Rebuttal

A multiresolution spatial parameterization for the estimation of fossil-fuel carbon dioxide emissions via atmospheric inversions

J. Ray, V. Yadav, A. M. Michalak, B. van Bloemen Waanders, and S. A. McKenna

Response to Reviewer # 1

At the very outset of the response, we would like to clarify that we view the work described in the manuscript as a methodological first step in the development of an inversion scheme that could be potentially used to estimate ffCO2 emissions. In this paper, we have concentrated on describing a model for very spatially heterogeneous ffCO2 emissions fields and how this feature could be used in an inverse problem (it requires a class of optimization methods – sparse reconstruction – that is not widely used in CO2 inversion studies). The inverse problem adopted a number of simplifications – e.g. no boundary fluxes into regional domain, very small model-data mismatch errors, an assumption that the ffCO2 concentration signal at a sensor could be isolated from the biospheric signal – and used a sensor network that is not specifically designed for targeting ffCO2 emissions. Consequently, we do not imply that our method could be quickly adapted for real-data inversions using current transport models, sensor networks and ffCO2 concentration measurement techniques.

It would be reasonable to believe that accurate ffCO2 estimation would require a new sensor network, placed closer to major sources of ffCO2 emissions. Further, while radiocarbon is one way of estimating ffCO2 concentrations at sensors, one could also consider pollutants from incomplete combustion e.g., CO, to "back-out" the ffCO2 signal. However, these topics, though integral to the question of ffCO2 emission estimation, are out of scope in a paper that, as the title suggests, is about a spatial parameterization for ffCO2 emission fields.

Many of the issues required to adapt our method to real-data inversions and/or regional inversions e.g., specification of boundary fluxes, determining their uncertainties and assessing their impact on ffCO2 emission estimates, are identical to those faced by inversion studies for biospheric CO2 fluxes. These issues have been investigated and addressed in the biospheric CO2 context (Gourdji et al, 2012), and may provide starting points for adapting our method (for ffCO2) to a real-world scenario. In addition, the complications introduced by transport (dispersion and dilution) and the impact of transport model errors are identical to those faced by biospheric CO2 inversion studies and may be solved in a similar manner – the papers by Chatterjee et al, 2012 and Gourdji et al, 2012, cited in our manuscript, discuss these issues to a greater extent. Consequently, we consider these topics outside the scope of this paper.

We will add this clarification in the Introduction section. See Pg. 6, paragraph 2.

The reviewer states: "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."*

Response: We have added a "Discussion" section where we elaborate on what the spatial parameterization could be used for (primarily for Observation System Simulation Experiments, determining the location of sensors, and the frequency with which measurements are obtained) the limitations of the tests performed in Sec. 5, as well as the impact of various numerical and boundary condition approximations on the emission estimates. See Sect. 6, last para.

The reviewer states: "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))."

Response: We will include it.

Major concerns

The reviewer states: "The proxy datasets may be strongly spatially correlated with energy use, but not necessarily with fossil-fuel emissions. Many large power plants 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."

Response: The reviewer is correct in saying that the use of nightlights and built-up area maps, which correlate with energy use, could lead to an inaccurate random field model. Specifically, we may omit a fine-scale wavelet that corresponds to the large point source. However, the random field model is multi-resolution, implying that another wavelet, at a coarser level, whose support covers the point source, could model it. In doing so, the point source gets "smeared" over a larger area and the estimate of its magnitude may incur an error. However, it will not be totally omitted from the inversion procedure, with its emissions apportioned to other non-neighboring sources. This is a consequence of the multiresolution nature of our MsRF.

In case accurate databases of large point sources exist e.g., CARMA, the impact of the point sources can simply be subtracted out. If another proxy such as infrared images exists, the wavelets in our MsRF could be augmented with the wavelets (of the same family) chosen using the second proxy. If neither exists, the large point sources are smeared, as described above.

We realize that proxies are imperfect markers of ffCO2 emissions. One of the experiments in Section 3 specifically investigates the modeling ramifications of using imperfect proxies to construct the spatial parameterization. Experiments in Section 5 investigate the usefulness of the final model.

We will add this to the "Discussion" section. See Pg. 24, paragraph 2.

The reviewer states: "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 CO2. 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 CO2 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."

Response: The primary difficulty in performing a realistic inversion (1 ppmv noise) is the placement of the measurement towers – they are far from sources of ffCO2 emissions, leading to a ffCO2 concentration signal that is usually no more 2 ppmv on any sensor. Adding a noise with an error variance of 1 ppmv makes them unusable. A true test of our method, under realistic conditions, would also require a sensor network designed and sited to measure ffCO2 emissions. Consequently, in this paper, we have chosen an idealized case and focused on developing the inversion methodology.

The reviewer states: "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 CO2 for North America. It is important to consider and discuss the potential complications of assigning fossil fuel CO2 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 CO2 mole fraction values would be needed along the boundaries of the emission estimation domain (2) accurate 4-dimensional information about fossil fuel CO2 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."*

Response: The reviewer is correct in stating that our method for regional inversions will suffer from the same boundary condition issues that other regional (biospheric CO2) inversion methods do. This is unavoidable in the absence of good boundary condition data. The choice of option (1) versus (2) above would depend upon where ffCO2 concentrations (to serve as boundary fluxes) were available (around the CONUS boundary or around the continent, as in Gourdji et al, 2012). In case the boundary fluxes were available at the CONUS boundary, we would use option (1). As in our paper, we would not estimate emissions outside CONUS and use compressive sensing to suppress estimates there. The impact of ffCO2 emissions from OCONUS on the measurements would be imposed by time-variant ffCO2 influx/efflux along the CONUS boundary. In principle this is no different than Gourdji et al, (2012). The question of LPDM particles failing to exist is identical to that faced (and addressed) in Gourdji et al, 2012. However, the issue of boundary fluxes and the impact of those uncertainties on the ffCO2 estimates are outside the scope of the paper.

This has been added to the 'Discussion' section. See Sect. 6, paragraph 2.

Specific concerns

Pg 1280:10" "I recommend expanding the discussion of the potential for using radiocarbon measurements as a (almost) direct tracer for fossil fuel CO2. Also note that accuracy of fossil fuel CO2 estimated from radiocarbon is ~1 ppm, measurements are limited because of cost, lack of technology, etc."

Response: Our manuscript describes a method to estimate ffCO2 emissions predicated on the availability of concentrations of ffCO2 measured at a set of sensors. Radiocarbon is one of the ways of obtaining that measurement, but it could potentially also be derived from joint measurements of pollutants such as CO. Issues related to the cost and feasibility of making radiocarbon measurements are outside the scope of the paper.

Pg 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."

Response: The complications introduced by transport are no different from those encountered by inversion studies focused on biospheric CO2 fluxes, and these have been addressed in literature. We would consider them to be out of scope of our paper.

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

Response: Yes they can. That is clear from the expression for the wavelet at scale *s* and translation *j.*

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

Response: Gothic W is the set of (I, j) indices of wavelets of scale s. The magnitude of the set (Gothic W within vertical bars) is the number of (I, j) pairs in the set. We will add this clarification in the manuscript. See Pg. 7, second-last line.

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

Response: A random field model allows one to generate arbitrary fields based on the values assumed by the model's parameters. Certain characteristics required of the fields can be encoded into the random field model. For example, if the modeled fields are required to be smooth, one can impose a spatial correlation between field values at different locations, e.g. adopt a Gaussian random field model. The correlation function's parameters can be used to control the degree of smoothness. If the modeled fields are known to be rough at certain locations, they can be modeled using wavelets, with fine wavelets restricted to the rough regions and the wavelet weights acting as the model parameters. These parameters can assume arbitrary values i.e., they are random variables, and thus the model can create random fields.

We have added this to the beginning of Sec. 2.1. See Pg. 8, paragraph 1.

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

Response: We will do so. See Pg. 9, end of paragraph 1.

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

Response: Yes there is. The finest wavelets, on the M = 6 hierarchy, are on the 6^{th} level, and have a support of 2 degree X 2 degree. s = 3 wavelets are 3 levels above, and are 2^3 x 2^3 larger i.e., 16 degree X 16 degree. We will add this example in the manuscript. Another reviewer also requested it. See Pg. 10, Sect. 3, paragraph 1.

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

Response: The two sentences were badly framed. What we meant was: "We will use wavelets selected using the (single) nightlight and BUA maps to estimate weekly ffCO2 emissions. Our tests above show that they model annually averaged Vulcan emissions adequately, and we assume that while the emissions wax and wane with time, their spatial distribution does not vary sufficiently to require a new wavelet selection. We base this assumption on ffCO2 emissions' correlation with human activities, and static sources like powerplants which do not display large spatial dislocations with time". We will revise the manuscript accordingly. See Pg. 12, last paragraph of Sect. 3.1.

Pg 1290:13 and Fig. 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 would be interesting perhaps along with a scatter plot."

Response: The relative error plot is not very informative. Locations with large emissions, as predicted by fv and fpr can be slightly offset; further, since the emission fields are so rough, neighboring locations can have drastically smaller emissions. This leads to division by (almost) zero problems, leading to very large relative errors. These are rare, but they increase the dynamic range of the relative error plots. We will, however, include scatter plots of fv and fpr for each grid cell, plotted against each other, in the online Supplementary Material. We see that while there is a strong correlation between the two, they are far from being identical i.e., while the prior fluxes are a "guess" for the true emissions, they are not particularly close.

Pg 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."

Response: The towers chosen belong to the network that existed in North America in 2008, and therefore represents a realistic network, although far from optimal for the purpose of fossil fuel flux estimation. This explanation has been added to the revised manuscript. See Pg. 14, last line.

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

Response: We will do so. It is Fig. 7 in the revised manuscript.

Pg 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."

Response: The reviewer is correct. We will add the rationale behind the 8-day estimation period in the manuscript. See second sentence of Sect. 4.1.

Pg 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 CO2 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."

Response: We agree with the reviewer that a radiocarbon-based inversion is a more straightforward application for our framework than joint biospheric-fossil-fuel CO2 inversion. However, the aim of the paper was to introduce a spatial parameterization, an accompanying sparse reconstruction method and provide evidence of their usefulness in an inversion. We have adopted a number of simplifications to do so, as the reviewer has pointed out. We used sensors from a network that existed in North America in 2008 and whose locations are ill-suited for ffCO2 emission estimation. A sensor network optimized for ffCO2 measurements does not currently exist. Consequently, we have focused on developing the methodology first, under idealized conditions. We introduce a number of methodological and modeling novelties - we could not find any use of wavelets, sparse reconstruction and non-negativity enforcement in the atmospheric inversion of rough (non-stationary) emission fields. We will add this rationale to the "Discussion" section in our revised manuscript. See Sect. 6, paragraph 2.

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

Response: We have added this to the revised manuscript. See Sect. 5, paragraph 2.

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

Response: We agree. We will include our rationale in the "Discussion" section, as stated above. See Sect. 6, paragraph 2.

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

Response: The two estimates differ slightly in the sense that strong ffCO2 sources may be estimated at slightly different locations. Since the spatial distribution of ffCO2 fields is rough, neighboring locations may have very different, i.e., small, emission estimates, leading to large relative errors (division by almost zero). Consequently we have not found relative error plots to be useful – the range of relative errors is set by a few grid-cells where these shifts occur.

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 ~1 ppm, corresponding to current radiocarbon capabilities."

Response: We fail to see what these transect would provide, since the fluxes are those before the imposition of non-negativity. As might be expected, the fluxes show positive and negative values. The negative fluxes and their frequency are small and are captured by the Cumulative Distribution Function that the reviewer has asked for. The spatial gradients can be seen in Fig. 6 (Fig. 8 in revised manuscript) and more figures are available in the Technical Report cited in the paper.

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

Response: We will reword to better reflect our accomplishments. See Sect. 7, paragraph 3.

Pg 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."

Response: We agree and will remove the discussion.

Fig. 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."

Response: The relative error plot is not very informative for the reasons described above for the comments for Pg. 1290:13 and Pg 1299:2-6. We will add a difference plot (between true and estimated emissions) to the paper. We will update the text to point of the comparison with Vulcan emissions. See Fig. 8 and Pg. 20, paragraph 2.

1280:5 & 1286:5 The reviewer found grammatical errors.

Response: Will be fixed in the manuscript

1289:8 "Why switch to delta from alpha used earlier?"

Response: alpha is normalized and is [0, 1]. Delta is not.

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

Response: We will provide a link to the copy of the Sandia technical report stored in Sandia's technical library. The software supplied with this paper is a more difficult challenge. We require permission from the US Department of Energy to license and distribute any software. This may not arrive before the paper is finalized and we may not be able to supply a codebase in time.

Fig. 5, Fig 7 and Fig 9: The reviewer points out that certain markings, axes in figures, colors used for plotting etc. are not very legible, and ought to be magnified, truncated etc. for readability.

Response: We will make the changes in the manuscript.

*Fig. 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.*

Response: We will correct the manuscript to reflect this. See Fig. 9 in revised manuscript.

References

[Gourdji et al, 2012] Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.

Response to Reviewer # 2

The reviewer states: "The use of nightlights versus BUA needs some more detail. The authors indicate that BUA uses nightlights and hence these are not independent, though land-cover maps are also used. Perhaps some further comment on whether or not that makes much difference would help. I worry that these are essentially the same thing and/or what differences do exist are hard to interpret via the results presented here."

Response: BUA maps bring very little prior information to the inversion (vis-à-vis nightlights), and sometimes lead to slightly worse reconstructions. This is discussed on Pg. 1298:10-25, but not succinctly stated anywhere. We will add this to the new 'Discussions' section that we have added to the paper. See Pg. 22, paragraph 2 and Sect. 6, paragraph 3.

The reviewer states: "This effort seems well-suited to explore questions of measurement siting....... What network of CO2 measurements might be more fossil-sensitive? This is independent of radiocarbon. . . there are perhaps locations where the fossil signal from Vulcan is well represented by the treatment here? That would offer some practical guidance to future network expansion or interpretation of OCO2 measurements. I don't expect the addition of these tests to the manuscript but some comment at the end would be a very useful addition to those readers pondering the practical utility of the approach."

Response: We have updated our manuscript to reflect this suggestion. See Pg. 6 paragraph 2 for context and Sect. 6, last para.

Our method can infer ffCO2 emissions provided informative measurements of ffCO2 concentrations are available. The current sensor network, sited with an eye towards biospheric CO2 fluxes, does not provide them. The reviewer suggests, quite correctly, our method can be used in OSSEs to design an ffCO2 sensor network. It can be used to determine locations and frequency of measurements of ffCO2 sensors, as well as the fidelities required of the atmospheric transport model and the measurements themselves (i.e., the size of the model-data mismatch). We will add this in the "Discussion" section. It can also be used to quantify the errors introduced into regional scale inversions by uncertainties in the inflow and outflow boundary conditions. These errors may be quite large, and may overwhelm the differences that exist between various ffCO2 inventories i.e., they may all be equally good as a source of prior information.

Page 1280, line 21: "I think a key reason why there have been fewer attempts at inverting for the fossil component of the carbon fluxes is the difficulty of observationally separating the fossil component in observed atmospheric CO2. Other that 14C, expensive and not comprehensively observed, there are few tracers that can offer much constraint. I agree that part of the problem is the underlying spatiotemporal variability but this is only half the problem."

Response: We agree and we will add this in the manuscript. In Sec. 1, we will state clearly that ffCO2 inversions are uncommon because of the cost of measuring ffCO2, primarily via radiocarbon. See Pg. 3, last para.

Page 1288, line 15: "be wary of CARMA – not a peer-reviewed dataset and has many problems."

Response: We agree and we will add this caveat at the end of Sec. 3.1. See Pg. 12, last paragraph of Sect. 3.1.

Page 1296, line 4: "Am I understanding c correctly in that it represents an aggregate total? If so, the aggregate totals in EDGAR aren't really EDGAR values but probably IEA country totals."

Response: The reviewer is correct that *c* represents an aggregate total. It is the ratio of the aggregate total of ffCO2 emissions to the aggregate total of radiances (for the nightlights) or percentages of built-up areas. We will add this clarification in Sec. 3.2, immediately after Eq. 4. See last paragraph of Sect. 3.2.

Response to Reviewer # 3

The reviewer states: "The authors obtain spatial sparsity fractions of about 80% for representing Vulcan emissions using wavelets. While these compression rates may, at first, seem phenomenal, they are not surprising given that most of the gridcells in the region of interest in Fig. (1b) do not contain significant levels of fossil fuel CO2 emissions. One alternative and naïve approach to wavelets would be to consider only those gridcells, or some aggregated set of gridcells (e.g. 4×4), that contain emissions above a specified level (e.g. >1% of the max). Another approach could be to prescribe spatial basis functions that have areas proportional to population (i.e. small areas for large metropolitan regions, and large areas for rural regions). I surmise that these naïve approaches would also lead to large sparsity fractions or reductions in dimension. To better illustrate the strengths of their wavelet approach, I recommend that the authors devise a naïve metric of sparsity and compare and contrast their numbers to this metric."

Response: The reviewer is correct inasmuch that there are other sparse representations of ffCO2 emissions e.g., simply retaining grid-cells above a threshold. However, what a collection of selected grid-cells will *not* provide is a random field model, i.e. a systematic way of constructing a field based on independent model parameters. The discussion in Section 3 (selection of wavelets) is not about sparsity per se but about a sparse random field model for ffCO2 emissions. A random field model could be constructed from the selected grid-cells by requiring that they be related to others in some fashion. One such relation could be to constrain the values of the grid-cells using a spatial variogram. This would result in a Gaussian random field, which are used when estimating biospheric CO2 fluxes (and which, as we described in Sec. 1, will not work for ffCO2 emissions). Alternatively, one could use a multi-resolution approach for the spatial model, using wavelets to develop spatial structures in a continuous-level-of-detail manner.

A sparse collection of grid-cells, with an inversion method treating each as an independent parameter, would not be very useful – measurements at sensors are not very sensitive to them individually. In contrast, the MsRF allows one to systematically perturb the weights of wavelets whose supports span the entire domain (which will have a substantial impact globally on the sensors) down to the fine-scale ones (2 degree x 2 degree supports), with very local impact. Ensuring that the MsRF is sparse (low-dimensional) is thus very significant in an inversion setting.

In our paper, we had assumed that simply using the term "random field model" would convey our intention of constructing fields in a systematic manner (for use inside an optimization algorithm, when solving an inverse problem). We will add this clarification - our explicit interest in constructing fields systematically and ensuring that we can do so using as few free parameters as possible – in the beginning of Sec. 2.1. See Pg. 8, Sect 2.1, paragraph 1.

The reviewer states: "On page 1300, lines 26-28, the authors note that the deterministic nature of their presented method is a drawback. Without quantified confidence intervals and uncertainties, it is difficult to ascertain the significance of the inversion results (e.g. as shown in Fig. 7). The authors should run additional inversion tests that vary ε2, ε3, and other relevant parameters, and then report on the sensitivity of their results to these variations. Furthermore, the manuscript should contain a discussion of the errors described in items 3 and 4 below."

The main source of uncertainty in our inversions is the sparse set of measurement towers and their locations (far from the source of ffCO2 emissions). In a realistic (not synthetic data) inversion, the boundary conditions will also contribute their uncertainty, as will the "error" model for the data-model discrepancy. The numerical parameters mentioned above are not significant sources of errors.

The two parameters, ε_2 and ε_3 are numerical tolerances that are set, as is common in numerical studies, by reducing them till the results (wavelet coefficient estimates) become insensitive to their values (variations drop below 1.0e-6). Selection of the correct setting for M_{cs} is more involved and is described in Ray et al, (2013); it involved increasing M_{cs} till the error in ffCO2 estimates fellow below a threshold (10%).

We will add the rationale for the values of the numerical tolerances, and the citations to Sec. 5. See Sect. 5, paragraph 2. In the new 'Discussion' section, we will enumerate the sources of uncertainties whose impacts our deterministic method does not quantify. See Sect. 6, paragraph 1.

The reviewer states: "Underreporting is a known and persistent bias in using inventory-based estimates for monitoring anthropogenic emissions. The authors should describe what happens to this important source of error when using nightlights and BUA as spatial proxies for inventories in their wavelet representation. Does this error become confounded with separate errors in the proxies and can it be attributed to the inventory post-inversion? In a similar vein, are there errors in the proxies (e.g. clouds obscuring nightlights) that become confounded with the inventory in the wavelet representation?"

Response: Underreporting will have no effect on the accuracy of the estimation. Underreporting of country-level emissions lead to a smaller value of *c* in Eq. 4, and a smaller f_{pr}. However, f_{pr} serves as normalization for wavelet weights **w**. A smaller *c* will not lead to any differences in the wavelet weights *relative to each other,* and consequently, have no impact on the minimization of the L1 norm. The normalizing prior, of course, has no impact when evaluating the (Y – G'w') constraint in Eq. 10.

Errors in nightlights and BUA can lead to smeared reconstructions. Errors in these proxies lead to an erroneous selection of wavelets in the MsRF. If we omit a fine-scale wavelet from the MsRF in a region with high ffCO2 emissions, it will be captured using a coarser wavelet that covers the region with the ffCO2 source (leading to a smeared source). If we select a wavelet in a region without significant ffCO2 emissions, the sparse reconstruction method will simply set its weight to zero. Complications can occur if an erroneously chosen wavelet has (1) a single sensor sited in its support AND (2) is far from all other sensors. The measurements at a sensor are immensely sensitive to emissions/fluxes in its vicinity. In an inversion, in this particular case, the measurements at the sensor could be attributed entirely to the emissions modeled by the erroneously chosen wavelet. Occlusion of nightlights by clouds is not a major issue since these proxies are annually averaged quantities.

We will add this in the new "Discussion" section of the paper. See Sect. 6, paragraph 1 and 3.

The reviewer states: "The inversions are performed assuming a perfect atmospheric model. In reality, atmospheric models contain biases and other imperfections that can severely limit the *ability to invert for regional scale surface emissions. The authors should describe how model imperfections could be included the inversion (e.g. as an extra term in Eq. 5) and how they might be confounded with other errors in their sparse wavelet representation."*

Response: Transport model errors are not unique to estimation of ffCO2 emissions; they are also encountered in the estimation of biospheric CO2 fluxes too, and are addressed by using observation-specific model-data error variances (the diagonal terms in Re, in Eq. 6). We have cited papers by Chatterjee et al, (2012) and Gourdji et al, (2012), where their calculation has been described in detail. Adapting these methods to our ffCO2 emission problem is outside the scope of the paper. Observation-specific model-data error variances would result in the rescaling of the constraint in Eq. [10]. We will add this to the "Discussion" section. See Sect. 6, paragraph 2.

The reviewer states "The inversion results for the U.S. shown in Fig. (7a) exhibit pronounced seasonality, with small error reductions during periods 7 and 27, and large error reductions offset by 2-3 months during periods 15 and 35. The time dependence of the inversion suggests the presence of multiple time scales of interest that do not seem to be represented in the inversion demonstration. Although the wavelet coefficients in Eq. 7 vary with time (i.e. they contain index k), the wavelets themselves do not (i.e. do not contain index k). Are the spatial distributions of the nightlight and BUA proxies fixed for the year? If so, would introducing time-varying spatial distributions of these proxies reduce this seasonality? Please respond and include appropriate discussion in the manuscript."

Response: The errors in the estimation are due to changes in the wind direction, which blow the ffCO2 away from the measurement locations. Having a new MsRF for each k would imply that ffCO2 emitting regions change significantly on a weekly (or seasonal) basis, to the point that one needs to select new wavelets. This is unlikely. More practically, nightlights and BUA maps are computed as annual averages to remove the effect of cloud cover, holidays etc. on nightlight radiances, and it would be difficult to construct time-varying prior emissions, from a feasibility point of view.

We do, of course, allow for multiple timescales; our emissions are estimated at 8-day resolution, which is sufficiently fine to capture any seasonal changes e.g., change of wind patterns, and we perform inversions for 360 days.

The reviewer states: "In a comment related to item 5, fossil fuel emissions also vary over multiple time scales (daily, weekly, monthly, and yearly). Although the manuscript adequately describes the various spatial scales (and "spatial" is specified in the title), the discussion of multiple time scales is haphazard. I recommend including this discussion in the manuscript and describing how the sparse wavelet technique can (or cannot) be extended to capture multiple time scales. Making a clearer distinction between multiple time and space scales will also be helpful."

Response: We are somewhat confused by the comment of haphazard modeling of time-scales of ffCO2 emissions, since our paper contains none. Emissions are estimated over 8-day periods, and each period is considered to be independent of the others. The MsRF is designed to address the non-stationary spatial nature of ffCO2 emissions. We have not addressed the temporal modeling of ffCO2 emissions because of the reason below.

Changing (seasonal) wind patterns (which blow ffCO2 away from our limited set of measurement sites) pose one of the stiffest challenges to accurate emission estimation, and is the issue that we have investigated here. The primary variation of ffCO2 emissions, as represented by the Vulcan dataset, is diurnal (approximately 2x, with an afternoon peak). A much smaller spatial variation occurs, comparatively very slowly, when emissions shift from the hot south to the cold north in winter. In this paper, we have focused on investigating whether the MsRF is useful for estimation purposes, given the seasonal nature of wind patterns. Seasonal processes are adequately resolved at weekly resolution. The seasonal variation of ffCO2 emissions occur at a timescale far longer than the time taken by ffCO2 to be transported across the US (roughly two weeks) and consequently we have not had to impose any kind of temporal correlation in the emissions to obtain our ffCO2 estimates.

We will add, in the new Discussion section, why we have not attempted to address the multiscale temporal nature of ffCO2 emissions. Sophisticated temporal modeling of ffCO2 emissions, to simultaneously capture both the diurnal and seasonal variations, could probably be performed using non-stationary correlation functions. Our MsRF was not designed to do this, and we conjecture that sparse reconstruction and wavelets would likely be overkill. See Sect. 6, second-last para.

There is a possibility that the reviewer may have been misled by our use of the term "non-stationary". We do not use it to mean time-variant or unsteady. Rather, we use it in the statistical sense. For a stationary function, statistical summaries (means, variances etc.) computed within a moving window defined inside the support of the function would remain the same. More practically, a function that can be characterized by a single time/length scale e.g., a sine wave, is stationary, whereas another, displaying different scales in different parts of its support, is not. Consequently, a time-series of smoothly varying temperature can be stationary, whereas the porosity field of a block of dry soil may not.

The reviewer states: "The manuscript attributes inversion differences to differences between EDGAR and Vulcan emissions. The authors should also compute and report the raw differences between these two emissions inventories before they are used in the inversion demo."

Response: We do not ascribe the estimation error to differences between EDGAR and Vulcan. The primary source of estimation error is the lack of informative measurements (too few towers, and winds tend to blow ffCO2 emissions away from them). The differences between EDGAR and Vulcan are a small source of error. We will mention these errors (between annually averaged Vulcan and EDGAR) in Sec. 5. See Sect. 5, paragraph 2.

The reviewer states: "The synthetic observations used in the inversion, which are first introduced on page 1291 and later discussed on page 1295, should be described more clearly and in more detail. Were the elements of the sensitivity matrix H generated for another problem and adapted for this manuscript or were they computed specifically for this paper? As a numerical verification test, do the sensitivities multiplied by the Vulcan fluxes equal the concentrations obtained from a single forward simulation using Vulcan (i.e. does y equal Hf as given in Eq. 5)? More information about the WRF setup would also be useful (What lateral boundary conditions were used to generate the winds? What physics packages options were used? and so on)."

Response: The H matrices used in this paper were generated for two previous studies (Gourdji et al, 2010; Gourdji et al, 2012), which describe in great detail the gridding and the WRF settings used to construct them. We cite the papers on Pg 1291:21, and fail to see what repeating the same details would contribute to our paper. However, we will update the paper to explicitly mention the references where details on the calculation of H (meshes, models and settings) can be obtained. See Sect. 4, paragraph 2.

The reviewer also asks "do the sensitivities multiplied by the Vulcan fluxes equal the concentrations obtained from a single forward simulation using Vulcan (i.e. does y equal Hf as given in Eq. 5)?" We presume he/she means "if the sensitivities multiplied by the estimated (NOT Vulcan) fluxes equal the concentrations obtained from a single forward simulation using Vulcan?" Yes, they do. This is equivalent to asking if the estimated fluxes reproduce the observations (y). The agreement is plotted for a few measurement towers in Fig. 9 (right).

The reviewer states: "The authors analyze and display (Fig. 3) the statistics of non-zero wavelet coefficients. To help with visualization, it may also be useful to display maps of a few of the major features obtained from the wavelet decomposition."

Response: We thank the reviewer for this excellent suggestion. We will do so in the revised version of the paper. It is Fig. 5 in the revised manuscript.

The reviewer states: "On page 1288, line 13, the authors incorrectly associate static sources with emissions from highways. While it is true that highways are fixed, the traffic flow along them is not. CO2 emissions from traffic are usually categorized as mobile and non-stationary."

Response: The reviewer is correct that emissions from highways are classified as mobile. Averaged over time, they appear as line sources whose strengths vary along the line. However, they do not move, and can be captured by the same set of wavelets. Hence we called them "static" sources.

The lines no longer exist in the revised manuscript.

The reviewer has some suggestions regarding rewording of some figure captions, in the interest of clarity

Response: We will make the suggested changes

The reviewer states: "Please add "et al" to the Friedlingstein reference on pages 1278 and 1303. Also, according to recent work (see Fig. 1a in Regnier et al, doi:10.1038/ngeo1830), fossil fuel emissions are not the largest net carbon flux at the atmosphere-surface interface. Please revise the second sentence in the Introduction accordingly."

Response: Thank you for pointing out the error in the Friedlingstein reference. We are a little confused by the Regnier reference. Fig 1a therein clearly shows that fossil-fuel emissions are indeed the largest NET carbon exchange between land and atmosphere.

References

[Chatterjee et al, 2012] Chatterjee, A., Michalak, A. M., Mueller, K. L., Anderson, J. L., and Yadav, V.: Towards reliable ensemble Kalman filter estimates, J. Geophys. Res.-Atmos., 117, D22306, doi:10.1029/2012JD018176, 2012.

[Gourdji et al, 2010] Gourdji, S.M., Hirsch, A.I., Mueller, K.L., Yadav, V., Andrews, A.E., and Michalak, A.M.,: Regional-scale geostatistical inverse modeling of North America CO2 fluxes: a synthetic data study", Atmospheric Chemistry and Physics, 10:6551-6167, 2010.

[Gourdji et al, 2012] Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, doi:10.5194/bg-9-457-2012, 2012.

[Ray et al, 2013] Ray, J., Lee, J., Lefantzi, S., Yadav, V., Michalak, A. M., Bloemen-Waanders, B., and McKenna, S. A.: A multiresolution spatial parametrization for the estimation of fossil-fuel carbon dioxide emissions via atmospheric inversions, SAND Report SAND2013-2919, Sandia National Laboratories, Livermore, CA 94551-0969, available at: http://www.sandia.gov/ jairay (last access: 3 February 2014), 2013.