Main questions and comments

There was some disagreement in the general appraisal of your manuscript in the referees' reports: Anonymous Referee #1 was overall positive about the manuscript and recommended publication after minor revision only, whereas Anonymous Referee #2 was more critical and asked for major revisions, although (s)he did not want to review the revised paper again.

I therefore decided to give the revised paper a more in-depth review by myself. I find that the criticisms of Anonymous Referee #1 have been addressed quite comprehensively in the revised manuscript, except for a few isolated points. This is, unfortunately not the case regarding some important points raised by Anonymous Referee #2, that are well discussed in the online replies, but that must then also be discussed in the manuscript.

Please find below a list of points that I would like to ask you to address in the next revised manuscript, that I invite you to submit. Some of them are minor points, but clarifying those right now will make readers' lives easier, as one will not have to go through the whole discussion to gather some important information that is missing in the paper or that remains unclear. I further add a few technical issues to address. Page numbers refer to the most recently uploaded manuscript version, except where explicitly stated otherwise. Please address all of the points mentioned below.

Thank you very much for your time and comments on this work. We have modified the manuscript in the new version by considering all the referees' comments and your points. We also added the missing part in the text. All the points you mentioned have been addressed in the following.

1. Do the fossil fuel CO2 emission data have some time resolution?

The current manuscript seems to say NO (p. 7, ll. 12-13: "There is no time structure in the fossil fuel emissions."), although I then do not see any need to cycle them repeatedly (p. 7, ll. 10-11). The FFDAS web site (http://hpcg.purdue.edu/FFDAS), on the other hand, indicates that their data are on a "0.1° x 0.1°, hourly grid." Please clarify this, specify if and how the FFCO2 data have been post-processed before use in the model (especially regarding the time dimension), the more since there is later a comparison with station data at hourly resolution (section 3.3). The text really must be unambiguous regarding this point.

The FFDAS fossil fuel CO_2 emission data have sub-annual time resolution.

The FFDAS emission products have different spatial/temporal resolutions. The FFDAS dataset used in this study are at 0.1degree/annual resolution. The product at 0.1degree/hourly resolution is obtained by multiplying the annual emission by the hourly fraction. This information is not included in the statement on the FFDAS website yet, since the website is still under construction. Given the fact that this study aims to quantify the impact of regridding approach by comparing simulated CO_2 concentrations not only at spatial distribution but their temporal variations, we chose to the FFDAS product with no time structure. The annual amount in each grid cell is divided by 2920 to obtain evenly distributed emissions at three-hour model resolution.

To make the point clear, we firstly added one sentence emphasizing this point in introduction section: "The bias is defined as spatial distribution and temporal variations of the simulated CO_2 concentration difference driven with two regridding fossil fuel emission inventories."

Then, in Methods section, we added sentences clarifying the resolution of the FFDAS emission data used in this study:

"FFDAS version 2.0 estimates fossil fuel emissions originally at 0.1˚/annual resolution over the globe. Based on the annual resolution, higher temporal resolution products at hourly scale are constructed by multiplying sub-annual fraction by annual total at each grid. Considering the goal of this study, we choose the product at 0.1˚/annual resolution to construct the fossil fuel emissions with no time structure. The annual amount in each grid cell is divided by 2920 to obtain evenly distributed emissions at three-hour model resolution."

2. Anonymous Referee #2 asked about potential regridding issues related to possibly different spatial resolutions of MERRA and PCTM. While this is clarified in the online Author's Comment, that information (and a short statement why this is important) is still missing in the revised manuscript. Please amend.

We have added statement clarifying this issue in Method section:

"The initial data product of GEOS-5 is at 0.7˚ longitude x 0.5˚ latitude with 72 hybrid vertical levels. Two coarser MERRA products are also produced by aggregating the high-resolution product to a resolution at 1.25˚ longitude x 1.25˚ latitude or 1.25˚ longitude x 1˚ latitude with 72 hybrid vertical levels (Rienecker et al., 2011; Reichle et al., 2011; Reichle et al., 2012). In atmospheric transport simulation and inversion system, a dynamical consistence problem might be introduced if the driving meteorology data doesn't match the transport model grid. However, this problem doesn't exist in this study, since the MERRA product used in this study is on the same grid as PCTM."

3. Anonymous Referee #1 pointed out that the Dutch city of Groningen is not a coastal city (p. 9, l. 21). I agree with this observation as Groningen is indeed located somewhat inland and I find that the reply given regarding this point is not clear.

Sorry that we forgot to correct the mistake in previous submitted version. We have removed "Groningen" from the sentence in the text. The modified sentence is: "A particularly notable example is the surface CO_2 concentration difference downwind of the cluster of large coastal western European cities, for example, London, Rotterdam, Barcelona and Rome."

4. The description of the characteristics of the TAP site is not entirely clear. It would first of all benefit from some streamlining: it is stated right at the beginning that TAP is in close proximity to a heavily populated coastal area (p. 11, l. 3); later on the same page, the location information is partly repeated, and completed. Why not give all of the relevant information right away (situation, names of nearby urban areas. The text indicates that the station is situated 300m offshore in the Tae-ahn Peninsula. I find this contradicting: either it is offshore, or it is in the Peninsula (i.e., on land). The information that I have been able to find rather indicates that it is on land (the site is of the type "vegetated peninsula") which would suggest to me that it is a land station. Please check this and correct if necessary. It would furthermore be helpful to state that TAP is located in the Republic of Korea.

Thanks for pointing out the mistake. We also double-checked the location and confirm that the TAP site locates on land of the Peninsula. We have corrected the statement.

We have put all relevant information together and modified the statement:

"TAP is a coastal station (36˚43′N, 126˚07′E) locating at the Tae-ahn Peninsula in the Republic of Korea. This site is in close proximity to two urban cities, Seosan and Taean. TAP is designated to ocean grid cell in PCTM resolution. The emissions on this grid are aggregated to adjacent land grid after regridding, and thus is located in the negative portion of the emission dipole"

5. The points raised by Anonymous Referee #2 in the two last paragraphs of her/his review (p. C1121 therein) are well addressed in the online Author's Comment. The points raised by Anonymous Referee #2 are, however, very important and also need to be discussed in the manuscript. Please complete the discussion in the manuscript on the basis of your reply.

Regarding the point raised by Referee #2 in the second last paragraph, we add the statement in two paragraphs of section 3.4:

"Atmospheric CO_2 inversion studies are also a good example of research that must overcome this potential problem. However, we don't consider the impact and uncertainty on atmospheric inversion in this study, since atmospheric inversions are not the only purpose for simulations of fossil fuel-like tracers. Many studies in atmospheric chemistry have the same need and consequently the same problem. But the study also does do something of direct use for an inversion. The fossil fuel is part of the prior flux. So in an atmospheric inversion this term represents a systematic uncertainty in the mapping of fossil fuel flux into the prior mismatches (prior simulation of concentration – observations). It can be seen that the effect is widespread and large compared to the measurement uncertainty usually used. Thus, this is enough to demonstrate significance for an inversion", and

"It should be noted that the reshuffling simply might transfer errors from one place to another. For example reshuffling emissions away from an oceanic gridpoint may leave a station in that grid cell further from emissions than it really should be. This is possible of course. This can only been investigated by separating the transport and relocation effects by using an on-line model. However, it is expected that this shuffling method could introduce land-ocean biases, since fixed fossil sources are almost entirely land-based and putting them in ocean gridpoints seems far more likely to introduce land-ocean biases as the inversion tries to correct a poorly transported signal from the wrong environment. In general, without further research testing the sensitivity of results to this technique, it is unclear to what extent this minimizes the fossil fuel CO_2 emissions regridding problem discussed in this study."

For the point raised by Referee #2 in the last paragraph in the comments, we added the statement in the paragraph in section 3.4 :

"It also should be pointed out that the fossil fuel emissions from planes and ships are not included in this study. Airborne emissions are unlikely to be strongly impacted by this problem since the differences in atmospheric physics between land and ocean decrease once above the boundary layer. While emissions from shipping do potentially suffer from this problem the fraction subject to misallocation will be small so the total problem is a small fraction of a small fraction. "

Technical details:

The Supplement currently only includes a single figure. Are there any compelling reasons for not including that figure directly into the main manuscript? Not mandatory, but, if we can simplify readers' lives…

We agree that it is more appropriate to include the figure in the main manuscript. We have put this figure in the Method section, which is specified as Fig. 1.

Throughout the paper: the correct SI abbreviation of kilogram is "kg", with a small "k" (not "Kg").

We have corrected the abbreviation by replacing “Kg” with “kg”.

Page 3, line 19: Reference Peylin et al. (2013) is still incorrect or missing.

This publication was already included in the reference list. We had made a mistake in the publishing year when citing this reference. We have corrected it.

Page 6, lines 13-14: Reference Gurney and Coltin (2014) is missing.

This work is still in preparation. We removed this reference from the text.

Page 7, line 21: please add "(longitude x latitude)" after "1.25° x 1.0°" at the end of the line.

We have added it.

Page 12, line 2: "largest value by about" should probably read "largest value of about".

We have modified it in the text.

Page 12, lines 7-9: "High-frequency signals are also shown through the year in the hourly timeseries for the TAP site, indicating the impact of synoptic-scale atmospheric transport." This sentence looks strange to me. Should it not better read "The hourly timeseries for the TAP site also contains high-frequency components throughout the year, indicating the impact of synoptic-scale atmospheric transport." or "... also shows high-frequency behavior throughout ...".

We have modified the sentence as: “The hourly timeseries of the TAP site also shows high-frequency behavior throughout the year, …”

Page 20, Figure 2: the scale annotations are barely readable.

We enlarged the scale to enable the figure to be readable.

Page 22, Figure 4: the annotations are barely readable. Please enlarge. I suggest to arrange the panels vertically instead of horizontally.

Thanks for suggestion. We enlarged the annotations and showed the panels vertically.

Page 23: the blue stars indicating the location of TAP are not well visible, especially on the left-hand panel. I suggest to draw it in white. The annotations of the scales are barely readable; please enlarge.

We drew the star in white. We also enlarged the scales to make it readable.

Non-public comments to the Author:
It would have been helpful to upload the revised manuscript first intrack-change mode; needless to say that it was quite tedious and time-consuming to make out where exactly the text was changed and where not.

Could you please upload the next revision in track-change mode (starting from the currently uploaded version)? Thank you.

Sensitivity of simulated CO_2 concentration to regridding of global fossil fuel CO_2 emissions

Xia Zhang[1], Kevin Robert Gurney[1,2], Peter Rayner[3], Yuping Liu[4], Salvi Asefi-Najafabady[1]

1. School of Life Sciences, Arizona State University, Tempe, AZ 85287

2. Global Institute of Sustainability, Arizona State University, Tempe, AZ 85287

3. School of Earth Sciences, University of Melbourne, 3010, Vic, Australia

4. Laboratory for Atmosphere, Science Systems and Applications, Inc., NASA Goddard Space Flight Center Code 614 Greenbelt, MD 20771

Abstract

Errors in the specification or utilization of fossil fuel CO_2 emissions within carbon budget or atmospheric CO_2 inverse studies can alias the estimation of biospheric and oceanic carbon exchange. A key component in the simulation of CO_2 concentrations arising from fossil fuel emissions is the spatial distribution of the emission near coastlines. Regridding of fossil fuel CO_2 emissions ($FFCO_2$) from fine to coarse grids to enable atmospheric transport simulations can give rise to mismatches between the emissions and simulated atmospheric dynamics which differ over land or water. For example, emissions originally emanating from the land are emitted from a gridcell for which the vertical mixing reflects the roughness and/or surface energy exchange of an ocean surface. We test this potential "dynamical inconsistency" by examining simulated global atmospheric CO_2 concentration driven by two different approaches to regridding fossil fuel CO_2 emissions. The two approaches are: (1) a commonly-used method that allocates emissions to gridcells with no attempt to ensure dynamical consistency with atmospheric transport; (2) an improved method that reallocates emissions to gridcells to ensure dynamically consistent results. Results show large spatial and temporal differences in the simulated CO_2

concentration when comparing these two approaches. The emissions difference ranges from -30.3 TgC/gridcell/yr (-3.39 ~~K~~kgC/m^2/yr) to +30.0 TgC/gridcell/yr (+2.6 ~~K~~kgC/m^2/yr) along coastal margins. Maximum simulated annual mean CO_2 concentration differences at the surface exceed ±6 ppm at various locations and times. Examination of the current CO_2 monitoring locations during the local afternoon, consistent with inversion modeling system sampling and measurement protocols, finds maximum hourly differences at 38 stations exceed ±0.10 ppm with individual station differences exceeding -32 ppm. The differences implied by not accounting for this dynamical consistency problem are largest at monitoring sites proximal to large coastal urban areas and point sources. These results suggest that studies comparing simulated to observed atmospheric CO_2 concentration, such as atmospheric CO_2 inversions, must take measures to correct for this potential problem and ensure flux and dynamical consistency.

1 Introduction

The terrestrial biosphere and oceans play a critical role in the global carbon cycle by removing approximately 5.1 PgC/yr of CO_2 out of the total emitted due to industrial activity and deforestation (Le Quéré et al., 2013). Quantification of the spatial and temporal patterns of this removal using atmospheric CO_2 inversions is an important approach for understanding the feedbacks between the carbon cycle and the climate system (e.g., Gurney et al., 2002). Atmospheric CO_2 inversions infer the ocean and biosphere uptake by solving a set of source-receptor relationships, with the fossil fuel CO_2 emissions acting as either a boundary condition with no uncertainty or as a "prior" flux for which some adjustment is allowed in the inversion process (Enting, 2002).

Global fossil fuel CO_2 emission data products are now being produced at spatial resolutions smaller than 10 km and time resolutions that resolve the diurnal cycle (Rayner et al., 2010; Oda and Maksyutov, 2011; Wang et al., 2013; Nassar et al., 2013). This, along with the increasing density of atmospheric CO_2 concentration observations, places new emphasis on a careful examination of the use and uncertainty associated with these high resolution fossil fuel CO_2 emission data products (Ciais et al., 2009; Gurney et al., 2005; Peylin et al., 2011; Nassar et al., 2013; Asefi-Najafabady et al., 2014). For example, Gurney et al. (2005) found a monthly regional bias of up to 50% in the biosphere's net carbon exchange caused by unaccounted variation in fossil fuel emissions. Peylin et al. (2011~~3~~) also showed a large response in simulated CO_2 concentration to the spatial and temporal resolution of fossil fuel emissions over Europe. Similarly, Nassar et al. (2013) confirmed the importance of hourly and weekly cycles in fossil fuel emissions to simulated CO_2 concentration levels. It is clear from these studies that the

specification of the fossil fuel CO_2 emissions is a critical component in efforts that either use fossil fuel emissions directly or as part of an atmospheric CO_2 inversion process.

In addition to concerns regarding the accuracy of the high-resolution fossil fuel CO_2 emission data products, there are elements of uncertainty in how they are used within atmospheric tracer transport schemes, either in forward simulation or inverse mode. Transport models typically distinguish the surface characteristics of a model gridcell in broad classes such as land versus water or urban versus rural. These classifications are important to both the emissions of $FFCO_2$ and atmospheric transport above and/or downwind of particular gridcells. For example, modeled atmospheric transport processes such as mixing with the planetary boundary layer, convection, synoptic flow, and even general circulation are influenced by the gridcell surface characteristics (e.g. surface roughness or energy budget).

Global tracer transport models usually discretize surface gridcells at a lower resolution than those of fossil fuel CO_2 emission data products produced in recent years and, thus, the emissions need to be aggregated to the coarser model resolution. In this process, the transport model gridcells with less than 50% land geography are usually designated as water gridcells. Emissions present on the finer $FFCO_2$ grid, resident within the coarser model water gridcell are thereby mixed into the atmosphere according to vertical mixing characteristics of ocean or lake transport dynamics. This inconsistency between the emissions and transport dynamics can cause bias both locally and downwind of the errant gridcell(s). This problem is particularly important for fossil fuel CO_2 emissions as they are notoriously large along coastal margins where population and infrastructure are dominant.

This study aims to quantify this bias arising from the regridding of fossil fuel CO_2 emissions in global tracer transport simulations. The bias is defined as spatial distribution and temporal

variations of the simulated CO_2 concentration difference driven with two regridded fossil fuel emission inventories. We do this by constructing two experiments: 1) using the typical regridding procedure in which emissions are left in gridcells defined by the majority surface geography and 2) proportionally shifting or "shuffling" these emissions to neighboring land gridcells to maintain the spatial integrity of the fossil fuel emissions while avoiding the emissions-transport inconsistency.

Although a similar phenomenon might be expected for inland urban areas where designation of urban versus rural gridcells may not align with surface emissions, the global tracer transport models used in this study do not attempt to resolve transport dynamics over urban versus rural areas. Thus, we restrict ourselves to the study of the land versus water misallocation problem.

Section 2 describes the fossil fuel CO_2 emission data product used in the simulations, the atmospheric transport model employed and the adjustment method used to regrid the emissions. Section 3 presents results highlighting the difference induced by the shuffling procedure. We examine differences in emissions and in concentrations, the latter performed at active CO_2 monitoring locations for which the shuffling influence is greatest. Section 4 presents our conclusions.

2 Methods

The impact of fossil fuel CO_2 emission regridding is tested here by examination of simulated CO_2 concentration driven by two different emission fields through an atmospheric transport model. The fossil fuel CO_2 emissions are aggregated from a 0.1˚ x 0.1˚ grid to a 1.25˚ x 1.0˚ transport model grid. One of these emission fields has the coastal gridcells "shuffled" to correct

for the regridding impact ("experiment") while the other is left in the original unshuffled condition ("control").

2.1 Fossil fuel CO_2 emissions

Fossil fuel CO_2 emissions from the Fossil Fuel Data Assimilation System (FFDAS) version 2.0 are used as the fossil fuel CO_2 emissions in this study (Asefi-Najafabady et al., 2014). The FFDAS emissions are produced on a 0.1˚ x 0.1˚ grid for every year spanning the 1997 to 2010 time period. We use emissions for 2002 in this study. The FFDAS is a data assimilation system that estimates the fossil fuel CO_2 emissions at every gridcell by solving a diagnostic model constrained by a series of spatially explicit observation datasets. The diagnostic model is the Kaya identity (Rayner et al., 2010) which decomposes emissions into population, economics, energy and carbon intensity terms. In FFDAS v2.0 the observational datasets are used to constrain elements in the Kaya decomposition. The FFDAS uses the remote sensing-based nighttime lights data product, gridded population and national sector-based fossil fuel CO_2 emissions from the International Energy Agency (IEA), and a recently constructed database of global power plant CO_2 emissions (Elvidge et al., 2009; Asefi-Najafabady et al., 2014~~; Gurney and Coltin, 2014~~).

FFDAS version 2.0 originally estimates fossil fuel emissions at 0.1˚/annual resolution over the globe. Based on the annual resolution, higher temporal resolution products at hourly scale are constructed by multiplying sub-annual fraction by annual total at each grid. Considering the goal of this study, we choose the product at 0.1˚/annual resolution to construct the fossil fuel emissions with no time structure. The annual amount in each grid cell is divided by 2920 to obtain evenly distributed emissions at three-hour model resolution.

2.2 Atmospheric transport model

This study uses a global tracer transport model, the Parameterized Chemical Transport Model (PCTM) to simulate the CO_2 concentration resulting from the FFDAS surface emissions (Kawa et al. 2004; 2010). The model uses dynamical fields from the Modern-Era Retrospective analysis for Research and Applications (MERRA) (Bosilovich, 2013), which is a NASA reanalysis for the satellite era using a new version of the Goddard Earth Observing System Data Assimilation System Version 5 (GEOS-5). The initial data product of GEOS-5 is at 0.7˚ longitude x 0.5˚ latitude with 72 hybrid vertical levels. Two coarser MERRA products are also produced by aggregating the high-resolution product to a resolution at 1.25˚ longitude x 1.25˚ latitude or 1.25˚ longitude x 1˚ latitude with 72 hybrid vertical levels (Rienecker et al., 2011; Reichle et al., 2011; Reichle et al., 2012). In atmospheric transport simulation and inversion system, a dynamical consistence problem might be introduced if the driving meteorology data doesn't match the transport model grid. However, this problem doesn't exist in this study, since the MERRA product used in this study is on the same grid as PCTM. The model uses a semi-Lagrangian advection scheme; the subgrid-scale transport includes convection and boundary layer turbulence processes. The model is run at 1.25˚ longitude x 1.0˚ latitude with 56 hybrid vertical levels. The vertical mixing profile in PCTM includes two dynamical processes: turbulent diffusion in the boundary layer and convection. The two processes are parameterized following the MERRA model – which differentiates the vertical mixing in the boundary layer over land and ocean by using different surface heating, radiation, moisture, roughness and other physical factors in the eddy diffusion coefficient (Kh scheme) (Louis et al., 1982; Lock et al., 2000; McGrath-Spangler and Molod, 2014). Considering the purpose of this study, a check of the diffusion coefficients of the MERRA meteorology is performed. The result shows a significant difference between land

and ocean planetary boundary layers, indicating the existence of different vertical mixing characteristics between the two boundaries (~~see figure in the Supplement~~ Fig. 1).

The simulation is run for four years, driven by 2002 MERRA meteorology and fossil fuel CO_2 surface emissions (cycled repeatedly). The MERRA meteorology has a 3-hour time resolution and a 7.5-minute time step is used in the model simulations. There is no time structure in the fossil fuel emissions. In the model simulations, tracers are propagated in the atmosphere to reach a state of equilibrium under the applied forcing. This is achieved with a four-year simulation in which the first three-year period is used for spin-up and the last year is used for analysis. The PCTM outputs hourly CO_2 concentration at every point in the three-dimensional grid. The annual mean surface CO_2 concentration field and hourly time series at GLOBALVIEW-CO_2 monitoring sites are analyzed (http://www.esrl.noaa.gov/gmd/ccgg/globalview/co2/) (Masarie and Tans, 1995).

2.3 Coastal "shuffling"

The FFDAS emissions are regridded from the original 0.1˚ x 0.1˚ resolution to the 1.25˚ longitude x 1.0˚ latitude resolution of the PCTM. The two grids have the same origin and hence, the coarser grid is overlaid onto the finer grid and the 0.1˚ gridcells are integrated, as needed. In the longitudinal direction, gridcell boundaries do not align and so area-weighting was used to distribute emissions.

The PCTM utilizes a gridded land/sea mask that is used to denote the character of the model surface (land versus ocean/lake). The designation is based on what constitutes the majority type within each gridcell. In order to maintain dynamical consistency with the land/sea mask, those gridcells that are considered ocean/lake by the mask but contain emissions integrated from the

0.1˚ degree emissions grid, are treated with a "shuffling" procedure. These gridcells will have the emitted quantities transferred to adjacent land gridcells according to weights assigned by the relative magnitude of those adjacent land gridcells (~~see~~ Fig. ~~1~~2). The weight is defined as the ratio of emissions in each of the designated adjacent gridcells to the sum of their emissions:

$$w_j = F_j / \sum_{i=1}^{N} F_i \quad (1)$$

where w_j is the weight of the jth land gridcell, F_j is its emissions, N is the total number of land gridcells to which emissions are transferred. Adjacent gridcells are defined as those that share a corner with the shuffled cell.

3 Results and discussion

3.1 Emissions difference

The shuffling procedure reallocates emissions along global coastlines but the impact on the final CO_2 fluxes is most pronounced where there are large coastal emissions associated with urban areas or large point sources. Fig. ~~2~~3 shows the difference in surface emissions between the control and experiment emission fields. The coastal locations with cities or large point sources exhibit an emissions "dipole". Positive values reflect the addition of emissions to land gridcells adjacent to those designated as ocean in the coarse grid land/sea mask while negative values reflect the removal of emissions from gridcells designated as ocean.

The largest emissions adjustments occur in coastal areas of the US Great Lakes, coastal Europe, China, India and Japan. The range of the emission difference varies from -30.3 TgC/gridcell/yr (-3.39 ~~K~~kgC/m^2/yr) to +30.0 TgC/gridcell/yr (+2.6 ~~K~~kgC/m^2/yr). To provide context, an emission difference of 30 TgC/gridcell/yr is equivalent to ~62% and ~13% of the annual total carbon

emissions for the Netherlands and Germany in 2002, respectively, but is only limited to a few gridcells in Eastern Asia. Most emission differences in land gridcells vary in between 0.001 TgC/gridcell/yr (0.0001 k~~K~~gC/m^2/yr) and 5.0 TgC/gridcell/yr (0.056 k~~K~~gC/m^2/yr). The summed magnitude of the emissions that are relocated from ocean to neighboring land gridcells is 674.5 TgC/yr, which is equivalent to ~10% of the global total fossil fuel CO_2 emissions in 2002.

3.2 CO_2 concentration difference

The atmospheric CO_2 concentration resulting from the control and experiment simulations offers additional insight into the impact of the regridding and coastal shuffling (Fig. ~~3~~4). Similar to the emissions difference, the simulated CO_2 concentrations in the lowest model layer show differences along coastlines where large urban centers or point sources are present. In contrast to the emission differences, the response of surface CO_2 concentration not only reflects the immediate local emission impact but also a downwind impact as the differing concentration fields are transported by atmospheric motion. A particularly notable example is the surface CO_2 concentration difference downwind of the cluster of large coastal western European cities, for example, London, Rotterdam, ~~Groningen,~~ Barcelona and Rome. Also evident are dipole patterns associated with many of the large CO_2 concentration differences along the coastline driven by the emission dipole explained in Section 3.1, with negative values over ocean gridcells and positive values over the adjacent land gridcells.

The annual mean concentration differences range from -6.60 ppm to +6.54 ppm at the gridcell scale. These CO_2 concentration differences should be placed in the context of well-known surface concentration gradients such as the north-south gradient in annual mean CO_2 concentration of ~4.0 ppm and northern hemisphere longitudinal gradients of ~1.5 ppm (Conway and Tans, 1999). These differences represent a potential bias in the simulated CO_2 signal at, or

downwind from, numerous locations associated with coastal/urban areas, and are the combined result of the differing emission distribution in the two experiments acted upon by the atmospheric transport.

3.3 Hourly CO_2 concentration

Here we examine the simulated CO_2 concentration differences at locations where CO_2 concentrations are directly monitored, in an attempt to provide more guidance to atmospheric CO_2 inversion studies that use these locations as the observational constraint to estimating carbon exchange between the ocean, land and atmosphere. An examination of the hourly time series of CO_2 concentration in the lowest model layer at GLOBALVIEW monitoring stations indicates that 169 stations (out of 313 total GLOBALVIEW stations) show hourly CO_2 concentration differences greater than ±0.10 ppm and 12 of these stations show differences that exceed ±2.0 ppm (Fig. 4). Most of the larger differences are located close to coastal urban areas and occur at night and the early morning hours. This is not surprising given the reduction in mixing between the free troposphere and the planetary boundary layer at these times.

The hourly differences at these 12 stations range from -32.1 ppm to +2.50 ppm. Tae-ahn Peninsula (TAP) has the largest response (-32.1 ppm). ~~mainly due to its close proximity to a heavily populated coastal area and the subsequent large gradients in the experiment versus control emissions.~~ Yonagunijima (YON) and Gosan (GSN) also show large responses, with maximum differences reaching +5.23 ppm and -4.43 ppm, respectively.

Given the fact that many atmospheric CO_2 inversions sample the simulated and observed CO_2 concentration as a local afternoon average, and the simulated maximum differences found here occur at varying times of day, greater insight can be gained by examining the simulated

differences during the afternoon. In this case, 38 surface stations show hourly CO_2 concentration differences exceeding a magnitude of ± 0.10 ppm during the local afternoon (12:00—18:00). Of the 38 stations, five (TAP, GSN, SCSN, YON and RYO) have a local afternoon mean difference ranging between 0.12 ppm and -4.58 ppm (Fig. 4) with the largest difference at Tae-ahn Peninsula (TAP) in South Korea.

The shift between a positive and negative bias shown in Fig. 4 owes to the fact that these coastal sites likely experience onshore and offshore airflow at different times and this changes which portion of the local emission dipole influences the monitoring location. The specific circumstances at the TAP station are a good example of how the transport acts upon the emission dipoles to either enhance or diminish the concentration differences seen in Fig. ~~5~~6. TAP is a coastal station (36˚43′N, 126˚07′E) locating at the Tae-ahn Peninsula in the Republic of Korea. This site is in close proximity to two urban cities, Seosan and Taean. TAP is designated to ocean grid cell in PCTM resolution. The emissions on this grid are aggregated to adjacent land grid after regridding, and thus is located in the negative portion of the emission dipole~~The TAP station is located approximately 300 m offshore in the Tae-ahn Peninsula and close to two large urban areas, Seosan and Hongseong, South Korea. The TAP monitoring station~~ is located in the negative portion of the emission dipole (emission difference: -24.1 TgC/gridcell/yr) corresponding to the positive emission portion on adjacent land gridcells, as displayed in Fig. ~~5~~6a. Consistently, the TAP site lies in the negative portion of the annual mean surface CO_2 concentration field (-6.60 ppm) opposing to the positive portion on land (Fig. ~~5~~6b). Timeseries of the hourly concentration difference for the TAP site shows the largest value ~~by~~ of about -32.1 ppm occurring on January 13th at 5:00 pm local time. PCTM wind fields show low wind speeds on January 12th (daily mean: <2 m/s) and in the daytime of January 13th (3.5 m/s) compared to

the much higher monthly mean value (8.4 m/s). The weak transport during this time period accentuates the difference between the two experiments by lessening the amount of horizontal mixing and dispersion of the dipole gradient in this location. The hourly timeseries of the TAP site also shows high~~High~~-frequency behavior ~~signals are also shown~~ throughout the year ~~in the hourly timeseries for the TAP site~~, indicating the impact of synoptic-scale atmospheric transport. Another feature to note is the seasonal pattern in the hourly CO_2 concentration difference time series, with larger absolute magnitudes appearing at RYO, YON and TAP in the spring and summer, indicating a seasonal contribution of atmospheric transport to the potential monitoring station bias. Further examination of the hourly time series also shows diurnal patterns in all 12 monitoring sites.

3.4 Implications for carbon cycle studies

Research in which simulated CO_2 concentrations are compared to observed must consider ways to avoid the potential bias introduced when regridding high-resolution fossil fuel CO_2 emissions to the lower-resolution grids typical of atmospheric transport models. Atmospheric CO_2 inversion studies are also a good example of research that must overcome this potential problem. However, we don't consider the impact and uncertainty on atmospheric inversion in this study, since atmospheric inversions are not the only purpose for simulations of fossil fuel-like tracers. Many studies in atmospheric chemistry have the same need and consequently the same problem. But the study also does do something of direct use for an inversion. The fossil fuel is part of the prior flux. So in an atmospheric inversion this term represents a systematic uncertainty in the mapping of fossil fuel flux into the prior mismatches (prior simulation of concentration – observations). It can be seen that the effect is widespread and large compared to the

measurement uncertainty usually used. Thus, this is enough to demonstrate significance for an inversion.

Utilizing the shuffling procedure outlined here is one way to minimize this potential bias in the spatial distribution of the fossil fuel CO_2 emissions. The goal is to maintain the localization of the large emission gradients that occur near coastlines due to the preponderance of large cities and point sources while simultaneously ensuring dynamic consistency between the emissions and modeled atmospheric transport.

Alternatively, modelers could use data selection procedures to minimize potential bias when choosing which CO_2 concentration observing sites to compare to simulated results (e. g., Law, 1996). Some inversion model system such as NOAA's CarbonTracker model sample only the afternoon daytime measurements at quasi-continuous stations to avoid times when the model boundary layer is less reliable (e.g. nighttime) (Peters, et al, 2007). Eliminating or de-emphasizing (via the assignment of large uncertainty) atmospheric CO_2 monitoring locations that are near, or strongly influence by, large fossil fuel CO_2 sources can reduce the potential for the emissions regridding problem. However, given that many global carbon cycle studies are observationally underconstrained, this choice does come with potentially large information loss. Given this fact, we recommend the use of an emissions shuffling procedure.

~~Many earth system models avail of "tiling" techniques which can assign more than one surface characteristic to a gridcell. Without further research testing the sensitivity of results to this technique, it is unclear to what extent this minimizes the fossil fuel CO_2 emissions regridding problem discussed in this study.~~ It also should be pointed out that the fossil fuel emissions from planes and ships are not included in this study. Airborne emissions are unlikely to be strongly impacted by this problem since the differences in atmospheric physics between land and ocean

decrease once above the boundary layer. While emissions from shipping do potentially suffer from this problem the fraction subject to misallocation will be small so the total problem is a small fraction of a small fraction.

It should be noted that the reshuffling simply might transfer errors from one place to another. For example reshuffling emissions away from an oceanic gridpoint may leave a station in that grid cell further from emissions than it really should be. This is possible of course. This can only been investigated by separating the transport and relocation effects by using an on-line model. However, it is expected that this shuffling method could introduce land-ocean biases, since fixed fossil sources are almost entirely land-based and putting them in ocean gridpoints seems far more likely to introduce land-ocean biases as the inversion tries to correct a poorly transported signal from the wrong environment. In general, without further research testing the sensitivity of results to this technique, it is unclear to what extent this minimizes the fossil fuel CO_2 emissions regridding problem discussed in this study.

4 Conclusions

This study tests the sensitivity of simulated CO_2 concentration to regridding of fossil fuel CO_2 emissions from a high resolution grid to a coarser global atmospheric transport model grid. Two experiments are conducted. The first regrids from the fine to coarse grid but with no post-regridding adjustment to those emitting gridcells that inevitably ends up in the ocean ("control"). The second experiment performs the same regridding process as 1) but moves or "shuffles" the ocean-based emissions to adjacent land gridcells in a proportional manner. The two experiments exhibit large fossil fuel CO_2 emissions differences in coastal regions, which range from -30.3 TgC/gridcell/yr (-3.39 ~~K~~kgC/m^2/yr) to +30.0 TgC/gridcell/yr (+2.6 ~~K~~kgC/m^2/yr) which, when

summed globally, are equivalent to 10% of the 2002 global total fossil fuel CO_2 emissions. After transport of these emissions through a global tracer transport model, these two experiments show simulated CO_2 concentration differences along the coastal margin in both the spatial and temporal domains. The resulting annual mean surface CO_2 concentration difference when examining all surface gridcells varies between -6.60 ppm to +6.54 ppm. At the hourly level, individual CO_2 concentration differences exceed ±0.10 ppm at 38 monitoring stations, with a maximum of -32.1 ppm at one monitoring locations. When examining local afternoon mean values (average of 12:00-18:00), which both modeling systems and monitoring protocols emphasize, the CO_2 concentration differences are as large as -4.58 ppm. These CO_2 concentration differences result from the shifted emissions acted upon by modeled meteorology and can result in biased flux estimation in atmospheric CO_2 inversions which rely on comparison of simulated to measured CO_2. This phenomenon is also potentially important in any study investigating source-receptor simulations such as those found in air quality and other trace gas research efforts.

5 Code availability

The Fortran code to regrid and reallocate the surface fossil fuel emissions flux to ensure the dynamical consistence between emission and global transport model is available from the corresponding author (email:Xia.Zhang11@asu.edu).

Acknowledgements

This work was supported by NASA CMS grant NNX12AP52G. Rayner is in receipt of an Australian Professorial Fellowship (DP1096309).

References

Asefi-Najafabady, Rayner, S., P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Elvidge, C., Baugh, K.: A new global gridded dataset of CO_2 emissions from fossil fuel combustion: Methodology, evaluation and analysis, submitted to J. Geophys. Res., 2014.

Bosilovich, M. G.: Regional climate and variability of NASA MERRA and recent reanalyses: U.S. summertime precipitation and temperature, J. Appl. Meteor. Climatol., **52**, 1939–1951, 2013.

Ciais, P., J., Paris, D., Marland, G., Peylin, P. , Piao, S. L., Levin, I., Pregger, T., Scholz, Y., Friedrich, R., Rivier, L., Houwelling, S., Schulze, E. D. and members of the CARBOEUROPE Synthesis Team (1): The European carbon balance revisited. Part 4: fossil fuel emissions, Global Change Biol., doi:10.1111/j.1365-2486.2009.02098.x. , 2011

Conway, T. J. and Tans P. P.: Development of the CO2 latitude gradient in recent decades, Global Biogeochem. Cy., *13*, 821-826, 1999.

Elvidge, C. D., Ziskin, D., Baugh, K. E., Tuttle, B. T., Ghosh, T., Pack, D. W., Erwin, E. H., Zhizhin, M.: A Fifteen Year Record of Global Natural Gas Flaring Derived from Satellite Data, Energies, 2 (3), 595-622, 2009.

Enting, I.: Inverse Problems in Atmospheric Constituent Transport, Cambridge Univ. Press, New York, 2002.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K, Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J.,

Taguchi, S., Takahashi, T., and Yuen, C.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626-630, 2002.

Kawa, S. R., Erickson, D. J. III, Pawson, S. and Zhu, Z.: Global CO_2 transport simulations using meteorological data from the NASA data assimilation system, J. Geophys. Res., 109, D18312, doi:10.1029/2004JD004554, 2004.

Kawa, S. R., Mao, J., Abshire, J. B., Collatz, G., J. and Weaver, C. J.: Simulation studies for a space-based CO_2 lidar, Tellus, 62B, 759-769, 2010.

Law, R.: The selection of model-generated data CO_2 data: a case study with seasonal biospheric sources, Tellus, *48B*, 474-486, 1996.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I. , Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viovy, M., Zaehle S. and Zeng, N.: The Global Carbon Budget 1959-2011. Earth Syst.Sci. Data, 5, 165-185, 2013.

Masarie, K.A. and Tans, P. P.: Extension and Integration of Atmospheric Carbon Dioxide Data into a Globally Consistent Measurement Record. J. Geophys. Res., 100, No. D6, p. 11593-11610, 1995.

Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M. and Smith, R. N. B.: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. Mon. Wea. Rev., 138, 3187–3199, 2000.

Louis, J., Tiedtke M., and J. Geleyn: A short history of the PBL parameterization at ECMWF. Proc. ECMWF Workshop on Planetary Boundary Layer Parameterization, Reading, United Kingdom, ECMWF, 59–80, 1982.

McGrath-Spangler, E. L. and Molod A.: Comparison of GEOS-5 AGCM planetary boundary layer depths computed with various definitions, Atmos. Chem. Phys., 14, 6717-6727, 2014S

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R. and Deng, F.: Improving the temporal and spatial distribution of CO_2 emissions from global fossil fuel emission data sets, J. Geophys. Res.: Atm, 118, 917-933, 2013.

Oda, T**.** and Maksyutov, S.: A very high-resolution (1 km×1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights, Atmos. Chem. Phys., 11, 543-556, 2011.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarle, K., Miller, J. B., Bruhwiler, L. M. P., Pėtron, G., Hirsch, A. I., Worthy, D. E. J., vander Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol M. C. and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, PNAS,104, 18925-18930, 2007.

Peylin, P., Houweling, S., Krol, M. C., Karstens, U., Rödenbeck, C., Geels, C., Vermeulen, A., Badawy, B., Aulagnier, C., Pregger, T., Delage, F., Pieterse, G., Ciais, P. and Heimann, M.: Importance of fossil fuel emission uncertainties over Europe for CO_2 modeling: model intercomparison, Atmos. Chem. Phys., 11, 6607-6622, 2011.

Rayner, P.J., Raupach, M.R., Paget, M., Peylin, P. and Koffi, E.: A new global gridded data set of CO_2 emissions from fossil fuel combustion: Methodology and evaluation, J. Geophys. Res., 115, D19306, doi:10.1029/2009JD013439, 2010.

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P. and Toure, A.: Assessment and enhancement of MERRA land surface hydrology estimates, J. Clim., 24, 6322-6338, doi:10.1175/JCLI-D-10-05033.1, 2011.

Reichle, R. H.: The MERRA-Land Data Product, version 1.1.GMAO Technical Report, NASA Global Modeling and Assimilation Office, Goddard Space Flight Center, Greenbelt, MD, USA. Available at http://gmao.gsfc.nasa.gov/reasearch/merra/ile specifications.php, 2012.

Rienecker and Coauthors, MERRA – NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Clim., 24, 3624-3648, doi:10:10.1175/JCLI-D-11-00015.1, 2011.

Wang, R., Tao, S., Ciais, P., Shen, H. Z., Huang, Y., Chen, H., Shen, G.F., Wang, B., Li, W., Zhang, Y.Y., Lu, Y., Zhu, D., Chen, Y.C., Liu, X.P., Wang, W.T., Liu, W.X., Li, B.G. and Piao, S. L.: High-resolution mapping of combustion processes and implications for CO_2 emissions, Atmos. Chem. Phys., 13, 5189-5203, 2013.

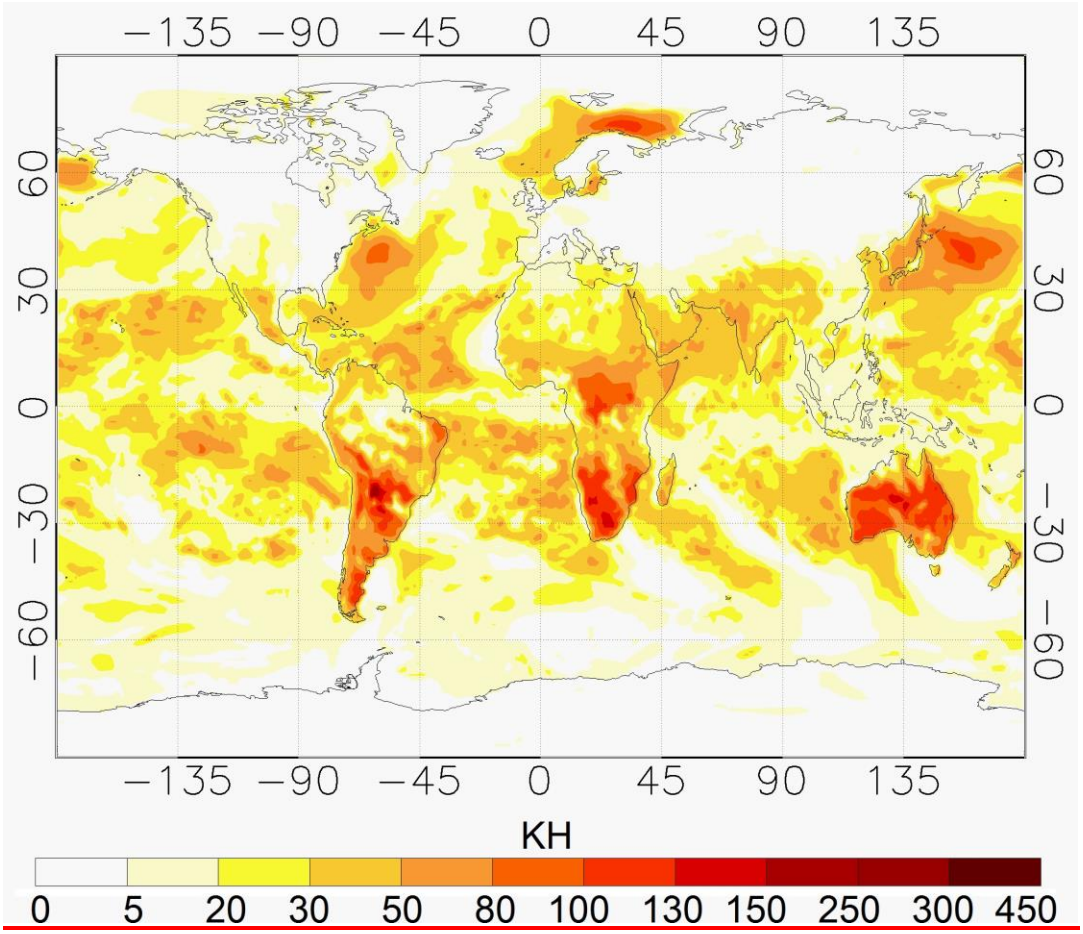


Fig. 1. Daily mean diffusion coefficient (KH) at 1.25˚ x 1.0 ˚ for July 30, 2002 at pressure level about ~950 hpa in MERRA reanalysis. The diffusion coefficient is determined using a K-diffusion scheme in MERRA modeling.

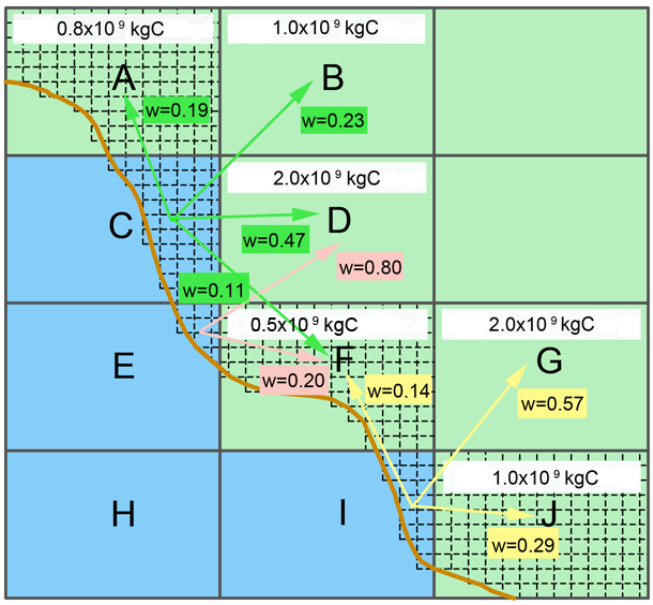
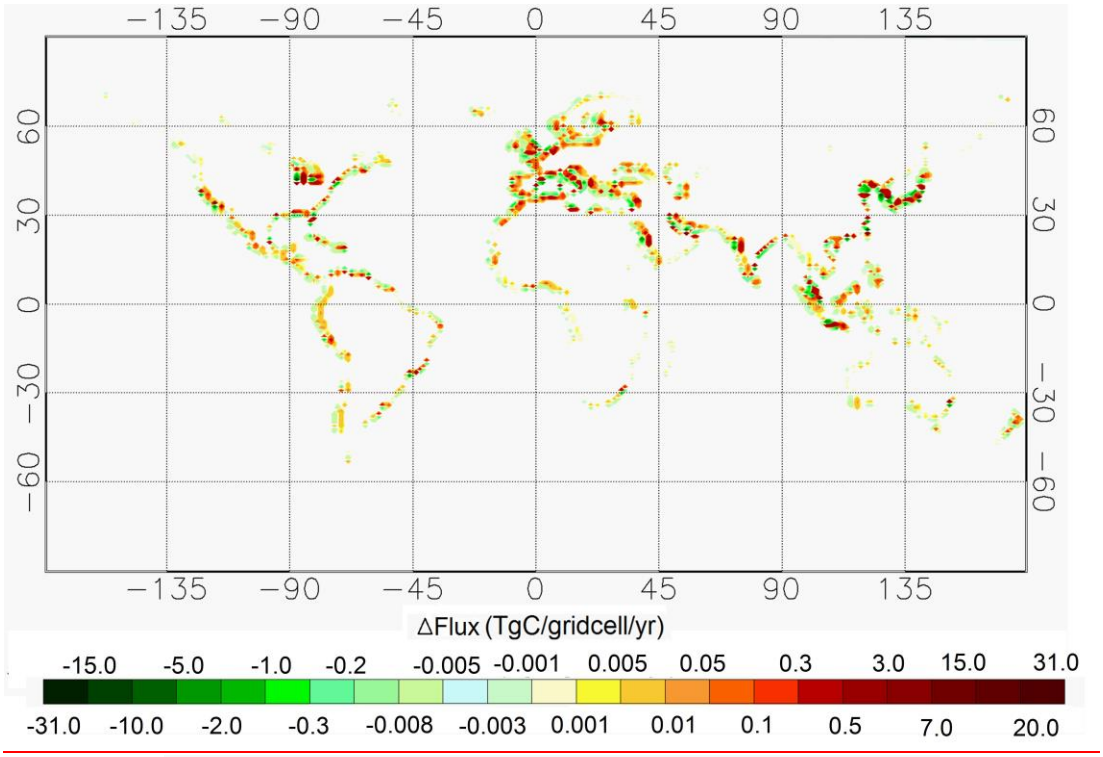


Fig.~~1~~2. Depiction of the "shuffling" procedure when regridding from a 0.1˚ x 0.1˚ to a 1.25˚ x 1.0˚ model grid. Capital black letters deote the coarser model grid (1.25˚ x 1.0˚). Gridcells outlined with dashed lines denote the finer model grid (0.1˚ x 0.1˚). Green denotes land, blue denotes water. Example emission values and weighting values (*w*) and the direction of the allocation are included.



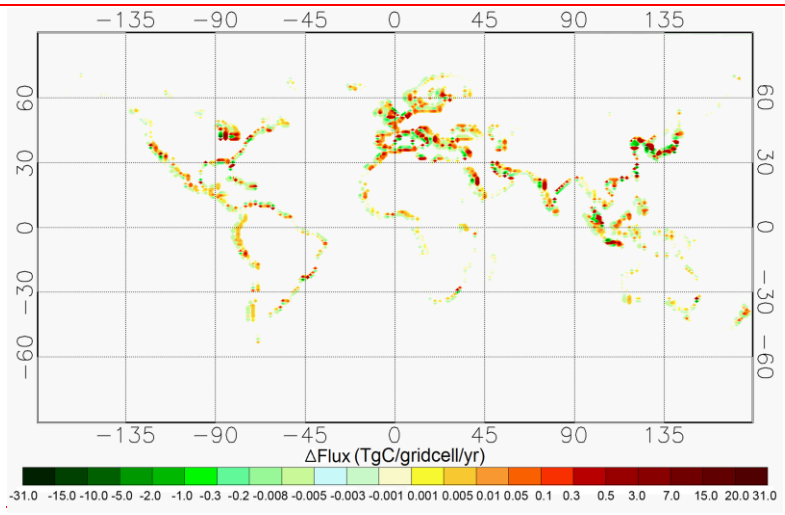


Fig. ~~2~~3. Difference between experiment and control fossil fuel CO_2 emissions. The difference is obtained by subtracting the control from the experiments. The emission values for some gridcells are not evident because the gridcells are saturated (beyond the color scale range).

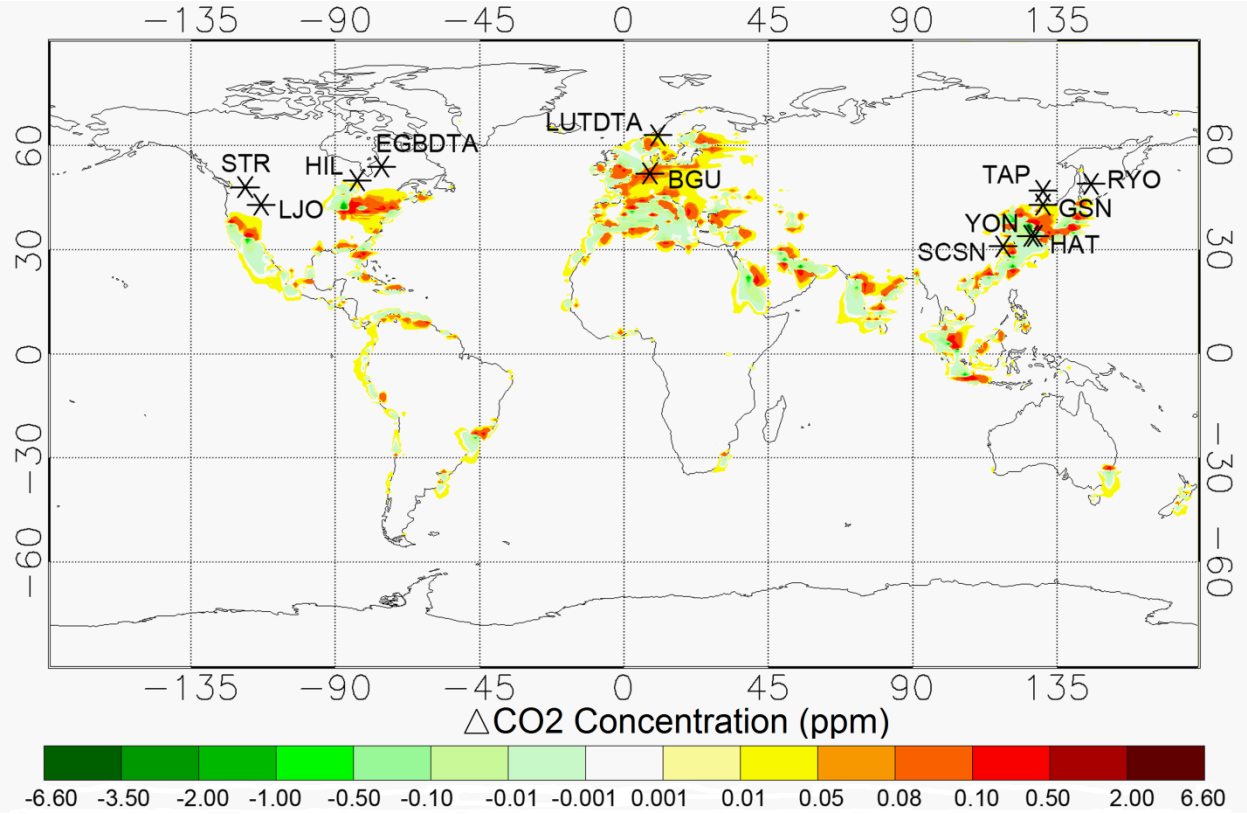


Fig. ~~3~~4. Simulated PCTM surface annual mean surface CO_2 concentration difference (experiment minus control, Units: ppm). The * in the figure denotes existing CO_2 monitoring locations where the annual mean CO_2 concentration difference exceeds 2 ppm.

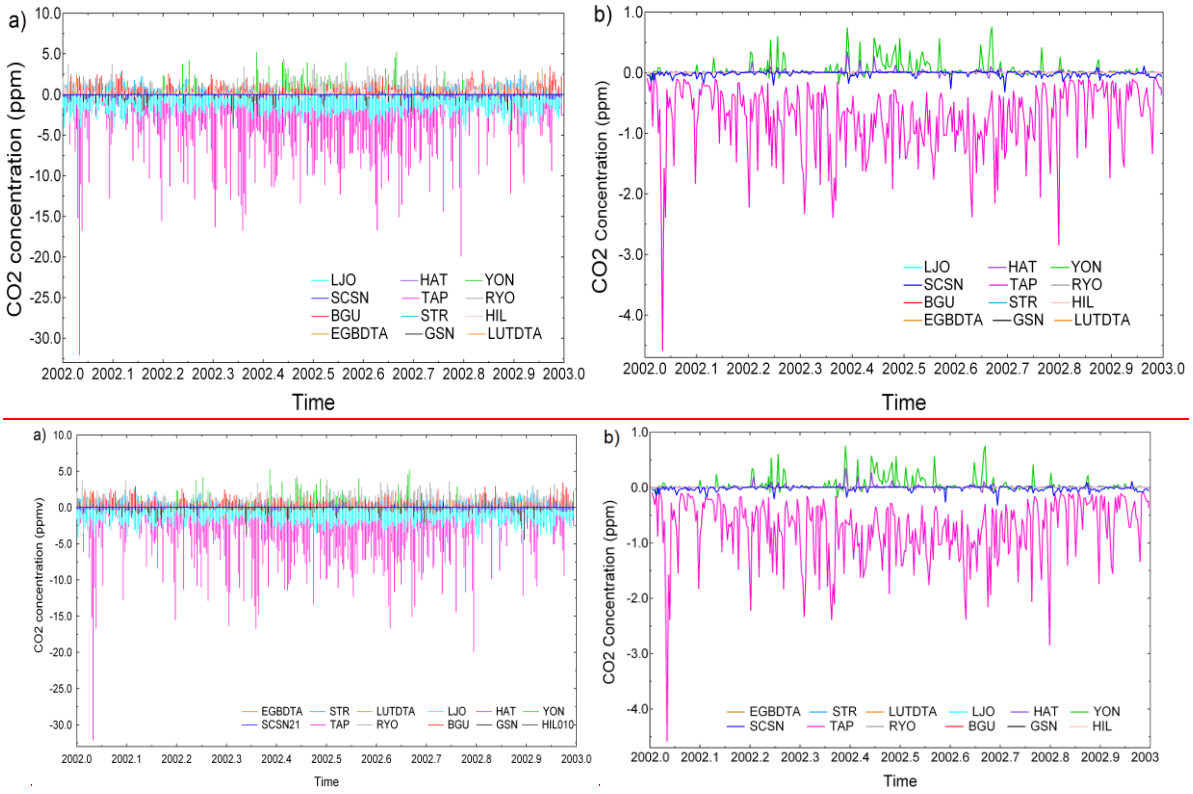


Fig. 45. Simulated PCTM surface CO_2 concentration difference (experiment minus control, Units: ppm) at the 12 GLOBALVIEW monitoring stations with the largest concentration difference. a) Hourly means CO_2 concentration difference; b) afternoon means (12:00-18:00) CO_2 concentration differences.

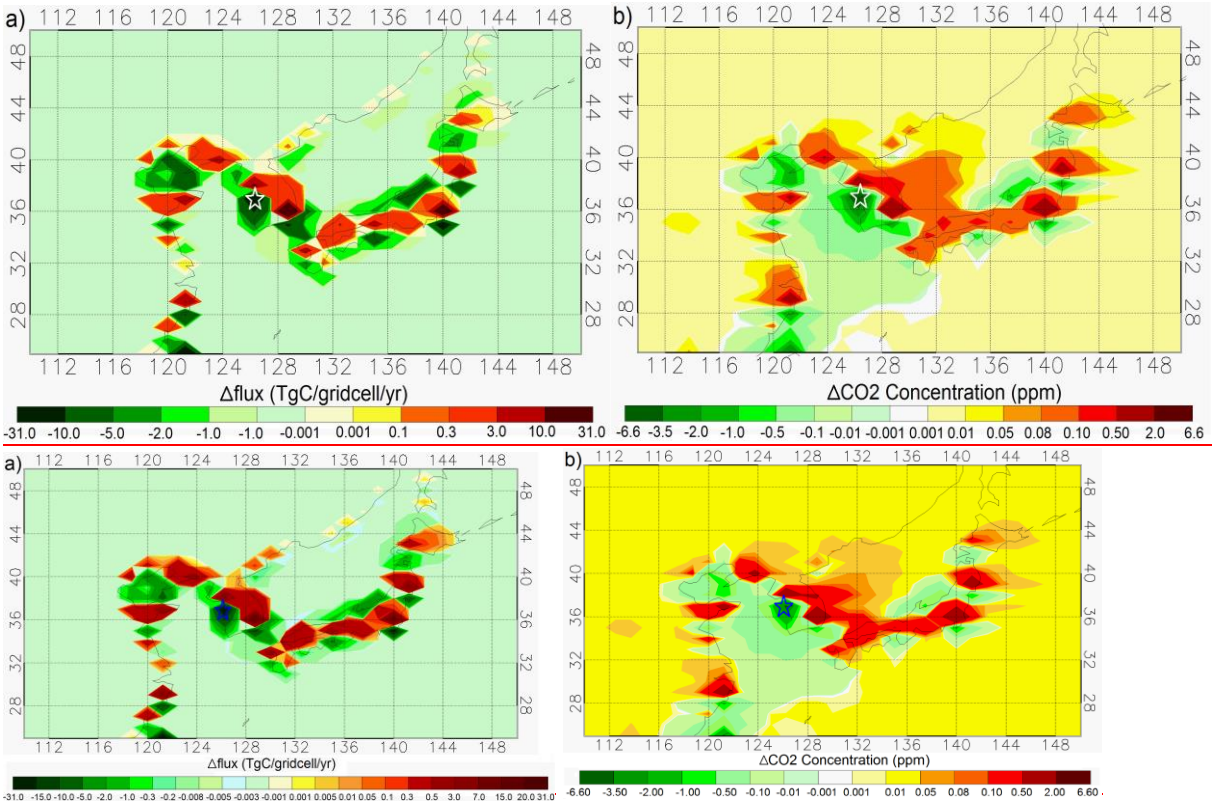


Fig. ~~5~~6. Regional fluxes difference and simulated surface CO_2 concentration differences (experiment minus control) and the location of GLOBALVIEW monitoring site TAP. a) Flux difference; b) concentration difference. Blue stars mark the location of the TAP site.