Simulation and analysis of XCO₂ in North China based on high accuracy surface modeling

Yu Liu^{1,2} · Tianxiang Yue^{1,2} · Lili Zhang³ · Na Zhao^{1,2} · Miaomiao Zhao^{1,2} · Yi Liu^{1,2}

Received: 24 April 2018 / Accepted: 29 June 2018 / Published online: 22 July 2018
© The Author(s) 2018

Abstract

As an important cause of global warming, CO₂ concentrations and their changes have aroused worldwide concern. Establishing explicit understanding of the spatial and temporal distributions of CO₂ concentrations at regional scale is a crucial technical problem for climate change research. High accuracy surface modeling (HASM) is employed in this paper using the output of the CO₂ concentrations from weather research and forecasting-chemistry (WRF-CHEM) as the driving fields, and the greenhouse gases observing satellite (GOSAT) retrieval XCO₂ data as the accuracy control conditions to obtain high accuracy XCO₂ fields. WRF-CHEM is an atmospheric chemical transport model designed for regional studies of CO₂ concentrations. Verified by ground- and space-based observations, WRF-CHEM has a limited ability to simulate the conditions of CO₂ concentrations. After conducting HASM, we obtain a higher accuracy distribution of the CO₂ in North China than those calculated using the classical Kriging and inverse distance weighted (IDW) interpolation methods, which were often used in past studies. The cross-validation also shows that the averaging mean absolute error (MAE) of the results from HASM is 1.12 ppmv, and the averaging root mean square error (RMSE) is 1.41 ppmv, both of which are lower than those of the Kriging and IDW methods. This study also analyses the space-time distributions and variations of the XCO₂ from the HASM results. This analysis shows that in February and March, there was the high value zone in the southern region of study area relating to heating in the winter and the dense population. The XCO₂ concentration decreased by the end of the heating period and during the growing period of April and May, and only some relatively high value zones continued to exist.

Keywords HASM · WRF-CHEM · GOSAT XCO₂ · XCO₂ simulation

Introduction

Global warming has drawn worldwide attention. As an important greenhouse gas, atmospheric CO₂ concentrations can profoundly affect the trends of developments various climate

scenarios, thus affecting national security and sustainable economic developments. According to the newest WMO Greenhouse Gas Bulletin (2017), globally averaged concentration of atmospheric CO₂ rose to 403.3 ppm in 2016, compared with the 280 ppm of the preindustrial era. Therefore, to

Responsible editor: Gerhard Lammel

✉ Tianxiang Yue
yue@reis.ac.cn

Yu Liu
liuyu.15b@igsnr.ac.cn

Lili Zhang
zhangll@reis.ac.cn

Na Zhao
zhaon@reis.ac.cn

Miaomiao Zhao
zhaomm@reis.ac.cn

Yi Liu
liuyi@reis.ac.cn

¹ State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

control CO₂ emissions and reduce the effects of human activity on climate warming, understanding the spatial and temporal distributions of atmospheric CO₂ is crucial.

A large number of ground-based observatories have been built around the world that use methods such as bottle sampling, spectroscopy, and eddy covariance to obtain information about CO₂ concentrations and carbon fluxes. However, due to technical and financial constraints, the distribution of surface observation stations is still too sparse, and the region-scale observation capabilities are limited. In recent decades, with the development of remote sensing technology, satellite-based observations have made up for some of the deficiencies of ground-based observations. The CO₂ concentration data retrieved from the satellite observation spectral data provide additional support for scientists studying the sources and sinks of CO₂. SCIAMACHY is the first satellite-based detector to be sensitive to the boundary layer and acquired a large amount of observations over its 10-year orbit. Despite its lower accuracy, the data from the SCIAMACHY instrument still provides an opportunity to study the behavior of the terrestrial biosphere and climate change (Wang et al. 2011; Barkley et al. 2007). At present, the satellites in orbit that specifically make CO₂ observations include GOSAT from Japan, OCO-2 from the USA, and TanSat from China. These satellites' objectives are to provide global, long-term, continuous monitoring of XCO₂ concentrations, to improve the measurement accuracies of carbon sources and sinks as well as of the regional scale CO₂ concentrations, and to improve the understanding of their distribution characteristics and evolutions (Turner et al. 2015; Hakkarainen et al. 2016; Fischer et al. 2017). TanSat, the third carbon satellite in the world, has been active for more than a year, and the relevant research teams are working on spectral inversions of its data (Zhang et al. 2017a). The work of this paper is mainly to prepare regional scale research by TanSat data for the future, after TanSat L2 retrieval data released.

Another viable method to obtain the high resolution spatial and temporal distributions of CO₂ concentrations is via the atmospheric chemical transport model, which uses meteorological data to represent atmosphere activities and CO₂ fluxes to identify different emission scenarios. The chemical transport model is widely used to study CO₂ on global and regional scales and as a comparison with observations (Shim et al. 2013; Wang et al. 2014). Additionally, by combining atmospheric chemical transport model and CO₂ observations, atmospheric CO₂ data assimilation has become a direct and effective approach to identify carbon sources and sinks (Peters et al. 2007; Peng et al. 2015; Tian et al. 2014). Since the model's result is an estimation of the real world, the performance of the model is limited by the accuracy of the priori flux field

and the model transport mechanism. Ground-based observations and satellite-based observations have high degrees of accuracy but have the drawbacks of limited spatial distributions and time spans (Yue et al. 2016a).

Based on the High Accuracy Surface Modeling (HASM) developed from the fundamental theorem of surfaces, the model simulation results are used as the driving fields, and the observational data as the optimal control conditions to obtain more accurate CO₂ distribution fields. In this paper, we first simulated the CO₂ distributions over North China using Weather Research and Forecasting - Chemistry (WRF-CHEM) as a regional atmospheric chemical transport model and evaluated the accuracy of the WRF-CHEM simulation results. Second, the CO₂ concentration field simulated by the WRF-CHEM and GOSAT inversion data is introduced into HASM to obtain XCO₂ concentrations. Then, the availability of HASM is verified by comparing its results with those of the classical interpolation methods commonly used in past studies (Xu et al. 2013; Liu et al. 2011). Finally, the distribution of XCO₂ from the HASM simulation results and the changes of the monthly evolutions are discussed.

Data and methods

Study region

The study area is located in the North China, between 34.3° N–43.5° N and 111° E–121.9° E (see Fig. 1). This region incorporates the Beijing-Tianjin-Hebei region, which is a key economic zone in northern China, and Shandong Province, which is a large economically important province with the second largest population in China. In addition, part of Henan Province (which has the third largest population), Shanxi Province, Liaoning Province, and Inner Mongolia, which have a variety of underlying surfaces, including grass, forests, farmland, and cities, are also included in the study area.

GOSAT XCO₂ data

GOSAT (Greenhouse gases Observing SATellite) was jointly designed by the Ministry of the Environment (MOE), the National Institute for Environment Studies (NIES), and the Japan Aerospace Exploration Agency (JAXA), and was successfully launched on January 23, 2009 by the Japanese Space Agency. This satellite is the first to be used to specifically monitor the concentrations of CO₂ and CH₄ from space, operating at a 666 km sun-synchronous orbit with a 3-day recurrence. The TANSO-FTS on the satellite can detect the gas-absorption spectra of the reflected light in the Short Wave InfraRed (SWIR) region (0.76, 1.6, and 2.0 μm) and Thermal InfraRed (TIR) band (from 5.5 to 14.3 μm) from the Earth's

Fig. 1 Land cover of the study region derived from ESA (European Space Agency) global land cover dataset. The red dots are the locations of the available GOSAT XCO₂ L2 data points during the study period (from February to May 2015). The blue dot is the location of the Shangdianzi Global Atmosphere Watch (GAW) Regional Station

surface. Using these spectra data, CO₂ and CH₄ can be retrieved with a footprint that measures 10.5 km along one side (Kuze et al. 2009; Kadyrov et al. 2009; Yokota et al. 2009).

In this study, the column-averaged dry-air fractions of the CO₂ (XCO₂) datasets released by the GOSAT project are used as the true values to verify the chemistry transport model outputs and drive the HASM. We chose the GOSAT FTS SWIR L2 data (hereafter GOSAT XCO₂) from February to May of 2015, considering the consecutiveness of the data availability in the study area, and the data were filtered by the screening procedures described in NIES GOSAT TANSO-FTS SWIR Level 2 Data Product Format Description. The distribution of valid GOSAT XCO₂ points is shown in Fig. 1, noting that each point may have been observed several times. The number of valid GOSAT XCO₂ measurements from February to May are 39, 37, 30, and 33 respectively. The website containing the data is https://data2.gosat.nies.go.jp/index_en.html.

WDCGG

The WDCGG (World Data Centre for Greenhouse Gases) is one of the data archiving and service centers under the GAW (Global Atmosphere Watch) program of WMO (World Meteorological Organization) that collects all observational data of greenhouse gases. Such as CO₂, CH₄, CFC, N₂O, from 378 stations all over the world. The WDCGG was founded in

October 1990 and operated at the Japan Meteorological Agency (JMA). The download website for the WDCGG data is <https://ds.data.jma.go.jp/gmd/wdogg/wdogg.html>. Based on the study period and study area, Shangdianzi Station is used as the ground observation point to verify the output of the chemistry transport model at the surface. Shangdianzi Station is located at 40.65° N, 117.12° E and is 150 km northeast of urban Beijing (see Fig. 1). This station is mainly influenced by land flux and anthropogenic emissions (Li et al. 2017).

WRF-CHEM

The WRF-CHEM (Weather Research and Forecasting-Chemistry) is a regional air quality model developed by the National Oceanic and Atmospheric Administration (NOAA) for biomass combustion, anthropogenic emissions, chemical vapor schemes, and aerosol solutions; the model includes a trace gas transport option and a subroutine to calculate the plume lifting (Simpson et al. 1995; Sandu et al. 2003; Liu et al. 2012). The main components of the WRF-CHEM involved in this study are the WRF and greenhouse gas module.

WRF

The WRF (Weather Research and Forecasting) model is a new generation of mesoscale weather forecasting model and was codeveloped by NCAR and NCEP. This model has a fully

compressible non-static equilibrium mode and contains a wealth of physical parameterization options (Wei et al. 2016; William et al. 2008). Due to its high prediction accuracy, strong portability, fast calculation, and easy maintenance, this model is widely used in meteorological research and businesses around the world. In WRF-CHEM, the role of the WRF is to provide real-time meteorological fields for chemical modules. Because the WRF can provide a meteorological element field with a high spatiotemporal resolution, WRF-CHEM can meet the current requirements for refined forecasting. This is one of the biggest advantages of WRF-CHEM.

Greenhouse gas module

The greenhouse gas module of WRF-CHEM was developed by the Max Planck Institute for Biogeochemistry (Beck et al. 2011). This module is able to simulate the distributions and transports of greenhouse gases (passive tracers such as CO₂, CH₄, and CO) with high resolutions and is able to obtain the initial field and boundary conditions required for simulations from the global atmospheric transport model (Ahmadov et al. 2007; Diao et al. 2015; Pillai et al. 2016). The model was originally called WRFHG (WRF Greenhouse Gas model). Starting at WRF-CHEM v3.4, the corresponding module of WRFHG was officially included in WRF-CHEM.

In the greenhouse gas module, the Vegetation Photosynthesis Respiration Model (VPRM) is a key element used to estimate the net ecosystem exchange (NEE), including the light-driven gross ecosystem exchange and the ecosystem respiration term driven by temperature.

$$NEE = -\lambda \times T_{scale} \times P_{scale} \times W_{scale} \times EVI \times \frac{1}{(1 + PAR/PAR_0)} \times PAR + \alpha \times T + \beta \quad (1)$$

In which $\lambda(\mu\text{mol CO}_2 \text{ m}^{-2}\text{s}^{-1}/(\mu\text{mol PAR m}^{-2}\text{s}^{-1}))$ is the maximum quantum yield, and $PAR_0(\mu\text{mol m}^{-2}\text{s}^{-1})$ is the half saturation value of the photosynthetically active radiation. EVI (Enhanced Vegetation Index) represents the ratio of the absorbed photosynthetically active radiation to the total photosynthetically active radiation. T_{scale} , P_{scale} , and W_{scale} , respectively, represent the characteristics of the leaf temperatures, leaf surface characteristics, and canopy water contents. The functions of T_{scale} , P_{scale} , and W_{scale} are shown as follows:

$$T_{scale} = \frac{(T - T_{min})(T - T_{max})}{[(T - T_{min})(T - T_{max}) - (T - T_{opt})^2]} \quad (2)$$

$$W_{scale} = \frac{1 + LSWI}{1 + LSWI_{max}} \quad (3)$$

$$P_{scale} = \frac{1 + LSWI}{2} \quad (4)$$

where $T(^{\circ}\text{C})$ is the temperature in atmosphere, and T_{min} , T_{max} , and T_{opt} represent the minimum, maximum, and optimum temperatures for photosynthesis. T_{scale} is set to 0 when the air temperature is less than T_{min} . LSWI (Land Surface Water Index) is the moisture content of the vegetation canopy, and $LSWI_{max}$ is the maximum LSWI value during the growing season in each grid cell. The value of P_{scale} depends on the growth stages of the vegetation, and for the evergreen forest, the value of P_{scale} is fixed at 1.0. In other cases, T_{scale} , P_{scale} , and W_{scale} range from 0.0 to 1.0. Compared with other models that treat breathing as an exponential function of temperature, VPRM reduces breathing to a linear function of temperature in which $\alpha(\mu\text{mol CO}_2 \text{ m}^{-2}\text{s}^{-1}/^{\circ}\text{C})$ and $\beta(\mu\text{mol CO}_2 \text{ m}^{-2}\text{s}^{-1})$ can be adjusted according to the observed data.

Running of WRF-CHEM

We use a $1^{\circ} \times 1$ reanalysis produce, i.e., the ERA-Interim data, with time intervals of 6 h as downloaded from the European Centre for Medium-Range Weather Forecasts (ECMWF) for the initial field and boundary conditions for WRF. The total surface CO₂ exchanges are calculated as follows:

$$F_t = F_{ant} + F_R + F_{GEE} + F_{fire} \quad (5)$$

where F_t denotes the total CO₂ flux. F_{ant} is the anthropogenic emissions obtained from the Emission Database for Global Atmospheric Research (EDGAR). F_R and F_{GEE} are the biospheric respiration and gross ecosystem exchange, respectively, and both are calculated via the VPRM in greenhouse gas module. F_{fire} is the biomass burning emissions provided by the Global Fire Emissions Database (GFED). The initial fields and boundary conditions of the CO₂ concentrations from different sources are adopted from CarbonTracker2016.

The input of the VPRM module includes four kinds of data. The temperature at 2 m and the downward shortwave flux at the ground surface can be provided by the coupled WRF. LSWI and EVI are calculated from the Terra MODIS satellite level-3 land product (MYD09A1: MODIS/Aqua Surface Reflectance 8-Day L3 Global 500 m SIN Grid V006) via the VPRM-Preprocessor tool.

The WRF-CHEM utilized is WRF-CHEM version 3.6, which is used to model from February to May 2015. During the study period, simulations were performed four times. Each simulation has a 6-h spin up for its meteorology and a 1-month run time period for the CO₂ transport. The model domain is centered at 39.0° N, 116.5° E, with a 10 km × 10 km horizontal resolution, 35 vertical layers in the terrain-following hydrostatic-pressure vertical coordinate system from the surface to 50 hPa, and an hourly output on the Lambert projection. The chosen physical parameterization schemes are WSM 5-class microphysics scheme, RRTM longwave radiation scheme, Goddard short wave radiation scheme, revised MM5 Monin-Obukhov

surface-layer scheme, unified Noah land-surface model, YSU boundary-layer scheme, and Grell 3D ensemble cumulus scheme. The dataset, components, and workflow involved in the operation of WRF-CHEM are shown in Fig. 2.

HASM

Based on the fundamental theorem of surface, a surface is uniquely determined by the first and second fundamental coefficients. The first fundamental coefficients of a surface describe the geometric properties of the surface, by which we can calculate the lengths of the curves, the angles of the tangent vectors, the areas of regions, and the geodesics on the surface. These geometric properties and objects are called the intrinsic geometric properties and are only determined via the first fundamental coefficients of a surface, depending on measurements that we can conduct while staying on the surface itself (Toponogov 2006). The second fundamental coefficients of a surface describe the local deformations of the surface, which can be observed while staying above the surface. In other word, these are the deviations of the relevant point on the surface from the tangent plane (Liseikin 2004).

Suppose a surface z can be represented by a function of x and y ; that is, $z=f(x, y)$. The first fundamental coefficients, E , F , and G , and the second fundamental coefficients, L , M , and N , are defined as follow:

$$E = 1 + f_x^2 \quad (6)$$

$$F = f_x \cdot f_y \quad (7)$$

$$G = 1 + f_y^2 \quad (8)$$

$$L = \frac{f_{xx}}{\sqrt{1 + f_x^2 + f_y^2}} \quad (9)$$

$$M = \frac{f_{xy}}{\sqrt{1 + f_x^2 + f_y^2}} \quad (10)$$

$$N = \frac{f_{yy}}{\sqrt{1 + f_x^2 + f_y^2}} \quad (11)$$

In which f_x is the first order partial derivative of the surface z with respect to the independent variable x . f_{xx} is the second order partial derivative of the surface z with respect to the independent variable x .

A similar definition applies for f_y and f_{yy} . f_{xy} is the second order mixed partial derivative of the surface z with respect to the independent variables x and y successively.

Based on previous research, an equation set called the Gauss equation was found to relate the intrinsic curvature of the surface to the derivatives of the Gauss map, namely, the first fundamental coefficients (Eqs. 6–8) and the second fundamental coefficients (Eq. 9–11) satisfy the following equation set:

$$f_{xx} = \Gamma_{11}^1 f_x + \Gamma_{11}^2 f_y + \frac{L}{\sqrt{E + G - 1}} \quad (12)$$

$$f_{xy} = \Gamma_{12}^1 f_x + \Gamma_{12}^2 f_y + \frac{M}{\sqrt{E + G - 1}} \quad (13)$$

$$f_{yy} = \Gamma_{22}^1 f_x + \Gamma_{22}^2 f_y + \frac{N}{\sqrt{E + G - 1}} \quad (14)$$

where

$$\Gamma_{11}^1 = \frac{GE_x - 2FF_x + FE_y}{2(EG - F^2)} \quad (15)$$

$$\Gamma_{12}^1 = \frac{GE_x - FG_x}{2(EG - F^2)} \quad (16)$$

$$\Gamma_{22}^1 = \frac{2GF_y - GG_x - FG_y}{2(EG - F^2)} \quad (17)$$

$$\Gamma_{11}^2 = \frac{2EF_x - EE_y - FE_x}{2(EG - F^2)} \quad (18)$$

Fig. 2 The dataset, components, and workflow involved in the operation of WRF-CHEM

Fig. 3 The workflow used to obtain the XCO₂ field based on HASM

$$\Gamma_{12}^2 = \frac{EG_x - FE_y}{2(EG - F^2)} \quad (19)$$

$$\Gamma_{22}^2 = \frac{EG_y - 2FF_y + FG_x}{2(EG - F^2)} \quad (20)$$

Fig. 4 The comparison of CO₂ values near the surface between simulations and observations. The red lines show the WRF-CHEM grid CO₂ simulations. The blue line is the observations of Shangdianzi Station from WDCGG

Table 1 Basic statistics of XCO₂ from WRF-CHEM and GOSAT

	2015.02		2015.03		2015.04		2015.05	
	WRF-CHEM	GOSAT XCO ₂	WRF-CHEM	GOSAT XCO ₂	WRF-CHEM	GOSAT XCO ₂	WRF-CHEM	GOSAT XCO ₂
Max (ppmv)	404.41	405.35	403.87	406.36	405.35	404.71	404.97	403.22
Min (ppmv)	400.92	396.43	401.58	395.60	402.97	398.09	402.01	398.27
Mean (ppmv)	402.17	400.21	402.41	399.59	403.84	400.65	403.83	400.87
Variance	1.04	2.74	0.57	2.96	0.72	1.71	0.66	1.31
Correlation coefficient	0.78		0.51		0.56		0.40	

$\Gamma_{11}^1, \Gamma_{12}^1, \Gamma_{22}^1, \Gamma_{11}^2, \Gamma_{12}^2$, and Γ_{22}^2 are the second type of Christoffel symbols and rely only on the first fundamental coefficients, E, F , and G , and their derivatives. In the process of solution of the Eqs. 12–14, the central difference method is used to displace the partial derivatives.

We mark the first, second, and third equations of Gauss equation set as a, b, and c. Then, HASM abc can be expressed as a constrained least-squares approximation.

$$\left\{ \min \left\| \begin{bmatrix} A \\ B \\ C \end{bmatrix} \cdot z^{(n-1)} - \begin{bmatrix} d \\ q \\ h \end{bmatrix} \right\|_2 \right. \quad (21)$$

$$\left. S \cdot z^{(n-1)} = k \right.$$

where the second equation of Eq. 21 is the constraint equation representing the sampling points information. A, B , and C are the coefficient matrix of the discrete equation form of the Gauss equation. d, q , and p are found on the right-hand side of the Gauss equation. S denotes the sampling matrix, and k denotes the sampling vector.

Equation 21 is a least-squares problem constrained by terrestrial sampling. The purpose of Eq. 21 is to confine the overall simulation error to a minimum value, while keeping the simulated value equal to the sample value at the sampling point. Taking full advantage of the sampling information is

also an effective way to ensure that the iterative formulation of HASM approaches the best simulation result.

HASM can also be transferred into an unconstrained least-squares approximation:

$$\begin{bmatrix} A \\ B \\ C \\ \lambda \cdot S \end{bmatrix} Z^{(n+1)} = \begin{bmatrix} A^T B^T C^T \lambda \cdot S^T \end{bmatrix} \begin{bmatrix} d^{(n)} \\ q^{(n)} \\ p^{(n)} \\ \lambda \cdot k \end{bmatrix} \quad (22)$$

where λ denotes the weight of the sampling point values, which refers the relative importances of the sampling points in the simulated field. λ could be a real number or a vector, depending on whether all sampling points are equally important or if each point has its own weight.

In the existing research, HASM have been applied for the study of soil properties (Shi et al., 2011, 2016), carbon storage (Wang et al., 2016; Yue et al. 2016b), and climate change (Yue et al. 2013; Zhao and Yue 2014). At a global

Fig. 5 Comparison of the three methods and GOSAT retrieval data on 2015.02. **a** HASM. **b** Kriging. **c** IDW

scale, Zhao (Zhao et al. 2017a, b) and Zhang (2017b) introduced HASM for XCO₂ simulations with plenty of sampling points. Yue et al. (2015) verified that HASM is an alternative approach to filling voids on XCO₂ surfaces from satellites. In this study, we force a regional scale simulation with a scant number of sampling points to investigate the performance of HASM. Based on HASM, the workflow used to obtain the XCO₂ field is shown in Fig. 3.

Results and discussion

Comparison between WRF-CHEM and WDCGG

Figure 4 shows the comparison between the WRF-CHEM CO₂ data and the ground-based measurements from Shangdianzi Station during the study period. Overall, the WRF-CHEM CO₂ simulations represent the fluctuations of the observation time series, with a high correlation

Fig. 6 Comparison of the three methods and GOSAT retrieval data on 2015.03. **a** HASM. **b** Kriging. **c** IDW

coefficient of approximately 0.94. The difference between the model and observations is small in the low value range (approximately 405 ppmv), whereas the bias becomes large when the observed CO₂ concentrations are high and the model values are lower. The reason for this difference may be that the data from the model is a mean value of a $0.1^{\circ} \times 0.1$ grid cell. Therefore, the simulations

results reflect the smooth characteristics. Due to the complexity of atmosphere and the lack of understanding of atmospheric motion, WRF includes many approximations, and the choice of many parameters is debatable. Because of these issues, WRF is unable to accurately simulate the wind field, thereby affecting CO₂ transmission and diffusion in the atmosphere. In addition, the emission data are

Fig. 7 Comparison of the three methods and GOSAT retrieval data on 2015.04. **a** HASM. **b** Kriging. **c** IDW

not very accurate, which also affects the distribution of CO₂ concentration.

Comparison between WRF-CHEM and GOSAT

Based on the information of the dates and longitudes/latitudes in the GOSAT retrieval dataset, we extract the simulation data for the same dates and positions to compare between the simulation and GOSAT XCO₂. Note that the original WRF-CHEM output is a layered CO₂ concentration.

Thus, here, WRF-CHEM XCO₂ is calculated based on the following function:

$$XCO_2 = \sum_{i=1}^N h_i u_i \quad (23)$$

where N is the number of layers in the original WRF-CHEM output, u_i presents the CO₂ concentrations as each layer, and h_i

Fig. 8 Comparison of the three methods and GOSAT retrieval data on 2015.05. **a** HASM. **b** Kriging. **c** IDW

presents the pressure weighting function which relates the layered CO₂ concentration to the profile-weighted average (Connor et al. 2008).

$$h_i = \left| \left(-p_i + \frac{p_{i+1} - p_i}{\ln\left(\frac{p_{i+1}}{p_i}\right)} \right) + \left(p_i - \frac{p_i - p_{i-1}}{\ln\left(\frac{p_i}{p_{i-1}}\right)} \right) \right| \frac{1}{p_{\text{surf}}} \quad (24)$$

where p_i is the pressure on each level, and p_{surf} is the surface pressure. For the upper or lower boundary layers, the function retains only the left or right items within the absolute operator.

The basic statistics of the XCO₂ from the WRF-CHEM and GOSAT retrieval data are shown in Table 1. In the 4-month study period, the simulated minimums are approximately 4–5 ppmv higher than those of the GOSAT XCO₂, and the average of the simulations is also approximately 2–3 ppmv higher. Considering XCO₂ is the general representation of the CO₂ content from the surface to the top of atmosphere, and the lower simulated values of the surface CO₂ concentration are shown in Fig. 4, the higher simulated values of XCO₂ from WRF-CHEM may come from excess CO₂ simulated in the troposphere. The simulated variances of each month in Table 1 are less than the observed variances, which is a common problem for transmission models (Lei et al. 2014). The correlation coefficients of WRF-CHEM XCO₂ and GOSAT XCO₂ are between 0.4 and 0.78, being highest in February and lowest in May.

From Fig. 4 and Table 1, WRF-CHEM has a certain degree of ability to simulate regional CO₂ concentrations. However, due to the limited emission inventories and the performances of the models, the simulation accuracies remain to be improved. In the next chapter, we adopt HASM to achieve our goal, such that the CO₂ concentrations of the WRF-CHEM model are taken as driving fields, and GOSAT XCO₂ is used as the accuracy control conditions.

Comparison with the Kriging and IDW methods

Taking the simulated XCO₂ fields from WRF-CHEM and the retrieval data from GOSAT as the driving field and accuracy control conditions, respectively, we operate HASM to update the XCO₂ field. Meanwhile, as the classic interpolation methods, Kriging and IDW methods are also involved in the experiments to compare with HASM. The XCO₂ data is calculated on a monthly basis due to limited retrieval data. For each month, we organize the cross-validation test, which means that one data point in GOSAT XCO₂ is selected as the test point, and the remaining points are used to drive the three methods. After this, the simulation values are extracted for the test points.

Comparisons of the three methods and the GOSAT retrieval data for each month are shown in Figs. 5, 6, 7, and 8. The fitting equations of HASM present the highest R^2 values among the three methods in all cases; meanwhile, IDW

performs the worst, except in May. Furthermore, we use two statistics to show the differences of the three methods:

$$\text{MAE} = \frac{1}{N} \sum_I^N |O_I - S_I| \quad (25)$$

$$\text{RMSE} = \sqrt{\frac{\sum_I^N (O_I - S_I)^2}{N-1}}$$

Table 2 shows the monthly mean absolute error (MAE) and root mean square error (RMSE) values of the three methods and indicates that HASM perform better than the classic interpolation methods. This is because that the HASM used in this paper is not a typical spatial interpolation method to construct new data points using only a finite set of known data points. The output of HASM normally contains information from both finite known data points and an approximate field. To some extent, the HASM used in this paper is more of a data fusion method.

Spatiotemporal distribution of XCO₂ from HASM

The monthly XCO₂ outputs from HASM are reproduced in Fig. 9. The distribution of XCO₂ varies by month. In February, the XCO₂ is much higher in the southern regions and not the northern regions of the study area. In Shandong Province and the southern part of Hebei Province, the XCO₂ exceeds 402 ppmv, in contrast to the values of less than 398 ppmv observed in Inner Mongolia. The spatial distribution trend in March is similar to that in February, although the high value zone contracts. This distribution phenomena may be related to the heating in the winter and the population distribution. In the northern part of study area, there is a small population and thus a lower energy consumption for heating. The middle of the study area has a great population, and thus, the XCO₂ is relatively high. Shandong and Henan provinces are the second and third most populous in China and contribute the most carbon emission for heating. Especially in southern Shandong, burning coal to keep warm in rural areas aggravates the carbon emissions. Besides, northerly winds prevailing in the winter reinforce this distribution.

Table 2 Mean absolute error (MAE) and root mean square error (RMSE) values of the three methods

	MAE (ppmv)			RMSE (ppmv)		
	HASM	Kriging	IDW	HASM	Kriging	IDW
2015.02	1.50	1.75	1.77	1.89	2.15	2.10
2015.03	1.53	1.48	1.48	1.98	2.00	2.04
2015.04	0.80	0.83	0.95	0.98	1.03	1.10
2015.05	0.65	0.76	0.72	0.80	1.02	0.88
mean	1.12	1.27	1.23	1.41	1.55	1.53

With vegetation growth in the terrestrial ecosystems and the end of heat use in April, the difference of the north-south distribution of XCO_2 rapidly decreases, and the high value zone (i.e., greater than 402 ppmv) disappeared. The concentration is further reduced in May. Only in Beijing and Tianjin as well as at the border of Shandong, Henan, and Hebei is the XCO_2 greater than 401 ppmv but less than 402 ppmv, which means that the carbon emissions in these areas are strongly influenced by human activity.

Conclusions

According to the fundamental theorem of surfaces, a surface is determined by the first and the second fundamental coefficients. In this paper, HASM is applied to obtain high precision simulated XCO_2 fields, using the simulated field outputs from the atmospheric chemical transport model as the approximate driving

fields and the observed data as the accuracy control points. The main conclusions of this paper are as follows: (a) The simulation results of the regional atmospheric transport model WRF-CHEM can reflect the changes and distributions of CO_2 concentrations to a certain extent. Although the precision of this model is poor, the result can provide driving field information for HASM for an approximate surface. (b) Comparing the output of HASM with those of the classical Kriging and IDW interpolation schemes shows that the simulation results of HASM have relatively higher accuracies during the study period. A cross-validation shows that the MAE from HASM is 1.12 ppmv and RMSE is 1.41 ppmv. (c) The CO_2 concentration field simulated by HASM has obvious spatial differentiations and changes with the seasonal changes. Due to heating emissions in the winter, densely populated areas have higher CO_2 concentrations. With the end of the heating period and the new growth of plants, the high concentration values rapidly decrease. At this time, the distribution of high concentration area is related to industrial activities.

Fig. 9 Spatiotemporal distribution of monthly XCO_2 values from HASM. **a** February. **b** March. **c** April. **d** May

Fig. 9 (continued)

Compared with the previous research of HASM, the few observed data are used in this study. Meanwhile, the accuracy of the WRF-CHEM simulation field is limited. These factors will transmit biases into the outputs of HASM. Therefore, increasing the number of observation points and improving the driving field accuracies are important methods to obtain more accurate CO₂ distributions using HASM. In addition, the GOSAT inversion data is used as the true values due to the lack of XCO₂ observations from ground-based observation stations within the study area. This study should be repeated as more data from ground- and space-based observations become available.

Acknowledgments We are grateful for the CO₂ observations from Shangdianzi Station, which were provided by the research teams on the WDCGG website, the XCO₂ data from the NIES GOSAT Projects, and the 300 m annual global land cover dataset from ESA. We also thank the WRF-GHG team for providing us with access to the WRF-GHG, VPRM-Preprocessor tool, and relevant data.

Funding information This work was supported by the National Natural Science Foundation of China (41590844), the National Natural Science Foundation of China (41421001), and the Innovation Project of LREIS (O88RA600YA).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ahmadv R, Gerbig C, Kretschmer R, Koerner S, Neininger B, Dolman AJ, Sarra C (2007) Mesoscale covariance of transport and CO₂ fluxes: evidence from observations and simulations using the WRF-VPRM coupled atmosphere biosphere model. *J Geophys Res-Atmos* 112:1–14

- Barkley MP, Monks PS, Hewitt AJ, Machida T, Desai A, Vinnichenko N, Nakazawa T, Arshinov MY, Fedoseev N, Watai T (2007) Assessing the near surface sensitivity of SCIAMACHY atmospheric CO₂ retrieved using (FSI) WFMDOAS. *Atmos Chem Phys* 7:3597–3619
- Beck V, Koch T, Kretschmer R, Marshall J, Ahmadov R, Gerbig C, Pillai D, Heimann M (2011) The WRF Greenhouse Gas Model (WRF-GHG). Technical Report No. 25, Max Planck Institute for Biogeochemistry, Jena, Germany
- Connor BJ, Boesch H, Toon G, Sen B, Miller C, Crisp D (2008) Orbiting carbon observatory: inverse method and prospective error analysis. *J Geophys Res-Atmos* 113:1–14
- Diao YW, Huang JP, Liu C, Cui J, Liu SD (2015) A modeling study of CO₂ flux and concentrations over the Yangtze River Delta using the WRF-GHG model. *Chin J Atmos Sci (in Chinese)* 39:849–860
- Fischer ML, Parazoo N, Brophy K, Cui XG, Jeong S, Liu JJ, Keeling R, Taylor TE, Gurney K, Oda T, Graven H (2017) Simulating estimation of California fossil fuel and biosphere carbon dioxide exchanges combining in-situ tower and satellite column observations. *J Geophys Res* 122:3653–3671
- Hakkaraianen J, Ialongo I, Tamminen J (2016) Direct space-based observations of anthropogenic CO₂ emission areas from OCO-2. *Geophys Res Lett* 43:11400–11406
- Kadyrov 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-Atmos* 114:1–12
- Kuze A, Suto H, Nakajima M, Hamazaki T (2009) Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the greenhouse gases observing satellite for greenhouse gases monitoring. *Appl Opt* 48:6716–6733
- Lei LP, Guan XH, Zeng ZC, Zhang B, Ru F, Bu R (2014) A comparison of atmospheric CO₂ concentration GOSAT-based observations and model simulations. *Sci China Earth Sci* 57:1393–1402
- Li R, Zhang MG, Chen LF, Kou XX, Skorokhod A (2017) CMAQ simulation of atmospheric CO₂ concentration in East Asia: comparison with GOSAT observations and ground measurements. *Atmos Environ* 160:176–185
- Liseikin VD (2004) A computational differential geometry approach to grid generation. Springer, Berlin
- Liu Y, Wang XF, Guo M, Tani H (2011) Mapping the FTS SWIR L2 product of XCO₂ and XCH₄ data from the GOSAT by the Kriging method—a case study in East Asia. *Int J Remote Sens* 33:3004–3025
- Liu X, Easter RC, Ghan SJ, Zaveri R, Rasch P, Shi X, Lamarque JF, Gettelman A, Morrison H, Vitt F, Conley A, Park S, Neale R, Hannay C, Ekman AML, Hess P, Mahowald N, Collins W, Iacono MJ, Bretherton CS, Flanner MG, Mitchell D (2012) Toward a minimal representation of aerosols in climate models: description and evaluation in the community atmosphere Model CAM5. *Geosci Model Dev* 5:709–739
- Peng Z, Zhang M, Kou X, Tian X, Ma X (2015) A regional carbon data assimilation system and its preliminary evaluation in East Asia. *Atmos Chem Phys* 15:1087–1104
- Peters W, Jacobson AR, Sweeney C, Andrews AE, Conway TJ, Masarie K, Miller JB, Bruhwiler LMP, Petron G, Hirsh AI, Worthy DEJ, Werf GR, Randerson JT, Wennberg PO, Krol MC, Tans PP (2007) An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proc Natl Acad Sci USA/PNAS* 104:18925–18930
- Pillai D, Buchwitz M, Gerbig C, Koch T, Reuter M, Bovensmann H, Marshall J, Burrows JP (2016) Tracking city CO₂ emissions from space using a high-resolution inverse modelling approach: a case study for Berlin, Germany. *Atmos Chem Phys* 16:9591–9610
- Sandu A, Daescu DN, Carmichael GR (2003) Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: part I - theory and software tools. *Atmos Environ* 37:5083–5096
- Shi WJ, Liu JY, Du ZP, Stein A, Yue TX (2011) Surface modelling of soil properties based on land use information. *Geoderma* 162:347–357
- Shi WJ, Yue TX, Du ZP, Wang Z, Li XW (2016) Surface modeling of soil antibiotics. *Sci Total Environ* 543:609–619
- Shim C, Lee J, Wang YX (2013) Effect of continental sources and sinks on the seasonal and latitudinal gradient of atmospheric carbon dioxide over East Asia. *Atmos Environ* 79(7):853–860
- Simpson D, Guenther A, Hewitt CN, Steinbrecher R (1995) Biogenic emissions in Europe. 1 Estimates and uncertainties. *J Geophys Res* 100D:22875–22890
- Tian X, Xie Z, Liu Y, Cai Z, Fu Y, Zhang H, Feng L (2014) A joint data assimilation system (Tan-Tracker) to simultaneously estimate surface CO₂ fluxes and 3-D atmospheric CO₂ concentrations from observations. *Atmos Chem Phys* 14:13281–13293
- Toponogov VA (2006) Differential geometry of curves and surfaces. Birkhaeuser Boston, New York
- Turner AJ, Jacob DJ, Wecht KJ, Maasakkers JD, Lundgren E, Andrews AE, Biraud SC, Boesch H, Bowman KW, Deutscher NM, Dubey MK, Griffith DWT, Hase F, Kuze A, Notholt J, Ohyama H, Parker R, Payne VH, Sussmann R, Sweeney C, Velasco VA, Warneke T, Wennberg PO, Wunch D (2015) Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. *Atmos Chem Phys* 15:4495–4536
- Wang K, Jiang H, Zhang XY, Zhou GM (2011) Analysis of spatial and temporal variations of carbon dioxide over China using SCIAMACHY satellite observations during 2003–2005. *Int J Remote Sens* 32:815–832
- Wang YA, Lu DR, Li Q, Duan MZ, Hu F, Hu SX (2014) Observed and simulated features of the CO₂ diurnal cycle in the boundary layer at Beijing and Hefei, China. *Chin Sci Bull* 59:1529–1535
- Wang YF, Yue TX, Lei YC, Du ZP, Zhao MW (2016) Uncertainty of forest biomass carbon patterns simulation on provincial scale: a case study in Jiangxi Province, China. *J Geogr Sci* 26:568–584
- Wei JF, Su H, Yang ZL (2016) Impact of moisture flux convergence and soil moisture on precipitation: a case study for the southern United States with implications for the globe. *Clim Dynam* 64(1–2):467–481
- William CS, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang XY, Wang W, Powers JG (2008) A description of the advanced research WRF version 3. Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder. NCAR technical note
- WMO Greenhouse Gas Bulletin (2017) The state of greenhouse gases in the atmosphere based on global observations through 2016. World Meteorological Organization, Geneva
- Xu YY, Ke CQ, Wang JL, Sun JL, Liu Y, Harris W, Kou C (2013) Satellite-derived estimations of spatial and seasonal variation in tropospheric carbon dioxide mass over China. *Ecol Evol* 3:4310–4325
- Yokota T, Yoshida Y, Eguchi N, Ota Y, Tanaka T, Watanabe H, Maksyutov S (2009) Global concentration of CO₂ and CH₄ retrieved from GOSAT: first preliminary results. *SOLA* 5:160–163
- Yue TX, Zhao N, Ramsey D, Wang CL, Fan ZM, Chen CF, Lu YM, Li BL (2013) Climate change trend in China, with improved accuracy. *Clim Chang* 120:137–151

- Yue TX, Zhao MW, Zhang XY (2015) A high-accuracy method for filling voids on remotely sensed XCO₂ surfaces and its verification. *J Clean Prod* 103:819–827
- Yue TX, Zhang LL, Zhao MW, Wang YF, Wilson JP (2016a) Space-and ground-based CO₂ measurements: a review. *Sci China Earth Sci* 59: 2089–2097
- Yue TX, Liu Y, Zhao MW, Du ZP, Zhao N (2016b) A fundamental theorem of Earth's surface modelling. *Environ Earth Sci* 75:1–12
- Zhang LL, Yue TX, Wilson JP, Wang DY, Zhao N, Liu Y, Liu DD, Du ZP, Wang YF, Lin C, Zheng YQ, Guo JH (2017a) Modelling of XCO₂ surface based on flight test of TanSat instruments. *Sensors* 16:1818
- Zhang LL, Yue TX, Wilson JP, Zhao N, Zhao YP, Du ZP, Liu Y (2017b) A comparison of satellite observations with the XCO₂ surface obtained by fusing TCCON measurements and GEOS-Chem model outputs. *Sci Total Environ* 601–602:1575–1590
- Zhao N, Yue TX (2014) A modification of HASM for interpolating precipitation in China. *Theor Appl Climatol* 116(1–2):273–285
- Zhao MW, Yue TX, Zhang XY, Sun JL, Jiang L, Wang C (2017a) Fusion of multi-source near-surface CO₂ concentration data based on high accuracy surface modeling. *Atmos Pollut Res* 8:1170–1178
- Zhao MW, Zhang XY, Yue TX, Wang C, Jiang L, Sun JL (2017b) A high-accuracy method for simulating the XCO₂ retrieval data. *Sci China Earth Sci* 60:143–155