill, B., Pichs, R., Riahi, K., Rose, S., Runci, P., Stouffer, R., van Vuuren, D., Weyant, J., Wilbanks, T., van Ypersele, J.P., Zurek, M., 2008. *Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies*. Technical Summary. Intergovernmental Panel on Climate Change, Geneva.
- Oberheitmann, A., 2010. A new Post-Kyoto climate Regime based on Per-Capita Cumulative emissions Rights—Rationale, architecture and quantitative assessment of the implication for the CO₂-emissions from China, India and the Annex-i countries by 2050. *Mitig. Adapt. Strat. Glob. Change* 15 (2), 137–168.
- Oberheitmann, A., 2013. Some remarks on the individual contribution to climate change. *Am. J. Clim. Change* 2, 198–202.
- Petsri, S., Chidthaisong, A., Pumijumnong, N., Wachrinrat, C., 2013. Greenhouse gas emissions and carbon stock changes in rubber tree plantations in Thailand from 1990 to 2004. *J. Clean. Prod.* 52, 61–70.
- Quarteroni, A., Sacco, R., Saleri, F., 2000. *Numerical Mathematics*. Springer, New York.
- Raisanen, J., 2007. How reliable are climate models? *Tellus* 59A, 2–29.
- Rayner, P.J., O'Brien, D.M., 2001. The utility of remotely sensed CO₂ concentration data in surface source inversions. *Geophys. Res. Lett.* 28, 175–178.
- Reuter, H.I., Nelson, A., Jarvis, A., 2007. An evaluation of void-filling interpolation methods for SRTM data. *Int. J. Geogr. Inf. Sci.* 21, 983–1008.
- Shi, W.J., Liu, J.Y., Song, Y.J., Du, Z.P., Chen, C.F., Yue, T.X., 2009. Surface modelling of soil pH. *Geoderma* 150 (1–2), 113–119.
- Shi, W.J., Liu, J.Y., Du, Z.P., Stein, A., Yue, T.X., 2011. Surface modelling of soil properties based on land use information. *Geoderma* 162 (3–4), 347–357.
- Toponogov, V.A., 2006. *Differential Geometry of Curves and Surfaces*. Birkhaeuser Boston, New York.
- West, T.O., Marland, G., 2002. A synthesis of carbon sequestration, carbon emissions, and net carbon flux in agriculture: comparing tillage practices in the United States. *Agric. Ecosyst. Environ.* 91, 217–232.
- Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., Yokota, T., 2011. Retrieval algorithm for CO₂ and CH₄ column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite. *Atmos. Meas. Tech.* 4, 717–734.
- Yue, T.X., Zhao, N., Yang, H., Song, Y.J., Du, Z.P., Fan, Z.M., Song, D.J., 2013a. The multi-grid method of high accuracy surface modelling and its validation. *Trans. GIS* 17 (6), 943–952.
- Yue, T.X., Zhao, N., Ramsey, D., Wang, C.L., Fan, Z.M., Chen, C.F., Lu, Y.M., Li, B.L., 2013b. Climate change trend in China, with improved accuracy. *Clim. Change* 120, 137–151.
- Yue, T.X., 2011. *Surface Modelling: High Accuracy and High Speed Methods*. CRC Press, New York.
- Yue, T.X., Song, D.J., Du, Z.P., Wang, W., 2010a. High-accuracy surface modelling and its application to DEM generation. *Int. J. Remote Sens.* 31 (8), 2205–2226.
- Yue, T.X., Chen, C.F., Li, B.L., 2010b. An adaptive method of high accuracy surface modeling and its application to simulating elevation surfaces. *Trans. GIS* 14 (5), 615–630.
- Yue, T.X., Wang, S.H., 2010. Adjustment computation of HASM: a high-accuracy and high-speed method. *Int. J. Geogr. Inf. Sci.* 24 (11), 1725–1743.
- Yue, T.X., Du, Z.P., Song, D.J., Gong, Y., 2007. A new method of surface modeling and its application to DEM construction. *Geomorphology* 91 (1–2), 161–172.
- Zhao, N., Yue, T.X., 2014. A modification of HASM for interpolating precipitation in China. *Theor. Appl. Climatol.* 116, 273–285.