<table>
<thead>
<tr>
<th>Title</th>
<th>Assessment of Global Carbon Dioxide Concentration Using MODIS and GOSAT Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Guo, Meng; Wang, Xiufeng; Li, Jing; Yi, Kunpeng; Zhong, Guosheng; Tani, Hiroshi</td>
</tr>
<tr>
<td>Citation</td>
<td>Sensors, 12(12): 16368-16389</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2012-12</td>
</tr>
<tr>
<td>Doc URL</td>
<td><a href="http://hdl.handle.net/2115/51823">http://hdl.handle.net/2115/51823</a></td>
</tr>
<tr>
<td>Type</td>
<td>article</td>
</tr>
<tr>
<td>File Information</td>
<td>Sen12-12_16368-16389.pdf</td>
</tr>
</tbody>
</table>
Assessment of Global Carbon Dioxide Concentration Using MODIS and GOSAT Data

Meng Guo 1,*, Xiufeng Wang 2, Jing Li 3, Kunpeng Yi 1, Guosheng Zhong 1 and Hiroshi Tani 2

1 Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan; E-Mails: yikp@env.agr.hokudai.ac.jp (K.Y.); zhgs@env.agr.hokudai.ac.jp (G.Z.)
2 Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan; E-Mails: wang@env.agr.hokudai.ac.jp (X.W.); tani@env.agr.hokudai.ac.jp (H.T.)
3 Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China; E-Mail: lijingsara@neigae.ac.cn

* Author to whom correspondence should be addressed; E-Mail: guomeng@env.agr.hokudai.ac.jp; Tel./Fax: +81-11-706-4174.

Received: 29 October 2012; in revised form: 22 November 2012 / Accepted: 23 November 2012 / Published: 26 November 2012

Abstract: Carbon dioxide (CO2) is the most important greenhouse gas (GHG) in the atmosphere and is the greatest contributor to global warming. CO2 concentration data are usually obtained from ground observation stations or from a small number of satellites. Because of the limited number of observations and the short time series of satellite data, it is difficult to monitor CO2 concentrations on regional or global scales for a long time. The use of the remote sensing data such as the Advanced Very High Resolution Radiometer (AVHRR) or Moderate Resolution Imaging Spectroradiometer (MODIS) data can overcome these problems, particularly in areas with low densities of CO2 concentration watch stations. A model based on temperature (MOD11C3), vegetation cover (MOD13C2 and MOD15A2) and productivity (MOD17A2) of MODIS (which we have named the TVP model) was developed in the current study to assess CO2 concentrations on a global scale. We assumed that CO2 concentration from the Thermal And Near infrared Sensor for carbon Observation (TANSO) aboard the Greenhouse gases Observing SAteLLite (GOSAT) are the true values and we used these values to check the TVP model accuracy. The results indicate that the accuracy of the TVP model is different in different continents: the greatest Pearson’s correlation coefficient (R2) was 0.75 in Eurasia (RMSE = 1.16) and South America (RMSE = 1.17); the lowest R2 was 0.57 in Australia (RMSE = 0.73). Compared with the TANSO-observed CO2 concentration (XCO2), we found that the
accuracy throughout the World is between −2.56~3.14 ppm. Potential sources of TVP model uncertainties were also analyzed and identified.

**Keywords:** MODIS; CO₂ concentration; GOSAT TANSO; LST; NDVI/EVI; LAI/FPAR; GPP/NPP

1. Introduction

Climate change is one of the great challenges of the 21st century [1]; the average surface temperature has increased by 0.74 °C over the past 100 years (1906~2005). Climate exerts the dominant influence on the spatial distribution of the major vegetation types on a global scale. In turn, vegetation cover affects the climate via the alteration of the physical characteristics of the land surface, including albedo, roughness, water conductivity and atmospheric gas composition [2]. The increasing concentration of greenhouse gases (GHG) in the atmosphere has been verified as the most important cause of global warming, which is a major environmental concern and a prominent research topic [3–5]. According to the World Data Center for Greenhouse Gases [6], the average global carbon dioxide (CO₂) concentration in 2010 was 389.0 ppm, which is 11.9 ppm greater than in 2004; this figure has increased 39% from the pre-industrial global level of 280.0 ppm. Observations of GHG in the atmosphere are important for predicting global climate change. However, it is difficult to observe the global variation of GHG because the direct sampling of gases, especially in the upper atmosphere, requires great effort and cost [7].

The increased atmospheric CO₂ emissions from the burning of fossil fuels [8,9] and land-use changes [10–12] are such that terrestrial and marine environments are able to absorb approximately one-half to three-quarters of the emitted CO₂ [13,14]. To understand the relationship between land-cover changes and GHG concentration, great efforts have been made to measure the GHG emissions from forests and agricultural systems in recent years [11,15–17], and numerous data from field measurements and laboratory studies have been accumulated. However, global and regional GHG emissions estimates are still far from reliable because of the large spatial and temporal variations of the emissions records on which they are based.

There has been an intensive effort in recent years to measure and model the carbon cycles between the terrestrial biosphere and the atmosphere using a combination of modeling, remote sensing and observation data from eddy covariance flux towers. The eddy covariance technique offers an alternative for continuously and steadily assessing the long-term ecosystem carbon exchange. Furthermore, carbon budgets and the effects of environmental controls have been quantified with this technique for many forest types across the continent [11,17–21]. However, these carbon estimates from the eddy covariance technique only represent fluxes at the scale of the tower footprint, ranging from a hundred meters to several kilometers. To quantify the net exchange of CO₂ over regions or continents, flux tower measurements need to be up-scaled to large areas [17]. In [22] the DeNitrification and DeComposition (DNDC) model, which focused on N₂O and CO₂ emissions, was developed. However, this model was mainly based on the cropping practices and soil conditions in the U.S. and China and cannot be used directly for the global carbon cycle [9,15]. An international
network [23] has been formed to support such large-scale studies and provides standardized methods
to develop “optical scaling” [24]. Today, more than 300 surface observation stations contribute to the
WDCGG and World Meteorological Organization (WMO) by providing GHG data observed by flask
sampling or in situ, measured worldwide at the Earth’s surface, from towers, onboard aircraft in the
atmosphere and in oceanic environments [25,26]. Although ground-based measurements of GHG are
highly accurate, they are sparse.

Launching satellites to collect GHG data can solve this issue quite well. At present, the Thermal
And Near infrared Sensor for carbon Observation (TANSO) aboard the Greenhouse gases Observing
SATellite (GOSAT), launched in 2009 by Japan, is the first instrument dedicated to record CO₂ and
methane (CH₄) concentrations (denoted XCO₂ and XCH₄, in ppm. XCO₂ and XCH₄ are column-averaged dry air mole fractions of atmospheric CO₂ and CH₄) [8,27]. The SCAnning Imaging
Absorption spectrometer for Atmospheric CHartographY (SCIAMACHY) aboard the ENVironmental
SATellite (ENVISAT), launched in 2002 and lost in 2012 by the European Space Agency (ESA), can
also retrieve GHG concentrations, but it is not specially designed for GHG retrieval and is not highly
accurate; the single-measurement precision is approximately 1~2%, with an intermonth scatter of
2.3% [28,29]. Moreover, before 2002, the GHG concentrations could not be measured directly using
the remote sensing technique.

To obtain CO₂ concentration data using remote sensing data, [30] used CO₂ fluxes with static
chamber techniques and the normalized difference vegetation index (NDVI) data from a charged
coupled device camera to analyze the relationship between NDVI and CO₂ fluxes in arctic Alaska and
found that the NDVI is not sufficient to estimate the carbon flux rate. Using remote sensing data and
field data, [31] have developed models to map air-sea CO₂ fluxes. To our knowledge, there is no
model that can be used to retrieve CO₂ concentrations from remote sensing data at a regional or global
scale. Therefore, new approaches are critically needed to capture regional variations in CO₂
concentration and to bridge a major gap between field and satellite observations [32,33].

In the present study, we assessed CO₂ concentration using Moderate Resolution Imaging
Spectroradiometer (MODIS) products globally from June 2009 to November 2011. The objectives
were: (1) to examine the relationship between the MODIS-derived indices and TANSO XCO₂; (2) to
develop a monthly CO₂ concentration model for Africa, Australia, Eurasia, North America and South
America; and (3) to analyze the model uncertainties. This study explored the remote sensing potential
for monitoring CO₂ concentration at regional and global scales.

2. Study Area and Period

Because of the different climates, vegetation covers and human activity intensity in various regions
of the World, it is difficult to derive CO₂ concentrations using a single metric. To increase the derived
accuracy of the CO₂ concentration, we divided the World into five regions according to the continents:
Africa, Australia, Eurasia, North America and South America (Figure 1). For each continent, we used
different variables to assess CO₂ concentrations. The annual NDVI in 2011 was 0.33, 0.36, 0.33, 0.41
and 0.59 for Africa, Australia, Eurasia, North America and South America, respectively. According to
the ESA 2009 GlobCover data [34], which has a 300-m resolution, the bare area (the sum of sparse
vegetation and bare area) rate in Africa, Australia, Eurasia, North America and South America is
36.77%, 41.79%, 31.20%, 10.9% and 15.3% respectively. The World’s largest desert, the Sahara, is located in northern Africa, leading to a lower vegetation cover in northern Africa. Australia is the driest continent in the World [35] and from Figure 1 we can see that the vegetation cover in all of Australia is very low. In Eurasia, the bare area (including the Arabian Desert and the Gobi Desert) is mainly distributed in western and central Asia. The Amazon River Basin in South America is mainly covered by evergreen tropical rainforests, which have a higher vegetation cover [36]. The total forest area in South America is approximately $9.5 \times 10^6$ km$^2$ [37]. Moreover, South America is located near the Equator, and the higher annual temperatures and precipitation contribute to the higher NDVI in this region.

Figure 1. Study area. This is the 2009 global land-cover data that were downloaded from the ESA. The boundary lines of each continent are clearly shown. This figure also shows that in northern Africa, western Eurasia and most of Australia, the bare area rates are very high. Northern and southeastern North America, northern South America, central Africa and northern Eurasia have higher vegetation cover. Because of the smaller area, the study areas did not include some countries in Southeast Asia.

This study was conducted for data from June 2009 to November 2011 because the GOSAT TANSO data from April 2009 to November 2011 were available, but no such data exists for May 2009. The MODIS products for this time period were all available.

3. Data and Methods

3.1. MODIS Products Used in This Study

Optical remote sensing systems measure the surface reflectance, the fraction of solar energy that is reflected by the Earth’s surface. For a given wavelength, different land cover types may have different reflectance [16]. As a part of the National Aeronautics and Space Administration (NASA)-centered
international Earth Observing System (EOS), two MODIS instruments (Terra and Aqua) have been launched (in 1999 and 2002, respectively) to provide information for global studies of atmospheric, land and ocean processes. The MODIS instrument strengths include its global coverage, high radiometric resolution and dynamic ranges and accurate calibration in the visible, near-infrared and thermal infrared bands [38].

MODIS, which is a key instrument aboard the Terra (AM) and Aqua (PM) satellites is one of the most reliable data sources on the global scale [21]. Terra and Aqua MODIS can view the entire Earth’s surface every 1 to 2 days, acquiring data in 36 spectral bands ranging in wavelength from 0.4 µm to 14.4 µm. The MODIS data have been used to generate science data products for more than eleven years and these science data products are available to the public free of charge. More than 10 types of products are retrieved from the 36 MODIS bands. In this study, we selected four product types, MOD11C3, MOD13C2, MOD15A2 and MOD17A2, to develop a CO₂ concentration prediction model.

3.1.1. MOD11C3 LST

The land surface temperature (LST) is retrieved from MODIS thermal bands data (bands 31 and 32) over the entire land surface of the Earth. The required surface emissivity in bands 31 and 32 are estimated from land cover types [39]. LST is derived from the energy balance at the soil-atmosphere interface and is therefore a crucial parameter for the environmental energy budget [39,40]. The atmospheric effects are corrected using the split-window algorithm, considering that the signal difference in the two TIR bands is caused by the atmospheric differential radiation absorption [41].

LST is a good indicator of the ground surface energy balance, as it is one of the key parameters in the physics of land-surface processes at regional and global scales [42]. LST is the result of the surface-atmosphere interactions and energy fluxes between the atmosphere and ground surface; therefore, it is widely used in climate, hydrologic, ecological and biogeochemical studies [11,41,43–46].

The MODIS LST accuracy is within 1 K in the range from 263 to 322 K under clear-sky conditions [47,48]. MODIS/Terra LST Monthly L3 Version 041 with a resolution of 0.05 degrees (MOD11C3) was used in this study.

3.1.2. MOD13C2 NDVI/EVI

The Vegetation Index (VI) has become a common tool for assessing the different aspects of plant processes while simultaneously determining spatial variations in vegetation cover. The VI utilizes the differential absorption properties of leaves in different spectrum ranges to measure the amount and health of green biomass [30]. Two types of MODIS VIs, the NDVI and the Enhanced Vegetation Index (EVI), are globally produced over land at 1-km resolutions at 16-day compositing periods. Whereas the NDVI is chlorophyll sensitive, the EVI is more responsive to canopy structural variations, including canopy type, plant physiognomy and canopy architecture [49].

The NDVI has proven to be a valuable tool in researching large-scale changes in plant or ecosystem processes associated with global change. However, the actual range of the NDVI varies depending on the instrument used, background reflectance and canopy structure. Beyond a certain canopy density, increasing amounts of green biomass make little difference in the NDVI [50]. The EVI is more functional for Near Infrared (NIR) reflectance than for red absorption; therefore, it does not become
saturated as rapidly as the NDVI in dense vegetation. Additionally, the EVI has been shown to be highly correlated with photosynthesis and plant transpiration [51]. The equations used in the NDVI and EVI are shown below (Equations 1 and 2):

\[
NDVI = \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + \rho_{\text{red}}}
\]

\[
EVI = 2.5 \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + (6\rho_{\text{red}} - 7.5\rho_{\text{blue}}) + 1}
\]

where \(\rho_{\text{nir}}, \rho_{\text{red}},\) and \(\rho_{\text{blue}}\) are the spectral reflectance in MODIS bands 2, 1 and 3, respectively [49,52].

The NDVI accuracy is within ±0.025, while that of the EVI is within ±0.015 and the accuracy of retrieving VI for a good quality day would be to within ±0.020 for NDVI and ±0.010 for EVI [53]. MODIS/Terra Vegetation Indices Monthly L3 Version 005 with a spatial resolution of 0.05 degrees (MOD13C2) were used in this study.

3.1.3. MOD15A2 LAI/FPAR

Earth observations by the MODIS instruments on the Terra and Aqua satellites are being used to derive global vegetation properties such as the leaf area index (LAI) and absorbed fraction of photosynthetically active radiation (FPAR) at a 1-km pixel resolution and are updated once each 8-day period throughout each calendar year [24,47].

LAI is a biophysical variable influencing transpiration and photosynthesis at the plant canopy scale and carbon, water and energy dynamics at the regional scale [54]. The LAI is useful in measuring carbon flux because of its strong influence on canopy energy balance and rates of gas exchange [55]. It has been suggested, based primarily on theoretical studies, that canopy reflectance data provide more reliable indicators of canopy photosynthetically active radiation (PAR) interception and potential photosynthesis rates than do purely structural parameters such as the LAI and biomass [56].

The computation of the LAI from satellite data is based on a known relationship between the fractions of incident light transmitted through a canopy and the light extinction coefficient (k) (Equation 3):

\[
LAI = \frac{-\ln(1 - FVC)}{k(\theta)}
\]

where the FVC (fractional vegetation cover) is defined as the vegetation-covered fraction of the ground viewed in the nadir direction and \(k(\theta)\) is the light extinction coefficient for a solar zenith angle \(\theta\). For a detailed description of LAI, please see the paper by [57].

Converting LAI to FPAR uses a simple Beer’s Law approach, which enables the computation of FPAR as a function of LAI and the canopy light extinction coefficient (Equation 4):

\[
FPAR = 0.95(1 - e^{-kLAI})
\]

The value of the extinction coefficient \(k\) is 0.5 [58].

The accuracy of LAI is 0.66 LAI units RMSE when all biomes are taken in account. If broadleaf forests are excluded, the accuracy in all the other biomes is 0.5 LAI units RMSE. The accuracy is
0.12 FPAR units RMSE [53]. MODIS/Terra LAI/FPAR 8-Day L4 version 005 (MOD15A2) with a resolution of 1-km was used in this study.

3.1.4. MOD17A2 GPP/NPP

The gross primary production (GPP) represents the capacity of the plants in an ecosystem to capture energy and carbon. The net primary production (NPP), the reduction of GPP after autotrophic respiration, is the net carbon stored as new plant material in an ecosystem, which supplies humans with various foods, fuels, fibers and construction materials [21]. Despite the strong theoretical and empirical relationship between surface reflectance and the FPAR, accurate estimates of vegetative productivity (GPP, NPP) depend strongly on the quality of the radiation inputs [59].

Global productivity can be estimated by combining remote sensing with carbon cycle processing [59]. The MODIS dataset provides the first operational, near-real-time calculation of the global GPP and NPP at a 250-m spatial resolution with daily coverage [10]. The MODIS GPP and NPP products have been validated as being able to capture spatial and temporal GPP and NPP patterns across various biomes and climate regimes, and the products are consistent with the ground flux tower-based GPP and field-observed NPP estimations [21].

The NPP product is designed to provide an accurate and regular measure of the production activity or growth of terrestrial vegetation. These products will have both theoretical and practical utility. The theoretical use is primarily to define the seasonally dynamic terrestrial surface CO\textsubscript{2} balance for global carbon cycle studies [60,61]. Previous studies used tower and MODIS GPP/NPP data to model the carbon cycle in forests [11,17,50,62] and determined the relationships between GPP/NPP and LAI, NDVI and climate [10,62,63].

In addition, the NPP/GPP ratio (NG) [64] was found to be constant across a range of CO\textsubscript{2} levels and temperature for herbaceous and woody plants. An assessment of the global patterns in the NG can enhance knowledge about features of the NG and benefit research on global change [21]. In the present study, we also attempt to use the GPP minus the NPP (GN) as a variable of the CO\textsubscript{2} concentration model.

MODIS global annual estimates of GPP and NPP are within 10.4% and 9.0% [53]. MOD17A2 Monthly GPP/NPP with a resolution of 1 km, downloaded from the University of Montana [65], was used in the present work.

3.2. GOSAT TANSO Data

The GOSAT was launched on January 23, 2009 from Tanegashima Island, Japan, and is a joint project of the Japan Aerospace Exploration Agency (JAXA), the Ministry of the Environment of Japan (MOE) and the National Institute for Environmental Studies (NIES) of Japan. The primary purpose of the GOSAT project is to accurately estimate the emissions and absorptions of GHG on a subcontinental scale to assist environmental administrations in evaluating the carbon balance of land-based ecosystems and to provide assessments of regional emissions and absorptions [66]. The GOSAT main sensor (TANSO) is a nadir-looking Fourier Transform Spectrometer (FTS) with four spectral bands, from the visible to thermal infrared. GOSAT is the World’s first spacecraft designed to measure XCO\textsubscript{2} and XCH\textsubscript{4}, the two major GHG concentrations, from SWIR bands with global coverage every three days [67]. GOSAT is a sun-synchronous orbit with a local overpass time of 13:00 [68,69]. It derives
the CO₂ and CH₄ column amounts by observing ground-surface scattered solar spectra in the 1.6 and 2.0 µm regions, which contain suitable CO₂ and CH₄ absorption bands, and the satellite also measures vertical profiles higher than approximately 2 km by measuring thermal emission spectra from the atmosphere [8].

TANSO FTS SWIR Level 2 Version 02.00 and 02.11 (stored column abundances of CO₂ and CH₄ retrieved from the radiance spectra in bands 1 through 3 of FTS) data from June 2009 to November 2011 were used in this study. The relative accuracy of TANSO FTS Level 2 data is 0.3–1% (1–4 ppm) for CO₂[70]. The GOSAT FTS SWIR Level 2 data are point data and the footprint is a circle with 10.5 km diameter at nadir [71]; thus, we obtained the monthly CO₂ concentration easily. Moreover, in this study, we assumed that the TANSO XCO₂ were the true value without bias. TANSO XCO₂ data were also used to check the accuracy of the modeled CO₂ concentration from MODIS derived products.

3.3. Data Processes

MOD11C3, MOD13C2 and MOD15A2 were downloaded from the NASA homepage [74], and MOD17A2 was downloaded from the University of Montana [65]. Regarding the spatial resolution of these MODIS products, MOD11C3 and MOD13C3 have 0.05-degree spatial resolutions; MOD15A2 and MOD17A2 have 1-km resolutions. All of these products are sinusoidally projected, and all data are monthly except MOD15A2 (8-day). To use these data in one model, we must unify them into the same projection, the same spatial resolution and the same temporal resolution.

First, all data were reprojected and masked from a sinusoidal projection to the Geographic WGS84 projection as a 0.05-degree pixels using the MODIS Reprojection Tool [75]. Then, MOD15A2 LAI/FPAR was monthly stacked. We employed the maximum value composite (MVC) method that selected the highest pixel value to represent the composting period [17,49].

The spatial variability of GPP and NPP over the global is enormous, ranging from approximately 1,000 g C m⁻² for evergreen tropical rain forests to less than 30 g C m⁻² for deserts. For bare areas such as the Sahara Desert in North Africa and deserts in central-eastern Asia and western China, there are no calculated GPP data, and the values were defined as 0 in the present study.

For the TANSO CO₂ data, we averaged all of the TANSO FTS Level 2 points in each month for each continent as the monthly value. From June 2009 to November 2011, there were 66,126 CO₂ points in Africa, and the point numbers were 21,088, 37,651, 11,796 and 10,504 in Australia, Eurasia, North America and South America, respectively.

3.4. Linear Regression Model

The correlation coefficient (R), also called the Pearson correlation coefficient, provides an index of the degree of correlation between two datasets. In the present work, R² was used to indicate the strength of the correlation between the modeled data set and the observations. The present work also employed the root mean square error (RMSE) to analyze the error between predictions and observations. The RMSE was used to evaluate the error of the predicted and the uncertainty relative to the TANSO-observed XCO₂. The RMSE and R² provide information on the regional performance and uncertainty of the modeled CO₂ concentration.
4. Results

4.1. Seasonal Variations of the MODIS-Derived Indices on Each Continent

MODIS-derived indices have different variations on different continents (Figure 2). In Eurasia and North America, all indices (LST, NPP, NDVI and LAI) change seasonally, with maximum values in summer (July or August) and minimum values in winter (January or February). The LST in Australia has the opposite trend from that found in Eurasia and North America, but the seasonal change is not as obvious. The fluctuation of NPP, NDVI and LAI are quite small in Africa, Australia and South America. South America always has the largest value of NPP, NDVI and LAI.

Figure 2. MODIS-derived indices trend of LST, NPP, NDVI and LAI from June 2009 to November 2011 in Africa, Australia, Eurasia, North America and South America.
Eurasia and North America are located in Northern Hemisphere and the seasonal sensitivity of plant green-up and snow events are the main reasons of the changes in the indices. Australia is located in Southern Hemisphere, with a large percentage of bare areas leads to the small fluctuation of NPP, NDVI and LAI. The Amazon River Basin in South America has the highest vegetation cover in the world; thus, the NPP, NDVI and LAI are all higher in South America. The LST is closely associated with land-cover type, especially in bare areas [39], which is the main reason for higher value in of this parameter in Africa and Australia.

4.2. Correlation Analysis between TANSO XCO₂ and MODIS-Derived Indices

The R² and RMSE between TANSO XCO₂ and MODIS-derived indices were analyzed across months on each continent. The following nine indices were considered in the analysis: LST (of MOD11C3), EVI and NDVI (of MOD13C2), LAI and FPAR (of MOD15A2), GPP, NPP, GN (GPP minus NPP) and NG (NPP/GPP) (of MOD17A2).

Different correlation coefficients between TANSO XCO₂ and MODIS-based indices were found on different continents (Table 1). In Africa, the highest correlation coefficients were between XCO₂ and NPP (R² = 0.61 and RMSE = 1.21). Although the values between XCO₂ and NDVI and LST were not very high (R² = 0.39 and 0.24, RMSE = 1.53 and 1.70, respectively), but they meet the model indices select standard (see Section 4.3). In Australia, because the higher rate of bare areas can increase the reflectance of background, the highest correlation coefficient was between XCO₂ and LST (R² = 0.41, RMSE = 0.85). Although the bare areas in Eurasia are also extensive, the forests in Siberia and southern China increase the correlation between XCO₂ and NDVI and EVI (R² = 0.44 and 0.42, RMSE = 1.73 and 1.76, respectively). In North America, the correlation coefficients between XCO₂ and NDVI reached R² = 0.65 (RMSE = 1.15) and the lowest values that were selected for the TVP model is R² = 0.49 (between XCO₂ and LST and NDVI). In South America, the highest correlation coefficients were between XCO₂ and LST (R² = 0.64, RMSE = 1.39). The index of NG (R² = 0.19, RMSE = 2.09) were also selected for the model build.
**Table 1.** Pearson’s correlation coefficient ($R^2$) and the root mean square error (RMSE) between TANSO XCO$_2$ and MODIS-derived indices. Gray masked indices are selected for the TVP model.

<table>
<thead>
<tr>
<th></th>
<th>LST</th>
<th>EVI</th>
<th>GPP</th>
<th>NPP</th>
<th>NG</th>
<th>GN</th>
<th>NDVI</th>
<th>FPAR</th>
<th>LAI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Africa</strong></td>
<td>R$^2$</td>
<td>0.24</td>
<td>0.09</td>
<td>0.47</td>
<td>0.61</td>
<td>0.40</td>
<td>0.01</td>
<td>0.39</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.007</td>
<td>0.104</td>
<td>0.000</td>
<td>0.000</td>
<td>0.576</td>
<td>0.000</td>
<td>0.008</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.70</td>
<td>1.86</td>
<td>1.42</td>
<td>1.21</td>
<td>1.50</td>
<td>1.94</td>
<td>1.53</td>
<td>1.72</td>
</tr>
<tr>
<td><strong>Eurasian</strong></td>
<td>R$^2$</td>
<td>0.22</td>
<td>0.42</td>
<td>0.31</td>
<td>0.26</td>
<td>0.01</td>
<td>0.35</td>
<td>0.44</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.009</td>
<td>0.000</td>
<td>0.002</td>
<td>0.004</td>
<td>0.543</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>2.04</td>
<td>1.76</td>
<td>1.92</td>
<td>1.98</td>
<td>2.29</td>
<td>1.87</td>
<td>1.73</td>
<td>1.86</td>
</tr>
<tr>
<td><strong>North America</strong></td>
<td>R$^2$</td>
<td>0.49</td>
<td>0.59</td>
<td>0.51</td>
<td>0.49</td>
<td>0.00</td>
<td>0.45</td>
<td>0.65</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.796</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.37</td>
<td>1.23</td>
<td>1.35</td>
<td>1.37</td>
<td>1.92</td>
<td>1.42</td>
<td>1.15</td>
<td>1.23</td>
</tr>
<tr>
<td><strong>Oceania</strong></td>
<td>R$^2$</td>
<td>0.41</td>
<td>0.13</td>
<td>0.00</td>
<td>0.16</td>
<td>0.35</td>
<td>0.20</td>
<td>0.38</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.00</td>
<td>0.05</td>
<td>0.99</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.85</td>
<td>1.04</td>
<td>1.11</td>
<td>1.02</td>
<td>0.90</td>
<td>1.00</td>
<td>0.88</td>
<td>1.05</td>
</tr>
<tr>
<td><strong>South America</strong></td>
<td>R$^2$</td>
<td>0.64</td>
<td>0.30</td>
<td>0.42</td>
<td>0.13</td>
<td>0.19</td>
<td>0.61</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.048</td>
<td>0.016</td>
<td>0.000</td>
<td>0.777</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.39</td>
<td>1.94</td>
<td>1.77</td>
<td>2.16</td>
<td>2.09</td>
<td>1.46</td>
<td>2.32</td>
<td>2.19</td>
</tr>
</tbody>
</table>

The different values indicate that CO$_2$ concentration was influenced by different factors associated with different climate and land-cover types [2]. Table 1 also shows that LST was important in all five continents, which indicates that LST is an important factor in CO$_2$ concentration research. With regard to the selected indices in the five continents, only Australia had an average RMSE lower than 1; Eurasia had the highest averaged RMSE (1.87).

### 4.3. Assessment of Modeled CO$_2$ Concentration

In this study, we used the MODIS products of temperature (MOD11C3), vegetation cover (MOD13C2 and MOD15A2) and productivity (MOD17A2) as variables to measure land surface CO$_2$ concentration; thus, we named the model TVP. To select the indices of the TVP model, first, we calculated the correlation coefficient between each of these indices and the TANSO XCO$_2$ on each continent (Table 1). Then, we selected indices according to higher $R^2$, lower RMSE and significance value (P); if P was less than 0.05 (significance greater than 95%) and the index increased the correlation coefficient in the multivariate linear regression of the TVP model, then the index was selected. Finally, we selected seven indices for Eurasia and North America, six indices for South America and five indices for Africa and Australia. The results are shown in Table 1. EVI and NDVI, GPP and NPP and LAI and FPAR are pairs of indices that measure the same parameter and can, in some cases, be substituted for one another. Researchers found that a small number of indices is not sufficient to provide an accurate estimate of CO$_2$ flux [30]; thus, we also tested the use of a single
index for a given parameter (e.g., NPP but not GPP in Africa and NDVI but not EVI in Eurasia), but we found that the adjusted $R^2$ value was lower than that of the TVP model. At last, the index number that been selected for TVP model is 5 in Africa and Australia, 7 in Eurasia and North America and 6 in South America. After analyzing the $R^2$ and RMSE between MODIS products and the TANSO XCO$_2$, we employed MODIS-derived indices that were selected for each continent as independent variables and TANSO XCO$_2$ as dependent variables of multivariate linear regression in Microsoft Office 2010 to model the CO$_2$ concentration on each continent. The TVP models for each continent are described below:

**Africa:**

$$CO_2 = 396.06 - 0.06 \times LST + 89.49 \times NDVI - 0.13 \times GPP + 0.23 \times NPP - 49.18 \times NG$$  (5)

**Australia:**

$$CO_2 = 362.73 + 0.07 \times LST + 29.02 \times NDVI + 0.03 \times NPP - 23.90 \times NG - 0.07 \times GN$$  (6)

**Eurasia:**

$$CO_2 = 277.93 + 0.40 \times LST + 95.04 \times EVI - 64.32 \times NDVI - 0.89 \times FPAR + 2.73 \times LAI - 0.03 \times GPP - 0.004 \times GN$$  (7)

**North America:**

$$CO_2 = 321.04 + 0.23 \times LST + 48.44 \times EVI - 27.26 \times NDVI - 0.52 \times FPAR + 1.32 \times LAI - 0.03 \times GPP - 0.01 \times GN$$  (8)

**South America:**

$$CO_2 = 634.59 - 0.49 \times LST - 9.49 \times EVI - 0.66 \times LAI + 0.11 \times GPP - 166.04 \times NG - 0.27 \times GN$$  (9)

TVP models were run on a monthly time scale using the data of LST, EVI, NDVI, LAI, FPAR, GPP, NPP, GN and NG to predict CO$_2$ concentrations for the five continents. The TVP model performance was then evaluated using scatter plots (Figure 3) of the predicted versus TANSO-observed CO$_2$ concentrations from June 2009 to November 2011.

Figure 3 shows that the correlation coefficient was not very high ($R^2 = 0.57, \text{RMSE} = 0.73$) in Australia. The TVP model provided better estimates of CO$_2$ concentration for Eurasia and South America ($R^2 = 0.75$ for both, RMSE = 1.16 and 1.17, respectively). The validation points of the TVP-modeled and TANSO-observed XCO$_2$ of Africa and North America were distributed closely around the 1:1 line (RMSE = 0.98 and 1.02, respectively). Satisfactory correlation coefficients were found in both Africa and North America ($R^2 = 0.69$ and 0.72, respectively). Thus, the results were very encouraging on all five continents. Notably, the higher $R^2$ values for the TVP-modeled CO$_2$ concentration in North America and South America do not necessarily indicate that the accuracy was higher for these continents. We found that in North and South America, the TANSO CO$_2$ points alone were inadequate (11,796 and 10,504, respectively) because the TANSO sensor is only able collect points under a clear sky (without clouds) [28]. As mentioned in Section 3.2 that the accuracy of TANSO XCO$_2$ is about 0.3~1% and it must increasing the uncertainty of the correlations between predicted and observed CO$_2$ concentrations if we presume TANSO XCO$_2$ as true value. Moreover, in
different continents the bias of TANSO XCO$_2$ is different [28,72,73,76], which can also increase the uncertainly of Figure 3. These uncertainties have not necessarily weakened this study, in terms of achieving the targeted objectives.

**Figure 3.** Comparison of TVP modeled with TANSO-observed CO$_2$ concentration in Africa (●), Australia (●), Eurasia (●), North America (●) and South America (●). The dotted blue line is the 1:1 line between TVP derived and TANSO observed XCO$_2$. The black line is the regression line.
Figure 3 also shows that the slope of predicted versus observed CO$_2$ concentration was significantly lower than the 1:1 line, suggesting that the TVP algorithm somewhat underestimates high CO$_2$ concentration and overestimates low CO$_2$ concentration. However, the over/underestimation of extreme values had only a small impact on the predicted value, with the largest RMSE at 1.17 in South America and the lowest at 0.73 in Australia.

5. Discussion

5.1. Seasonal Variation and the Differences between TVP-Modeled and TANSO-Observed CO$_2$ Concentration

Figure 4 shows the seasonal variation of CO$_2$ concentration and the differences between TVP-modeled and TANSO-observed CO$_2$ concentration. The TANSO XCO$_2$ time series showed a more pronounced seasonal variation and the extreme values appeared in different months on different continents. In Africa, the observed CO$_2$ concentration ranged between a maximum of 384.13 ppm and a minimum of 376.00 ppm between June 2009 and November 2011. The maximum and minimum appeared in March and September, respectively. The CO$_2$ concentration estimated by TVP model were overestimated before September 2010 and underestimated in 2011. The largest difference between the modeled and observed CO$_2$ concentration was $-2.43$ ppm, which occurred in September 2009. In Australia and South America, the maximum value of CO$_2$ concentration was usually in June or July and the minimum value appeared around December. Moreover, Figure 4 shows that the fluctuation of CO$_2$ concentration from June 2009 to December 2011 was very mild in Australia, where the difference between minimum and maximum value of TANSO XCO$_2$ was 4.17 ppm. This low fluctuation mainly results from the land-cover types in Australia, where bare areas covered 41.79%. The amount of CO$_2$ sequestered by desert soils is lower than by other soils [77]. Lack of plant photosynthetic is the main reason for the mildly seasonal change of CO$_2$ concentrations. The largest differences between TVP-modeled and TANSO-observed CO$_2$ concentration was 1.24 ppm in Australia and the average value is 0.62 and 0.95 ppm in Australia and South America, respectively. In Eurasia and North America, the minimum values of CO$_2$ concentration appeared in August and the maximum values appeared in April or March. The difference between minimum and maximum value of TANSO XCO$_2$ was 9.46 and 8.39 ppm in Eurasia and North America, respectively. In Eurasia, overestimating of CO$_2$ concentration by TVP model was found from June 2009 to April 2010 and underestimating appeared after December 2010. In North America, the overestimating happened from June 2009 to April 2010 (except December 2009) and from then on underestimating happened (except December 2010). The average difference between TANSO-observed and TVP-modeled CO$_2$ concentration was 0.97 and 0.84 ppm in Eurasia and North America, respectively. In Africa, the maximum overestimating of TVP model was 2.43 ppm in September 2009 and the maximum underestimating value was 1.86 ppm in February 2011.

Figure 4 also shows that in the study time scale, TANSO XCO$_2$ clearly increased in all the study areas. The maximum slope was in Eurasia (slope = 0.13) and the minimum slope was in Australia (slope = 0.09).
Figure 4. The seasonal variations of CO₂ concentrations on each continent and the differences of TVP-estimated CO₂ concentration compared with the TANSO-observed values. Hollow points are the TVP-modeled CO₂ concentration and solid points are the TANSO-observed values. The lines are the trend line of TANSO XCO₂.
5.2. Uncertainties and the Significances of the TVP Model

Some studies have used ground-based data to validate the CO₂ concentration estimated by TANSO and SCIAMACHY and found different results in different regions. TANSO XCO₂ with six ground-based Total Carbon Column Observing Network [78] sites around the world have been compared and found that the average bias of TANSO XCO₂ was 2.80 ppm [28]. Researchers also found that the computed errors of TANSO XCO₂ varied between 1.20 and 3.20 ppm [72]. With regard to the accuracy of SCIAMACHY CO₂ concentration, some researchers found that the global single-measurement precision amounted to 2.50 ppm [29,79]. In the present study, we used TANSO XCO₂ as the observed value to check the accuracy of TVP model and found that the accuracy in the study area was between −2.56~3.14 ppm. Australia exhibited the lowest difference, at −1.24~1.19 ppm, and the largest difference was in South America, at −2.09~3.14 ppm. It indicates that the accuracy of TVP model is similar with the SCIAMACHY and TANSO.

Despite the encouraging performance of the TVP model in estimating CO₂ concentration from MODIS products, these estimates still contain significant uncertainties and are a long way from reliably estimating CO₂ concentration from remote sensing data. Although random error is usually considered less serious than uncertainties [31], there can be many sources of error that create a high overall uncertainty. The CO₂ exchange between terrestrial ecosystems and the atmosphere has been mainly dominated by three factors: climate, land-cover type and human activities [4,80–82]. With regard to the TVP model in the present study, there are several sources of uncertainty associated with the CO₂ concentration estimates: uncertainty in the TANSO accuracy, uncertainty in land cover, uncertainty in the input data and uncertainty regarding the extent and nature of climate change in different regions. Thus, there remains a great deal of research is still required to ensure the accuracy of CO₂ concentration estimated based on remote sensing data.

Although the TVP model itself has some shortcomings but it is a meaningful explore in CO₂ calculate using remote sensing data. The same as used MODIS, we can also use NOAA AVHRR data combine with ground stations data to derive CO₂ and CH₄ concentrations at regional or global scale. Using the TVP model we can derive global CO₂ concentration from 1981 to 2002 when there was no GHG satellite. In fact, we can also use TVP model to identify the spatial distribution of GHG concentrations which is also an important issue nowadays. Moreover, in the future, with the launching
of other GHG observe satellites, e.g., Orbiting Carbon Observatory 2 (OCO-2), Carbon Monitoring Satellite (CarbonSat) and GOSAT-2, the accuracy of TVP model will increasing.

6. Summary and Conclusions

TANSO, onboard the GOSAT satellite launched on January 23, 2009, is the first and only orbiting greenhouse gas observing satellite. Although SCIAMACHY, onboard ENVISAT, has also been able to measure greenhouse gas concentrations since 2002, it is not specifically designed for studying GHG, and the data require calibration. Approximately 300 stations are far apart from accurately global GHG concentration measure. The present study demonstrates that MODIS derived indices have great potential for CO₂ concentration measurement on regional as well as global scales.

The assessment of CO₂ concentration from remote sensing data has been conceived and attempted in the present work. First, we examined the relationships between the MODIS derived indices and CO₂ concentration on each continent and found that CO₂ concentration is dominated by different factors in different continents. LST, NDVI and GPP are the main factors that affect the CO₂ concentration and their relative contributions are different in different areas. Second, we developed a TVP model for CO₂ concentration measurement based on the relationships between the MODIS derived indices and CO₂ concentration and got satisfactory results. The accuracy of TVP model is different on different continents: the highest correlation coefficient between the modeled and observed value is in Eurasia and South America (R² = 0.75, RMSE = 1.16 and 1.17, respectively); the lowest value is found in Australia (RMSE = 0.73). Compared with the TANSO XCO₂, we found that the accuracy across the study area is between −2.56~3.14 ppm. From the results we also found that the TVP model underestimates high CO₂ concentration and overestimates low values.

Increasing the accuracy of CO₂ concentration that derived from remote sensing data remains beyond current capabilities. The TVP model that proposed in this study also contains some uncertainties (described in Section 5.3). With the releasing of CO₂ concentration data by GOSAT project, validated data at different time scales should be used in future studies.

Acknowledgments

This study was supported by Scientific Research (B), 23405005 (PI: Xiufeng Wang, Hokkaido University, Japan) and National Natural Science Foundation of China (41201159, Study on the effect mechanism of commercial center pattern on traffic carbon emissions in Shenyang city. PI: Jing Li, Chinese Academy of Sciences, China). We also thank the GOSAT Project of Japan, NASA and ESA for providing data that were used in this study.

References


34. ESA. Available online: http://ionia1.esrin.esa.int/ (accessed on 15 October 2012).


74. ECHO. Available online: http://reverb.echo.nasa.gov/reverb/ (accessed on 15 October 2012).


78. TCCON. Available online: http://tccon.caltech.edu/ (accessed on 15 October 2012).


© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).