Anonymous Referee #3

General comment:

The authors developed an extended version of Stagewise Orthogonal Matching Pursuit (StOMP) method to estimate a multi-resolution field in a linear inverse modeling framework. The developed method is capable of including prior information and imposing non-negativity on the estimated field, which improve the accuracy of the estimation. The authors implemented an inversion to replicate a set of 8-day constant non-negative emission fields that were created using the Vulcan fossil fuel emission data product (Gurney et al. 2009) as a truth. The authors used a set of synthetic tower CO2 data (which is a radiocarbon-like tracer) that the authors have developed and used for their previous studies (e.g. Ray et al. 2014). The reproducibility of the truth emission fields seems to be fair. The text is well written. The authors had to describe so many technical details (as another RC pointed out), but I thought the authors did a really great job to summarize the methodological background of this study to invite a wide spectrum of the audience of GMD to this discussion. I believe the method developed in this study could be implemented and useful for some real applications. Also, the authors are providing a MatLab code. This is a very good for us, the audience of GMD, to keep track of their excellent modeling work.

However, it is a big pity to me that the authors did not choose a good test case to prove the practicability of the method developed. As also pointed out by another RC, the setup for their synthetic fossil fuel CO2 (ffCO2) emission inverse problem is too far from the reality. Because of that, I had a difficulty in discussing the feasibility of this method in a fossil fuel emission estimate. In my opinion, what the authors supported by the synthetic experiment is the replication of “non-negative emission fields” in a linear inverse problem, not ffCO2 emission fields. Thus, I would suggest to rework on the synthetic experiment and/or reword some of the text in the manuscript depending on what the authors would like to claim by this manuscript. In the following two sections, I’m trying to discuss several major concerns. Initially, I was listing points I wanted to discuss, but I learned that most of them have been raised by RCs for the author’s previous paper (Ray et al. 2014).

http://www.geosci-model-dev-discuss.net/7/1277/2014/gmdd-7-1277-2014-AR1.pdf (last access: Nov 20, 2014)

So I decided to make use of the *pdf as a reference. I understand that the *pdf was for discussions for Ray et al. (2014), but I think this is still fair to do as in anyways I would raise the exactly same questions/concerns and we probably don’t want to replicate the same conversations presented in the *pdf.

1. FFCO2 does not seem to be the best emission to be used for the synthetics study.

As acknowledged in Ray et al. (2014), ffCO2 emissions is very difficult to estimate given the existing observation network and also data collected there. The recent paper Shiga et al. (2014) (I see two of the authors are listed) also confirmed that as well. Shiga et al. (2014) also pointed out that the winter month have a better luck in estimating (detecting) ffCO2, but it would be very difficult the rest of the month. Probably this is something the authors should acknowledge in the manuscript. I also strongly feel that the authors need to convince us of the use of tower CO2 for ffCO2 estimation. Or since it was a synthetics study, the authors could come up with an ideal tower network for a ffCO2 emission inverse problem.

As we see in the manuscript, the use of existing tower network made the synthetic inversion setup far off from the reality. For example, the error assigned to the radiocarbon-like tracer was 0.1 ppm. This is too small compared to an actual C14 case as pointed out by the review discussion for Ray et al. (2014). It was acknowledged in Ray et al. (2014), but this small error was used again in this manuscript without any note. The authors mentioned the use of Carbon Monoxide (CO) in their response to defend the use of the 0.1 ppm error. I would not say no, but it is not clear to me how CO data from the existing tower data would help us to estimate ffCO2 emissions that well. If the authors meant to say the use of satellite-driven CO, that would become another problem as an inversion setup needs to be modified (also, data number would dramatically increase). My point here is that it is very questionable that such an ideal tower-based observational data become available sometime in near future, although I definitely think we push forward to make that happen as a community (in my opinion). The reality is we don't see any plan to expand the tower network to look at fossil fuel emissions (as far as I know). Perhaps I would be a little bit more convinced if the authors would have used observation from Orbiting Carbon Observatory 2 (OCO2) and/or Greenhouse Gases Observing Satellite (GOSAT) (although I don't know how to derive ffCO2 contribution in XCO2 data). Also, the emission imposed by Vulcan does not keep the nature of fossil fuel emission fields because of the simplification (averaging). Fossil fuel emissions are not constant over 8 days. We do sometime estimate monthly and/or weekly natural fluxes via inversion, but we do usually have diurnal cycle in forward modeling. Because of points I made above, I thought this manuscript is misleading when I read the abstract saying "We demonstrate the method on the estimation of ffCO2 emissions".

I would suggest to the authors to choose a different application to design a synthetic inversion. It is fair to claim that this manuscript is focusing on a method development. But the synthetic experiment needs to be reasonably close to the reality, especially if a particular target is assumed. If the authors want to stick to the ffCO2 inverse problem, the authors should design more realistic experiment, prove the feasibility of the method and show the reproducibility of the truth. Maybe one thing the authors easily could do is to reword "fossil fuel emissions" to "non-negative emission fields". After proving the reproducibility of the true field, the authors could discuss a possible application to fossil fuel emission estimation (like people do when they propose a numerical scheme). This correction should be fair enough to support the development of the inversion scheme for multi-resolution fields in a linear inverse problem.

2. *What is the real utility from this method in FFCO2 study?*

Apart from the synthetic study presented for a moment, I would like to discuss about the utility of the method presented in the manuscript (which I thought a major weakness of this manuscript). Again, I understand this manuscript is trying to advance the methodological aspect of future ffCO2 estimation method using atmospheric measurements (which I think it is great). I'm grad to see the authors mentioned current issues in ffCO2 inventory/modeling. But I was a little bit disappointed because this work is not really responsive to those issues. For instance, even in the the highly-idealized condition, the method does not give an accurate answer. Given by the nature of the method, this method could only offer a good approximate of ffCO2 fields, but not an accurate field. I thought this is a critical shortcoming. The authors claimed that the error in their estimation is around 5%. But the 5% is not the same thing as 5% two sigma uncertainty for reported national emissions. In real world, we need to deal with ffCO2 that has diurnal cycle while other strong signals from biosphere are present (and often we have a difficulty in decent angle them). In addition, the method is assuming a 1x1 degree field where we still don't resolve most of sectoral emission activities (at this spatial resolution, emission spatial proxy such as population and nightlights works pretty well). If I need a ffCO2 emission field, say for this year 2013 for US at a 1x1 degree, I would just project Vulcan emission data (or CDIAC 1x1 degree map) using fuel consumption data (I see this is essentially a very, very crude version (also w/o observation) of what the authors have done in this manuscript). In this way, I don't get any update on the emission spatial extent, but at least I could expect a very good estimate of the annual total emission. Given that, I'm not clear about what the method presented in this manuscript could offer to us. It seems to me that it is very difficult to expect the method to yield a very accurate emission field or update inventories. What we ultimately expect is to be able to tell the discrepancy between what is reported (or calculated) and what is measured using atmospheric observation (e.g. MRV). But it is not clear to me if the method could offer the opportunity as it seems to be difficult for atmosphere to tell a subtle difference in emissions especially at an aggregated 1x1 degree. Of course these do not need to be fully addressed, but I would be very curious how the authors envision to approach to the goal from where we are. I think these kind of discussion should be done to figure out what we really need to do, prior to getting into the technical details. We probably all agree on a need for an accurate ffCO2 estimate. If so, people would still prefer to use a very slow inversion method if the method yields an accurate answer. I thought the authors could easily work on more simple inverse problem for non-negative fields using an ideal observation network (and/or satellites) to estimate accurate emissions. I would thus suggest to rework the text and the synthetic experiment to be more responsive to ffCO2 issues if the authors would like to address a ffCO2 issues.

Line by line comments:

P5624, L13: This synthetic inversion was implemented in the highly-idealized synthetic world. Although the Vulcan data was used to create a truth field, this synthetic study is not fair enough to support the implementation of a fossil fuel emission estimation (see my general comment). This needs to be rephrased. Otherwise this could easily give an wrong impression to the audience of GMD.

P5624, P17: This is certainly an improvement achieved by this paper. But in my opinion, if the authors meant to develop a method to estimate FFCO2, the first thing the authors should try is to get an accurate emission estimate. Even if a method is very computationally heavy, people would be happy to use it to get an accurate estimate. As a study for FFCO2 estimation, this is not very appealing (at least, to me).

P5625, L16: I think the important thing here is how accurate the estimate is.

P5626, L11: I don't see the method developed here is going towards this goal.

P5626, L21: Since this method is based on the parameterization, it would be difficult to achieve an emission estimate that is accurate enough to updated or improve an emission inventory. This limitation needs to be acknowledged.

P5626, L26: It is unclear to me how this method could offer a method to give an accurate emission estimate using atmospheric measurements.

P5627, P19: I would like to learn the benefit this method could offer as opposed to the use of an gridded inventory plus fuel statistics (see my general comment).

P5628, L20: So here, the authors assumed that biosphere fluxes is perfect? Or ffCO2 can be solved as a separate problem?

P5630, L11: How would you specify Q here?

P5630, L20: Given the use of **f - fpr**, would you be able to "correct" an inventory using atmospheric observation?

P5640, L1: The synthetic observation is really a key in this study. This should be fully described. It is really misleading because CO2 concentration is actually a radiocarbon-like tracer, not normal CO2 concentration usually measured.

P5640, L2: Vulcan is averaged. This should be considered when you evaluate the numerical accuracy. I assume the authors did it right, but I was actually not very sure of it by this manuscript.

P5640, L7: Again, this CO2 is a synthetic radiocarbon-like tracer.

P5640, L7: The link to the Ray et al. (2013) is not active.

P5640, L22: So your "truth" is a 8-day averaged Vulcan. Correct? Did you use the averaged Vulcan for creating synthetic CO2 data or hourly Vulcan data? I want to make sure.

P5640, L24: So the afternoon selection is not applied.

P5641, L2: This is too small. Also, no transport uncertainties?

P5645, L3: Spell out BAO and MAP. I assume they are not used for the synthetic inversion. Correct? Why were those two sites selected? I saw this in a pessimistic way as the model is showing almost perfect fit to the observation while the emission estimate is still not perfect.

P5645, L1: Again, I think the accuracy of estimation needs to be discussed first if ffCO2 emission estimation is the ultimate target of this project.

P5645, L13: The authors should defend that why those small minor feature can be ignored while a very high accuracy is required to improve and/or update inventories.

P5645, L28: We should not have a trade off especially in this highly-idealized world. Again, we need an accurate estimate to address an issue in ffCO2 study.

P5646, L15: This means the method doesn't get both national total and spatial distribution right.

P5647, L16: Again, this needs to be rephrased.

P5648, L9: Two drawbacks are acknowledged in Ray et al. (2014) (See P1917, 1: Need for a good fossil fuel tracer, 2: uncertainty as also mentioned here), but the first one is not acknowledged.

P5649, Candes -> Candés

P5652, L3: Ray et al. (2013) is not accessible from the link indicated.

P6555, Figure 2. Compared to Table 2 from Asefi-Najafabady et al. (2014), the correlation is not very good. The FFDAS (Nightlight+Population) scores a correlation of 0.86. This is a part of the reason I don’t see a utility of this method. Especially this is a synthetic study, the author should seek the way to get much better score.

Table 2. Comparison of Sector-Specific 2002 FFDAS v2 Fossil Fuel CO_2 Emissions for the United States With Vulcan Fossil Fuel CO_2 Emissions Data Product From Gurney *et al.* [2009][a]

	FFDAS (NL and POP)		FFDAS (NL)		FFDAS (POP)	
Sectoral Comparison at 0.1°	SAD	Correl	SAD	Correl	SAD	Correl
FFDAS versus Vulcan all sectors	714	0.86	731	0.86	749	0.85
FFDAS other versus vulcan other	551	0.48	567	0.49	587	0.45
FFDAS other versus vulcan residential	740	0.69	791	0.69	733	0.84
FFDAS other versus vulcan onroad	552	0.82	587	0.83	577	0.76
FFDAS other versus vulcan nonroad	802	0.64	854	0.64	795	0.67
FFDAS other versus vulcan commercial	781	0.50	832	0.50	775	0.54
FFDAS other versus vulcan industrial	858	0.16	904	0.17	863	0.12
FFDAS other versus vulcan airports	820	0.25	871	0.24	815	0.18

[a]The FFDAS v2 results are constructed as only population-based (FFDAS (POP)), only nightlight-based (FFDAS (NL)), and combined population- and nightlight-based (FFDAS (NL and POP)). The summed absolute difference (SAD) and spatial correlation (correl) are used as the comparison metric.

P5658, Figure 5. Again, I thought this is very pessimistic (if I interpreted this figure correctly). Maybe you could add a forward modeling result using original Vulcan emissions and highlight the difference.

References:

Asefi-Najafabady et al. (2014) JGR
Shiga et al. (2014) GRL