Response to reviewers

A sparse reconstruction method for the estimation of multiresolution emission fields via atmospheric inversion

Ray, Lee, Yadav, Lefantzi, Michalak and van Bloemen Waanders, GMDD 7:5623-5659, 2014.

Response to Reviewer No. 2

We thank the reviewer for his/her suggestions and comments. Our responses to his/her suggestions are below.

Overall response: At the very outset we would like to clarify that our study is focused on investigating the algorithmic aspects of a sparse reconstruction method (based on Stagewise Orthogonal Matching Pursuit, StOMP) for estimating rough emission fields, such as that of fossil-fuel CO2 (ffCO2). A sparse reconstruction method is necessary since the spatial parameterization for rough fields tends to be high dimensional (many parameters). The parameters that can be estimated depend on the information content of the observation data, which can change with time/season. The method is not customized to a particular tracer, measurement network or a transport model. Customization to a tracer occurs when we choose a spatial parameterization (a wavelet-based random field model in this study) for use with our sparse reconstruction method. It also occurs when we choose an observational dataset. The method can accommodate prior information on the field being estimated, but only uses its spatial pattern; thus, by design, it is insensitive to under/overestimation of the emissions in the prior information.

The paper investigates which formulations of the inverse problem do and do not work, and explains why. It develops a metric (mutual coherence) to quantify the information content in the observations collected by our measurement network. It finds the information content lacking, which motivates the need to introduce prior information into the inverse problem. We then identify a way to do so; the obvious/intuitive ways do not work. We also show how a wavelet-based field model, designed for modeling fields in rectangular geometries, can be used to estimate emission fields in an irregular region $\mathcal{R}$ (the Lower 48 states of the US). Finally, we show how StOMP can be extended to enforce non-negativity on the estimates. Sparse reconstruction methods are typically not used in atmospheric inversions.

Our motivation to develop this method arose from a need to construct and/or validate gridded inventories of ffCO2. Fortunately, many gridded ffCO2 inventories are available and a wavelet-based spatial parameterization also exists. We demonstrated the method in an idealized, synthetic-data inversion. The idealizations include: (1) assuming ffCO2 to be a radiocarbon-like tracer and ignoring interference by biospheric CO2 which can make ffCO2 estimation impossible except in winter (see Shiga et al., [2014]); (2) using a model-data mismatch ε that is smaller than the one used in real-data inversions and (3) assuming the same distribution for ε for all towers (i.e., ignoring transport model errors). These idealizations allowed us to explore issues related to the algorithm and formulations in a relatively “clean” setting. We also use an observational dataset collected from a measurement network that was sited with biospheric CO2 fluxes, not ffCO2, in mind (the towers are usually far from locations with high ffCO2 emissions); a network for ffCO2 does not currently exist.

Due to these idealizations adopted in our test, we do not claim that the method can be used to estimate ffCO2 emissions fields in a realistic setting using measurement techniques and infrastructure that are currently available (or could be in the near future). At the very least, our method has to be extended to include the estimation of biospheric fluxes as well as larger and tower-dependent model – data mismatches. This is a substantial body of work and outside the scope of this study. In order to check how accurate the estimates would be, we would have to conduct an OSSE (Observational System Simulation Experiment) or design an ideal network. Our tests also provide no information on the best method to collect information on estimation of ffCO2 emissions over regional scales (tower, airplane transects etc.).

We check our inversion method using the following metrics:

1. As part of our algorithmic development, we modify StOMP to incorporate prior information to improve estimates. We check whether it indeed does so, since the information content of the observations are found to be poor.
2. The aim of sparse reconstruction is to estimate parameters that are supported by data (usually large spatial patterns in the emission field) and remove the details that are not. We check whether this "sparsification" characteristic of the algorithm is still present after including prior information.
3. Our method restricts emission fields in an irregular region $\mathcal{R}$ (while using a wavelet-based model); this incurs a computational cost that can be limited by a user-defined setting. We check if the behavior of the algorithm provides a principled way of computing that setting (e.g., if improvement of results shows a "diminishing returns" behavior with the computational cost).

Note that in this study we do not use the accuracy of the estimated emission fields as a metric for evaluating our inversion method. This is because accuracy of estimation is determined primarily by two factors (once we have specified a model – data mismatch): (1) the suitability of the spatial parameterization for the rough fields being estimated and (2) the information content of the observational dataset. In our previous paper [Ray et al, 2014] we fixed the observational data and used the accuracy of the emission estimates to gauge the quality of the spatial parameterization. The converse procedure – fixing the spatial parameterization and varying the quantity of observational data – is not a very useful direction for investigation, for our StOMP-based method, because of the following reasons:

1. The estimation accuracy of StOMP, as the quantity of observational data is varied, has been investigated in Donoho et al, [2012]
2. If the aim is to obtain a very accurate reconstruction of the ffCO2 field (when we have full discretion to design an ideal observation network/technique), then we are limited only by what the spatial parameterization can capture. As reported in our previous paper [Ray et al, 2014], the spatial parameterization with 1023 wavelet coefficients (parameters) has a relative error of 10% at the 1-degree resolution; this would be recovered (modulo the small model – data mismatch) in case of an ideal network. If we retain all wavelet coefficients that can be described on a 1-degree mesh in the spatial parameterization (4096 coefficients), the reconstruction will be perfect (modulo the model – data mismatch).

Atmospheric inversion could be a way of estimating/verifying self-reported $ffCO_2$ emissions in countries where the uncertainty is high. The uncertainty in emission reports from China is estimated to be 15-20% [Andres et al, 2012], though studies based on the TRACE-P campaign proposed a 54% revision of inventory estimates for 2000 [Suntharalingam et al, 2004] (it was officially revised upwards by 23% between 2006-2007). Other countries have larger variations. These uncertainties affect inventories, but do not affect our inversion method (see paragraph 1). Even if our variable of interest were to be total emissions over a region (nation or province), estimating a spatially variable emission field before spatially aggregating it reduces the aggregation error. However in order to do this, a measurement infrastructure designed with $ffCO_2$ in mind is a requirement. Its size will be determined by whether we are interested estimating total national emissions or we seek fine scale details.

In addition, as mentioned above, our method can be used with other tracers provided we have a spatial parameterization for them.

The introduction section in our paper does not describe the idealized nature of our tests or the limits/caveats on the conclusions that can be drawn from them. It also does not describe the reasoning behind the metrics that were adopted for evaluating our algorithm. We will add them in the revised paper.

Detailed comments

Issue # 1: Then reviewer states "The summary of section 2 states that "mutual incoherence may offer analytical in- sight into the quality of observations and uniqueness of solutions". But in the text, I could not find a proof of this point. It would be very useful for the community if the authors could add examples to illustrate how this method can detect bad quality observations, and show the uniqueness of flux solutions."

We have expressed ourselves badly. What we meant was "mutual incoherence may provide an analytical metric for the quality of observations and consequently, solutions". We will change the sentence in the updated manuscript.

It was not our intention to provide proofs that show low mutual coherence (incoherence) leads to informative observations and is a necessary condition for obtaining a unique solution (without the use of prior information). Necessary conditions for a unique solution also involve a property called Restricted Isometry. Proofs on the necessary conditions for accurate sparse reconstruction can be found in the literature on compressive sensing.

We have, however, calculated the mutual coherence between our transport matrix ("sensing matrix" in compressive sensing terms) and our bases and shown them to be far inferior to the ones achieved in compressive sensing. This is primarily due to where the measurement towers are placed (the network was designed with an eye towards biospheric CO_2 fluxes, not $ffCO_2$). Given the lower information content in our observations, the conventional compressive sensing way of solving the inverse problem (i.e., without prior information, except sparsity of representation using wavelets) provided poor estimates (Approach A). We did not investigate the non-uniqueness of solutions.

Note that mutual coherence does NOT help us identify "bad quality" observations in the sense that interference from an anomalous source or a faulty instrument corrupts them.

Instead it helps us identify if the measurement towers "intercept" ffCO2 emissions transported by the wind and thus gather information on them. Given a network of measurement towers and a transport model, mutual coherence provides a metric for the quality of the observations that the network could offer without any noise in the measurements. Noise reduces the quality of the observations further.

Issue # 2: The reviewer states "The purpose of the proposed method is to estimate fossil fuel CO2 emissions. As stated in the conclusion, the fossil fuel CO2 emission is not the only type of emission in nature. I would like to see an example of estimating fossil fuel emissions with this method in the presence of inaccurate biosphere fluxes. It is possible that this method could not work over the entire year, but could possibly work in some months of the year (e.g., January). The authors could then discuss the challenges of estimating fossil fuel emissions in a more realistic scenario."

We thank the reviewer for this suggestion; it provides us with a better structure for explaining what we did, and what remains to be done. The aim of the paper, as described in "Overall response" is to present an inversion algorithm, for rough emission fields, that is insensitive to over/under-estimation in prior beliefs regarding the emissions in question. Performing a real-data inversion, using uncertain biospheric fluxes, would be outside the scope of this paper.

References

[Andres et al, 2012] R. J. Andres et al, "A synthesis of carbon dioxide emissions from fossil-fuel combustion", Biogeosciences, 9, 1845-1871, 2012. doi: 10.5194/bg-9-1845-2012.

[Donoho et al, 2013] D. Donoho et al, "Sparse solutions of underdetermined linear equations by Stagewise Orthogonal Matching Pursuit", IEEE Transactions on Information Theory, 58, 1094-1121, 2012.

[Ray et al, 2014] J. Ray et al, "A multiresolution spatial parameterization for the estimation of fossil-fuel carbon dioxide emissions via atmospheric inversions", Geoscientific Model Development, Geosci. Model Dev., 7, 1901-1918, 2014. doi:10.5194/gmd-7-1901-2014.

[Shiga et al, 2014] Y. P. Shiga et al, "Detecting fossil fuel emissions patterns from sub-continental regions using North American in-situ CO2 measurements", Geophysical Research Letters, 41(12):4381-4388, 2014.

[Suntharalingam et al, 2004] P. Suntharalingam et al, "Improved quantification of Chinese carbon fluxes using CO2/CO correlations in Asian outflows", Journal of Geophysical Research, 109, D18S18, 2004. doi:10.1029/2003JD004362.