Geosci. Model Dev. Discuss., 7, C203–C209, 2014
www.geosci-model-dev-discuss.net/7/C203/2014/

Geoscientific Model Development Discussions
Open Access

GMDD
7, C203–C209, 2014

Interactive comment on "A multiresolution spatial parameterization for the estimation of fossil-fuel carbon dioxide emissions via atmospheric inversions" *by* J. Ray et al.

Anonymous Referee #1

Received and published: 28 March 2014

This paper describes a novel approach for estimating fossil fuel emissions from proxy datasets such as nightlights and built-up-area maps, with estimates further informed by inverse modeling. The numerical tests considered use hypothetical measurements of fossil fuel CO2 from a network of instrumented towers distributed across the continental US.

The mathematical framework is rather complicated but is carefully described and can be understood with reasonable effort. However, the numerical tests presented are not realistic, and it is unclear from the present manuscript how this tool would be applied in practice. Ideally this should be addressed by applying the model to a more realistic case study as described below. If that is not possible, then the limitations of the numerical tests described in section 5 should be clearly described. The introduction or discussion should include a more detailed description of questions/applications for which this framework is appropriate. One such application could be an Observing System Simulation Experiment to evaluate the utility of a dense radiocarbon sampling network or the utility of continuous radiocarbon measurements that have to date been demonstrated with sufficient accuracy only in the laboratory.

Model code is provided in Matlab via a website. I'm sorry that I did not have a chance to download this, but I wonder whether it includes the early steps in the wavelet analysis (i.e., corresponding to eqn (3)). If so, that would enable motivated readers who are new to wavelet applications to gain practical understanding. Matlab code is helpful, but open source code would be even better.

Major concerns:

1) The proxy datasets may be strongly spatially correlated with energy use, but not necessarily with fossil-fuel emissions. Many large power plants are in the US are located far from the urban areas they serve. For example, large power plants in Wyoming and Ohio serve customers in distant urban areas. In the case of remotely located large power plants, the nightlight and built-up-area index would be unlikely to have intensity proportional to the emissions. In the US and certain other nations, detailed emissions data are available for large point sources. How could the framework be modified to take advantage of such information? E.g., could these point source emissions be subtracted prior to the inversion)? Is there another proxy dataset that could provide information about large point sources (e.g., perhaps high-resolution thermal imagery?) For areas where reliable emissions point source data are not available, might large point sources complicate or confound the analysis? Please address this in the introduction and/or discussion.

2) Fossil fuel CO2 cannot be directly measured. This is acknowledged in the manuscript but treated rather blithely. Radiocarbon measurements provide the most

direct measurement-based constraint available for separating biological and fossil fuel CO_2. Radiocarbon is a powerful tracer, but unfortunately measurements are expensive and are being made on discrete samples at a subset of the 35 tower measurement sites considered here at a rate of 3 midday samples per week. The errors on fossil fuel CO_2 estimated from radiocarbon are ~1ppm (J.B. Miller et al., JGR, 2012). The sampling frequency is overestimated by an order of magnitude and the measurement uncertainty is grossly underestimated by the numerical tests considered here. A technique for continuous measurement of radiocarbon has been demonstrated in the laboratory (D. Murnick, O. Dogru, and E. Ilkmen, 14C Analysis via Intracavity Optogalvanic Spectroscopy, Nucl. Instrum. Methods Phys. Res B., 2010 April 1; 268(7-8): 708–711. doi:10.1016/j.nimb.2009.10.010.), but field deployment of continuous radiocarbon sensors has not been demonstrated. Operational autonomous field operation will not be plausible for many years. I am curious whether a more realistic numerical test representative of currently available or plausibly augmented radiocarbon data (e.g. 10 - 35 towers, 3-7 mid-afternoon samples per week, 1 ppm measurement errors) would provide a useful constraint if aggregated over a long time period, e.g. 1 year, and limited to the region where the footprints show sensitivity.

3) In any regional inversion, boundary values need to be estimated and may have large uncertainty. Gourdji et al., (2012) showed that boundary/initial condition errors are potentially large enough to preclude reliable quantification of the net annual ecosystem uptake of CO_2 for North America. It is important to consider and discuss the potential complications of assigning fossil fuel CO_2 boundary values for the region where fluxes are being estimated, i.e. here the boundaries of CONUS. This seems especially complicated here, given that the impact of emissions from areas outside CONUS but within the rectangular domain would need to be taken into account. The compressive scaling strategy to exclude emissions outside of CONUS as described in the paper is appropriate for the idealistic case considered (synthetic obs), but in a real-data study either (1) accurate 4-dimensional fossil fuel CO_2 mole fraction values would be needed along the boundaries of the emission estimation domain (2) accurate 4-dimensional informa-

GMDD
7, C203–C209, 2014

tion about fossil fuel CO_2 mole fractions along the boundaries of the continent along with a correction for emissions within the rectangular domain but outside the emission estimation area. Other complications arise if a significant number of LPDM particles fail to exit the domain.

Specific comments:

1280:10 I recommend expanding the discussion of the potential for using radiocarbon measurements as an (almost) direct tracer for fossil fuel CO_2. Also note that accuracy of fossil fuel CO_2 estimated from radiocarbon is $\sim$1 ppm, measurements are limited because of cost, lack of technology, etc.

1280: After discussing radiocarbon, add a short paragraph about the atmospheric transport model describing signatures of emissions are dispersed and diluted and possible errors in simulated transport.

1283:8 Can wavelets be scaled up as well as shrunk?

1283:2nd to last line I don't understand what is meant where GothicW(s) then | GothicW(s)| (i.e., the |'s don't appear in the equation, but do appear in the description).

1284: Consider defining "random field" and briefly explaining why a random field is useful for representing complex emission maps

1285: eqn 10 Explicitly define || ||p notation here instead of or in addition to where it defined on pg 1288 ln 5.

1286:16 Is thare length scale associated with s=3 (i.e. in degrees lat/lon)? Struggling a bit to understand how wavelets manifest in physical space.

1288: 12 Why does sentence 2 ("Thus, while we ...") follow from sentence 1 ("Note that the sparse nature...")?

1290:13 & Figure 5 It looks like the magnitudes of the errors are similar to the magnitudes of the emissions themselves. It would be nice to include a plot of relative error

GMDD

7, C203–C209, 2014

would be interesting perhaps along with a scatter plot.

1291:2 Briefly explain here or in the introduction why you are using these 35 tower locations that are ill-suited for fossil fuel estimation. It is sufficient to state that the footprints were available from earlier studies and that it was convenient to use them for method development.

Section 4.2 It would be nice to include the figure from Ray (2013) showing a CDF to illustrate impact of non-negativity.

1292:1st paragraph 8 days seems a short timescale for estimating emissions. Was this selected to minimize aggregation error when computing monthly averages? Even annual estimates would be useful for some applications.

1295:18-22 As already noted, a radiocarbon-based inversion seems a much more straightforward application for this framework than extending the wavelet approach for simultaneous estimation of bio and fossil fluxes. However, the measurement density is much larger than will be possible for radiocarbon anytime in the next two years. The measurement errors for fossil fuel co2 are ~1ppm. Chatterjee et al. (2012) errors of 0.1 ppm seem optimistic even for total COÂň2 inversion unless model transport errors are somehow accounted for elsewhere (not an issue for synthetic data studies, but potentially important for real-data inversion). NOAA's CarbonTracker (www.carbontracker.gov) uses much larger sigma values for continental sites though they were assigned somewhat arbitrarily (Peters et al., PNAS, 2007). Also boundary value errors would be significant in any regional inversion.

1296:3 How different is the value of c when Edgar is used instead of Vulcan? How different are the total emissions?

1296:6 Please note also that 0.1 ppm is already very optimistic.

1299:2-6 and Fig 9 A relative error plots would be useful in addition to difference plot shown in Figure 9a.

GMDD

7, C203–C209, 2014

Sec 5 other comments: A figure showing one or more longitudinal transects of E, fpr, fv, and F (before non-negativity) would be interesting. Perhaps for a 32-day period. It would be nice to see the extent to which sharp spatial transitions are or are/not resolved for these quantities. If at all possible, it would be useful to include another more realistic case study with much sparser data (e.g. daily or thrice weekly samples) and with errors $\sim$1 ppm, corresponding to current radiocarbon capabilities.

1300:18-21: This overstates what has been demonstrated in the current study.

1301:5 Briefly describe what is meant by a dictionary or omit. The discussion of how the framework could be extended to account for biospheric fluxes is too brief to be of much use.

Figure 6. Please consider showing difference plots (from truth) and/or relative errors in addition to estimated emissions. Also, perhaps it would be more useful to show a 32-day average rather than an 8 day average.

Technical comments:

For Copernicus journals, I have found that it is advantageous to stack 2-panel figures vertically because they will be larger in the final typeset paper.

1280:5 unnecessary commas after "benefits" and "limitations"

1286: 5 "a priori using prior information" is redundant

1289:18 why switch to delta from alpha used earlier

1301:11 Ray (2013) is a link to the first author's individual webpage at Sandia. It does not seem like a particularly robust long-term repository for the MATLAB code. Perhaps a static version could be included as a supplement to the paper.

Ray et al. 2013 SAND Report reference: Is there a long-term repository for DOE technical reports that would perhaps be a better long-term link than the $\sim$jairay webpage?

GMDD
7, C203–C209, 2014

Fig 5. Very small in pdf and legible when viewed digitally only with much magnification. Middle panel x-axis range could be truncated to show only values with delta > 10-4. Perhaps move 5c to a separate figure along with additional panels showing relative error, which would be helpful.

Figure 7. RHS, legend notation seems inconsistent with caption (fk vs Ek?). I'm not convinced that incorporating yobs *clearly* improves the spatial agreement for the 8 day time periods, but agree that 32 day periods show substantial improvement.

Figure 8: Details could be seen more clearly if legend were removed and y-axis range truncated. Blue and black are hard to distinguish in printed copy, perhaps use black and green or grey instead?

Figure 9: Replace "Observations" with "Synthetic observations". Again blue and black are difficult to distinguish in printed copy.

Interactive comment on Geosci. Model Dev. Discuss., 7, 1277, 2014.